



Grid-Enabled Remote Instrumentation with Distributed Control and Computation

Thursday 2 March 2006 15:00 (30 minutes)

1 GRIDCC Applications and Requirements

The GRIDCC project [1], sponsored by the European Union under contract number 511381, and launched in September 2004, endeavors to integrate scientific and general-purpose instruments within the Grid. The motivation is to exploit the Grid opportunities for secure, collaborative work of distributed teams and to utilize the Grid's massive memory and computing resources for the storage and processing of data generated by scientific equipment. The GRIDCC project focuses its attention on eight applications, four of which will be fully integrated, tested and deployed on the Grid.

The PowerGrid will support the remote monitoring and control of thousands of small power generators; while the Control and Monitoring of HEP experiments aims to enable remote control and monitoring of the CMS detector at CERN. The (Far) Remote Operation of Accelerator Facility is an application for the full operation of a remote accelerator in Trieste, Italy; and the Grid-based Intrusion Detection System aims to provide detection and trace-back of flow-based DoS attacks using aggregated data collected from multiple routers. The other set of relevant applications includes: meteorology, neurophysiology, handling of device farms for measurements in telecommunications laboratories, and geophysiology [2][5].

The project, by nature, requires the availability of software components that allow for time-bounded and secure interactions, while operating instrumentation in a collaborative environment. In addition to the classical request/response Grid service interaction model, a considerable amount of information needs to be streamed from the instrument back to the user. The time-bounded interactions, dictated either by the instrument sensitivity and the accompanying requirement for careful handling and fast response to extreme conditions, or by the applications themselves, lead to the need for the establishment of SLAs for QoS or other guarantees, with support for compensation and rollback. The idea of collaboration and resource sharing, inherent in the Grid, is also extended and adapted to allow the share of unique instruments among users who are geographically dispersed, and who normally would not have access to such –usually rare and/or expensive – equipment.

2 GRIDCC and gLite

To cater for the diversity of instruments and the critical nature of the equipment being handled, the GRIDCC middleware platform relies on Web Service (WS) technologies, and sustains a Service Level Agreement (SLA) infrastructure, alongside enforcement of Quality of Service (QoS) guarantees. The GRIDCC middleware architecture is fully described in [2].

A number of gLite software components are extremely relevant to the GRIDCC middleware architecture, which is designed to comprise various novel middleware components to complement them. Firstly, we plan to perform job scheduling and bookkeeping via the WMS and specifically the WMProxy, and the LBProxy [2]. We also plan to rely on the Agreement Service for SLA signalling and for triggering

resource-level reservations [2]—this is essential to enforce SLA guarantees. In addition, we plan to test and possibly extend CREAM, as explained in the following Section.

The WSDL interface, exposed by the gLite WMS, streamlines job submission in a number of different scenarios: direct invocation by the Virtual Control Room (VCR) - the GRIDCC portal; direct submission onto preselected CEs via the GRIDCC Workflow Management System (WfMS); and indirectly, utilising the WMS's builtin scheduling capabilities, either as a single submission or part of a workflow [2]. The WfMS and VCR are described in more detail in Section 3.

Data gathered from IEs need to be stored, in MSS services. Consequently, data storage will be delegated to gLite SEs exposing SRM-compliant interfaces. VOMS and proxy-renewal services will be used. For authentication and authorization, it is foreseen to support both X.509 certificates and the Kerberos framework. The latter will be used when low response times are required.

Finally, for QoS performance monitoring, as it is experienced by GRIDCC users and services, we require the integration of service monitoring tools and services providing information about network performance, such as the gLite Network performance Monitoring framework.

3 GRIDCC Middleware

The gap between GRIDCC's requirements and gLite's existing service support, will be filled by a number of GRIDCC solutions, which leverage the existing gLite functionality.

The need for instrument support, necessitated the development of a new grid component, the Instrument Element (IE). The IE's naming and design reflect its similarity to gLite's SE and CE. The IE provides a Grid interface to a physical instrument or set of instruments, and should allow the user to control and access instrument data [2]. To cater for the varied needs of instrumentation, the IE also has local automated management and storage capacity [2].

The desire for QoS and SLA support is provided for by the following Execution Service components. The gLite AS will be extended to establish SLAs with the IE, and the IE will need to enforce such SLAs. To achieve this, the IE conceptual model and schema need to be defined in order to publish information about the instrument-specific properties.

The GRIDCC Workflow Management System (WfMS) provides an interface for users to submit workflows, which can orchestrate WS calls to underlying services [3]. The WfMS may also need to choreograph further steps into workflows, such as the SLA negotiation and logging steps, to facilitate the satisfaction of, possibly complex, QoS demands from the user [3]. It is also responsible for monitoring running workflows and responding to workflow events - such as contacting a user if QoS demands can no longer be satisfied [2].

The Virtual Control Room (VCR), supports a user Grid portal for the underlying services, in particular to: request SLAs from the AS; steer and monitor an IE; and submit workflows to the WfMS [2][3][4]. Additionally, the VCR provides a multi-user collaborative online environment, wherein remote users and support staff, share control of and troubleshoot IEs [2][4].

4 Extending gLite

To fulfill the GRIDCC application requirements, a number of gLite functionality extensions would be useful for successful middleware integration. Firstly, information about IEs needs to be made available by the information services. Secondly, in order to enforce upper-bounded execution times, the reservation of CEs and IEs needs to be supported. To this end, we will extend the AS, by adding CE and IE-specific SLA templates. Reservation needs to be triggered and enforced by elements at the fabric-layer. For this reason, we envisage the addition of a new operation to the WSDL interface exposed by CREAM, allowing the invocation of reservation operations. As mentioned above, GRIDCC, needs for QoS to be enforced at both the single-task and workflow level. The WMS already supports some workflow functionality; however, the WMS can only process workflows involving job execution tasks. We foresee the need to merge the functionality of the GRIDCC WfMS with the gLite WMS, to benefit from the existing WMS capabilities and avoid duplication of work.

References

[1] The GRIDCC Project home page: <http://www.gridcc.org>.

- [2] The GRIDCC Architecture –Architecture of Services for a Grid Enabled Remote Instrument Infrastructure (<http://www.gridcc.org/getfile.php?id=1382>).
- [3] D4.1 Basic Release R1, GRIDCC Project Deliverable GRIDCC-D4.1, May 2005 (https://ulisse.elettra.trieste.it/tutos_gridcc/php/file/file_show.php?id=1418)
- [4] Multipurpose Collaborative Environment, GRIDCC Project Deliverable GRIDCC-D5_2, Sept 2005 (https://ulisse.elettra.trieste.it/tutos_gridcc/php/file/file_show.php?id=1408)
- [5] SPECIFIC TARGETED RESEARCH OR INNOVATION PROJECT –Annex I - “Description of Work”, May 2004 (<http://www.gridcc.org>)
- [6] EGEE Middleware Architecture and planning, EGEE Project, Deliverable EGEE-DJRA1.1-594698-v1.0, Jul 2005 (<https://edms.cern.ch/document/594698/>).

Authors: KOTSOKALIS, Constantinos (GRNET); DICKENS, Luke (Imperial College); FERRARI, Tiziana (INFN)

Presenter: DICKENS, Luke (Imperial College)

Session Classification: 2c: Special type of jobs (MPI, SDJ, interactive jobs, ...) - Information systems

Track Classification: Special type of jobs (MPI, SDJ, interactive jobs, ...) and information systems