Global Grid User Support - Building a worldwide distributed user support infrastructure

T Antoni, W Bühler, H Dres, G Grein and M Roth

Forschungszentrum Karlsruhe, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

torsten.antoni@kit.edu

Abstract. The organization and management of the user support in a global e-science computing infrastructure such as EGEE 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 troubleshooting the grid middleware services, a standard centralized 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. 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. We will describe the ongoing efforts to build a worldwide grid user support infrastructure in the framework of WLCG by achieving interoperability between the EGEE and OSG user support systems. 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 Organizations (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, EGEE attempts to meet these expectations.

2. The EGEE project

The Enabling Grids for E-sciencE project (EGEE) [1] brings together scientists and engineers from more than 240 institutions in 45 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 on 1 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 36,000 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 30,000 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 structure 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.

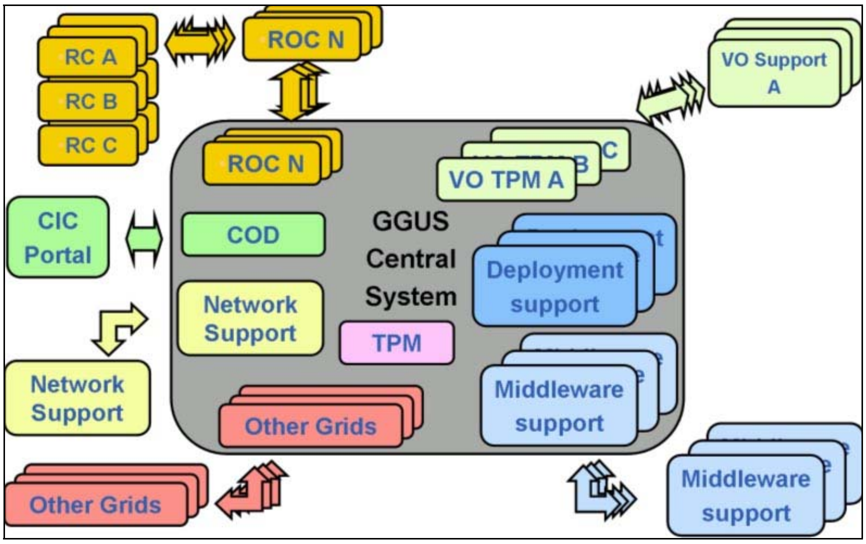

Figure 1. The GGUS central application and its interfaces to external support systems

All project wide support units are present 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.

Additionally the system has the possibility to connect to other grid infrastructures' support systems. This is very important since most of the users of grids do not only use one but several grid infrastructures to do their computing. Such an interface is currently in place between EGEE and Open Science Grid (OSG) [2], who are both heavily used for offline computing and storage of LHC [3] data and simulations in the framework of the WLCG project [4].

The management of the overall support infrastructure and the development and operations of the central helpdesk are done at the Institut für wissenschaftliches Rechnen (IWR) [5] of Forschungszentrum Karlsruhe [6].

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 he can contact the central helpdesk, called Global Grid User Support (GGUS) [7] 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.

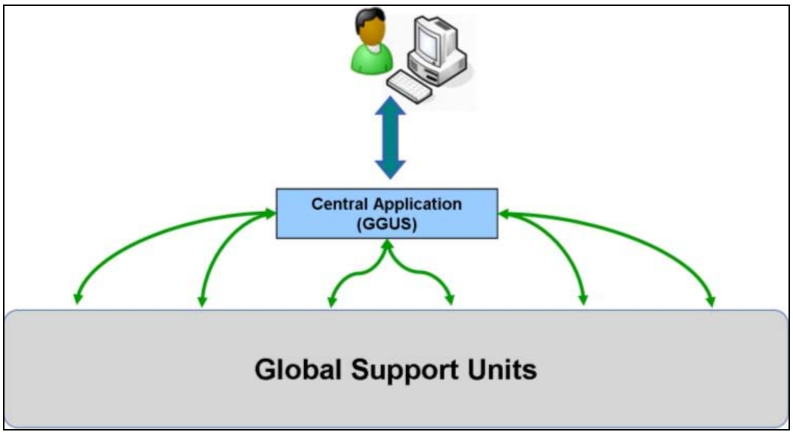

Figure 2. Using the central helpdesk to submit a request

The second possibility for the user is to contact a regional helpdesk or his VO's support infrastructure. Users who are used to the regional infrastructure that was 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 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. The same applies for a VO user filing his request with the VO internal user support. See Figure 3.

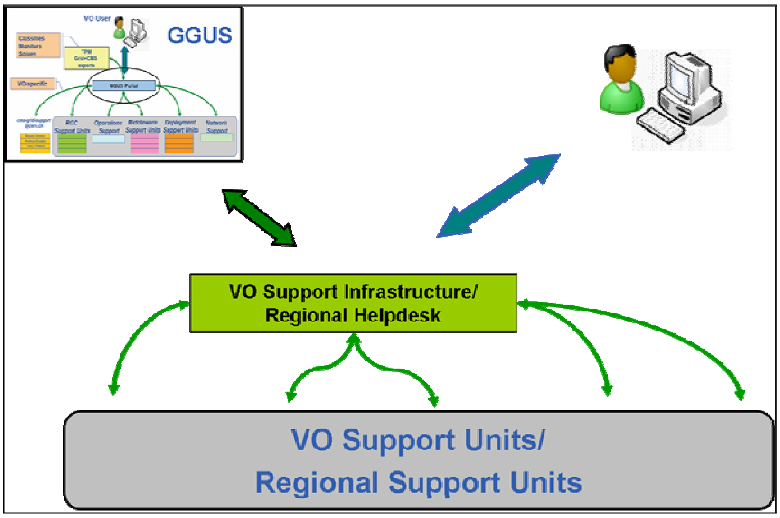

Figure 3. Using a regional or VO helpdesk to submit a request

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 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 works according to a weekly rota. For some Virtual Organisations (VOs) there also exists a group of

VO TPM, who together with the generic TMP 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 just resources coming from the operations part of the project.

Behind this first line support are two major categories of support units. On the one side there are support units dealing with services necessary for a proper working of the grid. For generic grid services these are mainly located at the Regional Operations Centres (ROC) or the resource centres connected to the ROCs. 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.

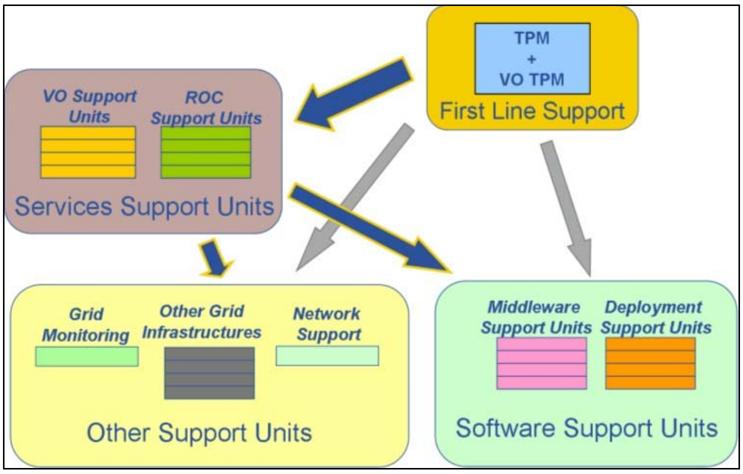

Figure 4. Support units in GGUS and main workflows

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 agree 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. More information on the user and operations support processes can be found in [8], [9], [10] and [11].

6. The GGUS system

6.1. Technical description

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

In 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 Action Request controls the entire workflow and the read/write communication with Oracle. Remedy also provides Web Services, which allow for easy data exchange with many of the interfaced ROCs and support units.

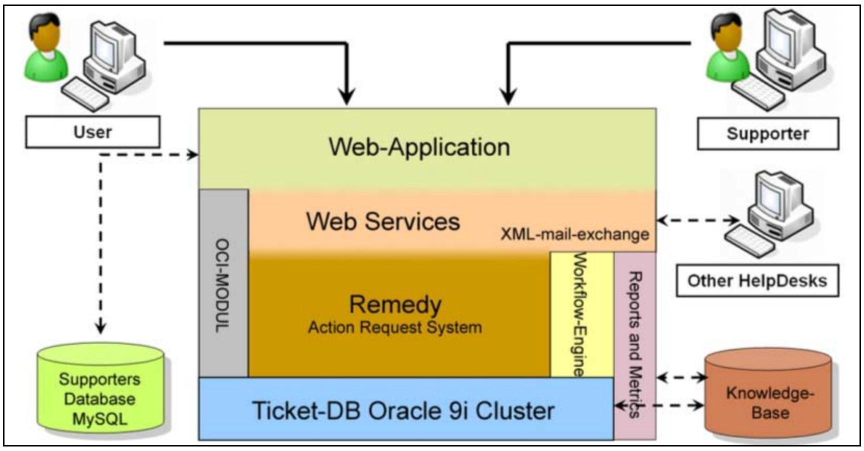

Figure 5. GGUS application model

6.2. The Backup Concept

As the GGUS central application is supposed to run on a 24/7 basis and the complexity of the built-in components may lead to system errors or outages, there is a need to backup the major parts of the system.

Two identical installations, a production and a backup system exist. They work independently from each other and have a separate connection to the redundant Oracle ticket database cluster.

Currently various concepts are being evaluated to guarantee the immediate switch from production to backup machine in case of failure.

Also administrative work such as update installations can then be done on the second machine while the first one is in charge of the production role. Hence, the duration of downtimes can be minimized.

6.3. Remedy Action Request System / Web Services

Finally the Remedy Action Request System Server provides the structure and the workflow for the ticket system. The mid-tier, as part of the ARS Server, gives the interface to the GGUS Web portal. It handles the web service calls as shown in Figure 6.

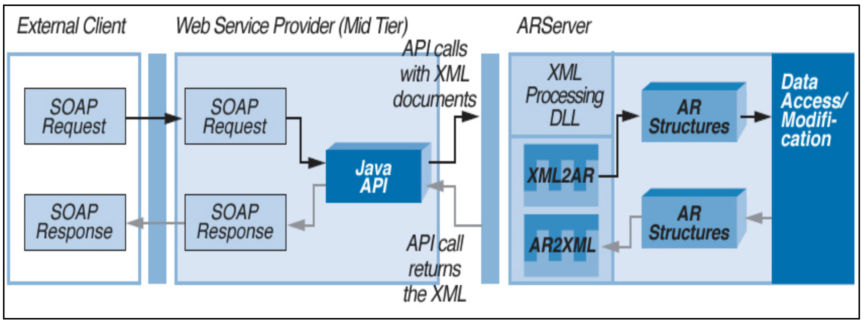

Figure 6. External Web Service call to Remedy ARS

Web services enable AR System functionality to be available over the Web (publishing), and enable the AR System applications to access third-party web services.

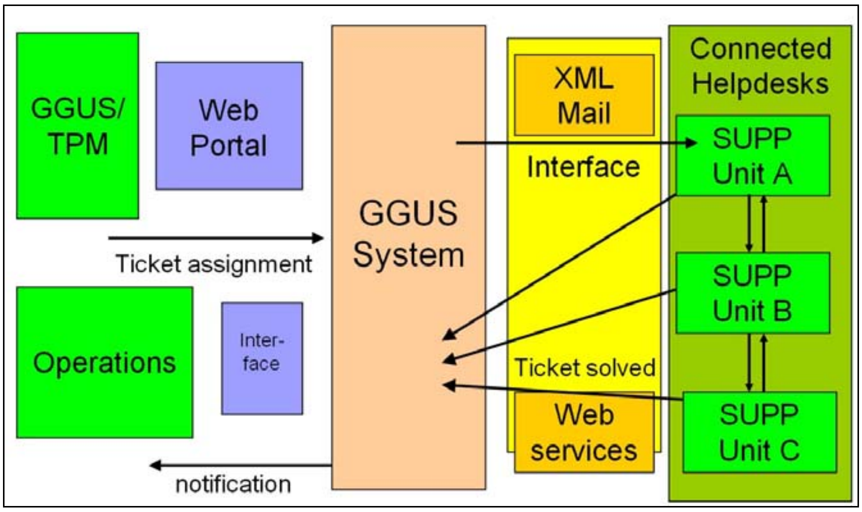

Figure 7. Schematic view on data exchange

Further description of the GGUS system can be found in [8], [9], [10] and [11].

6.4. The web portal

The web portal has been designed as the main entry point for users and supporters dealing with EGEE specific requests. An information section offers a wide range of useful documentation and access to ticket metric reports, various wikis and FAQs. It is freely accessible for everyone without any restriction. Other sections of the web portal require an authentication dependent on the audience. Basically there is a distinction between two groups, which are users and supporters.

After successful authentication (certificate or login/password) users are allowed to create and submit GGUS tickets, they can keep track of them and subscribe to other tickets which might be of interest. Special rights are required to access the supporters' interface, amongst other data the specification of the DN-String of their digital certificate is needed to finish a successful registration/application as GGUS supporter. Together with a powerful search engine the support interface offers all features which are necessary for the ticket handling process.

6.5. Recent developments in GGUS

Recently the following main changes in the web portal have been implemented:

In the supporters web interface three different mechanisms are provided to create a ticket relation. The first is called master/slave relation which means that a master ticket has to be defined first and then more tickets (slaves) can be linked to this ticket. Once the master is solved all the slaves will be solved automatically too. If you have marked a ticket as slave you cannot modify it before the master is solved. The vice versa relation is called parent/child relation: The parent ticket is solved at the time when all its children are solved. Finally there is the cross reference feature which allows referring amongst all the tickets in the system.

New ticket states have been introduced which enable a more meaningful classification. Tickets are classified by open states (assigned, in progress, new, on hold, reopened, waiting for reply) and terminal states (solved, unsolved, verified). New states are "on hold", used for requests whose solution date cannot be determined at the moment and "verified" as status for requests whose solutions have been checked and agreed on by the user.

GGUS is providing special reports about ticket statistics/metrics which have been requested by various groups as a basis for their escalation management.

In the download section of the GGUS portal reports for EMT, the ROCs, TPM and tickets submitted during weekends are being made available. The update intervals have been agreed with the concerned support groups, mostly on a weekly basis.

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 8.

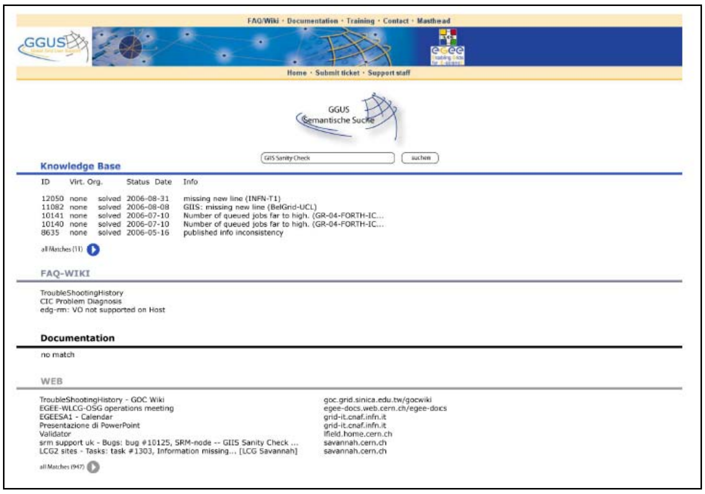

Figure 8. Possible layout of the result display of a semantic search over various data sources in the GGUS portal

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 9). 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.

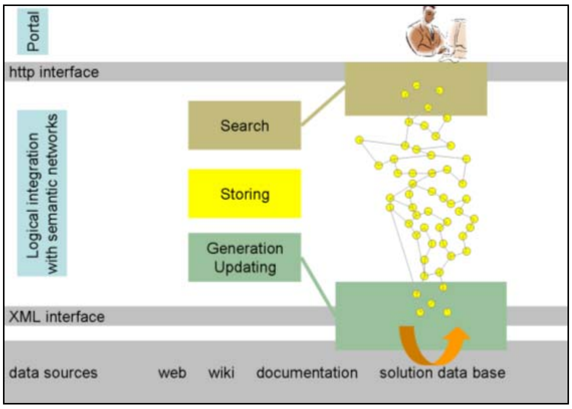

Figure 9. Schematic view of a semantic knowledge base

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, wikis, documents on the fileserver as well as relevant websites. This poses a problem to the user's demand to satisfy his information needs 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.

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 (e.g.: "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.

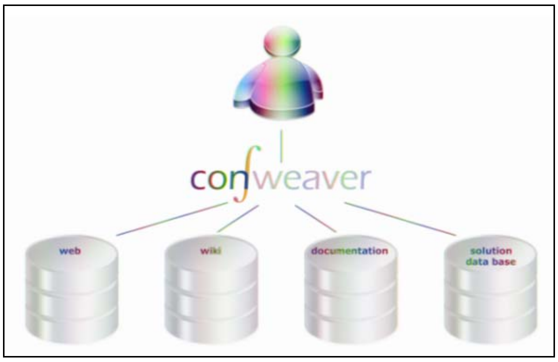

Figure 10. User with one single point of access based on automatically generated knowledge network

7.3. Technical Solution

The technical solution for the above mentioned approach is based on the ConWeaver workbench [12] that has been developed at Fraunhofer Institut für Graphische Datenverarbeitung [13] 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 different 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), application and other aspects. Also, the data obviously vary over the project environments. This suggests that the technology must vary from one setting to the next one, too. Essentially, this means that the data analysis methods applied to detect terminology, 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 respect to the knowledge model (objects: such as concepts, components, persons, projects, customers; relations “concept is broader term of sub-concept”, “person is project manager of project” according to which its objects 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 relational databases.

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.

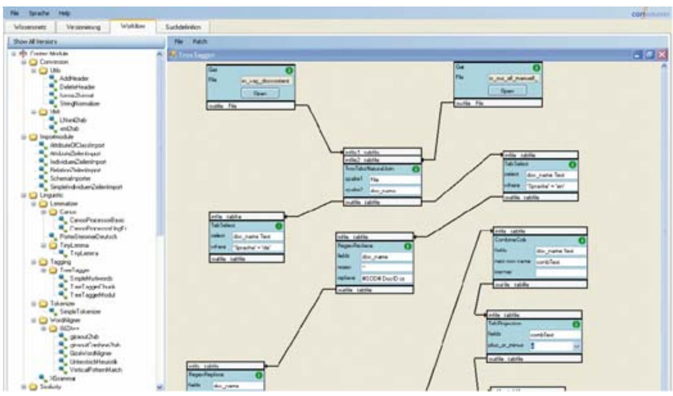

Figure 11. ConWeaver workbench - data analysis workflow

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

We would like to acknowledge the work and thank all the people, teams, activities and institutions who have contributed to the success of GGUS. This work has been funded by the EGEE projects.

References

[1] http://www.eu-egee.org
[2] http://www.opensciencegrid.org
[3] http://lhc.web.cern.ch/lhc
[4] http://lcg.web.cern.ch/LCG
[5] http://www.ggus.org
[6] http://www.fzk.de
[7] http://www.fzk.de/iwr
[8] Unterstützung der Grid-Benutzer: Global Grid User Support (GGUS) - weltweit rund um die Uhr. Nachrichten - Forschungszentrum Karlsruhe, 36(2004) S.148-152
[9] 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 Ex-perience in LHC Computing Grid. Proceedings of Computing in High Energy and Nuclear Physics (CHEP06), Mumbai, India, February 2006
[10] 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
[11] T. Antoni, W. Bühler, H. Dres, G. Grein, T. Kamps, R. Kupsch, M. Roth, R. Stenzel: Knowledge Management and Semantics in Global Grid User Support. German e-Science Conference, Baden-Baden, May 2007
[12] http://www.conweaver.de
[13] http://www.igd.fhg.de