CMS Tier structure and operation of the experiment-specific tasks in Germany

A. Nowack on behalf of the German CMS community
III. Physikalisches Insitut B, Physikzentrum, RWTH Aachen University, 52056 Aachen, Germany
E-mail: andreas.nowack@cern.ch

Abstract. In Germany, several university institutes and research centres take part in the CMS experiment. Concerning the data analysis, a couple of computing centres at different Tier levels, ranging from Tier 1 to Tier 3, exists at these places. The German Tier 1 centre GridKa at the research centre at Karlsruhe serves all four LHC experiments as well as four non-LHC experiments. With respect to the CMS experiment, GridKa is mainly involved in central tasks. The Tier 2 centre in Germany consists of two sites, one at the research centre DESY at Hamburg and one at RWTH Aachen University, forming a federated Tier 2 centre. Both parts cover different aspects of a Tier 2 centre. The German Tier 3 centres are located at the research centre DESY at Hamburg, at RWTH Aachen University, and at the University of Karlsruhe. Furthermore the building of a German user analysis facility is planned. Since the CMS community in German is rather small, a good cooperation between the different sites is essential. This cooperation includes physical topics as well as technical and operational issues. All available communication channels such as email, phone, monthly video conferences, and regular personal meetings are used. For example, the distribution of data sets is coordinated globally within Germany. Also the CMS-specific services such as the data transfer tool PhEDEx or the Monte Carlo production are operated by people from different sites in order to spread the knowledge widely and increase the redundancy in terms of operators.

1. Introduction
This paper describes the structure of the CMS Tier centres in Germany and the cooperation inside the German CMS community.

2. Sites in Germany
In Germany, several university institutes and research centres take part in the CMS experiment [1]. These are institutes of the RWTH Aachen University [2] and of the universities of Hamburg [3] and Karlsruhe [4] as well as the centres DESY (Deutsches Elektronen-Synchrotron [5], located at Hamburg and Zeuthen) and Forschungszentrum Karlsruhe [6].

Computing centres at different Tier levels exist at these sites. Their locations and their Tier levels are shown in figure 1. The German Tier 1 centre is called GridKa [7] and is located at the research centre at Karlsruhe. It serves all four LHC experiments as well as four non-LHC experiments. According to the CMS computing model [8–10], typical tasks are data distribution, skimming, re-reconstruction, and archiving of data.

The German Tier 2 centre is formed by a federation of DESY and RWTH Aachen University. Both parts cover different tasks of a Tier 2 centre and are located at different places. The task...
of the computing centre of DESY is the provision of infrastructure for user analyses, which requires sufficient storage capacity for event data. The computing-intensive tasks—Monte Carlo production and detector calibration and alignment—are processed at RWTH.

Tier 3 centres and/or local resources are available at all German sites. They are usable for analyses of the German and/or local users.

Furthermore, the building of a national analysis facility (NAF) at DESY is planned. The NAF will provide interactive and batch analysis environments for the German users. It will be an important part of the analysis infrastructure and will allow the users to process large data sets in parallel.

Table 1 shows an overview of the planned computing resources in Germany for CMS in 2008.
3. German CMS community
The German CMS community is spread over six locations and seven institutes or research centres. The relative size of the different groups is shown in the pie chart in figure 2. In total, about 90 people belong to this community, which is about 5% of the number of physicists in the whole CMS community.

Since the tasks are manifold and the number of people is limited, a good cooperation between the different sites is essential. This cooperation includes physical topics as well as technical and operational issues.

Certain physical topics are addressed at several sites while other topics are treated only at one place. In the first case, an agile communication about the topic itself takes place. In the latter case, the cooperation is reduced to common fields like general analysis strategies or background studies.

On the technical and operational level, the overlap between the topics is even larger. All sites have to provide computing facilities and certain services. Therefore, a large synergy can be gained from the cooperation.

4. Communication tools
Communication between the different groups at the various sites is very important.

4.1. Video conferences
Monthly video conferences (using VRVS [11] and EVO [12]) guarantee that all relevant information can be exchanged between all partners. The topics discussed range from physics analyses to technical and operational problems or improvements of the Tier centres.

4.2. Personal meetings
At least twice a year, all members meet personally at one site for two to three days to discuss the current status and topics which are not suitable for video conferences, e.g. workshops about analysis tools.

4.3. Mailing lists
There are several mailing lists to contact e.g. all members or all site administrators.

4.4. Web pages/wiki
Information of general interest is collected on dedicated web pages in a wiki.

4.5. Private communication (eMail, phone)
People who are working on similar topics often contact each other to discuss and solve common problems.

4.6. Global grid user support
GGUS [13] can contact a team of German CMS experts to solve CMS-specific problems in Germany.

5. Privileges for German users
All German CMS members are allowed to use resources at other German sites. On the technical level, this is accomplished using a special VOMS group (“cms:/cms/dcms”). Thus, each site can recognize German users on the basis of their certificate and grant them higher priority in the batch system or access to dedicated storage systems.
The user can decide whether he/she wants to use such privileges or whether he/she wants to act as an ordinary CMS user by generating a VOMS proxy certificate with or without the group extension, respectively.

6. Monte Carlo production
The different groups in Germany need various Monte Carlo samples e.g. to study systematic effects or to simulate particles beyond the Standard Model, which are not covered by the official Monte Carlo production. In order to ease the work for each person, a centralized service has been established to produce such requests on the German sites. This service uses the same framework [14–16] as the official Monte Carlo production but different databases. Choosing the particular database, a user can analyse the data with the CMS Remote Analysis Builder (CRAB [17]), which hides the interaction with the Grid.

Centralizing the effort gives the benefit that all sites are aware of the already produced data sets and that the priority of different productions can be adjusted taking all interests in Germany into account.

7. Data transfer and distribution of data sets
PhEDEx [18] is an advanced tool by CMS, that handles the transfer of data sets between the different Tier centres using the usual transfer protocols. Each Tier site has to run PhEDEx clients in order to be able to receive data sets.

Since PhEDEx is a very important tool, a smooth operation is needed. All sites run their PhEDEx clients on VO boxes. These machines can be accessed by all people in the VOMS group “cms:/cms/dcms” with the role “cmsphedex”. In case of urgent problems, every German PhEDEx administrator is able to log in at every site and treat the problem.

In order to use the available resources optimally, the data sets which are stored at the German sites and the places where they are stored are coordinated within the German community. The coordination is done using the mailing lists and during the conferences and meetings.

8. Conclusion
The cooperation inside the German CMS community is very fruitful since many years. Since there are many similar tasks at the different locations, a lot of synergy is gained from this cooperation.

Acknowledgments
The author thanks all German CMS members who contributed to the preparation of the poster and/or this paper. This project is funded by the German Federal Ministry of Education and Research.

References
[1] Compact muon solenoid URL http://cms.cern.ch
[2] Faculty of physics of RWTH Aachen University URL http://www.physik.rwth-aachen.de
[3] University of Hamburg: Institute of experimental physics URL http://www.exp.desy.de/institut/
[4] University of Karlsruhe: Institut für Experimentelle Kernphysik (Institute of experimental nuclear physics) URL http://www-ekp.physik.uni-karlsruhe.de
[5] DESY URL http://www.desy.de
[6] Forschungszentrum Karlsruhe GmbH URL http://www.fzk.de
[7] Grid Computing Centre Karlsruhe URL http://www.gridka.de
[8] Della Negra M et al. 2005 CMS computing: technical design report CERN-LHCC-2005-023 (Geneva: CERN)
[9] Eck C et al. 2005 LHC computing grid: technical design report CERN-LHCC-2005-024 (Geneva: CERN)
[10] Lassila-Perini K 2007 CMS workbook: computing model URL https://twiki.cern.ch/twiki/bin/view/CMS/WorkBookComputingModel
[11] Virtual room videoconferencing system URL http://www.vrvs.org
[12] Enabling virtual organizations URL http://evo.vrvs.org
[13] Global grid user support URL http://www.ggus.org
[14] ProdAgent URL https://twiki.cern.ch/twiki/bin/view/CMS/ProdAgent
[15] Data set bookkeeping system URL https://twiki.cern.ch/twiki/bin/view/CMS/DBS-TDR
[16] Data location service URL https://twiki.cern.ch/twiki/bin/view/CMS/DLS
[17] CMS remote analysis builder URL http://cmsdoc.cern.ch/cms/ww/ww/Crab/
[18] Physics experiment data export URL http://cmsdoc.cern.ch/cms/aprom/phedex/