A distributed approach for a regional grid operation centre

Dr. S. Hermann, Dr. H. Marten, J. v. Wezel

Forschungszentrum Karlsruhe, IWR, 76344 Eggenstein-Leopoldsh., Germany

E-mail: sven.hermann@iwr.fzk.de

Abstract. Forschungszentrum Karlsruhe is one of the largest science and engineering research institutions in Europe. The resource centre GridKa as part of this science centre is building up a Tier-1 centre for the LHC project. Embedded in the European grid initiative EGEE, GridKa also manages the ROC (regional operation centre) for the German Swiss region. A ROC is responsible for regional coordination, organisation and support of the sustainable grid operations infrastructure in EGEE. A particularity of this specific ROC is its distributed organisational structure. The German Swiss operations community, currently consisting of 16 resource centres, is supported and organised by a team of six partners. On the one hand, such a decentral approach tends to result in a larger organisational and managerial effort. On the other hand, grid knowledge and expertise is naturally spread among partners with various backgrounds. Temporary staff shortage e.g. during regional vacation times can be compensated for a lot easier. Grid technology is developed and deployed with a broader usability and used by more science communities compared with a central approach. Different aspects of this regional specific organisation structures as part of the global EGEE grid operations as well as roadmaps for upcoming ROC services and tasks will be highlighted. The successful integration of grid tools for operations and support leads to a sustainable structure of the German Swiss region within a world wide grid community.

1. Introduction

The EGEE (Enabling Grids for E-sciencE) project comprises over 70 institutions in 27 European countries. In recent years the project led to the construction of a multi-science grid infrastructure for the ERA (European Research Area). The largest user community with challenging applications for this grid is the LCG (LHC Computing Grid) group. Despite the fact that the EGEE grid is a prototype of this kind of distributed computing resources it already delivered a sustainable infrastructure which will be heavily used after the LHC (Large Hadron Collider) start-up in 2008. The EGEE project has successfully coordinated and supported the build-up of a production infrastructure based on the grid middleware gLite 3. The recent middleware release gLite 3.1 improved the reliability of grid services inside the EGEE grid and introduces Scientific Linux 4 support. As part of the SA1 activity, operations inside the EGEE project is organised in 12 regions. One of these regions is the German Swiss Region shown in Figure 1. Currently this region consists of 16 sites (one of them still to be certified). Together these sites are operating more than 3800 CPUs, more than 2300 TB of disk space and three mass storage systems. Six large resource centres are forming the distributed German Swiss ROC (Regional Operation Centre) as funded EGEE partners inside the project, and its distributed organisational structure is particularly described in the following.
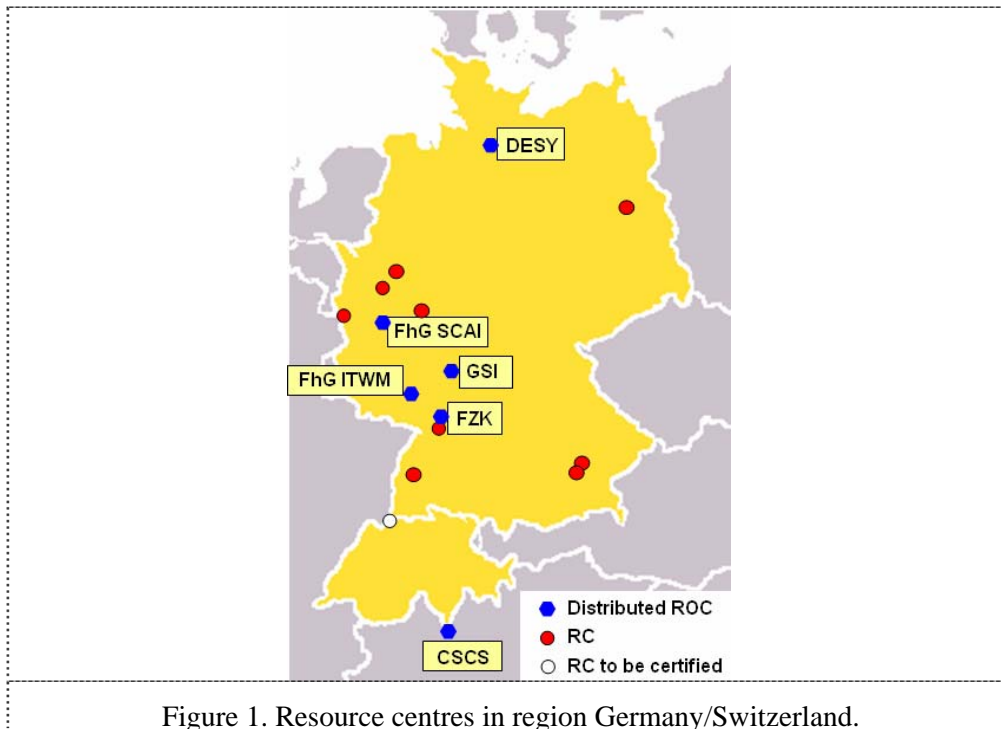
Figure 1. Resource centres in region Germany/Switzerland.

2. Distributed coordination, organisation, support

A main team at the LCG Tier-1 centre GridKa at Forschungszentrum Karlsruhe is managing and coordinating the work for the German Swiss ROC. EGEE organises weekly telephone conferences which include global partners, to discuss and work on actual problems in grid operations. The team of SA1 Germany-Switzerland is steadily contributing to and monitoring these meetings. Information relevant for the regional sites is forwarded, and feedback is gathered and fed into the discussions. Regional operational problems are also regularly communicated and discussed in bi-weekly regional telephone meetings:

http://indico.cern.ch/categoryDisplay.py?categId=334

Furthermore, a regional SA1 support team distributed over five partners is taking shifts in a bi-weekly rotation. This "ROC DECH On Duty" team is following up on trouble tickets coming via the central GGUS portal [1] and helps sites to solve problems. Trouble tickets are also forwarded to global support units e.g. for middleware issues if needed. This way the SA1 team keeps an overview of the situation in the region. Complicated issues get special attention in the team meetings. Discussions in those team meetings show that the status of information among the different partners is equally good, thus the communication channels seem to be working well.

Scalable global core services like GOC-DB, SAM Monitoring, CIC- and GGUS Portal are in place and working in production mode. The GGUS portal and the overall user support is coordinated for the whole EGEE project as part of the DECH SA1 activity. Each region has its own regional support platform. For region DECH this is

https://dech-support.fzk.de

All connected resource centres follow up regional or local site problems reported by users, administrators or the global COD Team (CIC On Duty). SA1 DECH is contributing to COD with a

distributed team. Six partners call up their experts in two distributed regional teams, using chat clients, email and telephone to keep in touch and spot recent problems on a global scope.

Figure 2. The six partners of the distributed ROC Germany/Switzerland.
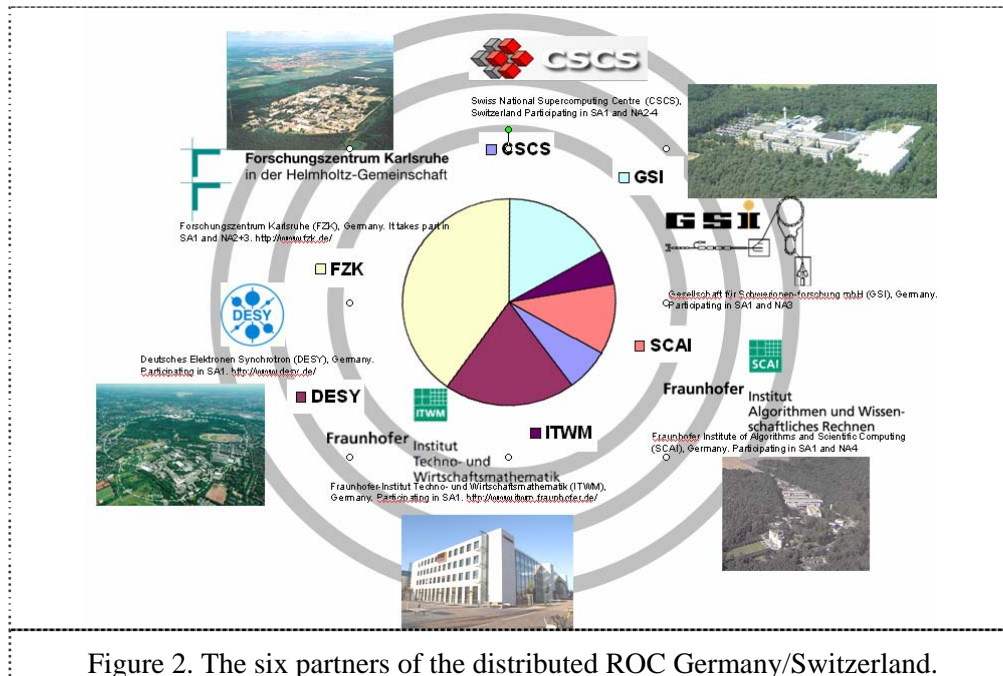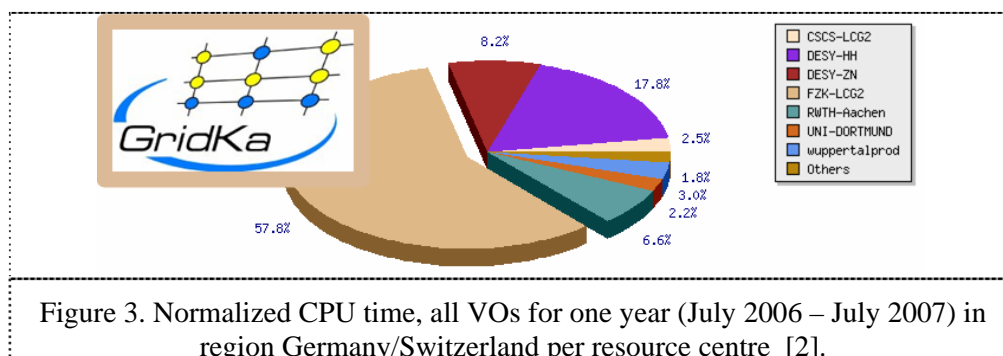
Many elaborated tools for a better distributed organisation have been developed and are continuously improved and updated:
- Public Webpages of EGEE Germany Switzerland http://www.eu-egee.de
- German Swiss SA1 Collaboration Wiki https://twiki.cscs.ch/twiki/bin/view/DECH/WebHome
- Shift schedules for "ROC DECH On Duty" http://egee.fzk.de/sa1/
- Regional Site Functional Test Portal https://sam-fzk.gridka.de/sft/lastreport.cgi

3. Distributed Resources and Partners

Another example for a broader experience in grid technology is a set of different approaches to fabric management (like e.g. Quattor and cfengine) at the sites. All German ROC partners are using the PPS (pre-production system) to test new gLite releases for their local environment and to offer the VOs an opportunity to test their software carefully in combination with new middleware before a release goes to production. Similarly to the preproduction in the German Swiss federation, the regional CPU resources are distributed too: about 40% are contributed by the partners of the GridKa Tier 1 (see Figure 3).

Figure 3. Normalized CPU time, all VOs for one year (July 2006 – July 2007) in region Germany/Switzerland per resource centre [2].

There's also a variety of Linux flavours used in production: 11 out of 16 sites are using Scientific Linux, but Debian, CENT OS and SUSE play a well-established role in the regional grid structure. Communication and coordination of activities in the German Swiss region need special emphasis to gain common experience in a steadily growing grid computing infra structure.

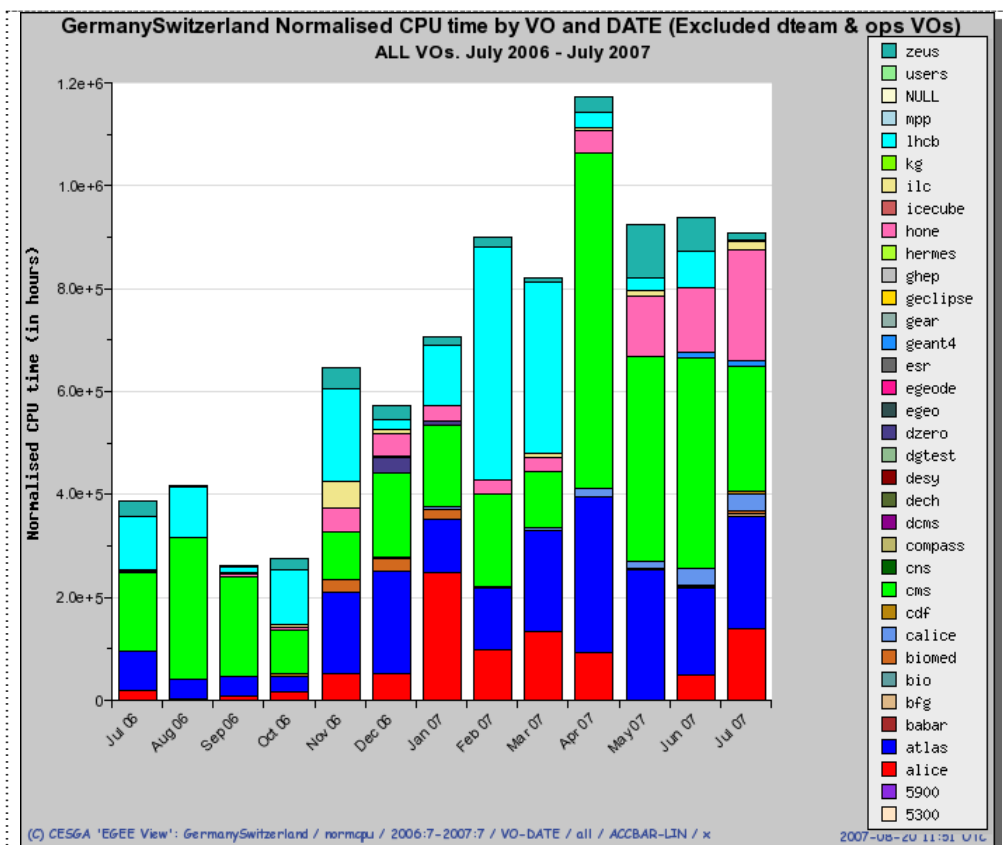## 4. Large Spectrum of VOs in production

Figure 1. Normalized CPU time for one year (July 2006 – July 2007) in region Germany/Switzerland per VO [2].

With its distributed structure, the German Swiss ROC broadens grid knowledge through a larger user community and comprises more application and science communities, than it would in a central approach. About 30 regional and global virtual organizations (VO) use the German Swiss production grid infrastructure for their computations. Of course the LHC VOs span the largest fraction (with about 80% of CPU usage). In addition there are many local user groups either from HEP or from very different scientific areas, which will once be important for a common regional grid structure. Collaboration with Swiss Grid and D-Grid has started also in terms of new regional VOs who start using gLite as part of their middleware setup.

An important task is to facilitate new scientific user groups to test grid technology for their needs prior to becoming a full member of the grid community. As a regional contribution to this task the dedicated DECH VO is in place and is broadly supported in the German Swiss region: The DECH VO is supported by nearly all sites of the federation (DESY, SCAI, GSI, FZK, universities of Dortmund, Aachen, Freiburg, Wuppertal, Karlsruhe, …). The VO's purpose is to give users outside of established VOs the possibility to use the gLite infrastructure for deployment of new applications. In addition the

VO is usable for working with a grid environment (test case for students) and for providing test capabilities for the German Swiss federation. Some examples for applications yet tested in this VO are
- ROOT deployment (C++ API)
- Detector simulations based on Geant4, Geant3, VMC
- Porting of a Bio "Unicore" Workflow to EGEE
- Gilda Portal for BIO Workflow (student work)

Another important use case of the VO has been the successful organisation of admin and user middleware courses during GridKa School 2007 [3] where the VO has been extensively used.

5. Conclusion

A distributed approach for a regional operations centre has led to a wide acceptance of the regional SA1 management in the German Swiss EGEE grid community. A decentral approach increases the organisational and managerial effort however, grid knowledge and expertise is naturally spread among partners with various backgrounds. Temporary manpower shortage e.g during local vacation times can be compensated for much easier resulting in a more reliable customer support. The most important advantage of the distributed ROC and the broad coordination structure for grid operations in the German Swiss region is an exchange of experiences and a more even distribution of know-how in a steadily growing grid computing infrastructure. For the future of the distributed German Swiss ROC it will be important to improve the presentation of the ROC in the World Wide Web. This is needed to make information easily accessible to site administrators and users. Furthermore existing tools and portals should be unified and consolidated. The situation becomes more complex since the national German grid initiative D-Grid has taken up the task to integrate and support dozens of communities and projects. These include demand for a large variety of grid technologies that need to be combined with the EGEE infrastructure in the near future. Already today, several sites including GridKa have started to support applications running on top of gLite, Globus or Unicore using one and the same compute hardware Following a phase of optimizing grid services and establishing on-duty teams to increase the reliability of the organizational structure, manually offered services through the on-duty teams will have to be aggregated or automated to save costs and improve response times.

References

[1] GGUS portal http://www.ggus.org

[2] *CESGA EGEE Accounting Portal*
http://www3.egee.cesga.es/gridsite/accounting/CESGA/egee_view.php?

[3] *GridKa School 2007* http://gks07.fzk.de/