Proceedings of the Third Conference of the EELA Project
R. Gavela, B. Marechal, R. Barbera et al. (Eds.)
CIEMAT 2007

Grid Monitoring in the EUChinaGrid Infrastructure

Lanxin Ma

European Organization for Nuclear Research (CERN), Geneva - Switzerland
Institute of High Energy Physics (IHEP), Beijing - China
Lanxin.Ma@cern.ch

Abstract

The EUChinaGrid project is funded in the 6th framework program of the European Union. The project aims to support the interconnection and interoperability of Grids between Europe and China. To provide stable grid services, improve the reliability of the grid infrastructure and provide stakeholders with views of the resources allowing to understand the current and historical status of the service packages like SAM, GridICE and GStat are used to monitor the EUChinaGrid infrastructure. In this paper, we discuss the use of these tools which are able to check if a given grid service works as expected for a given user or set of users on the different resources available on a grid.

This paper also aims to evaluate the currently existing grid monitoring status in the EUChinaGrid infrastructure and provides thoughts about what could be improved in the future to increase the quality and quantity of monitoring information.

1. Introduction

The EUChinaGrid project is funded by the EU in its 6th framework program. The project aims to extend European grids to China and supports the interconnection and interoperability of the existing European/EGEE[1] and Chinese/CNGrid[2] sites. Until now, the EUChinaGrid infrastructure contains 12 sites based on gLite[3] in China and in 3 European countries (Greece, Italy and Poland). According to GStat, there are around 2000 CPUs and 70 TB of storage space in the EUChinaGrid infrastructure supporting many applications, such as HEP (LHC experiments: CMS, ATLAS), Astrophysics (ARGO-YBJ), biology, etc. So it is important to provide stable grid services, improve the reliability of the grid infrastructure and provide stakeholders with views of the infrastructure allowing to understand the current and historical status of the services. For this purpose, SAM (Service Availability Monitoring), GridICE and GStat are used to monitor the EUChinaGrid infrastructure.

CNGrid stands for China National Grid with its middleware named GOS (Grid Operating System) based on web services. CNGrid contains 10 sites in China and their GOS is still totally apart from the gLite infrastructure. The interoperability between gLite and GOS is being developed and has not been put into production yet. So, the EUChinaGrid monitoring covers at the moment only the sites based on gLite.

2. Global Grid Monitoring: the EGEE/WLCG Case

Due to the situation described above only gLite based monitoring tools are discussed here.

gLite was born from the collaborative efforts of more than 80 people in 12 different academic and industrial research centers as part of the EGEE project, gLite provides a bleeding-edge, best-of-breed framework for building grid applications, tapping into the power of distributed computing and storage resources across the Internet. The gLite software distribution is an integrated set of components designed to enable resource sharing. In other words, this is middleware for building a grid. gLite middleware is currently deployed on hundreds of sites as part of the EGEE project and enables global science in a number of disciplines, notably serving the WLCG(Worldwide LHC Computing Grid) project.

2.1 SAM (Service Availability Monitoring)

SAM[4] is a grid services monitoring tool and provides site independent, centralized and uniform monitoring for all grid services. It is the main source of monitoring information for high-level grid operations and is being used in the validation of sites and services via calculating their availability metrics.

SAM has a server/client structure. The architecture of SAM is shown in Figure 1, which displays its components logically divided into three independent layers: input, data storage and processing and output.

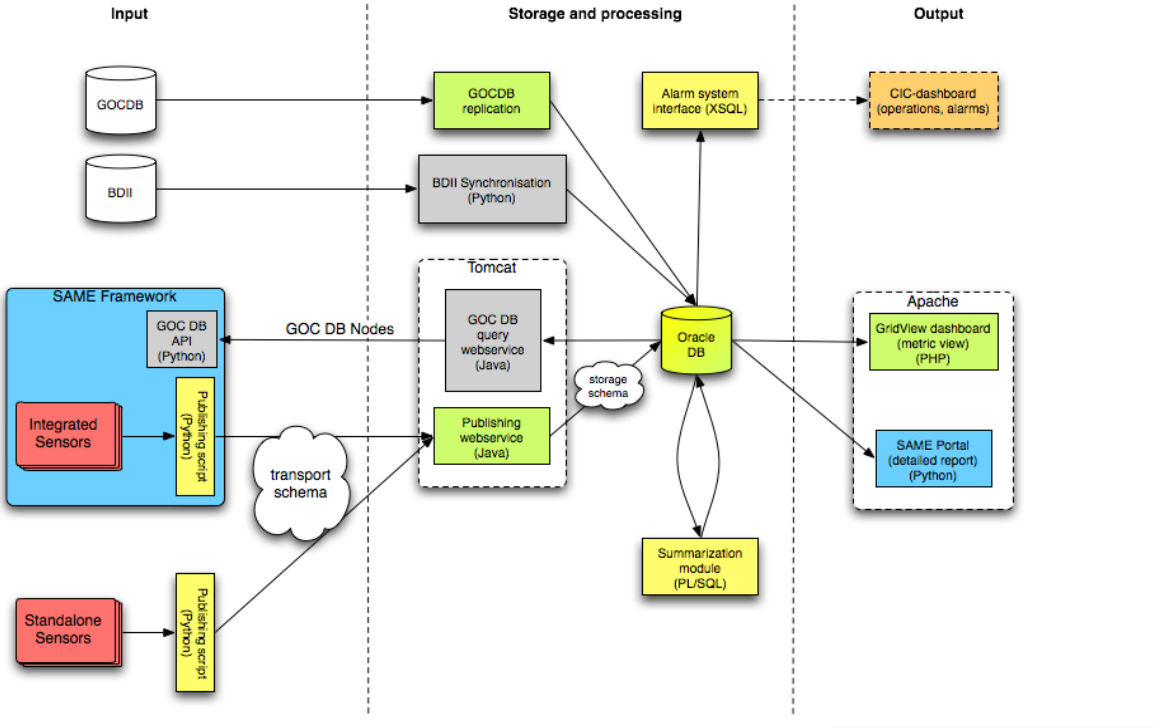

Figure1: SAM architecture

Input

The input part mostly consists of components responsible for synchronizing data from top BDII and GOCDB tables to the SAM server, executing regular tests against all grid services and delivering/publishing results. There are two possibilities: either tests are executed and re-

sults published by the default submission component or the equivalent functionality is provided by standalone monitoring tools that publish the results into SAM in a predefined way[5].

This framework provides a uniform platform for executing the tests and publishing test results to the central database. It has command line utilities to perform queries and to publish test results to the underlying database through web services.

All the sensors in SAM are plug-in modules that communicate with the framework using a fixed protocol. In the design of SAM two levels of hierarchy were introduced: sensors as containers and tests as individual code units (executables) which usually produce a single result record.

Storage and processing

The components of SAM which are responsible for collecting monitoring data, storage and post-processing are installed on a central machine called SAM Server. The relational database is the heart of the server and contains all the data: grid infrastructure description (sites, nodes, services, VOs and relations between them), test results, test criticality, availability metrics and application configuration.

Output

The presentation layer of SAM contains a number of components that are accessing the SAM database directly or indirectly (through XML data exports) and may even be part of external systems. The following are the three most important components. The SAM Portal is a reporting tool written in Python that displays individual test results by VO, service type, and region, as an HTML table with possibility of showing history of test results and the detailed log of test execution. The GridView visualization portal shows configurable availability plots (intervals, VO-wise, site-wise, etc). The COD Dashboard is an external portal for Grid operations in EGEE/WLCG which is a front end to the alarm and ticketing systems.

2.2 GridICE

GridICE[6] is a distributed monitoring tool designed for Grid systems. It monitors the grid resources available at each site, and across the grid as a whole. Its design is based on requirements given by different types of users, each of them dealing with a different abstraction level of a grid: the Virtual Organization level, the Grid Operation Center level, the Site Administration level and the End-User level. It also gives statistics according to historical traces, supports a customized graph generation, gives run-time alerts and has a powerful and complete web-based interface for data presentation.

The architecture[7] of GridICE is structured in 5 layers, from initial producers of monitoring data to final consumers of these data: Measurement Service, Publisher Service, Data Collector Service, Detection/Notification and Data Analyzer Services and Presentation Service. A GridICE server queries a set of nodes to collect data about system information, such as CPU load, CPU type and memory size etc.; Grid service information, such as CE ID, queued jobs etc.; network information such as packet loss; and job usage such as CPU time, Wall time etc.

2.3 GStat

GStat[8] is a centralized monitoring service, which monitors the worldwide Grid Information System. Its primary goal is to detect faults, verify the validity of and display useful data from the Information System.

GStat periodically queries and collects information from the EGEE information system. The retrieved information is processed by an analysis framework that checks failures and

errors, generates web-accessible reports that present aggregated detailed views and history plots. GStat is a useful monitoring tool for ROCs and grid site admins for site troubleshooting.

3. EUChinaGrid Monitoring

3.1 Monitoring Tools in the EUChinaGrid Infrastructure

As most of EUChinaGrid sites are EGEE production sites and all sites run gLite, SAM, GridICE and GStat are adopted as the main grid monitoring tools in the EUChinaGrid infrastructure.

SAM is used to monitor grid services. The EUChinaGrid SAM[9] server is located in IHEP in Beijing. SAM periodically submits test jobs to each node to test the grid services and publishes the test results. The EUChinaGrid SAM monitors different VOs, such as the euchina VO, dteam VO etc. The EUChinaGrid SAM portal shows the service status of each node for each site. Figure 2 is a screen shot of the EUChinaGrid SAM portal. We can follow the test log to track and trouble shoot sites which show an error status.

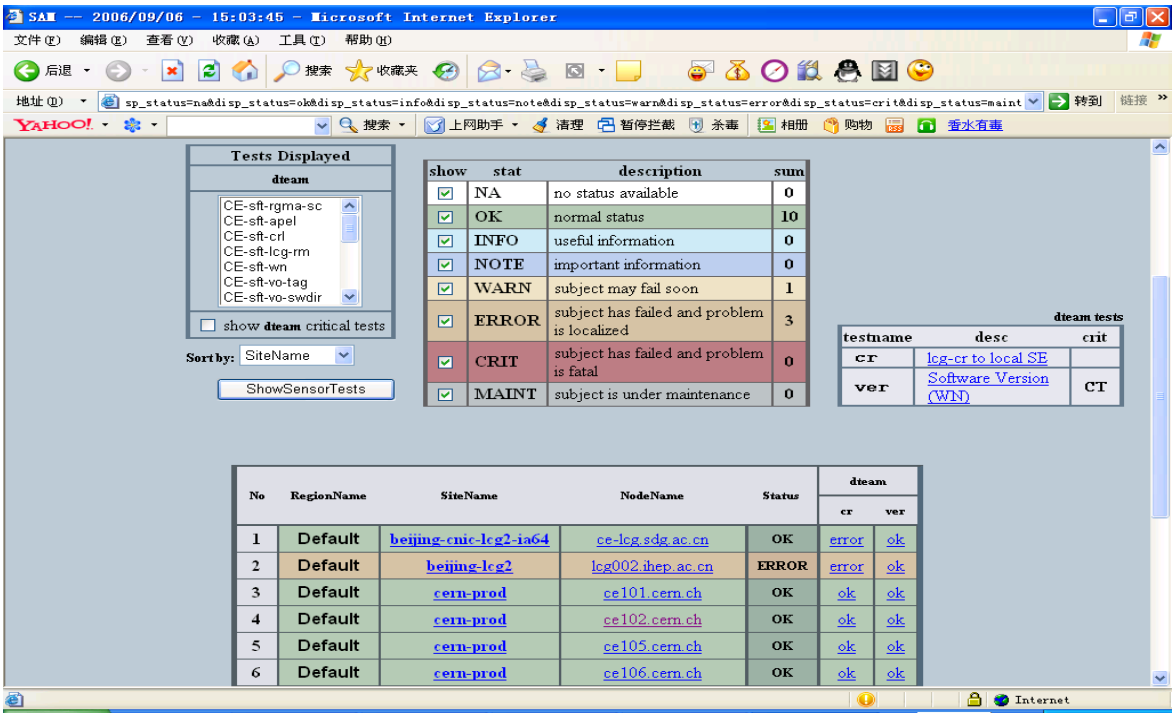

Figure 2: EUChinaGrid SAM Portal

The EUChinaGrid GridICE[10] server is managed at the INFN computing centre CNAF in Bologna, Italy. It shows hosts, jobs and information on computing resources. The GridICE portal shows country info based on the GridICE detection mechanism, administrative info from GOCDB, such as the name of the ROC the sites belong to, the number of CEs, SEs, WNn, CPUs, running and waiting jobs and the site job load as a measure how busy the site is.

GStat provides information about a site, grid services, software and VOs supported at that site. The EUChinaGrid GStat monitoring[11] uses servers running in CNAF and at ASGC in Taipei.

3.2 Grid Operations in EUChinaGrid

The EUChinaGrid infrastructure is based on EGEE/WLCG grid services provided by 12 sites in China and 3 countries in Europe. There are 58 users in the euchina vo until now. The EUChinaGrid ROC (Regional Operation Center)[12] follows EGEE/WLCG operational procedures to support the EUChinaGrid infrastructure and is responsible for the availability and performance of the grid services of EUChinaGrid. The duty of the EUChinaGrid ROC is to certify new sites, to monitor all the sites, to assure the quality of the services, to report existing problems to relevant sites and track the problem status using the EUChinaGrid ticket system, to help sites to solve their problems and to finally validate the solution. The EUChinaGrid ROC on duty is performed by members of EUChinaGrid in CERN and in INFN-CNAF using the monitoring tools discussed above: SAM, GridICE and GStat.

If a new site wants to join EUChinaGrid, the site has to pass the SAM tests to assure that the site and the grid services are configured correctly. After passing the tests, the site can accept jobs.

The EUChinaGrid ROC on duty uses SAM to detect grid service problems. If the status of a service shows an error in SAM, then a ticket is sent to the site admins of the relevant site. Sometimes we have to use SAM and GridICE at the same time to judge if a service really runs wrongly. For example, when a queue in CE is in “Draining” status, then the relevant job submission test is failing. This can only be investigated in GridICE as SAM cannot tell us the queue status.

SAM monitors the status of the grid services. GridICE provides some fabric info and queue, job and resource status etc. GStat monitors a lot of things, but GStat does not give the status of the services or the interoperation of the different services.

3.3 EUChinaGrid Ticket System and Problem Tracking Procedure

The EUChinaGrid ticket system[13] is run by the Italian Academic and Research Network (GARR), in Rome, Italy, based on the Xoops[14] system which is an easy to use dynamic web content management system written in PHP. Figure 3 shows a screen shot of the EUChinaGrid ticket system portal.

The EUChinaGrid ticket system organizes ticket distribution in different departments. There are the SAM department, VOMS department, helpdesk department, site department, etc. We can submit, view and update tickets and send notification to service admins via the ticket system. It is easy to track the problem. The EUChinaGrid ticket system can also work as a ROC support tool. We can schedule ROC shifts and log shift notes via the system.

The problem tracking procedure consists of creating tickets in the EUChinaGrid ticket system by the EUChinaGrid ROC whenever any problem related with sites is detected by monitoring tools and the site admins receive a notification about the problem. Site admins try to solve the problem and update the tickets when the problem status is changed. If a site admin cannot solve a problem, (s)he can ask for help from the ROC or others till the problem is solved. The ROC will finally validate the solution and close the ticket after the problem is solved.

3.4 SmokePing in the EUChinaGrid infrastructure

SmokePing[15] is a deluxe latency measurement tool which is not developed by EGEE. It can measure, store and display latency, latency distribution and packet loss. SmokePing uses RRDtool[16] to maintain a longterm data-store and to draw pretty graphs, giving up to the minute information on the state of each network connection.

One object of the EUChinaGrid is to study IPv6 on the available and foreseen network connectivity to promote new high bandwidth links between Europe and China or Asia in gen-

eral and to study the available Grid Middleware for an IPv6 network and the interaction between Grid Services and IPv4-IPv6 communication.

For the above purpose, SmokePing is used to measure network availability and connection for IPv4/IPv6 test bed and general network in the EUChinaGrid infrastructure. SmokePing server is managed by Roma-3 university in Rome, Italy. The EUChinaGrid SmokePing portal[17] shows the RTT info, the rate of average package loss etc. between the nodes.

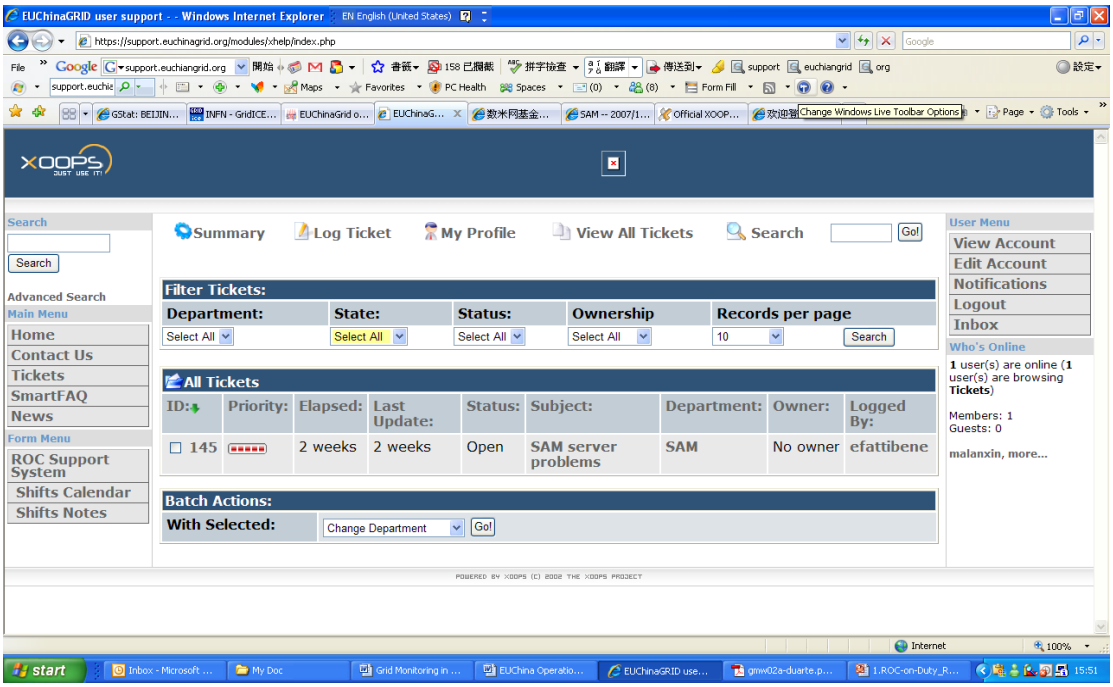

Figure 3: EUChinaGrid Ticket System Portal

4. Other Monitoring Tools

Nagios[18] is a host and service monitor designed to inform you of network problems before your clients, end-users or managers do. It can monitor network services (SMTP, POP3, HTTP, NNTP, PING, etc.), host resources (processor load, disk and memory usage, running processes, log files, etc.) and environmental factors such as temperature. Nagios can notify predefined contacts when service or host problems occur and get resolved (via email, pager, or other user-defined methods). Optional escalation of host and service notifications to different contact groups is also available. Nagios has a web interface for viewing current network status, notification and problem history, log file, and so on. Nagios has strong functionality, providing views of fabric information and network information between sites or within a site.

Ganglia[19] is a scalable distributed monitoring system for high-performance computing systems such as clusters and grids. It allows the user to remotely view live or historical statistics (such as CPU load averages or network utilization) for all machines that being monitored.

5. Conclusion

SAM, GridCE and GStat are very useful grid monitoring tools. Until now, several critical problems have been detected by those tools in the EUChinaGrid infrastructure and were solved in time so that the quality of the grid services could be assured. A good example is that the LHC experiments applications, WISDOM and many other applications run successfully in the EUChinaGrid infrastructure.

In the author's personal view, the grid services, jobs and queues and some fabric information etc. have been monitored well by SAM, GridICE and GStat and SmokePing tell us if the network works properly between the nodes. It may be better if we could extend some features of either SmokePing, Nagios or Ganglia, for example, that could monitor ports of services. This may make it easier for site admins to diagnose the grid problem.

References

[1] EGEE *http://public.eu-egee.org/*

[2] CNGrid *http://www.cngrid.org/web/guest/home*

[3] EGEE gLite *http://glite.web.cern.ch/glite/*

[4] SAM *http://goc.grid.sinica.edu.tw/gocwiki/Service_Availability_Monitoring*

[5] A. Duarte, P. Nyczyk, A. Retico and D. Vicinanza, "Global Grid Monitoring: The EGEE/WLCG Case", HPDC Meeting in Monterey Bay in June 2007

[6] GridICE *http://gridice.forge.cnaf.infn.it/*

[7] S. Andreozzi, N. De Bortoli, S. Fantinel, A. Ghiselli, G.L. Rubini, G. Tortone and M.C. Vistoli, "GridICE: a Monitoring Service for Grid Systems", In Future Generation Computer Systems Journal, Elsevier, 21(4):559-571, 2005.

[8] GStat *http://goc.grid.sinica.edu.tw/gstat/*

[9] EUChinaGrid SAM *https://euchina-sam.ihep.ac.cn/sam/sam.py*

[10] EUChinaGrid GridICE *http://euchina-gridice.cnaf.infn.it:50080/gridice/site/site.php*

[11] EUChinaGrid GStat *http://goc.grid.sinica.edu.tw/gstat/euchina/*

[12] EUChinaGrid Grid ROC *http://roc.euchinagrid.org/home.php*

[13] EUChinaGrid ticket system *https://support.euchinagrid.org/*

[14] Xoops *http://www.xoops.org*

[15] SmokePing *http://oss.oetiker.ch/smokeping/index.en.html*

[16] RRD tool *http://oss.oetiker.ch/rrdtool/*

[17] EUChinaGrid SmokePing *http://netsmoke.roma3.infn.it/cgi-bin/smokeping.cgi*

[18] Nagios *http://www.nagios.org/*

[19] Ganglia *http://www.ganglia.info/*