

Various Runtime Environments in Grid by Means of Virtualization of Working Nodes

Andrey P. Demichev

E-mail: demichev@theory.sinp.msu.ru

Ilya N. Gorbunov

E-mail: gorb44@mail.ru

Viacheslav A. Ilyin

E-mail: ilyin@theory.sinp.msu.ru

Alexander P. Kryukov*

E-mail: kryukov@theory.sinp.msu.ru

Lev V. Shamardin

E-mail: shamardin@theory.sinp.msu.ru

*Skobeltsyn Institute of Nuclear Physics,
Lomonosov Moscow State University (MSU SINP),
119991 Moscow, Russian Federation*

In the framework of this work an approach for batch processing of computer jobs prepared for various runtime environments in grid is proposed. This method is based on the virtualization (using virtual machines) of working nodes of grid resource centers and enables executing applied jobs irrespective of runtime environment it has been initially developed for. In particular, jobs developed for execution in the environment of the widespread OS MS Windows, can be processed in resource centres of the grid system whose working nodes operate under OS Linux.

*XII Advanced Computing and Analysis Techniques in Physics Research
Erice, Italy
3-7 November, 2008*

* Speaker

1.Introduction

Nowadays grid systems are used for calculations and data processing in various applied areas such as biomedicine, nanotechnology and materials science, cosmophysics and high energy physics as well as in a number of industrial and commercial areas. However, one of the basic problems costing on a way to wide use of grid systems is related to the fact that applied jobs, as a rule, are developed for execution in a definite runtime environment specified by type and version of operating systems (OS), auxiliary software (libraries), type of file system, presence or absence of facilities for parallel computing, etc. On the other hand, working nodes (WN) in the resource grid centers (where the jobs are executed) operate under control of a certain OS and offer a fixed runtime environment. Therefore if applied jobs were not developed initially for the particular runtime environment of the WNs, they cannot be directly processed in the grid.

For huge scientific or industrial projects (e.g., such as Large Hadron Collider, LHC [1]) it is quite reasonable that all WN of a grid system offer only a single specific runtime environment because in the framework of such projects a special applied software can be developed for use in the given grid. Such situation is not acceptable when grid is used for solving problems in the framework of small and medium size projects. The point is that frequently users working for such projects already have their own software for solving applied problems or wish to use third party software for this aim and just need computing resources. For various reasons (e.g., due to license and/or financial restrictions) in most cases it is impossible to adapt available applied software for the runtime environment of a grid system which can provide the required computing resources. Therefore the development of technology for offering in grid runtime environment on demand is of great importance.

The main objective of this work is to propose a method for batch processing in grid of jobs prepared for various runtime environments. In particular, jobs developed for MS Windows, should be able to be processed by WNs operating under OS Linux. The work is intended mainly for EGEE/RDIG [2] grid infrastructure. Therefore the implementation is compatible with the gLite [3] middleware.

Among related works it is worth mentioning the project Virtual Workspaces [4] in the framework of Globus Alliance community. The aim of this project is creation of special grid service named Virtual Workspace (VW) for deployment and configuring virtual machines in grid working nodes so that authorized users may communicate with them. However, Virtual Workspace is intended, basically, for an interactive communication of users with the virtual environments. In particular, it can be used as a portal [5] for Amazon Elastic Compute Cloud [6]. On the other hand, in large grid systems the principal (and often single) mode is a batch mode. Our approach provides availability of various runtime environments in the batch mode and therefore more suitable for use in large grid systems such as EGEE/RDIG.

2. Batch processing in grid of jobs prepared for various runtime environments

The method is based on virtualization of WNs by using virtual machines (VM). For this aim the Xen [7] virtual machine monitor is used both in paravirtualization and full virtualization modes, the latter being possible owing to hardware support (Intel VT / AMD-V). As it is specified above, the implementation of the method uses the gLite middleware. An advantage of our approach is that enabling job processing in various runtime environments requires a *minimal* gLite modification. In fact, only the gLite WNs must be modified. Namely, the modification is reduced to the following operations:

- installation of virtual machine monitor (VMM) Xen;
- appropriate modification of the wrapping script `jobwrapper`;
- installation of the two new scripts:
 - `jobwrapper-vee`;
 - `xenrun`.

In addition, the Computing Element (CE) of the gLite middleware should be appropriately configured: namely, a file `host.list` must be added in the directory `/opt/edg/var/info/host` with the list of runtime environments available at the given grid site. Disk images with these runtime environments must be available in WN's file system. The content of the file `host.list` should be of the form:

```
VEE-REname1
VEE-REname2
...
```

Here `VEE-REname1` , `VEE-REname2` , ... are the names of files with disk images of the runtime environments available at the given site.

Use case for the modified system is as follows. A user specifies in the JDL-file [8] required runtime environment by using the Glue schema [9] attribute `GlueHostApplicationSoftwareRunTimeEnvironment`. A sample of such JDL-file is given below:

```
Executable = "testJob.exe";
StdOutput = "testJob.out";
StdError = "testJob.err";
InputSandbox = { "./testJob.exe", "job.jdl" };
OutputSandbox = { "testJob.out", "testJob.err" };
Requirements = Member( "VEE-REname1",
other.GlueHostApplicationSoftwareRunTimeEnvironment );
```

Then the gLite Workload Management System (WMS) with the help of information system finds an appropriate Resource Center with CE matching the requirements in the JDL-file and submits the job to this CE. The CE processes the job in the usual (gLite) manner, transfers it to the local batch system and eventually to a modified WN. The algorithm of job processing by such a modified WN is depicted on Fig.1.

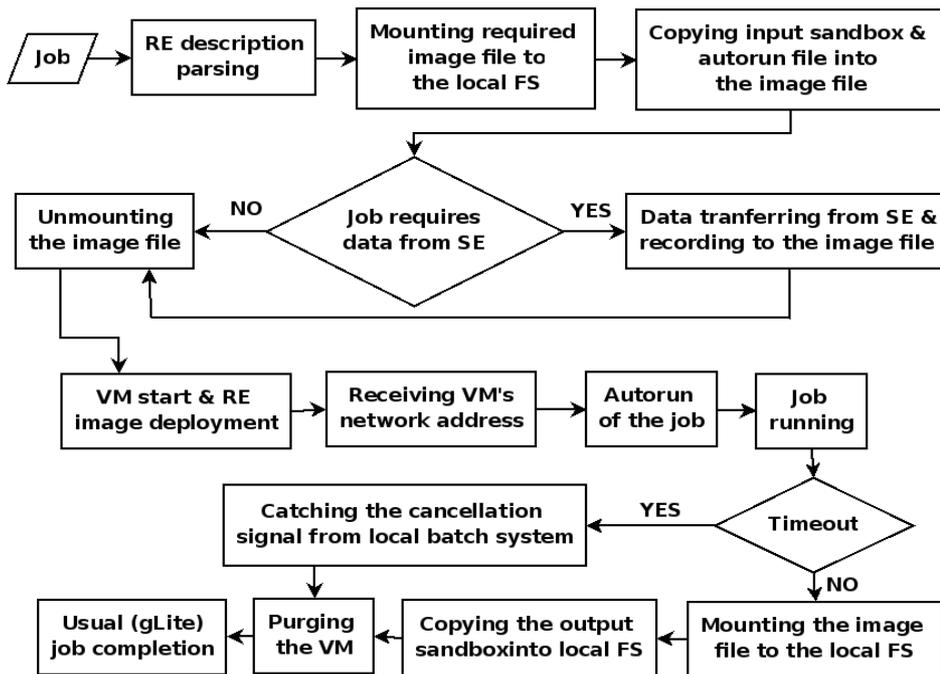


Fig. 1. Algorithm of job processing by a modified WN

After processing a job in the required runtime environment of the modified WN, results can be obtained by means of the usual gLite facilities.

It is worth mentioning that it is desirable to add some additional functionality to the depicted algorithm and to the current implementation of the system for job submission in various runtime environments. Among such functionalities are capability to deploy custom runtime environments on working nodes and capability of data transfer while job running. This work is in progress.

3. Conclusion

Development of the proposed method for batch processing of jobs prepared for various runtime environments can be crucial for creation of grid infrastructures offering computer resources for solving problems in various scientific, engineering, innovative and industrial areas such as nanotechnology, thermonuclear power, biomedicine, geoscience, etc. A wide spectrum of the requirements to runtime environments from users working in these areas can be satisfied on the basis of the virtualization technology. Thus the approach suggested provides a substantial extension of the class of applied problems that can be solved with the aid of global grid infrastructures. This is an important step to further development of grids and their wide commercial applications.

This work was partially supported by The Federal Agency for Science and Innovation (GC-02.514.11.4072 and GC-NSh-1456.2008.2), SKIF-GRID Program of the Belarusian-Russian Union State (GC-SS-2/07) and RFBR (grants No. 07-07-12023 and No. 08-07-00294).

References

- [1] The LHC project: <http://lhc.web.cern.ch>
- [2] EGEE Project: <http://www.eu-egee.org>; RDIG Project: <http://egee-rdig.ru>
- [3] gLite Middleware Project: <http://glite.web.cern.ch/glite>
- [4] Virtual Workspaces Project: <http://workspace.globus.org>;
http://dev.globus.org/wiki/Incubator/Virtual_Workspaces
- [5] Virtual Workspaces Project (IAQ): <http://workspace.globus.org/vm/TP1.3.1/iaq.html>
- [6] Amazon Elastic Compute Cloud: <http://www.amazon.com/b/?node=201590011>
- [7] Xen Project: <http://www.xen.org>
- [8] Job Description Language HowTo, *DataGrid-01-TEN-0102-0_2*, http://server11.infn.it/workload-grid/docs/DataGrid-01-TEN-0102-0_2-Document.doc;
JDL Attributes Specification, *EGEE-JRA1-TEC-555796-JDL-Attributes-v0-6*,
<https://edms.cern.ch/file/555796/1>
- [9] GLUE Schema: <http://glueschema.forge.cnaf.infn.it>, <http://forge.ogf.org/sf/projects/glue-wg>