

Monitoring and Accounting for GridNNN Project

S.D. Belov, V.V. Korenkov, S.V. Mitsyn, D.A. Oleynik, A.S. Petrosyan, R.N. Semenov, A.V. Uzhinskiy

Laboratory of Information Technologies, JINR, Dubna

Abstract. Contemporary distributed systems like computational grids are complex technical systems, therefore to keep an eye on its state and to count the consumption of computational resources special automated tools should be used. In this paper we discuss the experience of development of monitoring and accounting system for the project GridNNN [1, 2], which aim is to provide grid infrastructure for National Nanotechnology Network in Russia. The grid middleware used in GridNNN project is apartially based on well known packages like Globus Toolkit 4 [3] and VOMS [4], and to fit the needs of the specific application area several grid services were developed from scratch. In such conditions, special monitoring and accounting system was created within the project.

The monitoring is a rather general concept. Most common tasks we deal with are:

- Continuous watching for the state of grid services both common for all infrastructure and in a particular Resource Center;
- Obtaining information on resources (slots number, operation system, hardware architecture, special software packages) and their utilization;
- Access control rules for the resources by Virtual Organizations and groups inside them;
- Execution monitoring, tasks and jobs submission, state changes and return codes;
- Resource usage information (especially CPU consumption);
- Watching the quotas for resource usage by Virtual Organizations.

For effective control, planning and faults detection it is important to know not only a current state of the grid infrastructure but also to keep track of the state's history.

Introduction

The aim of GridNNN project is a creation and support for national nanotechnology network of Russia. The main goal of the project is to provide an effective access to the distributed computational, informational and networking facilities for nanotechnology science and industry. The base middleware of particular GridNNN services (like MDS, GRAM) is Globus Toolkit 4. Some services are developed by the project team (e.g. job handling tool Pilot [5], Information Index [6] based on Globus MDS, GRAM connection with non-standard Local Resource Management System, Web User Interface, etc.).

Operation of GridNNN infrastructure, unlike many huge grid projects, is more centralized in

a sense: there are about 15-30 resource centers (supercomputers) controlled by two operation centers, the main and the backup one (having the same set of central services). The infrastructure has one central information index where all information providers have to publish their data. Another difference from most grid projects is a great variety of different Virtual Organizations (VOs) and their software environments.

Monitoring subsystem overview

Typical computational jobs in GridNNN are parallel and use MPI technology. They demand a huge volume of computation, but do require storage or transfer of considerable amount of data. Therefore monitoring activities in this area are primarily concentrated on tasks and jobs tracking (computational jobs are parts of a task and could be interconnected with each other). Jobs monitoring is naturally associated with billing (or accounting) features: it is important to know who, when and where uses the project's resources.

To choose the resource center to submit jobs to, the end user should have a possibility to know main characteristics of the environment there. The most important of them are supported VOs, hardware architecture, number of total available and free slots for running jobs, operation system version, special software packages and so on.

Then, the next main point is an overall information on the state of the infrastructure. Now a simple test to check whether the infrastructure services are alive, was prepared within the monitoring subsystem. Along with it was created geoinformation real-time visualization of systems operation based on Google Earth, which allows seeing all jobs and tasks events on 3D globe in real time [8]. This feature is of impress and often used to make graphic presentations.

Also there are several significant tacks related to monitoring but not included to the monitoring subsystem and supported by other teams in the project. Firstly, these are the so-called RAT-tests (resource availability tests), when sample tasks are periodically submitted to each computational resource in the system. A second type of regular checks is the examination of information published by sites to their local information indexes (and then to the central information index).

Information gathering for monitoring and accounting

All the information of services entry points and sites provided by special service named Service for Registration of Resources and Grid Services (SR-RGS). It is the main place where such data are available and where it originates from. SR-RGS provides data as a reply for simple HTTPS queries.

The information System (including central and local information indexes) contains both slow changing and dynamic information on the resources. Site publishes many clue parameters like job available job slots, system architecture type, OS version, list of special software packages, VO access information this changes not so frequent. But there is a small piece of information that contains changing real-time information on state of job queues and available job slots on the site.

There are two types of Information System central services. The first one is based on Globus MDS 4, and WSRF queries are used to perform a request [6]. The second one is based on the GridNNN's development called Infosys2 and has a RESTful interface. Both types of information indexes have the same information schema (based on extended GLUE 1.3) and should contain and provide identical data gathered from the resource cen-

ters.

Along with obtaining information from the Information Indexes, there are special functional test. For each type of service in the project there is a small simple test just to check if the service is available and is responding to queries specific to it (e.g. Information Index should return non-empty response to a WSRF or HTTPS queries).

All the collected information coming from Information Indexes, SR-RGS and simple service tests is then handled and stored to the monitoring database. The monitoring web interface is used to present both real time and historical information. Data flows in monitoring are shown on the Fig. 1.

Jobs monitoring information and accounting data are taken mainly from Pilot (job execution service). Pilot servers publish special accounting log containing all the events occur with tasks and their jobs (starting from task submission, sending jobs to the particular resources and to the task finishing or termination). Monitoring service is querying for new events every minute, and then parses result came in JSON format [7]. Obtained events information (task, job, user, VO, start and finish time) is linked with the same events which is already in the database, forming the states of tasks and jobs in question. Accounting data (mainly consumed CPU

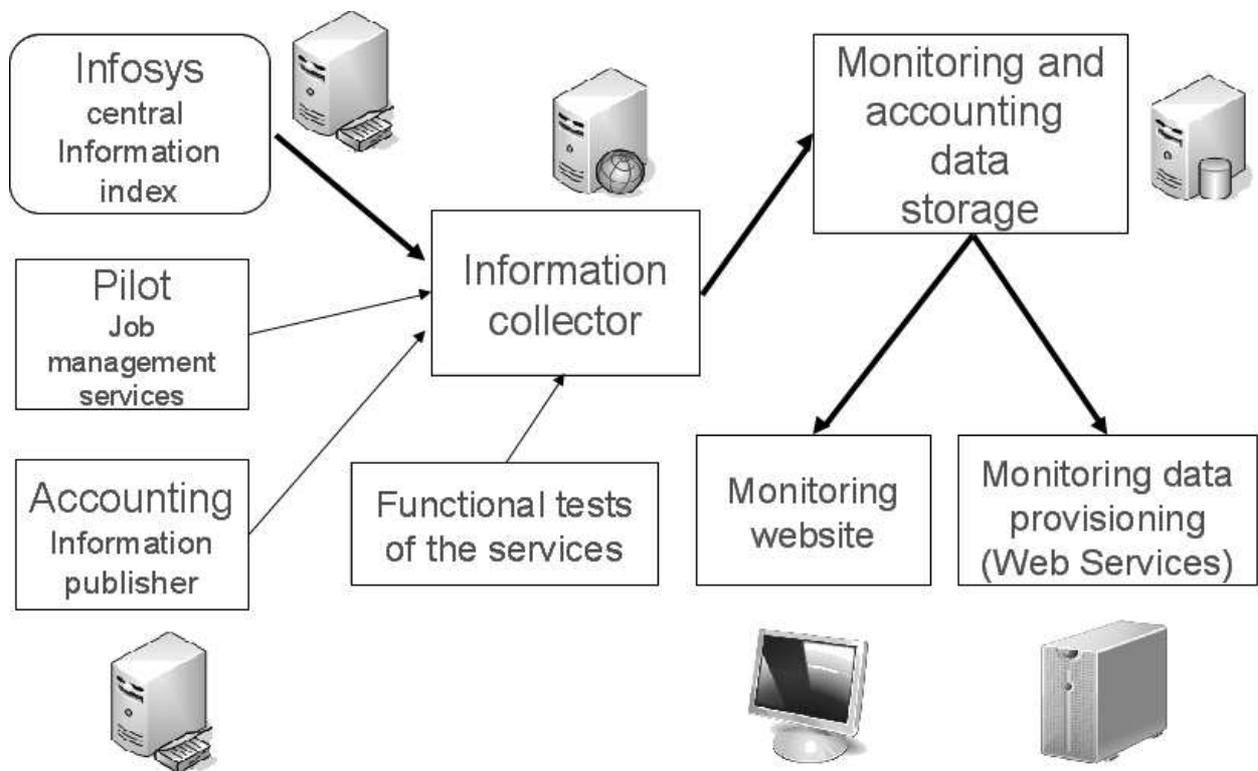


Figure 1: Monitoring and accounting data flow

time) are to be taken from local Grid Resource Allocation Managers (GRAMs) in the resource centers via the special service GACCT. Then the accounting data is to be linked with the job information from Pilot servers in the database. At the end, a full aggregated accounting information is prepared in the database and is available to the end-user via a web-interface.

Data representation for monitoring and accounting

For the representation of collected information, the monitoring service has a web-interface [9]. The main parameters to be displayed on the site are states of computational jobs queues, resource characteristics, operation systems version and so on (see previous paragraph). Accounting information is available on the site as several report views with tables and diagrams by resources, users and VOs.

Real time jobs monitoring allows displaying on 3D globe how and where jobs are started and finished [8]. A special script periodically (each 10 minutes) prepares information on job events based on Accounting DB and makes a KML file to use it in visualization in Google Earth. It is probably the best and the most spectacular way of infrastructures operation demonstration to the wide community.

Summary

GridNNN project has its own features and peculiarities, and it differs from other grid projects. Therefore for making monitoring and accounting for it, along with common means and approaches some special developments are required to fit the project. Speaking of the monitoring and accounting subsystem, it is ready and available on the web [9].

The project has been already finished; the full-functional grid infrastructure of GridNNN is in production state. The program packages created are

used in newly developed Russian national projects and are still evolving.

References

- [1] V.A. Ilyin et al., Design and Development of Grid-infrastructure for National Nanotechnology Network, Distributed Computing and Grid-Technologies in Science and Education: Proceedings of the 4th Intern. Conf. (Dubna, June 28-July 3, 2010). Dubna, JINR, D-11-2010-140, pp. 357-363, 2010.
- [2] GridNNN project site (in Russian), <http://ngrid.ru>
- [3] I. Foster, Globus Toolkit Version 4: Software for Service-Oriented Systems, IFIP International Conference on Network and Parallel Computing, Springer-Verlag LNCS 3779, pp. 2-13, 2005.
- [4] VOMS project home page, <http://voms.forge.cnaf.infn.it/>
- [5] L. Shamardin et al, GridNNN Job Execution Service: a RESTful Grid Service, Distributed Computing and Grid-Technologies in Science and Education: Proceedings of the 4th Intern. Conf. (Dubna, June 28-July 3, 2010). Dubna, JINR, D-11-2010-140, pp. 215-219, 2010.
- [6] M.M. Stepanova et al, Information System of GridNNN, Distributed Computing and Grid-Technologies in Science and Education: Proceedings of the 4th Intern. Conf. (Dubna, June 28-July 3, 2010). Dubna, JINR, D-11-2010-140, pp. 423-428, 2010.
- [7] JavaScript Object Notation (JSON), <http://json.org>
- [8] S. Mitsyn, S. Belov, Development of Real-time Visualization Service Based on Google Earth for GridNNN Project, Distributed Computing and Grid-Technologies in Science and Education: Proceedings of the 4th Intern. Conf. (Dubna, June 28-July 3, 2010). Dubna, JINR, D-11-2010-140, pp. 47-50, 2010.
- [9] GridNNN monitoring and accounting site, <http://mon.ngrid.ru>