PANDA Grid – a Tool for Physics

D. Protopopescu

School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, UK

K. Schwarz

Gesellschaft für Schwerionenforschung, Planckstr. 1, D-64291 Darmstadt, Germany

E-mail: `k.schwarz@gsi.de`

Abstract. PANDA Grid is the computing tool of the $\bar{\text{P}}$ANDA experiment at FAIR with concerted efforts dedicated to evolving it beyond passive computing infrastructure, into a complete and transparent solution for physics simulation, reconstruction and analysis, a *tool* right at the fingertips of the physicist. $\bar{\text{P}}$ANDA's position within the larger FAIR community, synergies with other FAIR experiments and with ALICE@LHC, together with recent progress are reported.

1. GSI, FAIR and $\bar{\text{P}}$ANDA

GSI (Gesellschaft für Schwerionenforschung) operates a large and in many aspects worldwide unique accelerator facility and employs more than one thousand people. Researchers from Europe and from around the world conduct here experiments extending from nuclear and atomic physics to plasma and materials research, and encompassing biophysics and cancer therapy. GSI is probably best-known for the discovery of six new chemical elements and the development of a novel type of tumor therapy using ion beams.

Centered on GSI, in the years to come an international structure named FAIR (Facility for Antiprotons and Ion Research) will evolve. The international FAIR GmbH was founded on October 4, 2010 and the first beam is expected by 2016. Roughly 3,000 scientists from more than 40 countries are already working on the planning of the experiment and accelerator facilities. FAIR will practically use GSI as injector, but two new synchrotrons SIS100 and SIS300, a Super-FRS and several storage rings among which one that will be used for antimatter research will be built.

$\bar{\text{P}}$ANDA, located on the HESR storage ring, is one of the key experiments at FAIR. $\bar{\text{P}}$ANDA stands for "Anti-**P**roton[1] **AN**nihilations at **DA**rmstadt" and will study via proton-antiproton collisions various topics around the weak and strong forces, exotic states of matter, and the structure of hadrons. For this purpose, a versatile detector is being built, that will be able to provide precise trajectory reconstruction, energy and momentum measurements and very efficient identification of charged particles. The design and construction of this detector is being done by the $\bar{\text{P}}$ANDA Collaboration, which comprises more than 450 scientists from 17 countries.

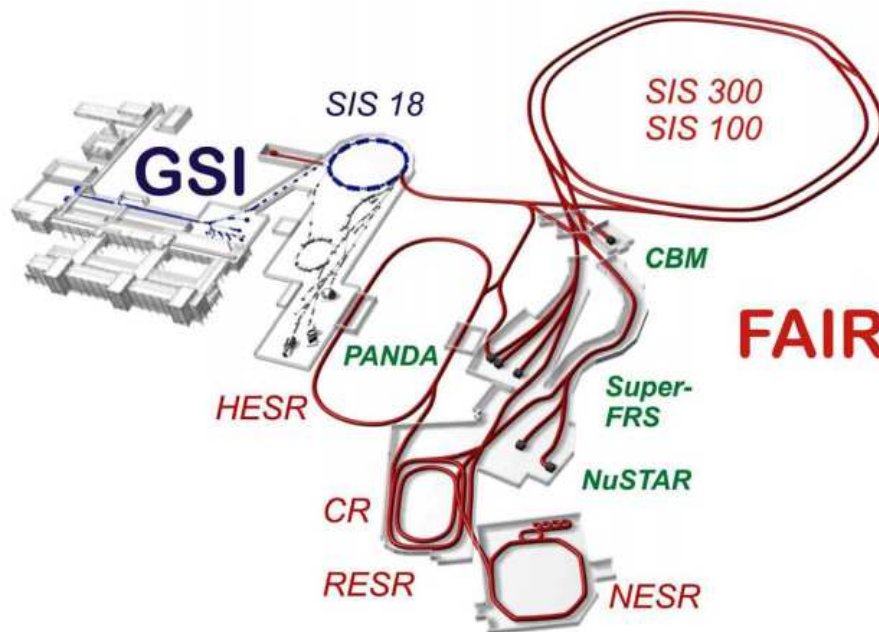[1] The antiproton is customarily represented by a barred $\bar{p}$

Figure 1. The FAIR accelerator and storage ring complex

2. P̄ANDAGrid

Since it is not possible to add all needed computing resources at a single place experiments such as P̄ANDA [1] or ALICE [2], with complex final states and very high data acquisition rates would not stand a chance without a Grid model for their computational needs. The P̄ANDA experiment will use PANDA Grid as its computing infrastructure for simulations, reconstruction and data analysis [3] .

PANDA Grid has started being deployed in 2004, and has become now practically the *first up-and-running component* of the P̄ANDA experiment while continually growing and adding new sites. PANDA Grid encompasses now 19 sites from 13 institutes in 9 countries, including EGEE/EGI sites integrated via VOBox tools.

PANDA Grid employs AliEn *middleware* [4], developed by the ALICE collaboration at CERN but used now by a few other FAIR experiments. The data analysis *software*, that stands on top of the *middleware*, is centered on PandaRoot [5], the
core framework developed at GSI IT [6] in collaboration with P̄ANDA and other experiments of the FAIR community. Grid monitoring and data production supervision are done via MonALISA [7].

PANDA Grid serves now as a benchmark environment for PandaRoot, is used to run theoretical calculations, simulations and other jobs related to detector design and prototyping, run weekly builds to test latest software releases on various platforms, and store, reanalyze and share data within the P̄ANDA community.

3. Components and Capabilities

The core component of PANDA Grid, the AliEn (**Ali**ce **En**vironment) middleware, is a lightweight Open Source Grid Framework built around other Open Source components using the combination of a Web Service and Distributed Agent Model [4]. It started within the ALICE Off-line Project at CERN and was meant primarily as the production environment for simulation, reconstruction, and analysis of physics data of the ALICE Experiment. It has been

used in production at CERN for more than ten years and during the last half decade, several FAIR experiments started employing this framework as well (notably P̄ANDA and CBM [8]).
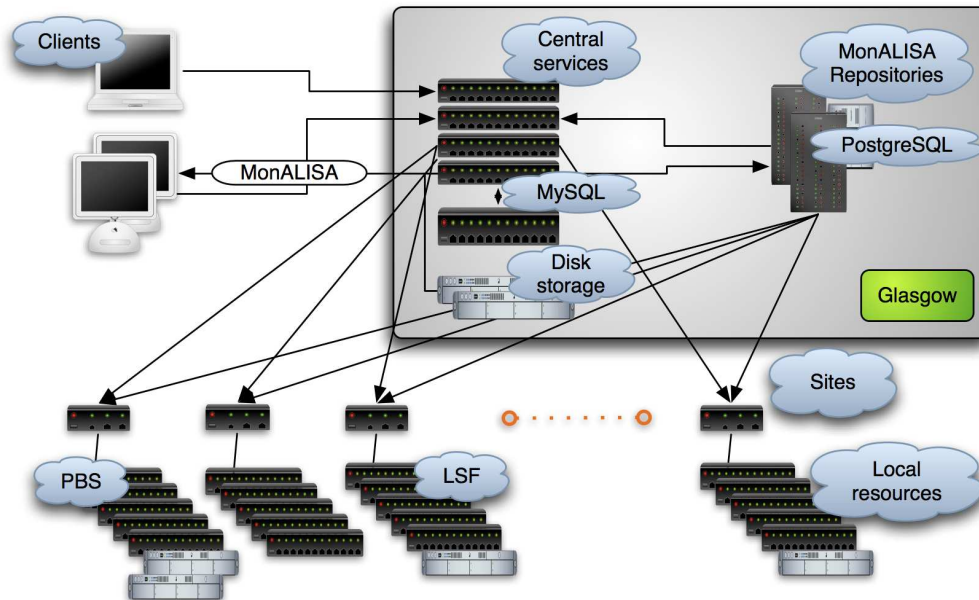

Figure 2. PANDA Grid components scheme. Central services and databases are located at the University of Glasgow, UK

The MonALISA system is an ensemble of agent-based subsystems registered as dynamic services, that are able to collaborate and cooperate in performing a wide range of information gathering and processing tasks. This is a very scalable system, designed to easily integrate existing monitoring tools and procedures and to provide this information in a dynamic, customized, self describing way to any other services or clients [7]. MonALISA can monitor all aspects of complex systems (system information on computer clusters and nodes, network status, performance of applications and services), can interact with any other services to provide in near real-time customized information based on monitoring information, can provide secure, remote administration for services and applications. MonALISA provides agents to supervise applications, restart or reconfigure them and to notify other services when certain conditions are detected. MonALISA provides a very user-friendly and highly customizable Graphical User Interfaces to visualize complex information and is ideal for global monitoring repositories of distributed Virtual Organizations. MonALISA provides system and task monitoring, remote administration and control, data production management via the PANDA ML Repository [9].

PandaRoot is the simulation and reconstruction software framework for the P̄ANDA experiment. It is an extension on FairRoot [6], a framework jointly developed by GSI IT, CBM and P̄ANDA. The FairRoot framework is fully based on ROOT [10] and the Virtual Monte-Carlo (VMC) interface [11]. The user can create simulated data and/or perform analysis within a common framework. The FairRoot framework delivers base classes which enable users to construct their detectors or analysis tasks in a simple way, and it delivers in addition general functionality like track visualization. An interface for reading magnetic field maps is also implemented. PandaRoot then adds PANDA-specific physics and detector geometry. The software is distributed on PANDA Grid in two source packages, FairRoot and PandaRoot, compiled on sites.

In terms of hardware resources, GSI and JINR are at present the our largest sites, but GSI and later on FAIR will very likely be the major contributors to PANDA Grid. Estimates for FAIR computing resources forecasted for 2018 are: CPU 100 MSI2k, Online 33 PByte, Archive 33 PByte. GSI is also engaged in LHC computing since 2001: part of WLCG (ALICE-Tier-2) is located at GSI, GSI is developing software and operating procedures for the LHC, and GSI scientists are actively involved in LCH boards and committees. This allows for a great deal of expertise transfer between LCH and FAIR since the data rates and duty cycles are of the same order of magnitude. Synergy effects between FAIR, $\bar{\text{P}}$ANDA, CBM and the AliEn community are created by sharing goals, soft- and middleware.

4. A Tool for Physics

Although customarily referred to as 'infrastructure', PANDA Grid is being developed beyond a passive computing infrastructure. Concerted efforts from a heterogeneous team of physicists and computer scientists are focused on making PANDA Grid a *'tool for physics'*, that is a *complete and transparent solution for physics simulation, reconstruction and analysis*, with minimal overhead for the user (*i.e.* the physicist).

The whole $\bar{\text{P}}$ANDA software for data simulation, reconstruction and analysis is based on PandaRoot. In the $\bar{\text{P}}$ANDA computing model, the *user* software is mainly based on ROOT macros using modular precompiled libraries from PandaRoot. The PANDA Grid infrastructure is used for software distribution via the package manager service (PackMan) provided by the middleware layer. We have now completely integrated and automatic system for package installation, compilation and testing, supervised directly by PandaRoot developers, such that the user has to only customize a couple of scripts before pursuing his/her own physics.

AliEn is the perfect tool for the job, as it already provides advanced data production tools at the middleware level (various mechanisms for job split and merging, catalogue tags and triggers, file collections, automatic data replication across multiple SEs [4]). The addition of the MonALISA integration layer brings a range of supervision tools, command mechanisms and browser-based administration interfaces that make PANDA Grid easy to manage and intuitive to use.

5. Recent Progress

Close contact between the subgroups participating in this effort is maintained via periodical Data Challenges, numerous ad-hoc meetings as well as via dedicated bi-annual workshops [13], hosted in turn by the various participant institutes. Four PANDA Grid workshops were held during 2009-2010, in the Philippines, Italy, UK and Switzerland, the last one being a joint ALICE-$\bar{\text{P}}$ANDA meeting hosted at CERN.

Data Challenges (DC) [14], during which large quantities of data is produced via a full chain of simulation, digitization and reconstruction are a tool for stress-testing the middleware, software and site storage and network capabilities. The jobs run during the DCs are either theoretical calculations or simulations used in the process of design and prototyping of the $\bar{\text{P}}$ANDA detector.

Significant effort has been dedicated during the last years to perfecting the organizational framework for administration and usage of PANDA Grid. A *production management* scheme has been set in place (and tested during the DCs), complete with user, data storage and scheduling schemes. Since the start, there has been a mutually beneficial collaboration in terms of software development and Grid deployment between ALICE and the FAIR experiments using AliEn, with frequent joint meetings since 2008 [15].

References

[1] $\bar{\text{P}}$ANDA experiment portal, http://www-panda.gsi.de

[2] The ALICE experiment at the CERN LHC, The ALICE Collaboration *et. al*, 2008, JINST 3 S08002

[3] D. Protopopescu *et. al*, PANDA Grid Activities, GSI Scientific Report 2008

[4] P. Buncic, A.J. Peters, P. Saiz and J.F. Grosse-Oetringhaus, The architecture of the AliEn system, CHEP 2004, Interlaken, Switzerland (2004) 440

[5] PandaRoot documentation at http://panda-wiki.gsi.de/cgi-bin/view/Computing/PandaRoot

[6] M. Al-Turany *et. al*, Status of the FairRoot Simulation and Analysis framework, INSTR. METH. 59, GSI Scientific Report 2008

[7] MonALISA portal, http://monalisa.cern.ch

[8] CBM experiment, http://www.gsi.de/forschung/fair_experiments/CBM

[9] MonALISA Repository for PANDA VO, http://mlr2.gla.ac.uk

[10] The ROOT framework, http://root.cern.ch

[11] R. Brun *et al.*, Virtual Monte Carlo, CHEP La Jolla, CA, 2003, p. 24-28

[12] GSI documentation Wiki at http://panda-wiki.gsi.de/cgi-bin/view/Computing/PandaGrid

[13] PANDA Grid Workshops website, http://nuclear.gla.ac.uk/grid-workshop/

[14] see 'Data Challenges' links in [12]

[15] M. Al-Turany *et. al*, Grid Activities at GSI, GSI Scientific Report 2008