



Contribution ID: 73

Type: **Oral contribution**

GridICE monitoring for the EGEE infrastructure

Thursday 2 March 2006 15:35 (15 minutes)

Grid computing is concerned with the virtualization, integration and management of services and resources in a distributed, heterogeneous environment that supports collections of users and resources across traditional administrative and organizational domains.

One aspect of particular importance is Grid monitoring, that is the activity of measuring significant Grid resource-related parameters in order to analyze usage, behavior and performance of a Grid system. The monitoring activity can also help in the detection of fault situations, contract violations and user-defined events.

In the framework of the EGEE (Enabling Grid for E-science) project, the Grid monitoring system called GridICE has been consolidated and extended in its functionalities in order to meet requirements from three main categories of users: Grid operators, site administrators and Virtual Organization (VO) managers. Besides the specific needs of these categories, GridICE offers a common sensing, collection and presentation framework enabling to share common features, while also offering user-specific needs.

A first common aspect to the different users is the set of measurements to be performed. Typically, there is a wide number of base measurements that are of interest for all parties, while a small number is specific to them. What makes the difference is the aggregation criteria required to present the monitoring information. This aspect is intrinsic to the multidimensional nature of monitoring data. Example of aggregation dimensions identified in GridICE are: the physical dimension referring to geographical location of resources, the Virtual Organization (VO) dimension, the time dimension and the resource identifier dimension.

As an example, considering the entity 'host' and the measure 'number of started processes in down state', the Grid operator can be interested in accessing the sum of the measurement values for all the core machines (e.g., workload manager, computing element, storage element) in the whole infrastructure, while the Virtual Organization manager can be interested in the sum of the measurement values for all the core machines that are authorized to the VO members. Finally, the site administrator can be interested in accessing the sum of the measurement values for all machines part of its site.

Another aspect that is common to all the consumers is being able to start from summary views and to drill down to details. This feature can enable to verify the composition of virtual pools or to sketch the sources of problems.

As regards the distribution of monitoring data, GridICE follows a 2-level hierarchical model: the intra-site level is within the domain of an administrative site and aims at collecting the monitoring data at a single logical repository; the inter-site level is across sites and enables the Grid-wide access to the site repository. The former is typically performed by a fabric monitoring service, while the latter is performed via the Grid Information Service. In this sense, the two levels are totally decoupled and different fabric monitoring services can be adapted to publish monitoring data to GridICE, though the proposed default solution is the CERN Lemon tool.

Considering the sensing activity, GridICE adopts the whole set of measures defined in the GLUE Schema 1.2, further it provides extensions to cover new requirements. The extensions include a more complete host-level characterization, Grid jobs related attributes and summary info for batch systems (e.g., number of total slots, number of worker nodes that are down).

The development activity in the EGEE project has focused on the following aspects: the redesign of the presentation level took into consideration the usability principles and compliance with W3C standards; sensors for measuring parameters related to Grid job have been re-engineered to scale to the number of jobs envisioned by big sites (e.g., LCG Tier 1 centers); new sensors have been written to deal with summary information for computing farms; stability and reliability of both server and sensors.

The deployment activity covers the whole EGEE framework with several server instances supporting the work of different Grid sub-domains (e.g., whole EGEE Grid domain, ROC domain, national domain). Other Grid projects have adopted GridICE for monitoring their resources (e.g., EUMedGrid, EUChinaGRID, EELA).

As regards the user experience, GridICE has proven to be useful to different users in different ways. For instance, Grid operators have summary views for aspects such as information sources status and host status. Site administrators appreciate the job monitoring capability showing the status and computing activity of the jobs accepted in the managed resources. VO managers use GridICE to verify the available resources and their status before to start the submission of a huge number of jobs. Finally, GridICE has been positively adopted in dissemination activities.

While GridICE has reached a good maturity level in the EGEE project, many challenges are still open in the dynamic area of Grid systems. The short term plans are: (1) as regards the discovery process, there is the need to finalize the transition from the MDS-based information service to the gLite service discovery plus publisher services such as R-GMA producers and CEMon; (2) integration with information present in the Grid Operation Center (GOC) database for accessing resource planned downtime and other management information; (3) tailored sensors for the workload management service; (4) sensors for measuring data transfer activities across Grid sites.

References:

Dissemination website: <http://grid.infn.it/gridice>

Publications:

<http://grid.infn.it/gridice/index.php/Research/Publications>

Summary

C. Aiftimiei, S. Andreozzi, G. Cuscela, N. De Bortoli, G. Donvito, S. Fantinel, E. Fattibene, G. Misurelli, A. Pierro, G.L. Rubini,

G.Tortone.

Primary authors: Mr PIERRO, Antonio (INFN-Bari); Mrs AIFTIMIEL, Cristina (INFN-LNL); Mr FATTIBENE, Enrico (INFN-CNAF); Mr TORTONE, Gennaro (INFN-Napoli); Mr DONVITO, Giacinto (INFN-Bari); Mr MISURELLI, Giuseppe (INFN-CNAF); Mr CUSCELA, Guido (INFN-Bari); Ms DE BORTOLI, Natascia (INFN-Napoli); Mr ANDREOZZI, Sergio (INFN-CNAF); Mr FANTINEL, Sergio (INFN-LNL)

Presenter: Mr ANDREOZZI, Sergio (INFN-CNAF)

Session Classification: 2d: VO tools - Portals

Track Classification: VO management - Portals