The Architecture of the ClusterGrid Brokering System

Péter Stefán, Ferenc Szalai, Gábor Vitéz
NIIF/HUNGARNET
H-1132 Budapest, Victor Hugo 18-22
Phone: +36 1 4503070, E-mail: grid-tech@niif.hu, Web: www.clustergrid.iif.hu

Abstract

The ClusterGrid broker software refers to a software system that intends to give a lightweight middleware alternative to Globus Toolkit. In the ClusterGrid brokering system (referred to as broker) developed and maintained by NIIFI the classic grid functions such as MDS, GridFTP or GRAM are implemented in a special way.

The key idea of the broker is that each message conducted among the different grid role-players, i.e. functions executed on virtual or real machines, is accomplished through atomic communication units, transactions, executed over HTTP/POST. The broker structurally consists of

- the execution component, and
- the submit component.

The submit component contains three major subsystems, the scheduler component, the transaction subsystem, and the entry side GIS.

The scheduler is responsible for bringing the unutilised resources (clusters or supercomputers) and pre-schedule jobs together and dispatching them after the schedule decision. The scheduler is built up in a modular way, i.e. different scheduler modules can be inserted into the core component which communicate with it through XML descriptions. The scheduler modules can be written in any language.

The transaction subsystem helps in registering, maintaining and executing the different transactions. There are four major transaction types: submit, info, remove and getout, representing job removal, job and resource information transfer, job removal and job output transfer respectively.

The entry side GIS along with the local master side GIS component are responsible for propagating the resource and job status information from the execution nodes toward the submitters, based on resource update advertisement messages provided by the local resource manager. Update messages are collected on every submitter and stored on local databases.

The execution component of the grid apart from GIS includes a simple interface to the local scheduler allowing smooth job transfer between the grid layer and the resource layer.

Being in production since 2003 the new broker still faces further development: putting the communication channel onto SOAP, working out service gateways among different grid architectures, getting more services involved are just a couple of the possible directions that will be addressed in the future.