Comparison of European Grid Projects

Jarek Nabrzyski, Ariel Oleksiak (PSNC)

Project:

GEMSS

Area:

Monitoring and Performance

Table of Contents

1. Introduction

1.1. Objective and Structure

This document is one of thirteen templates that have common goal to gather information related to main European Grid Projects in order to make their accurate comparison in the framework of GRIDSTART initiative. We believe that the participation of particular projects members in preparation of this document will allow comparing all activities in a credible and exhaustive way.

The proposed structure of the description consists of two parts. The former is concerned with the general overview and architecture together with the contents of layers (the first template). The latter includes the main components of the Grid infrastructure (remaining 12 templates). Since information regarding the project architecture is to be quite general, more detailed description should be provided in the review of the main aspects of the Grid infrastructure. In order to prepare uniform description for each project, we identify the important issues that have to, should or can be included into particular components. Common issues for all components and these specific for this component are briefly described in the next section.

We ask you to proceed according to this schema. However, a feedback is obviously welcome. For some projects the document has been partially completed on the basis of descriptions found at the official web pages. In this case, we ask you to revise already filled in sections, correct and complete them if necessary.

You should take into consideration future plans while you fill in particular sections. Actually they are even more important then the current state of the project components. If you are not going to design some elements in the scope of the project at all, please, note it in the proper section.

1.2. Uniform description

All the descriptions of the Grid infrastructure components are divided into three parts: **General** section includes main requirements and functionality, **Details** section relates to the issues specific for particular component and **External** defines its connections with other components and users.

As it was mentioned above, some of the issues are common for all components or at least repeat for many of them. Such issues, appearing for many or even all areas are shortly characterized below.

In **General** section:

Main requirements determine the objectives and requirements of the workpackage or the software module responsible for the design of functionality related to the particular domain of the Grid infrastructure.

Functionality contains a set of operations provided by the project in the given area.

In **External** section:

Interfaces define services, SDKs, APIs and so forth which can be used in order to access the functionality of the component.

Low level Grid middleware is the middleware providing basic Grid functionality as for example Globus or UNICORE.

Relations with other components determine components that utilize or are utilized by component being described as well as data and information flow between them.

Issues that are specific for this particular domain of the Grid infrastructure are presented in the sequel. Some of them, which we consider to be clear, have been skipped, however, if they turn out to be vague, please, contact the authors of this document (ariel@man.poznan.pl).

The **Details** section contains two tightly related components of grid environment: monitoring of running jobs, state of resources etc. and modeling, measurement and prediction of performance.

Metrics are used in order to express accurately some aspects of program behavior, for example, CPU or memory usage. They can be local (specific for particular machine) or global (comparable between different machines) and this information should be included in **Local** and **Global** paragraphs. **Definition**, in turn, describes exact structure of the metric (i.e., information included into the metric).

Components of monitoring system describe architecture of monitoring system, e.g. its division into local and global monitors.

Instrumentation contains techniques used to gather data from a program while it executes.

2. Monitoring and Performance

2.1. General

The GEMSS project will not develop a general Grid monitoring and performance analysis system.

However, the GEMSS infrastructure will provide basic support for enabling users to monitor jobs executing on remote Grid hosts. The GEMSS design is still discussed by the project and not all issues have been finalized yet.

Moreover, for each medical application service considered within GEMSS, an application-specific performance estimator will be provided which will allow to estimate the time required for executing a user job on a specific Grid site.

- **Main requirements**

  Provide support for the monitoring of jobs (GEMSS applications) executing on remote Grid sites.

  Provide support for estimating the time required for executing a GEMSS application service on a specific Grid site depending on meta-data about a service request provided by the user.

- **Functionality**

  Support for job monitoring and performance estimation will be provided to client-side application components via generic interfaces (WSDL, Java API, …) These interfaces are currently being defined.

2.2.Details

- **Metrics**

  The following metrics will be used to measure the success of service provision, and invoke penalty clauses etc. if appropriate.

  - Service completion time
  - Service throughput
  - Service security aspects
  - Service availability

  Definition

  Service completion time
  The time interval (elapsed time) between the request of a specific Grid service and the completion of the Grid service.

  Service throughput
  The number of service requests that can be completed in a specific amount of time.

  Service security aspects
  Supported aspects of security in the areas of authorization, authentication, data anonymization, data encryption and intrusion detection. For GEMSS it is likely that maximum security will be applied at all times due to the sensitive nature of medical data.

  Service availability
  The percentage of successfully executed services under the specified quality of service constraints.

  Metrics used in the project
  see above.

  Local

  Global

- **Components of monitoring system**
  not applicable

- **Instrumentation techniques**
  not applicable

  Program

  Processor

Operating System

Network

- **Performance specific issues**

  Methods for performance model description
  For each GEMSS application an application-specific set of parameters will be specified which will be fed into an application-specific performance model. These parameters will be specified based on appropriate XML schemas, which are currently being defined

  Measurement

  Estimation and prediction

  Bottleneck analysis

  Application benchmarking

- **Visualization**
  not applicable

  Types of data visualized

  Performance

  Metrics

  Events

  Other

  Visualization techniques

2.3.External

- **Interfaces**
  not yet specified

- **Relations with other components**