

Status of the Grid Computing for the ALICE Experiment in the Czech Republic

D Adamová¹, J Chudoba², T Kouba², P Mendez Lorenzo³, P Saiz³, J Švec²,
J Hampl¹

¹Nuclear Physics Institute of the AS CR Rez/Prague

²Institute of Physics of the AS CR Prague

³CERN - European Organization for Nuclear Research

E-mail: adamova@ujf.cas.cz

Abstract: The Czech Republic (CR) has been participating in the LHC Computing Grid project (LCG) ever since 2003 and gradually, a middle-sized Tier-2 center has been built in Prague, delivering computing services for national HEP experiments groups including the ALICE project at the LHC. We present a brief overview of the computing activities and services being performed in the CR for the ALICE experiment.

1. Introduction

The Computing model of the ALICE experiment at the CERN LHC* [1, 2] relies on the ALICE Computing Grid, a global distributed infrastructure of resources for data storage and processing. As of today, this project is made of over 80 sites spanning 5 continents (Africa, Asia, Europe, North and South America), involving 6 Tier-1 centers and more than 70 Tier-2 centers [3]. Altogether, the resources provided by the ALICE sites represent about 10,000 CPUs and 10 PB of distributed storage, and the gradual upscale of this capacity is ongoing. About 1-2% of these resources have been provided, for several years, by the Tier-2 center in Prague, Czech Republic (CR) [4] (cf. Figures 1, 2). The Prague site was one of the first centers participating in the ALICE Grid project; it joined the project in 2003, when the ALICE Grid involved only 13 centers and a few hundred of CPUs.

63. NIKHEF	erf.nikhef.nl		-	
64. NSC	alien-ce.smokerings.nsc.liu.se	v2-16.68	13.05.2009 11:57	
65. OSC	websvcs03.osc.edu	v2-15.53	13.05.2009 11:47	
66. PANDA	130.209.45.237		-	
67. PNPI	gt1.pnpi.nw.ru	v2-16.60	13.05.2009 12:01	
68. Poznan	vobox.reef.man.poznan.pl	v2-16.56	13.05.2009 11:59	
69. Prague	goliasx31.farm.particle.cz	v2-16.59	13.05.2009 11:48	
70. RAL	lcv0597.gridpp.rl.ac.uk	v2-16.31	13.05.2009 11:58	
71. RRC-KI	house.grid.kiae.ru	v2-16.59	13.05.2009 11:46	

Figure 1. Status of AliEn services at ALICE sites (a part of the table showing the Prague site)

* For the list of acronyms used in this article, see the Appendix.



Figure 2 (color online). Map of ALICE sites in Europe

2. Regional Computing Center for Particle Physics in Prague

The Czech entry point into the ALICE Grid is the computer farm Golias (cf. Figure 3), located in Prague at the Institute of Physics AS CR (FZU). The availability of the farm for ALICE has been ensured by the Czech ALICE group’s investments into the farm hardware. Accordingly, the computing services have been provided for the processing of the ALICE production and analysis jobs at the farm [4].



3. Hardware resources

The computing center/farm Golias is the biggest site in the CR providing computing and storage services for Particle Physics experiments (apart from ALICE, most of the CPU time is usually consumed by the Tevatron experiment D0 and the LHC experiment ATLAS), and also for Solid State Physics and Astrophysics. The computing and storage resources [5] represented about 450 CPUs and 50 TB of disc space in the end of 2008 and will be gradually expanded to about 1550 CPUs (corresponding to 3,3 MSI2k) and 216 TB of disc space during 2009. The new hardware includes e.g. the IBM iDataPlex (672 cores) and different systems SGI Altix. The farm has excellent network connectivity (see Figure 4, cf. also [6]) with other institutions, both national and worldwide (1 to 10 Gb/s).

Figure 3 (color online). The farm Golias

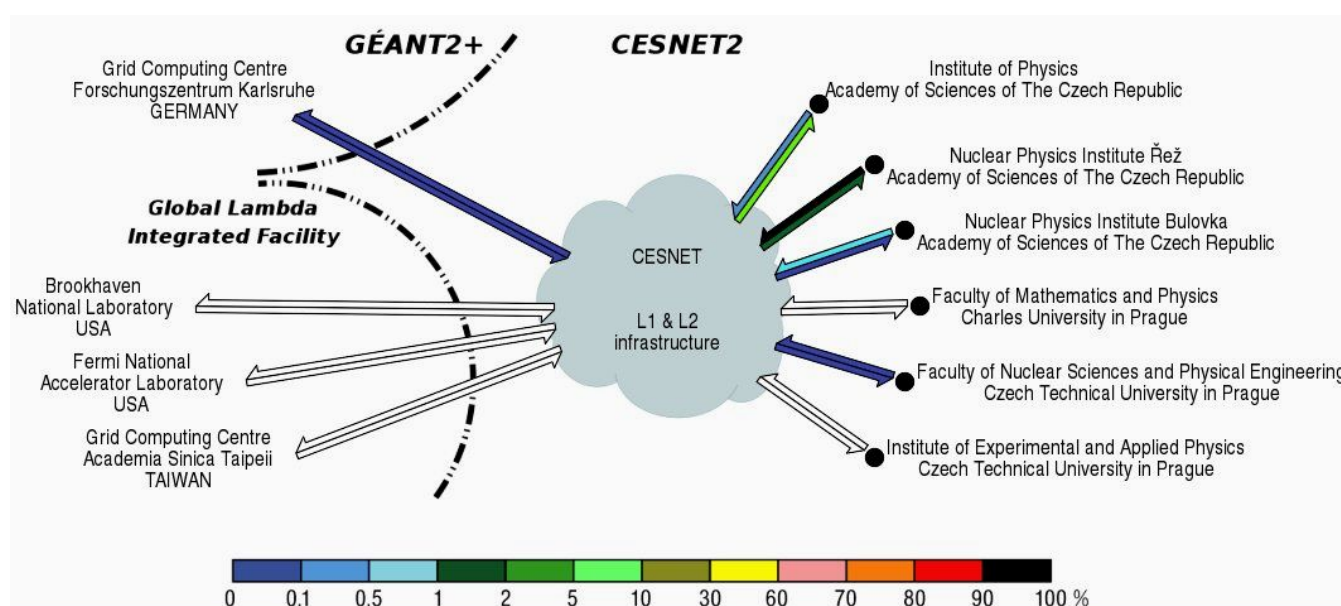


Figure 4 (color online). Topology and traffic of CESNET2, the Prague-based network facility for HEP

4. Grid connectivity

The Golias farm is integrated via the installed middleware components into the LCG. The installed gLite grid middleware components include CE, SE, UI, MON box, site BDII and have been configured with YAIM, integrated into Cfengine used for the management of the local site changes. Since 2005, the site is a certified Tier-2 center of the LCG project. In 2008, the Czech Republic represented by the FZU/farm Golias signed the Memorandum of Understanding of the Worldwide LHC Computing Grid Collaboration (WLCG) (cf. Figure 5).

Czech Rep., FZU, Prague	2008	2009	2010	2011	2012	2013	Split 2009	ALICE	ATLAS
CPU (kSI2K)	164	376	637	996	1517	2235	Offered	220	156
							% of Total	2%	1%
Disk (Tbytes)	35	72	201	367	584	867	Offered	35	37
							% of Total	1%	0%
Nominal WAN (Mbits/sec)	10000	10000	10000	10000	10000	10000			

Figure 5 (color online). Czech Republic pledges to WLCG

5. Job management and monitoring

Job management at the Golias farm is performed using the local resource management system PBSPro (current version 9.2). In preparation and testing are the resource manager TORQUE and the cluster scheduler Maui. Hardware, software and network status are extensively monitored using many standard packages including Nagios, Munin, MRTG, IPAC, SAM and locally developed custom tools [5] (see e.g. Figure 6).

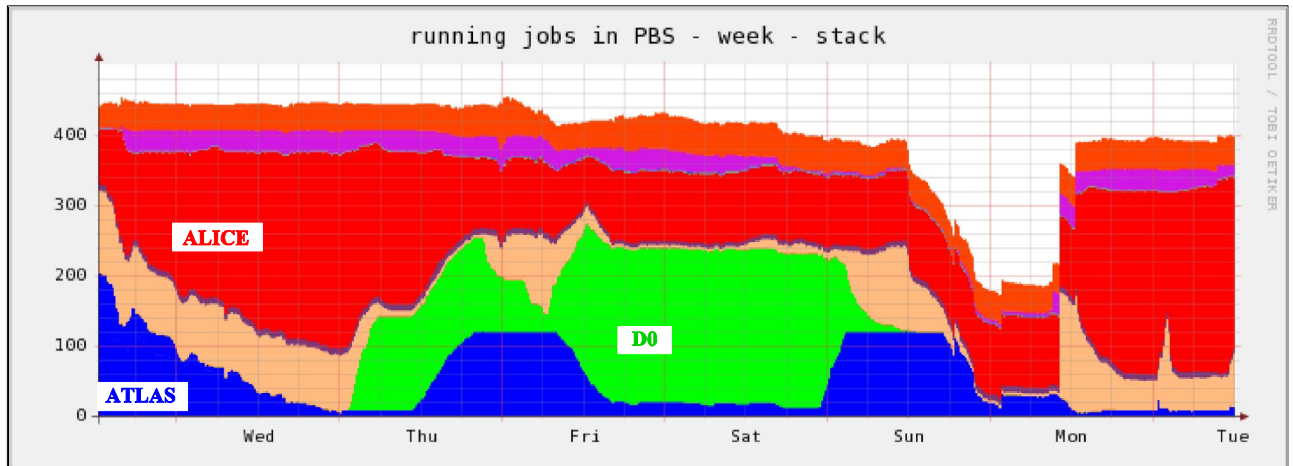


Figure 6 (color online). Monitoring of jobs in PBS:

Jobs on Golias in the period 24.02.2009 – 03.03.2009 (ALICE jobs indicated by red color)

6. Physics Data Challenges

As one of the Tier-2 centers, the Prague farm Golias provides computational and storage services to perform Monte Carlo (MC) simulations of the p+p and Pb+Pb collisions and for end user analysis [7]. Similar to the other LHC experiments, ALICE has been running massive MC simulations during comprehensive tests of its computing model and Grid infrastructure aimed at assessing the physics performance of the detector, the so-called Physics Data Challenges (PDC). The first distributed MC productions date back to 2003. Between April 2006 and October 2008, the PDC has been running in a quasi-permanent mode, and since then have followed individual MC productions for the First Physics program [8]. The data files generated during the PDCs are stored on the Grid and are available for further processing. For example, during the period April 2006 – July 2008, roughly 330 millions of events with various physics content have been produced.

7. ALICE PDCs in Prague

Grid computing for ALICE in the CR has started in 2003. The integration into the ALICE Grid system is provided by a combination of the (LCG) gLite and AliEn [9] middleware components. One member of the local group is appointed as the local ALICE production manager, responsible for the production of simulated data, control of the analysis jobs and maintenance/upgrade of the AliEn middleware.

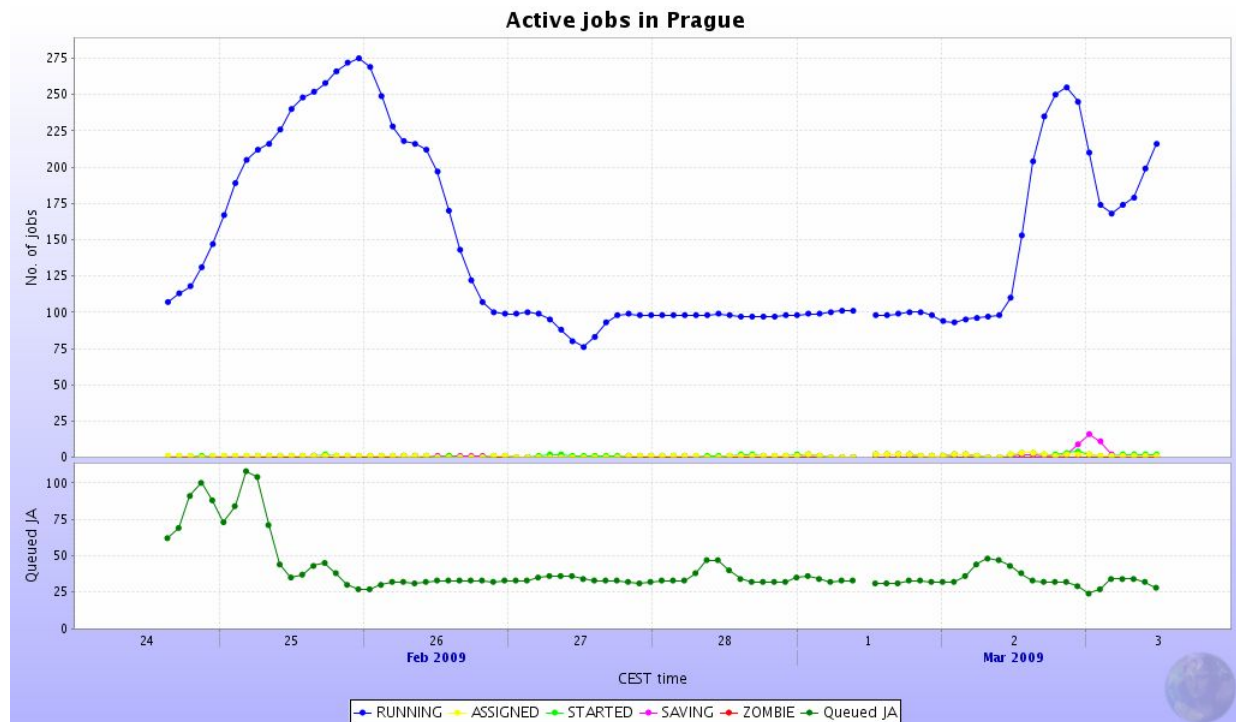


Figure 7 (color online). ALICE Grid monitoring:

Active jobs in Prague in the period 24.02.2009 - 03.03.2009 (max. 278 jobs)

8. Performance of the Prague site

The number of jobs executed at the Golias farm during the ALICE PDCs has been most of the time higher than the officially pledged resources would guarantee (cf. Figure 7). The number of completed production jobs was up to 4 – 7% of the overall ALICE production during PDCs 2004 – 2006. After that the relative contribution from the CR decreased to about 1,5% (see Figure 8, cf. e.g. [10]), which was due to the increasing number of ALICE sites participating in the PDC and to the corresponding increase of the global resources. During the PDC 2007 - 2008, the number of jobs successfully executed has been more than 100,000 (~ 1,5% of the whole production). This was achieved by maintaining a steady job submission and usage of all available farm CPUs for the ALICE jobs. An important feature of the systematic job monitoring is the control of the log files produced by the jobs, which contributes to the debugging and validation of AliEn and of the ALICE production software AliRoot.

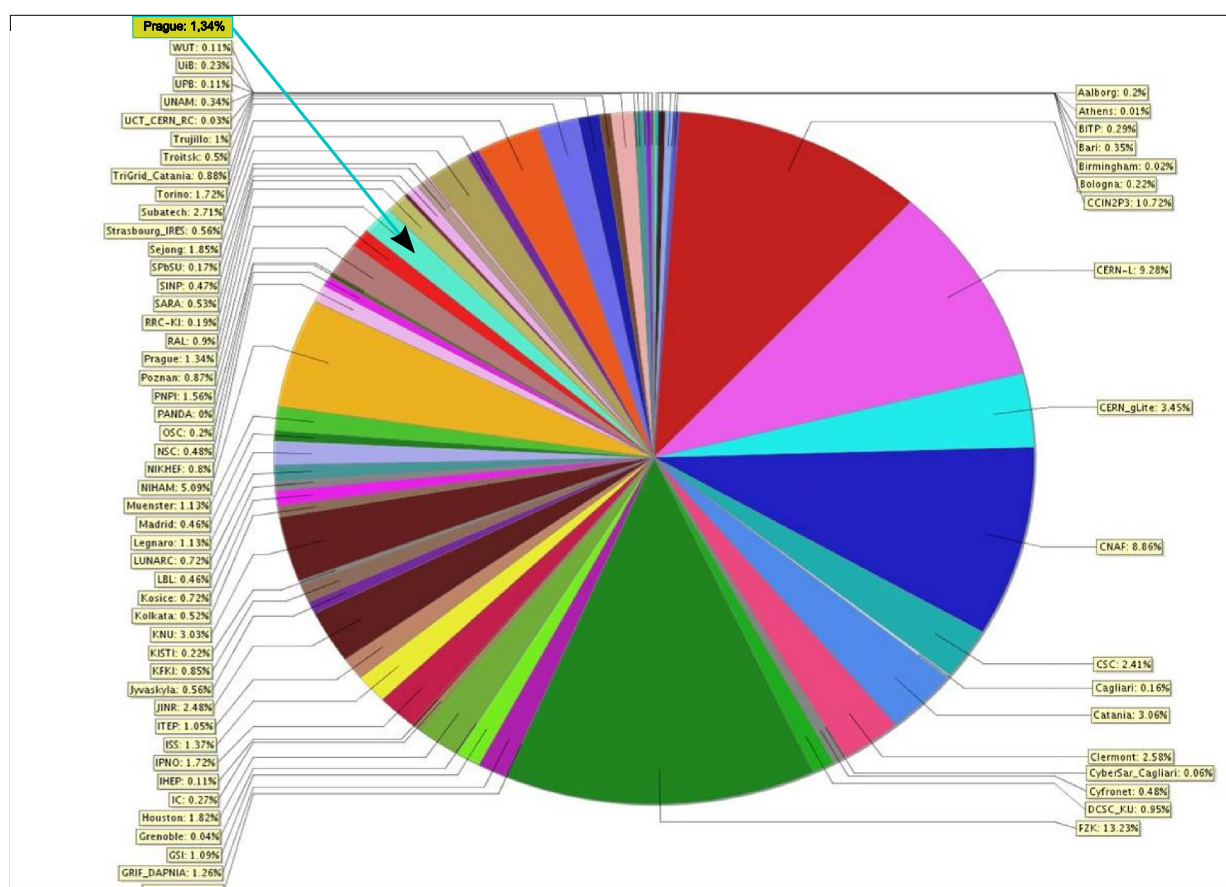


Figure 8 (color online). Relative site contributions in 2007 - 2008 (Prague 1.34%)

9. The ALICE storage

During the last 3 - 4 years, ALICE put a considerable effort into the building of a distributed storage system. Currently, ALICE operates more than 50 storage system endpoints distributed over 30 sites, providing 4 different flavors of storage solutions with the total capacity ~ 10 PB [11]. The Tier-2 centers are required to provide the disc-based storage, and are recommended to follow a rule 1TB of disc storage per 2 - 3 CPUs.

The Czech group has been building a storage cluster for ALICE for about 2 years. As a result, two Storage Elements (SE) installed as pure Xrootd SE are available: ALICE::Prague::Disk (2 TB) and ALICE::Prague::Disk2 (20 TB) (see Figure 9). This small storage cluster is a step towards building a distributed WLCG Tier-2 center for HEP experiments in the CR: it is located at a site about 20 km far from the farm Golias. The good performance of the cluster is ensured by the excellent network connectivity (combined 2 Gb/s public and 1 Gb/s E2E links) (cf. Figure 10). It is planned to upgrade this Czech storage cluster for ALICE with another 25TB server.

37. NDGF - DCACHE	ALICE::NDGF::DCACHE	OK	32.18 TB	36.18 TB	32.18 TB	112.4%	439,304	srn
38. NDGF - DCACHE_SINK	ALICE::NDGF::DCACHE_SINK	OK	837.9 TB	271 GB	837.9 TB	0.032%	3,488	srn
39. NDGF - DCACHE_TAPE	ALICE::NDGF::DCACHE_TAPE	OK	24.18 TB	31.07 GB	24.18 TB	0.125%	3,370	srn
40. NIHAM - FILE	ALICE::NIHAM::FILE	OK	85.53 TB	31.65 TB	85.53 TB	37.01%	6,504,491	File
41. PNPI - DPM	ALICE::PNPI::DPM	OK	27.34 TB	37.39 GB	27.31 TB	0.134%	26,829	SRM
42. Prague - DISK	ALICE::Prague::DISK	OK	1.268 TB	720.3 GB	577.6 GB	55.5%	194,225	File
43. Prague - DISK2	ALICE::Prague::DISK2	OK	34.59 TB	9.55 TB	25.04 TB	27.61%	948,227	File
44. RAL - CASTOR2	ALICE::RAL::CASTOR2	OK	931.2 GB	22.48 GB	931.2 GB	2.414%	9,736	CASTOR
45. RAL - CASTOR2_SINK	ALICE::RAL::CASTOR2_SINK	OK	90.95 PB	29 KB	90.95 PB	-	1	CASTOR
46. RRC-KI - DPM	ALICE::RRC-KI::DPM	OK	113.3 TB	8.23 TB	105.1 TB	7.265%	249,305	SRM

Figure 9 (color online). Status of the ALICE Storage Elements
(a part of the table showing the Prague SEs)

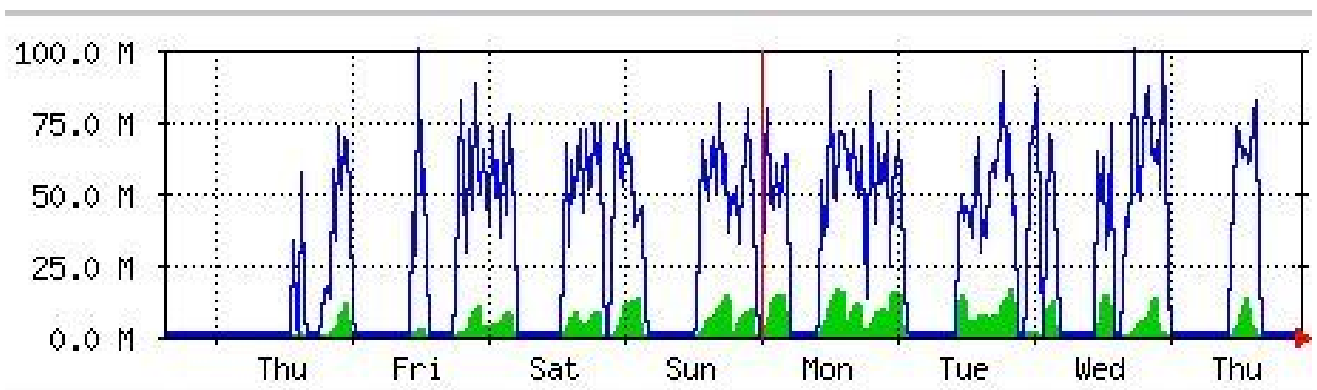


Figure 10 (color online). Load on the E2E link Golias – Prague SE, 26.02 – 05.03.2009
(max. out 99.6 MB/s)

10. Summary

The CR has been an active participant into the ALICE Grid project ever since 2003. The entry point into the ALICE Grid has been the computing farm Golias in Prague, a medium sized Tier-2 center fully integrated into the LCG and ALICE environment. The local ALICE group has been officially offering 1-2% of the overall ALICE computing resources; however the delivered computing performance has often been exceeding these resources.

The Prague center will, according to the Tier hierarchical system, continue the production of Monte Carlo simulated data and the processing of the user analysis jobs also after operation of the LHC is resumed in October 2009. To deliver the computing performance pledged by the CR to the WLCG MoU, we plan to perform regular upgrades of the hardware resources dedicated to ALICE and to maintain the current reliability of the computing services.

ACKNOWLEDGEMENTS

We highly appreciate the critical reading of the manuscript by Federico Carminati and Latchezar Betev (CERN) and by the members of the ALICE Editorial Board.
The work was supported by the MSMT CR contracts No. 1P04LA211 and LC 07048.

Appendix: a list of used acronyms

ALICE	A Large Ion Collider Experiment (at the CERN LHC).
AliEn	ALICE distributed computing ENVIRONMENT.
AliRoot	ALICE software package for data simulation/reconstruction/analysis.
BDII	Berkeley Database Information Index. Information service used on Grid.
CE	Computing Element.
Cfengine	configuration management system for Unix.
CPU	Central Processing Unit. In this paper a synonym for a processor core.
EGEE	Enabling Grids for E-science. Europe's leading grid computing project.
gLite	the EGEE-developed middleware for grid computing.
HEP	High Energy Physics.
IPAC	a Linux tool for the bandwidth monitoring based on the iptables/ipchains.
LDAP	Lightweight Directory Access Protocol.
LHC	Large Hadron Collider (at CERN).
LCG	LHC Computing Grid.
Maui	a job scheduler for clusters and supercomputers.
MON box	gLite monitoring system collector server.
MRTG	Multi Router Traffic Grapher. A tool to monitor the traffic load on network links.
Munin	a monitoring tool for survey of computer clusters.
Nagios	a tool for comprehensive network monitoring.
PBSPRO	Portable Batch System. A workload management system.
Torque	a resource manager for control over batch jobs.
SAM	Service Availability Monitoring.
SE	Storage Element.
YAIM	a tool for configuration of Grid Services.

REFERENCES

- [1] ALICE Collaboration: ALICE: Physics Performance Report, Volume I., J. Phys. G30 (2004) 1517; Volume II., J. Phys. G32 (2006) 1295.
- [2] ALICE Experiment Computing TDR, <http://aliceinfo.cern.ch/Collaboration/Documents/TDR/Computing.html>.
- [3] ALICE Sites Performance, <http://pcalimonitor.cern.ch/reports/index.jsp>
- [4] Regional computing center for particle physics, <http://www.particle.cz/farm/index.aspx?lang=en>.
- [5] J. Svec: The Status of Prague Computing Farm, http://gemma.ujf.cas.cz/~adamova/pres/FARM/seminar_svec_Dec2008.ppt.
- [6] CESNET2: experimental facility for High Energy Physics, <http://www.ces.net/netreport/hep-cesnet-experimental-facility/>
- [7] LHC Computing Grid: *Technical Design Report*, <http://lcg.web.cern.ch/LCG/tdr/>
- [8] ALICE MC Production Cycles, http://pcalimonitor.cern.ch/job_details.jsp.
- [9] Saiz, P. et al., AliEn - ALICE environment on the GRID, Nucl. Instrum. Meth., A502 (2003) 437-440; <http://alien.cern.ch/twiki/bin/view/AliEn/Home>.
- [10] F. Carminati and L. Betev : The ALICE distributed computing, <http://niham.nipne.ro/aliceworkshop08/download/Sibiu0808fca.pdf>.
- [11] ALICE Storage Elements, <http://pcalimonitor.cern.ch/stats?page=SE/table>.