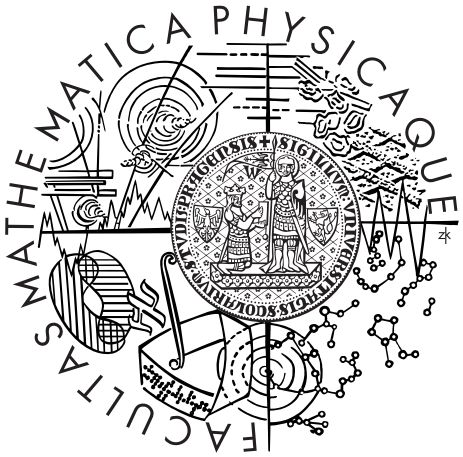
Charles University in Prague
Faculty of Mathematics and Physics

DOCTORAL THESIS



Jaroslava Schovancová

Measurement of the top quark properties

Institute of Physics, Academy of Sciences of the Czech Republic

Supervisor of the doctoral thesis: RNDr. Jiří Chudoba, Ph.D.
Study programme: 4F9 – Subnuclear Physics

Prague 2016

To my family.

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In Geneva, Switzerland in March 2016 Jaroslava Schovancová

Název práce: Měření vlastností top kvarku

Autor: Mgr. Jaroslava Schovancová

Katedra: Fyzikální ústav Akademie věd České republiky, v. v. i.

Vedoucí disertační práce: RNDr. Jiří Chudoba, Ph.D., Fyzikální ústav Akademie věd České republiky, v. v. i.

Abstrakt: Tato disertační práce se zabývá měřením diferenciálního účinného průřezu systému páru top-antitop na experimentu ATLAS jako funkce příčné hybnosti, hmoty a rapidity systému páru kvarků top-antitop. Při měřeni byl analyzován vzorek přibližně 4.7 fb^{-1} dat nabraných v roce 2011 při proton-protonových srážkách při energii srážky v těžišťovém systému 7 TeV. Tvary spekter diferenciálního účinného průřezu v systému $\mathrm{t\bar{t}}$ jsou konzistentní s předpovědí Standardního modelu a dobře popsány dostupnými generátory.
Dále byly prezentovány příspěvky do ATLAS Distributed Computingu, zejména v oblastech operace a provozu, monitorování, automatizace, jakož i vývoje Workload Management Systému PanDA. Tyto aktivity mají za cíl zlepšit efektivitu a usnadnit využití distribuovaných výpočetních zdrojů a tím přispívají k podpoře fyzikálního programu experimentu ATLAS.

Klíčová slova: ATLAS, top kvark, diferenciální účinný průřez, chybějící příčná hybnost, ATLAS Distributed Computing

Title: Measurement of the top quark properties

Author: Mgr. Jaroslava Schovancová

Department: Institute of Physics, Academy of Sciences of the Czech Republic

Supervisor: RNDr. Jiří Chudoba, Ph.D., Institute of Physics, Academy of Sciences of the Czech Republic

Abstract: This Thesis presents the ATLAS experiment measurement of the top quark differential cross-section as a function of p_{T}, mass and rapidity of the $\mathrm{t\bar{t}}$ system. A sample of approx. 4.7 fb^{-1} of the 2011 pp collision data at the center-of-mass energy of 7 TeV was analyzed. The differential spectra shapes in the $\mathrm{t\bar{t}}$ system are consistent with the Standard Model and reasonably described by the event generators.
Several activities in the scope of ATLAS Distributed Computing are presented, particularly in the area of operations, monitoring, automation, and development of the PanDA Workload Management System. Such activities aim to improve efficiency and facilitate the use of distributed computing resources and therefore contribute in supporting the ATLAS Physics Program.

Keywords: ATLAS, top quark, differential cross-section, missing transverse energy, ATLAS Distributed Computing

Acknowledgments

I would like to thank to my supervisor Jiří Chudoba for providing me the opportunity to explore beauties of distributed computing in the past decade, and to join the ATLAS Collaboration.

I would like to thank to many people who have helped me with discussions, ideas, encouragements and suggestions throughout the years of my PhD studies. While it is impossible to list you all, I would like to mention a few names. My thanks belong above all

to Alessandro Di Girolamo for being my friend and mentor over the past 5 years, providing me with guidance and discussing various topics related to the ATLAS Distributed Computing and beyond, and be a never-ending source of motivation and encouragements to finish this Thesis;

to Simone Campana for helping me realize that I can achieve some cool things, countless fruitful discussions over the years, and for sharing his knowledge of QCD while helping me study for my Doctoral exam;

to Jiří Kvita for providing me guidance, for fruitful discussions, and for words of encouragement during my work on studies of the top differential cross-section;

to Elena Korolkova for numerous helpful discussions and interactions, and kind words of encouragement throughout the years;

to Alexei Klimentov and the ADC Management for giving me the opportunity to come to CERN in 2011 to work in ADC and entrusting me with coordination of ADC Shifts and ADC Monitoring;

to Kaushik De and the ADC Management for giving me the opportunity to return to CERN in 2014 to become responsible for the ATLAS Central Services Operations and Security;

to Tomáš Javůrek and Martina Javůrková for helping me study for my Doctoral exam, for being great source of motivation for me, and for many discussions about physics and ATLAS Distributed Computing we had over the years;

to Tadashi Maeno for many friendly discussions about PanDA Workload Management System;

to Prof. Jiří Hořejší for giving me the opportunity and accepting me to this program, despite my originally purely Astrophysics background, and to Prof. Jiří Chýla for encouragement to write this Thesis;

to David Cameron for reading the manuscript and helping me in my article battle.

My life at CERN and elsewhere would not have been that much enjoyable without colleagues and friends from the Czech and Slovak ATLAS teams, namely Tomáš Kouba, Tomáš Javůrek, Martina Javůrková, Jaroslav Günther, Jiří Kvita, David Šálek, Zdeněk Hubáček, Pavol Federič, as well as my friends and colleagues from CERN IT Alessandro Di Girolamo, Simone Campana, Edward Karavakis, Pablo Saiz. Many thanks to you for a myriad of discussions about physics, ATLAS Distributed Computing or WLCG, and for being there for me!

My travels and stays at CERN were kindly supported by grants of the Ministry of Education, Youth and Sports of the Czech Republic, LA08032. Support from the Academy of Sciences of the Czech Republic, from ACEOLE: A Marie Curie ITN Project, from the grants SVV-2010-261 309 and GAUK-316911 of the Charles University in Prague, and from the US DOE ASCR and US DOE HEP Computing are greatly acknowledged.

Last but not least, many thanks belong to my dear parents and my family, who have tirelessly supported me throughout the the many years of my education.

Preface

"The results today are only possible because of the extraordinary performance of the accelerators, including the infrastructure, the experiments, and the Grid computing," were the words of the CERN Director General Rolf Heuer on the 4th July 2012 when the observation of a new particle consistent with a Higgs Boson was disclosed [1]. The end result of all the investments into building and operate the CERN's Large Hadron Collider is the data that is recorded and the knowledge it can be squeezed out. The value enclosed in the data is explored on the global computing infrastructure.

ATLAS Computing was essential for the reconstruction and analysis of the ATLAS data that lead to the discovery of the Standard Model Higgs boson in 2012 [2]. The great performance of ATLAS Computing and its contribution to the ATLAS Physics program was recognized during the Higgs seminar [1], as shown in Fig. 1.

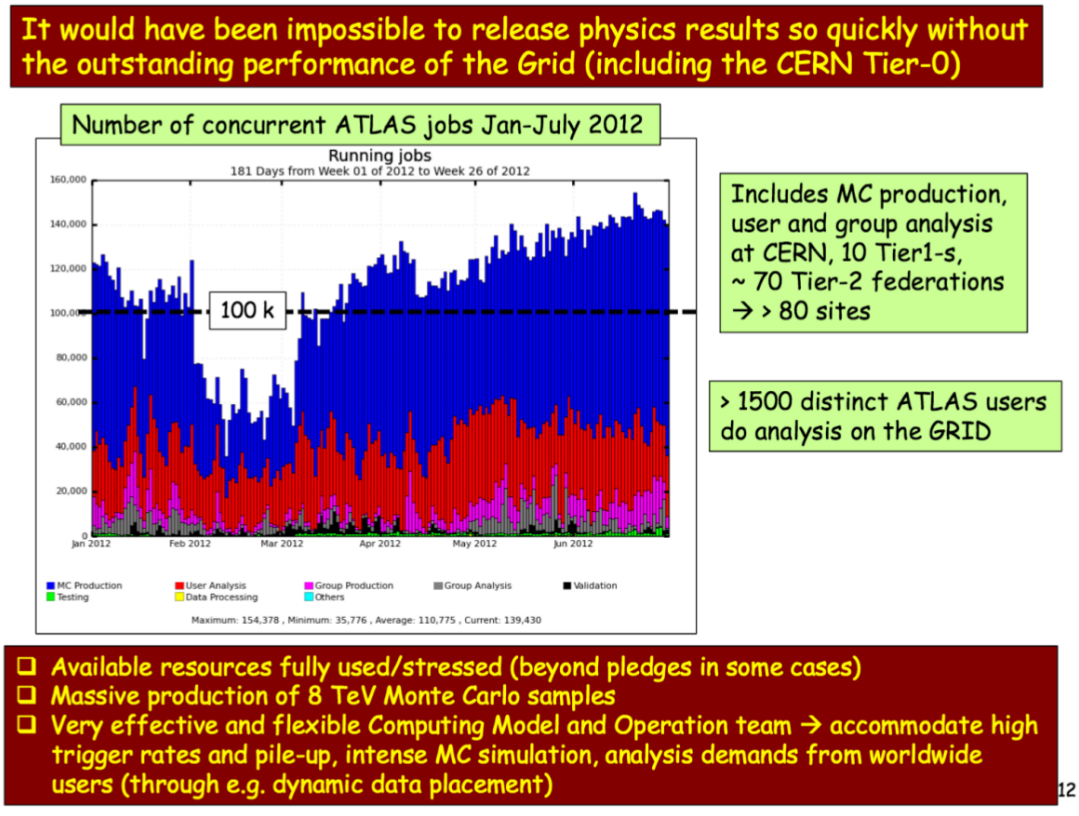


Figure 1: ATLAS Computing contributed to the quick release of physics results related to the observation of a new particle consistent with a Higgs Boson [2]. This slide was part of a talk presented by then ATLAS Spokesperson Fabiola Gianotti at the Higgs seminar [1] on 4th July 2012.

ATLAS computing provides ways to achieve ATLAS physics program goals, through development of SW products, conveying access to data and MC simulations, and efficient utilization of the ATLAS Distributed Computing resources. Operation of such a complex system is a challenging task, however an essential one for production of high-quality physics result of the ATLAS Collaboration.

The main focus of this Thesis is the "Measurement of top quark properties" in particular with the data collected with the ATLAS experiment at CERN during LHC Run I. The contributions to this topic are coming from various areas of the ATLAS activities, specifically from ATLAS Distributed Computing, Combined Physics and Performance program with emphasis of missing transverse energy performance in early LHC Run I analyses of Top quark physics, and Top Quark Physics with emphasis of measurement of differential cross-section of top-antitop pair system.

The Thesis is structured in the following way:

The concepts of the CERN's Large Hadron Collider and the ATLAS experiment are presented in Chapters 1 and 2. Chapter 3 brings a brief overview of the Standard Model of Particle Physics. The Top quark physics and its properties are outlined in Chapter 4, together with a summary of measurements of top quark properties carried out by the ATLAS experiment.

ATLAS Distributed Computing Overview of ATLAS Distributed Computing (ADC), and contributions to its mission are described in Chapters 5 through 8.

- Chapter 5 describes basics of ATLAS Computing, and further details on ADC tasks, challenges of the ADC Operations with the emphasis of the ADC shifts, and provides overview of ADC Monitoring tools used during LHC Run I.

- Chapter 6 the suite of monitoring and accounting tools developed during LHC Run I is presented. The author of this Thesis contributed as a coordinator of the ADC Monitoring project, responsible for the evolution and development for over 2 years during LHC Run I, and as a developer of several monitoring and accounting tools. One of the monitoring tools in particular, the Production Task Monitor, played a significant role in monitoring progress of data processing and generation of the Monte Carlo simulations for the Top quark group during the LHC Run I.

- Chapter 7 describes contributions to the automation effort of the ADC Operations, in order to save manpower on repetitive tasks, and focus on the more advanced topics. The automation effort is paying off by being able to utilize available computing resources more efficiently for the data processing and analysis activities carried out by the ATLAS physics groups,

and decreasing the waiting time needed to finish analysis by individual ATLAS physicists.

- Chapter 8 describes contributions to the PanDA Workload Management System, which is developed by the ATLAS Collaboration by end of 2015 being used not only by the ATLAS experiment, but used and evaluated by several other experiments as well, e.g. Alpha Magnetic Spectrometer (AMS), Large Synoptic Survey Telescope (LSST), Common Muon and Proton Apparatus for Structure and Spectroscopy (COMPASS). The wider-spread of usage of PanDA WMS beyond ATLAS was made possible in part by contributions to the PanDA server and PanDA monitor carried out by the author of this Thesis. PanDA monitor is the entry point for ATLAS physicists to monitor progress of their analyses on the world-wide distributed computing resources.

Combined Physics and Performance Contributions to the program of the Combined Physics and Performance group, in particular performance studies of the missing transverse energy in 2010 through 2012, with emphasis on Top Quark Physics are listed in Chapter 9.

Top Quark Physics Contributions to the measurement of the top-pair differential cross-section carried out in 2012 with the top quark signal in data collected with the ATLAS detector at $\sqrt{s}$ =7 TeV are described in Chapter 10.

The author of this Thesis had the opportunity to present an overview of ATLAS Top physics results at a conference [3].

Publication activity Results of the various contributions to these activities have been documented in 2 internal notes and 21 papers listed towards the end in the chapter *Selected List of Publications*. The author has also been a co-author of more than 530 ATLAS Collaboration publications since October 2008.

Contents

List of Abbreviations

ADC ATLAS Distributed Computing. iv

ADCoS ADC Operations Shift. 32

AFT Analysis Functional Test. 62

AGIS ATLAS Grid Information System. 60

AOD Analysis Object Data. 81

ATLAS A Toroidal LHC ApparatuS. 7

BLUE Best Linear Unbiased Estimator. 91

CE Computing Element. 64

CKM Cabbibo-Kobayashi-Maskawa matrix. 16

CPUtime Amount of time spent on processing instructions of a program of a computational job running on a worker node, as opposed to less CPU-intensive phases such as waiting for input/output. 111

CSC Cathode Strip Chambers. 12

DAQ Data Acquisition System. 29

DAST Distributed Analysis Support Team. 33

DDM Distributed Data Management. 29

ECAL ATLAS Electromagnetic Calorimeter. 10

EF Event Filter of the ATLAS Trigger System. 12

EM Electromagnetic Interactions. 16

EW Electro-Weak Interactions. 15

FCAL ATLAS Forward Calorimeter. 11

FTS File Transfer Service. 38

GGUS Global Grid User Support. 34

GOCDB Grid Operation Center DataBase. 61

GUID Globally Unique IDentifier. 38

HCAL ATLAS Hadronic Calorimeter. 11

HEC Hadronic End-Cap. 11

HLT High-Level Trigger. 13

HPC High-Performance Computing. 31

HS06 HepSpec06, unit of CPU benchmarking. 111

ID Inner Detector. 9

IP Interaction Point. 8

JSON JavaScript Object Notation. 56

JVF Jet Vertex Fraction. 94

L1 Level-1 of the ATLAS Trigger System. 12

L2 Level-2 of the ATLAS Trigger System. 12

LEP Large Electron-Positron Collider. 5

LHC Large Hadron Collider. 5

LS1 LHC Long Shutdown 1. 30

MDT Monitored Drift Tubes. 12

MET Missing Transverse Energy ($E_{\mathrm{T}}^{\mathrm{miss}}$). 77

OIM OSG Information Management System. 61

PanDA Production and Distributed Analysis Workload Management System. 30

PFT Production Functional Test. 62

PS Proton Synchrotron. 5

QCD Quantum ChromoDynamics Theory. 16

ROOT A data analysis framework developed at CERN. 82

SAM Service Availability Monitor. 61

SCT Semi-Conductor Tracker. 9

SE Storage Element. 64

SM Standard Model. 15

SPS Super Proton Synchrotron. 5

SURL Storage Uniform Resource Locator. 38

TRT Transition Radiation Tracker. 9

VO Virtual Organization. 31

WallTime Time allocated for a single computational job on a worker node. 110

WLCG World-wide LHC Computing Grid. 31

WMS Workload Management System. 30

1. Large Hadron Collider

The Large Hadron Collider (LHC) [4], [5] is a hadron accelerator and collider located at the CERN, the European Organization for Nuclear Research, near Geneva, Switzerland. The LHC was built in the former LEP [6] tunnel, which has circumference of almost 27 km. The LHC magnets produce a magnetic field intensity of over 8 T, used to bend trajectories of particles circulating inside the accelerator.

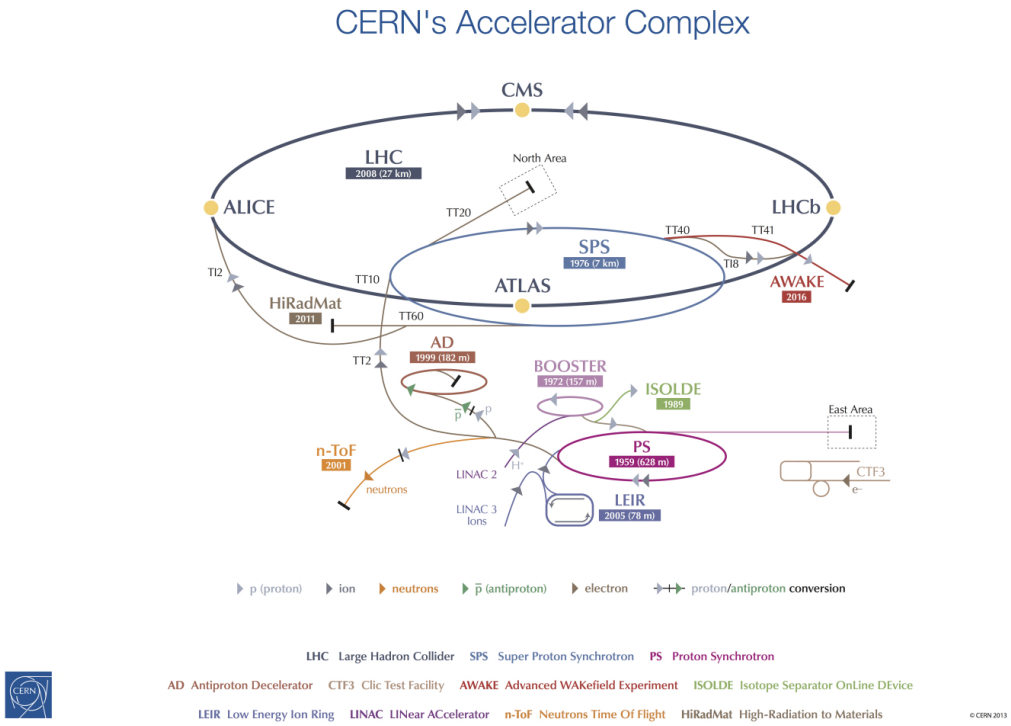


Figure 1.1: Schematic view of the CERN accelerator complex [7].

The accelerator complex (in Fig. 1.1) consists of a series of particle accelerators with a progressively increased energy of the protons (or ions). Protons are generated from ionizing hydrogen gas in an electric field. The protons are then accelerated to 50 MeV using a linear accelerator (LINAC2), boosted by the two circular accelerators (BOOSTER) to 1 GeV. Later the protons enter the Proton Synchrotron (PS), where are accelerated to 26 GeV. Upon acceleration in PS, the proton beams are injected to the Super Proton Synchrotron (SPS) to be accelerated to the minimum energy required by the LHC before injection, 450 GeV. There are two beams circulating in the LHC in the opposite directions and accelerating to the desired energy. When such a desired energy is reached, the beams are collided in 4 different interaction points, where 4 major LHC experiments are located.

A spectacular start of the LHC activity occurred on September 10th 2008.

However, after 9 days there was an incident [8] induced by the faulty magnet connection, which resulted in a large helium leak to the LHC tunnel, which destroyed several magnets. Upon LHC repair there was a upper limit set for the intensity of the electrical currents that create the magnetic field, resulting in the maximum energy of protons to 3.5 GeV. The first 7 TeV collisions took place in March 2010 and continued until the end of 2011. In February 2012 the LHC started data taking with center-of-mass energy of 8 TeV, which was used until the end of LHC Run I in February 2013. The LHC machine upgrades performed during the LHC Long Shutdown I enabled restart of data-taking of LHC Run II in April 2015, with center-of-mass energy of 13 TeV from May 2015. Analysis part of this Thesis focuses on data taken in 2010-2011 at the center-of-mass energy of 7 TeV.

2. ATLAS Detector

The ATLAS (A Toroidal LHC AparatuS) [9] detector is one of the four major LHC experiments. ATLAS is a general purpose detector designed for the physics program to explore a wide range of physics processes at the LHC: with the highly energetic collisions delivered by the LHC not only high precision measurements of the Standard Model parameters are performed, but new particle searches and searches for physics beyond the Standard Model are performed.

The physics program of the ATLAS experiment covers many interesting area of the High-Energy Physics: study asymmetry between the behavior of matter and antimatter known as CP violation as a part of the B-mesons physics group, study properties of the top quark, precise measurements of the Standard Model, search for new particles and new physics phenomena, such as research of broken supersymmetry, or acquiring inside in the nature of dark matter.

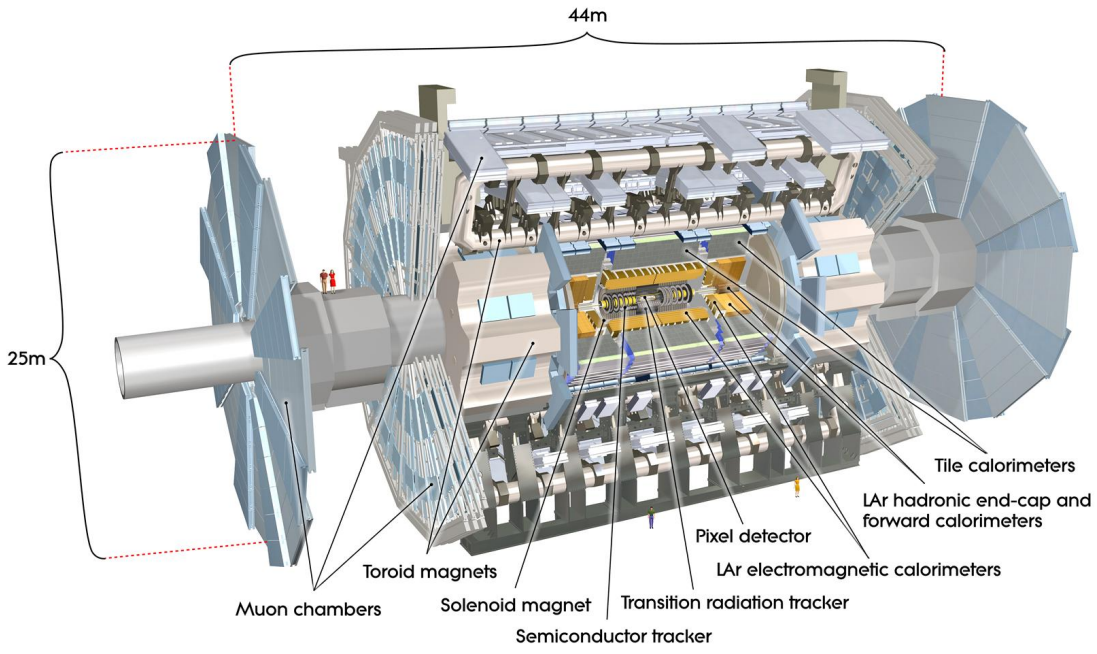


Figure 2.1: Computer generated image of the whole ATLAS detector [10].

One of the most challenging and visible goals of the ATLAS experiment was to investigate missing piece of the Standard Model mosaic, the Higgs boson. In July 2012 decades of search for the Higgs boson particle peaked in the Higgs boson discovery by the ATLAS [2] and CMS [11] experiments, and we transitioned from the Higgs boson search to Higgs sector study phase. In 2013 theoretical physicists François Englert and Peter W. Higgs were jointly awarded the Nobel Prize in Physics 2013 [12]:

> for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which

recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider.

The 3D illustration of the whole ATLAS detector [13] is shown in Fig. 2.1. The ATLAS detector is a forward-backward symmetric in regards to the Interaction Point (IP). The ATLAS detector is 44 m long and 22 m in diameter, its weight its about 7000 tons. All the particles produced at the IP where the beam particles collide are identified in the particle detectors designed with three different detecting subsystems: the Inner Detector, the Calorimeters, and the Muon spectrometer. Each of the three sub-detectors specializes in measurement of different characteristics of the particles, thus serves for their detection and identification. Overview of the sub-detectors is provided in the subsequent sections of this chapter.

2.1 ATLAS coordinates system

The ATLAS coordinate system (in Fig. 2.2) is a right-handed system with the origin in the interaction point. The z axis is defined by the beam axis, and the plane transverse to the beam line is the xy plane. The x axis points to the center of the LHC tunnel, and the y axis points upwards. The azimuthal angle ϕ is measured from the positive x axis, and increases clock-wise looking in the positive z direction. The polar angle θ is the angle relative to the beam direction.

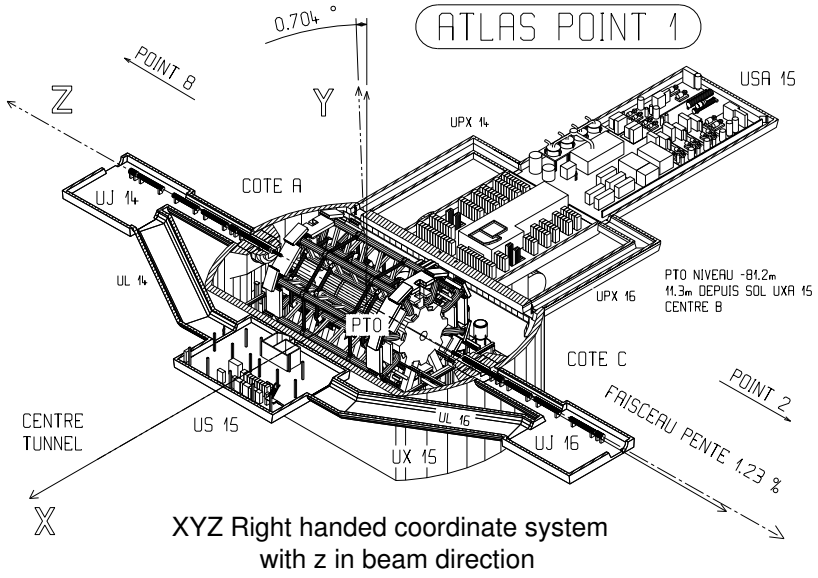


Figure 2.2: Global ATLAS coordination system, from [14].

A set of important parameters that describe particles inside the ATLAS detector contains in addition the transverse momentum p_{T} and pseudo-rapidity η.

The transverse momentum p_{T} is the particle's momentum projected to the xy plane. The pseudo-rapidity η of a particle from the primary vertex is defined as

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right). \tag{2.1}$$

Distance between two particles in the $\eta - \phi$ space is described by the cone size ΔR as

$$\Delta \mathrm{R} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{2.2}$$

2.2 Inner detector

The Inner Detector (ID, in Fig. 2.3) is the sub-detector located closest to the interaction point, covering regions of pseudo-rapidities of $|\eta| < 2.5$. Main purpose of the ID is to reconstruct trajectory and measure momentum of charged particles. The ID is immersed in a 2 T solenoidal magnetic field. The location of ID comes with great density of diversity of the particles, therefore isolation of a single track requires high granularity. In addition, the ID has to be able to cope with the very fast bunch crossing frequency of 40 MHz. The Inner Detector consists of three different parts: the Pixel Detector, the Semi-Conductor Tracker (SCT), and the Transition Radiation Tracker (TRT).

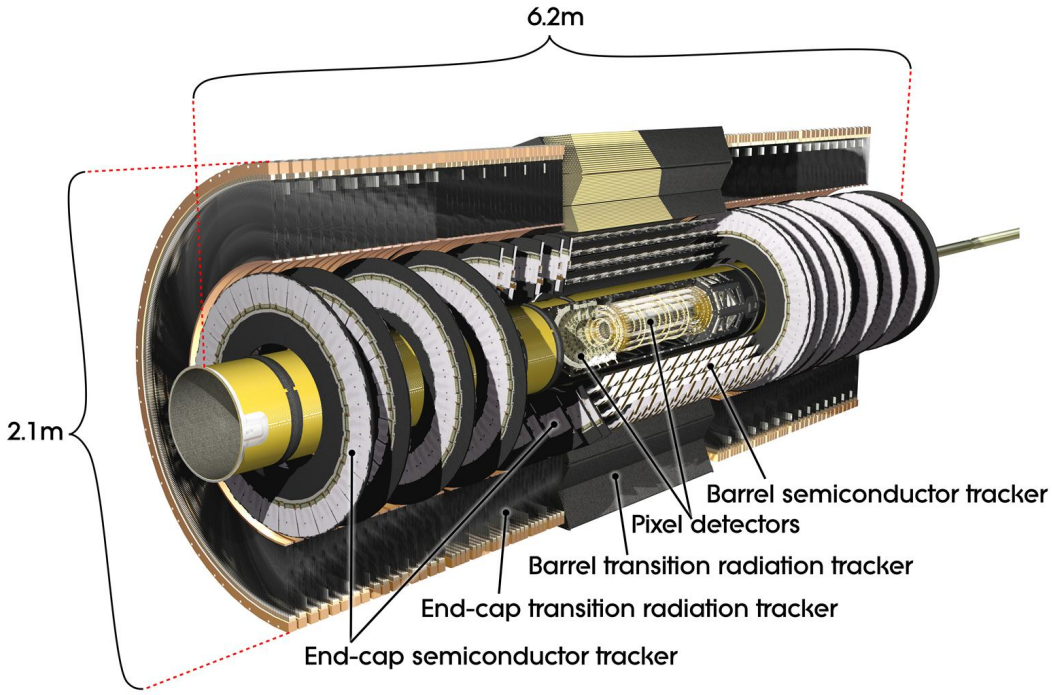


Figure 2.3: Computer generated image of the ATLAS inner detector [15].

2.2.1 Pixel detector

The Pixel detector is the innermost component of the ATLAS experiment. The smallest units of the Pixel detector are square pixels of 400 μm along the z axis and 50 μm in the xy plane. Each of the pixels is working as a diode driven in a reverse-bias mode. Typical track crosses three pixel layers. In total, there are about 80 millions pixels that provide good spacial resolution and exact vertex reconstruction. The bending plane of the magnetic field (which is identical with the xy plane) offers excellent position measurement precision of 12 μm, while spacial resolution in the z direction is roughly 66 – 77 μm.

2.2.2 Semi-conductor tracker

The SCT consists of the long narrow silicon strips with dimensions 80 μm $\times$ 12 cm, which serve as the precision tracking detector. The SCT system consists of four concentric double-layers designed to provide 8 measurements per track in the intermediate radial range. The spacial accuracy in the Rϕ plane is 17 μm, and in the z direction the resolution is 580 μm. There are approximately 6.3 million read-out channels in the SCT.

2.2.3 Transition radiation tracker

The TRT is located in the outermost layer of the ID. The TRT consists of straw tubes with a diameter of 4 mm. The drift chambers are filled with a Xenon mixture. Using the transition radiation measurements, the TRT allow electron identification. The TRT intrinsic accuracy is 170 μm per straw. The detector provides a large measurements per track (typically 36), therefore the combined resolution is better than 50 μm. There are approximately 351 thousands read-out channels in the TRT.

2.3 Calorimetry

The ATLAS calorimeter (in Fig. 2.4) system was designed to reconstruct energy of electrons, photons, and jets, and to ensure good measurements of the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$. It consists of two sampling detectors with the full ϕ-symmetry: the inner electromagnetic calorimeter, and the outer hadronic calorimeter.

2.3.1 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) is a lead/liquid argon (LAr) detector that consists of a barrel ($|\eta| < 1.475$) and two end-caps ($1.375 < |\eta| < 3.2$) housed in 3 cryostats. The LAr identifies electrons, positrons, and photons. The LAr measures energy and position of these particles with high accuracy.

The ECAL uses accordion-shaped absorbers and kapton electrodes, which collect the charge of ionization electrons. Such a structure provides the complete

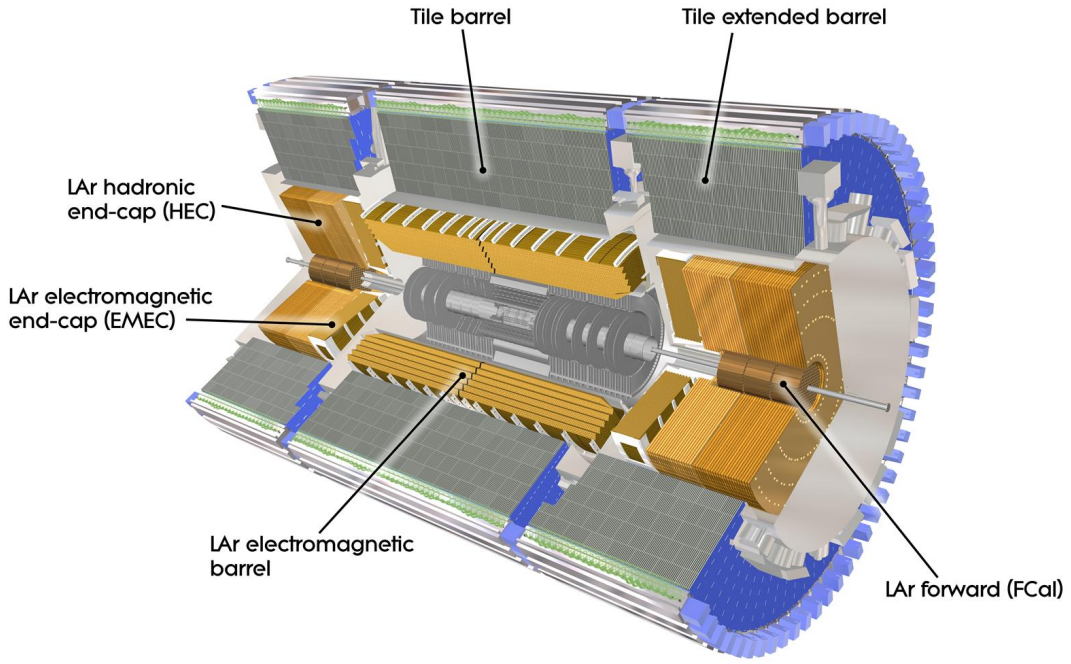


Figure 2.4: Computer Generated image of the ATLAS calorimeter [16].

ϕ-symmetry without azimuthal cracks. The total thickness of the detector is larger than 24 X_0 (where X_0 is the radiation length) in the barrel, and larger than 26 X_0 for the end-caps. There are approximately 190 thousands read-out channels in the ECAL.

2.3.2 Hadronic calorimeter

The main purpose of the hadronic calorimeter (HCAL) is identification and reconstruction of jets, as well as measurement of their energy losses, therefore measurement of the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$. The missing transverse energy is carried away by the particles without interaction with the detector (e.g. neutrinos), it is very important quantity for many channels. The $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement and reconstruction for the Top physics is discussed further in Chapter 9.

The HCAL system consists of two different detector technologies: the tile technology in the barrel part, and the LAr technology in the Hadronic End-Cap (HEC) and forward region (FCAL).

The Tile calorimeter is a steel/plastic scintillator plates detector that covers pseudo-rapidity range of $|\eta| < 1.7$, operating at the room temperature. At larger pseudo-rapidities the higher radiation resistance is needed, the LAr detectors operate: the HEC covers pseudo-rapidities range $1.5 < |\eta| < 3.2$, while the FCAL covers pseudo-rapidities range $3.1 < |\eta| < 3.9$. Both HEC and FCAL are housed in the same cryostat as the EM end-caps.

2.4 Muon spectrometer

The Muon Spectrometer (in Fig. 2.5) is the outermost component of the ATLAS detector. It also takes up the largest volume. The Muon Spectrometer was designed to measure the deflection of muon tracks in the superconducting air-core toroid magnets up to $|\eta| < 2.7$, and to trigger on the muon tracks in the range $|\eta| < 2.4$. In the majority of the η range of muon spectrometer, the Monitored Drift Tubes (MDT) provide high-precision tracking. At large pseudo-rapidities and closer to the IP the higher-granularity Cathode Strip Chambers (CSC) are used.

The toroidal field bends the track in polar angle θ, while the solenoidal field in the inner detector curves the path of a particle in the azimuthal angle ϕ. Thus the standalone momentum measurement is performed in the $R - z$ projection. The z coordinate is measured in the barrel ($|\eta| < 1$, while the R is measured in the transition ($1.0 < |\eta| < 1.4$) and the end-cap ($1.4 < |\eta| < 2.7$) regions.

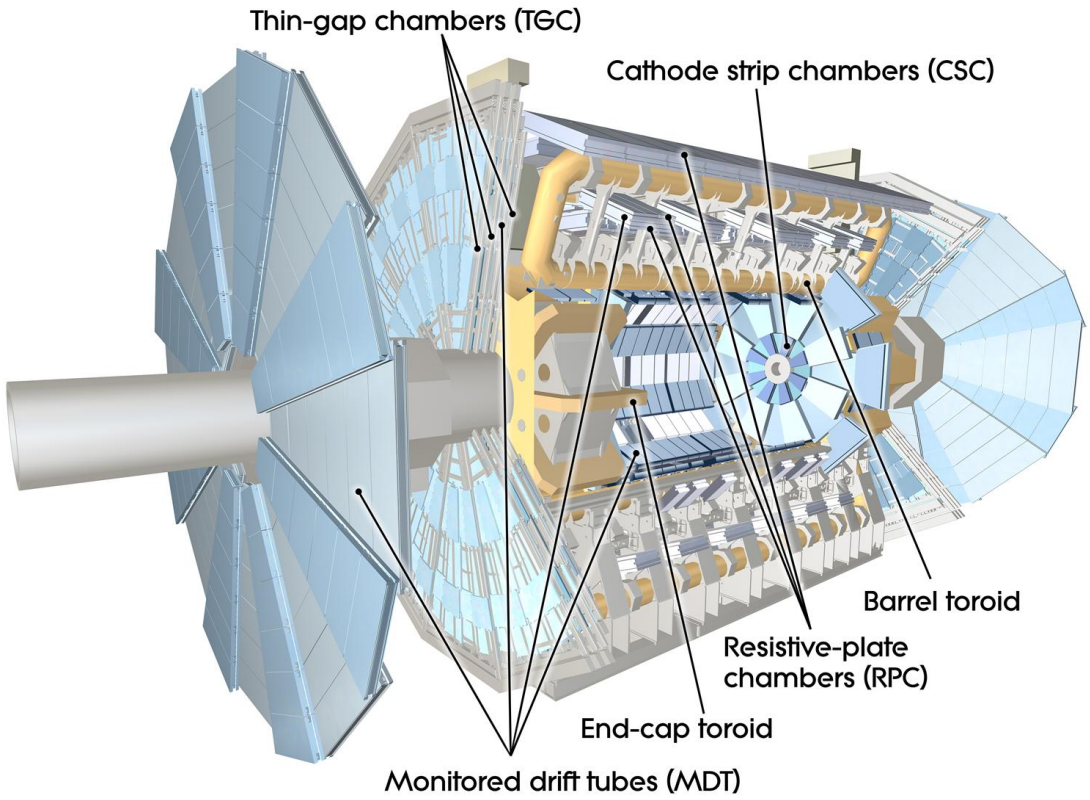


Figure 2.5: Computer generated image of the ATLAS Muons subsystem [17].

2.5 Trigger system

The ATLAS Trigger System provides 3 levels of online event selection: Level-1 (L1), Level-2 (L2), and the Event Filter (EF). The L2 and EF together compose

the High-Level Trigger (HLT), and are implemented entirely in the software, while the L1 operates on customized hardware boards.

The purpose of the trigger system is reduction of the bunch-crossing rate of 40 MHz to roughly 100 Hz for permanent storage and physics analysis. The L1 trigger makes an initial selection by exploiting the calorimeter system and muon detectors with reduced granularity information. The L1 trigger searches for high p_{T} muons, electrons or photons, jets, and τ leptons decaying into hadrons and large missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$. The L1 accept rate is about 75 kHz within the decision time of less than 2.5 μs. The L2 trigger uses full-granularity data from all detectors, and is able to combine information from the different sub-detectors. The L2 trigger reduces the event rate below 3.5 kHz, and the average processing time is 10 ms. Upon the L2 selection, the full event is sent to the EF level, that is based on the offline software. The output event rate to be recorded for the offline analysis is about 100 Hz, which corresponds to an output data rate of 100 MB/s.

The ATLAS Trigger and Data Acquisition system, as well as further steps of offline data processing are summarized in section 5.1, and the various aspects of the ATLAS Computing are described in Chapters 5 through 8.

3. The Standard Model

The Standard Model (SM) of strong and electro-weak (EW) interactions satisfactory describes the fundamental particle physics, its current formulation was finalized in the second half of the 20th century.

The SM describes constituents of matter, and interactions among them. A schematic depiction of elementary particles and some of their properties (mass, charge, spin) is shown in Table 3.1. According to the SM, the matter consists of quarks (u, d, s, c, t, b) and leptons (e, μ, τ, and lepton neutrinos ν_e, ν_μ, ν_τ), and their anti--particles anti-quarks $(\bar{u}, \bar{d}, \bar{s}, \bar{c}, \bar{t}, \bar{b})$ and anti-leptons ($\bar{e}, \bar{\mu}, \bar{\tau}$, and lepton anti-neutrinos $\bar{\nu}_e, \bar{\nu}_\mu, \bar{\nu}_\tau$).

Standard Model of elementary particles – mass, charge, spin					
Particle		Mass	Charge	Spin	Flavor
u	(up)	≈ 2.3 MeV			
c	(charm)	≈ 1.275 GeV	2/3	1/2	Quarks
t	(top)	≈ 173.34 GeV			
d	(down)	≈ 4.8 MeV			
s	(strange)	≈ 95 MeV	-1/3	1/2	Quarks
b	(bottom)	≈ 4.18 GeV			
e	(electron)	0.511 MeV			
μ	(muon)	105.7 MeV	-1	1/2	Leptons
τ	(tau)	1.777 GeV			
ν_e	(electron neutrino)	< 2.2 eV			
ν_μ	(muon neutrino)	< 0.17 MeV	0	1/2	Lepton neutrinos
ν_τ	(tau neutrino)	< 15.5 MeV			
g	(gluon)	0	0	1	
γ	(photon)	0	0	1	Gauge bosons
Z	(Z boson)	91.2 GeV	0	1	
W	(W boson)	80.4 GeV	± 1	1	
H	(Higgs boson)	≈ 126 GeV	0	0	Higgs boson

Table 3.1: Mass, charge, and spin of the Standard Model elementary particles [18].

Quarks and leptons have spin number[1] equal to 1/2, thus obey to Fermi statistics. The neutrinos interact only through the weak force, and their weak interaction eigen-states differ from mass eigen-states. Quarks, as well as leptons, are organized in three generations, each generation consisting of two quark particles of different flavor. The quarks differ from the leptons due to interaction through strong nuclear force and they form bound states (e.g. proton consists of

[1]Anti-particles of quarks and leptons have spin number equal to -1/2, and obey to Fermi statistics as well.

two u quarks and one d quark). Through EW interactions, the transition between an upper quark and any of the down quarks (d, s, b) is possible, however this is not observed for leptons.

The Cabbibo-Kobayashi-Maskawa (CKM) matrix V_{CKM} describes the probability for a quark of one flavor to change into a quark of another flavor:

$$V_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{3.1}$$

The single top quark production cross-section measurement enables direct determination of $|V_{tb}|$ without assumption of CKM matrix unitarity. The combined cross-section measurement $\left(2.76^{+0.58}_{-0.47}\right)$ pb performed by the DØ and CDF experiments [19] suggests that $|V_{tb}| = 0.88 \pm 0.07$. The limit was set to $|V_{tb}| > 0.77$ at the 95 % CL. The most precise determination of the CKM matrix elements comes from a global fit using all the available measurements, and constraints on the SM, e.g. three generation unitarity. The global fit result for the magnitudes of all CKM matrix elements reads [18]:

$$V_{\mathrm{CKM}} = \begin{pmatrix} 0.97425 \pm 0.00022 & 0.2253 \pm 0.0008 & (4.13 \pm 0.49) \times 10^{-3} \\ 0.225 \pm 0.008 & 0.986 \pm 0.016 & (41.1 \pm 1.3) \times 10^{-3} \\ (8.4 \pm 0.6) \times 10^{-3} & (40.0 \pm 2.7) \times 10^{-3} & 1.021 \pm 0.032 \end{pmatrix}. \tag{3.2}$$

The interactions among particles are mediated by the gauge intermediate bosons: the intermediate bosons have spin number 1, thus obey to Bose-Einstein statistics. Photon is mediating electromagnetic (EM) interaction, the W and Z bosons mediate the weak interaction, and gluons mediate strong interaction.

The Higgs boson plays a special role in the Standard Model. It is a scalar boson introduced in order to explain the mass of the elementary particles, as a consequence originating from the non-zero mean vacuum value of the Higgs field. The Higgs boson is the latest discovered elementary particle of the Standard Model as it was observed by the ATLAS [2] and CMS [11] experiments in 2012.

The interaction of particles in the Standard Model are described with a relativistic quantum field theory consistent with both special relativity and quantum mechanics theories. The SM combines the EW theory based on the gauge group $SU(2)_L \times U(1)_Y$, where quarks and leptons form left-handed doublets and right-handed singlets of the $SU(2)$ group, and the $SU(3)_C$ group that describes strong interactions in the scope of the quantum chromodynamics (QCD) theory. The SM group structure describes hypercharge (Y), weak isospin (L), and colour charge (C) gauge groups.

The Standard Model theory is in excellent agreement with experimental results. However, it leaves quite some unanswered questions and cannot be considered a final theory of the particle physics. The SM does not describe the gravitational interaction, does not explain hierarchy of particle masses, and cannot satisfactory explain the origin of the CP violation. There are theories going "Beyond the Standard Model" that attempt to complement the Standard Model, e.g. supersymmetry [20], superstring theory [21], and the extra dimensions theory [22].

4. Top Quark

4.1 Top quark physics

The top quark is an elementary fermion discovered by the CDF[23] and DØ[24] experiments at the Tevatron Collider in 1995. Its discovery completed the three-generation structure of the Standard Model.

The top quark phenomenology is driven by its high mass (≈ 173 GeV, [18]):

- Studies of electro-weak symmetry breaking: The top quark has a large mass, it is the only fermion with Yukawa coupling to the Higgs boson close to 1.
- Testing QCD: due to its large mass, the top quark production is a short-distance process. In addition, due to the small strong coupling constant $\alpha_S(m_t^2) \approx 0.1$ the rapidly converging behavior of the perturbative expansion constitutes the top quark an ideal perturbative object to test QCD.
- Constraint on SM Higgs mass: The accurate measurement of top quark mass m_t helps constraining the mass of the Standard Model Higgs boson m_H, as m_t enters the electro-weak observables through loop corrections that behave as powers of m_t.
- Test V-A behavior of the Standard Model: The top quark has a very short lifetime, decays before hadronization. Spin characteristics of top quark are passed onto decay products. This helps us to study and test the V-A behavior of the Standard Model.
- New physics Beyond the Standard Model: The top quark cross-section is sensitive to new physics through resonant production of top-antitop pairs. Top quarks constitute a significant background process to many models Beyond the Standard Model.

For these reasons top quark plays a special role in the Standard Model and many of its extensions. Knowing accurately various properties of the top quark, such as mass, couplings, production cross-section, decay branching ratios, can help to learn even more key information about fundamental interactions at the electro-weak breaking scale, and beyond.

4.2 Top quark production at LHC

Top quarks can be produced either through the strong interaction as top-antitop pair $t\bar{t}$, or through the electro-weak interaction as a single top quark. Feynman diagrams of the leading order of the $t\bar{t}$ production are shown in Fig. 4.1, for the single top production in Fig. 4.2. The Large Hadron Collider has been designed to produce millions of $t\bar{t}$ pairs per year (with $\sim 10\text{fb}^{-1}$ of collected data), dubbing the LHC the Top factory. QCD predictions of hard-scattering cross-sections of

various processes as a function of center-of-mass energy in proton-(anti)proton collisions at Tevatron and LHC are depicted in Fig. 4.3.

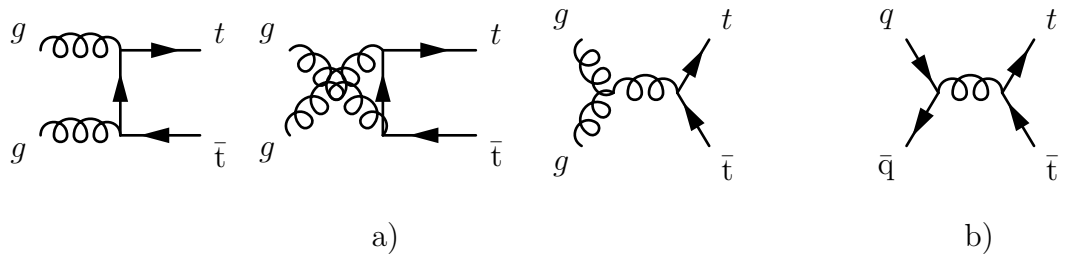


Figure 4.1: Top quark pair production through gluon fusion (a) and quark- anti-quark annihilation (b).

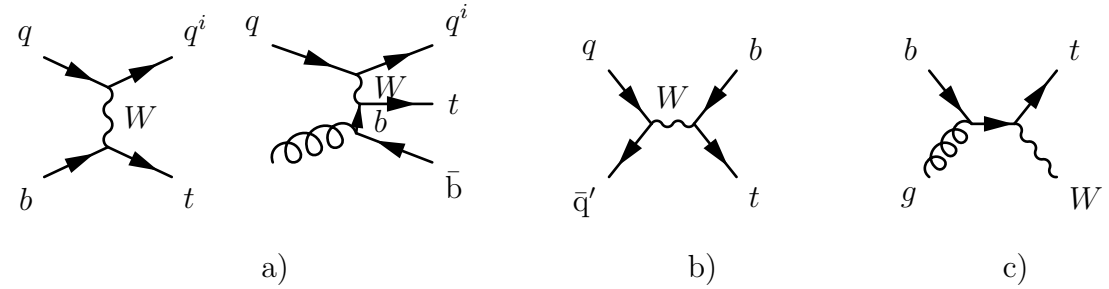


Figure 4.2: Single top quark production through W-gluon fusion "t-channel"[1](a), W "s-channel" (b), Wt production (c).

4.2.1 Top quark pair production

The top-antitop pair production at the LHC, shown in Fig. 4.1, is dominated by gluon fusion (gg) by 90% of time [26] in the case of $\sqrt{s} = 14$ TeV. The remaining 10% the top-antitop pair is produced from quark-antiquark ($q\bar{q}$) annihilation. At Tevatron during Run II the production ratios were roughly swapped.

4.2.2 Single top quark production

The Standard Model top quark can be produced also through electro-weak single top production in three distinct channels:

- $t-$channel (Fig. 4.2a): a virtual time-like ($q^2 \geq 0$) W boson is created by the quark-antiquark pair $q\bar{q}$ annihilation and decays into a top and a bottom quarks.

- $s-$channel (Fig. 4.2b): a virtual space-like ($q^2 \leq 0$) W boson interacts with a bottom quark, in order to produce a top quark.

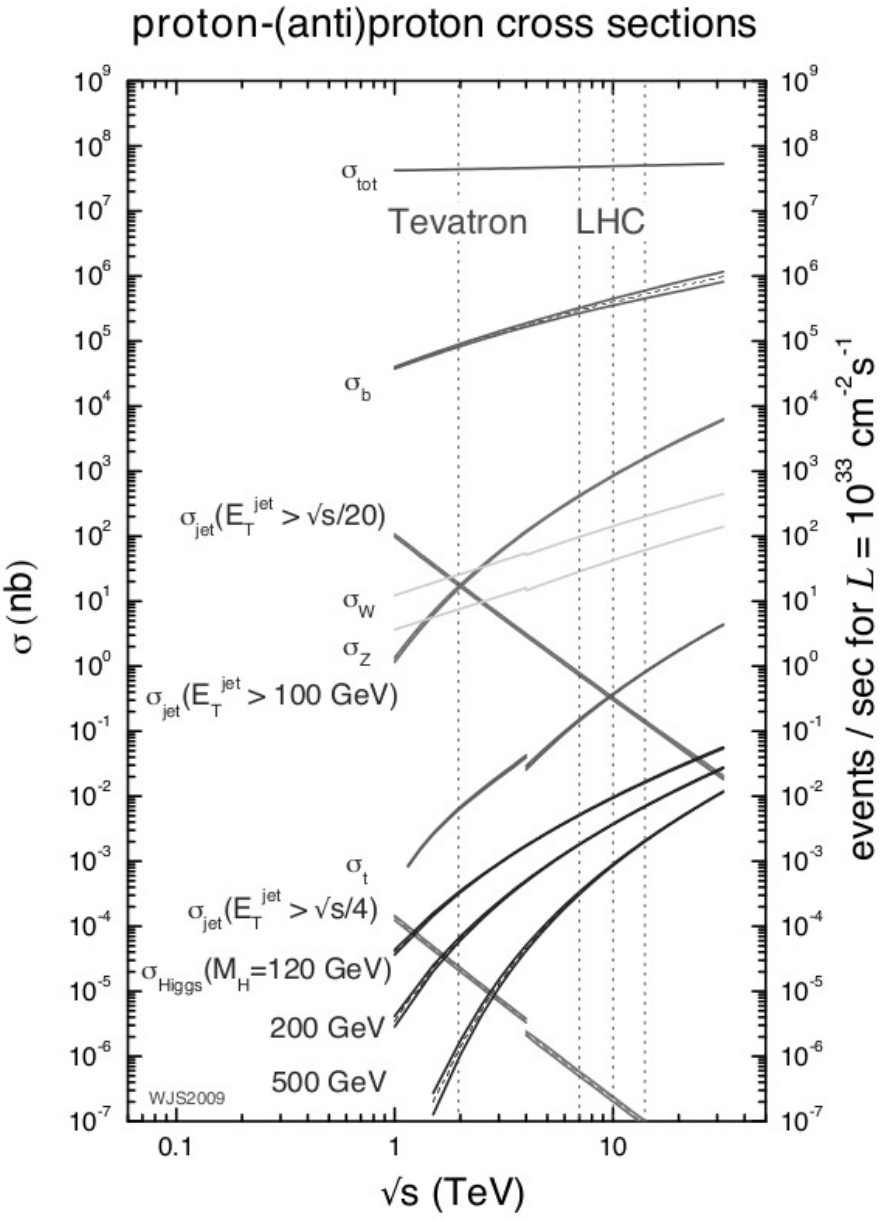


Figure 4.3: QCD predictions of hard-scattering cross-sections of various processes in proton-(anti)proton collisions at Tevatron and LHC [25] as a function of center-of-mass energy $\sqrt{s}$. Discontinuities in the various curves designate transition between proton-antiproton collisions at Tevatron to proton-proton collisions at LHC.

- Wt associated production (Fig. 4.2c): a real W boson and an initial parton (b quark) produce a top quark.

The single top quark produced in electro-weak interactions (Fig. 4.2) at hadron colliders proceeds through W-gluon fusion, W* process, or Wt production.

4.3 Top quark decay

The Standard Model top quark decays almost exclusively to a W boson and a b quark, $t \to Wb$. Top quark can decay into W boson and a quark: bottom, strange or down, however decays to s or d quark are suppressed – the corresponding elements of the CKM matrix (3.2) read [18]:

$$|V_{tb}| = 1.021 \pm 0.032 \quad (4.1)$$
$$|V_{ts}| = (40.0 \pm 2.7) \times 10^{-3} \quad (4.2)$$
$$|V_{td}| = (8.4 \pm 0.6) \times 10^{-3} \quad (4.3)$$

A W boson decays in about 1/3 cases [18] to a charged lepton and a neutrino, the three lepton flavors are produced at approximately equal rate. In about 2/3 cases W boson decays into a quark-antiquark pair, ratios of different $q\bar{q}$ pairs are given by the corresponding CKM matrix elements. In particular, W decay into b quark is suppressed by the CKM mechanism, rendering W boson source of only the light quarks.

4.3.1 Top quark decay width

The decay width of the dominant top quark decay process $t \to Wb$ can be expressed in next-to-leading-order [27] as

$$\Gamma_{t\to Wb} = \frac{G_F m_t^3}{8\sqrt{2}\pi}\left(1 - \frac{M_W^2}{m_t^2}\right)^2\left(1 + 2\frac{M_W^2}{m_t^2}\right)\left[1 - \frac{2\alpha_s}{3\pi}\left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right], \quad (4.4)$$

where G_F is Fermi coupling constant, M_W is the mass of W boson. In the calculation of the top quark decay width the terms of the order m_b^2/m_t^2, α_s^2, and $\alpha_s M_W^2/m_t^2$ were neglected. For a top quark with mass $m_t = 172.5$ GeV/c^2, α_s evaluated at the Z boson scale, predicted value of top quark decay width given by (4.4) is approximately 1.5 GeV.

4.3.2 Top quark decay modes

Top quark decay can be characterized by the number of W bosons that decay leptonically. The three decay signatures follow:

[1] Diagram 2.3 (a) on the right is a correction to the one on the left: the process on the left occurs if the b quark is considered intrinsic, i.e. a distribution function of b quark in proton exists, while the diagram on the right assumes b quark is not intrinsic and only 4 quarks are considered intrinsic (u, d, s, c).

- Semileptonic channel (ℓ + jets) represents about 4/9 of the $t\bar{t}$ decays. Presence of a single lepton with high transverse momentum p_T helps suppressing the SM W+jets and QCD backgrounds. Transverse momentum of a neutrino can be reconstructed in this case, since it is the only source of missing energy (E_T^{miss}) of the signal event.
- Fully hadronic channel represents about 4/9 of the $t\bar{t}$ decays. Both W bosons decay hadronically, which gives rise to at least 6 jets in the event: four light jets from the W boson decay, and two b-jets from decaying top quarks. This channel lacks high p_T lepton, and the signal is difficult to distinguish from the SM QCD multijets production, that is expected to be orders of magnitude larger than the signal. In addition, the large combinatorial background imposes challenge on top mass reconstruction.
- Dileptonic channel represents about 1/9 of the $t\bar{t}$ decays. Both W bosons decay into a lepton and neutrino pair, producing an event with 2 charged leptons, 2 neutrinos, and 2 b-jets. This decay channel is signed by two high p_T leptons and presence of E_T^{miss} in the final state. This mode in theory allows for a clean sample of top quark events, however it is challenging for top quark reconstruction on the ATLAS experiment, since the two neutrinos escape detection.

4.4 ATLAS top quark physics

4.4.1 Top quark properties

Top quark is a $Q = 2/3$, $T_3 = +1/2$ member[2] of a weak-isospin doublet together with the bottom quark. The Standard Model predicts a very short mean lifetime [26] of a top quark to be roughly 5×10^{-25} s, which is about 1/20 of the strong interaction timescale, thus top quark does not manage to form hadrons.

Top mass

The top quark has the highest mass of all observed elementary particles. One of the goals of the ATLAS top quark physics program is the precision measurement of the top quark mass m_t, as it is an essential parameter of the Standard Model. It has been determined with a great precision from top pair events from Tevatron and LHC experiments. The first combination of the Tevatron and LHC measurements of the top quark mass [28] yields

$$m_t = 173.34 \pm 0.27(\text{stat.}) \pm 0.71(\text{syst.}) \text{ GeV/c}^2. \quad (4.5)$$

The Tevatron data corresponds to an integrated luminosity of up to 8.7 fb^{-1} of proton-antiproton collisions of Run II at $\sqrt{s} = 1.96$ TeV. The LHC data[3] corresponds to an integrated luminosity of up to 4.9 fb^{-1} of proton-proton collisions

[2] Q is electric charge, T_3 is the 3rd component of weak-isospin T.

[3] Analysis presented in this Thesis was performed with the LHC Run I data sample collected by the ATLAS experiment at the center-of-mass energy of $\sqrt{s} = 7$ TeV. This chapter provides

at $\sqrt{s} = 7$ TeV. The combination consists of measurements in the different final states: $t\bar{t} \to$ lepton+jets, $t\bar{t} \to$ dilepton, $t\bar{t} \to$ all jets, $t\bar{t} \to E_{\mathrm{T}}^{\mathrm{miss}}$ + jets[4].

The ATLAS Collaboration measured the top mass in the semi-leptonic, dilepton, and fully hadronic channels. The data corresponds to an integrated luminosity of 4.6 fb^{-1} of proton-proton collisions at $\sqrt{s} = 7$ TeV.

The semi-leptonic channel analysis [31] uses a three-dimensional template technique which determines the top quark mass together with a global jet energy scale factor (JSF), and a relative b-to-light-jet energy scale factor (bJSF). The term b-jet refers to a jet originating from a b-quark, while light-jets refer to jets originating from u, d, c, s-quarks, and gluons. Semi-leptonic channel top mass is measured to be

$$m_{\mathrm{t}}^{\mathrm{semi-leptonic}} = 172.33 \pm 0.75(\mathrm{stat.} + \mathrm{JSF} + \mathrm{bJSF}) \pm 1.02(\mathrm{syst.})\ \mathrm{GeV/c^2}. \quad (4.6)$$

The dilepton channel analysis [31] uses a one-dimensional template method with the $m_{\ell b}$ observable, defined as average invariant mass of the two lepton + b-jet pairs in each event. Dilepton channel top mass is measured to be

$$m_{\mathrm{t}}^{\mathrm{dilepton}} = 173.79 \pm 0.54(\mathrm{stat.}) \pm 1.30(\mathrm{syst.})\ \mathrm{GeV/c^2}. \quad (4.7)$$

Combination of the semi-leptonic channel result and the dilepton channel result presented in [31] yields

$$m_{\mathrm{top}} = 172.99 \pm 0.48(\mathrm{stat.}) \pm 0.78(\mathrm{syst.})\ \mathrm{GeV/c^2} \quad (4.8)$$

with a total uncertainty of 0.91 GeV.

In the fully hadronic channel analysis [32] events with at least 6 jets are selected. The analysis uses template fits to the ratio of three-jet to dijet mass. The three-jet mass is calculated from the three jets produced in a top-quark decay. Dijet mass is calculated from the two jets produced in the W boson decay. The fit-obtained top-quark mass measurement is less sensitive to the uncertainty in the energy measurement of the jets. A binned likelihood fit yields a top-quark mass

$$m_{\mathrm{t}}^{\mathrm{fully\ hadronic}} = 175.1 \pm 1.4(\mathrm{stat.}) \pm 1.2(\mathrm{syst.})\ \mathrm{GeV/c^2}. \quad (4.9)$$

In addition to the top quark mass measurement in different channels, the ATLAS Collaboration measured also difference between top mass and antitop mass [33]. The data corresponds to an integrated luminosity of 4.7 fb^{-1} of proton-proton collisions at $\sqrt{s} = 7$ TeV. Semi-leptonic channel events are considered, for each event the mass difference between the top and the antitop quark candidate is calculated. The background contribution is reduced by requirement of 2 b-tags.

---

overview of measurements of top quark properties performed with the ATLAS experiment on full Run I data sample – if a result at $\sqrt{s} = 8$ TeV is available, the $\sqrt{s} = 7$ TeV result is not listed.

[4]The 4th final state $t\bar{t} \to E_{\mathrm{T}}^{\mathrm{miss}}$ + jets measurements were pursued with the CDF experiment: cross-section measurement [29], mass measurement [30]. Such a measurement was not performed at the ATLAS experiment and is not discussed further in this Thesis.

The per-event mass differences are fitted with a maximum likelihood fit, and the top-antitop mass difference yields

$$\Delta m \equiv m_{\mathrm{t}} - m_{\bar{\mathrm{t}}} = 0.67 \pm 0.61(\mathrm{stat.}) \pm 0.41(\mathrm{syst.})\ \mathrm{GeV/c^2}. \tag{4.10}$$

This measurement is consistent with the Standard Model expectation of no mass difference.

Top-antitop charge asymmetry

The ATLAS Collaboration measured the top-antitop charge asymmetry. The data corresponds to an integrated luminosity of 4.7 fb^{-1} of proton-proton collisions at $\sqrt{s} = 7$ TeV. Dileptonic events are selected. Two observables are studied: one based on identified charged leptons, $A_{\mathrm{C}}^{\ell\ell}$, and the other one based on reconstructed $\mathrm{t\bar{t}}$ final state, $A_{\mathrm{C}}^{\mathrm{t\bar{t}}}$. The observable $A_{\mathrm{C}}^{\ell\ell}$ is defined as an asymmetry between positively and negatively charged leptons (electrons and muons) in the dilepton decays of the $\mathrm{t\bar{t}}$ pairs:

$$A_{\mathrm{C}}^{\ell\ell} = \frac{N(\Delta|\eta| > 0) - N(\Delta|\eta| < 0)}{N(\Delta|\eta| > 0) + N(\Delta|\eta| < 0)}, \tag{4.11}$$

where

$$\Delta|\eta| = |\eta_{\ell^+}| - |\eta_{\ell^-}|, \tag{4.12}$$

η_{ℓ^+} (η_{ℓ^-}) is the pseudo-rapidity of the positively (negatively) charged lepton and N is the number of events with positive or negative $\Delta|\eta|$.

Observable $A_{\mathrm{C}}^{\mathrm{t\bar{t}}}$ corresponds to the asymmetry in top quark and antitop quark rapidities:

$$A_{\mathrm{C}}^{\mathrm{t\bar{t}}} = \frac{N(\Delta|y| > 0) - N(\Delta|y| < 0)}{N(\Delta|y| > 0) + N(\Delta|y| < 0)}, \tag{4.13}$$

where

$$\Delta|y| = |y_{\mathrm{t}}| - |y_{\bar{\mathrm{t}}}|, \tag{4.14}$$

y_{t} ($y_{\bar{\mathrm{t}}}$) is the rapidity of the top (antitop) quark and N is the number of events with positive or negative $\Delta|y|$.

The charge asymmetries are measured [34] as

$$A_{\mathrm{C}}^{\ell\ell} = 0.024 \pm 0.015(\mathrm{stat.}) \pm 0.009(\mathrm{syst.}), \tag{4.15}$$

$$A_{\mathrm{C}}^{\mathrm{t\bar{t}}} = 0.021 \pm 0.025(\mathrm{stat.}) \pm 0.017(\mathrm{syst.}). \tag{4.16}$$

This measurement is consistent with the Standard Model predictions.

Top quark polarizations

The ATLAS Collaboration measured top quark polarization in the top-antitop events [35] from proton-proton collisions at $\sqrt{s} = 7$ TeV, using data of an integrated luminosity of 4.7 fb^{-1}. Final states with one or two isolated leptons (electrons or muons) and jets have been considered. Two measurements of the product of the leptonic spin-analyzing power and the top quark polarization,

$\alpha_\ell P$, were performed. An assumption has been made that the polarization is introduced by either a CP conserving (CPC) or a CP violating (CPV) production process. The measurements are in good agreement with the Standard Model prediction of negligible top quark polarization [36]:

$$\alpha_\ell P_{\mathrm{CPC}} = -0.035 \pm 0.014\ (\mathrm{stat.}) \pm 0.037\ (\mathrm{syst.}), \tag{4.17}$$
$$\alpha_\ell P_{\mathrm{CPV}} = -0.020 \pm 0.016\ (\mathrm{stat.})\ ^{+0.013}_{-0.017}\ (\mathrm{syst.}). \tag{4.18}$$

4.4.2 Top quark production mechanisms

Top pair production cross-section

The top-antitop pair production cross-section at $\sqrt{s} = 8$ TeV was measured [37] to be

$$\sigma_{\mathrm{t\bar{t},\ ATLAS}} = 260 \pm 1\ (\mathrm{stat.})^{+20}_{-21}\ (\mathrm{syst.}) \pm 8\ (\mathrm{lumi}) \pm 4\ (\mathrm{beam})\ \mathrm{pb}. \tag{4.19}$$

Assuming $m_\mathrm{t} = 172.5\ \mathrm{GeV/c^2}$, this measurement is in a good agreement with a theoretical NNLO prediction ([38] - [39])

$$\sigma_{\mathrm{t\bar{t},\ theory}} = 253^{+13}_{-15}\ \mathrm{pb}. \tag{4.20}$$

Single top production cross-section

Measurement of cross-section of single top production is a more challenging task in comparison to measurement of cross-section of the $\mathrm{t\bar{t}}$ pair production, due to higher backgrounds and a less characteristic signature of the single top events. The cross-sections of single top production processes are smaller than those of $\mathrm{t\bar{t}}$.

The approximate NNLO precision calculations of the single top production predictions in t-channel, assuming $m_\mathrm{t} = 173\ \mathrm{GeV/c^2}$ and $\sqrt{s} = 7$ TeV [40], follow:

$$\sigma_{\mathrm{t-channel,\ NNLO,\ top}} = 41.7^{+1.6}_{-0.2} \pm 0.8\ \mathrm{pb}, \tag{4.21}$$
$$\sigma_{\mathrm{t-channel,\ NNLO,\ antitop}} = 22.5 \pm 0.5^{+0.7}_{-0.9}\ \mathrm{pb}. \tag{4.22}$$

The first uncertainty is due to scale variation, the second is the PDF uncertainty.

The ATLAS Collaboration measured the single top cross-section in the t-channel [41] at the pp collision energy $\sqrt{s} = 7$ TeV to be in a good agreement with the Standard Model prediction:

$$\sigma_{\mathrm{t-channel,\ ATLAS,\ top}} = 41.9^{+1.8}_{-0.9}\ \mathrm{pb}, \tag{4.23}$$
$$\sigma_{\mathrm{t-channel,\ ATLAS,\ antitop}} = 22.7^{+0.9}_{-1.0}\ \mathrm{pb}. \tag{4.24}$$

Top mass $m_\mathrm{t} = 172.5\ \mathrm{GeV/c^2}$ has been assumed.

Measurement of the s-channel process of the single top production is more difficult than the t-channel process measurement due to a much smaller signal-to-background ratio. Approximate NNLO calculations of the s-channel process

cross-section, assuming $m_{\mathrm{t}} = 173$ GeV/c^2, [42] predict:

$$\sigma_{\mathrm{s-channel,\ NNLO,\ top,\ }\sqrt{s}\mathrm{=7\ TeV}} = 3.17 \pm 0.06^{+0.13}_{-0.10}\ \mathrm{pb}, \quad (4.25)$$

$$\sigma_{\mathrm{s-channel,\ NNLO,\ top,\ }\sqrt{s}\mathrm{=10\ TeV}} = 5.16 \pm 0.09^{+0.20}_{-0.14}\ \mathrm{pb}, \quad (4.26)$$

$$\sigma_{\mathrm{s-channel,\ NNLO,\ top,\ }\sqrt{s}\mathrm{=7\ TeV}} = 1.42 \pm 0.01^{+0.06}_{-0.07}\ \mathrm{pb}, \quad (4.27)$$

$$\sigma_{\mathrm{s-channel,\ NNLO,\ top,\ }\sqrt{s}\mathrm{=10\ TeV}} = 2.48 \pm 0.02^{+0.09}_{-0.13}\ \mathrm{pb}. \quad (4.28)$$

The first uncertainty is due to scale variation, the second is the PDF uncertainty, at the 90% CL.

The ATLAS Experiment performed search for s-channel single top-quark production in pp collisions at $\sqrt{s} = 8$ TeV. Analysis of data corresponding to integrated luminosity of 20.3 fb^{-1} leads to an upper limit on the s-channel single top-quark production cross-section of 14.6 pb at the 95% CL [43]. The data fit gives a cross-section consistent with the Standard Model expectations:

$$\sigma_{\mathrm{s-channel,\ ATLAS\ fit}} = 5.0 \pm 4.3\ \mathrm{pb}. \quad (4.29)$$

The approximate NNLO calculations of the cross-section of a single top-quark production in the Wt channel [44] in pp collision, assuming $m_{\mathrm{t}} = 172.5$ GeV/c^2 and $\sqrt{s} = 7$ TeV, predict

$$\sigma_{\mathrm{Wt\ channel,\ NNLO}} = 15.74^{+1.06}_{-1.08}\ \mathrm{pb}. \quad (4.30)$$

The ATLAS Collaboration presented evidence for the associated production of a W boson and a top quark using 2.05 fb^{-1} of pp collision data at $\sqrt{s} = 7$ TeV [45]. The corresponding cross-section measurement is in a good agreement with the Standard Model prediction:

$$\sigma_{\mathrm{Wt\ channel,\ ATLAS}} = 16.8 \pm 2.9\ (\mathrm{stat.}) \pm 4.9\ (\mathrm{syst.})\ \mathrm{pb}. \quad (4.31)$$

Single top production via flavor-changing neutral currents

A search for the production of single top-quarks via flavor-changing neutral-current has been performed [46], using data from proton-proton collisions at $\sqrt{s} = 7$ TeV of an integrated luminosity of 2.05 fb^{-1}. Semileptonic top quark decay candidate events are classified as a signal or a background-like events with help of several kinematic variables. The candidate events serve as an input of a neural network. No signal has been observed in the neural network output distribution, if Bayesian upper limit is placed on the production cross-section. The observed upper limit at 95 % CL on the cross-section multiplied by the branching fraction of t $\rightarrow$ Wb is measured to be

$$\sigma_{\mathrm{qg}\rightarrow\mathrm{t}} \times \mathcal{B}(\mathrm{t} \rightarrow \mathrm{Wb}) < 3.9\ \mathrm{pb}. \quad (4.32)$$

Search for $\mathrm{t\bar{t}}$ resonance in semi-leptonic events with highly boosted top

A search for resonant production of high-mass top-quark pairs [47] with lepton+jets final state events, designed for the topology arising from the decay of

highly boosted top quarks, is performed on 2.05 fb^{-1} of proton-proton collisions data at $\sqrt{s} = 7$ TeV. The observed ttbar invariant mass spectrum is found to be compatible with the Standard Model prediction and 95% CL upper limits are derived on the ttbar production rate through new massive states. An upper limit of 0.7 pb is set on the production cross section times branching fraction of a narrow 1 TeV resonance. A Kaluza-Klein gluon with a mass smaller than 1.5 TeV is excluded.

Top-antitop pair differential cross-section

Measurement of $t\bar{t}$ pair differential cross-section is discussed in Chapter 10.

4.4.3 Search for new physics

Top quark spin correlations and squark search

A search for pair production of squarks [48] with masses close to the top quark mass decaying to predominantly right-handed top quarks and a light neutralino, the lightest supersymmetric particle, has been performed from proton-proton collisions at $\sqrt{s} = 8$ TeV, using data of an integrated luminosity of 20.3 fb^{-1}. Top squarks with masses between the top quark mass and 191 GeV are excluded at the 95% CL.

Measurement of correlation between the top and the antitop quark spins was extracted from dilepton $t\bar{t}$ events. The difference in azimuthal angle between the two charged leptons in the laboratory frame was used. The measurement was performed on the same dataset as the top squark search. In the helicity basis the measured degree of correlation corresponds to

$$A_{\text{helicity}} = 0.38 \pm 0.04 \tag{4.33}$$

This measurement is in agreement with the Standard Model prediction.

Search for $t\bar{t}$ resonance in semi-leptonic events and Z'

A search for new particles that decay into top-antitop quark pairs $t\bar{t}$ [49] is performed with the ATLAS experiment an integrated luminosity of 4.7 fb^{-1} of proton–proton collision data collected at $\sqrt{s} = 7$ TeV. In the $t\bar{t} \to$ WbWb decay, the lepton plus jets final state is used, where one W boson decays leptonically and the other hadronically. The $t\bar{t}$ system is reconstructed using small- and large-radius jets. The large-radius jets were supplemented by a jet substructure analysis.

Local excesses in the number of data events compared to the Standard Model expectations were searched in the $t\bar{t}$ invariant mass spectrum. No evidence for a $t\bar{t}$ resonance has been found and 95% CL limits on the production rate are determined for massive states predicted in two benchmark models. The upper limits on the cross section times branching ratio of a narrow Z' resonance range from 5.1 pb for a boson mass of 0.5 TeV to 0.03 pb for a mass of 3 TeV. A narrow leptophobic topcolor Z' resonance with a mass below 1.74 TeV is excluded. Limits are also derived for a broad color-octet resonance with $\Gamma/m = 15.3\%$. A

Kaluza–Klein excitation of the gluon in a Randall–Sundrum model is excluded for masses below 2.07 TeV.

Search for $t\bar{t}$ resonance in fully hadronic events and Z'

A search for resonances produced in $\sqrt{s} = 7$ TeV proton-proton collisions and decaying into top-antitop quark pairs [50] is based on data of an integrated luminosity of 4.7 fb^{-1}. This analysis takes into account events where the top quark decay produces two massive jets with large transverse momenta. Jets originating from top quark and jets originating from light quarks and gluons are separated by two techniques that rely on jet substructure. Additionally, each massive jet is required to have evidence of an associated b quark decay. The data is consistent with the Standard Model.

Limits on the production cross section times branching fraction of a Z' boson and a Kaluza-Klein gluon resonance can be set. These limits exclude, at the 95% CL, Z' bosons with masses 0.70 - 1.00 TeV as well as 1.28 - 1.32 TeV, and Kaluza-Klein gluons with masses 0.70 - 1.62 TeV.

Search for tb resonance and W'$_{\mathrm{R}}$

A search for tb resonance [51] in 1.04 fb^{-1} of proton-proton collisions data at $\sqrt{s} =$ 7 TeV has been performed. Events with a lepton, missing transverse momentum, and two jets have been selected and the invariant mass of the corresponding final state has been reconstructed. The search exploits the shape of the tb invariant mass distribution compared to the expected Standard Model backgrounds.

This search was benchmarked with a model of right-handed W'$_{\mathrm{R}}$ with Standard Model-like couplings. No statistically significant excess of events is observed in data. The upper limits on the cross-section times the branching ratio of W'$_{\mathrm{R}}$ resonances at 95 % CL lie in the range 6.1 – 1.0 pb for W'$_{\mathrm{R}}$ masses ranging from 0.5 to 2.0 TeV, imposing limit on the lower bound of the allowed mass of W'$_{\mathrm{R}}$ at the 95% CL to be

$$m_{W'_{\mathrm{R}}} > 1.13 \text{ TeV}. \tag{4.34}$$

5. ATLAS Computing

5.1 Data processing and analysis for ATLAS Top Quark Physics

The ATLAS detector, described further in Chapter 2, consists of four major components: inner detector, calorimeter, muon spectrometer, and magnet system. The ATLAS detector is designed to register nearly one billion of proton-proton collisions per second, that generate combined data volume of more than 60 PB/s. Nevertheless, only a very small fraction of these events contains intriguing characteristics that may suggest that a new discovery is on the road ahead, waiting to be uncovered.

In order to reduce this enormous data flow to levels manageable by the available computing resources the ATLAS experiment uses a specialized multi-level computing system, the trigger system, to read-out and select events with differentiating characteristics. Only the events interesting for physics analyses make the cut and pass through the trigger system. The trigger system filters 100 events per second out of one billion others.

Data describing the physics events that made it through tough trigger selection are then channeled from the detectors to the permanent storage by the ATLAS Data Acquisition System (DAQ) [52].

The first-pass processing of the raw data received from the ATLAS DAQ is performed on the ATLAS Tier-0 [53] facility at CERN. The CERN Tier-0 facility has to be highly-available, and provide short turn-around and response times. The processed raw data and derived data are archived on the Tier-0 mass storage system.

Derived data products are distributed from Tier-0 to the Tier-1 and Tier-2 centers around the world, to be further processed and analyzed at over 120 computing centers (Tier-1, Tier-2, and Tier-3, [54]) of the ADC resources that are distributed world-wide.

The ATLAS Distributed Data Management (DDM) system [55], [56] covers all activities that are related to data storage and distribution: data distribution following the rules prescribed by the ATLAS Computing Model and ad-hoc requests from the ATLAS physicists, data consistency checks, obsolete data deletion, and operations-driven activities such as data loss declaration upon degradation or failure of the storage system at any of the computing centers. At the beginning of LHC Run II (June 2015) the DDM system manages about 150 PB of data (about 25 PB of raw data).

Data processing consists of running Monte Carlo simulations, and processing the raw and derived data. It is performed by the ATLAS Physics groups or by individual ATLAS physicists, in order to leverage the available computing resources efficiently, and to grant all the analyzers access to the data of an expected quality, resulting from prior expertise and physics performance groups recommendations, as soon as possible after the events have been collected by the ATLAS detector.

Data processing and analysis is efficiently managed by the ATLAS workload management system PanDA (Production and Distibuted Analysis) [57], [58]. Central and group-specific data processing is initiated in the Production System [59], which interacts with different ADC products (DDM, WMS, meta-data catalog) and accommodates the data processing workflow. PanDA provides ATLAS physicists with command-line client tools to submit and manage their distributed analysis, and a monitoring application to help track the progress, and debug issues of the analysis.

Data processing and analysis is aggregated in tasks. Aggregation is defined by the task submitter, i.e. physicists. The task submitter defines what is the nature of the task activity, names input and output data, and specifies executable and its configuration. The task submitter can choose between specifying from what resources the task will benefit, and leaving the decision to the scheduler component of the WMS.

Selection of proper executable and its configuration is connected to the task activity (e.g. physics analysis, data derivation, MC simulation, data merging, SW installation and validation). The ATLAS Collaboration developed ATLAS offline software Athena [60], that enables access to all objects interesting to physics analysis, from different data formats.

During the LHC Run I and LHC Long Shutdown 1 (LS1) many analysis groups developed a variety of custom analysis packages. The custom packages helped the analysis teams to pursue the analysis goals with a particular format of derived data. Progress in the custom analysis packages development reflected in further development of the Athena framework during the LHC Long Shutdown I, and led to a unified way of storage of analysis object data.

The custom analysis packages benefit from the ROOT analysis framework [61]: ROOT installation is available for all the operating systems running on nowadays personal laptops or workstations, and distributed computing resources. The analysis packages in addition benefit from the common flat-ntuple data format, containing only slimmed events with properties that are relevant for a particular analysis in branches. User friendly access to the data and tuned analysis packages contribute to the much needed fast turn-around to produce high quality results of the ATLAS Physics program.

5.2 ATLAS Computing Resources

5.2.1 Computing Resources Overview

The ADC [62] infrastructure is a complex and heterogeneous system: The ATLAS grid resources (CPU resources, storage systems, network links) are spread over more than 120 computing centers distributed worldwide. ATLAS grid computing centers host over 150 PB storage either on disk or tape systems, with different flavors of storage systems, and heterogeneous CPU resources available to accommodate over 160k job slots. ATLAS grid sites are organized within three different flavors of grid: EGI, OSG, and NorduGrid.

5.2.2 Tiers of resources

ATLAS benefits from the World-wide LHC Computing Grid (WLCG) collaboration. The WLCG is made of four layers, the Tiers [54]: Tier-0, Tier-1, Tier-2, Tier3. Each tier provides a specific set of services.

- Tier-0 is located at CERN. Tier-0 is responsible for the safe-keeping of the raw data, and for the first-pass reconstruction, as well as distribution of raw data and reconstruction output to the Tier-1s and Tier-2s, and reprocessing of data during LHC downtimes.
- These are 13 large computer centres with sufficient storage capacity and with round-the-clock support for the Grid. The ATLAS experiment used 10 of them during the LHC Run I, and 11 of them during the LS1 and LHC Run II. The Tier-1 centers are responsible for the safe-keeping of a proportional share of raw and reconstructed data, large-scale reprocessing and safe-keeping of corresponding output, distribution of data to Tier-2s and safe-keeping of a share of simulated data produced at these Tier-2s.
- The Tier-2s are typically co-located with universities and other scientific institutes, which can store sufficient data and provide adequate computing power for specific analysis tasks. They handle analysis requirements and proportional share of simulated event production and reconstruction.
- The Tier-3s are smaller sites without a pledge to WLCG nor the ATLAS experiment. Their sizes and provided services vary significantly.

5.2.3 Clouds

During the LHC Run I the ATLAS sites were grouped in 11 clouds, during LS1 there were 12 clouds[1]. Each cloud is a set of grid sites. The most powerful site in each cloud is a Tier-1 site, with exception of the CERN cloud, where the most powerful site is a Tier-0 site, accompanied by Tier-3 sites. There are usually several Tier-2 sites as a part of the cloud. Cloud may host also Tier-3 grid sites.

5.3 ATLAS Distributed Computing

ATLAS Distributed Computing [63], [62] operates computing resources of the ATLAS experiment distributed world-wide, and covers a set of projects and a set of activities that enable utilization of such resources. The ATLAS Computing Model [64], [65] describes design and goals of the ATLAS Computing.

ADC activities consist of ADC Operations, Integration & Commissioning, Monitoring & Accounting, DDM Operations, Databases, Computing Shifts, Central Services & Security, Virtual Organization (VO) Management, SW deployment, and Data Processing Activities: Tier-0 Operations, Distributed Analysis and Production, Cloud Computing, High-Performance Computing (HPC).

[1] As of 2015 the twelve ATLAS clouds are CA, CERN, DE, ES, FR, IT, ND, NL, RU, TW, UK, US.

ADC projects consist of ADC Development, DDM Development, Metadata, AGIS, Data Processing Projects: Tier-0 Development, Event Service, Production System, Computing Facilities, and Workload Management System PanDA.

The next chapters provide more details about activities and projects to which the author of this Thesis had the opportunity to contribute, either as a developer, or as an activity expert or coordinator. Section 5.4 details contribution to the Computing Shifts. Chapter 6 covers contributions to the Monitoring and Accounting activities. Chapter 7 summarizes contributions to Automation of ADC Operations. Chapter 8 provides overview of contributions to the development and operation of the Workload Management System PanDA. As noted in the Preface, all these contributions in the ATLAS Computing domain contributed to the great success of the ATLAS Physics program.

5.4 ATLAS Distributed Computing Operations

The ADC Operations team has to be able to easily identify issues with the infrastructure, and to address these issues, in order to provide a high quality of service to the ATLAS Collaboration. Such a challenging task, to spot and recognize what is going on with the ADC resources, and to address or escalate the issue, is performed by the various teams that take part in the ADC Operations activity: activity and project experts, Computing Shifters, and Cloud support teams[2]. The ADC Operations teams have a set of monitoring tools and functional test probes at their disposal.

In this Section we describe several areas of operation of ADC resources. From 2007 to 2015 author of this Thesis contributed to the following projects:

1. ADC operations: contributed to all the ADC shifts, gained experience in various aspects of ADC operations, and became one of the ADC experts. Coordinated the ATLAS Distributed Computing Operations Shift (ADCoS, [66]) team, and coordinated and was responsible for all the ADC shift teams.

2. ADC monitoring: contributed to the activity as a developer and coordinator. These contributions are described in Chapter 6.

3. ADC automation: contributed to the automation effort of the ADC Operations. These contributions are described in Chapter 7.

4. was responsible for operations and security of the ATLAS Central Services located at CERN.

Projects were presented at several international conferences [66], [67], [68], and [69].

[2]Cloud support teams are composed of local site administrators who also happen to be experts in ATLAS Distributed Computing and are actively sharing their expertise with the other team members.

5.4.1 ATLAS Distributed Computing Shifts during LHC Run I

This section describes structure and goals of the ADC shift teams during the LHC Run I. In order to provide reliable and efficient quality of service to the ATLAS Collaboration, the system has to be monitored and its functionality regularly tested. There are three areas covered by three different teams of shifters, who work from different places with different distance from an activity expert: the Distributed Analysis Support Team, the Comp@P1 Shift Team, and the ATLAS Distributed Computing Operations Shift Team.

In this section we describe structure of the Shift Teams, their mutual communication, issue reporting and lessons learned during the LHC Run I data taking era with the ATLAS Experiment.

Organization

The ADCoS Team was organized in 2008 to support 24/7 simulation production of the ATLAS experiment, data reprocessing and data management operations at more than a hundred of computing centers located world-wide. In order to support the ATLAS Physics Analysis the ATLAS Distributed Analysis Support Team (DAST) was formed in 2008. In addition, a team to support data processing at Tier-0 and data export from Tier-0 to Tier-1s, the Comp@P1, was formed in 2009. The ADCoS and DAST teams are distributed world-wide and work remotely from their home institutions.

The User analysis activity is monitored by the DAST. The DAST provides first line support to grid-related questions of the ATLAS physicists.

The data processing at Tier-0 and data export from Tier-0 to Tier-1s and calibration Tier-2s have been monitored by the Comp@P1 Shift Team located at CERN in the ATLAS Control Room.

The Grid production activity, including the data transfers among Tier-1s, Tier-2s, and Tier-3 sites, data processing, data reprocessing, Physics Group production, and Monte Carlo Simulations, and Functional Testing, is monitored by the ADCoS Team.

The ADCoS and DAST are world-wide distributed teams organized in three different time zones, natural to the location of the team members: the Asia-Pacific time zone (00-08 hrs CE(S)T), the Europe time zone (08-16 hrs CE(S)T), and the Americas time zone (16-24 hrs CE(S)T). The Comp@P1 shift, which is located at CERN, follows similar time zone scheme.

The WLCG sites supporting the ATLAS Virtual Organization are organized in 11 areas called clouds. Each cloud organizes a team of experts, the Cloud support team, which can provide first line support for the site administrators in ATLAS matters.

In addition to the ATLAS Distributed Computing shift teams and the Cloud support teams, there is one ATLAS Distributed Computing Manager on Duty, who can address or escalate issues with the core services of the ATLAS Distributed Computing.

Training

The Shifter training consists of three phases: the shift tutorial, self-study of shift instructions, and the supervised "Trainee" shift.

The shift tutorials are organized with different frequency reflecting the current needs of the corresponding shift team. The tutorial frequency varies from once per year to once per month for the ATLAS Distributed Computing Shifts.

Each ATLAS Distributed Computing shift team has its own documentation page to describe the shift workflow, to describe procedures, and to propose solutions or a list of steps to be followed in order to escalate the issue to activity expert attention, for every category of issues. The instructions point to every monitoring [70] or command-line tool, which can provide useful information in order to debug an issue. The instructions are available to every member of the ATLAS Experiment. The Shift instructions are regularly updated. The self-study of shift instructions gives the Trainee shifter a first hands-on glance of the Shift activity and duties.

During the supervised Trainee shift the Trainee Shifter can perfect her/his skills in issue troubleshooting, so that she/he is fairly familiar with the Shift duties for his solo shift. At any point of the training or the solo shift the Shifters are encouraged to ask questions to the Shift experts.

Communication

The ATLAS Distributed Computing Shifters use different technologies to communicate with others on duty. Each of the three Shift Teams use mailing lists for intra-/inter- shift communication. In addition to e-mail communication the Shift teams can instantly communicate in the ATLAS Distributed Computing Virtual Control Room based on the XMMP protocol. Every message to a mailing list or to the Virtual Control Room is archived, therefore can be used as a knowledge bank for the Shift Team.

The ATLAS Distributed Computing Shift Teams use Global Grid User Support (GGUS) [71] tool to escalate issues with services provided by the sites to the site administrators. In addition, when a site is failing to perform well in an activity, the site is excluded from that activity and the exclusion and subsequent actions are recorded in a ticket[3]. Any time an ATLAS Distributed Computing Shifter takes an action, such an action is then recorded in an electronic logbook. Every ATLAS Distributed Computing Shifter, expert, or Cloud Squad member can then access a full record of manual actions which affect his/her area of interest.

In the first two years of data taking there were approximately 200 GGUS tickets per month created to address issues at ATLAS sites, over 97 % of the GGUS tickets were created by the ATLAS Distributed Computing Shifters. The ATLAS Distributed Computing Operations are trying to increase level of automation of the repetitive tasks, in order to decrease the need for repetitive manual operations

[3] A ticket is a bug report in a ticket tracking system, aimed to the activity expert or developer. There were several different ticket tracking systems used during LHC Run I and LS1.

taken by site administrators, experts, or Shifters.

Issue Reports and Follow-ups

Every failure of any subsystem of the three areas of activities is reported so that proper action is taken by the service provider and the issue is fixed. Software bugs were reported to the responsible experts via software ticket trackers[4]. GGUS system is used to report site-related issues to the sites, and as a communication channel between ADC Operations and sites.

Anytime an ADCoS Shifter takes an action, this action is to be recorded as an eLog entry. The ADC eLogbook contains history of any kind of action taken by ADC Operations members. At the end of the daily shift ADCoS Shifter submits a daily report with the most important issues of the day so that the following ADCoS Shifter on duty is informed and can continue in problem chasing from the point where previous ADCoS Shifter stopped. The daily reports are aggregated in weekly reports presented at ADCoS Weekly meetings.

Summary: ATLAS Distributed Computing Shifts

The ATLAS Distributed Computing Shift Teams have been ensuring reliability and quality of service of the ATLAS Distributed Computing system since the start of LHC Run I data taking. ADC Shift Teams monitor ADC subsystems 24 hrs per day, 7 days per week. The shift teams report or escalate observed issues, and perform service tests when necessary.

5.5 ATLAS Distributed Computing Monitoring at a glance

ATLAS Distributed Computing Monitoring tools cover every aspect of the day-to-day work of ADC Operations. In this section we describe how the ADC activities were monitored, as of May 2012, after the full 2 years of LHC data taking. As the requirements on ADC evolve, the monitoring tools respond with similar evolution in order to address all operational needs. The ADC Monitoring suite consists of real-time monitoring tools, and of long-term accounting tools. In this section we describe the state-of-the-art of monitoring of key areas of ATLAS Distributed Computing: ATLAS Tier-0 at CERN, Distributed Data Management, data processing, status of database system essential for ATLAS Computing, and status of sites and services hosted at sites.

The real-time monitoring tools address needs of the day-to-day work of the ADC Operations team – the activity experts, the distributed computing shifters, and the site administrators.

The accounting tools [72] aggregate summaries of the real-time information and store it for long-term series analysis and visualization. The accounting tools

[4]Originally through the LCG Savannah portal during the LHC Run I, which was superseded by the Jira portal during LS1.

provide visualization of resource utilization for the ATLAS Management and ATLAS National contacts.

The architecture of the ADC Monitoring tools has been standardized to a multi-layer design with clear separation of data access and visualization. Such approach favors modular design of software products, and allows for long-term maintainability.

The data access layer exposes data from possible various sources in a predefined structure in different formats, e.g. JSON (JavaScript Object Notation), XML (Extensible Markup Language), CSV (comma-separated values), utilizing variety of open-source or in-house made frameworks, e.g. Dashboard framework [73], PanDA (Production and Distributed Analysis System) web platform [74]. Monitoring data is stored primarily in Oracle database at CERN.

The visualization layer exposes monitoring data in form of graphics, or tables using various JavaScript libraries, such as jQuery and its plug-ins, `xbrowse` [75] or `hBrowse` [76] JavaScript libraries. The multi-layer design not only enables unified visual identity across all ADC Monitoring, but opens a variety of ways to share content.

5.5.1 ATLAS Tier-0 at CERN

Early in the LHC Run I the raw data acquired with the ATLAS detector were recorded at a nominal rate of 200 Hz. The average uncompressed event size is 1.6 MB; after compression it is 0.8 MB. Data is stored on tape directly at the Tier-0. In 2011, ATLAS decided to increase the rate of event recording up to 400 Hz. The raw data recorded with the ATLAS detector is processed at the Tier-0 and then the raw data and the derived data are exported from the CERN Tier-0 to Tier-1 centers and several calibration Tier-2 centres. The data processing activity is monitored with the *conTZole* dashboard [63]. In Fig 5.1 the view of the overall status of the data processing at Tier-0, status of the underlying batch system, and status of data registration to the ATLAS Distributed Data Management system is shown. In addition, several views detailing the data processing progress are provided. The *conTZole* dashboard is one of the many monitoring tools based on the Experiment Dashboard framework [73].

5.5.2 Distributed Data Management

Since the very first collisions at 900 GeV in 2010 and at 7 TeV in 2011 the ATLAS data reconstruction at Tier-0 has produced over 10 PB of raw and derived data. This data is registered to the ATLAS Distributed Data Management system (DDM) [55] and distributed to the ATLAS sites worldwide. The data distribution status and progress is monitored with the DDM dashboard. In the first 2 full years of LHC data taking the DDM dashboard UI has been significantly improved to address several requests, e.g. ability to visually distinguish between source and destination failures (source/destination matrix replaced the destination-oriented list), introduction of flexible filtering based on string patterns, and flexible grouping of DDM endpoints. The backend of the DDM dashboard had to cope with

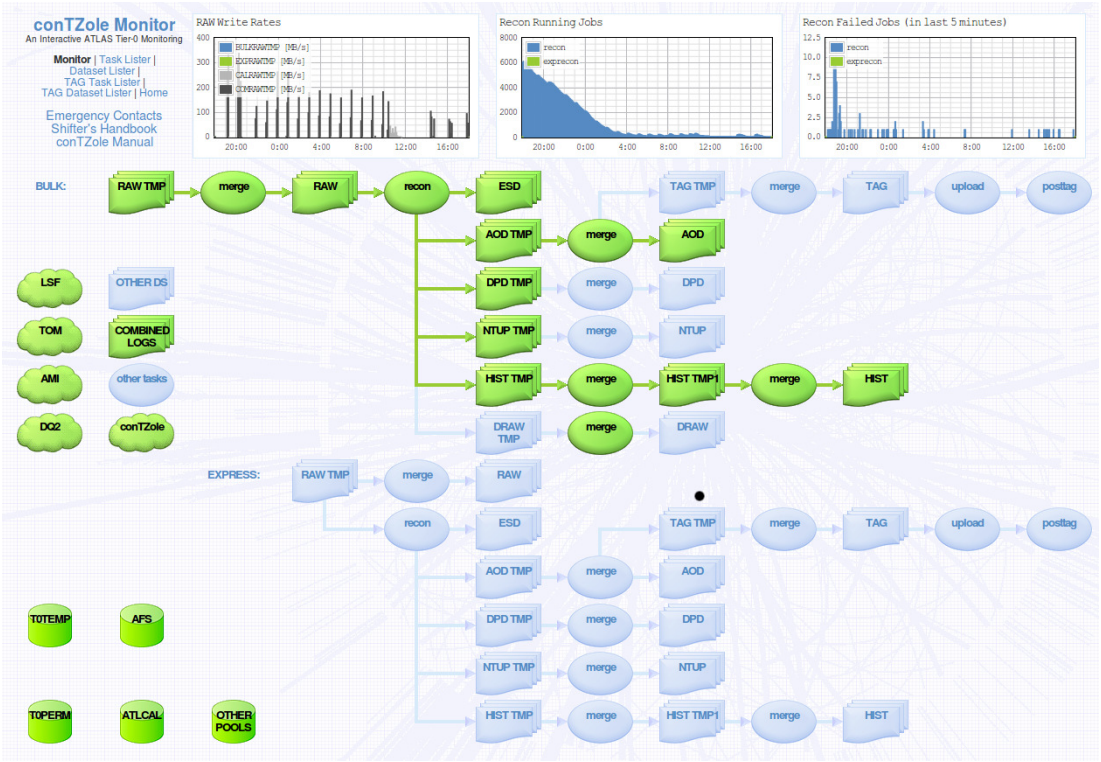


Figure 5.1: Overview of the data processing at the ATLAS Tier-0 at CERN shown by the *conTZole* dashboard. The top-left plot shows uncompressed RAW data write rates: physics streams (blue), express stream (green), calibration streams (grey), and commissioning data (black). The top-middle and top-right plots show the evolution of the number of running and failed reconstruction jobs.

larger data volumes, and generation of the visualization data is faster and scales better. Currently, the DDM dashboard is able to process and visualize the amount of transfer events (start/end) with a rate of 100 Hz. The average callback load is around 30 Hz. In Fig. 5.2 the source/destination matrix is shown. This matrix shows the efficiency of the transfers between each pair of groups of the storage endpoints, the throughput of such transfers and the number of successes and errors. It also offers a view to show data on the registration process: number of datasets, number of files, and number of errors. Sources and destinations are grouped. The first level of grouping is per cloud (group of geographically close Tier-2 and Tier-3 sites which are connected to the same Tier-1), followed by grouping endpoints on the level of sites. The most detailed group lists all DDM endpoints (DDM spacetokens) provided by a site. Thanks to the elegant way of grouping storage endpoints, it is very straightforward to spot a failing endpoint and to see whether that endpoint is failing to transfer data as a source, or as a destination, or as both. The DDM dashboard matrix can be adjusted in terms of time interval of the presented data, filtering sites as sources or destinations, based on filtering by cloud, tier, site name, or spacetoken, as well as filtering by one or more transfer activities.

Each cell of the transfer matrix leads to a special view with details about the

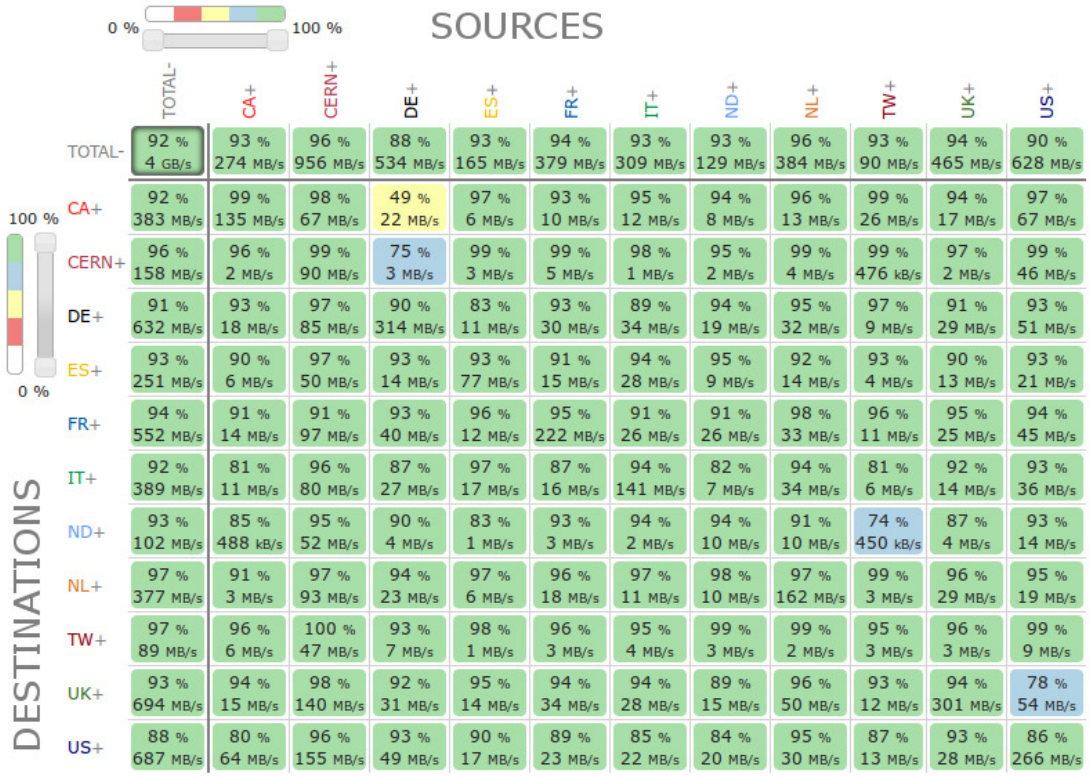


Figure 5.2: ATLAS Distributed Data Management transfer matrix. The current transfer efficiency and throughput between each group of storage endpoints is shown.

transfers. For erroneous transfers, this detailed view provides groups of errors determined by the error pattern, number of occurrences, and a detailed list of attempts of failed transfers with a list of transfer metadata such as file transfer placement time, file name, exit code and status of the transfer, GUID, counter of attempts, transfer requester identifier, source and destination SURL, transfer ID, FTS channel, error message, and activity name. Such a set of detailed information is very useful to provide in a bug report, especially to clarify if the issue is at the source, at the destination or just for a pair of sites (usually a network issue).

The DDM dashboard provides not only textual information, but also a set of plots with history trends. There are sets of Efficiency, Throughput, Transfer Successes, and Transfer Failures history plots available for Sources, Destinations, and Activities. The overall throughput of the DDM system from early 2010 to mid-2012 is shown in Fig. 5.3.

The DDM dashboard is based on the Experiment Dashboard framework. The DDM dashboard uses the *xbrowse* framework [75] to visualize data. Plots are made with the *Highcharts* library. A WLCG Transfers Dashboard used across the LHC experiments is built using the same architecture as the DDM dashboard.

If data transfer or Grid jobs indicate a missing, inaccessible or corrupted file at the source, data consistency checks are performed. The consistency monitoring consists of a set of tables listing DDM endpoints with suspicious data. For each

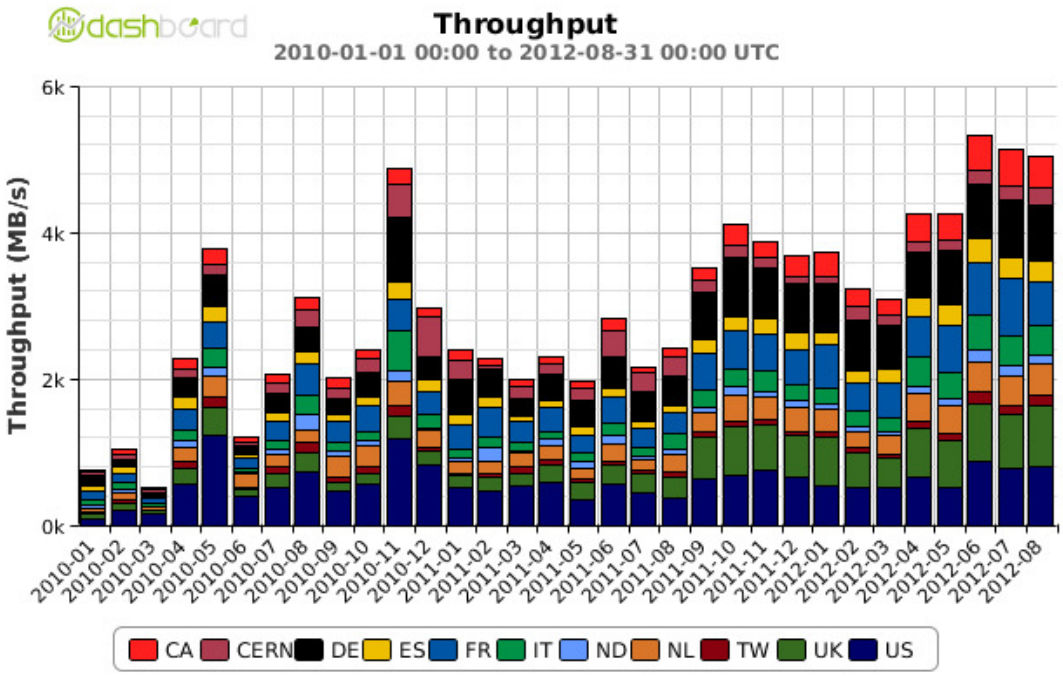


Figure 5.3: Monthly averaged throughput of the ATLAS Distributed Data Management system since beginning of 2010.

problematic file, a list of the SURL, GUID and dataset name is provided and the status of the recovery procedure is displayed.

In order to increase the number of dataset replicas according to its access frequency, data popularity is monitored. Weekly reports with plots of the number of file accesses vs. the number of file replicas are then provided, as well as time series plots for dataset popularity per dataset project name or per user.

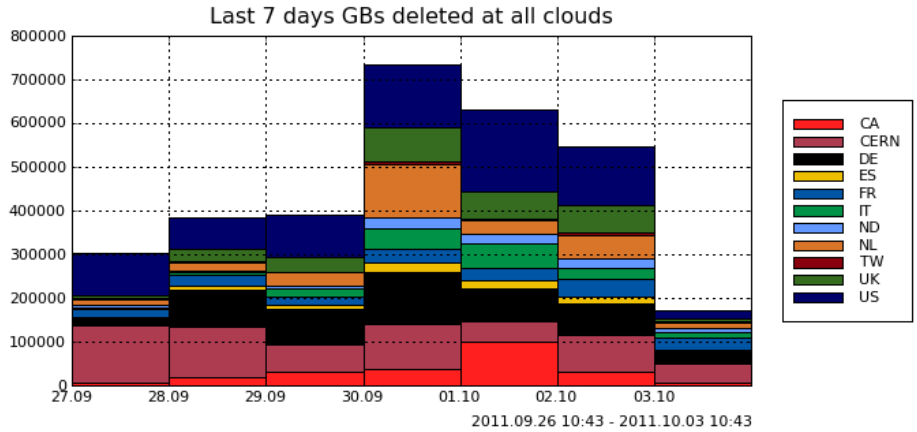


Figure 5.4: ATLAS DDM deletion activity during a busy week is shown: amount of data deleted – peak around 750 TB deleted on 30th September 2011.

The DDM deletion activity is mainly triggered by transient data cleaning or to reduce the number of dataset replicas. There are on average 3 M files to be deleted

every day. In Fig. 5.4 the deletion activity during a busy week is shown: number of files deleted – peak around 6 M deleted files per day on 2nd October 2011, and amount of data deleted – peak around 750 TB deleted on 30th September 2011.

In order to prevent ATLAS from oversubscribing data to storage endpoints ATLAS monitors how much space the site pledged to the experiment and how much space is remaining at a site (Fig. 5.5). Storage accounting reports are available to the ATLAS Collaboration.

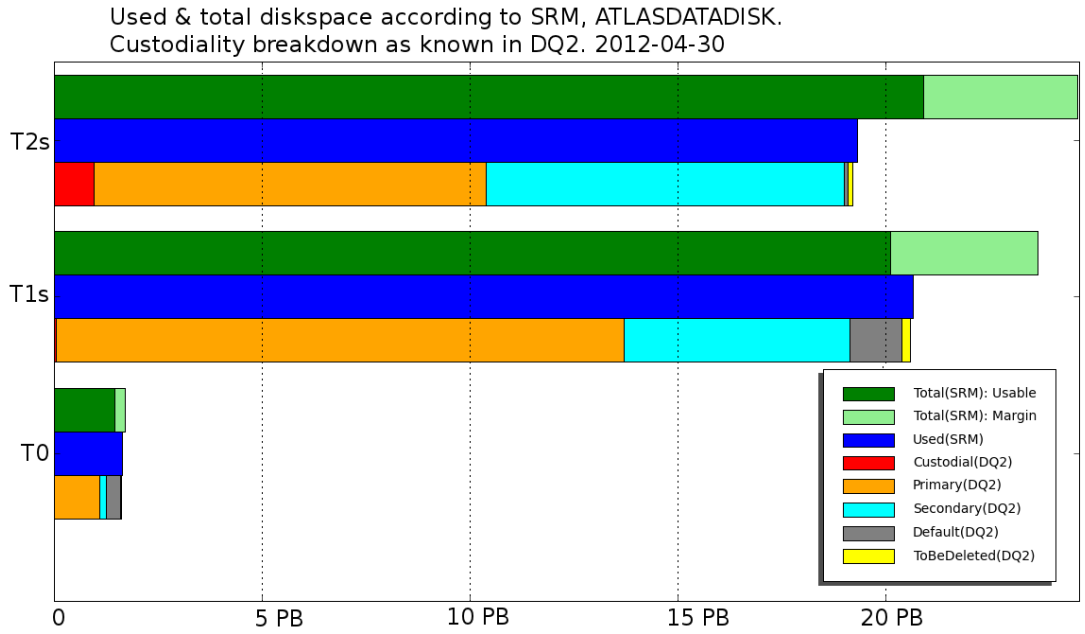


Figure 5.5: ATLASDATADISK usage as of April 2012. Custodiality of data and Tier category of sites is distinguished.

There are three possible ways to replicate data in the ATLAS DDM systems. Firstly, the data is replicated centrally according to the pledged resources availability. Such a replication follows WLCG Memorandum of Understanding (MoU) pledges of the sites. The second possibility is a replication triggered by an active subscription of data by physics groups or by physicists to a particular grid site. The third possibility is PanDA Dynamic Data Placement (PD2P) [74], based on data popularity and availability in several places. PD2P is triggered by user jobs running on the grid. The aim of all data replication policies is to make data available to the ATLAS physicists in a timely manner. The data transfers are monitored with the DDM dashboard. In addition, accounting reports and weekly activity reports are provided.

Functional testing of the DDM system at a marginal level is an important activity. DDM functional tests [77] are submitted regularly with 1 week period probing availability of a storage endpoint to write and read data with files of different sizes ranging from several MB (small files), through several hundred MB (medium files), to several GB (large files). In addition to storage sanity checks, such a functional test provides information about network throughput between each pair of disk storage endpoints used by ATLAS. In Fig. 5.6 an example of

network throughput evolution between the ATLAS Tier-0 at CERN and an ATLAS Tier-1 site is presented. More precise network link properties measurement can be accomplished with the perfSONAR-PS [78] infrastructure.

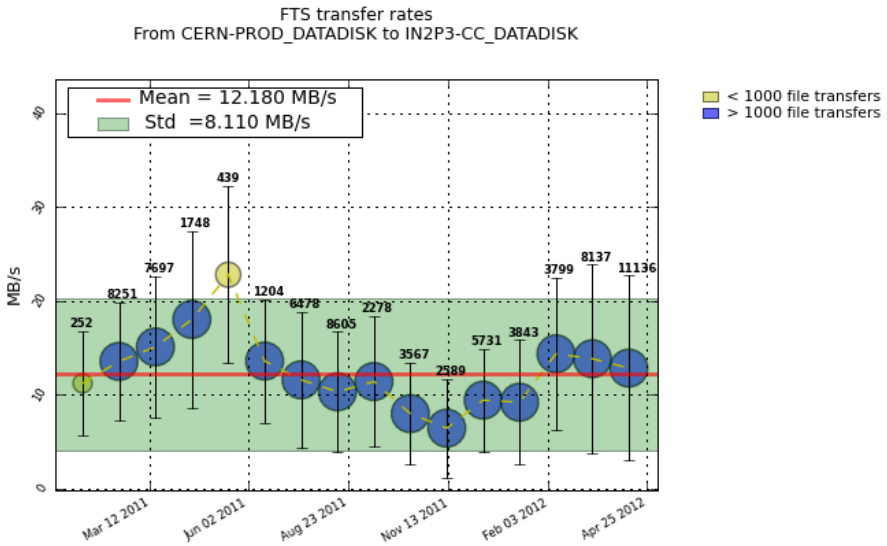


Figure 5.6: Time evolution of the network throughput between 2 DDM endpoints. Measured with large files (size of several GB).

5.5.3 Data processing

The ATLAS Experiment uses PanDA [57] as the workload management backend. ATLAS is able to run over 160k production and analysis jobs simultaneously on the grid, averaging 700k-1.5M jobs per day. There are three points of view for instant monitoring of data processing: site-centric, production task-centric, and user task-centric.

The site-centric view is provided by the Panda Monitor [74]. Panda Monitor provides detailed information about the workload management backend. Primarily, it provides information about jobs, and about site performance with respect to running jobs from the Production and Analysis activities.

The task-centric monitoring for production and analysis activities [79] comes with many useful summaries to monitor the progress of a task. Summaries are provided as a comprehensive table or as a plot. The comprehensive table contains basic task description information, task requester, task progress information (duration, progress status, processing time). The summary plots describe key metrics of the task progress: evolution of job status with failures distinction, evolution of number of job attempts to better track job tails, distribution of processing time, summary of failures. Task-centric monitoring uses the Experiment Dashboard and hBrowse [76] frameworks.

The instant job monitoring tools provide all necessary information to follow up possible job failures and to file a bug report.

The workload accounting is visualized in the Historical Views [80] dashboard. The Historical Views dashboard provides views which focus on accounting in-

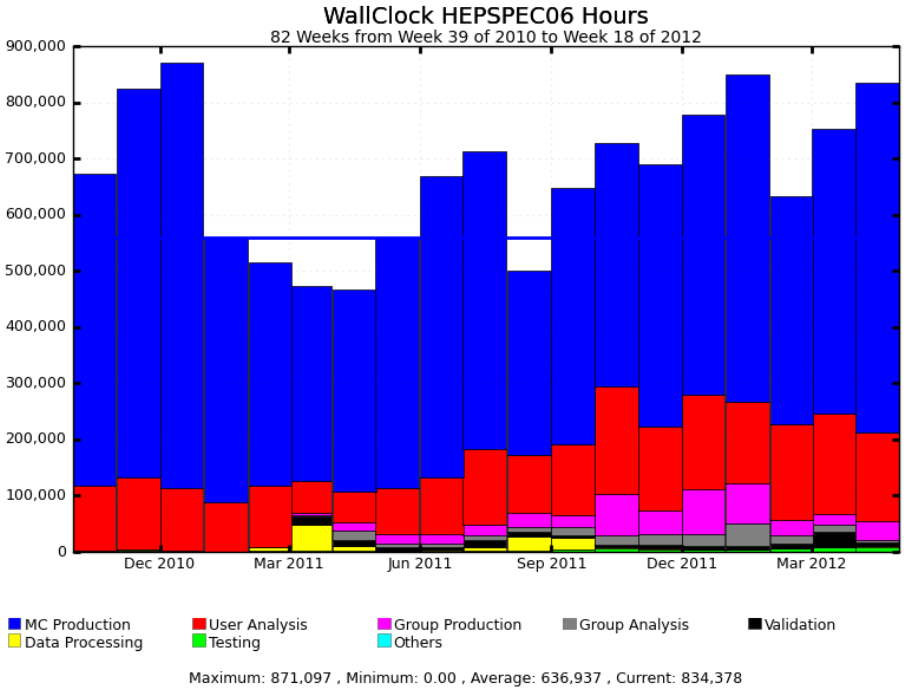


Figure 5.7: ATLAS grid resources utilization since late 2010. This plot shows walltime (in HEPSPEC06 hours) spent on all ATLAS sites. Data bins are aggregated by activity. Blue bars denote the Monte Carlo production, red bars User analysis, violet bars Group production, grey bars Group analysis, yellow bars Data processing, and green bars Testing activity.

formation about completed, submitted, pending and running jobs, CPU consumption and efficiency, processed data, success/failures, activities, and resource utilization. The Historical Views dashboard provides a set of powerful filtering options: filter by sites (selected by Tier level, cloud, country, or simply a list of sites), by physics groups requesting a workload to be processed, by input data types (e.g. AOD, ESD, NTUP, RAW and their derivatives) and data projects (e.g. data11_7TeV, mc11_7TeV, etc.), by cloud, by activity (e.g. Data processing, Production triggered by a physics group, Analysis triggered by a physics group, Monte Carlo simulations, User analysis, Validation and Testing). The filters allow the user to define the time interval for the plots, time binning, and grouping by all the filters listed above. The plots in each of the views contain the time evolution plots as instant and cumulative evolution in time, and a pie chart with drill down to categories given by the filter. The Historical Views dashboard provides a useful source of information for data mining the performance of the ATLAS grid resources, as well as a platform to interconnect information about the computing resources from different sources, e.g. the information about jobs is taken from PanDA, while information about pledged CPU resources is taken from WLCG REBUS, and the site topology information is taken from the ATLAS Grid Information System (AGIS) [81]. An example of such an interconnection of the CPU resources information is shown in Fig. 5.7: time evolution plot of

the walltime (in HEPSPEC06 hours) spent on all ATLAS grid sites from early 2010 to mid-2012 is shown, the line of amount of walltime pledged by the sites is depicted. In addition, this plot shows relative share of different activities: User analysis, Monte Carlo simulations, Group production and analysis, Data processing, Testing and Validation. Each of the bins show the monthly-averaged sum. One can easily see the steep increase in the User analysis activity share around the summer conferences of 2011.

5.5.4 Databases

The ATLAS Distributed Computing products store their status information and metadata in a central database, which is served by an Oracle 11g cluster hosted at CERN. The cluster is monitored by the DB Administration experts. ATLAS Distributed Computing shifters and experts can monitor the status of the databases hosted by the Oracle 11g cluster on a simple dashboard. This DB dashboard shows the status of each of the databases, distinguishing service degradation with a simple color scheme. The DB dashboard also provides a very compact set of parameters to describe the cluster health. More detailed information can be retrieved from the Service Level Status (SLS) [82] monitoring, which is referred from the DB dashboard.

The ATLAS grid jobs need to access several databases to retrieve metadata essential for data processing. At the beginning of 2010, there was a replica of such a database at the ATLAS Tier-0 at CERN and each of 10 Tier-1s. Only direct access to this database from a hosting site was possible. Subsequently ATLAS started using Frontier [83] in order to optimize data access and cache the content at the ATLAS Tier-0 at CERN and every Tier-1. In addition, a squid cache is deployed at every site, so that the metadata is cached, and access to metadata at Frontiers at Tier-1 sites is further optimized.

Frontier and Squid monitoring consists of a set of common monitoring tools shared between the ATLAS and CMS experiments. Frontier monitoring consists of an SLS probe which is used mainly by the ADC shifters and experts, and the more detailed awstats information targeting the Frontier experts. Squid monitoring consists of a standardized set of MRTG plots provided for the squids at every site. An example of such monitoring is shown in Fig. 5.8.

5.5.5 Status of Sites or Services

In 2010 ATLAS started to use the Site Status Board (SSB) [84] to aggregate monitoring information from many different sources, and visualize the status of ATLAS grid sites on a single page. The ATLAS SSB [85], [86] proved to be a very useful tool providing not only the current status of the ATLAS Distributed Computing infrastructure, but also historical trends of the aggregated monitoring data. The core feature of the ATLAS SSB, status information aggregated on a single page, is used by different groups of consumers ranging from the ADC Shifters, through experts, to ATLAS Site administrators. In Fig. 5.9 an example of the Cloud view is shown. A simple color scheme used in the ATLAS SSB

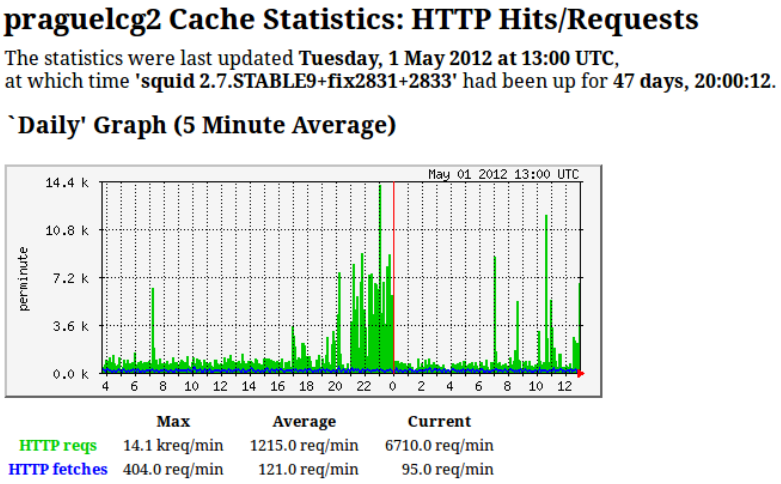


Figure 5.8: Monitoring of a Squid instance at a Tier-2 site *praguelcg2*.

views enables the user to rapidly spot problems and focus on their escalation or solution. The ATLAS SSB provides links to the original monitoring tool, where the shifter, or expert, or site administrator can get more detailed information about the metric status.

The ATLAS SSB views provide aggregated information for a variety of different metrics, such as site topology information, downtime, status in DDM, status in workload management, status in functional tests (HammerCloud [87], SAM/Nagios [88], DDM FT), exclusion status for various activities, number of critical software releases installed at a site and progress of SW installation.

Site Info					Panda Efficiency					
Tier	Cloud	DDM DT - Status	Panda Analysis status	Panda Production status	Analy Activated Jobs	Analy Running Jobs	Analy Efficiency 12h [%]	Prod Activated Jobs	Prod Running Jobs	Prod Efficiency 12h [%]
T2D	DE	online	online	online	111	536	38	296	174	99
T2	DE	online	online	online	2836	791	86	461	798	99
T2D	DE	online	online	online	1001	573	94	628	1146	100
T2D	DE	online	online	online	1720	256	84	358	225	100
T1	DE	online	online	online	12026	594	84	4798	1652	100
T2D	DE	blacklisted	online	online	2302	82	94	1676	1382	95
T2	DE	online	brokeroff	online	0	2	100	0	8	100
T2D	DE	online	online	online	206	78	0	211	37	2
T2D	DE	online	online	online	3124	316	84	1848	1289	99
T2	DE	online	online	online	2428	841	90	458	783	94
T2	DE	online	test	online	2	5	28	0	8	85
T3	DE	online	online	online	0	26	85	68	133	81

Figure 5.9: ATLAS Site Status Board – Cloud view with aggregated status of several ATLAS sites. A simple color scheme is used to easily distinguish between services which are running fine (green), those which may need the attention of a site administrator (yellow), those which are failing or excluded (red) and those which are inactive at a particular site (grey).

The ATLAS SSB history page provides a simple history plot, a ranking plot, a history table with plot and list of all possible states, and a plot with the number

of sites entering each of the states. Each plot can be configured with a powerful set of filters: time interval, group of sites (one or multiple sites, filtering sites by their Tier level or cloud), and choice of time-binning. An example plot showing ATLAS Site Usability in Production and Analysis for the ATLAS Tier-0 at CERN and all the Tier-1, and Tier-2 sites since August 2011 is shown in Fig. 5.10. The ATLAS Site Status Board is one of the applications based on the Experiment Dashboard framework.

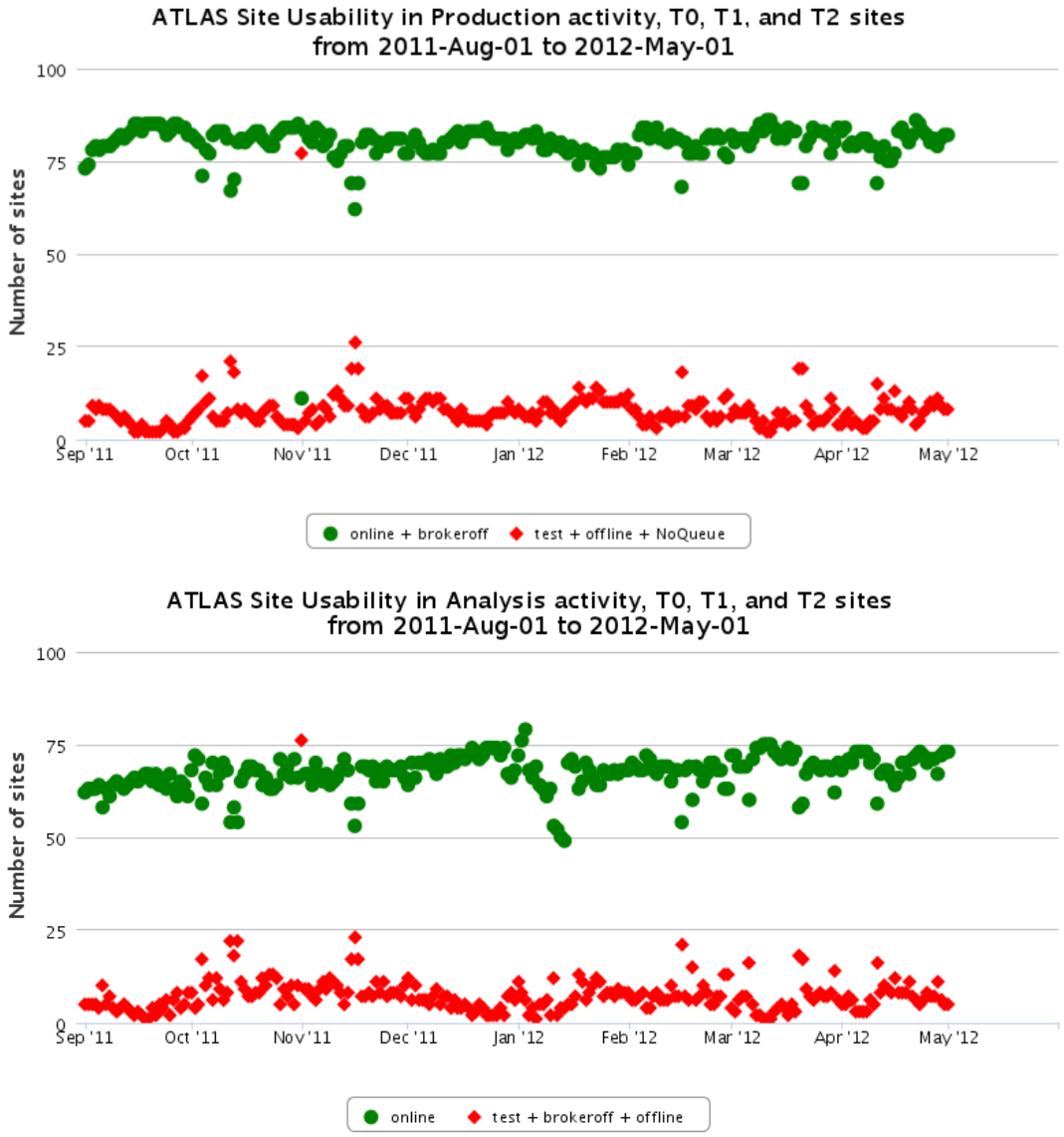


Figure 5.10: ATLAS Site Status Board – sum of the number of the ATLAS Tier-1, and Tier-2 sites (including the ATLAS Tier-0 at CERN) available for Production activity (top) and for Analysis activity (bottom).

ATLAS uses the Site Usability Monitor (SUM) [89] to visualize the latest and historic results of the SAM/Nagios tests since February 2012. The ATLAS SUM provides availability and reliability plots for the critical CE and SE metrics,

as a quality or as a ranking plot. These plots are provided for availability and reliability of a site, of a service registered in GOCDB [90] or OIM DB [91], and for test results. In Fig. 5.11 an example of a site reliability plot for the ATLAS Tier-0 at CERN and ATLAS Tier-1 sites is shown. The visualization part benefits from the Experiment Dashboard framework, while the data is provided by the SAM framework API.

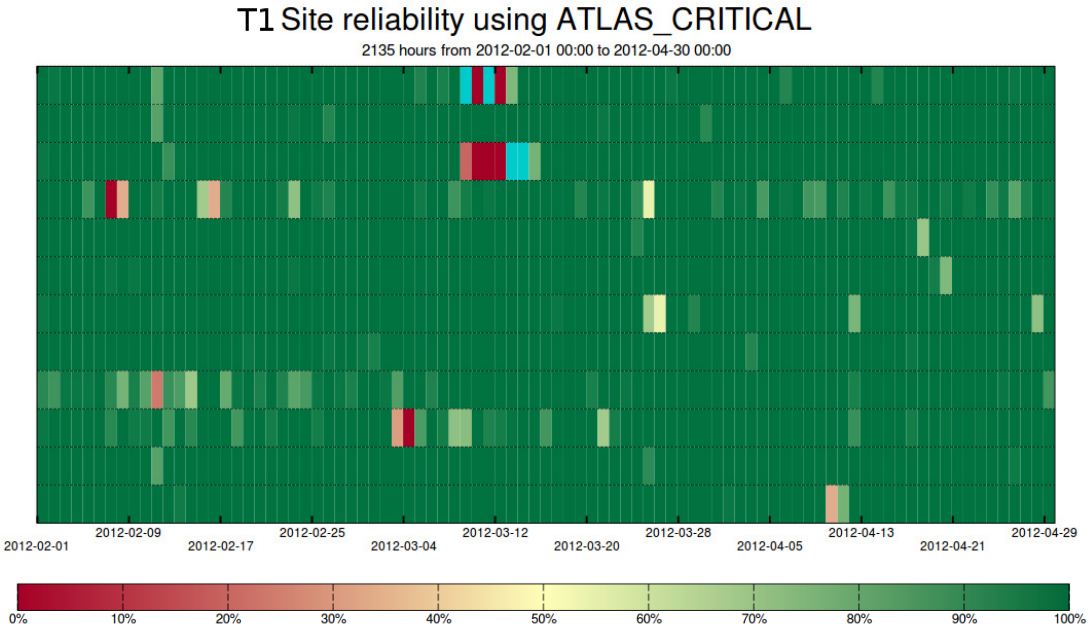


Figure 5.11: ATLAS Site Usability Monitor – Reliability of the ATLAS Tier-0 at CERN and ATLAS Tier-1 sites based on results of critical SAM/Nagios tests since February 2012. Site reliability ranges from 100 % (green) to 0 % (red). The blue cell denotes an "unknown" reliability, which corresponds to missing test results during a period of more than 24 hours, e.g. due to resources overloaded by the production workload (tests have lower priority than production activities).

ATLAS Distributed Computing workload management uses pilot jobs submitted to ATLAS grid sites. The pilot jobs then execute useful job payloads. Pilot job submission is handled by a pilot factory. The pilot factory can be operated centrally or by an expert local to the cloud. Each site defines several queues. The pilot factory then submits jobs to these queues. The health of each of the queues, submission, runtime, and exit status is monitored for each of the pilot factories.

In 2010 ATLAS increased the number of ATLAS grid sites available for physics analysis with the addition of the Tier-3 sites. A Tier-3 site is a site which provides non-pledged computing resources (CPU, storage, network connectivity) to the local community, but a Grid one can be used centrally by ATLAS in an opportunistic way. Such a Tier-3 site may be completely off-grid, or can be part of the ATLAS grid site family. The non-grid Tier-3 sites are monitored with the T3MON package [92], which provides a set of very useful monitoring tools for site administrators. The T3MON package monitors local computing resources at a off-grid Tier-3 site. It provides a global aggregation of the monitoring probes.

6. ATLAS Distributed Computing Monitoring during LHC Run I

In this chapter we describe how ADC resources are monitored, in order to ensure their efficient utilization. In 2011-2015 author of this Thesis contributed to the following projects:

1. Coordinator of the ADC Monitoring for over 2 years during LHC Run I. During the term the suite of the ADC Monitoring tools was enriched, and a more standardized approach of how to organize and implement user interfaces and APIs has been adopted. Accounting applications aiming at the Management of ATLAS and ATLAS Computing, the funding agencies, and other resource reporting bodies, have been introduced.

2. Developer of several monitoring applications: Site Usability Monitor, PanDA Dynamic Data Placement Monitor, PanDA Brokerage Monitor, and the Production Task Monitor. The latest one was extensively used by the Top physics group managers to follow data processing and Monte Carlo simulations of the Top physics group.

3. Co-project leader of ATLAS Site Status Board project, and a developer of Site Status Board monitoring metrics.

Projects were presented at several international conferences [70], [93], [85], [86], [7], [94], [72], [89].

An overview of ADC monitoring suite as of May 2012 was provided in Section 5.5. Here in Section 6.1, we summarize the standardization efforts conducted in the first 2 full years of LHC data taking. In Section 6.2, we briefly outline future challenges for the monitoring after the LHC Run I. In Section 6.3 the ATLAS Production Task Monitor is further described.

6.1 Standardization

The ATLAS Distributed Computing Monitoring tools significantly evolved over the past 5 years. Historically, there were many tools which originated at different times with different technologies available at those times and used a variety of visualization elements. Over the past two years the monitoring tools converged in terms of data communication and data visualization. Data is provided by a backend in the lightweight data-interchange format of JSON (JavaScript Object Notation). JSON data is then rendered by a visualization framework based on the JavaScript library jQuery and some of its plug-ins. Examples of the visualization frameworks used are the *xbrowse* framework (DDM dashboard, WLCG Transfers dashboard) and the *hBrowse* framework (Production and User task

monitoring). Examples of plotting libraries used are *Highcharts* (Production and User task monitoring), *graphtool* (Historical Views dashboard), *flot* (Panda monitor), *Google charts* (PD2P monitoring). The ATLAS Distributed Computing Monitoring tools converged also in terms of visual identity. All the monitoring tools follow common color schemes based on the sRGB color space.

6.2 Future challenges

Identification of common solutions in the monitoring field can help us to save manpower needed to develop and maintain the monitoring tools. Another concern is and will be the scalability of a solution. In order to decrease the load on production databases we are going in two distinct ways: distribution of load on multiple clusters, and moving from relational DBs to NoSQL DB solutions. One of the major concerns for the future is not only to be able to monitor every possible activity of the ATLAS Distributed Computing, but also to be able to preserve the monitoring data or their summaries on a long-term scale.

The ATLAS Experiment has already been collecting data since the beginning of the LHC Run I. The ATLAS data is processed and analyzed at more than 120 grid sites distributed worldwide. Such a complex system is continuously monitored. The ATLAS Distributed Computing successfully fulfills its mission to deliver data to the ATLAS physicists. The available monitoring tools help to identify possible issues in a timely manner, as well as mine long-term data series, or monitor and effectively utilize available computational resources. The ATLAS Distributed Computing Monitoring tools provide a comprehensive way to monitor the infrastructure, identify and address issues, and to improve automation of repetitive tasks, as well as ways to provide ways to present accounting information desired for resource utilization reports on many levels. The ADC Monitoring tools have to maintain a high level of quality, and be flexible enough to support the high energy physics community by adapting to the ever-evolving environment. The future challenges to face come with advances in data storage technologies, and the need to preserve data.

6.3 ATLAS Production Task Monitoring

In this section the ATLAS Production Task Monitoring tool is described. This monitoring tool that has been developed between 2011 and 2013. The focus of this section is in particular on the features developed between January and April 2013 by the author of this Thesis.

During the LHC Run I and most of LS1 PanDA WMS provided a job-oriented workflow, while the production system provided a task-oriented workflow. This approach has been superseded during the 2nd part of the LS1, when both PanDA and production system became task-oriented.

A task is defined by the input and output datasets, transformation (to distinguish how the data will be processed, or which MC generator to use for simulation), and version of the SW to be used for the execution. A production campaign

consists of many tasks.

Upon submission of a task to the production system, task definition is communicated to PanDA WMS, where the workload is split into computational jobs. Depending on the task definition, in particular on size of input dataset, and transformation, the number of jobs of a task can range from several tens to several tens of thousands.

According to the ATLAS Physics Program needs and goals established by the ATLAS Physics Coordination, different tasks have different priorities assigned, therefore different amount of resources available over a period of time. The time span of a task lifetime can range from several hours up to many months as seen from Fig. 6.1.

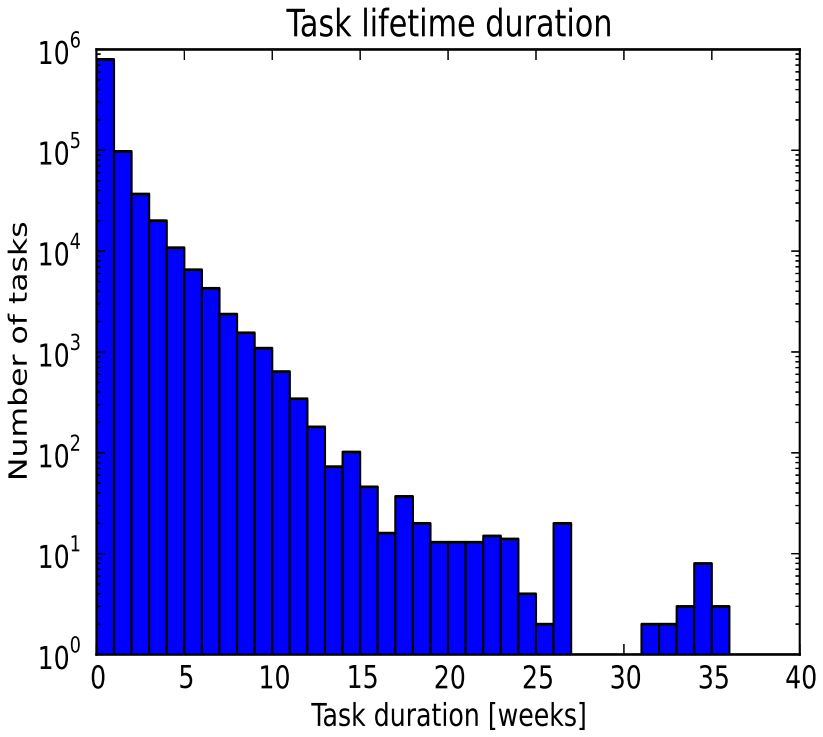


Figure 6.1: Distribution of task life-span during LHC Run I.

About 10 % of jobs fail with various causes: e.g. bug in SW, infrastructure issue, job running for too long in a job slot defined with a lifetime shorter than the actual job lifetime, etc. Since late phases of the LHC Run I, PanDA has been enhanced to automatically retry some of the jobs failing with a well defined error, e.g. in case of a site-related failure, job is re-assigned to a different site where the job gets a 2nd chance to finish successfully.

Despite a series of improvements in different ADC projects, some of the job failures cannot be automatically removed, and thus the failures lead to the need to bring experts (being it computing or physics expert) into play at some point

to address the issue. The ADC Operations teams and the experts then to follow progress of the tasks and jobs.

The most challenging part of a campaign execution is chasing the tails of the tasks, i.e. follow the last few percent of jobs that have failed to run. When there are tens of tasks in a campaign, chasing the task tails and avoiding omitting one last task is a very demanding job, performed by the production managers.

The task tails progress can hardly be monitored from the PanDA job-oriented monitoring, this is when a task-oriented monitoring tool comes handy. For ATLAS, such a task-oriented monitoring tool was ATLAS Production Task Monitor.

In February and March 2013 the ATLAS Production Task Monitor went through a phase of rapid development, that introduced much requested and needed features to run ATLAS Top D3PD production for all 2012 data and mc12_8TeV samples, and for production of Higgs to Tau Tau embedding simulations. The monitoring and other improvements lead to impressive progress on these two production campaigns, it was appreciated not only by the ATLAS Top D3PD production coordinator, but by the ATLAS Physics Coordination as well.

6.3.1 Technologies

The ATLAS Production Task Monitor is one of the applications utilizing the Experiment Dashboard framework developed in CERN IT Department, by the former Experiment Support Group. The user interface benefits from the hBrowse [76] framework. The ATLAS Production Task Monitor provides access to the task-oriented data not only via the user interface, but through an API as well.

6.3.2 Task Filters

In order to navigate among tens of thousands tasks, filters come very handy. In the ATLAS Production Task Monitor, filter consists of filter fields. Filter fields are organized into filter field groups. The resulting filter is applied as the logical AND operation on values of active filter groups. The ATLAS Production Task Monitor provides a comprehensive summary (Fig. 6.2) of applied filter fields for the current view. The available filter groups and fields are described in Appendix A.1, and depicted in Fig. 6.3.

The Filter feature of the ATLAS Production Task Monitor is equipped with a set of predefined labels and hints. Predefined labels and hints are available on hover for Filter labels, input filter fields, and table column headers.

Figure 6.2: Example of a filter summary.

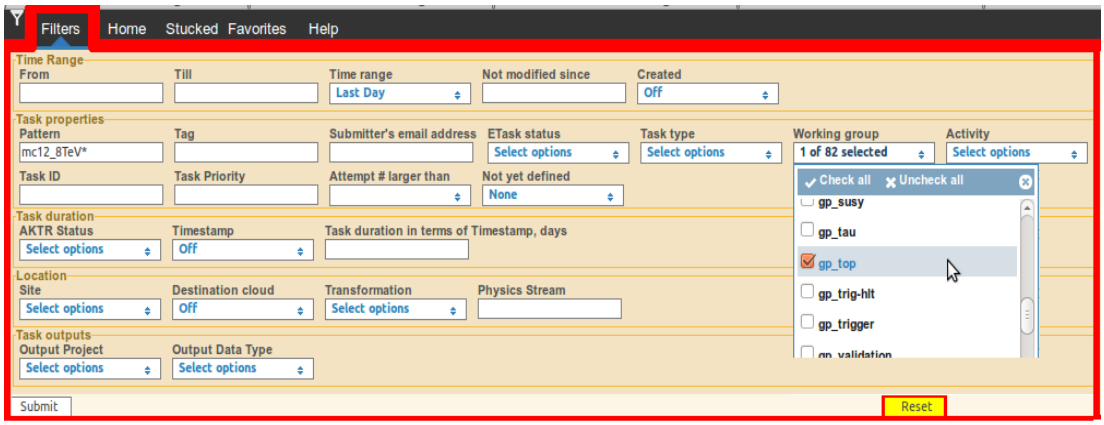


Figure 6.3: Filter fields of the ATLAS Production Task Monitor.

6.3.3 Favorites

The ATLAS Production Task Monitor provides a short list of pre-selected task filters, known as Favorites. These filters are defined by properties of important campaigns or their steps. The Favorites filters serve as a starting point for the Group production managers and ADC Operations teams to focus on a small well defined subset of all active tasks, and chase issues with such tasks. An example of Favorites from February 2013 is shown in Fig. 6.4.

The pre-selection of Favorites has been managed by the ATLAS Production Task Monitor developer, and the Group production managers. The definition is stored in a JSON dictionary, available as a part of the documentation TWiki page available to all the ATLAS Collaborators.

6.3.4 Home view

The Home view of ATLAS Production Task Monitor contains 3 elements: the menu bar, the filter area, and the content area. The content area consists of the Data tab, and the Summary tab.

Home view – Data tab

The Data tab shows a table with the list of tasks that fulfill the applied Filter. The table contains header, footer, and the data part. Example of such a table is shown in Fig. 6.5 and Fig. 6.6. The Home view is further described in Appendix A.2.

6.3.5 Task details

The table in the Data tab of the Home view provides overview of a list of tasks. In addition, more detailed information about a single Task is shown upon expanding the Task row in the said table. The details of a single task details consist of

- the original *Data table row*, with link to list of jobs of the task,
- *TaskName*,

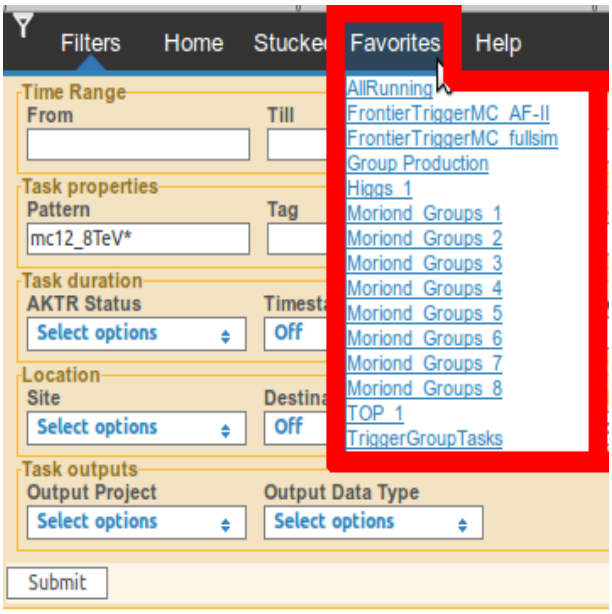
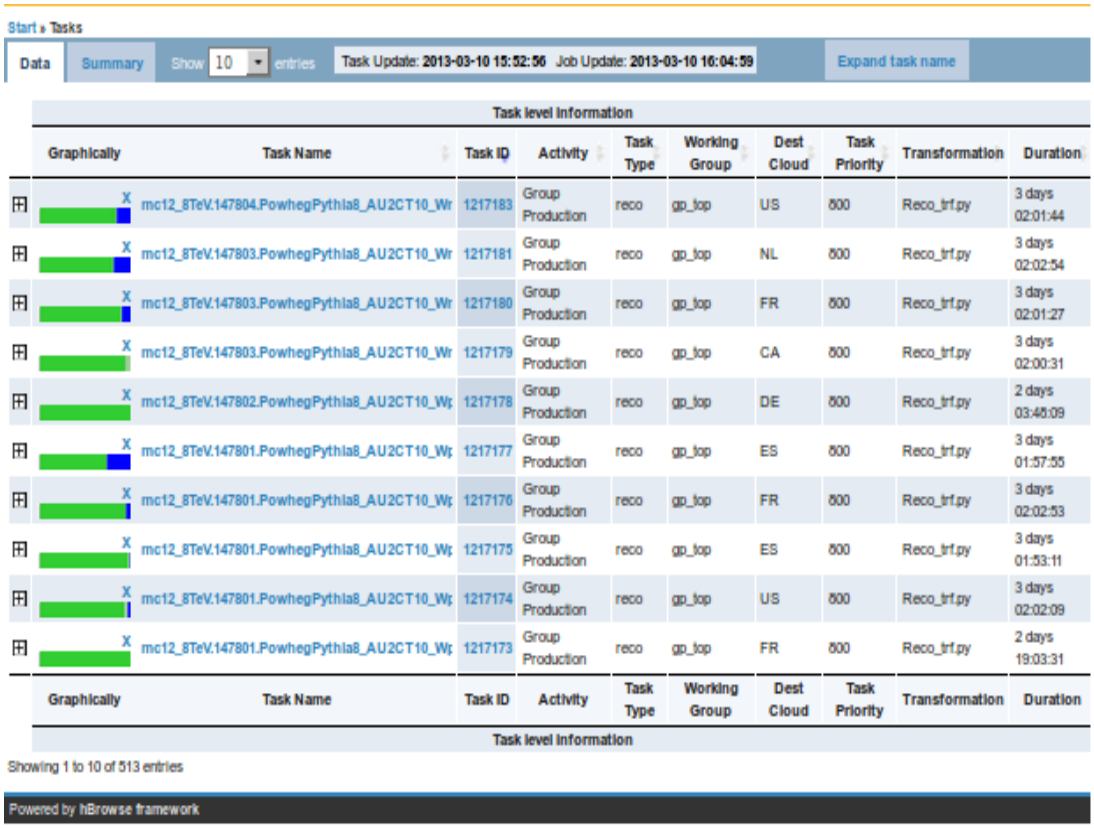


Figure 6.4: "Favorites" pre-selected filters of the ATLAS Production Task Monitor available in February 2013, with emphasis on important production campaigns in preparation for the conference Rencontres de Moriond 2013.

- *Failed*: number of failed jobs of the task,
- *TaskMonitorId*: internal string of the ProdTaskMonitoring application, composed of prefix panda_prod, submitter's e-mail address, and the Task ID,
- *TaskCreatedTimeStamp*: timestamp of the task creation,
- *Application*: Name and version of the application used,
- *NEvents*: Number of input events processed by successful jobs of the task,
- *ModificationTime*: Task modification timestamp,
- *DboardModificationTime*: Internal timestamp of latest update of the task-related information in the Dashboard DB,
- *InputCollection*: Name of the input dataset/container of the task,
- *Site*: List of ATLAS Site names on which the task has run,
- *InputDataFiles*: Number of input files,
- *InputFileType*: Data Type of the input,
- *InputFileBytes*: Size of the input,
- *InputFileProject*: Data Project of the input,



Start » Tasks

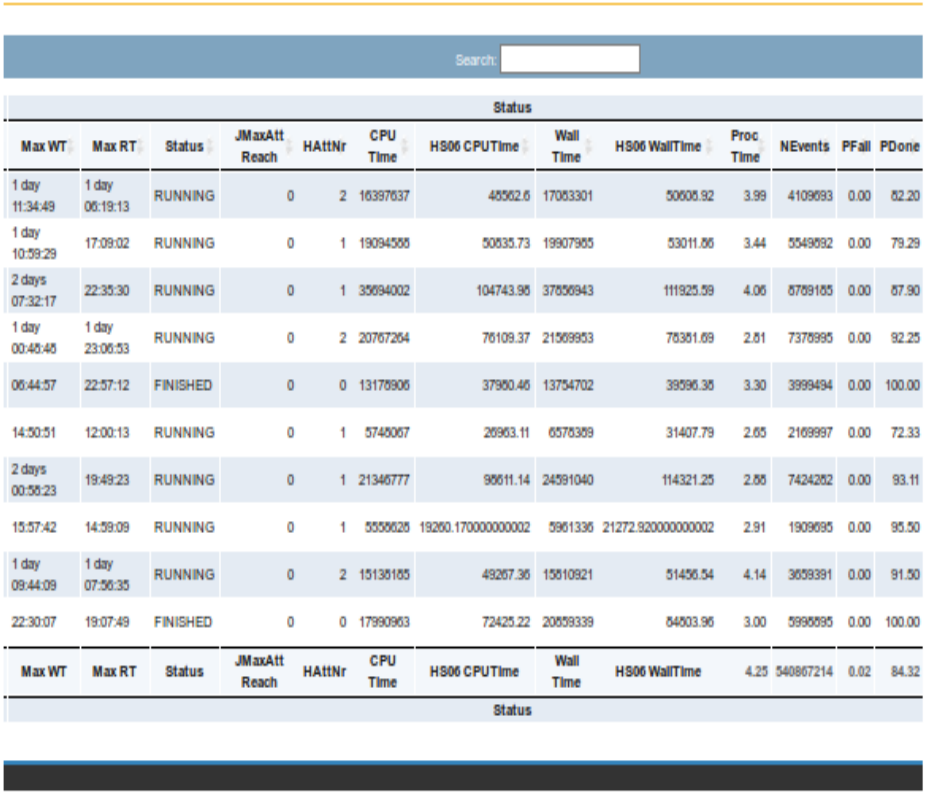
Data | Summary | Show 10 entries | Task Update: 2013-03-10 15:52:56 Job Update: 2013-03-10 16:04:59 | Expand task name

Task level information

Graphically	Task Name	Task ID	Activity	Task Type	Working Group	Dest Cloud	Task Priority	Transformation	Duration
	mc12_8TeV.147804.PowhegPythia8_AU2CT10_Wr	1217183	Group Production	reco	gp_top	US	800	Reco_trf.py	3 days 02:01:44
	mc12_8TeV.147803.PowhegPythia8_AU2CT10_Wr	1217181	Group Production	reco	gp_top	NL	800	Reco_trf.py	3 days 02:02:54
	mc12_8TeV.147803.PowhegPythia8_AU2CT10_Wr	1217180	Group Production	reco	gp_top	FR	800	Reco_trf.py	3 days 02:01:27
	mc12_8TeV.147803.PowhegPythia8_AU2CT10_Wr	1217179	Group Production	reco	gp_top	CA	800	Reco_trf.py	3 days 02:00:31
	mc12_8TeV.147802.PowhegPythia8_AU2CT10_Wp	1217178	Group Production	reco	gp_top	DE	800	Reco_trf.py	2 days 03:46:09
	mc12_8TeV.147801.PowhegPythia8_AU2CT10_Wp	1217177	Group Production	reco	gp_top	ES	800	Reco_trf.py	3 days 01:57:55
	mc12_8TeV.147801.PowhegPythia8_AU2CT10_Wp	1217176	Group Production	reco	gp_top	FR	800	Reco_trf.py	3 days 02:02:53
	mc12_8TeV.147801.PowhegPythia8_AU2CT10_Wp	1217175	Group Production	reco	gp_top	ES	800	Reco_trf.py	3 days 01:53:11
	mc12_8TeV.147801.PowhegPythia8_AU2CT10_Wp	1217174	Group Production	reco	gp_top	US	800	Reco_trf.py	3 days 02:02:09
	mc12_8TeV.147801.PowhegPythia8_AU2CT10_Wp	1217173	Group Production	reco	gp_top	FR	800	Reco_trf.py	2 days 19:03:31

Showing 1 to 10 of 513 entries

Powered by hBrowse framework



Search:

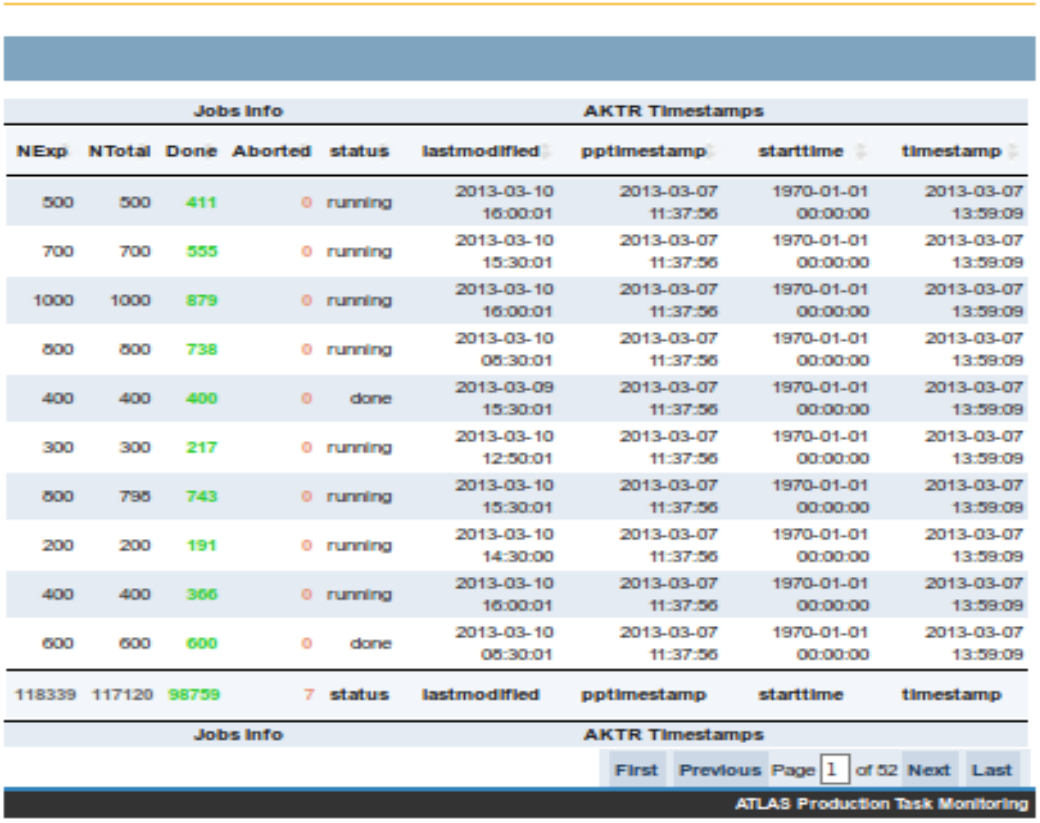
Status

Max WT	Max RT	Status	JMaxAtt Reach	HAttNr	CPU Time	HS06 CPUTime	Wall Time	HS06 WallTime	Proc Time	NEvents	PFail	PDone
1 day 11:34:49	1 day 06:19:13	RUNNING	0	2	16397637	46562.6	17053301	50606.92	3.99	4109693	0.00	82.20
1 day 10:59:29	17:09:02	RUNNING	0	1	19094566	50535.73	19907965	53011.66	3.44	5549692	0.00	79.29
2 days 07:32:17	22:35:30	RUNNING	0	1	35694002	104743.95	37656943	111925.59	4.06	5759165	0.00	67.90
1 day 00:46:46	1 day 23:06:53	RUNNING	0	2	20767264	76109.37	21569953	78361.69	2.61	7376995	0.00	92.25
06:44:57	22:57:12	FINISHED	0	0	13176906	37960.46	13754702	39596.36	3.30	3999494	0.00	100.00
14:50:51	12:00:13	RUNNING	0	1	5746067	26963.11	6576369	31407.79	2.65	2169997	0.00	72.33
2 days 00:56:23	19:49:23	RUNNING	0	1	21346777	96611.14	24591040	114321.25	2.66	7424262	0.00	93.11
15:57:42	14:59:09	RUNNING	0	1	5556626	19260.170000000002	5961336	21272.920000000002	2.91	1909695	0.00	95.50
1 day 09:44:09	1 day 07:56:35	RUNNING	0	2	15136165	49267.36	15610921	51456.54	4.14	3659391	0.00	91.50
22:30:07	19:07:49	FINISHED	0	0	17990963	72425.22	20659339	84603.96	3.00	5995695	0.00	100.00

Max WT	Max RT	Status	JMaxAtt Reach	HAttNr	CPU Time	HS06 CPUTime	Wall Time	HS06 WallTime	4.25	540867214	0.02	84.32

Status

Figure 6.5: Part I: Task list table in the ATLAS Production Task Monitor.



Jobs Info					AKTR Timestamps			
NExp	NTotal	Done	Aborted	status	lastmodified	pptimestamp	starttime	timestamp
500	500	411	0	running	2013-03-10 16:00:01	2013-03-07 11:37:56	1970-01-01 00:00:00	2013-03-07 13:59:09
700	700	555	0	running	2013-03-10 15:30:01	2013-03-07 11:37:56	1970-01-01 00:00:00	2013-03-07 13:59:09
1000	1000	879	0	running	2013-03-10 16:00:01	2013-03-07 11:37:56	1970-01-01 00:00:00	2013-03-07 13:59:09
800	800	738	0	running	2013-03-10 05:30:01	2013-03-07 11:37:56	1970-01-01 00:00:00	2013-03-07 13:59:09
400	400	400	0	done	2013-03-09 15:30:01	2013-03-07 11:37:56	1970-01-01 00:00:00	2013-03-07 13:59:09
300	300	217	0	running	2013-03-10 12:50:01	2013-03-07 11:37:56	1970-01-01 00:00:00	2013-03-07 13:59:09
800	795	743	0	running	2013-03-10 15:30:01	2013-03-07 11:37:56	1970-01-01 00:00:00	2013-03-07 13:59:09
200	200	191	0	running	2013-03-10 14:30:00	2013-03-07 11:37:56	1970-01-01 00:00:00	2013-03-07 13:59:09
400	400	366	0	running	2013-03-10 16:00:01	2013-03-07 11:37:56	1970-01-01 00:00:00	2013-03-07 13:59:09
600	600	600	0	done	2013-03-10 05:30:01	2013-03-07 11:37:56	1970-01-01 00:00:00	2013-03-07 13:59:09
118339	117120	98759	7	status	lastmodified	pptimestamp	starttime	timestamp
Jobs Info					AKTR Timestamps			

First Previous Page 1 of 52 Next Last

ATLAS Production Task Monitoring

Figure 6.6: Part II: Task list table in the ATLAS Production Task Monitor.

- *Plots*:
  - Plot of Evolution of job states,
  - Cumulative plot of Evolution of job states,
  - Plot Distribution of attempt number of jobs,
  - Plot Distribution of processing time per event,
  - Plot Current Error Summary: Error summary for the latest attempt of jobs,
  - Plot Total Error Summary.
- Summary table with number of failed jobs per category (column), for latest attempt of the jobs, and in total. The error categories, that specify at which stage of an ideal job lifetime the job failed, are:
  - TaskBuffer,
  - Supervisor,
  - JobDispatcher,
  - Transformation,
  - Pilot,

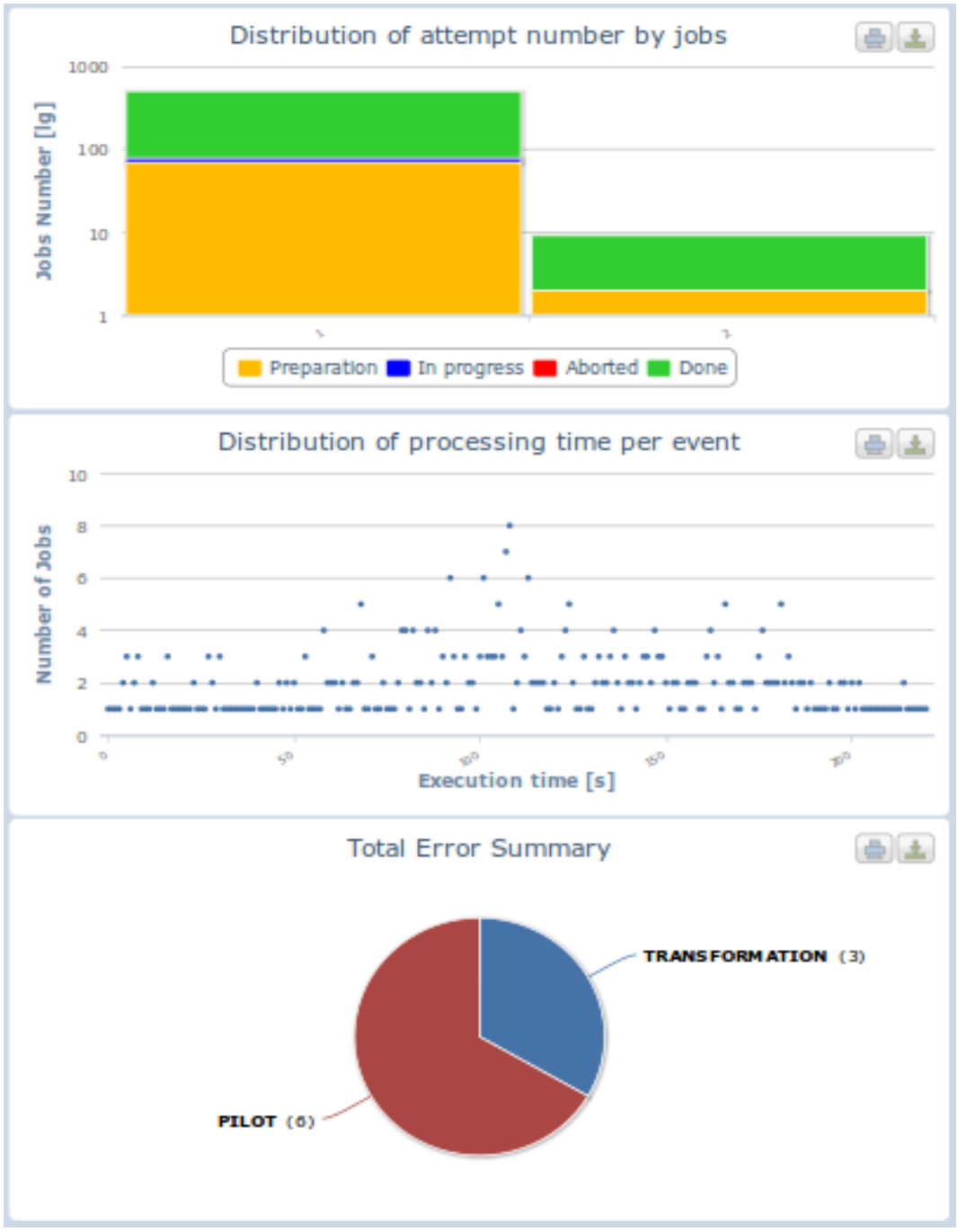


Figure 6.7: Plots describing details of a task progress.

  - DDM,
  - Brokerage,
  - Execution.

Example plots describing progress of a single task are shown in Fig. 6.7.

6.3.6 View of jobs per task

The view of jobs per task provides summary information about job instances of a particular task, and link to job monitoring with much more detailed information about the jobs. This view is available from the main task table view, by clicking on the Task Name or the letter X in the Graphically column.

The view of jobs consists of the Jobs Filters, and the content part, which is either Jobs Data tab, or Jobs Summary tab. Further description of this view is provided in Appendix A.3.

6.3.7 Task API

In order to enable programmatic access to production task data, an API has been implemented. It exposes information about list of tasks, and list of jobs, in the JSON format. Such an information can be used for further data mining by the Production task manager.

The JSON responses contain similar information as the data tables (of Task and Jobs), and in addition to the raw data the aggregated values are included.

6.3.8 Problematic tasks view

When a Production manager submits tens to thousands of tasks at the same time, it it extremely easy to overlook a handful of tasks that do not behave as expected. Such tasks may get stalled due to mis-configuration, or a typo in a task parameter definition. The Problematic tasks view helps to identify tasks which stay within the task definition phase, without having any jobs defined. It exposes such faulty tasks, resulting in decreased delay between the physics task definition, and delivery of the processed data or generated MC simulations to the physicists for further analysis.

The problematic tasks can be filtered in the similar way as the tasks of the Task view. The Problematic task view contains table with the list of tasks, with the following columns:

- Task Name,
- Task ID,
- Activity,
- Task Type,
- Working Group,

- Destination Cloud,
- Task Priority,
- Status,
- Last Modified Timestamp,
- PostProduction Timestamp,
- Start Time Timestamp.

7. ATLAS Distributed Computing Automation

A description of automation of repetitive tasks of several ADC activities is given in this chapter. These tasks were automated in order to improve efficiency of usage of the ATLAS computing resources, and to optimize utilization of available manpower of the ADC Operations team. In 2011-2014 author of this Thesis contributed to the following projects:

1. **Prototyped topology of ADC resources**, unifying naming conventions of resources in different systems: the workload and distributed data management systems and the grid information systems used different names and naming conventions for ADC resources. Prototype of the topology was adopted by the the ATLAS Grid Information System (described in Section 7.4), and since then has been significantly extended to address evolving needs of ADC.

2. **Developed Switcher1**, automated tool to exclude and recover distributed compute resources for downtime of an underlying computing element, further described in Section 7.5.

3. **Prototyped Switcher2**, acting on PanDA resources, and taking into account also downtimes of related storage element.

4. **Centralized PanDA exclusion** project: simplify exclusion and recovery resources by defining policies that simplified the way how the exclusion actuators (Switcher and HammerCloud) cooperate.

5. **Storage Area Automatic Blacklisting (SAAB)** project: exclusion and recovery based on results of high-granularity storage functional tests [95].

6. Very early phase of the project of collecting network performance information in order to improve brokerage decision for data replication and for job dispatcher in PanDA.

Projects were presented at several international conferences [96], [97].

In Section 7.2 we describe why is it essential to invest into automation. In Section 7.3 we discuss validation of ATLAS grid sites. In Section 7.4, we summarize key features of the ATLAS Grid Information System for the automation efforts. In Section 7.5 we describe the automatic exclusion tool Switcher.

7.1 Introduction

The ADC infrastructure was described in Section 5.2. A challenging task to accommodate requests for the monitoring of the ADC infrastructure is addressed by the ADC Monitoring team [70], and was discussed in Chapter 6. An even

more challenging task to monitor the ADC infrastructure is covered by the ADC Operations teams: various ADC Shift teams [68] discussed in Section 5.4.1. ATLAS sites may or may not be part of three ATLAS Activities: Data transfers, Data processing, Distributed analysis. Data available in the ADC Monitoring tools helps to improve automation [98] of repetitive tasks, and direct the expert manpower to address more urgent issues.

ATLAS site validation with various functional tests is described in Section 7.3.

The ATLAS Grid Information System (AGIS) [81] contains detailed description of ATLAS computing resources, together with the information about resource status, and scheduled downtimes and unscheduled outages. AGIS is discussed in Section 7.4.

Aggregation of information about resource status from AGIS, automatic probes acting on resource downtime, and results of functional tests gives rise to a centralized exclusion tool. One of its parts, the Switcher, is described in Section 7.5.

7.2 Repetitive tasks and need for automation

We aim to maximize utilization of available ATLAS Distributed Computing resources by minimizing their downtime, and to expose only grid resources which provide reliable services to the ATLAS Physicists. In order to achieve this goal grid services at the sites are continuously monitored and validated. The whole ATLAS Distributed Computing infrastructure is monitored 24x7 by several teams of shifters with different responsibilities, as described in Section 5.4.1.

Whenever a shifter on duty identifies an issue with the ADC infrastructure, the shifter creates a bug report to expert, or activity requester, or to the site. There were over 11,700 GGUS [71] tickets created to the ATLAS grid sites in 2010-2015. The number of GGUS tickets per year, together with the daily average, is listed in Table 7.1.

Year	GGUS tickets created	Daily average
2010	2955	8.1
2011	2627	7.2
2012	2037	5.6
2013	1657	4.5
2014	1397	3.8
2015	1095	3.0

Table 7.1: Evolution of number of GGUS tickets created since 2010.

This amount of bug reports represents huge manual effort carried out by the ADC Operations team, ranging from the deep issue investigation by the ADC Shifter, creation of the bug report, addressing the issue by the site administrator or activity expert, resulting in functional testing of the reported service, and bringing the service back into production for ATLAS Activities. The amount

of manual work is the main motivation for automation troubleshooting of well known issues, and their prevention.

Over the past 6 years, in 2010-2015, the average number of GGUS tickets created daily decreased by a factor of 2.7, from 8.1 tickets per day on average in 2010, to 3 tickets per day on average in 2015. The number of shift teams reduced from 3 to 2 after Run I, leaving only DAST for distributed analysis support, and ADCoS for troubleshooting all other ADC activities. Effort invested into automation of ADC operations contributed to decrease of amount of issues the shift teams have to manually report, leaving more time to focus on issues that have not been automated yet.

7.3 Site Validation

The ATLAS experiment runs a steady flow of functional tests at the ADC resources in order to validate functionality of an ATLAS site for various Activities. The amount of resources needed for the test is marginal with respect to normal ATLAS Activity at a site, yet the functional tests provide significant value in spotting issues early. The fraction of functional tests with respect to the overall activity is of the order of percent (1-2 %). This section describes procedures for the validation, and details on what functional tests are performed.

7.3.1 Service exclusion and recovery

ATLAS sites can declare downtimes in common site information systems. Such a downtime then propagates to AGIS. When a site is on downtime or is failing functional tests, it is excluded from affected activity according to the ATLAS Site Exclusion Policy. Once downtime is over or the issue is fixed, functionality of the site is tested, and the site is recovered and enabled for an activity after it passes tests conducted automatically, or by the Shift Team. The schema of the service exclusion and recovery based on downtime published in the ATLAS Grid Information System is depicted in Fig. 7.1.

7.3.2 Service functional tests

ATLAS uses the WLCG Service Availability Monitor (SAM) framework [88], [89] to test resources [89] registered in GOCDB or OIM. Currently, the SAM test results are used as additional sanity check when manual recovery of a service is necessary.

Evaluation of SAM functional tests of storage elements leads to the tests enhancements, resulting in high-granularity tests. Automatic actions [95] are taken based on the high-granularity tests.

7.3.3 Validation of software installation

ATLAS SW installation system validates each release just installed at the site. WLCG Site Availability Monitor probes availability of various services at sites,

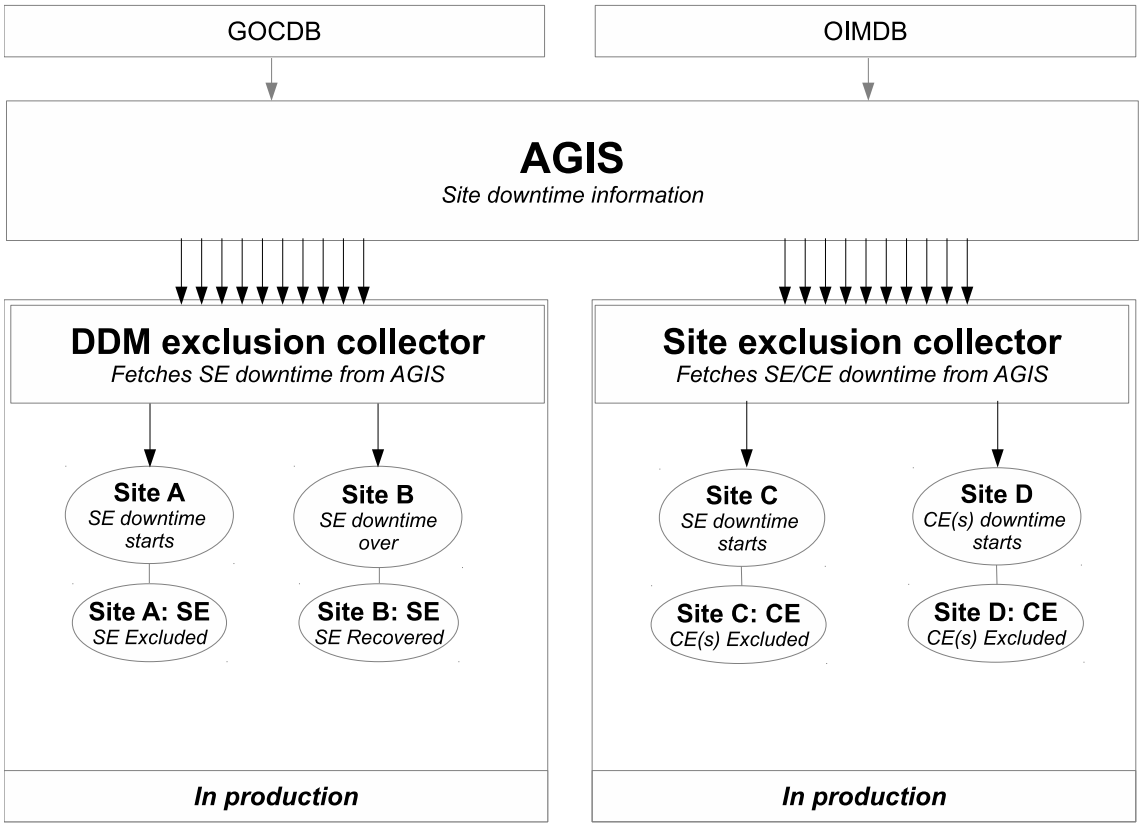


Figure 7.1: Schema of the information flow for the service exclusion and recovery based on downtime published in the ATLAS Grid Information System.

e.g. computing and storage resources. Monitoring information from different sources is aggregated in the ATLAS Site Status Board [85], [86].

7.3.4 Distributed Analysis and Data Processing functional tests

Analysis Functional Test (AFT) and Production Functional Test (PFT) represent functional tests for Distributed Analysis and Data Processing activities, both AFT and PFT run in the HammerCloud [87] framework. HammerCloud automatically excludes sites from analysis activity based on test failures, and recovers them once a certain number of tests in a row succeed. There are at least three flavors of analysis/production test jobs running on every analysis/production site every hour. The AFT/PFT test job examines correctness of the job environment, availability of required SW release, condition data, data stage-in from a storage element, data stage-out to a storage element, and registration of output in the DDM system catalog.

The ATLAS experiment uses the HammerCloud [87] framework to test how a site performs in the Data processing and Distributed analysis Activities. The HammerCloud test jobs simulate the behavior of a regular ATLAS data processing or analysis job. The HammerCloud uses the same environment as usual ATLAS

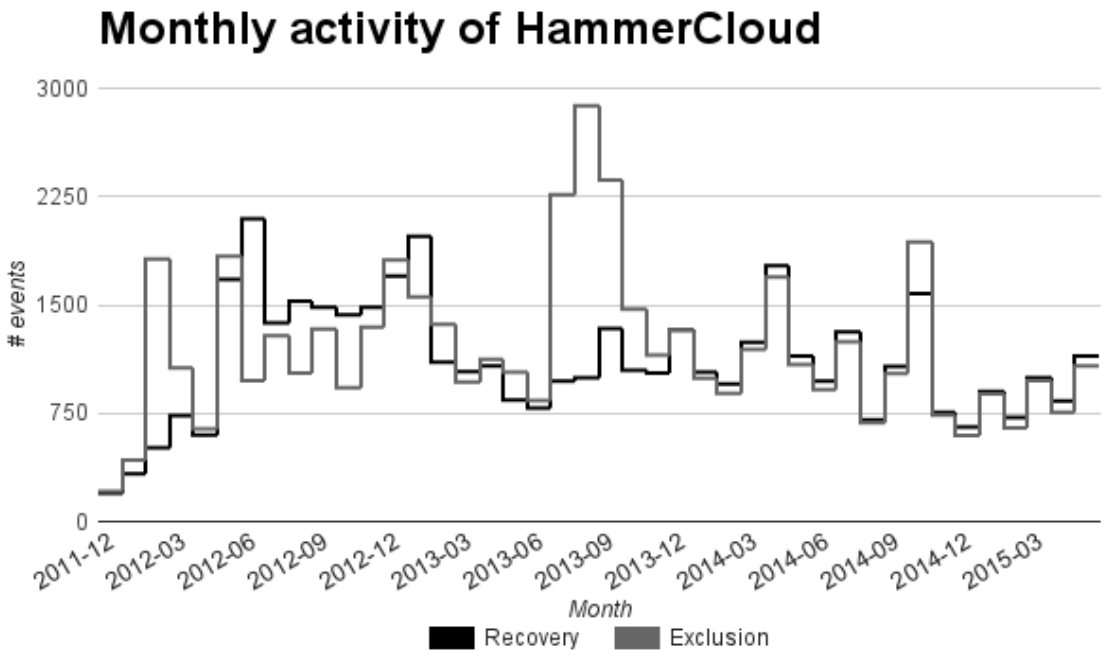


Figure 7.2: Time evolution plot of monthly activity (number of exclusion and recovery events) of HammerCloud from December 2011 to May 2015.

jobs, access input data and installed SW in the same way, and stages out the output data in the same way. The HammerCloud then provides a very useful probe in the site health for real Activities. When several HammerCloud tests fail, site is excluded from an Activity for a period of time, and recovered for that Activity only once a set of jobs in a row succeeds. The HammerCloud takes ca 300 exclusion/recovery actions per week. Monthly averaged activity of HammerCloud is shown in Fig. 7.2. The HammerCloud framework is used as the recovery framework for the Switcher exclusions, both tools cooperate as a part of the Centralized PanDA exclusion project.

7.3.5 DDM functional tests

DDM services at sites are validated through the DDM functional tests. Functional tests track how well a site can transfer data, and how fast files of different sizes can be transferred. Functional tests run between Tier-1 and Tier-2 sites from the same cloud, as well as between any pair of sites no matter where they are located. The DDM functional tests represent less than 1 % of data throughput, therefore do not affect performance of the complex DDM system.

7.3.6 Network link functional tests

Monitoring information about network link status became part of the automatic decision to determine the path for data transfers in the ATLAS distributed data management system [55], [99], and for job brokerage in PanDA [100].

The ATLAS experiment probes NxN endpoint-to-endpoint transfers functionality with the Sonar [77] test. The purpose of this testing is to find the optimal path for the transfers. In the past, ATLAS used strictly hierarchical topology of DDM endpoints.

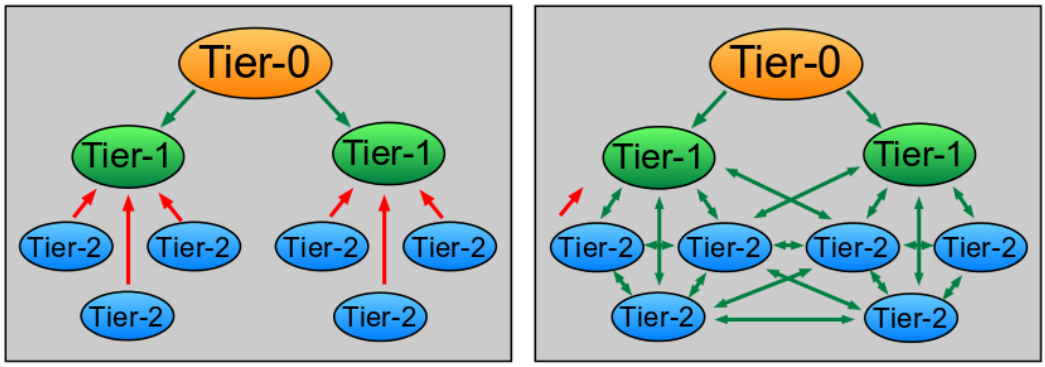


Figure 7.3: The evolution of hierarchical relationship between Tiers (left) to a less hierarchical model used today (right). The image originates from [101].

In the past a transfer between 2 Tier-2 sites, which belong to different clouds, was possible only through 2 Tier-1 sites, transfer path then was Tier-2 (Cloud A) → Tier-1 (Cloud A) → Tier-1 (Cloud B) → Tier-2 (Cloud B). This transfer path may not be very optimal, due to 3 additional sites being filled with data on the way.

Since LS1 the topology is no longer strictly hierarchical and ATLAS relaxes a bit the strictly hierarchical tier mode, and direct transfers are enabled between Tier-2 sites from different clouds with very good network connectivity. Evolution of hierarchy is depicted in Fig. 7.3.

As of 2012 there were 36 WLCG sites [102] taking part in the LHCONE [103] network project. Such sites are running perfSonar [78] tests. As of end of 2015 there are no automatic actions taken based on perfSonar test results.

7.4 ATLAS Grid Information System

7.4.1 AGIS Overview

The ATLAS Grid Information System (AGIS) [81] collects site information from the Grid Operations Center Data Base (GOCDB) [90] and the Open Science Grid (OSG) Information Management System (OIM) [91], and exposes it in a way convenient to the experiment. The AGIS provides topology information about the ATLAS grid sites, about services at sites, about downtimes of those services, it is the primary source of such information for the ATLAS experiment.

This unique information collection available in AGIS enables the ATLAS experiment to map between physical resources (e.g. CEs, SEs) and ATLAS activity endpoints (PanDA [57] queues workload management endpoint, DDM [55] space-

token[1] endpoints), and an additional logical layer in AGIS provides availability information of an ATLAS Activity at a particular site based on availability of subsequent physical resources at that site.

In addition, AGIS holds information about status of ATLAS computing resources, and their scheduled downtimes and unscheduled outages. Availability of this information in a single place facilitated creation of automatic probes which not only take storage and computing elements out of production for the duration of downtime, but had drained them in advance of a downtime as well. Draining of resources in advance helps to expose only fully functional resources to the ATLAS Physicists, and results in less manual interventions performed by the shift teams and site administrators.

7.4.2 Collectors that benefit from AGIS

Having written what useful set of information AGIS provides, ATLAS benefits from several collectors, which collect downtime information for ATLAS activity endpoints, and exclude those activity endpoints for downtime period from corresponding ATLAS Activities.

First example of such a collector is the DDM collector, which excludes DDM spacetoken(s) from Data transfer activity, with granularity of sub-activities such as write/read/deletion, when a downtime of underlying SE starts, and re-enables those DDM spacetokens for the sub-activities once the SE downtime is over.

The DDM collector takes automatic action ca 30 times per week. The main benefit of both collectors is saving ADC Operations manpower when a site declares unscheduled downtime, secondary benefit is for scheduled downtime.

Second example of a collector is the DDM space collector, which based on DDM spacetoken occupancy excludes a DDM spacetoken for write when a very small fraction of its size (several TBs) is remaining free to be used. When a fraction of used space at that DDM spacetoken is cleaned, at least up to the limit which enables uninterrupted ATLAS Activities at that site, DDM spacetoken is enabled for writing again.

Third example of a collector taking action when a service is on downtime, is the Switcher, described in Section 7.5.

7.5 Switcher – automatic exclusion and recovery of compute resources

7.5.1 ADC resources topology

In early 2011 the AGIS was still very far from being a production-quality Grid information system for ATLAS as it is today. AGIS collected information about ATLAS sites from different information systems: GOCDB, OIM. This information consisted mainly of service name, in-production/testbed status, and scheduled and unscheduled downtimes for the service. In addition to that AGIS collected

[1] A spacetoken is a logical area that corresponds to an area reserved on a SE.

information about ATLAS Activity endpoints, such as DDM spacetokens, and PanDA sites/resources/queues. However, AGIS did not provide any additional information regarding relations between services at a site, nor between services and ATLAS Activity endpoints at the time.

Throughout 2011 the author of this Thesis prototyped a mapping that described PanDA queues by aggregating information from several sources and adding activity-specific flags which were nowhere provided before. The elements describing each of the PanDA queues were:

- site name in external information system (GOCDB, OIM),
- site name in ATLAS,
- names of PanDA queue, resource, and site,
- site name in ATLAS DDM system,
- cloud to which the ATLAS site belongs,
- country in which the ATLAS site is located,
- flag that distinguishes analysis, production, and testbed-only PanDA resources,
- tier level of the ATLAS site

This mapping between PanDA queue and DDM endpoint and PanDA resource categorization were adopted by AGIS, which then became the sole provider of this information that resembles the first version of ATLAS Distributed Computing resources topology.

Later AGIS provided a mapping between ATLAS service endpoints (PanDA queue, DDM spacetoken) and the underlying site services (SEs, CEs). Together with the available information about the site service downtime, the ADC resources topology and the service downtime were the first two building blocks for Switcher.

7.5.2 Manual exclusion of PanDA queues

Early in the LHC Run I the ADCoS shifters had to manually exclude PanDA queues when the jobs started failing there, or per request of the site administrator. At that time 2010 the HammerCloud framework was being developed, and it started being used for AFT and for PFT around the end of the first half of LHC Run I. Any functional testing of PanDA resources in the first half of LHC Run I that followed PanDA queue exclusion has to be started manually by the ADCoS on duty. This manual operation was error prone, submission of a test job required a certain expertise in ATLAS job definition. When a test job failed, another one had to be manually sent. Only after the test job finished successfully, the PanDA queue was manually set online by the ADCoS on duty. The whole PanDA queue testing procedure was very manual and due to several-steps nature there were unnecessary delays in its completion, however, those were hard to avoid. In some cases a further interaction with the site administrator was conducted, in order to fix an issue uncovered by the test job.

7.5.3 Switcher design

In order to take huge work load off the shifter teams and ADC experts, and to increase automation level of the exclusion, test and recovery procedure, the implementation of the ATLAS Site Exclusion Policy gave rise to the automatic exclusion tool Switcher.

The ATLAS Site Exclusion takes into account performance of a site in the ATLAS activities. Performance is determined by results of the activity functional tests, and by the activity efficiency itself. It is monitored in the ATLAS Site Status Board [85], [86] which is described in Chapter 6.

The Switcher collects information relevant to PanDA queues (later PanDA resources) from AGIS: in particular name and status of PanDA queue, name of DDM endpoint, a list of SEs for the DDM endpoint, and downtimes of those SEs, and combines it. Based on this information it calculates a decision regarding what status the PanDA queue should be set, and executes the decision. Decisions per ATLAS cloud were aggregated and the cloud support team was notified, as well as ADCoS team and ADC expert.

The status decision follows a state-less protocol, therefore any action is taken based on utmost up-to-date information from AGIS, as opposed to action taken based on previous/current state of the PanDA queue and underlying services.

In 2011 the average duration of a production and analysis job was approximately 8 hours. This number is a baseline for Switcher1 decision taking timeline. Eight hours in advance of a service downtime start Switcher takes decision to disable a PanDA queue for the activity.

Under normal operational conditions the PanDA queue status is set to `online`, all relevant jobs are accepted. When a queue is set to `offline` state, no new jobs are accepted by the PanDA queue, and the already running jobs are allowed to finish. PanDA queue status `brokeroff` has the same relevance for production and analysis PanDA queues: only jobs that are forced to run in the PanDA queue are allowed to run, this PanDA queue is left out of automatic PanDA WMS brokerage decisions. For a PanDA queue status `test` only a very limited number of specialized test jobs are allowed to run at the PanDA queue. Time evolution of PanDA queue state changes decided and executed by the Switcher is listed in Table 7.2.

PanDA queue status	Analysis queue status in time	Production queue status in time
online	up until T - 8 hrs	up until T - 8 hrs
brokeroff	[T - 8 hrs; T - 4 hrs]	n/a
offline	[T - 4 hrs; $T + D$]	[T - 8 hrs; $T + D$]
test	[$T + D$; $T + D + ST$]	
online	after the successful test	

Table 7.2: Workflow of PanDA queue status change set by the Switcher. T denotes the time of downtime start, D denotes duration of the downtime, ST denotes time to successfully test the PanDA queue.

For the duration of the downtime of a SE the PanDA queue is disabled and does not accept new jobs. In case a running job did not finish before the start of downtime T, the job will most probably fail. This is part of a calculated risk. Upon the end of the SE downtime the Switcher initiates PanDA queue recovery. In early days of Switcher this meant that the PanDA queue is set to test and the ADCoS team was notified that the queue has to be manually tested. Later, as HammerCloud AFT became available for functional tests of analysis queues, and HammerCloud PFT became available for functional tests of production queues, the manual intervention of the ADCoS team to run the test jobs was no longer required. In 2012 Switcher and HammerCloud completed a fully automated system for PanDA queue exclusion and recovery based on a downtime of a SE serving the said PanDA queue, without a need for manual exclusion, test, recovery of a PanDA queue.

Switcher2 design

Early in the LS1 a major cleanup and simplification of PanDA site $\leftrightarrow$ PanDA resource $\leftrightarrow$ PanDA queue was performed, leaving us with only one PanDA queue per a PanDA resource. PanDA resource is the entity used for brokerage of ATLAS jobs, therefore this simplification campaign was a natural evolution. The simplified structure of PanDA resources hierarchy gave rise to the second generation of Switcher, the Switcher2. In addition, a clear mapping between a PanDA resource and all underlying SEs and CEs is available from AGIS.

Switcher2 takes and executes decisions on PanDA resource status, based on downtimes of underlying SEs *and* CEs. When a PanDA resource is served by one or multiple SE endpoints, downtime of any of the SE endpoints is considered as "storage is unavailable" for the PanDA resource. When a PanDA resource is served by one or multiple CEs, all the CEs have to be on downtime at the same time for the PanDA resource to be considered as "compute is unavailable".

The status change timeline for Switcher2 is slightly different from that of Switcher1, Switcher2 timelines are listed in Table 7.3.

The notion of *too short*, *normally long* and a *very long* downtimes was introduced. A *too short* downtime is a downtime of a SE or a CE that would result in less than 4 hours of downtime of a PanDA resource for an activity. A *normally long* downtime is a downtime of a SE or a CE that would result in less than 48 hours of downtime of a PanDA resource for an activity. A *very long* downtime is a downtime of a SE or a CE that would result in at least 48 hours of downtime of a PanDA resource for an activity. For a too short downtime of CEs or a SE the PanDA resource is not excluded by the Switcher2, since the site itself is draining the CE resources in advance.

The PanDA resource recovery in Switcher2 is fully automated: Switcher2 notifies HammerCloud that a test job submission is necessary for the PanDA resource in question, and the HammerCloud runs a PFT/AFT job at the production/analysis PanDA resource in question. Once the PFT/AFT test job is completed successfully, the PanDA resource in question is recovered by HammerCloud. No manual intervention from the ADCoS team is needed. The average

PanDA res.	Analysis resource status in time	
status	CE/SE normal length, or CE, too long	SE, too long
online	up until T - 8 hrs	up until T - 120 hrs
brokeroff	[T - 8 hrs; T - 4 hrs]	[T - 120 hrs; T - 72 hrs]
offline	[T - 4 hrs; $T + D$]	[T - 72 hrs; $T + D$]
test	[$T + D$; $T + D + ST$]	
online	after the successful test	

PanDA res.	Production resource status in time	
status	CE/SE normal length, or CE, too long	SE, too long
online	up until T - 12 hrs	up until T - 48 hrs
brokeroff	n/a	n/a
offline	[T - 12 hrs; $T + D$]	[T - 48 hrs; $T + D$]
test	[$T + D$; $T + D + ST$]	
online	after the successful test	

Table 7.3: Workflow of PanDA resource status change set by the Switcher2. T denotes the time of downtime start, D denotes duration of the downtime, ST denotes time to successfully test the PanDA resource.

test duration is approximately 2 hours in 2015.

7.5.4 Switcher activity

Switcher runs once every 20 minutes. The Switcher and Switcher2 monthly activity is shown in Fig. 7.4. The exclusion and recovery actions contribute 50% each, therefore they are not distinguished. From the author's personal experience with ADCoS shift, introduction of Switcher and later cooperation of Switcher(2) and HammerCloud acting as the centralized exclusion tool saved approximately 20% of shifters' time, time which can be and since a long time is spent on investigation of issues that have yet not been reported.

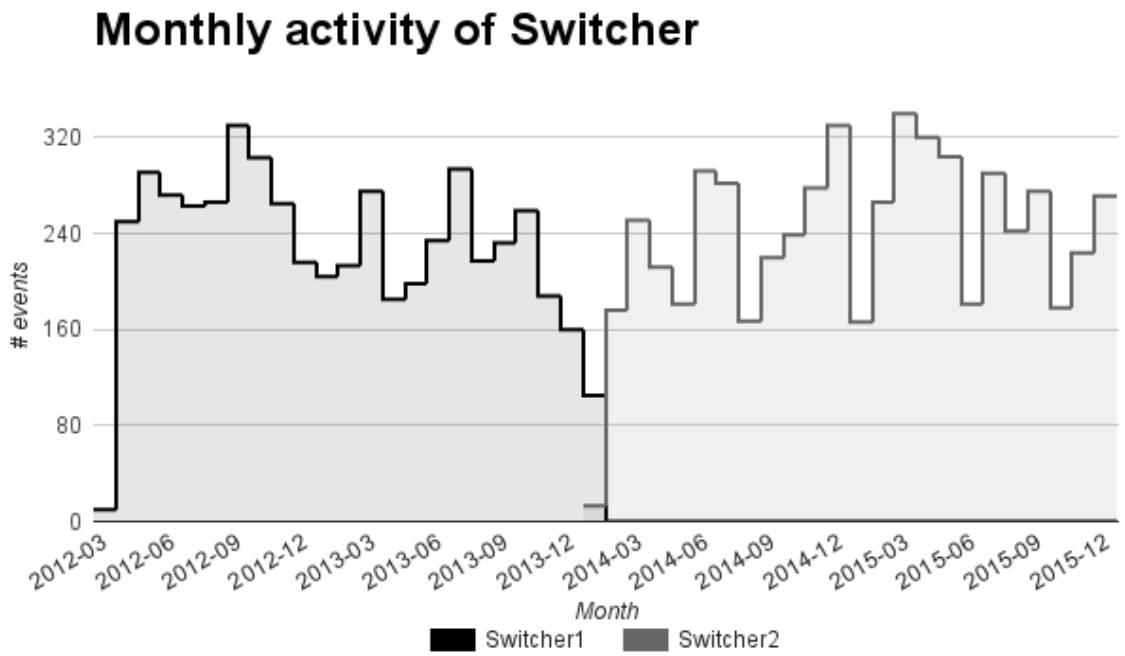


Figure 7.4: Time evolution plot of monthly activity of Switcher1 from March 2012 to January 2014, and of Switcher2 in 2014 and 2015.

8. ATLAS PanDA Workload Management System

The PanDA (Production and Distributed Analysis) Workload Management System [57] plays a key role in the infrastructure of the ATLAS Distributed Computing [104]. ATLAS PanDA manages Monte Carlo simulations and data reprocessing jobs, as well as physics analysis carried out by physics groups or individual physicists. As of March 2014 ATLAS PanDA manages up to 1.5 M jobs per day with 1400 distinct users submitting their analysis jobs through PanDA. PanDA has performed very well during the ATLAS data taking period in the LHC Run I [100]. During the LS1 period various parts of PanDA have been enhanced in order to ensure smooth operation during future data taking periods, and to address needs of users in a variety of science fields of the growing community of the PanDA ecosystem.

In June 2013 – October 2014 author of this Thesis contributed to the following projects:

1. Introduced **multiple DB backend support in PanDA server**, that simplified development of PanDA server for different scientific communities and contributed to a rapid growth of number of PanDA instances round the world.

2. Set up and operated **OSG BigPanDA instance**, that enabled introduction of the The Large Synoptic Survey Telescope (LSST) experiment to grid. We successfully demonstrated that a LSST production workflow can be executed on several distributed grid sites and managed together by the LSST production system and PanDA WMS.

3. Designed and prototyped **BigPanDA monitoring**, implemented several modules that are used by the ATLAS experiment to date. Design enhancements of the PanDA Monitoring are described in Section 8.1.

4. Prototyped the **HTCondor monitoring API**, that utilized the VO-neutral monitoring package of PanDA to monitor a completely different application, the HTCondor [105], and provided a rest-ful API to input the monitoring data.

5. Introduced **PanDA build system** and **BigPanDA monitoring QA suite**, in order to improve development and operational experience with the PanDA instance, and to improve reliability of PanDA monitoring.

The projects were presented at several international conferences: [58], [106], [107], [108], [109], [110].

8.1 PanDA Monitoring

For many years the PanDA Workload Management System has been the basis for distributed production and analysis for the ATLAS experiment at the LHC. Since the start of data taking PanDA usage has ramped up steadily, with up to 1 million completed jobs per day in 2013, and up to 1.5 million completed jobs per day in 2014. The associated monitoring data volume has been rising as well, to levels that present a new set of challenges in the areas of database scalability and monitoring system performance and efficiency. Outside of ATLAS, the PanDA system is also being used in projects like AMS, LSST and a few others. It is currently undergoing a significant redesign, both of the core server components responsible for workload management, brokerage and data access, and of the monitoring part, which is critically important for efficient execution of the workflow in a way that is transparent to the user and also provides an effective set of tools for operational support. The next generation of the PanDA Monitoring System [107] is designed based on a proven, scalable, industry-standard Web Framework – Django. This allows us to achieve significant versatility and possibilities of customization, which is important to cover the needs of the growing community of PanDA users in a variety of science and technology areas. We describe the design principles of the core Web application, the UI layout of the presentation layer, and the challenges that must be met in order to continue the necessary support of the ATLAS experiment while expanding the scope of applications handled by PanDA.

8.1.1 The original ATLAS PanDA monitoring

The current ATLAS PanDA monitoring has evolved over many years. It provides a real-time and short-term-history monitoring tool for PanDA system.

The primary goals of PanDA monitoring are the rapid identification of failures, and monitoring of progress of a distributed physics analysis from the submission of the set of jobs to finalization of the latest running job. Ease in spotting failures allows users to focus on the most serious failures first, collect as much information as possible, and raise the issue with an expert or a responsible person, so that the issue can be resolved.

The PanDA monitor provides very useful summary views with emphasis to roles: a physicist, who runs distributed analysis jobs on the ATLAS distributed computing resources; production manager, who manages large-scale Monte Carlo simulation or data processing campaign on behalf of a physics group or the whole experiment; site administrator, who is monitoring performance of the experiment running jobs on the site; distributed computing operations personnel, who are monitoring health of the overall ATLAS distributed computing resources and chasing failures in a timely manner.

Members of each of the role groups can drill down from the distilled summary information through information about sets of jobs, to very detailed description of a single PanDA job, access log files of that job, or navigate to other ATLAS monitoring tools to follow up various other aspects of distributed computing and

monitor all other parts of the distributed computing infrastructure. The jobs monitoring mentioned in Chapter 7 is performed with the PanDA monitor.

8.1.2 The next generation of ATLAS PanDA monitoring

The next generation of PanDA monitoring is developed as a part of the project "Next Generation Workload Management and Analysis System for BigData" [106]. The project gives us the opportunity to factorize and generalize PanDA monitoring. The aim of the effort is to design generic components and APIs of a monitoring service, which can be customized to address needs of the various experiments of the PanDA ecosystem. The generic components and APIs of the next generation of PanDA monitoring, the BigPanDA monitoring (BigPanDAmon), are described in the following sub-sections.

BigPanDA monitoring at a glance

BigPanDAmon is a modular monitoring package that clearly separates the data access layer and visualization of the data. The monitoring is built around common key objects of the PanDA system, such as PanDA job or PanDA resource. For each of the objects a set of views is available. The view visualizes data in form of tables, plots, and lists. The data access layer provides pre-filtered information describing a set of objects, tailored to the corresponding view. The data access layer benefits from a group of REST API resources.

BigPanDA monitoring backend

BigPanDAmon is a web application built in the Django Web Framework. On the backend it uses Django Web Framework, and several Django plugin libraries, such as Django REST framework, and Django DataTables view.

The monitoring application data is stored in a relational database. The two main database backends supported in BigPanDAmon core are MySQL and Oracle. As Django Web Framework supports even more DB backends (SQLite, PostgreSQL in addition to MySQL and Oracle), the choice of DB backend rests with the community of the experiment/project, which can fully leverage available resources in a flexible way.

Basic object selection is done with Django Querysets. In special case of the model of PanDA job the Queryset has been replaced by its generalization, Querychain. DB queries are optimized using the raw SQL code.

BigPanDA monitoring modules

The BigPanDAmon instance for any of the experiments/projects is composed of two parts: the core part with functionality around the common PanDA objects, and the experiment/project-specific part which extends the common core. The schema of the BigPanDAmon site is depicted in Fig. 8.1.

The common core consists of several Modules which encapsulate functionality to present objects; of UI elements which directly represent either elements of the

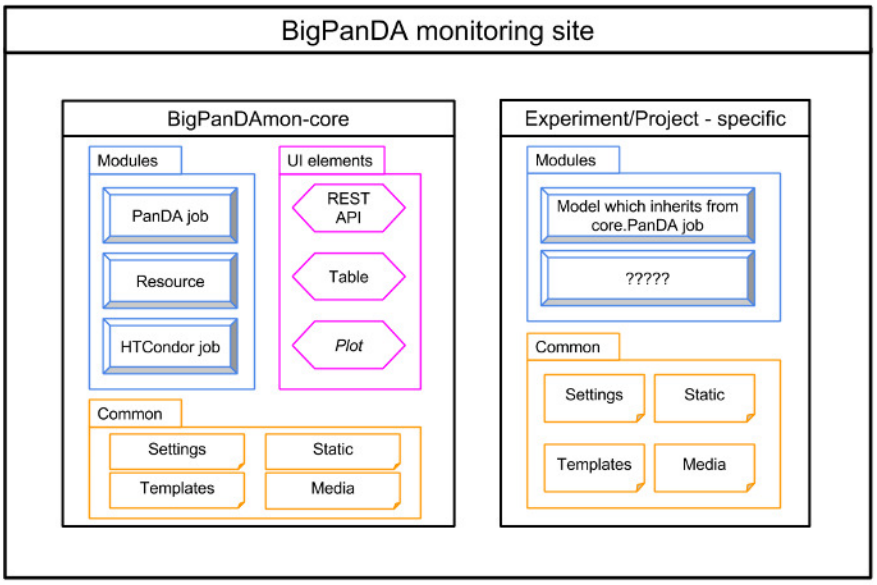


Figure 8.1: Module structure of BigPanDA monitoring package for an experiment/project.

monitoring user interfaces, or provide support for the visual part of the monitoring; and of the Common configuration part not only with the Django application settings, but with templates, static files, and common user-uploaded content.

BigPanDA monitoring customization

The BigPanDAmon frontend consists of a set of views which engage javascript library jQuery and its plugins, such as DataTables and Highcharts.

The core feature of design of monitoring package with clear separation of data layer and visualization lies in the flexibility and ease of customization of the user interfaces of the monitoring. Each experiment/project can extend the provided BigPanDAmon frontend, resulting in a different look-and-feel for each of the BigPanDAmon sites.

Modularity of the BigPanDAmon core allows for experiment/project-specific extension of any of the parts of the core: from feature enhancement of the views around a particular Model, through introduction of a completely new Model central to the PanDA instance of that particular experiment, to modifications or re-organizations of the monitoring views.

The UI elements, such as tables, plots, or APIs can be generalized to visualize information about a completely new Model. The BigPanDA monitoring package can serve as a base framework for both PanDA and external-to-PanDA monitoring, as discussed later in case of the HTCondor monitoring.

BigPanDA monitoring configuration

BigPanDAmon application configuration extends the usual Django application configuration. It separates quasi-static general project configuration, configura-

tion of machine-specific environment with sensible default values, and settings to override machine-specific environment. The core configuration can be extended or overridden in the experiment/project-specific configuration.

8.1.3 REST APIs

BigPanDAmon REST API follows API design recommendation and best practices of Apigee.com [111]. The frontend views are supplied with data from the RESTful APIs. In combination with X509 authentication framework common in the Grid computing world we can benefit from a REST API to not only list available resources, but to create new ones, modify them, or delete them as well.

HTCondor monitoring

HTCondor [105] monitoring is an excellent example of the PanDA monitoring generalization in the BigPanDAmon package. The evolving PanDA-HTCondor REST API supplies data to the HTCondor monitoring interface, which is one of the BigPanDAmon family members.

There are several views of the HTCondor monitoring. Each of the views exposes instances of the HTCondor job object. The HTCondor REST API supplies data with information about the HTCondor jobs to those views. The views are equipped with a filter that enables a user to limit selection of HTCondor jobs desired to monitor. The filter selection reflects in the URL of the view, endorsing multi-user collaborative environment which leverages the BigPanDAmon package. The HTCondor API resources are described in Table 8.1.

Resource	`/v2/api-auth/htcondor/jobs/`	
HTTP verb	Purpose	Description
`POST`	create	Bulk create new HTCondor jobs.
`GET`	read	Bulk list new HTCondor jobs.
`PUT`	update	Bulk update new HTCondor jobs.
`DELETE`	delete	Bulk delete new HTCondor jobs.
Resource	`/v2/api-auth/htcondor/jobs/<ID>`	
HTTP verb	Purpose	Description
`POST`	create	*Error.*
`GET`	read	Show HTCondor job `<ID>`.
`PUT`	update	If exists, update HTCondor job `<ID>`. If it does not exist, `<ID>`.
`DELETE`	delete	Delete HTCondor job `<ID>`.

Table 8.1: REST API resources of the HTCondor monitoring with BigPanDAmon package.

The prototype of HTCondor monitor visualized status of several thousands HTCondor jobs, and used MySQL as a DB backend.

9. Missing Transverse Energy

In this chapter contributions to the studies on transverse energy reconstruction and performance used in Top Physics analyses in 2010 [112] and in 2011 [113] are described. In 2010 the missing tranverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ performance has been studied in the context of re-observation of the top quark signal in data collected with the ATLAS detector at $\sqrt{s} = 7$ TeV.

9.1 Introduction

The missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ is one of the essential quantities that discriminates $\mathrm{t\bar{t}}$ events in the dilepton and lepton+jets channels from backgrounds which do not contain $E_{\mathrm{T}}^{\mathrm{miss}}$, e.g. Drell-Yan processes and QCD multi-jets.

In case of $E_{\mathrm{T}}^{\mathrm{miss}}$ both its magnitude and correlation between its direction and direction of leptons or jets in the event can be used as a rejection criteria. The $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution is of great importance, because the degree of separation depends on it.

There are many theories beyond the Standard Model that predict signatures with large $E_{\mathrm{T}}^{\mathrm{miss}}$, where $\mathrm{t\bar{t}}$ is the dominant background. A good understanding of this kinematic variable, and in particular of the tails of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution, is essential not only for the Top Physics, but for the "Beyond the Standard Model" physics searches as well.

9.2 $E_{\mathrm{T}}^{\mathrm{miss}}$ composition

9.2.1 $E_{\mathrm{T}}^{\mathrm{miss}}$ composition in the early 2010 analyses

In early days of top quark re-observation Athena release 15 has been used. The ATLAS Jet/Etmiss Working Group recommended to use the Cell-based $E_{\mathrm{T}}^{\mathrm{miss}}$ algorithm [114] that uses topological clustering at the *l*ocal hadronic scale (MET_LocHadTopo) with $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstructed inside $|\eta| < 4.5$. Additional muon corrections to $E_{\mathrm{T}}^{\mathrm{miss}}$ were taken into account for early physics analyses.

The reduced coverage in $|\eta|$ rejected inner ring of ATLAS Forward Calorimeter FCAL (spans up to $|\eta| < 4.9$), raising need for ϕ modulation and introduction of corrections to $E_{\mathrm{T}}^{\mathrm{miss}}$.

However, computation of MET_LocHadTopo does not account for jets calibrated at the EM-scale with subsequent Numerical Inversion calculation (EM +JES). Therefore a new algorithm, Met_RefFinal_muID_emjes, has been proposed. It uses Object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ algorithm described in [115], which allows a coherent treatment of jets in $E_{\mathrm{T}}^{\mathrm{miss}}$ computation.

Calculation of Object-based $E_{\mathrm{T}}^{\mathrm{miss}}$ uses calibrated cells that belong to identified high-p_{T} objects in the order of electrons, photons, taus, jets and muons. This approach allows for replacement of the initial cell energies with a more refined calibration. In case a cell belongs to more than 1 objects, only the first association

is taken into account in order to avoid double-counting of the cells in $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. The remaining cells that do not belong to any of the high-p_{T} objects are then included in the CellOut term of $E_{\mathrm{T}}^{\mathrm{miss}}$.

In the early-data top analyses the $E_{\mathrm{T}}^{\mathrm{miss}}$ contributions were calculated [112] as follows:

$$\not{E}_{x,y}^{\mathrm{RefFinal}} = \not{E}_{x,y}^{\mathrm{RefEle}} + \not{E}_{x,y}^{\mathrm{RefJet}} + \not{E}_{x,y}^{\mathrm{RefMuon}} + \not{E}_{x,y}^{\mathrm{RefCellOut}}, \tag{9.1}$$

where

$$E_{\mathrm{T}}^{\mathrm{miss}} = \sqrt{\left(E_x^{\mathrm{miss}}\right)^2 + \left(E_y^{\mathrm{miss}}\right)^2}. \tag{9.2}$$

Each of the terms is calculated from the negative sum of the calibrated cells inside the object. The calorimeter cells are calibrated at the EM-scale. Contributions originating from electrons are included in the RefEle term using electrons passing the isEM::medium definition [116]. Cells that belong to a jet are scaled with the Numerical Inversion scaling factor of the jet. MuId muons contribute to the RefMuon term. The terms for photons and taus were not included since calibrations for these objects were not available at the time and their contributions are small for $\mathrm{t\bar{t}}$ events. Their energies are effectively taken into account either in the RefJet or the CellOut terms. The MET_Cryo terms that are supposed to take into account energy losses in the cryostat were left out as well, since they have not yet been validated at the time.

Composition of $E_{\mathrm{T}}^{\mathrm{miss}}$ for each of its terms and Met_RefFinal_muID_emjes is shown in Fig. 9.1.

9.2.2 $E_{\mathrm{T}}^{\mathrm{miss}}$ composition in Winter 2011 analyses

For the Top analyses aiming at the Winter 2011 conferences Athena release 16 has been used. The ATLAS Jet/Etmiss Working Group recommended to use the objected-based $E_{\mathrm{T}}^{\mathrm{miss}}$ with the topological clusters calibrated to the EM scale, the MET_RefFinal_EM. The Top Group approaches the $E_{\mathrm{T}}^{\mathrm{miss}}$ computations in the similar way: using $E_{\mathrm{T}}^{\mathrm{miss}}$ object definitions that are consistent with the Top object reconstruction definition, resulting in utilization of the MET_RefFinal_EM_tightewtm. This $E_{\mathrm{T}}^{\mathrm{miss}}$ definition uses calorimeter cells whose calibration has been governed by what first high-p_{T} object they have been associated with. These objects are electrons, photons, taus, jets, soft jets, and muons. Ordering of the objects determines order of association. The remaining energy from cells without association to a high-p_{T} object is included in the CellOut term, and calibrated to the EM scale.

In Top analyses for Winter 2011 conferences, the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation [113] follows:

$$E_{x,y}^{\mathrm{miss}} = \not{E}_{x,y}^{\mathrm{RefEle}} + \not{E}_{x,y}^{\mathrm{RefPhoton}} + \not{E}_{x,y}^{\mathrm{RefTau}} + \not{E}_{x,y}^{\mathrm{RefJet}} + \not{E}_{x,y}^{\mathrm{RefSoftJet}} + \not{E}_{x,y}^{\mathrm{RefMuon}} + \not{E}_{x,y}^{\mathrm{CellOut}}. \tag{9.3}$$

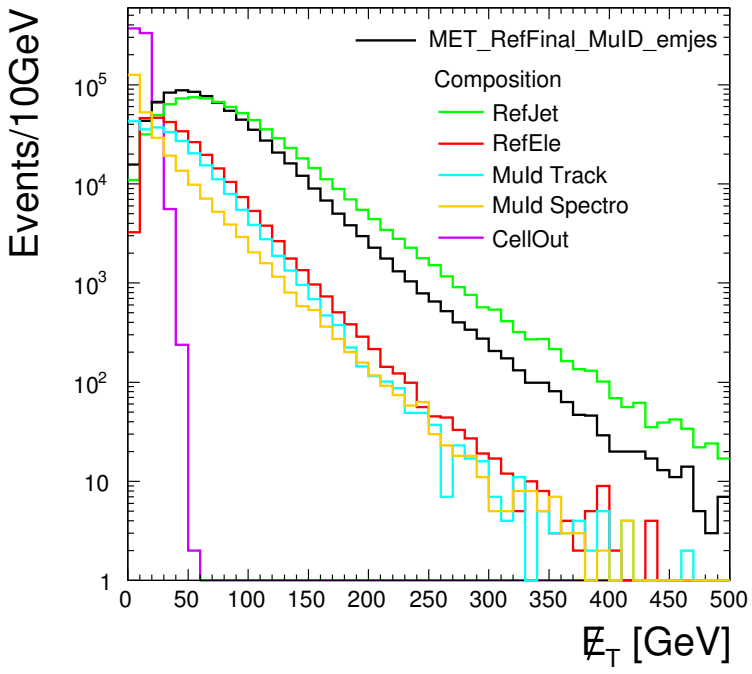


Figure 9.1: Composition of the various $E_{\mathrm{T}}^{\mathrm{miss}}$ terms contributing to MET_REFFINAL_MUID_EMJES in $t\bar{t}$ events in the early analyses, from [112].

The electron term in the $E_{\mathrm{T}}^{\mathrm{miss}}$ is using electrons from the ElectronAODCollection satisfying the ISEM::TIGHT_WITHTRACKMATCH definition with $p_{\mathrm{T}} > 10$ GeV. The energy scale of electrons includes all the electron correction factors except the correction of out-of-cluster.

The photon term in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation is using photons from the PhotonAODCollection, where the photons taken into account in the $E_{\mathrm{T}}^{\mathrm{miss}}$ have to pass the ISEM::TIGHT with a $p_{\mathrm{T}} > 10$ GeV, and the selected photons are added into the $E_{\mathrm{T}}^{\mathrm{miss}}$ at the EM scale.

The τ objects that are included in $E_{\mathrm{T}}^{\mathrm{miss}}$ at EM scale, if they satisfy the TAUCUTSAFETIGHT, come from the TauRecContainer.

The jets can be included upon satisfying 2 criteria: either be one of the refined jets included into the $E_{\mathrm{T}}^{\mathrm{miss}}$ at the EM+JES scale, or one of the soft jets included at the EM scale. The refined jet is any jet from `AntiKt4TopoEMJets` with $p_{\mathrm{T}} > 20$ GeV. The soft jet is any jet with 7 GeV $\leq p_{\mathrm{T}} \leq 20$ GeV.

The muon term in the $E_{\mathrm{T}}^{\mathrm{miss}}$ is using p_{T} of muons from the MuidMuonCollection, for the full acceptance range of the muon spectrometers $|\eta| < 2.7$. All combined muons within $|\eta| < 2.5$ are included in the $E_{\mathrm{T}}^{\mathrm{miss}}$. In addition, the muon term in the $E_{\mathrm{T}}^{\mathrm{miss}}$ contains isolated muons (MET_MU_TRACK) and non-isolated muons (MET_MU_SPECTRO) as well. The MET_MU_TRACK muons require

that the tracks have to be isolated from all `AntiKt4TopoEMJets` (cone size=0.4) by

$$dr = \sqrt{(\delta\phi)^2 + (\delta\eta)^2} = 0.3, \tag{9.4}$$

and includes the muon energy deposited in the calorimeter in the CellOut term. The MET_Mu_Spectro muons have the energy deposited in the calorimeter included in the jet term. More details regarding the muon term are provided in [112].

The $E_{\mathrm{T}}^{\mathrm{miss}}$ and the $E_{\mathrm{T}}^{\mathrm{miss}}$ composition in the $t\bar{t}$ e+jets and μ+jets channels is illustrated in Fig. 9.2.

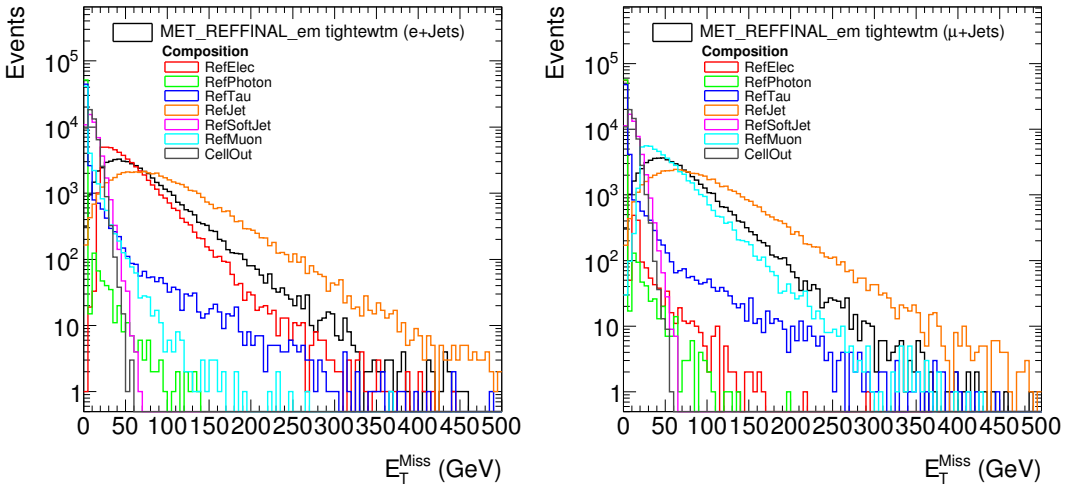


Figure 9.2: The total $E_{\mathrm{T}}^{\mathrm{miss}}$ and the $E_{\mathrm{T}}^{\mathrm{miss}}$ contributions from the electron, photon, τ, jets, soft jets, muons and the CellOut terms are shown for the e+jets selection (left) and μ+jets (right) selection on $t\bar{t}$ MC, from [113].

9.3 Uncertainty for Jet Definition used in the $E_{\mathrm{T}}^{\mathrm{miss}}$ Calculation

Uncertainty for jet definition used in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation using data collected by the ATLAS detector in 2010 at a center of mass energy $\sqrt{s} = 7$ TeV is studied. Samples of simulated $t\bar{t}$ and the corresponding background events in pp collisions at $\sqrt{s} = 7$ TeV were generated using several MC generators. The events generated samples were passed through a full simulation of the ATLAS detector and trigger using the GEANT4 simulation toolkit [117]. The events were reconstructed and selected using the same analysis tools that are used for data. Athena release 16 has been used to reconstruct and analyse MC and data samples.

MC samples generated with different levels of pile-up were used in the study. Samples with no pile-up and with pile-up imitated conditions of structure of LHC bunches of the last 2010 data taking period. The simulated bunch structure

consisted of double trains with 225 ns separation. Each train contains 8 filled bunches with bunch spacing of 150 ns, and the average number of interactions per bunch crossing was set to 2.2.

The jet collection `AntiKt4TopoEMJets` (JET_* collection in D3PD) used in the calculation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ has a pile-up correction, which is over corrected. On the other hand, jet collection `CorrectedAntiKt4TopoEMJets` (JET_COR_ANTIKT4-TOPOEMJETS_* in D3PD) have no pile-up correction applied. Both jet collections are considered incorrect by the Jet/ETmiss group and should not be used for the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. There are three effects that contribute to the discrepancy in $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation:

1. Total energy associated with the jet is reduced by the pile-up subtraction.

2. Reduction of the total energy or p_{T} referable to the pile-up removal may cause that jet fails to pass the p_{T} cut of 20 GeV at the EM+JES scale and is included at the EM scale instead.

3. The JES scale determines difference between the jets in the two collections.

Uncertainty for jet definition used in $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation applies in the following manner: firstly, match jets from both collections at the EM scale. For every matched jet the $\{x, y\}$-scale is defined as ratio of projection of p_{T} of the `CorrectedAntiKt4TopoEMJets` and projection of p_{T} of the `AntiKt4TopoEMJets`.

When any of the matched jets fails to pass the cut of 20 GeV, projection of p_{T} is replaced by projection of E in EM scale for every jet flavour that fails to pass the cut of 20 GeV.

Then pile-up over-correction (weighted `AntiKt4TopoEMJets` p_{T} projection to x, y) is removed (i.e. term with oposite sign is added) from the ${E_{\mathrm{T}}^{\mathrm{miss}}}_{\{x,y\}}$ and the correction term is subtracted (weighted `AntiKt4TopoEMJets` p_{T} projection to x, y scaled with the corresponding $\{x, y\}$-scale from previous step).

Finally, the corrected ${E_{\mathrm{T}}^{\mathrm{miss}}}_{\mathrm{corr}}$ reads

$$ {E_{\mathrm{T}}^{\mathrm{miss}}}_{\mathrm{corr}} = \sqrt{\left(E_{x,\mathrm{corr}}^{\mathrm{miss}}\right)^2 + \left(E_{y,\mathrm{corr}}^{\mathrm{miss}}\right)^2} \quad . \tag{9.5} $$

9.4 Validation of $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction algorithms

In this section the $E_{\mathrm{T}}^{\mathrm{miss}}$ validation tool is described, `TopMetValidation` package, developed in Summer 2010 in order to validate Top $E_{\mathrm{T}}^{\mathrm{miss}}$ terms reconstruction algorithms. This tool has been used to not only to compare effects of different calibrations on the $E_{\mathrm{T}}^{\mathrm{miss}}$ terms, but to spot possible bugs in the reconstruction algorithms that have been implemented in the rapidly evolving Top physics software packages.

The `TopMetValidation` package is based on Athena framework, it is derived from the package `TopInputsValidation-00-00-05`. It implements object selection, validation, and histogramming part. It uses ATLAS AOD data format of MC samples `mc09_7TeV.105200.T1_McAtNlo_Jimmy` as input, and produces a

ROOT file with a set of histograms describing different $E_{\mathrm{T}}^{\mathrm{miss}}$ terms. The output file is processed with a customized ROOT macro in order to create support document with the sets of plots for easy evaluation of $E_{\mathrm{T}}^{\mathrm{miss}}$ properties, and visible comparison of possible changes. The whole `TopMetValidation` package is available to the ATLAS Collaboration in a SVN repository.

The `TopMetValidation` package has been used either for comparison of different scale approaches, or to compare the same scale approaches in different SW releases used by the Top Working Group for physics analyses in Summer 2010 – during the top quark re-observation period – or later.

The `TopMetValidation` package provides access to monitoring of various $E_{\mathrm{T}}^{\mathrm{miss}}$ terms as they become available throughout different Athena Releases: `MET_RefFinal`, `MET_RefEle`, `MET_RefPhotons`, `MET_RefTau`, `MET_RefJet`, `MET_SoftJets`, `MET_RefMuon`, `MET_CellOut`, `MET_Base`, `MET_Base0`, `MET_Calib`, `MET_CorrTopo`, `MET_CorrTopoTower`, `MET_CorrTower`, `MET_Cryo`, `MET_CryoCone`, `MET_DM_All`, `MET_DM_Crack1`, `MET_DM_Crack2`, `MET_DM_Cryo`, `MET_Final`, `MET_LocHadTopo`, `MET_LocHadTopoObj`, `MET_Muid`, `MET_Muid_Spectro`, `MET_Muid_Track`, `MET_Muon`, `MET_MuonBoy`, `MET_MuonBoy_Spectro`, `MET_MuonBoy_Track`, `MET_MuonMuid`, `MET_Muon_Total_Muid`, `MET_RefMuon_Muid`, `MET_RefMuon_Track`, `MET_RefMuon_Track_Muid`, `MET_Topo`, `MET_TopoObj`, `MET_TopoTowerObj`, `MET_TowerObj`, `MET_Track`, `MET_Truth`, `MET_Truth_PileUp`.

Each of the $E_{\mathrm{T}}^{\mathrm{miss}}$ terms may or may not have its variant without scaling, with EM scale, with EM+JES scale, and their evolutions. Each $E_{\mathrm{T}}^{\mathrm{miss}}$ term comes with information about its properties such as E_{T}, $\Sigma\ E_{\mathrm{T}}$, $E_{\mathrm{T},x}$, $E_{\mathrm{T},y}$, ϕ.

The list of 10 $E_{\mathrm{T}}^{\mathrm{miss}}$ validation campaigns performed between Summer 2010 and early 2011 follows:

1. Athena Production Release 15.6.10.7
   SW tags:
   - `MissingET-03-02-27`,
   - `MissingETEvent-00-04-10`,
   - `TopPhysD2PDMaker-00-02-46`,

   Input dataset: `mc09_7TeV.105200.T1_McAtNlo_Jimmy.merge.AOD.e510_s765_s767_r1302_r1306`,
   Purpose: comparison of EM and EM+JES scale for $E_{\mathrm{T}}^{\mathrm{miss}}$ terms.

2. Athena Production Release 15.6.10.7
   SW tags:
   - `MissingET-03-02-35`,
   - `MissingETEvent-00-04-10`,
   - `TopPhysD2PDMaker-00-02-48`,

   Input dataset: `mc09_7TeV.105200.T1_McAtNlo_Jimmy.merge.AOD.e510_s765_s767_r1302_r1306`,
   Purpose: comparison of EM and EM+JES scale for $E_{\mathrm{T}}^{\mathrm{miss}}$ terms.

3. Releases TopPhys-15.6.11.3.1, TopPhys-15.6.12.3.1, TopPhys-15.6.12.4.1
   Input dataset: `mc09_7TeV.105200.T1_McAtNlo_Jimmy.merge.AOD.e510_s765_s767_r1302_r1306`,
   Purpose: comparison of EM and EM+JES scale for $E_{\mathrm{T}}^{\mathrm{miss}}$ terms.

4. Releases TopPhys-15.6.12.3.1, TopPhys-15.6.12.4.1
   Input dataset: `mc09_7TeV.105200.T1_McAtNlo_Jimmy.merge.AOD.e510_s765_s767_r1302_r1306`,
   Purpose: monitoring of changes in behavior of `MET_RefFinal_antikt_emjes` term between TopPhys-15.6.12.3.1 and TopPhys-15.6.12.4.1.

5. Releases TopPhysCache-15.6.12.4.1, TopPhysCache-15.6.12.4.2
   Input dataset: `mc09_7TeV.105200.T1_McAtNlo_Jimmy.merge.AOD.e510_s765_s767_r1302_r1306`,
   Purpose: comparison of EM and EM+JES scale for $E_{\mathrm{T}}^{\mathrm{miss}}$ terms, monitoring of changes in behavior of various $E_{\mathrm{T}}^{\mathrm{miss}}$ terms between 2 releases.

6. Releases TopPhysCache-15.6.12.4.2, TopPhysCache-15.6.12.7.1
   Input dataset: `mc09_7TeV.105200.T1_McAtNlo_Jimmy.merge.AOD.e510_s765_s767_r1302_r1306`,
   Purpose: comparison of EM and EM+JES scale for $E_{\mathrm{T}}^{\mathrm{miss}}$ terms, monitoring of changes in behavior of various $E_{\mathrm{T}}^{\mathrm{miss}}$ terms between 2 releases.

7. Athena Production Release 16.0.2.2
   Input dataset: `valid1.105200.T1_McAtNlo_Jimmy.recon.AOD.e603_s932_r1633`,
   Purpose: comparison of EM and EM+JES scale for $E_{\mathrm{T}}^{\mathrm{miss}}$ terms.

8. Athena Production Release 16.0.2.2, TopPhysCache-15.6.12.7.1
   Input datasets:
   `valid1.105200.T1_McAtNlo_Jimmy.recon.AOD.e603_s932_r1633`,
   `mc09_10TeV.105807.JF35_pythia_jet_filter.evgen.EVNT.e469_tid095268.`
   Purpose: comparison of `MET_RefFinal_antikt_emjes` term from release 15 and `MET_RefFinal_em` term from release 16, looking for a bug in `MET_RefEle` term, comparison of the `MET_Muid_Track` and `MET_Muid_Spectro` terms between these two releases.

9. Athena Production Release 16.0.3.2
   SW tag: `TopInputsSelection-00-03-08`,
   Input dataset:
   `user.jsearcy.mc10_7TeV.105200.T1_McAtNlo_Jimmy.merge.AOD.e598_s933_s946_r1652_r1700.D3PD.v4`
   Purpose: validation of $E_{\mathrm{T}}^{\mathrm{miss}}$ terms in the new Top D3PD format.

10. TopPhys-16.0.3.3.2
    Input datasets:
    `mc10_7TeV.106088.McAtNloZmumu_no_filter.merge.AOD`

.e613_s933_s946_r1831_r1700/,
mc10_7TeV.107660.AlpgenJimmyZmumuNp0_pt20.merge.AOD
.e600_s933_s946_r1831_r1700/,
mc10_7TeV.107661.AlpgenJimmyZmumuNp1_pt20.merge.AOD
.e600_s933_s946_r1831_r1700/,
mc10_7TeV.107662.AlpgenJimmyZmumuNp2_pt20.merge.AOD
.e600_s933_s946_r1831_r1700/,
mc10_7TeV.107663.AlpgenJimmyZmumuNp3_pt20.merge.AOD
.e600_s933_s946_r1831_r1700/,
mc10_7TeV.107664.AlpgenJimmyZmumuNp4_pt20.merge.AOD
.e600_s933_s946_r1831_r1700/,
mc10_7TeV.107665.AlpgenJimmyZmumuNp5_pt20.merge.AOD
.e600_s933_s946_r1831_r1700/,
Purpose: validation of $E_{\mathrm{T}}^{\mathrm{miss}}$ terms in Z$\rightarrow \mu\mu + 0$ jets and Z$\rightarrow \mu\mu+$ 1 or 2 jets MC samples, due to potential bug in jets terms in $E_{\mathrm{T}}^{\mathrm{miss}}$.

Examples of the comparison plots for the 9th validation campaign mentioned earlier are shown in Fig. 9.3, Fig. 9.4, and Fig. 9.5.

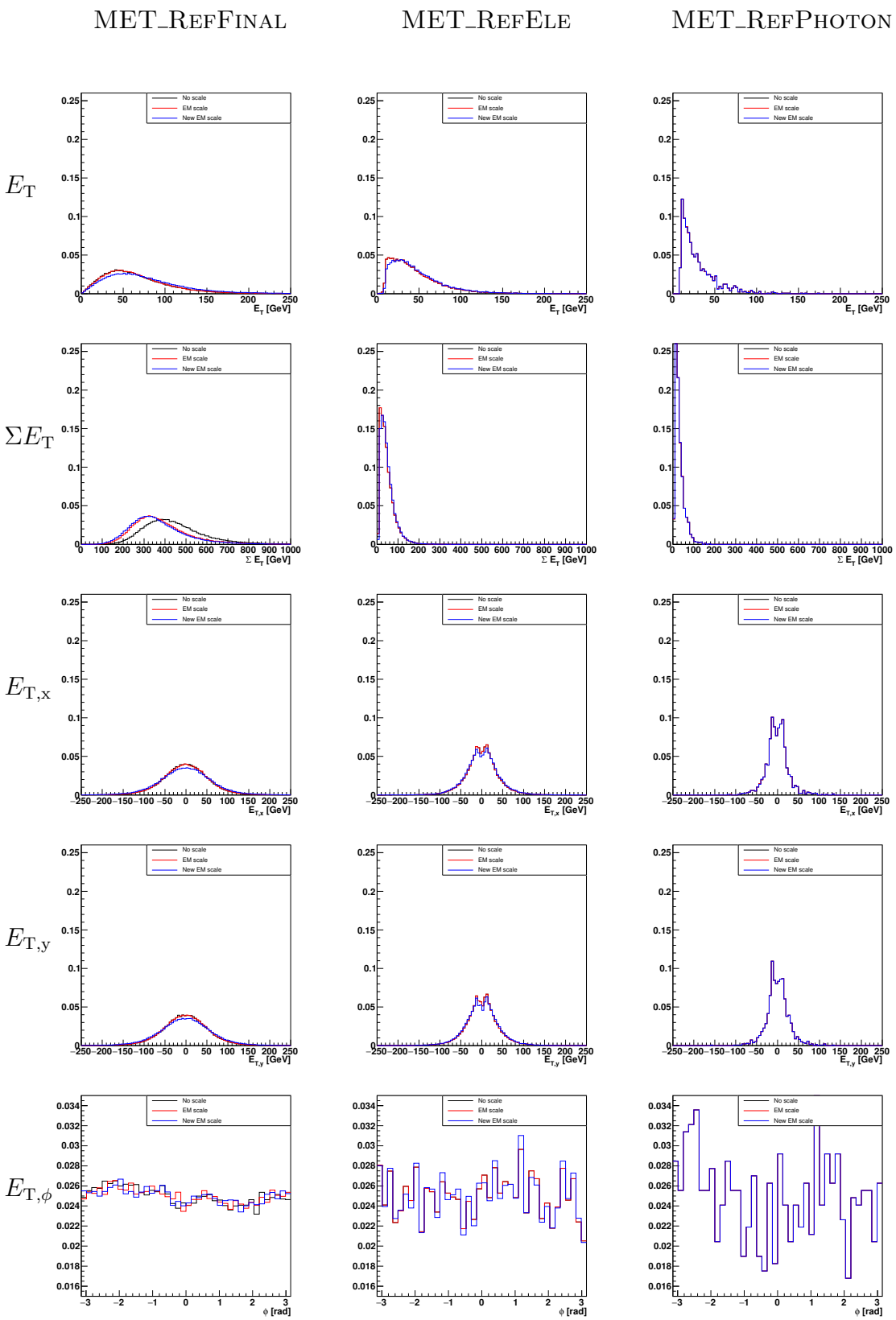


Figure 9.3: Properties of the MET_RefFinal, MET_RefEle, and MET_RefPhoton terms of $E_{\mathrm{T}}^{\mathrm{miss}}$. Comparison of different scales: no scale, the original EM scale, and the new EM scale in Athena release 16.0.3.2.

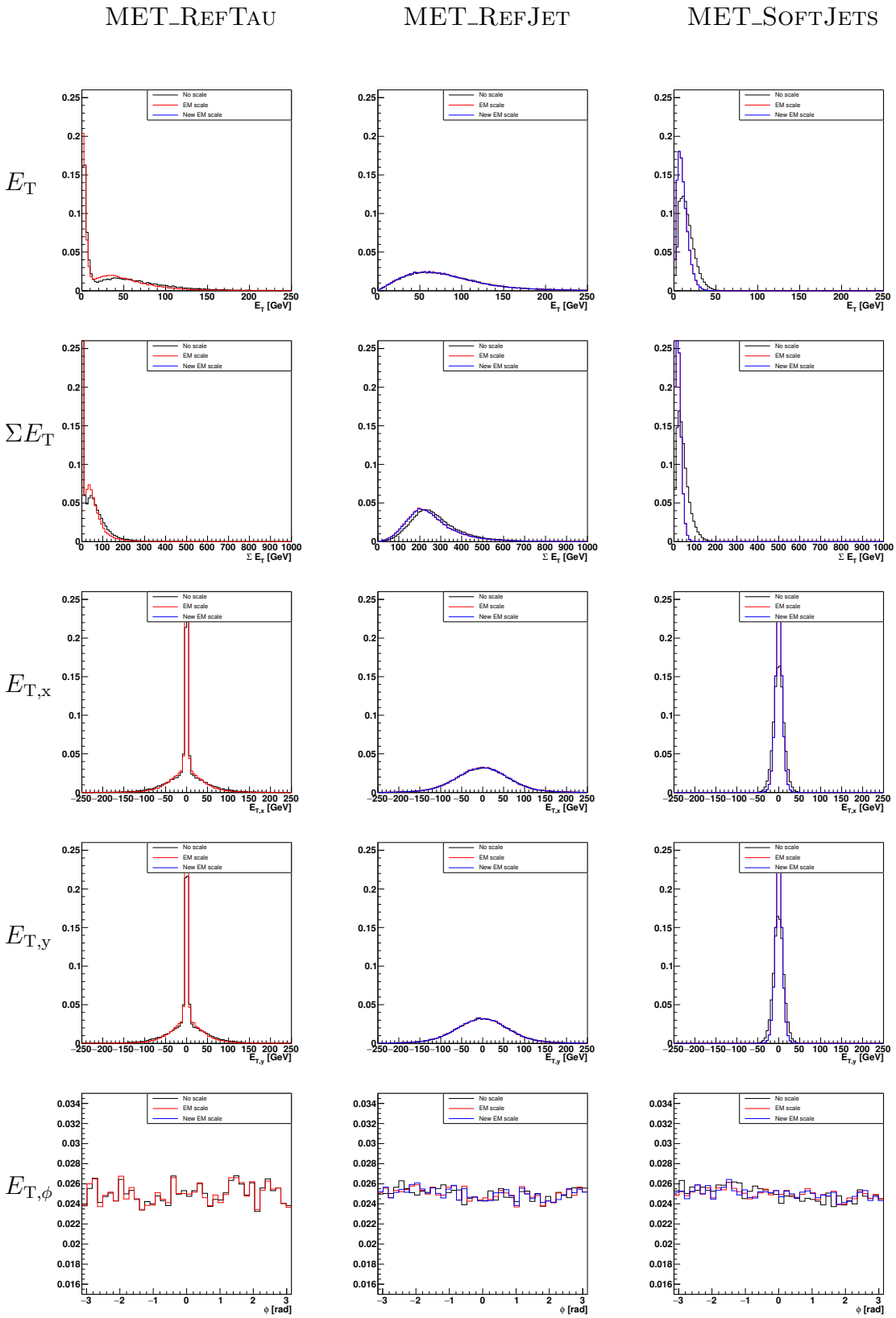


Figure 9.4: Properties of the MET_RefTau, MET_RefJet, and MET_SoftJets terms of $E_{\mathrm{T}}^{\mathrm{miss}}$. Comparison of different scales: no scale, the original EM scale, and the new EM scale in Athena release 16.0.3.2.

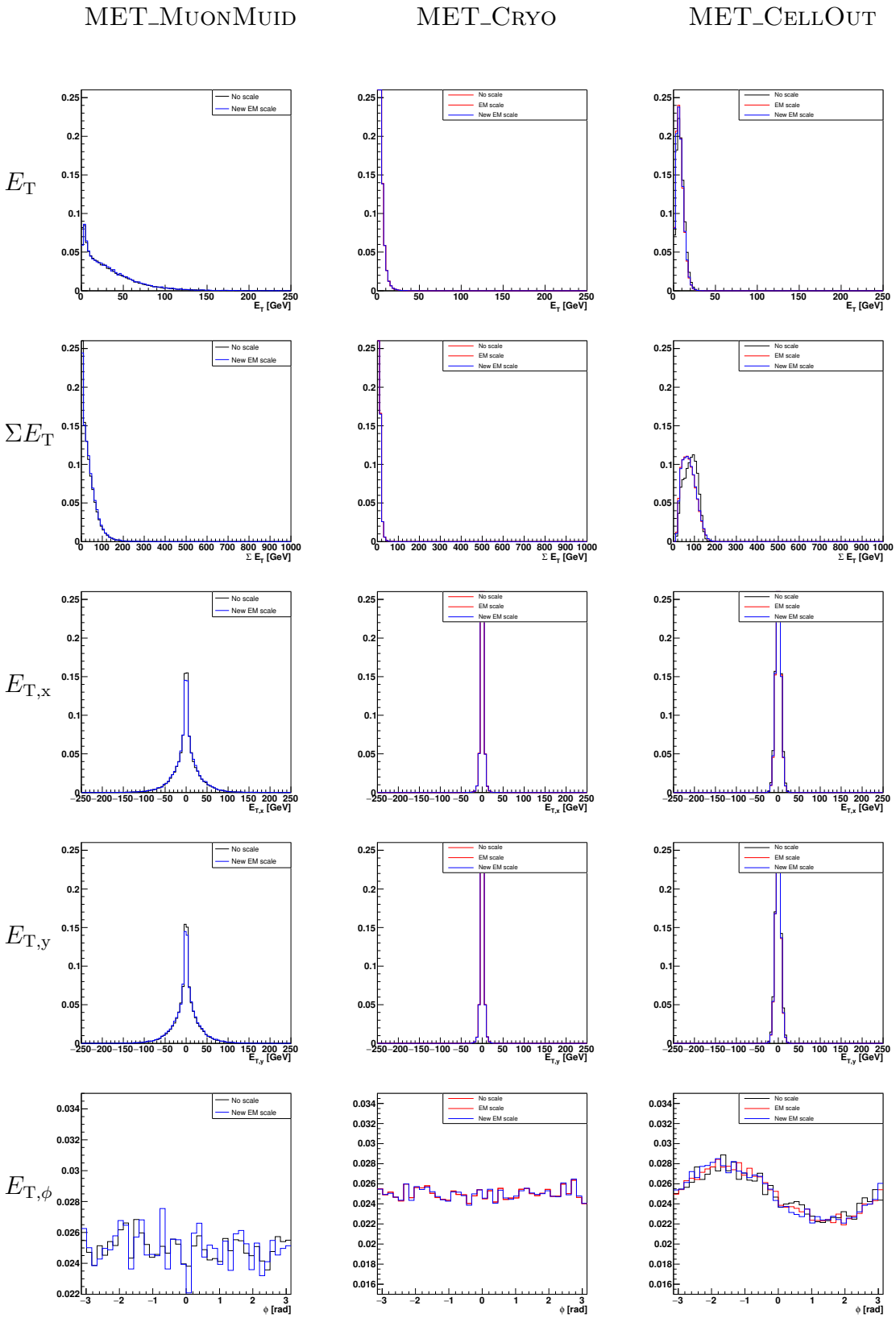


Figure 9.5: Properties of the MET_MuonMuid, MET_Cryo, and MET_CellOut terms of $E_{\mathrm{T}}^{\mathrm{miss}}$. Comparison of different scales: no scale, the original EM scale, and the new EM scale for MET_CellOut, and no scale vs. the new EM scale for MET_MuonMuid, MET_Cryo terms in Athena release 16.0.3.2.

10. Top differential cross-section

This chapter summarizes contribution to the measurement of the top quark differential cross-section as a function of p_{T}, mass and rapidity of the $\mathrm{t\bar{t}}$ system, presented in note [118]. The study was performed using pp collisions data at $\sqrt{s} = 7$ TeV collected by the ATLAS detector in 2011 and corresponding to an integrated luminosity of 4.7 fb^{-1}. The $\mathrm{t\bar{t}}$ events were selected in the lepton (electron or muon) + jets channel. Due to the large $\mathrm{t\bar{t}}$ cross-section at LHC, such measurements allow to study in details the properties of the top quark production and decay, enabling precise test of perturbative QCD.

The core contribution consisted of development of the `TopNtupleAnalysis` package, and of implementation of the event selection in this analysis package for the analysis during the period of rapidly changing selection criteria in 2012.

10.1 `TopNtupleAnalysis` package

Architecture. The `TopNtupleAnalysis` package is an analysis package based on ATLAS RootCore tool, that enables standalone analysis of ATLAS data, i.e. outside the Athena [60] framework, using the ROOT data analysis framework [61]. The `TopNtupleAnalysis` package core is written in C++.

Data and MC analysis. The `TopNtupleAnalysis` package interfaces Monte Carlo simulation files and data files in a similar way, therefore enabling comparison of simulated expectations to the reality of 2011 *pp* stable beams collisions with $\mathrm{t\bar{t}}$ events collected during the 2011 data taking.

Inputs and outputs. The `TopNtupleAnalysis` package uses input data in the format of a flat ROOT ntuple (`D3PD`). Output files are ROOT-readable files as well, e.g. file containing histograms for easy plotting of the resulting figures.

Configuration. The `TopNtupleAnalysis` package provides a configuration object that interfaces configuration of any of the sub-modules of the analysis package from a set of text configuration files, avoiding the need to re-compile the whole package when a single parameter has to be modified. An example of configuration option is shown in Fig. 10.1.

Usability and automation. The chosen configuration approach enabled us to run parameter phase-space studies in an automated way: Besides the core, the `TopNtupleAnalysis` package comes with a set of scripts to automate creation of input file lists, and to run analysis on data and MC with variation of input parameters. The automation effort helped us to minimize unfortunate accidental human errors that could set the analysis back, and to run the analysis jobs in an uniform way that ensures results reproducibility by all the team members, as well as clearly document recipes for the analysis run.

```
##############################################
### TopNtupleAnalysis configuration example ###
##############################################

### Distinguish data from MC
analysis.IsMC:          0

### Good runs selection:
analysis.GRList:configs/Top_GRL_H4.xml

### Selection cuts
# btag cuts:
analysis.TaggerToUse:   MV1

### Binning for histogram
# Tracks:
analysis.etaBinEdges:   -2.5 -1.6 -0.8  0.0  0.8  1.6 2.5
```

Figure 10.1: Examples of a very few of the available `TopNtupleAnalysis` package configuration options.

10.2 Analysis Flow

Analysis presented in [118] follows several steps:

1. Trigger and Good Run List: Analysis uses events that are selected online from triggers highly efficient for the $t\bar{t}$ signals. Only events that come from data taking runs when all the relevant parts of the detector functioned well are taken into account. The standard Top Working Group Good Run List was used.

2. Object Definition: Good identification efficiency and selection of particles from the $t\bar{t}$ sample is the core requirement on object definition. On contrary, particles coming from background processes need to be rejected.

3. Event Selection: Events are reconstructed from the offline objects, and then selected from these reconstructed objects.

4. Reconstruction of the $t\bar{t}$ system: A kinematic fit is performed in order to assign objects in the selected event to a hypothesis that the particles observed in the event come from a $t\bar{t}$ system.

5. Background Determination: Background to the top quark signature was modeled using a mixture of data-driven and Monte Carlo methods. Background contribution is estimated bin-by-bin.

6. Uncertainty estimation: Uncertainty is estimated on each bin in the kinematic distribution, based on various experimental and theoretical uncertainties.

7. Unfolding: Upon the background determination, the measured cross-section can be calculated bin-by-bin[1] as a function of various kinematic variables. However, the measured kinematic distributions are smeared by the detector resolution and acceptance, another approach to unfold the measured distribution to the 'truth' level was used, allowing for easier comparison with the theoretical predictions.

8. Combination: Each channel is treated independently, due to different background and systematics. Later the decay modes are combined using the BLUE [119] method.

Detailed explanation of each of the steps is provided in the note [118].

10.3 Data and Simulation Samples

10.3.1 Data Samples

This analysis used proton-proton collision data collected in 2011 with the ATLAS detector, at the center-of-mass energy $\sqrt{s} = 7$ TeV with an integrated luminosity of 4.7 fb^{-1}. The full data sample is divided into 11 periods. These data taking periods and a corresponding integrated luminosity collected during these periods is listed in Table 10.1.

In 2011 data taking the pile-up conditions increased dramatically. The data for this analysis was collected using single electron and single muon triggers, listed in Table 10.2. The trigger conditions changed during the data taking in 2011, in order to accommodate higher pile-up conditions.

10.3.2 Simulation Samples

This analysis used Monte Carlo simulations samples produced as a part of the MC11 simulation campaign. With respect to the previous campaign MC10, the MC11 campaign provided several improvements:

- improved handling of pile-up, when pile-up was taken from data distributions and weighted at the correct level for the 2011 data,
- geometry updates were included,
- conditions data were included in the simulation,

[1]The binning of each variable has been optimized on simulation based on the resolution of the reconstructed distribution in fine bins of the corresponding truth variable. The resolution curve was extracted and fitted with a 2nd-order polynomial, and the bin edges were extracted by finding the point where the resolution function evaluated at the midpoint of the bin is equal to the width of the bin.

Period	Integrated Luminosity
B	$17\,\mathrm{pb}^{-1}$
D	$179\,\mathrm{pb}^{-1}$
E	$50\,\mathrm{pb}^{-1}$
F	$152\,\mathrm{pb}^{-1}$
G	$560\,\mathrm{pb}^{-1}$
H	$278\,\mathrm{pb}^{-1}$
I	$399\,\mathrm{pb}^{-1}$
J	$232.9\,\mathrm{pb}^{-1}$
K	$660.2\,\mathrm{pb}^{-1}$
L	$1568\,\mathrm{pb}^{-1}$
M	$1121\,\mathrm{pb}^{-1}$

Table 10.1: Luminosity by data period for the 2011 data taking.

Channel	**Period**	**Trigger**
e+jets	B-H	`EF-e20-medium`
e+jets	I-K	`EF-e22-medium`
e+jets	L-M	`EF-e22vh-medium1`
μ+jets	B-I	`EF-mu18`
μ+jets	J-M	`EF-mu18-medium`

Table 10.2: Triggers for the 2011 data in the e+jets or μ+jets channel.

- parton distribution functions and several physics parameters were updated,
- the top quark mass value was taken to be 172.5 GeV.

The simulation samples of the $t\bar{t}$ production and single top production were generated using the MC@NLO (version 4.01) and using the CT10 PDF. The HERWIG [120] (version 6.520) and JIMMY [121] (version 4.31) generators were used to hadronize events and have a complete underlying event model, with the CT10 HERWIG and JIMMY ATLAS AUET2 simulation tune [122]. Production of the vector bosons was simulated using ALPGEN with HERWIG and JIMMY generators.

List of various simulation samples used for this analysis is summarized in Table 10.3.

10.4 Object Definition

Definition of analysis objects used for selection follows the Top working group criteria for analysis of 2011 data, as detailed in [123].

Data Sample	Generator	Dataset Number	Cross-Section [pb]	k-factor
$t\bar{t}$	MC@NLO	105200	79.01	1.146
single top		108340	7.12	0.98
W+0 parton	Alpgen + Herwig, Jimmy	107680	6930.50	1.196
W+1 parton		107681	1305.30	
W+2 parton		107682	378.13	
W+3 parton		107683	101.86	
W+4 parton		107584	25.68	
W+5 parton		107685	6.88	
W+bb+0 NP	Alpgen + Herwig, Jimmy	107280	47.35	1.200
W+bb+1 NP		107281	35.76	
W+bb+2 NP		107282	17.33	
W+bb+3 NP		107283	7.61	
W+c+0j		117293	644.4	
W+c+1j		117294	205.0	
W+c+2j		117295	50.8	
W+c+3j		117296	11.4	
W+c+4j		117297	2.8	
W+cc+0j		117284	127.5	
W+cc+1j		117285	104.7	
W+cc+2j		117286	52.1	
W+cc+3j		117287	16.9	
Z+0 parton	Alpgen + Herwig, Jimmy	107650	668.7	1.25
Z+1 parton		107651	134.8	
Z+2 parton		107652	40.3	
Z+3 parton		107653	11.2	
Z+4 parton		107554	2.5	
Z+5 parton		107655	0.77	
Z+bb+0 parton		109300	6.6	
Z+bb+1 parton		109301	2.5	
Z+bb+2 parton		109302	0.9	
Z+bb+3 parton		109303	0.4	
WW+0 parton	Alpgen + Herwig, Jimmy	107100	2.1	1.26
WW+1 parton		107101	1.0	
WW+2 parton		107102	0.5	
WW+3 parton		107103	0.18	
WZ+0 parton		107104	0.7	
WZ+1 parton		107105	0.4	
WZ+2 parton		107106	0.2	
WZ+3 parton		107107	0.1	
ZZ+0 parton		107108	0.5	
ZZ+1 parton		107109	0.2	
ZZ+2 parton		107110	0.09	
ZZ+3 parton		107111	0.03	

Table 10.3: Summary table of Monte Carlo samples used, generator, dataset number, cross-section, and k-factor.

10.4.1 Muons

Muon objects were selected with the MuID algorithm. In addition, the muon candidates have to pass these requirements:

- Muons have to be combined between the inner detector and the muon spectrometer.
- Muons have to be within the detector acceptance of the inner detector and the muon spectrometer: $|\eta| < 2.5$.
- Muons have to have transverse momentum $p_{\mathrm{T}} > 20$ GeV.
- Muons have to pass inner detector track quality selection of the Muon Combined Performance group.
- Transverse energy in a code with radius $\Delta\mathrm{R} = 0.2$ around the muon has to be less than 4 GeV. Transverse energy is measured by energy deposits in the calorimeter.
- The sum of p_{T} of tracks in a cone with radius $\Delta\mathrm{R} = 0.3$ around the muon has to be less than 2.5 GeV.
- Muons have to be well separated from any high transverse energy jet: $\Delta\mathrm{R}(\mu, \mathrm{jet}) > 0.4$ for any jet with transverse energy $E_{\mathrm{T}} > 25$ GeV and jet vertex fraction (JVF) has to be $|\mathrm{JVF}| > 0.75$. JVF corresponds to the fraction of tracks associated to the jet and coming from the primary vertex in order to suppress the effect of the in-time pile-up.

10.4.2 Electrons

Electrons were reconstructed as energy deposits (clusters, `cl`) in the electromagnetic calorimeter, associated to the tracks in the inner detector. In addition, the electron candidates have to pass these requirements:

- Electrons have to have a transverse energy of $E_{\mathrm{T}} > 25$ GeV and their cluster size has to be $|\eta_{\mathrm{cl}}| < 2.47$, while the region between the barrel and the endcap is excluded ($1.37 < |\eta_{\mathrm{cl}}| < 1.52$). Electrons have to further satisfy the Tight++ criteria, in order to assure good separation between an isolated electron and a jet.
- Azimuthal position difference between the cluster in the 2nd layer of the calorimeter and the track in the inner detector has to be $\Delta\phi < 0.2$.
- Cluster's ratio E/p has to satisfy constrains that depend on η and E_{T}.
- In the transition radiation tracker, the difference between measured hits and expected hits has to be less than 15.
- A η dependent selection is performed on a fraction of tracks in the transition radiation tracker that are high-threshold hits.

- Electrons that matched converted photons were rejected.

- Jets with axis within $\Delta R = 0.2$ from the electron candidate are removed from the event. Upon application of jet-electron overlap removal, if there is yet another jet within $\Delta R = 0.4$ with $p_T > 20$ GeV, the electron candidate is rejected.

10.4.3 Jets

Jets were reconstructed with the anti-k_T algorithm [124], [125]. Jet search was performed on topological clusters in the calorimeter, evaluated at the electromagnetic (EM) scale and calibrated afterwards. Pile-up subtraction scheme was applied to jets at EM scale. Jets are then calibrated to the hadronic scale. Jet quality criteria were applied and 'bad jets' removed. Jets with jet vertex fraction $|JVF| < 0.75$ are rejected. Jets have to have transverse energy $E_T > 25$ GeV and $|\eta| < 2.5$.

10.4.4 b-tagging

b-tagging of a jet is used to reduce backgrounds. A combination of 3 b-tagging algorithms was used: JetFitter, IP3D, SV1. Weights of these algorithms and p_T and η of the jet are used as input of the MV1-tagger, in order to determine a single discriminant variable. Efficiency of MV1-tagger is tuned to select b-jets with a 60% efficiency.

10.4.5 Missing Transverse Energy

The E_T^{miss} is calculated using MET_REFFINAL algorithm [126]. More details about E_T^{miss} selection are provided in [123].

10.5 Backgrounds

The following backgrounds were considered in this analysis:

- W boson production in association with multiple jets (W+jets),

- Z boson production in association with multiple jets (Z+jets),

- single top production,

- multijet production,

- diboson production (WW,WZ,ZZ).

10.6 Selection of $t\bar{t}$ events in the ℓ + jets channel

This analysis followed selection criteria prescribed by the Top Working Group for the 2011 data analyses [123]. The same selection criteria and various correction factors added to Monte Carlo as in the inclusive cross-section measurement [127] were used in this analysis.

The event selection of this analysis focused on the ℓ+jets channel with these signatures:

- an isolated high momentum lepton,
- missing transverse momentum,
- multiple jets.

The high-p_{T} lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$ are both signature of a W boson decaying into a lepton and neutrino. The jet requirement helps with identification of a jet originating from a hadronically decaying W boson, and hadronization of the b quark from the top quark decay. In addition to jet selection the b-tagging is one of the requirements, since each of the $t\bar{t}$ events is expected to have 2 b-jets.

10.6.1 Trigger

This analysis benefited from single lepton triggers which ran without pre-scale during the 2011 data taking.

The muon channel analysis required the event to pass the `EF-mu18` trigger chain:

- events pass the `L1-MU10` trigger: the event has a muon with $p_{\mathrm{T}} > 10$ GeV at L1,
- at L2 this is combined with an inner detector track,
- at the EF level the event has a combined muon with a $p_{\mathrm{T}} > 18$ GeV.

The electron channel analysis required the event to pass the `EF-e20-medium` (`EF-e22-medium`) trigger chain for the data-taking periods B-H (I-M):

- an electromagnetic cluster at L1 with $E_{\mathrm{T}} > 14$ GeV,
- shower shape requirements are satisfied and a track match are required at L2,
- an EF electron with $E_{\mathrm{T}} > 20$ GeV in the data-taking periods B-H, or $E_{\mathrm{T}} > 22$ GeV in the periods I-M.

10.6.2 Muon channel

The μ+jets channel analysis required the following conditions to be fulfilled:

- One primary vertex with more than four tracks attached to it.
- Exactly one good muon passing the selection described above. Events with more than one good muon were rejected in order to maintain orthogonality to the dilepton analysis.
- No good electron in the event, to keep the μ+jets channel analysis orthogonal to the e+jets channel analysis.
- The reconstructed muon is matched to the trigger muon.
- Events with any bad jets were rejected.
- At least four jets passing the jet selection described above.
- Missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV.
- Missing transverse energy plus the transverse mass greater than 60 GeV. The transverse mass was evaluated using the reconstructed lepton and the missing transverse energy.
- At least one b-tagged jet (i.e. tagged as originating from a b-quark) following the description above.

10.6.3 Electron channel

The e+jets channel analysis required the following conditions to be fulfilled:

- One primary vertex with more than four tracks attached to it.
- Exactly one good electron passing the selection described above. Events with more than one good electron were rejected in order to maintain orthogonality to the dilepton analysis.
- No good muons in the event, to keep the e+jets channel analysis orthogonal to the μ+jets channel analysis.
- The reconstructed electron is matched to the trigger electron.
- Events with any bad jets were rejected.
- At least four jets passing the jet selection described above.
- Missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV.
- Transverse mass greater than 30 GeV. The transverse mass was evaluated using the reconstructed lepton and the missing transverse energy.
- At least one b-tagged jet tagged following the description above.

10.6.4 Event yields of the Monte Carlo samples

The event yields after various event selection steps are listed in Table 10.4 for the μ+jets channel, and in Table 10.5 for the e+jets channel.

Selection	Raw count	Weighted	Pileup Reweighting	Cumul. Eff. [%]
Initial	14879305	11503845	11501109.5	
Trigger	4245347	3291101	3294362.8	28.6
Primary Vertex Requirement	4245347	3291101	3294362.8	28.6
One good muon	3179878	2470396	2473004.1	21.5
Exactly one muon	3016136	2470396	2345932.2	20.4
Channel Orthogonality	2857157	2343430	2221304.2	19.3
Trigger Match	2810420	2181832	2184433.3	19.0
Muon overlap removal	2810420	2181832	2184433.3	19.0
Jet Cleaning	2774955	2154413	2157237.4	18.75
2 or more jets	2612517	2027735	2030740.3	17.7
3 or more jets	2107809	1632085	1634712.8	14.2
4 or more jets	1234467	946558	946557.0	8.2
$E_{\mathrm{T}}^{\mathrm{miss}}$ selection	1127418	865084	866469.0	7.5
Transverse mass selection	1066462	816966	818388.9	7.1
b-tagging	942526	724554	725840.1	6.3

Table 10.4: Event selection results on $\mathrm{t\bar{t}}$ Monte Carlo in μ+jets channel.

Selection	Raw count	Weighted	Pileup Reweighting	Cumul. Eff. [%]
Initial Sample	14879305	11503845	11501109.5	
Trigger Requirement	3930180	3050538	3055641.3	26.5
Primary Vertex Requirement	3927659	3048535	3053632.6	26.5
At least one electron	2405220	1878200	1877506.4	16.3
Only one electron	2342834	1829390	1828680.6	15.9
Channel Orthogonality	2142174	1672334	1671612.6	14.5
Trigger Match	2140365	1670969	1670261.3	14.5
Electron and μ overlap	2140265	1670887	1670178.6	14.4
Jet Cleaning	2114236	1650416	1649873.6	14.5
2 or more jets	1994849	1556535	1556297.2	13.5
3 or more jets	1608688	1251882	1251871.9	10.8
4 or more jets	942007	724951	725377.4	6.3
$E_{\mathrm{T}}^{\mathrm{miss}}$ selection	750713	578675	578986.2	5.0
Transverse mass selection	640642	492538	492914.6	4.2
One or more b-tagged jet	567262	437064	437413.9	3.8

Table 10.5: Event selection results on $\mathrm{t\bar{t}}$ Monte Carlo in e+jets channel.

10.7 Systematics uncertainties

There is a number of systematic uncertainties that effect the precision of the presented measurements. The systematics can be categorized in several ways. The systematics sources and their overview is documented in [123].

One can distinguish between systematics affecting both μ+jets and e+jets channels, and systematics evaluated on the full simulation or on dedicated fast simulation samples. In particular, systematics affecting signal and background modeling, detector modeling and the luminosity uncertainty were discussed in [118].

10.8 Cross-section unfolding

10.8.1 Cross-section unfolding for $\frac{\mathrm{d}\sigma}{\mathrm{d}M_{\mathrm{t\bar{t}}}}$

This section presents overview of results of the measurement of the $\mathrm{t\bar{t}}$ production cross-section as a function of the mass of the $\mathrm{t\bar{t}}$ pair. In Fig. 10.2 the reconstructed mass of the $\mathrm{t\bar{t}}$ system in the μ+jets and e+jets channels is shown, after all the selection cuts and in the mass bins used in the unfolding analysis. The final unfolded spectra of the normalized differential cross-section is shown in Fig. 10.3.

The plots show an overall good agreement between data and prediction. Both NLO prediction from MCFM and NLO + NNLL [128] calculation agree well with data within uncertainties. The MCFM prediction, evaluated using the CT10 PDF set, includes the uncertainty from the simultaneously varied renormalization and factorizations scales from the nominal $\mu = m_t$ to $\mu \in m_t/2; 2m_t$, and also at the dynamic scale choice of $\mu = m_{\mathrm{t\bar{t}}}/2$.

For each bin of $M_{\mathrm{t\bar{t}}}$ the unfolded cross-section and normalized cross-section are listed in Table 10.6 and Table 10.7.

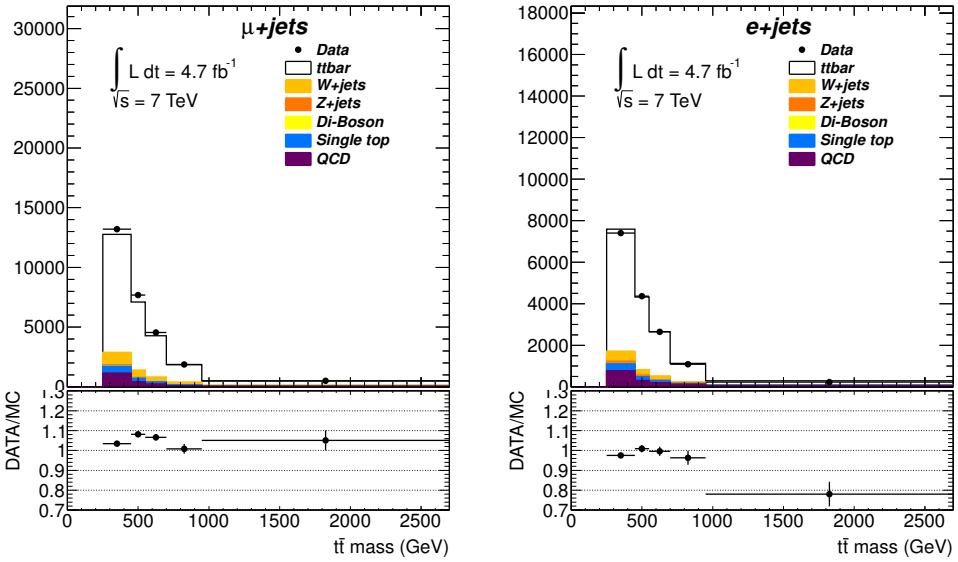


Figure 10.2: Reconstructed $M_{\mathrm{t\bar{t}}}$ in μ+jets and the e+jets channels, from [118].

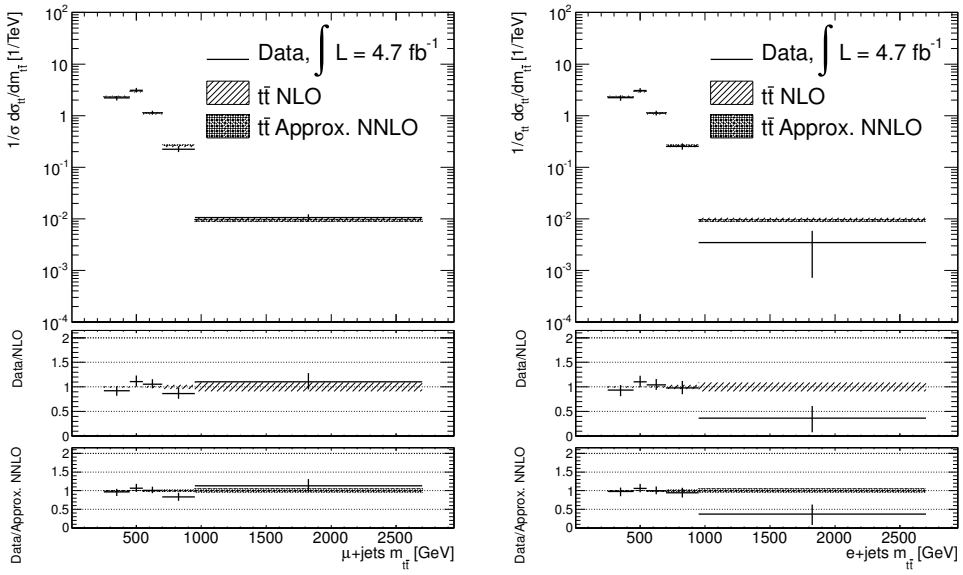


Figure 10.3: Unfolded normalized differential cross-section in μ+jets and the e+jets channels for the $M_{\mathrm{t\bar{t}}}$ spectrum, compared to the NLO prediction from MCFM, from [118].

$m_{t\bar{t}}$ [GeV]	$\mathrm{d}\sigma/\mathrm{d}m_{t\bar{t}}$[fb/GeV] ($\mu$+jets)	$\mathrm{d}\sigma/\mathrm{d}m_{t\bar{t}}$[fb/GeV] ($e$+jets)	$\mathrm{d}\sigma/\mathrm{d}m_{t\bar{t}}$[fb/GeV] ($\ell$+jets)
250 – 450	358 +120.2 / -81.38	352.3 +154.8 / -104.8	358.1 +121.1 / -80.79
450 – 550	515.2 +192.6 / -132.5	492.1 +159 / -118.1	492.6 +160.4 / -117.9
550 – 700	188.3 +61.2 / -42.84	179.4 +52.59 / -40.21	177.1 +51.69 / -40.86
700 – 950	36.99 +13.76 / -9.268	39.96 +14.04 / -9.314	39.59 +13.7 / -8.935
950 – 2700	1.739 +0.657 / -0.4267	0.5426 +0.4823 / -0.441	1.088 +0.4742 / -0.3508

Integrated cross-section (μ+jets): 163.7 ± 27.22 pb
Integrated cross-section (e+jets): 157.5 ± 30.39 pb
Integrated cross-section (ℓ+jets): 159.2 ± 25.65 pb
Integrated cross-section (Theory): $152.0\ ^{+27.7}_{-20.9}$ pb

Table 10.6: Summary of unfolded cross-sections in each bin of $M_{\mathrm{t\bar{t}}}$, using pseudo experiments including folding with full detector simulation, followed by the unfolding procedure, from [118].

Relative cross-section

$m_{t\bar{t}}$ [GeV]	$1/\sigma\ \mathrm{d}\sigma/\mathrm{d}m_{t\bar{t}}$($\mu$+jets)	$1/\sigma\ \mathrm{d}\sigma/\mathrm{d}m_{t\bar{t}}$($e$+jets)	$1/\sigma\ \mathrm{d}\sigma/m_{t\bar{t}}$($\ell$+jets)
250 – 450	2.208 +0.1939 / -0.2501	2.239 +0.2473 / -0.2945	2.224 +0.1513 / -0.2113
450 – 550	3.104 +0.3395 / -0.2871	3.089 +0.3472 / -0.3116	3.095 +0.2405 / -0.196
550 – 700	1.131 +0.1103 / -0.08806	1.116 +0.1278 / -0.1067	1.126 +0.1027 / -0.07969
700 – 950	0.2235 +0.03026 / -0.02714	0.2533 +0.03651 / -0.03362	0.2349 +0.02619 / -0.02246
950 – 2700	0.01055 +0.001684 / -0.001478	0.003471 +0.002389 / -0.002754	0.008647 +0.001392 / -0.001366

Table 10.7: Summary of unfolded relative cross-sections in each bin of $M_{\mathrm{t\bar{t}}}$, using pseudo experiments including folding with full detector simulation, followed by the unfolding procedure, from [118].

10.8.2 Cross-section unfolding for $\frac{d\sigma}{dy_{t\bar{t}}}$

This section presents overview of results of the measurement of the $t\bar{t}$ production cross-section as a function of the rapidity of the $t\bar{t}$ pair. In Fig. 10.4 the reconstructed rapidity y of the $t\bar{t}$ system in the μ+jets and e+jets channels is shown, after all the selection cuts and in the y bins used in the unfolding analysis. The final unfolded spectra of the normalized differential cross-section is shown in Fig. 10.5.

The measured differential cross-section is in good agreement with the predictions from MCFM agree well with data within uncertainties. The MCFM prediction, evaluated using the CT10 PDF set, includes the uncertainty from the simultaneously varied renormalization and factorizations scales from the nominal $\mu = m_t$ to $\mu \in m_t/2; 2m_t$, and also at the dynamic scale choice of $\mu = m_{t\bar{t}}/2$.

For each bin of y the unfolded cross-section and normalized cross-section are listed in Table 10.8 and Table 10.9.

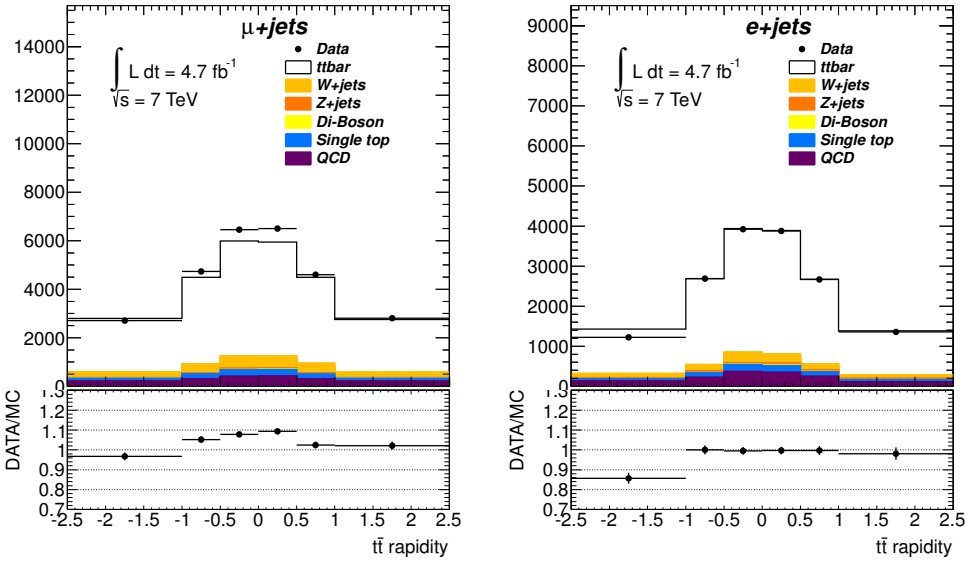


Figure 10.4: Reconstructed $y_{t\bar{t}}$ in μ+jets and the e+jets channels, from [118].

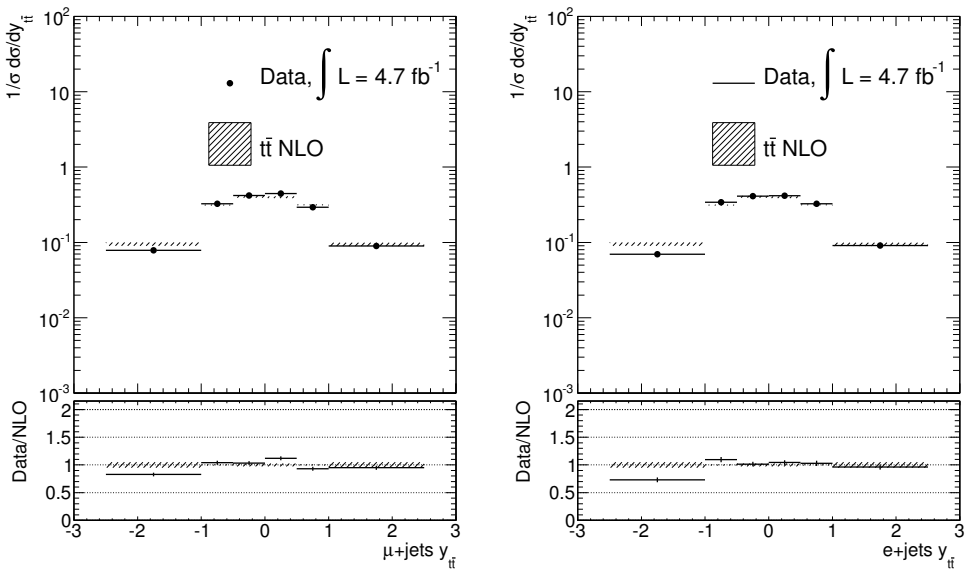


Figure 10.5: Unfolded normalized differential cross-section in μ+jets and the e+jets channels for the $y_{\mathrm{t\bar{t}}}$ spectrum, compared to the NLO predictions from MCFM, from [118].

$y_{t\bar{t}}$	$\mathrm{d}\sigma/\mathrm{d}y_{t\bar{t}}$[pb/GeV] ($\mu$+jets)	$\mathrm{d}\sigma/\mathrm{d}y_{t\bar{t}}$[pb/GeV] ($e$+jets)	$\mathrm{d}\sigma/\mathrm{d}y_{t\bar{t}}$[pb/GeV] ($\ell$+jets)
-2.5 – -1.0	13.53 +5.326 / -3.603	11.83 +4.444 / -3.165	11.05 +4.118 / -3.292
-1.0 – -0.5	54.37 +17.53 / -13.13	53.03 +19.21 / -13.65	54.18 +17.66 / -13.24
-0.5 – 0.0	70.08 +22.98 / -14.63	67.47 +23.20 / -16.49	70.61 +22.56 / -14.79
0.0 – 0.5	73.05 +23.78 / -17.63	64.57 +21.26 / -16.16	66.12 +21.77 / -16.15
0.5 – 1.0	49.15 +15.36 / -10.46	52.79 +18.59 / -12.96	47.46 +14.76 / -10.15
1.0 – 2.5	15.34 +5.884 / -3.990	14.89 +5.485 / -3.922	15.16 +5.607 / -3.917

Integrated cross-section (μ+jets): 166.6 ± 19.89 pb
Integrated cross-section (e+jets): 159.0 ± 19.95 pb
Integrated cross-section (l+jets): 158.5 ± 18.91 pb

Table 10.8: Summary of unfolded cross-sections in each bin of y, using pseudo experiments including folding with full detector simulation, followed by the unfolding procedure, from [118].

Relative cross-section

$y_{t\bar{t}}$	$1/\sigma\ \mathrm{d}\sigma/\mathrm{d}y_{t\bar{t}}$($\mu$+jets)	$1/\sigma\ \mathrm{d}\sigma/\mathrm{d}y_{t\bar{t}}$($e$+jets)	$1/\sigma\ \mathrm{d}\sigma/\mathrm{d}y_{t\bar{t}}$($\ell$+jets)
-2.5 – -1.0	0.08069 +0.005421 / -0.005083	0.07399 +0.005296 / -0.005463	0.07693 +0.004072 / -0.004160
-1.0 – -0.5	0.3220 +0.01730 / -0.01649	0.3306 +0.01808 / -0.01845	0.3262 +0.01228 / -0.01195
-0.5 – 0.0	0.4221 +0.02340 / -0.02280	0.4202 +0.02061 / -0.01981	0.4208 +0.01657 / -0.01574
0.0 – 0.5	0.4341 +0.02145 / -0.02232	0.4012 +0.01801 / -0.01999	0.4127 +0.01554 / -0.01509
0.5 – 1.0	0.2949 +0.01644 / -0.01609	0.3300 +0.01872 / -0.01805	0.3125 +0.01219 / -0.01194
1.0 – 2.5	0.09168 +0.004794 / -0.005012	0.09293 +0.005773 / -0.005517	0.09226 +0.003942 / -0.003976

Table 10.9: Summary of unfolded relative cross-sections in each bin of y, using pseudo experiments including folding with full detector simulation, followed by the unfolding procedure, from [118].

10.8.3 Cross-section unfolding for $\frac{d\sigma}{dp_T^{t\bar{t}}}$

This section presents overview of results of the measurement of the $t\bar{t}$ production cross-section as a function of the transverse momentum of the $t\bar{t}$ pair. In Fig. 10.6 the final unfolded relative differential cross-section in the μ+jets and e+jets channels is shown. The final unfolded spectra of the relative differential cross obtained by combining the μ+jets and e+jets channels is shown in Fig. 10.7.

The measured differential cross section has quite large discrepancy respect to the NLO prediction from MCFM in the lowest $p_T^{t\bar{t}}$ bins. This is due to the fact that such kinematic variables is very sensitive to extra gluon radiation and the parton shower, which is not included in MCFM, gives important corrections in the low $p_T^{t\bar{t}}$region. At larger $p_T^{t\bar{t}}$, where the parton shower corrections become negligible the agreement is restored. On the other hand the predictions from MC@NLO, which include the parton showers, are in reasonable agreement with the measured spectra in the whole $p_T^{t\bar{t}}$range.

The MCFM prediction, evaluated using the CT10 PDF set, includes the uncertainty from the simultaneously varied renormalization and factorizations scales from the nominal $\mu = m_t$ to $\mu \in m_t/2; 2m_t$, and also at the dynamic scale choice of $\mu = m_{t\bar{t}}/2$. The MC@NLO uncertainties include only those of statistical nature related to the finite generator statistics.

For each bin of p_T the relative differential cross-section is listed in Table 10.10.

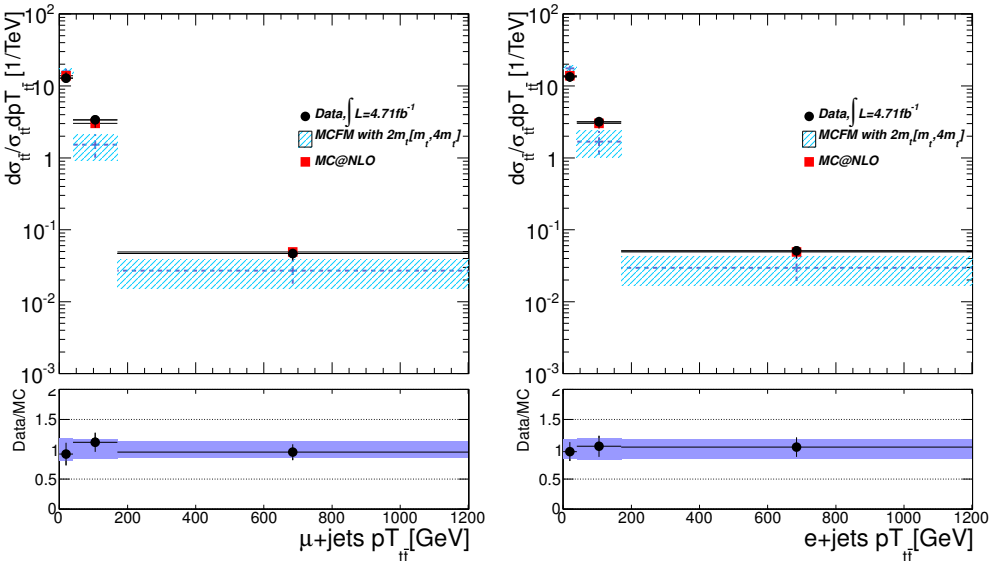


Figure 10.6: Unfolded relative differential cross-section in the μ+jets and the e+jets channels, from [118].

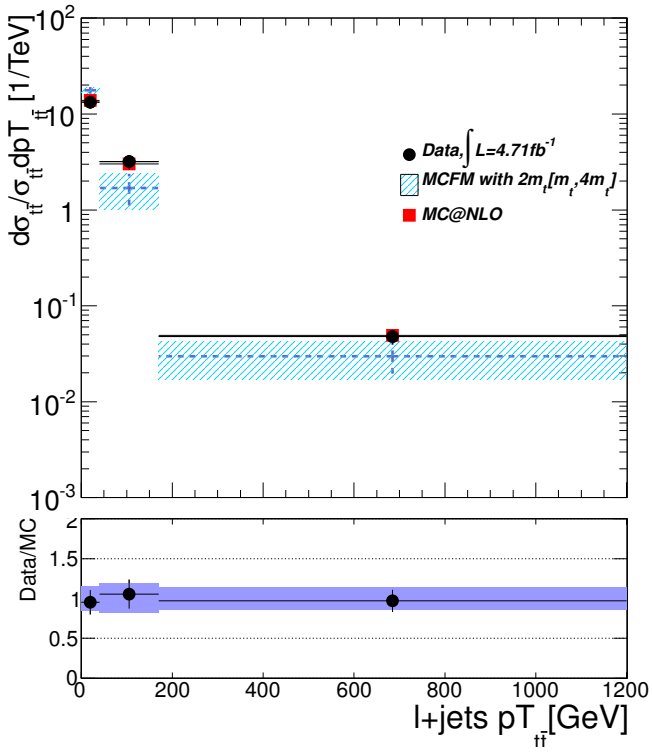


Figure 10.7: Unfolded relative differential cross-section obtained combining the μ+jets and the e+jets channels, from [118].

$p_{\mathrm{T}}^{\mathrm{tt}}$ [GeV]	$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}^{\mathrm{tt}}}$ [1/TeV] (μ+jets)	$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}^{\mathrm{tt}}}$ [1/TeV] (e+jets)	$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}^{\mathrm{tt}}}$ [1/TeV] (ℓ+jets)
0–40	13+3/-2	13+2/-2	13+2/-2
40–170	3.4+0.5/-0.7	3.2+0.5/-0.6	3.2+0.5-0.6
170–1200	0.047+0.007/-0.009	0.051+0.008/-0.011	0.048+0.007/-0.008

Table 10.10: Relative differential cross-section as function of $p_{\mathrm{T}}^{\mathrm{t\bar{t}}}$ measured in the μ+jets , e+jets and the combined channels, from [118].

11. Summary

Overview of contributions to ATLAS Distributed Computing, Combined Physics and Performance, and Top Quark Physics is presented.

Contributions to the ATLAS Distributed Computing The contributions to the ATLAS Distributed Computing influenced how efficiently the whole ATLAS Collaboration can produce high-quality physics results. The wide spectrum of contributions range from *ADC operations*, through *ADC Monitoring* development and coordination during LHC Run I, *ADC automation* efforts, to development of the *PanDA Workload Management System*.

- The *ADC operations* team deals with monitoring and troubleshooting the ADC resources on the daily basis, with the aim to ensure that the physics groups can efficiently generate Monte Carlo simulation samples and process and analyze data; operating infrastructure for the ATLAS Central Services, ranging from services essential for the ATLAS detector operation, for determining data quality, through monitoring of the Tier-0 facilities, to systems supporting the mission of the ATLAS Distributed Computing.

- The *ADC Monitoring* tools play the key role in monitoring progress of Monte Carlo simulation and data processing campaigns carried out by the ATLAS physics groups and individual physicists.

- The *ADC automation* efforts are saving manpower on repetitive tasks, routing the expert manpower to more advanced topics of the ADC Operations, and to improve efficiency of utilization of available ATLAS Distributed Computing resources for the data processing and analysis by the ATLAS physics groups.

- The *PanDA Workload Management System* enables world-wide distributed analysis of data essential to the ATLAS physics program, and provides tools used by the ATLAS physicists on a daily basis in order to perform physics analyses.

Contributions to the Combined Physics and Performance – the Top Missing Transverse Energy Studies of the composition of the Missing Transverse Energy (MET) in the early 2010 through 2012 Top physics analyses were presented. A study of uncertainty for jet definition used in the MET calculation was presented. A tool used for validation of MET reconstruction algorithms was developed and presented.

Contributions to the Top Quark Physics A summary of contributions to the measurement of the top quark differential cross-section as a function of p_{T}, mass and rapidity of the $\mathrm{t\bar{t}}$ system was presented. Analysis package `TopNtupleAnalysis` was developed and presented. Measurement of shapes of differential distributions in the $\mathrm{t\bar{t}}$ system were presented.

The simulations are in a reasonable agreement with the observed shapes. The differential spectra shapes in the top-antitop pair system are consistent with the Standard Model and reasonably described by the event generators.

Further analyses based on a larger statistic sample of the full LHC Run I and LHC Run II samples are performed with the aim to improve statistical errors and study systematic effect in more details, however, they are not discussed in this Thesis.

A. ATLAS Production Task Monitoring – further description

A.1 Description of Task Filters

This section provides additional description of the Task Filters available in the ATLAS Production Task Monitoring, previously mentioned in Section 6.3.2.

A.1.1 Time Range Filter Group

- *From*, *To*: Select date for the custom Time range, overrides current Time range selection. If either of the *From* or *To* field value is missing, the time range is automatically set to 1 day.

- *Time range*: Select tasks active within the selected Time range. Available options are Last Day, Last 2 Days, Last 3 Days, Last Week, Last 2 Weeks, Last Month.

- *Not modified since*: Select active tasks which have not been modified since the chosen date.

- *Created*: Select only active tasks which have been created on the selected date.

A.1.2 Task Properties Filter Group

- *Pattern*: Pattern string be used in SQL query to search tasks by specified task name pattern. Wild-card characters (`*`) can be used along with logical AND (`&&`) and OR (`||`) operators. Example: `*m716* || *Np5PDF*`.

- *Tag*: Tag part of the Pattern string be used in SQL query to search tasks by specified task name pattern. Tag string is considered as a logical AND addition to the Pattern string. Wild card characters (`*`) can be used can be used along with logical AND (`&&`) and OR (`||`) operators. Example: `p1400 || p1401`.

- *Submitter's email address* as provided in the task definition phase.

- *ETask status*: ProdSys Task status (from the ETASK table). Examples of possible values: `LOST, ABORTED, SUBMITTING, RUNNING, HOLDING`, `TOBEFINISHED, FINISHED, DONE`.

- *Task type*. Examples of possible values: `bstream, evgen, evgentx, filter, merge, pile, prod_test, reco, reprocessing, simul, validation`.

- *Working group*: Working group as provided in the task definition phase. Examples of possible values: `ap_top, gp_top, gr_top, top`.
- *Activity*: ADC Activity (as in the Historical views) of the desired task selection. Examples of possible values: `CAF Processing`, `Data Processing`, `Group Analysis`, `Group Production`, `MC Production`, `MC Reconstruction`, `MC Reconstruction (XP)`, `MC Simulation`, `MC Simulation (XP)`, `Others`, `T0 Processing`, `Testing`, `User Analysis`, `Validation`.
- *Task ID*: Select 1 particular task. All other filters are abandoned.
- *Task Priority*: Filter tasks by priority. Examples of syntax:
  - `NNN:` ... select all tasks with priority $\geq$ `NNN`,
  - `:NNN` ... select all tasks with priority $\leq$ `NNN`,
  - `NNN` ... select all tasks with priority exactly `NNN`.
- *Attempt # larger than*: Select min attempt number of jobs in tasks.
- *Not yet defined*: Filter field to show only tasks which do not have jobs defined yet.

A.1.3 Task Duration Filter Group

- *AKTR Status*: Task status as in the AKTR DB, the `T_TASK_REQUEST` table. Examples of possible values:
  - `waiting`: task is waiting for input, input is not ready (e.g. another task has to run, dq2-put required etc.) in the place yet, task has not passed through Bamboo yet, input dataset definition does not exist yet.
  - `pending`: waiting to be processed by the scripts (AKTR programs have not yet finished registrations to DQ2), input is ready, no scouts defined.
  - `submitting`: sent the scout jobs, has not been assigned to cloud by Bamboo yet.
  - `submitted`: all jobs have been submitted, got cloud assigned by Bamboo.
  - `running`
  - `failed`
  - `finished`: task is done, but not all jobs were successful
  - `done`: task is done, all jobs were successful
  - `aborted`
  - `obsolete`: task output has been obsoleted

  - `lost`: task output has been lost

- `Timestamp`: Select timestamp as recorded in the AKTR DB. Examples of possible values:
  - `AKTRlastmodified`: task info updated at that time.
  - `AKTRpptimestamp`: time when post-production info was changed.
  - `AKTRstarttime`: timestamp starttime.
  - `AKTRtimestamp`: last record update time.
- `Task duration in terms of Timestamp, days`: Selected AKTR Timestamp has not been modified since `N` days.

A.1.4 Location Filter Group

- `Site`: ATLAS Site name where the task is running.
- `Destination cloud`: Cloud to which the task has been assigned.
- `Transformation`: Transformation used to process the task, e.g. `Reco_trf.py`.
- `Physics Stream`: Physics Stream, 3rd field in the output dataset name [129], e.g. `physics_Muons`, or `AlpgenJimmy_AUET2CTEQ6L1_ZeeNp2`. Wildcard characters (`*`) can be used can be used along with logical AND (`&&`) and OR (`||`) operators. Example: `physics_Muons || physics_JetT*uEtmiss`.

A.1.5 Task Outputs Filter Group

- `Output Project`: Output Project, e.g. `data12_8TeV`.
- `Output Data Type`: Output Data Type, e.g. `NTUP_TOP`.

A.2 Description of Home view

This section provides additional description of the Home view available in the ATLAS Production Task Monitoring, previously mentioned in Section 6.3.4.

A.2.1 Table header

The table header contains following elements:

- Switch to the Data tab, denoted with text *Data*.
- Switch to the Summary tab, denoted with text *Summary*.
- Selector of the page size, denoted with text *Show N entries*, where `N` is 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 200, All.

- Timestamp of latest *Task Update* and *Job Update* of the selection of tasks.
- Button *Expand task name*, to expand and shrink the shown task name. By default only part of the task name is shown.
- Full-text *Search* input box. It enables full-text search of a string in the table. The Search input box does not filter tasks, it only shows subset tasks, and does not influence aggregated values in the Data table footer.

Data table header consists of 2 rows. The first row shows the column groups: Task level information, Status, Jobs Info, and AKTR Timestamps. The second row shows the actual column names. Each column can be sorted in ascending or descending order by clicking on the corresponding arrow to the end of the column name. Sorting by multiple columns can be achieved with simultaneous pressing of the SHIFT key and clicking on corresponding arrows. Hint with the column meaning is shown on hover above the column names.

A.2.2 Tabular data

In the table a single task is represented by 1 row. The basic task information is shown by default, one can show supporting details by clicking on the + sign next to the task name. The table contains these columns:

- *Graphically*: Graphical representation of the task progress, and link to List of jobs of a particular task which can be opened in a new browser tab (link is marked with the letter X).
- *Task Name*: Click on the Task Name loads List of jobs of a particular task in the very same browser tab. To load list of jobs in a new browser tab, please click on letter X in the Graphically column. Task Name can be filtered with the Pattern and Tag fields. Can be expanded/shrinked with the "Expand task name" button.
- *Task ID*: Can be filtered with the Task ID field.
- *Activity*: Can be filtered with the Activity field.
- *Task Type*: Can be filtered with the Task Type field.
- *Working Group*: Can be filtered with the Working Group.
- *Dest Cloud*: Destination Cloud, Can be filtered with the Dest Cloud field.
- *Task Priority*: Can be filtered with the Task Priority field.
- *Transformation*: Can be filtered with the Task Priority field.
- *Duration*: Time between task creation and last job modification time stamp. The format is: dd hh:mi:ss (day hour:minute:second).
- *Max WT*: Maximum WallTime spent in a single job of the task.

- *Max RT*: Maximum RunTime spent in a single job of the task.
- *Status*: Can be filtered with the ETASK Status field.
- *JMaxAtt Reach*: Number of jobs of that task which reached maximal attempt number defined per task.
- *AttNr*: Highest attempt number of a single job of the task.
- *CPU Time*: CPUtime in seconds spent by successful jobs of the task.
- *HS06 CPUTime*: HS06 [130] CPUtime hours spent by successful jobs of the task.
- *Wall Time*: WallTime in seconds spent by successful jobs of the task.
- *HS06 WallTime*: HS06 WallTime hours spent by successful jobs of the task.
- *Proc Time*: Processing Time, average CPUtime in seconds spent by successful jobs of the task to process 1 input event.
- *NEvents*: Number of input events processed by successful jobs.
- *PFail*: Percentage of failed jobs.
- *PDone*: Percentage of finished jobs.
- *NExp*: Number of expected jobs. Can differ from number of total jobs.
- *NTotal*: Number of total jobs defined for the task.
- *Done*: Number of finished jobs.
- *Aborted*: Number of aborted jobs.
- *(AKTR) status*: Can be filtered with the AKTR Status field.
- *(AKTR) lastmodified*: Can be filtered with the AKTR Timestamp field.
- *(AKTR) pptimestamp*: Can be filtered with the AKTR Timestamp field.
- *(AKTR) starttime*: Can be filtered with the AKTR Timestamp field.
- *(AKTR) timestamp*: Can be filtered with the AKTR Timestamp field.

A.2.3 Table footer

The table footer contains following elements:

- Information about number of entries (tasks) shown, and the total number of tasks fulfilling the filter selection, e.g. *Showing 1 to 10 of 504 entries*.
- Pagination: buttons *First, Previous, Next, Last*, and information about which page is shown and how many pages are available, e.g. *Page 1 of 51*, the page number can be set/reset by hand.

The table footer consists of 1 row that contains Column names of the visible columns. Some of the column names are replaced by average or sum value of that column, those are the following columns:

- *Proc Time*: Average CPUtime in seconds spent by successful jobs of all the selected tasks (even those listed on previous/next pages) to process 1 input event.
- *NEvents*: Sum of number of input events processed by successful jobs of all the selected tasks.
- *PFail*: Percentage of failed jobs of all the selected tasks.
- *PDone*: Percentage of finished jobs of all the selected tasks.
- *NExp*: Number of expected jobs of all the selected tasks, can differ from number of total jobs.
- *NTotal*: Number of total jobs defined for all the selected tasks.
- *Done*: Number of finished jobs of all the selected tasks.
- *Aborted*: Number of aborted jobs of all the selected tasks.

A.2.4 Summary tab

The Summary tab of the Home view contains groups of plots. These plots show different quantities, summarized for the selected list of tasks. By default, when opening the Summary tab, only plots of the first group are visible. The other groups of plots are listed on the bottom of the Summary tab page. The plots open up and are made visible upon click on the group name. There are following groups of plots available in the Summary tab:

- *Summary Plots*: Evolution of number of jobs with different job states, Task status distribution by date of last change, Task status overview, Overview of task assignment to clouds, Distribution of number of created tasks per time bin.
- *Activities and groups*: Working Group, Activity.

- *Processing*: Task priority histogram.
- *Processing time distribution*: Spread of processing time by group, Spread of processing time by activity, Distribution of processing time by Working group, Distribution of processing time by Activity, Spread of processing time by Task Type, Distribution of processing time by Task Type.
- *Trigger group*: Reco Task Timining, Merge Job Timining.
- *Inputs*: Distributions of average number of input files per task and average input size per task, by Working group, Activity, Task Type, Input File Type; Distributions of average number of input files per average job and average input size per average job, by Working group, Activity, Task Type, Input File Type.
- *Processing walltime distribution*: Spread of processing walltime by Working Group or by Activity, Distribution of processing walltime by Working Group or by Activity, Spread of processing walltime by Task Type, Distribution of processing walltime by Task Type.
- *Successful/failed job distribution after M-th attempt*: Distribution of number of successful or failed jobs per attempt, by Site, by Tier, by Destination Cloud, by Activity, by Working Group, by Task Type, Distribution of number of failed jobs per attempt by Error summary.

A.3 Description of View of jobs per task

This section provides additional description of the View of jobs per task available in the ATLAS Production Task Monitoring, previously mentioned in Section 6.3.6.

A.3.1 Jobs Filters

In the similar way as the Task Filter, the Jobs Filter consists of filter fields. Resulting filter is applied as the logical *AND* operation on values of active filter groups. The available Jobs Filters fields are as follows:

- *Task ID*.
- *Job Name(pattern)*: Pattern string be used in SQL query to search tasks by specified task name pattern. Wild-card characters (`*`) can be used along with logical AND (`&&`) and OR (`||`) operators.
- *AttemptNr*: Filter jobs by Attempt number. Examples of syntax:
  - `NNN:` ... select all jobs with Attempt number $\geq$ `NNN`,
  - `:NNN` ... select all jobs with Attempt number $\leq$ `NNN`,
  - `NNN` ... select jobs with Attempt number exactly equal to `NNN`.

- *Status*: Job status as in the PanDA monitor.
- *Site*: ATLAS Site name where the job ran.
- *WaitTime*: Time difference between timestamp when the job started *running*, and when it was *defined*.
- *RunTime*: Running time of the job.

A.3.2 Jobs Data tab

The Jobs Data tab contains summary information about the task, and a table with a list of jobs.

The Summary information about the task consists of:

- Task ID.
- Task Owner.
- Task Name.
- Link to the Task information page in the PanDA monitor.
- Link to the Task request page in the PanDA monitor.

The table with list of jobs has similar filtering, sorting and pagination features as the Task Data table. Each row contains the following columns:

- *+*: get list of attempts of the job and further information upon click. By default only the information about the latest job attempt is shown. For each job the resubmission history is shown. For each job attempt the following information is available:
  - *PandaId*: Link to the particular job page in the PanDA monitor.
  - *Error Code or Error Details*.
  - *Job Status*.
  - *Queue*: specifies resources to which a particular job was assigned.
  - *Submitted*: timestamp of the submission.
  - *Started*: timestamp of the start of the job execution.
  - *Duration*: duration of job lifetime.
  - *Finished*: timestamp of the end of the job execution.
- *PandaId*: Link to the particular job page in the PanDA monitor.
- *Job Name*.
- *Job Status*: as in the PanDA monitor.
- *AttemptNr*: specifies how many retries this job has been through.

- *MaxAtt*: specifies maximal attempt number of all jobs of the task.
- *Site*: specifies a resource where the job is executed.
- *WaitTime*: time difference between job submission and the start of job execution.
- *RunTime*.
- *CPUtime*.
- *NEvents Processed*: Number of input events processed by the successful job.
- *ProcTime [s/evt]*: Average CPUtime spent to process 1 of the NEvents by the successful job.
- *Submitted*: Timestamp when the job was submitted.
- *Started*: Timestamp when the job started running.
- *Finished*: Timestamp when the job finished running.

A.3.3 Jobs Summary tab

The Jobs Summary tab contains 3 plots:

- Status overview plot of the latest attempts of jobs of the task.
- Cumulative plot of time evolution of number of processed input events in time. Only input events processed by successful jobs are shown.
- Distribution of number of jobs in different state per site.

Example plots describing progress of a single task in terms of job summary are shown in Fig. A.1.

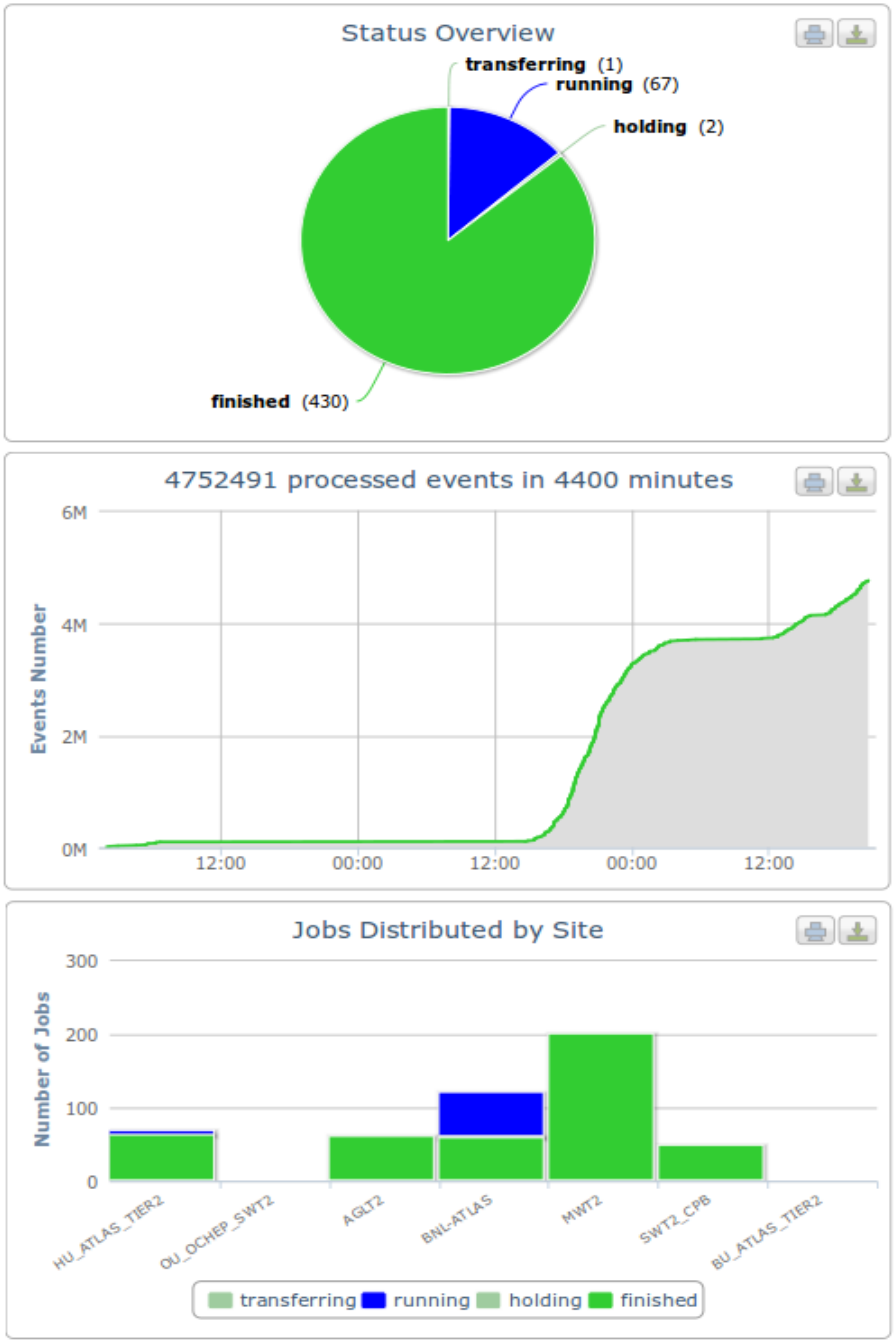


Figure A.1: Plots describing details of a task progress, with emphasis on jobs.

Selected List of Publications

This chapter provides a list of 23 selected papers co-authored by Jaroslava Schovancová. The full list of Jaroslava's publications is available in the INSPIRE-HEP database: `http://inspirehep.net/author/profile/J.Schovancova.1`. As of March 31st, 2016 it contains 537 publications with the ATLAS Collaboration, 45 publications with the Pierre Auger Collaboration, and 2 astrophysics publications.

[1] S Cabrera Urban, M J Costa, N Ghodbane, I M Nugent, S Adomeit, J F Arguin, C Bernius, D Boumediene, J Brau, D Cinca, C Doglioni, C Escobar, K Flavia Loureiro, D Gillberg, J Godfrey, Z Greenwood, P Haefner, D Hellmich, J Kroseberg, H Liao, A C Martyniuk, D Lopez Mateos, S Marti i Garcia, R Moles, M Moren Llacer, A Mc Carn, M Neubauer, H Okawa, D O'Neil, D Pallin, T Perez, M A Pleier, S Protopopescu, R Rajagopalan, J Schovancová, A Schwartzman, J Searcy, A Taffard, Y Takahashi, M Tomoto, M F Watson, and U K Yang. Jets, Missing Transverse Energy and Taus for Top Physics Analyses in Release 16 with the 2010 Dataset. Technical Report ATL-COM-PHYS-2011-132, CERN, Geneva, Feb 2011.

[2] W Bell, L Bellagamba, C Bernard, P Berta, K Bilsback, K Black, T Childers, R Di Sipio, L Dell'Asta, M Franchini, C Gabaldon, M Yamada, J Kanzaki, J Kvita, R Romano, A Sady, J Schovancová, F Spano', R Spighi, J Sjolin, M Talby, N Tannoury, and I Watson. Measurement of top quark pair differential cross-section with ATLAS in pp collisions at $\sqrt{s}$=7 TeV: Measurement of top quark pair differential cross-section with ATLAS using 5/fb of data collected in 2011 and the rel.17 version of the Atlas software . Technical Report ATL-COM-PHYS-2012-1137, CERN, Geneva, Jul 2012.

[3] J Schovancová. ATLAS Top results. Technical Report ATL-PHYS-SLIDE-2012-542, Oct 2012. Talk presented at the LHC days in 2012, Split, Croatia, 1-6 October 2012.

[4] K De, X Espinal, A Forti, E Korolkova, K Leffhalm, P A Love, J Schovancová, and Y Smirnov. ATLAS Distributed Computing Operations Shift Team Experience. *J. Phys.: Conf. Ser.* **331** (2011) 072045.

[5] J Schovancová, A Di Girolamo, J Elmsheuser, S Jézéquel, G Negri, N Ozturk, H Sakamoto, M Slater, Y Smirnov, I Ueda, D C Van Der Ster, and J Yu. ATLAS Distributed Computing Shift Operation in the first 2 full years of LHC data taking. *J. Phys.: Conf. Ser.* **396** (2012) 032122.

[6] A Sedov, A Di Girolamo, G Negri, H Sakamoto, J Schovancová, Y Smirnov, A Vartapetian, and J Yu. ATLAS Distributed Computing Operation Shift Teams experience during the discovery year and beginning of the Long Shutdown 1. *J. Phys.: Conf. Ser.* **513** (2014) 032085.

[7] L Sargsyan, J Andreeva, S Campana, E Karavakis, L Kokoszkiewicz, P Saiz, J Schovancová, and D Tuckett. ATLAS job monitoring in the Dashboard Framework. *J. Phys.: Conf. Ser.* **396** (2012) 032094.

[8] J Schovancová. ATLAS Distributed Computing Monitoring tools after full 2 years of LHC data taking. *J. Phys.: Conf. Ser.* **396** (2012) 032095.

[9] J Schovancová, S Campana, A Di Girolamo, S Jezequel, I Ueda, and T Wenaus. ATLAS Distributed Computing Monitoring tools during the LHC Run I. *Journal of Physics: Conference Series*, 513(3):032084, 2014.

[10] L Sargsyan, J Andreeva, M Jha, E Karavakis, L Kokoszkiewicz, P Saiz, J Schovancová, and D Tuckett. Dashboard Task Monitor for managing ATLAS user analysis on the Grid. *J. Phys.: Conf. Ser.* **513** (2014) 032083.

[11] E Karavakis, J Andreeva, S Campana, S Gayazov, S Jezequel, P Saiz, L Sargsyan, J Schovancová, and I Ueda. Common accounting system for monitoring the ATLAS Distributed Computing resources. *J. Phys.: Conf. Ser.* **513** (2014) 062024.

[12] J Schovancová. ATLAS Distributed Computing. Technical report, Mar 2011. Poster presented at the Poster Session of the 105th LHCC Meeting at CERN, Geneva, Switzerland, 23 March 2011.

[13] C Borrego, A Di Girolamo, X Espinal, L Rinaldi, J Schovancová, J Andreeva, M M Nowotka, and P Saiz. Aggregated monitoring and automatic site exclusion of the ATLAS computing activities: the ATLAS Site Status Board. Technical Report ATL-SOFT-PROC-2011-039, CERN, Geneva, Apr 2011. Paper submitted to the 5th Iberian Grid Infrastructure Conference, Santander, Spain, 8 - 10 Jun 2011.

[14] J Schovancová. ATLAS Distributed Computing on the way to the automatic site exclusion. Number ATL-SOFT-PROC-2012-001, Geneva, Jan 2012. Invited talk presented at the NEC'2011 XXIII International Symposium on Nuclear Electronics & Computing, Varna, Bulgaria, 12-19 September 2011.

[15] J Andreeva, C Borrego Iglesias, S Campana, A Di Girolamo, X Espinal Curull, I Dzhunov, S Gayazov, E Magradze, M M Nowotka, L Rinaldi, P Saiz, J Schovancová, G A Stewart, and M Wright. Automating ATLAS Computing Operations using the Site Status Board. *J. Phys.: Conf. Ser.* **396** (2012) 032072.

[16] J Schovancová, F H Barreiro Megino, C Borrego, S Campana, A Di Girolamo, J Elmsheuser, J Hejbal, T Kouba, F Legger, E Magradze, R Medrano Llamas, G Negri, L Rinaldi, G Sciacca, C Serfon, and D C Van Der Ster. ATLAS Distributed Computing Automation. Technical Report ATL-SOFT-PROC-2012-067, CERN, Geneva, Sep 2012. Talk presented at the 5th International Conference Distributed Computing and Grid-technologies in Science and Education, Dubna, Russian Federation, 16 - 20 Jul 2012.

[17] S A Tupputi, A Di Girolamo, T Kouba, and J Schovancová. Automating usability of ATLAS Distributed Computing resources. *J. Phys.: Conf. Ser.* **513** (2014) 032098.

[18] J Schovancová, K De, A Klimentov, T Maeno, P Nilsson, D Oleynik, S Panitkin, A Petrosyan, A Vaniachine, T Wenaus, and D Yu. PanDA Beyond ATLAS: Workload Management for Data Intensive Science. Technical Report ATL-SOFT-SLIDE-2013-860, Oct 2013. Talk presented at HEPiX Fall 2013 Workshop, 27th Oct - 1st Nov 2013, Ann Arbor, MI, USA.

[19] J Schovancová, K De, A Klimentov, P Love, M Potekhin, and T Wenaus. The next generation of the ATLAS PanDA Monitoring System. *PoS*(ISGC2014)035.

[20] T Maeno, K De, A Klimentov, P Nilsson, D Oleynik, S Panitkin, A Petrosyan, J Schovancová, A Vaniachine, T Wenaus, and D Yu. Evolution of the ATLAS PanDA Workload Management System for Exascale Computational Science. *J. Phys.: Conf. Ser.* **513** (2014) 032062.

[21] M Borodin, K De, S Jha, D Golubkov, A Klimentov, T Maeno, P Nilsson, D Oleynik, S Panitkin, A Petrosyan, J Schovancová, A Vaniachine, and T Wenaus. PanDA Beyond ATLAS : A Scalable Workload Management System For Data Intensive Science. Technical Report ATL-SOFT-SLIDE-2014-117, Mar 2014.

[22] K De, A Klimentov, D Oleynik, S Panitkin, A Petrosyan, A Vaniachine, T Wenaus, and J Schovancová. Integration of PanDA workload management system with Titan supercomputer at OLCF. Technical Report ATL-SOFT-PROC-2015-040, CERN, Geneva, May 2015. Proceedings paper of the talk presented at CHEP2015 - International Conference on Computing in High Energy and Nuclear Physics (CHEP), Okinawa, Japan, 13-17 Apr 2015.

[23] K De, A Klimentov, T Maeno, P Nilsson, D Oleynik, S Panitkin, A Petrosyan, J Schovancová, A Vaniachine, and T Wenaus. The Future of PanDA in ATLAS Distributed Computing. Technical Report ATL-SOFT-PROC-2015-047, CERN, Geneva, May 2015. Proceedings paper of the talk presented at CHEP2015 - International Conference on Computing in High Energy and Nuclear Physics (CHEP), Okinawa, Japan, 13-17 Apr 2015.

References

[1] Latest update in the search for the Higgs boson. Jul 2012. https://cds.cern.ch/record/1459565.

[2] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys.Lett.*, B716:1–29, 2012.

[3] J Schovancová. ATLAS Top results. (ATL-PHYS-SLIDE-2012-542), Oct 2012. Talk presented at the LHC days in 2012, Split, Croatia, 1-6 October 2012.

[4] O S Brüning, P Collier, P Lebrun, S Myers, R Ostojic, J Poole, and P Proudlock. *LHC Design Report*. CERN, Geneva, 2004.

[5] L R Evans and P Bryant. LHC Machine. *J. Instrum.*, 3:S08001. 164 p, 2008. This report is an abridged version of the LHC Design Report (CERN-2004-003).

[6] S Myers and E Picasso. The design, construction and commissioning of the CERN large electron–positron collider. *Contemporary Physics*, 31(6):387–403, 1990.

[7] J Haffner. The CERN accelerator complex. Complexe des accélérateurs du CERN. Oct 2013. General Photo, OPEN-PHO-ACCEL-2013-056, http://cds.cern.ch/record/1621894.

[8] T Taylor. Incident at the CERN LHC: The event, the repair and lessons to drawn. *TEION KOGAKU*, 45:344–351, 2010.

[9] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[10] J Pequenao. Computer generated image of the whole ATLAS detector. General Photo, CERN-GE-0803012, http://cds.cern.ch/record/1095924, Mar 2008.

[11] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30 – 61, 2012.

[12] The Nobel Foundation. The Nobel Prize in Physics 2013. http://www.nobelprize.org/nobel_prizes/physics/laureates/2013/, 2013.

[13] ATLAS Collaboration. *ATLAS technical coordination: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1999.

[14] ATLAS Collaboration. *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1997.

[15] J Pequenao. Computer generated image of the ATLAS inner detector. General Photo, CERN-GE-0803014, http://cds.cern.ch/record/1095926, Mar 2008.

[16] J Pequenao. Computer Generated image of the ATLAS calorimeter. General Photo, CERN-GE-0803015, http://cds.cern.ch/record/1095927, Mar 2008.

[17] J Pequenao. Computer generated image of the ATLAS Muons subsystem. General Photo, CERN-GE-0803017, http://cds.cern.ch/record/1095929, Mar 2008.

[18] K A Olive and Particle Data Group. Review of Particle Physics. *Chin.Phys.*, C38:090001, 2014.

[19] Tevatron Electroweak Working Group. Combination of CDF and DØ Measurements of the Single Top Production Cross Section. 2009.

[20] A Bilal. Introduction to supersymmetry. 2001. arXiv:hep-th/0101055.

[21] E Kiritsis. Introduction to superstring theory. *Leuven, Belgium: Leuven Univ. Pr. (1998) 315 p*, 1997. arXiv:hep-th/9709062.

[22] L Randall and R Sundrum. A Large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.*, 83:3370–3373, 1999.

[23] CDF Collaboration. Observation of top quark production in $\overline{p}p$ collisions with the collider detector at fermilab. *Phys. Rev. Lett.*, 74:2626–2631, Apr 1995.

[24] DØ Collaboration. Search for high mass top quark production in $p\overline{p}$ collisions at $\sqrt{s} =$ 1.8 TeV. *Phys. Rev. Lett.*, 74:2422–2426, Mar 1995.

[25] S Catani. Aspects of QCD, from the Tevatron to the LHC. 2000. arXiv:hep-ph/0005233.

[26] A Quadt. Top quark physics at hadron colliders. *The European Physical Journal C - Particles and Fields*, 48(3):835–1000, 2006.

[27] M Jeżabek and J H Kühn. QCD corrections to semileptonic decays of heavy quarks. *Nuclear Physics B*, 314(1):1 – 6, 1989.

[28] ATLAS Collaboration, CDF Collaboration, CMS Collaboration, and DØ Collaboration. First combination of Tevatron and LHC measurements of the top-quark mass. Technical Report arXiv:1403.4427. ATLAS-CONF-2014-008. CDF-NOTE-11071. CMS-PAS-TOP-13-014. D0-NOTE-6416, CERN, Geneva, Mar 2014.

[29] G Compostella, D Lucchesi, and S Amerio. Measurement of the $t\bar{t}$ production cross-section in the MET+jets channel (2.2/fb). Technical report, 2009.

[30] CDF Collaboration. Top-Quark Mass Measurement Using Events with Missing Transverse Energy and Jets at CDF. *Phys. Rev. Lett.*, 107:232002, Nov 2011.

[31] ATLAS Collaboration. Measurement of the top quark mass in the $t\bar{t} \to$ lepton + jets and $t\bar{t} \to$ dilepton channels using $\sqrt{s} = 7$ TeV ATLAS data. *Eur. Phys. J. C*, 75(arXiv:1503.05427. CERN-PH-EP-2015-050):330. 35 p, Mar 2015.

[32] ATLAS Collaboration. Measurement of the top-quark mass in the fully hadronic decay channel from ATLAS data at $\sqrt{s} = 7\,\text{TeV}$. *Eur.Phys.J.*, C75(4):158, 2015.

[33] ATLAS Collaboration. Measurement of the mass difference between top and anti-top quarks in pp collisions at $\sqrt(s) = 7$ TeV using the ATLAS detector. *Phys.Lett.*, B728:363–379, 2014.

[34] ATLAS Collaboration. Measurement of the charge asymmetry in dileptonic decays of top quark pairs in *pp* collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector. 2015.

[35] ATLAS Collaboration. Measurement of Top Quark Polarization in Top-Antitop Events from Proton-Proton Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector. *Phys.Rev.Lett.*, 111(23):232002, 2013.

[36] W Bernreuther and Z-G Si. Top quark spin correlations and polarization at the LHC: Standard model predictions and effects of anomalous top chromo moments. *Physics Letters B*, 725(1–3):115 – 122, 2013.

[37] ATLAS Collaboration. Measurement of the top pair production cross-section in 8 TeV proton-proton collisions using kinematic information in the lepton+jets final state with ATLAS. 2015.

[38] M Cacciari, M Czakon, M Mangano, A Mitov, and P Nason. Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation. *Phys. Lett.*, B710:612–622, 2012.

[39] M Czakon and A Mitov. Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders. *Comput. Phys. Commun.*, 185:2930, 2014.

[40] N Kidonakis. Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production. *Phys.Rev.*, D83:091503, 2011.

[41] ATLAS Collaboration. Comprehensive measurements of t-channel single top-quark production cross sections at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys.Rev.*, D90(11):112006, 2014.

[42] N Kidonakis. NNLL resummation for s-channel single top quark production. *Phys.Rev.*, D81:054028, 2010.

[43] ATLAS Collaboration. Search for s-channel single top-quark production in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Phys.Lett.*, B740:118–136, 2015.

[44] N Kidonakis. Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-. *Phys.Rev.*, D82:054018, 2010.

[45] ATLAS Collaboration. Evidence for the associated production of a W boson and a top quark in ATLAS at $\sqrt{s} = 7$ TeV. *Phys.Lett.*, B716:142–159, 2012.

[46] ATLAS Collaboration. Search for FCNC single top-quark production at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys.Lett.*, B712:351–369, 2012.

[47] ATLAS Collaboration. A search for $t\bar{t}$ resonances in lepton+jets events with highly boosted top quarks collected in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP*, 1209:041, 2012.

[48] ATLAS Collaboration. Measurement of Spin Correlation in Top-Antitop Quark Events and Search for Top Squark Pair Production in pp Collisions at $\sqrt{s} = 8$ TeV Using the ATLAS Detector. *Phys.Rev.Lett.*, 114(14):142001, 2015.

[49] ATLAS Collaboration. Search for $t\bar{t}$ resonances in the lepton plus jets final state with ATLAS using 4.7 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV. *Phys.Rev.*, D88(1):012004, 2013.

[50] ATLAS Collaboration. Search for resonances decaying into top-quark pairs using fully hadronic decays in pp collisions with ATLAS at $\sqrt{s} = 7$ TeV. *JHEP*, 1301:116, 2013.

[51] ATLAS Collaboration. Search for tb resonances in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys.Rev.Lett.*, 109:081801, 2012.

[52] M E Pozo Astigarraga. Evolution of the ATLAS Trigger and Data Acquisition System. Technical Report ATL-DAQ-PROC-2014-025, CERN, Geneva, Sep 2014.

[53] M Elsing, M Goossens, A Nairz, and G Negri. The ATLAS Tier-0: Overview and operational experience. *J. Phys.: Conf. Ser.*, 219:072011, 2010.

[54] WLCG tier centres. http://wlcg-public.web.cern.ch/tier-centres.

[55] M Branco, D Cameron, B Gaidioz, V Garonne, B Koblitz, M Lassnig, R Rocha, P Salgado, and T Wenaus. Managing ATLAS data on a petabyte-scale with DQ2. *J. Phys.: Conf. Ser.*, 119:062017, 2008.

[56] C Serfon, M Barisits, T Beermann, V Garonne, L Goossens, M Lassnig, A Nairz, and R Vigne. Rucio, the next-generation Data Management system in ATLAS. Technical Report ATL-SOFT-PROC-2014-009, CERN, Geneva, Oct 2014.

[57] T Maeno. Panda: distributed production and distributed analysis system for atlas. *Journal of Physics: Conference Series*, 119(6):062036, 2008.

[58] K De, A Klimentov, T Maeno, P Nilsson, D Oleynik, S Panitkin, A Petrosyan, J Schovancová, A Vaniachine, and T Wenaus. The Future of PanDA in ATLAS Distributed Computing. Technical Report ATL-SOFT-PROC-2015-047, CERN, Geneva, May 2015. Proceedings paper of the talk presented at CHEP2015 - International Conference on Computing in High Energy and Nuclear Physics (CHEP), Okinawa, Japan, 13-17 Apr 2015.

[59] M Borodin, K De, J E García Navarro, D Golubkov, A Klimentov, T Maeno, and A Vaniachine. Scaling up ATLAS production system for the LHC Run 2 and beyond: project ProdSys2. Technical Report ATL-SOFT-PROC-2015-054, CERN, Geneva, May 2015.

[60] P Calafiura, W Lavrijsen, C Leggett, M Marino, and D Quarrie. The athena control framework in production, new developments and lessons learned. In *Computing in high energy physics and nuclear physics. Proceedings, Conference, CHEP'04, Interlaken, Switzerland, September 27-October 1, 2004*, pages 456–458, 2005.

[61] I Antcheva, M Ballintijn, B Bellenot, M Biskup, R Brun, N Buncic, Ph Canal, D Casadei, O Couet, V Fine, L Franco, G Ganis, A Gheata, D Gonzalez Maline, M Goto, J Iwaszkiewicz, A Kreshuk, D Marcos Segura, R Maunder, L Moneta, A Naumann, E Offermann, V Onuchin, S Panacek, F Rademakers, P Russo, and M Tadel. ROOT — A C++ framework for petabyte data storage, statistical analysis and visualization. *Comput. Phys. Commun.*, 182(6):1384–1385, 2011.

[62] S Jézéquel, G A Stewart, and the ATLAS Collaboration. Atlas distributed computing operations: Experience and improvements after 2 full years of data-taking. *Journal of Physics: Conference Series*, 396(3):032058, 2012.

[63] I Ueda and the ATLAS Collaboration. ATLAS operations: Experience and evolution in the data taking era. *Journal of Physics: Conference Series*, 331(7):072034, 2011.

[64] ATLAS Collaboration. *ATLAS Computing: technical design report*. Technical Design Report ATLAS. CERN, Geneva, 2005.

[65] I Bird, P Buncic, F Carminati, M Cattaneo, P Clarke, I Fisk, M Girone, J Harvey, B Kersevan, P Mato, R Mount, and B Panzer-Steindel. Update of the Computing Models of the WLCG and the LHC Experiments. Technical Report CERN-LHCC-2014-014. LCG-TDR-002, CERN, Geneva, Apr 2014.

[66] K De, X Espinal, A Forti, E Korolkova, K Leffhalm, P Love, J Schovancová, Y Smirnov, and the ATLAS Collaboration. Atlas distributed computing operations shift team experience. *Journal of Physics: Conference Series*, 331(7):072045, 2011.

[67] J Schovancová. ATLAS Distributed Computing. Technical report, Mar 2011. Poster presented at the Poster Session of the 105th LHCC Meeting at CERN, Geneva, Switzerland, 23 March 2011.

[68] J Schovancová, A Di Girolamo, J Elmsheuser, S Jézéquel, G Negri, N Ozturk, H Sakamoto, M Slater, Y Smirnov, I Ueda, D C Van Der Ster, and J Yu. ATLAS Distributed Computing Shift Operation in the first 2 full years of LHC data taking. *Journal of Physics: Conference Series*, 396(3):032122, 2012.

[69] A Sedov, A Di Girolamo, G Negri, H Sakamoto, J Schovancová, Y Smirnov, A Vartapetian, and J Yu. ATLAS Distributed Computing Operation Shift Teams experience during the discovery year and beginning of the Long Shutdown 1. *Journal of Physics: Conference Series*, 513(3):032085, 2014.

[70] J Schovancová. ATLAS Distributed Computing Monitoring tools after full 2 years of LHC data taking. *Journal of Physics: Conference Series*, 396(3):032095, 2012.

[71] T Antoni, W Bühler, H Dres, G Grein, and M Roth. Global grid user support—building a worldwide distributed user support infrastructure. *Journal of Physics: Conference Series*, 119(5):052002, 2008.

[72] E Karavakis, J Andreeva, S Campana, S Gayazov, S Jezequel, P Saiz, L Sargsyan, J Schovancová, and I Ueda. Common accounting system for monitoring the ATLAS Distributed Computing resources. *Journal of Physics: Conference Series*, 513(6):062024, 2014.

[73] J Andreeva, S Belforte, M Boehm, A Casajus, J Flix, B Gaidioz, C Grigoras, L Kokoszkiewicz, E Lanciotti, R Rocha, P Saiz, R Santinelli, I Sidorova, A Sciabà, and A Tsaregorodtsev. Dashboard applications to monitor experiment activities at sites. *Journal of Physics: Conference Series*, 219(6):062003, 2010.

[74] T Maeno, K De, T Wenaus, P Nilsson, R Walker, A Stradling, V Fine, M Potekhin, S Panitkin, and G Compostella. Evolution of the atlas panda production and distributed analysis system. *Journal of Physics: Conference Series*, 396(3):032071, 2012.

[75] D Tuckett, J Andreeva, I Dzhunov, E Karavakis, L Kokoszkiewicz, M Nowotka, and P Saiz. Designing and developing portable large-scale javascript web applications within the experiment dashboard framework. *Journal of Physics: Conference Series*, 396(5):052069, 2012.

[76] L Kokoszkiewicz, J Andreeva, I Dzhunov, E Karavakis, M Lamanna, J Moscicki, and L Sargsyan. hbrowse – generic framework for hierarchical data visualization. *Proceedings of Science (EGICF12-EMITC2)*, 062, 2012.

[77] S Campana, F Barreiro Megino, S Jezequel, G Negri, C Serfon, and I Ueda. Evolving atlas computing for today's networks. *Journal of Physics: Conference Series*, 396(3):032019, 2012.

[78] P Laurens, S McKee, A Lake, H Severini, T Wlodek, S Wolff, and J Zurawski. Monitoring the us atlas network infrastructure with perfsonar-ps. *Journal of Physics: Conference Series*, 396(4):042038, 2012.

[79] L Sargsyan, J Andreeva, S Campana, E Karavakis, L Kokoszkiewicz, P Saiz, J Schovancová, and D Tuckett. ATLAS job monitoring in the Dashboard Framework. *Journal of Physics: Conference Series*, 396(3):032094, 2012.

[80] E Karavakis, J Andreeva, M Cinquilli, I Dzhunov, M Kenyon, L Kokoszkiewicz, and P Saiz. User-centric monitoring of the analysis and production activities within the atlas and cms virtual organisations using the experiment dashboard system. *Proceedings of Science (EGICF12-EMITC2)*, 110, 2012.

[81] A Anisenkov, S Belov, A Di Girolamo, S Gayazov, K Klimentov, D Oleynik, and A Senchenko. Agis: The atlas grid information system. *Journal of Physics: Conference Series*, 396(3):032006, 2012.

[82] S Lopienski. Service level status – a new real-time status display for it services. *Journal of Physics: Conference Series*, 119(5):052025, 2008.

[83] A L Dewhurst, D Barberis, F Bujor, J de Stefano, D Dykstra, D Front, E Gallas, C F Gamboa, F Luehring, and R Walker. Evolution of grid-wide access to database resident information in atlas using frontier. *Journal of Physics: Conference Series*, 396(5):052025, 2012.

[84] P Saiz, J Andreeva, I Dzhunov, E Karavakis, L Kokoszkiewicz, M Nowotka, D Tuckett, A Di Girolamo, A Sciaba, J Schovancová, J Flix, and P Kreuzer. Collaborative development. case study of the development of flexible monitoring applications. *Journal of Physics: Conference Series*, 396, 2012.

[85] C Borrego, A Di Girolamo, X Espinal, L Rinaldi, J Schovancová, J Andreeva, M M Nowotka, and P Saiz. Aggregated monitoring and automatic site exclusion of the ATLAS computing activities: the ATLAS Site Status Board. Technical Report ATL-SOFT-PROC-2011-039, CERN, Geneva, Apr 2011. Paper submitted to the 5th Iberian Grid Infrastructure Conference, Santander, Spain, 8 - 10 Jun 2011.

[86] J Andreeva, C Borrego Iglesias, S Campana, A Di Girolamo, X Espinal Curull, I Dzhunov, S Gayazov, E Magradze, M M Nowotka, L Rinaldi, P Saiz,

J Schovancová, G A Stewart, and M Wright. Automating ATLAS Computing Operations using the Site Status Board. *Journal of Physics: Conference Series*, 396(3):032072, 2012.

[87] D C van der Ster, J Elmsheuser, M Úbeda García, and M Paladin. Hammercloud: A stress testing system for distributed analysis. *Journal of Physics: Conference Series*, 331(7):072036, 2011.

[88] P Andrade, M Babik, K Bhatt, P Chand, D Collados, V Duggal, P Fuente, E Imamagic, P Joshi, R Kalmady, U Karnani, V Kumar, J Tarragon, W Lapka, and C Triantafyllidis. Service availability monitoring framework based on commodity software. *Journal of Physics: Conference Series*, 396(3):032008, 2012.

[89] J Andreeva, P Dhara, A Di Girolamo, A Kakkar, M Litmaath, N Magini, G Negri, S Ramachandran, S Roiser, P Saiz, M D Saiz Santos, B Sarkar, J Schovancová, A Sciabà, and A Wakankar. New solutions for large scale functional tests in the wlcg infrastructure with sam/nagios: the experiments experience. *Journal of Physics: Conference Series*, 396(3):032100, 2012.

[90] G Mathieu, A Richards, J Gordon, C Del Cano Novales, P Colclough, and M Viljoen. Gocdb, a topology repository for a worldwide grid infrastructure. *Journal of Physics: Conference Series*, 219(6):062021, 2010.

[91] R Pordes, M Altunay, P Avery, A Bejan, K Blackburn, A Blatecky, R Gardner, B Kramer, M Livny, J McGee, M Potekhin, R Quick, D Olson, A Roy, C Sehgal, T Wenaus, M Wilde, and F Würthwein. New science on the open science grid. *Journal of Physics: Conference Series*, 125(1):012070, 2008.

[92] J Andreeva, D Benjamin, A Klimentov, V Korenkov, D Oleynik, S Panitkin, and A Petrosyan. Tier-3 Monitoring Software Suite (T3MON) proposal. Technical Report ATL-SOFT-PUB-2011-001, CERN, Geneva, Mar 2011.

[93] J Schovancová, S Campana, A Di Girolamo, S Jezequel, I Ueda, and T Wenaus. ATLAS Distributed Computing Monitoring tools during the LHC Run I. *Journal of Physics: Conference Series*, 513(3):032084, 2014.

[94] L Sargsyan, J Andreeva, M Jha, E Karavakis, L Kokoszkiewicz, P Saiz, J Schovancová, and D Tuckett. Dashboard Task Monitor for managing ATLAS user analysis on the Grid. *Journal of Physics: Conference Series*, 513(3):032083, 2014.

[95] S A Tupputi, A Di Girolamo, T Kouba, and J Schovancová. Automating usability of ATLAS Distributed Computing resources. *Journal of Physics: Conference Series*, 513(3):032098, 2014.

[96] J Schovancová. ATLAS Distributed Computing on the way to the automatic site exclusion. (ATL-SOFT-PROC-2012-001), Jan 2012. Invited talk presented at the NEC'2011 XXIII International Symposium on Nuclear Electronics & Computing, Varna, Bulgaria, 12-19 September 2011.

[97] J Schovancová, F H Barreiro Megino, C Borrego, S Campana, A Di Girolamo, J Elmsheuser, J Hejbal, T Kouba, F Legger, E Magradze, R Medrano Llamas, G Negri, L Rinaldi, G Sciacca, C Serfon, and D C Van Der Ster. ATLAS Distributed Computing Automation. (ATL-SOFT-PROC-2012-067), Sep 2012. Talk presented at the 5th International Conference Distributed Computing and Grid-technologies in Science and Education, Dubna, Russian Federation, 16 - 20 Jul 2012.

[98] J Schovancová, F H Barreiro Megino, C Borrego, S Campana, A Di Girolamo, J Elmsheuser, J Hejbal, T Kouba, F Legger, E Magradze, R Medrano Llamas, G Negri, L Rinaldi, G Sciacca, C Serfon, and D C Van Der Ster. ATLAS Distributed Computing Automation. (ATL-SOFT-PROC-2012-067), Sep 2012. Talk presented at the 5th International Conference Distributed Computing and Grid-technologies in Science and Education, Dubna, Russian Federation, 16 - 20 Jul 2012.

[99] V Garonne, R Vigne, G A Stewart, M Barisits, T Beermann, M Lassnig, C Serfon, L Goossens, A Nairz, and the ATLAS Collaboration. Rucio – the next generation of large scale distributed system for atlas data management. *Journal of Physics: Conference Series*, 513(4):042021, 2014.

[100] T Maeno, K De, A Klimentov, P Nilsson, D Oleynik, S Panitkin, A Petrosyan, J Schovancová, A Vaniachine, T Wenaus, D Yu, and the Atlas Collaboration. Evolution of the atlas panda workload management system for exascale computational science. *Journal of Physics: Conference Series*, 513(3):032062, 2014.

[101] D Britton and S L Lloyd. How to deal with petabytes of data: the LHC Grid project. *Reports on Progress in Physics*, 77(6):065902, 2014.

[102] K Bos and I Fisk. A Use Case for the LHC Open Network Environment. Technical report, 2011. http://lhcone.web.cern.ch/node/23.

[103] I Fisk. New computing models and LHCONE. Computing in High Energy and Nuclear Physics (CHEP) 2012. May 2012. http://cds.cern.ch/record/1460610.

[104] S Campana and the ATLAS Collaboration. Evolution of the atlas distributed computing system during the lhc long shutdown. *Journal of Physics: Conference Series*, 513(3):032016, 2014.

[105] D Thain, T Tannenbaum, and M Livny. Distributed computing in practice: The condor experience: Research articles. *Concurr. Comput. : Pract. Exper.*, 17(2-4):323–356, February 2005.

[106] J Schovancová, K De, A Klimentov, T Maeno, P Nilsson, D Oleynik, S Panitkin, A Petrosyan, A Vaniachine, T Wenaus, and D Yu. PanDA Beyond ATLAS: Workload Management for Data Intensive Science. Technical Report ATL-SOFT-SLIDE-2013-860, Oct 2013. Talk presented at HEPiX Fall 2013 Workshop, 27th Oct - 1st Nov 2013, Ann Arbor, MI, USA.

[107] J Schovancová, K De, A Klimentov, P Love, M Potekhin, and T Wenaus. The next generation of the ATLAS PanDA Monitoring System. Technical Report 035. Talk presented at the International Symposium on Grids and Clouds (ISGC) 2014, 23-28th March 2014, Taipei, Taiwan.

[108] T Maeno, K De, A Klimentov, P Nilsson, D Oleynik, S Panitkin, A Petrosyan, J Schovancová, A Vaniachine, T Wenaus, and D Yu. Evolution of the ATLAS PanDA Workload Management System for Exascale Computational Science. *Journal of Physics: Conference Series*, 513(3):032062, 2014.

[109] M Borodin, K De, S Jha, D Golubkov, A Klimentov, T Maeno, P Nilsson, D Oleynik, S Panitkin, A Petrosyan, J Schovancová, A Vaniachine, and T Wenaus. PanDA Beyond ATLAS : A Scalable Workload Management System For Data Intensive Science. Technical Report ATL-SOFT-SLIDE-2014-117, Mar 2014.

[110] K De, A Klimentov, D Oleynik, S Panitkin, A Petrosyan, A Vaniachine, T Wenaus, and J Schovancová. Integration of PanDA workload management system with Titan supercomputer at OLCF. Technical Report ATL-SOFT-PROC-2015-040, CERN, Geneva, May 2015. Proceedings paper of the talk presented at CHEP2015 - International Conference on Computing in High Energy and Nuclear Physics (CHEP), Okinawa, Japan, 13-17 Apr 2015.

[111] Apigee.com. API best practices. Technical report, Apigee.com, 2014.

[112] A Canepa, D Cinca, S Farrell, H Garitaonandia Elejabarrieta, U Grundler, K Johns, V Kaushik, H Khandanya, X Lei, JR Lessard, T Liss, I Nugent, H Okawa, D Pallin, A Taffard, B Toggerson, and M Watson. Missing Transverse Energy for Top Physics analyses with early ATLAS data at $\sqrt{s}$=7 TeV. Technical Report ATL-COM-PHYS-2010-821, CERN, Geneva, Oct 2010.

[113] S Cabrera Urban, M J Costa, N Ghodbane, I M Nugent, S Adomeit, J F Arguin, C Bernius, D Boumediene, J Brau, D Cinca, C Doglioni, C Escobar, K Flavia Loureiro, D Gillberg, J Godfrey, Z Greenwood, P Haefner, D Hellmich, J Kroseberg, H Liao, A C Martyniuk, D Lopez Mateos, S Marti i Garcia, R Moles, M Moren Llacer, A Mc Carn, M Neubauer, H Okawa, D O'Neil, D Pallin, T Perez, M A Pleier, S Protopopescu, R Rajagopalan, J Schovancová, A Schwartzman, J Searcy, A Taffard, Y Takahashi, M Tomoto, M F Watson, and U K Yang. Jets, Missing Transverse Energy and Taus for Top Physics Analyses in Release 16 with the 2010 Dataset. Technical Report ATL-COM-PHYS-2011-132, CERN, Geneva, Feb 2011.

[114] ATLAS Collaboration. *Expected performance of the ATLAS experiment: detector, trigger and physics.* CERN, Geneva, 2009.

[115] ATLAS Collaboration. Performance of the Missing Transverse Energy Reconstruction and Calibration in Proton-Proton Collisions at a Center-of-Mass Energy of 7 TeV with the ATLAS Detector. Technical Report ATLAS-CONF-2010-057, CERN, Geneva, Jul 2010.

[116] ATLAS Collaboration. Electron and photon reconstruction and identification in ATLAS: expected performance at high energy and results at 900 GeV. Technical Report ATLAS-CONF-2010-005, CERN, Geneva, 2010.

[117] S Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[118] W Bell, L Bellagamba, C Bernard, P Berta, K Bilsback, K Black, T Childers, R Di Sipio, L Dell'Asta, M Franchini, C Gabaldon, M Yamada, J Kanzaki, J Kvita, R Romano, A Sady, J Schovancová, F Spano', R Spighi, J Sjolin, M Talby, N Tannoury, and I Watson. Measurement of top quark pair differential cross section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV: Measurement of top quark pair differential cross section with ATLAS using 5/fb of data collected in 2011 and the rel.17 version of the ATLAS software. Technical Report ATL-COM-PHYS-2012-1137, CERN, Geneva, Jul 2012.

[119] CDF Collaboration. Combination of CDF single top quark searches with 2.2 fb^{-1} of data. Technical Report CDF Note 9251, Mar 2008.

[120] Gennaro Corcella, I G Knowles, G Marchesini, S Moretti, K Odagiri, Peter Richardson, Michael H Seymour, and Bryan R Webber. HERWIG 6: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes). *J. High Energy Phys.*, 01(hep-ph/0011363. CAVENDISH-HEP-99-03. CERN-TH-2000-284. RAL-TR-2000-048):010. 93 p, Nov 2000.

[121] J Butterworth, J R Forshaw, and Michael H Seymour. Multiparton interactions in photoproduction at HERA. *Z. Phys. C*, 72(hep-ph/9601371. CERN-TH-95-82. MC-TH-96-05. UCL-HEP-96-02):637–646. 22 p, Jan 1996.

[122] ATLAS Collaboration. New ATLAS event generator tunes to 2010 data. Technical Report ATL-PHYS-PUB-2011-008, CERN, Geneva, Apr 2011.

[123] B Acharya, S Adomeit, M Aoki, B Alvarez, F Balli, W H Bell, S Bilski, K Becker, D Benjamin, C Bernard, K Black, S Calvet, R Camacho, Y Coadou, G Cortiana, G Compostella, C Doglioni, F Derue, K Finelli, N Ghodbane, K J Grahn, S Head, D Hirschbuehl, V Kaushik, T Kendziorra, O M Kind, A Krasznahorkay, T Kuhl, H C Lee, A Lister, K Loureiro, L Mijovic, J M Morris, R Moles Valls, O Nackenhorst, M zur Nedden, M Owen, M Pinamonti, K Rosbach, M Rudolph, G Salamanna, U de Sanctis, J Schwindling, J Searcy, L Serkin, E Shabalina, K Shaw, J Sjolin, R Soualah, D B Ta, T Theveneaux-Pelzer, M Vreeswijk, I J Watson, and H Zhu. Object selection and calibration, background estimations and MC

samples for the Summer 2012 Top Quark analyses with 2011 data. Technical Report ATL-COM-PHYS-2012-499, CERN, Geneva, May 2012.

[124] M Cacciari and G P Salam. Dispelling the myth for the jet-finder. *Physics Letters B*, 641(1):57 – 61, 2006.

[125] M Cacciari, G P Salam, and G Soyez. The anti-k_t jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063, 2008.

[126] Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy $\sqrt{s} = 7$ TeV. Technical Report ATLAS-CONF-2010-054, CERN, Geneva, Jul 2010.

[127] ATLAS Collaboration. Measurement of the $t\bar{t}$ production cross-section in pp collisions at $\sqrt{s} = 7$ TeV using kinematic information of lepton+jets events. Technical Report ATLAS-CONF-2011-121, CERN, Geneva, Aug 2011.

[128] V Ahrens, A Ferroglia, M Neubert, B D Pecjak, and L L Yang. Renormalization-Group Improved Predictions for Top-Quark Pair Production at Hadron Colliders. *JHEP*, 09:097, 2010.

[129] S Albrand, D Barberis, F Gianotti, C Guyot, R Hawkings, I Hinchliffe, B Heinemann, A Höcker, A Klimentov, G Lehmann, P Nevski, and H von der Schmitt. ATLAS Dataset Nomenclature. Technical Report ATL-GEN-INT-2007-001. ATL-COM-GEN-2007-003, CERN, Geneva, Nov 2007.

[130] HEP-SPEC06 benchmark. https://w3.hepix.org/benchmarks/.