Knowledge Management and Semantics in Global Grid User Support

T. Antoni[1], W. Bühler[1], H. Dres[1], G. Grein[1],
T. Kamps[2], R. Kupsch[1], M. Roth[1], R.Stenzel[2]

(1) Institute for Scientific Computing, Forschungszentrum Karlsruhe GmbH,
Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

(2) Fraunhofer Institut für Grafische Datenverarbeitung
Fraunhoferstraße 5, 64283 Darmstadt, Germany

Abstract

The organisation and management of the user support in a global e-science computing infrastructure such as EGEE (Enabling Grids for E-sciencE), a series of EU projects, is one of the challenges of the Grid. Given the widely distributed nature of the organisation, and the spread of expertise for installing, configuring, managing and trouble-shooting the Grid middleware services, a standard centralised model could not be deployed in EGEE. This paper presents the model used in EGEE for building a reliable infrastructure for user, virtual organisation and operations support. A short overview of EGEE is given. The model for supporting a production quality infrastructure for scientific applications will be described in detail. The advantages of the chosen model will be presented and the possible difficulties will be discussed. In this paper we will also describe a scheme of how knowledge management can be used in Grid user support and first steps towards a realisation in the framework of the EGEE user support infrastructure.

1 Introduction

Providing adequate user support in a Grid environment is a very challenging task due to the distributed nature of the Grid. The variety of users and the variety of Virtual Organisations (VO) with a wide range of applications in use add further to the challenge. Wherever a user is located and whatever the problem experienced, a user expects certain levels of service. With the Global Grid User Support (GGUS) infrastructure [1], EGEE [2] attempts to meet these expectations.

2 The EGEE project

The Enabling Grids for E-sciencE project brings together scientists and engineers from more than 100 institutions in 32 countries world-wide to provide a seamless Grid infrastructure for e-Science that is available to scientists 24 hours a day. Conceived from the start as a four-year project, the second two-year phase started in April 2006, and is funded by the European Commission.

Expanding from originally two scientific fields, high energy physics and life sciences, EGEE now integrates applications from many other scientific fields, ranging from geology to computational chemistry. Generally, the

EGEE Grid infrastructure is ideal for any scientific research especially where the time and resources needed for running the applications are considered impractical when using traditional IT infrastructures.

The EGEE Grid consists of over 20,000 CPU available to users 24 hours a day, 7 days a week, in addition to about 5 Petabytes (5 million Gigabytes) of storage, and maintains 20,000 concurrent jobs on average. Having such resources available changes the way scientific research takes place. The end use depends on the users' needs: large storage capacity, the bandwidth that the infrastructure provides, or the sheer computing power available.

3   The EGEE user support model

The model according to which the user support infrastructure in EGEE is built is connected to the federative model of the EGEE project. To build a successful support infrastructure it was important not to substitute things that were already in place but to integrate them into the project wide structure. Therefore technically, the most important part of the user support system is the central helpdesk application that brings together all the already existing helpdesks and tracking systems existing throughout the project. Through an interface to this system all other helpdesks can communicate. A schematic view of this is shown in Figure 1.
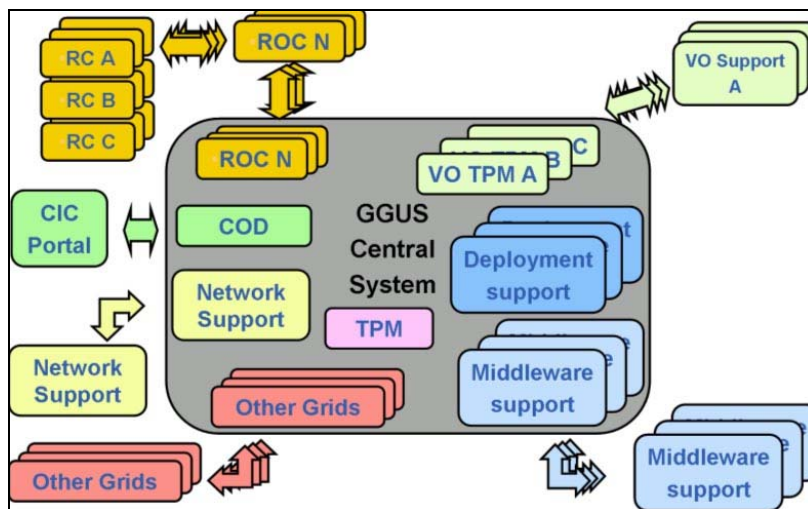

Figure 1: The GGUS central application
and its interfaces to external support systems

All project wide support units are represented in the central system although their main work might be done outside of the system. The interfaces between these systems make sure that in the central system the state of all the problems is always synchronised with the local helpdesk where work is being done on the problems.

4   User interaction with the support infrastructure

A user has two main possibilities to get in contact with the EGEE user support infrastructure. The first is to contact the central helpdesk, called

Global Grid User Support (GGUS) with his problem. He can do this via a web portal or by e-mail (Figure 2). Through the central application his request will then be routed to the responsible support unit. This might mean a transfer to a different helpdesk system for the ticket but this is transparent to the user. He can always track the progress of the ticket through the central system.
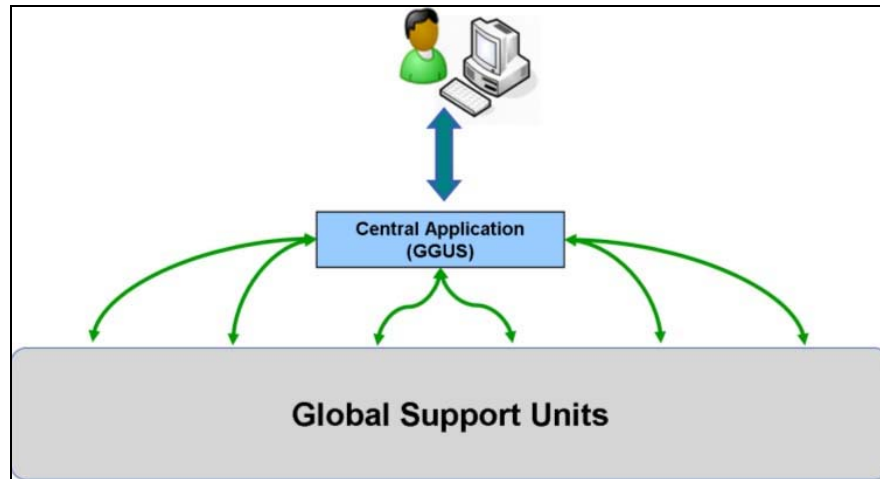

Figure 2: Using the central helpdesk to submit a request

The second possibility for the user is to contact a regional helpdesk. Users who are familiar with the regional infrastructure that has been already in place do not need to change their habits. The regional first line support will triage the problem and in case it is a local problem assign it to the responsible regional support unit. Should the problem be out of the scope of the regional support infrastructure or be of general interest, it will be transferred to the central helpdesk where the process described in the last paragraph will then be used for handling the further progress of the ticket.
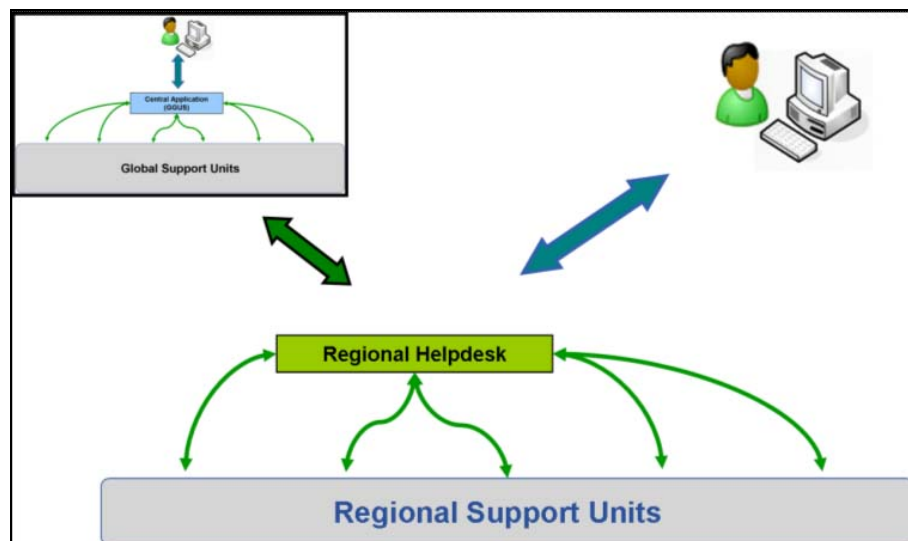

Figure 3: Using a regional helpdesk to submit a request

5 Internal process for handling requests

This paragraph describes, once a request has entered the central helpdesk, which steps are taken to solve the problem. The problem is first triaged by a first line support group which is called TPM (Ticket Process Managers). This group consists of Grid experts from the different federations in EGEE and work according to a weekly rota. For some Virtual Organisations (VOs) there also exists a group of VO TPM who together with the generic TPM categorise the problems into VO specific and non VO specific problems. This is crucial as with the rapid increase of VOs in EGEE and therefore also an increasing variety of VO specific software it will become more and more difficult to have a knowledgeable first line support with resources only coming from the operations part of the project.

Behind this first line support there are two major categories of support units. On one side support units are dealing with services necessary for a proper working of the Grid. For generic Grid services these are mainly located at the regional operations centres or the resource centres connected to the ROCs (Regional Operations Centres). For VO specific software these services are within the responsibility of the respective VOs.

The second category consists of software support units, dealing with all of the middleware components as well as with other tools used to work on the Grid like for example monitoring. These tools are not part of the basic middleware but nevertheless indispensable for a functional infrastructure.

In addition to these two main groups of support units there are a few other special groups that do not fit into this scheme of services and software support units, e.g. network operations and Grid monitoring. Figure 4 shows a schematic view of the process described in this paragraph.
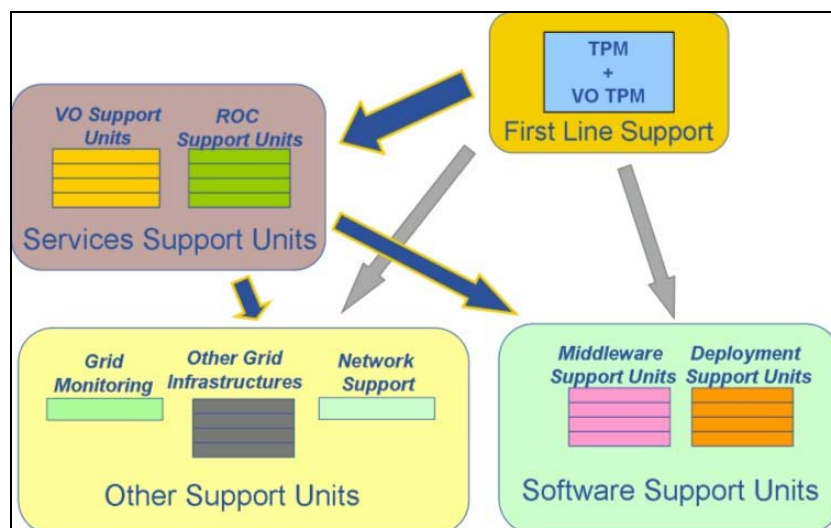

Figure 4: Support units in GGUS and main workflows

The whole infrastructure as described above is in place and operational, but of course there are tasks in this activity that are not completely finished or will always be ongoing.

Regarding the technical side of the system, the usability can still be improved to make it easier for the supporters to work with the system. Additionally it will be necessary to adapt the system to changes in the structure of the project and new requirements, e.g. new middleware components will call for new support units dealing with problems related to these components.

Defining the processes and procedures behind the system and agreeing on these with the people who make up the support units is the much more challenging task in this activity. There are several key areas in the overall process that are vital for the functioning of the complete support infrastructure. It is important that the first line support is in a position to successfully analyse the incoming problems, so that they can either be solved immediately or routed to the correct support unit. To achieve this it is necessary for the VOs to be involved in the first line support, if not being solely responsible for the triage of problems coming from their user communities. Otherwise the system will not be scaleable to a large number of VOs. Another issue is to ensure the responsiveness of the support units. There are several things that can be done to improve this. Of course this issue is connected to the usability of the system; if supporters have to spend less time on a ticket they will be more willing to work on tickets and will be able to do so more efficiently.

6   The GGUS system

The GGUS system consists of the following components:

- Web application based on a LAMP (Linux/Apache/MySQL/PHP) environment, using SSL (Secure Socket Layer)

- Remedy Action Request System (ARS), a development environment for automating Service Management business processes

- Oracle 9i database

On the GGUS web portal there are pages which are freely accessible and others which require a digital certificate or a GGUS login account. The Oracle Call-Interface (OCI) performs the communication between the web front-end and the ticket database. Remedy AR System controls the entire workflow and the read/write communication with Oracle. Remedy AR System also provides Web Services, which allow for easy data exchange with many of the interfaced ROCs and support units. A more detailed description of the GGUS system can be found in [3] and [5].

7   Using knowledge management and semantics

To help users to get quick answers to their questions and to aid the supporters in solving problems and providing correct solutions to the users, the GGUS helpdesk will be using knowledge management techniques and a semantic search mechanism. The aim is to automatically provide the supporter with possible solutions to user problems by analysing the help requests semantically and look for similar problems in the various data sources that are available, e.g. the knowledge base of known problems and their solutions, documentation, Wikis and websites.

Additionally there will be a search interface that can be used to directly search for a solution to a problem without submitting a help request. A possible layout of this can be seen in Figure 7.
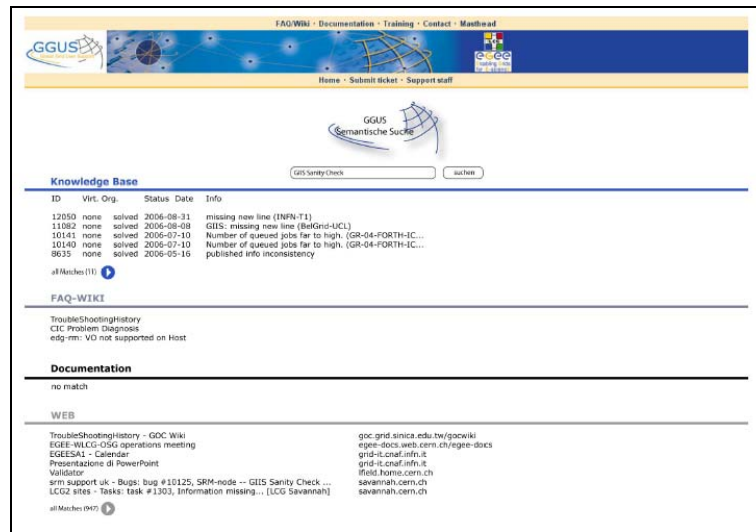

Figure 7: Possible layout of the result display of a semantic search over various data sources in the GGUS portal

The process of setting up this as a semantic knowledge base starts with analysing and integrating the existing material (help requests and solutions, documentation, websites) with semantic methods to identify the relevant vocabulary, phrases and their relations. This is done automatically by using linguistic and statistical methods. The additional information retrieved in this manner together with the original data then constitutes the semantic knowledge base, which has to be kept up-to-date continuously (see Figure 8). It can then be used to achieve much more precise results in a search or to compare new problem reports with already known problems and give hints at a solution.
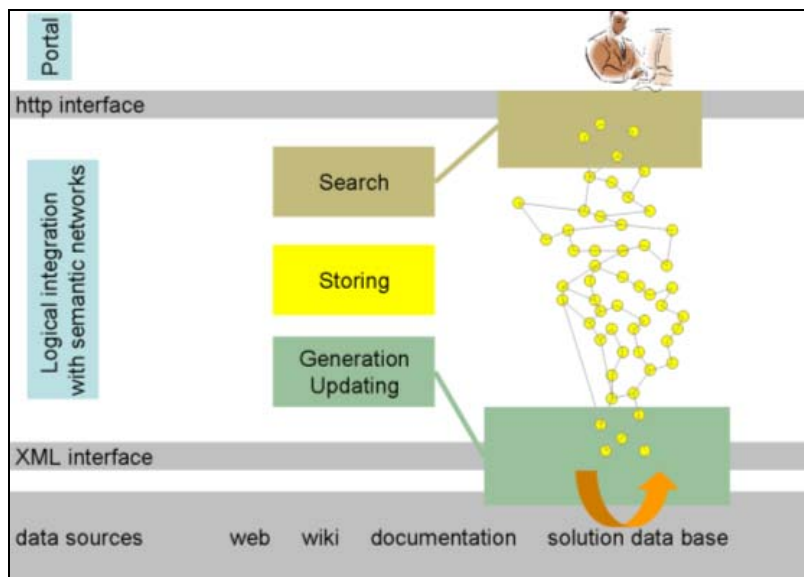

Figure 8: Schematic view of a semantic knowledge base

Next we will describe in more detail the methods used to semantically analyse the relevant data for setting up a knowledge base for Grid related problems in the EGEE context.

7.1 Challenges

As mentioned above the information concerning technical problems resides in various technical environments such as the knowledge base of known problems, the wikis, documents on the fileserver as well as relevant websites. This poses a problem to the user's demand to satisfy his information need quickly and easily. In case he is searching for a solution on a particular problem he might be forced to do so in different repositories. This inconvenience is even growing with the increasing number of documents fed into the repositories over time. Another problem the user might encounter is related to the terminology used to indicate particular problems and/or solutions. As the documents are created by different authors in different contexts they may refer to the same issue by giving it different names. A third issue worth being mentioned here is related to the fact that solutions for particular problems might be solutions to other problems as well. A difficulty for the user is, however, to match these problems, especially when the population of documents is large.

Figure 9: User confronted with different interfaces
to different data repositories

7.2 Approach

Considering the challenges discussed above a major step towards a convenient search function would be to provide a semantic integration of the different document repositories. This requires, first, an extraction of the relevant terminology, and second, a networking of the terminology by means of semantic relations. Among those are taxonomic relations (is broader term of) but also similarity relations such as synonyms and associations. By networking the terminology in this way it is possible to achieve a knowledge network that acts as a semantic index structure for the overall search space. It thus integrates all the repositories and at the same time allows for flexibility in the formulation of queries. As documents may be elements of the knowledge network themselves the generated similarity relations among documents may be used to find similar solutions. A key requirement for the approach is, however, that the knowledge network is generated automatically out of the data stored in the repositories. At the same time the network should be updated automatically as the document space is continuously growing. Another important aspect is that search function itself is easy to use.
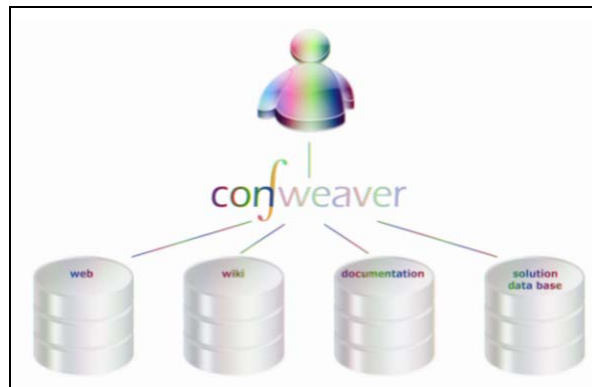
Figure 10: User with one single point of access based on
automatically generated knowledge network

7.3   Technical Solution

The technical solution for the above mentioned approach is based on the
ConWeaver workbench [4] that has been developed at Fraunhofer Institut für
Graphische Datenverarbeitung in Darmstadt. The basic rationale behind the
technology is to automate the construction of semantic index structures by
adapting the technology to the data. How can this be achieved? Data in dif-
ferent repositories may be distinguished with respect to different dimensions.
They may vary in structure (text vs. database structure), model (database
structure A vs. database structure B) data format (pdf, rtf, etc.), size (Mega
Byte vs Giga Byte vs Tera Byte), business application and other aspects.
Also, the data obviously vary over the corporate environments. This suggests
that the technology must vary from one corporate setting to the next one, too.
Essentially, this means that the data analysis methods applied to detect ter-
minology, relations, classifications, etc. must be adapted and configured in
the way they compute optimum results. To this purpose ConWeaver allows
for the construction of data analysis workflows that are designed with re-
spect to the knowledge model (objects: such as concepts, products, persons,
projects, customers, competitors; relations "concept is broader term of sub-
concept", "person is project manager of project" according to which its ob-
jects need to be instantiated. Once the workflow is designed it populates the
knowledge network automatically by taking the data of the repositories as its
input and transforming them into a semantic net using the analysis methods
defined in the workflow. An important aspect concerning the storage of the
knowledge network is that it can be saved and searched using standard rela-
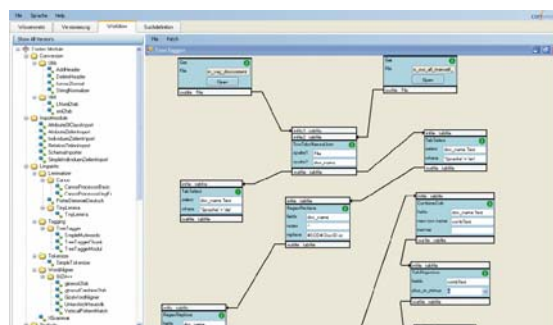tional databases.

Figure 11: ConWeaver workbench - data analysis workflow

A prototype of a semantic search and an instant help mechanism for supporters will be included into the existing GGUS portal in the near future.

8   Conclusions

We believe that the GGUS system is now ready for duty. It is expected that there will be a larger number of tickets passing through the system as the LHC VOs move towards production at full scale. It is also expected that the number of Virtual Organisations will grow as the work of EGEE-II proceeds. There will also be an increase in the number of support units involved with GGUS and an increase in the number of RCs. Adapting the support infrastructure to this growing environment will be the challenge of the near future.

The current model foresees only basic interfaces to the existing VO user support infrastructures. A move towards a further integration of the existing structures is needed for a better overall user support. Hence, the plan is to understand in detail how VOs provide support to their users so that GGUS can better cater for their needs.

The scalability of GGUS is constrained by the availability of supporters. This will rapidly become a constraint to growth unless more dedicated supporters are found.

The use of knowledge management techniques will be a major improvement of the GGUS system and will help make it even more useful and easy to handle for user as well as supporters.

Acknowledgements

References

1. https://www.ggus.org
2. http://www.eu-egee.org/
3. T. Antoni, F. Donno, H. Dres, G. Grein, G. Mathieu, A. Mills, D. Spence, P. Strange, M. Tsai, M. Verlato: Global Grid User Support, The Model and Experience in LHC Computing Grid. Proceedings of Computing in High Energy and Nuclear Physics (CHEP06), Mumbai, India, February 2006
4. http://www.conweaver.de
5. Unterstützung der Grid-Benutzer: Global Grid User Support (GGUS) - weltweit rund um die Uhr. Nachrichten - Forschungszentrum Karlsruhe, 36(2004) S.148-152
6. T. Antoni, W. Bühler, H. Dres, G. Grein, R. Kupsch, M. Roth: Global Grid User Support - Building a worldwide distributed user support infrastructure. German e-Science Conference, Baden-Baden, May 2007