A system for automatic installation of libraries in gLite Grids

Monday, 11 April 2011 09:00 (8 hours)

Overview

The gLite infrastructure does not provide tools to automatically install libraries required to run specific applications. Static linking of applications is the standard procedure currently in use. Complex development environments require a manual installation of the requested software and this operation must be done by the local administrators. This problem arises also when different versions of already installed libraries have to be used.

Therefore, the introduction of a tool that automatically provides for missing libraries and software environments is an added value to the gLite platform. This tool operates in a transparent way; users are only requested to declare their needs whereas all aspects related to the system management are delegated to the tool. It is in light of these considerations that we propose our tool.

Impact

The adoption of tools providing automatic installation of requested software, removes any human intervention by local administrators; end users can carry out this task on their own. So access to the Grid infrastructure does not need any mediation.

On the other hand the task the end user must carry out is very simple; users just specify the set of libraries requested for their jobs.

As the repository service is used, it is populated by more and more libraries that can be employed by others too. This enables the collection of statistics concerning the utilization of the various libraries (those used more frequently, the number of times they are installed, etc.), so libraries that are used more often become good candidates to be officially supported and deployed along with the standard WN build.

The adoption of this system allows the porting in the Grid of jobs running on complex environments, as well as of applications not developed by end users. Because end users are not requested to interact with system administrators, they are able to directly and immediately try and test different solutions.

Description of the work

The suggested solution is a VO-centric system to be integrated into the gLite platform, thereby extending its functionality. The implementation of this system requires some minor but important modifications of the CE and WN, still without altering the structure of gLite

components. The modifications regard the pre-execution scripts whose purpose is to set up the runtime environment.

In the following paragraphs the way of working of the system is illustrated, starting from the job submission all the way to the job execution in the WN.

Through the "Environment"attribute of the JDL, two variables are set up:

WN_ENV_SOFTWARE whose value is a colon-separated list of requested software packages;
WN_ENV_REPOSIORY containing the URL of the repository service, where information on requested packages is stored.

When jobs are submitted the values of these variables are not checked by the WMS and the job is dispatched to the CE according to the standard mechanism. To make CEs able to support this tool the pre-execution script has to be modified to take into account the new variables. When a specific WN takes charge of the execution, the values of the two variables are examined to determine if the associated packages are already installed. If not, their installation is tried: the URLs of the missing packages are obtained by invoking the Repository Service specified in the variable WN_ENV_REPOSITORY; with this information the packages are retrieved, downloaded and finally installed. It is worth noting that package installations are on a per-VO basis; jobs belonging to other VOs and running on the specific WN cannot use the newly installed packages.

All VO-specific packages installed in this way can be removed by the local system administrators in case of need (e.g. shortage of available disk space). Removed packages will be automatically re-installed if new jobs scheduled to run on the WN require them.

Conclusions

The question of which libraries and/or operational environments are available in the Grid, is perceived as one of the most important problems by end users. At present no tool capable of solving this problem, is included in the gLite middleware. The most popular workarounds available consist in installing software packages on a per-site basis, and then publishing the corresponding tag in the other.GlueHostApplicationSoftwareRunTimeEnvironment list. Alternatively, specific users adopted ad-hoc solutions like the manual installation of all prerequisites, before launching the core job.

The first approach is disliked by local system administrators. The second one forces the user to become knowledgeable in system administration, beyond his or her main interest which is just to run the code. This is one of the most critical aspects that can seriously hamper a wide diffusion of the Grid.

Primary author: Mr SPONZA, Massimo (OATs - INAF)

Co-authors: Mr VUERLI, Claudio (OATS - INAF); Mr TAFFONI, Giuliano (IASF - INAF); Mr COZZINI, Stefano (CNR - IOM Semocritos)

Presenter: Mr SPONZA, Massimo (OATs - INAF)

Session Classification: Posters

Track Classification: Poster