E-Science Activities at GSI

K. Schwarz for the e-science group/Scientific Computing[1]
[1]GSI, Darmstadt, Germany

This article describes the work of the GSI e-science Group with the aim to operate an ALICE tier2 centre within the global environment of the LHC Computing Grid and to prototype a distributed computing environment for FAIR.

ALICE tier2 centre at GSI and ALICE Grid in Germany

The ALICE tier2 centre and the National Analysis Facility at GSI provide a computing infrastructure for ALICE Grid and for the local usage of the German ALICE groups. New data-sets are being transferred to GSI continuously and are processed on the local batch farm via daily running analysis trains. Disk space is provided via the Cluster File System Lustre. The storage resources pledged to the global ALICE community (550 TB) are provided via a Grid Storage Element which consists of an xrootd daemon running on top of the Lustre file system. Data stored on the Grid Storage Element can also be read from local ALICE jobs. The remaining Lustre capacity (4.7 PB) is used for local storage but has to be shared with other experiment groups of GSI/FAIR. Throughout the year GSI participates in centrally managed ALICE Grid productions and data analysis activities, but also analysis jobs of individual users are running on the ALICE tier2 centre. Since recently the routing problem between the GSI batch farm and the local ALICE Storage Element has been solved the average job efficiency of Grid jobs running at the ALICE tier2 centre is close to that of other high level tier 2 centres. The overall job share in 2013 contributed by GSI tier2 and Forschungszentrum Karlsruhe (ALICE tier1 centre) as well as the HHLR compute cluster at Goethe University in Frankfurt has been 11% of all ALICE Grid jobs worldwide. This corresponds well with the pledged CPU resources for 2013: 7000 HEP-SPEC06 for GSI tier2 (4% of the global T2 requirements) and 30000 HEP-SPEC06 for FZK (30% of the global T1 requirements)

CRISP and LSDMA

The Cluster of Research Infrastructures for Synergies in Physics (CRISP) project is a collaboration between different institutions and facilities related to physics research. GSI participated in Work Package 16 and contributed in developing a pan-European system for unique identification. As a prototype solution Umbrella has been presented. The work of this group concentrated on bridging solutions between the Umbrella system and X509 certificates as used in Grid communities as well as between Umbrella and EduGAIN.

The work of the Data Life Cycle Lab ”Structure of Matter/FAIR” within the portfolio project ”Large Scale Data Management and Analysis (LSDMA)” is being defined in close collaboration with the FAIR experiments. Main work topics are parallel and distributed computing as well as global data federations.

KOSI

Within the context of the KOSI program (Kooperativer Studiengang Informatik) from Hochschule Darmstadt there is a computing project in close collaboration with the GSI theory group. It explores the synergetic use of different software packages like Mathematica and MathCode from Wolfram with Geneva from Gemfony Scientific and with Lapack, ScaLapack and GSL on the GSI computer clusters. A first application on the chiral extrapolation of baryon masses has been worked out [1].

Preparation for FAIR

In order to be able to prepare a distributed computing concept for FAIR experience is being accumulated starting with currently existing computing environments. Currently the focus lies on the AliEn middleware since know-how is available due to the ALICE tier2 centre at GSI. Based on a gap analysis it will be decided if the FAIR computing requirements can be mapped on further developed current systems or if major components if not all have to be replaced. Well-grounded decisions can be done by evaluating the performance of test beds. Therefore PANDA-Grid came into existence and is now in successful operation since 2004. It has been extended to currently 15 sites in Asia, Europe, and North America. Among these 3 sites are running via WLCG infrastructure. The most recent site which joined is Orsay in France and also the GSI Icarus cluster has been interfaced to PandaGrid. The good collaboration between ALICE Offline and the PandaGrid project is being continued, so recently a common workshop in Torino took place. The central PandaGrid services as well as the central Grid data base (MySQL) are running at GSI while the central MonaLisa monitoring repository and a backup database are maintained in Torino.

References

[1] M.F.M. Lutz, K. Schwarz, R. Bavontaweepanya, and C. Kobdaj, “On finite volume effects in the chiral extrapolation of baryon masses”, this report.