Enabling OMNeT++-based Simulations on Grid Systems

M. Kozlovszky
MTA SZTAKI
Victor Hugo u. 18-22, H-1518,
P.O.Box 63., Budapest, Hungary
+36-1-2796072
m.kozlovszky@sztaki.hu

A. Balasko
MTA SZTAKI
Victor Hugo u. 18-22, H-1518,
P.O.Box 63., Budapest, Hungary
+36-1-2796072
balasko@sztaki.hu

A. Varga
OpenSim Ltd.
Szőkö köz 11
H-1032 Budapest, Hungary
+36 1 3882763
andras@omnetpp.org

ABSTRACT
Simulations are typically computationally intensive problems, and lend themselves for execution on large-scale PC clusters or grids. Using grid infrastructure for discrete event simulation is currently not prevalent, but making grid technology easily accessible to simulation users can change that picture significantly. In this paper we give a detailed overview how the OMNeT++ simulation framework was ported onto a gLite-based grid infrastructure. The porting of the simulation framework to the grid infrastructure was supported by the GASUC Team of the EGEE III project. Later on in the paper we show an example grid service which is able to execute queuing network simulations, and assess its performance on the grid.

General Terms
Algorithms, Performance, Design.

Keywords
Grid, OMNeT++ simulation framework, gridification, grid infrastructure

1. INTRODUCTION
Distributed computer environments are using different technologies to coordinate their resources, such as processing power, job scheduling, and storage consumption. Porting legacy applications onto grid infrastructures is called "gridification". Porting the OMNeT++ framework to grid is useful for users running simulations with large-scale Parameter Studies (PS). Speeding up simulations with large parameter studies can be done by launching the applications on large computer farms. Potential users or beneficiary community of the grid-enabled application can be anyone running simulations with large parameter studies (the total execution time should be well over a few hours for a user to feel the need for grid; for example, when one simulation run takes 10 minutes and 60 runs are needed for a study, that's enough incentive to put in that extra effort needed for grid execution).

1.1 OMNeT++
OMNeT++ is a public-source, component-based, modular, discrete event simulation environment. Due to its strong GUI support and embeddable simulation kernel it is frequently used for simulation of communication networks, IT systems, queuing networks and for various business processes. OMNeT++ provides mainly component architecture for models. Components (modules) are programmed in C++, and then assembled into larger compound models using a high-level language (NED). OMNeT++ supports different platforms, such as Linux, various Unix-like systems and Windows (XP and 2000).

1.2 Related efforts
There have been several related efforts to exploit the computing power of clusters or grids for running simulations. Remote OMNeT++[11][12] is a now discontinued effort for distributing simulations onto remote processing hosts; the software was built on Java technology. A Remote OMNeT++ system is composed of Clients with GUI to control models and simulation runs of a user, of Processing Hosts to execute the simulations, and of Data Warehouses to store simulation models and results.

Akaroa-2 [10] is an implementation of the MRIP (Multiple Replications In Parallel) principle, which can be used to speed up steady-state simulations. Akaroa runs multiple instances of the same simulation program (but with different seeds) simultaneously on different processors, e.g. on nodes of a computing cluster, and a central process monitors certain output variables of the simulation. When Akaroa decides that it has enough observations to form an estimate of the required accuracy of all variables, it halts the simulation. When using n processors, simulations need to run only roughly 1/n times the required sequential execution time. Support for Akaroa is integrated into OMNeT++ [9] and is also available for ns-2 and Ptolemy Classic.

RSerPool (Reliable Server Pooling) [7] is an IETF protocol for server pool management and access, and it can also help in exploiting multiple computers for running simulations. The SimProcTC toolchain [8] provides scripts and makefiles for distributing OMNeT++ simulations to remote hosts for execution. SimProcTC uses an RSerPool implementation to find processing hosts (PE) that are suitable for running simulations, then proceeds to upload a tarball containing the simulation files (binaries plus...
Grid computing differs from conventional cluster computing systems by its heterogeneous, loosely coupled and geographically dispersed infrastructure. There are two major grid infrastructure development directions: service grids and desktop grids. Desktop grids are built on the concept of best effort, and are often based on Boinc [27] or XtremWeb [30] technologies (examples: SETI@Home [28] and EDGeS [29]); service grids provide a 24/7 computing service. Service grid infrastructures are potentially more suitable for running large-scale discrete event simulations.

One service grid example is EGEE (Enabling Grids for E-science), the largest multi-disciplinary grid infrastructure in the world, which brings together more than 140 institutions to produce a reliable and scalable computing resource available to the European and global research community. At present, it consists of approximately 300 sites in 50 countries and gives its 10,000 users access to 80,000 CPU cores around the clock.

Another large-scale grid is operated by the SEE-GRID-SCI (SEE-GRID eInfrastructure for regional eScience) project, which spans across 14 countries with 35 sites and 3000 CPUs. Other pan-European grid efforts include GRIDSTART, DAMIEN, DATAGRID, EUROGRID, CROSSGRID and DATATAG. In addition, several universities and research institutes operate specialized grids, and clusters running grid software; one example is the LHC Computing Grid [18].

Grid systems are composed of a set of integrated services: job queuing, scheduling policy, priority scheme, resource discovery and monitoring, resource management, data storage and management, security, user authentication and credentials management. Grid software places user-submitted jobs into a queue, chooses when and where to run the jobs based upon a policy, carefully monitors their progress, and ultimately informs the user upon completion.

Grid systems run various grid software packages. The Globus Toolkit [16], EGEE gLite [17], ARC (aka. NorduGrid middleware) [19], Sun Grid Engine [20], Condor [21] are all open-source, widely deployed grid middleware packages, themselves usually composed of several other, more fundamental service components. Deploying these toolkits on a set of computers usually requires substantial time and administrative skills.

Real-world adoption of grids is hindered by several factors. First, grid middleware functionality is most often exposed in the form of complex command-line interfaces, which poses challenges to average grid users who are used to point-and-click graphical user interfaces. Second, the command interfaces of different grid middleware packages differ significantly, so experience with one grid does not immediately help with another. Third, users have to get a login individually for each of the grids they intend to use.

Portal software solutions with web based easy-to-use user interface (such as the P-GRADE Grid Portal) have been developed to address the above issues and to help grid adoption.

1.4 P-GRADE Grid Portal

The P-GRADE Grid Portal (Parallel Grid Run-time and Application Development Environment) [4][14][15] is an open source, service-rich, workflow-oriented graphical grid front-end. It supports workflows composed of sequential jobs, parallel jobs and application services. The P-GRADE Portal hides the complexity of the grids through its high-level graphical web interface, and it can be used to develop, execute and monitor workflow applications on grid systems built with Globus, EGEE (LCG or gLite) and ARC middleware technologies. P-GRADE Portal installations typically provide the user with access to several grids, using a single login.

More than ten large-scale P-GRADE Portals are operating and serving the user communities of international multi-institutional grids and grid based virtual organizations in Europe and in the U.S. The list of these installations and information on how to get user accounts to the corresponding portals is available on the P-GRADE Portal website [14].

Non-grid applications need to be "gridified" for use with P-GRADE Portal. The gridification process includes developing a specialized web interface extension to P-GRADE Portal, specifically tailored for the application's needs. Sections 2 and 4 of this paper describe the gridification process of OMNeT++ queueing simulations. Section 5 presents the resulting web interface where users can submit simulations and download the results.

1.5 The GASuC Team

The gridification of OMNeT++ simulations was done by the GASuC team at MTA SZTAKI, Hungary. GASuC (Grid Application Support Centre) is supported by the EGEE III project¹, provides assistance to grid users and application developers during the application gridification process. GASuC helps identify and apply best patterns, practices, tools and infrastructures in order to decrease the porting time and get application code running on production grids. The generic support model of GASuC consists of 8 main steps: 1. Contact phase, 2. Pre-selection phase, 3. Analysis phase, 4. Planning phase, 5. Prototyping phase, 6. Testing phase, 7. Execution phase, 8. Dissemination and feedback phase. In case of OMNeT++, these phases were extended with more contacts during analysis and planning phases.

¹ EGEE-III is an FP7 project funded by the European Commission under contract INFSO-RI-222667.
1.5.1 Workflow concept of the P-GRADE Portal
The P-GRADE Grid Portal environment is using a DAG (directed acyclic graph) based workflow concept (shown in Figure 1).

In a generic workflow, nodes (shown in Figure 1. as large gray squares) represent jobs, which are basically batch programs to be executed on a computing element. Ports (shown here as small squares around the large ones) represent input/output files the jobs receiving or producing. Arcs between ports represent file transfer operations. The basic semantics of the DAG-based workflow is that a job can be executed if and only if all of its input files are available.

![Figure 1. Example workflow in P-GRADE Portal.](image)

1.5.2 Parameter Study support in P-GRADE Portal
In our experience, user communities have shown substantial interest in being able to run programs parallel with different input files. P-GRADE Portal supports this kind of parallelization called Parameter Study, or Parameter Sweep on a high level.

The original “job” idea has been extended by two special jobs called Generator and Collector to facilitate the development of Parameter Study type workflows in P-GRADE Portal. The Generator job is used to generate the input files for all parallel jobs automatically (called Automatic Parameter Input Generator), or by a user-uploaded application (called Normal Generator), while the Collector will run after all parallel execution and collects all parallel outputs [5].

All jobs connected to a Generator will run in as many instances as input files generated by the Generator. During the gridification of OMNeT++ simulations, we have exploited these possibilities of the P-GRADE Portal.

2. Gridification process
Main steps of the gridification process are shown on Figure 2. The compilation of the source code on the UI machine assures that the application will work on all the gLite-based Computing Elements in the grid infrastructure. Standalone, single run tests showed that we need to have the compiled binary code, the predefined NED and INI files as inputs for the porting of the simulation into a grid workflow.

We have also investigated the time and storage consumption and running environment requirements (e.g. shared libraries: libstdc++) of the simulator. For the first grid tests (single job tests), we have written a JDL (Job Description Language) file and created the workflow graph of the simulation in the P-GRADE Portal environment. The parallel job submission into the grid environment needs to have parameter assignment of the generated parameters. P-GRADE Portal’s PS workflow components and OMNeT++’s parameter study solution (a detailed description about the PS handling in OMNeT++ follows in Section 3.) were integrated together to create a grid-aware OMNeT++ simulator environment and to realize a complex grid workflow as a proof of concept. Later on, we have developed the web-based grid user interface for OMNeT++ simulator framework with the help of the Application Specific Module of the P-GRADE portal. On this web interface, end-users can upload their network topology and configuration files for queuing network simulations, and they can submit the simulation into the grid environment with large parameter fields.

![Figure 2. Gridification steps of the simulator.](image)

3. Parameter Study handling in OMNeT++
3.1 Overview
In a simulation study, we create a simulation model, then perform experiments on it in order to learn about its behavior under various conditions. OMNeT++ emphasizes separation of concerns, which in this case means that the model and its input parameters are in physically separate files. The simulation model is represented by C++ code (and the libraries/executables built from them) and NED files describing the topology and fixed parameters. Input parameters for the experiments are in INI files, which also contain configuration settings for the simulator.
In the OMNeT++ 4.0 version, an INI file (the default name is omnetpp.ini) may contain several named configurations. Every named configuration contains parameter assignments, each of which typically sets more than one model parameters, using wildcards (i.e. **.tcp.mss=1024 sets the MSS parameter for all TCP instances in the model). Named configurations may also contain simulator settings (stopping condition, seeds, etc), and they can also inherit from each other which makes it possible to factor out common parts. For every simulation run, a named configuration needs to be specified where all parameters and settings will be taken from.

A named configuration may also contain iterations and repetitions, to support parameter studies. Thus, a configuration may describe several simulation runs. They are identified with run numbers, in the range 0..n-1. Each run may be executed independent of the others. To perform a specific run, the user has to tell the configuration name and the run number to the simulation runtime, typically via command-line arguments. The simulation runtime may also be asked about the number of runs a particular configuration generates, or to enumerate the settings for each run.

By default, result files will be created in the result/subdirectory of the simulation. Each run generates separate files, by default named after the configuration and the run number. Result files are self-describing, and contain enough metadata so that later it can be identified how they were created.

Every run also gets a globally unique identifier, called a run ID, concatenated from the configuration name and run number, date/time, process ID, etc. Run IDs are also recorded into the files, and can be used (for example) to relate different files created by the same run.

### 3.2 Iterations
The value part of a parameter assignment in the INI file may contain iterations. Syntax of an iteration is ${valuelist} or ${name=valuelist}, where valuelist is a comma-separated list of values, and value is some arbitrary text (typically a numeric constant). There is also a shorthand for generating numeric sequences, with the syntax start..end or start..end step increment. Each value generates a separate simulation run, where the value gets textually substituted into the parameter assignment before the parameter value gets interpreted. If there is more than one iteration in a named configuration, they are combined with Cartesian product, i.e. they essentially form nested loops. It is also possible to specify a constraint expression, to take a subset of the Cartesian product. In the constraint expression, one can refer the current iteration values with the syntax $\{name\} or $0, $1, etc. This iteration variable syntax also makes it possible to refer to an iteration value in parameter assignments other than which defined it.

A special parallel iteration syntax makes it also possible that an iteration does not form a Cartesian product with another iteration, but gets advanced in lockstep with it. Given the iterations $\{N=10,20,50,100\}$ and $\{D=\text{"low"},\text{"medium"},\text{"high"},\text{"veryhigh"}!N\}$, four simulation runs will be generated, with N and D being (10,"low"), (20,"medium"), (50, "high"), and (100, "veryhigh"). There are also predefined variables: $\{configname\}, $\{runnumber\}, $\{runid\}, $\{datetime\}, $\{processid\}, etc. They can be useful for generating file names, for example.

#### 3.3 Repeating runs with different seeds
It is directly supported to perform several runs with the same parameters but different random number seeds. There are two configuration options related to this: repeat and seed-set. The first one specifies how many times a run needs to be repeated. For example, repeat=10 causes every combination of iteration variables to be repeated 10 times, with the $\{repetition\}$ predefined variable being the loop counter; thus, repeat=10 is essentially equivalent to adding $\{repetition=0..9\}$ to the ini file. The $\{repetition\}$ loop always becomes the innermost loop.

The seed-set configuration option affects seed selection. Every simulation uses one or more random number generators (as configured by the num-rngs option), for which the simulation kernel can automatically generate seeds. The first simulation run may use one set of seeds (seed set 0), the second run may use a second set (seed set 1), and so on. All automatic seeds generate random number sequences that are far apart in the RNG’s cycle, so they will never overlap during simulations. (OMNeT++ uses Mersenne Twister, which has a cycle length of $2^{19937}$.)

#### 3.4 How to run
To run a parameter study expressed with iterations as described above, one first needs to find out how many runs the iterations unroll to. This is achieved by running the simulation executable with the $\text{\texttt{--x <configname>}}$ command-line option, which in turn prints out the n number of runs in the given named configuration. Individual runs can then be executed by running the simulation program with the $\text{\texttt{--c <configname> -r <runnumber>}}$ options, where runnumber is an integer in the 0..n-1 range. The fact that each run can be executed independently enables them to be submitted to grids as well. Output files will be named after the configuration and the run number (although this is configurable as well).

#### 3.5 Experiment-measurement-replication
OMNeT++ uses the concepts experiment, measurement and replication to organize simulation results generated by batch executions or several batches of executions. During a simulation study, a person prepares several experiments. The purpose of an experiment is to find out the answer to questions like "how does the number of nodes affect response times in the network?" By default, an experiment is represented by a named configuration in the ini file, but it is also possible to group several configurations into a single experiment. For each experiment, several measurements are performed on the simulation model, and each measurement runs the simulation model with a different parameter settings. To eliminate the bias introduced by the particular random number stream used for the simulation, several replications of every measurement are run with different random number seeds. Measurements map to iteration variables in the ini file, and replications map to the repeat option.

In order to make result analysis tools aware of the above concepts, OMNeT++ saves an experiment label, measurement label and

---

2 Run IDs are only globally unique within reasonable limits, because we held it more important that they are both short and readable.

3 This macro-style implementation was chosen because in OMNeT++ it would be impractical to iterate over individual parameters directly, given that each parameter assignment in an INI file typically assigns several parameters together, using wildcard expressions. Macros also provide more flexibility, much in the same way as C macros can do things not possible with C functions.
replication label into result files. The labels can be configured in
the ini file, but the suitable defaults are "$\text{configname}$$","\text{iterationvars}$$" and the "$\text{repetition}, seed-set=<seedset>$$" text.

4. Grid workflow structure
The grid-aware queuing model workflow was developed using the
P-GRADE Portal’s Workflow Editor. A simple queuing model
diagram (Tandem Queue) is depicted in Figure 3.

![Tandem Queue Diagram](image)

Figure 3. Simple queuing model diagram from OMNeT++
To enable the execution of the model as a parameter study
application, we have used the “Automatic Parameter Input
Generator”, and the “Collector” elements of the P-GRADE Portal.

![Workflow Diagram](image)

Figure 4. The resulting workflow diagram with PS support in
the P-GRADE Portal’s Workflow Editor
The upper left box (shown in Figure 4.) is the Automatic
Parameter Input Generator job, which automatically generates
index numbers, and separates the runs by providing instructions
which run should be executed. The generated indices become
arguments to the $-r runnumber> command-line option of the
OMNeT++ simulation executable. This job takes little CPU time,
so it always runs on the machine where the P-GRADE Portal was
installed.

The second box in the middle contains the application in the form
of a pre-compiled binary, the input parameter and topology files
(INI and NED files) and some additional libraries. This box will
be executed in parallel on different machines, once for each of the
index files generated by the Parameter Input Generator. These
processes perform the simulations, and produce OMNeT++ vector
and scalar result files. The result files are collected by the third,
“Collector” box (lower right). At the end of the workflow, the
Collector job automatically compresses the results, and enables
users to download the compressed file for further user analysis.
The user interface has an additional feature that users can
download a part of the output files during the execution.

This workflow structure is fully hidden from the end-users by a
web-based user interface.

5. User interface development for grid-aware
OMNeT++ models
We have created an easy-to-use service on the web for queuing
model simulations. On this web interface end users, who are not
grid professionals, are able to upload their configuration and
topology files, setup and submit their model into the grid
infrastructure. (In this setup, the end user cannot access and
change the compiled binary.) The results produced by the
submitted simulations are downloadable by normal web browser
when they are finished. The developed web-based (portlet) user
interface has three main views:

- The “OMNeT++ Measurement Manager” portlet.
This is where the end user can choose among the available
OMNeT++-based model workflows (shown in Figure 5).

![OMNeT++ Measurement Manager](image)

Figure 5. OMNeT++ Measurement Manager portlet
- The “Upload inputs” portlet.
This is where the end user can upload the queuing topology
(NED) and configuration parameter (omnetpp.ini) files, and select
the indexes or interval of indexes to run (shown in Figure 6).

![Upload inputs portlet](image)

Figure 6. Upload inputs portlet
- The “Execution and download” portlet.
This is where the end user can execute the model by a single
button (which means basically a complex workflow submission
procedure into the grid infrastructure) with its parameter field.
(shown in Figure 7).
As the output files are generated in parallel during the execution, users can download them by clicking “Download” button. It is also possible to download the (partial set of) result files while some jobs are still running or waiting for execution.

6. Performance
The EGEE[24] and SEE-GRID-SCI[25] grids were used for submission of the developed workflow model. We have used several times a simple queuing model to test our porting concept. In our largest parameter study demo, the simulation time of the used queuing model was only 1 minutes, however the parameter field of the model generated 5000 “runs”, which would last about 80 hours to execute on a single average computer. We have clustered the runs with the help of the “Automatic Parameter Input Generator” into smaller (100-run) chunks, and submitted the model into the grid infrastructure. The grid broker allocated 50 computers for these chunks from the grid as it was requested. The results of the parallel grid submissions were available after 3 hours, instead of 3.5 days. Each running generated 25-225 kBytes of data (the compressed sca files), size of the final output was about 6MB.

7. Future plans
Future plans include allowing for arbitrary OMNeT++ simulations to be run on the grid. When the user's simulation is built exclusively from unmodified components of a published model framework (for example the INET Framework), the corresponding binaries and NED files would already be pre-deployed on the grid and the user would only have to upload the network NED file(s) and INI files to run the simulation.

In the more common case when the user's simulation involves nonstandard components or C++ code modifications, the user will need to get access to the official P-GRADE Portal, and will need to have a certificate for running applications in the grid environment. Then the user has to develop the workflow, and a compiled OMNeT++ simulation model has to be included into it. Finally the user can use all of features of P-GRADE Portal to run and manage the simulation.

8. Conclusion
We have successfully integrated the P-GRADE Grid Portal environment with the OMNeT++ simulation framework to enable additional, already existing, large-scale grid resources to the simulation user community. We have used the parameter study features from both (P-GRADE Portal and OMNeT++) environments to create a generic workflow solution for OMNeT++-based simulation models. We have shown that large-scale grid infrastructure can provide significant performance increase for OMNeT++-based simulations. As a proof of concept, we have created from the developed queuing workflow a web-based service, where a configurable queuing model can be submitted by the end-user into the grid infrastructure without any grid knowledge. If the developed workflow and the service solution can grow beyond the concept phase, they have the potential to become a useful tool for developers and end-users of simulations.

9. REFERENCES


