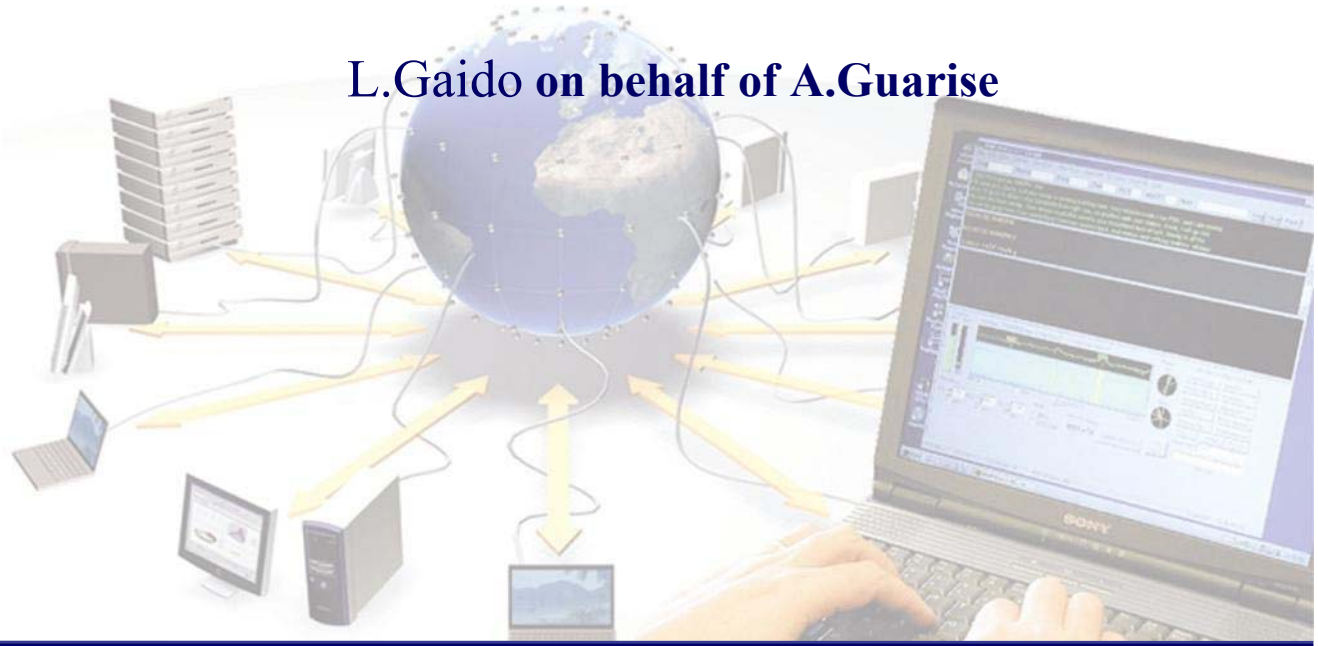
Enabling Grids for E-science in Europe

www.eu-egee.org

LCG Workshop
November 2004, CERN

DGAS Grid accounting

L.Gaido **on behalf of A.Guarise**



EGEE is a project funded by the European Union under contract INFSO-RI-508833

DGAS: Overview

DGAS (DataGrid Accounting System) is a grid infrastructure that implements two accounting approaches:

- Resource Usage Accounting
- Economic Accounting

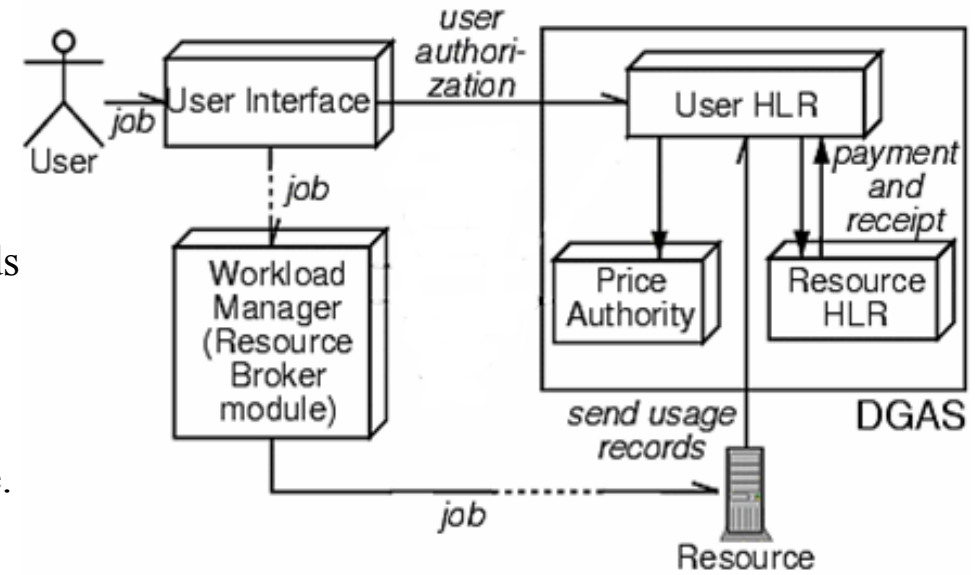
The first one means collecting and aggregating the resource usage records performed by grid users over grid resources.

The second one means assigning a price to grid resources and a cost to the operations performed by users on the grid and charge for them.

DGAS was designed as a a true ‘Grid’ accounting infrastructure. So it aims to be distributed, scalable, secure and fault tolerant and to provide the needed tools to VOs, resource managers and single users.

DGAS: Simplified Workflow

- The user submits the job and concurrently authorises its accounting
- The WMS submits the job to the CE
- A metering system on the CE collects the Usage Records
- The metering system sends the usage records to the accounting server where the user is registered (HLR). The UR are signed with user credentials and encrypted.
- If Economic Accounting is needed a Price Authority is queried to retrieve the CE price. The price can either be set by the resource's administrator or automatically computed using ad-hoc algorithms.
- If Economic Accounting is needed a cost is assigned to the job.
- The usage records are archived on the user's HLR, and on the HLR where the CE is registered.



- The Economic Accounting can be turned on by the resource administrator when needed. In this case a cost will be assigned to the jobs executed on the CE. Therefore the user account will be charged and the resource (CE) one credited for the corresponding amount.
- If Economic Accounting is turned off only the Usage Records for the job will be reported to the Accounting System.

DGAS: the VO's point of view

- The **basic accounting** needs for a VO manager are:
  - Be able to track the grid usage for each user belonging to the Virtual Organization.
  - Be able to track the grid usage for group of users belonging to the Virtual Organization.
  - Be able to get statistics of the grid usage for the whole VO.
- The **Economic Accounting** needs:
  - Be able to assign individual quotas to users.
  - Be able to track grid money transactions.
- Users belonging to a Virtual Organisation need to be registered on a DGAS server as well. With this registration an *account* for the user is created on a DGAS accounting server, the so-called HLR (***Home Location Register***). This can be seen as a virtual bank account for that user on the grid.
  - The user accounts on an HLR can be created by importing the information from an LDAP VO server or manually inserted using command line tools.

DGAS: the VO's point of view

The minimum granularity of the information archived within the DGAS RDMBS is the single grid job. The useful information available for each job is:

- GridJobId.
- DN of the user who submitted the job.
- CEId of the resource where the job was executed.
- the URL of the accounting server where the resource is registered; the usage record is archived for the resource administrator's use.
- Usage Records for the job (currently: CpuTime, WallClockTime, Physical and Virtual used Memory).
- The cost (if any) charged for the job.

The queries that the VO manager can perform on the RDBMS are:

- List of the jobs submitted by a given user over a given period of time.
  - The retrieved information consist in a list of records with the mentioned field.
- Aggregation of the usage records for the job submitted by a single user, a group of users or by the entire VO (over a given period of time).The information available with these type of queries is:
  - Total number of submitted jobs, used cpuTime, wallClockTime, Phisical and Virtual Memory, Total cost charged.

DGAS: the Site Manager's point of view

The **basic accounting** needs for a Site Manager are:

- Be able to track the usage of his/her resources performed by Grid Users.
- Be able to get statistics of the usage of his/her resources or a group of them.

The **Economic Accounting** needs:

- Be able to assign prices to his/her resources.

Also for the Grid resources, there is the need to register the resources on an HLR service.

- The resource accounts on an HLR can be created by importing the information from an LDAP BDII server or manually inserted using command line tools.

DGAS: the Site Manager's point of view

The minimum granularity is that of the single grid job. The available information for each job is:

- GridJobId.
- DN of the user who submitted the job.
- CEId of the resource where the job ran.
- the URL of the accounting server where the resource is registered; the usage record is archived for the resource administrator's use
- Usage Records of this job (currently CpuTime, WallClockTime, Physical and virtual memory used).
- The cost (if any) charged for the job.

The queries that the Site manager can perform on the RDBMS are:

- List of the jobs submitted to a given CE over a given period of time.
- Aggregated usage records for the job submitted to a single CE, a group of CE or the entire set of resources (over a given period of time).
  - The information available with these type of queries are: the total number of submitted jobs, the cpuTime, the wallClockTime, the Phisical and Virtual Memory and the total cost charged in that period.

DGAS: the User's point of view

The **basic accounting** needs for a User are:

- Be able to track the usage of his/her jobs submitted to Grid Resources.
- Be sure of the confidentiality of his/her work.

The **Economic Accounting** needs:

- Be able to check the cost of his/her jobs.
- Be able to check the price of a Grid resource *before submitting* a job.
- Be sure that the usage records charged to him originate from his/her jobs.

Since users will not be granted direct access to the DGAS server where their account are registered, a remote interface is provided. This interface (available on a Grid User Interface) allows them to perform simple queries to the DGAS server:

- Retrieve account information, including the aggregate of Usage Records consumed or the costs already charged.
- Retrieve information for their jobs with detailed Usage Records (cpuTime, WallTime, memory…), the cost charged for the job, the CEId of the CE and the DGAS server where that CE is registered.

DGAS: the User's point of view

For queries concerning information about a single job or a job for a given user it is important to grant a proper level of authorisation:

- A typical user would only be granted access to his/her own accounting information, i.e. he/she will not be able to access other account's information.
- If needed an authorisation mechanism may grant special privileges to some users and allow them to retrieve information regarding other people or resources as well.

Economic Accounting & Brokering

DGAS implements the possibility to “virtually” charge the users for their work on the Grid:

- The cost of a job is a function of the resource usage and the price of the resource.
- Since the grid resources may have different prices it is possible that the same job submitted to different resources results in different costs charged to the user.
- A dedicated service to compute (and archive) prices for the Grid resources has been developed. This service is known as **Price Authority**.

Apart from obvious reason (i.e. effectively charging for the grid usage), the economic accounting may also be used to make a *better use* of the grid by means of algorithms that automatically compute the resource prices; this implies “training” the users (or the broker on their behalf) to choose the cheapest resource.

See http://www.to.infn.it/grid/accounting

DGAS

For more information please contact

Andrea Guarise
andrea.guarise@to.infn.it