

# Ranking and Performance Exploration of Grid Infrastructures: An Interactive Approach

George Tsouloupas Marios D. Dikaiakos  
Dept. of Computer Science,  
University of Cyprus  
1678, Nicosia, Cyprus  
{georget,mdd}@ucy.ac.cy

**Abstract**—In the ever-growing Grid infrastructures picking the right resources is not an easy task. There are situations where finding enough idle CPU's that satisfy a set of minimum nominal hardware requirements is simply not enough. In many such occasions, users or VO managers need to be involved in the resource selection process, either by excluding resources or by indicating preference, more so when resources are not for free. To support users in this task, we propose an interactive and user-driven approach to a performance-based ranking of Grid resources. Using the GridBench framework and the provided suite of micro-benchmarks, the user is able to interactively explore and rank what is currently available on the Grid in order to pinpoint the right resources. Through this framework, the user is able to interactively compose ranked lists of Grid resources using custom ranking functions that are based on low-level measurements.

## I. INTRODUCTION

In most large-scale production Grids today, resource allocation is based on static or dynamic information provided by the resources via Grid Information Services. This information generally refers to the free or total CPU-count of each resource, their CPU nominal speed and the size of main memory, available bandwidth, etc. This approach works well in infrastructures where the main goal is to achieve a high overall job throughput and users do not have strict requirements on the *performance* of resources. In several scenarios, however, end-users need to “shop around” for Grid computing resources and many VO operators want to audit the *real* availability, configuration status and performance of their providers’ computing resources. In such cases, the information published by resource providers and Grid monitoring systems does not provide sufficient detail and accuracy. Grid users need, instead, the capability to define and configure on-demand different kinds of *end-to-end* benchmarks [2] or tests tailored to the characteristics of individual resources and/or to the measurements sought. Grid users also need the capability to administer such tests, analyze and compare their results in an *interactive* fashion.

To do this we need to: (i) Determine a core, yet extensible, set of minimally intrusive probes, that can deliver metrics to effectively characterize the performance capacity of Grid resources; (ii) Provide indicators on the quality of obtained measurements and possibly filter them; and (iii) Facilitate the specification, configuration, administration and analysis

of measurements that will result in the selection of the right resources.

To address these issues we have built a set of tools on top of GridBench [3], [5], a framework for evaluating the performance of Grid resources. In previous work [3], [5], [4] we have described the architecture and generic functionality of the GridBench framework. We provide an outline for an interactive, user-driven approach to ranking Grid resources, with user-specified metrics, custom ranking functions and ranking models. We ran our experiments on the EGEE [1] infrastructure.

## II. GRIDBENCH: TOOL DESCRIPTION

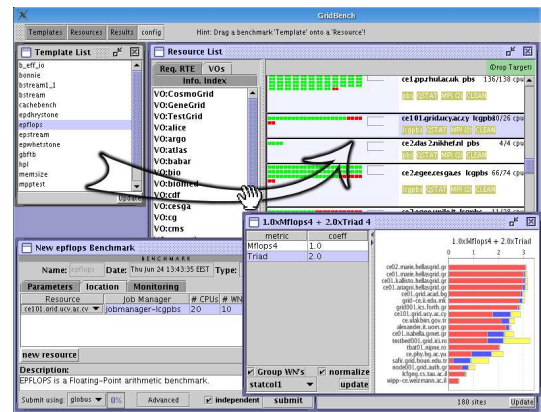


Fig. 1. The GridBench Graphical User Interface

GridBench is a tool which aims to automate the tasks of testing and benchmarking Grid resources where users can define, execute and manage tests and benchmarks. The tool is comprised of (i) a set of web-services that deal with the execution of tests and benchmarks specified in the XML-based GridBench Definition Language, and (ii) a user-friendly graphical interface (Figure 1) that simplifies the definition and execution of benchmarks and tests, as well as browsing and analysis of results. The interface provides tools for analysis through the easy construction of custom graphs from archived results. The tool allows the user to interact by providing real-time resource information, allowing the selective submission

and customization of tests or benchmarks and providing feedback on test/benchmark status. In Figure 1 we observe the list of available benchmarks (top left), the available resources (right), the benchmark customization panel (bottom left), as well as some generated charts.

**Metric Quality:** The underlying infrastructure prohibits targeted measurements of specific worker-nodes (cluster-nodes). We tackle this by sampling the resource. The tool provides two indicators of *metric quality*: (i) *count* – the number of measurements that are used in calculating the reported values of each metric, and (ii) *Sample ratio* – which reflects just how representative of the whole resource the currently available measurements are.

**Metrics and Benchmarks:** We have a selected a core set of micro-benchmarks to be integrated into GridBench, which will provide the necessary low-level metrics. This set of benchmarks provides metrics such as Floating Point operations per second, integer operations per second, memory bandwidth, memory cache performance, disk bandwidth as well as network interconnect latency and bisection bandwidth.

### III. USER-DRIVEN RANKING

The GridBench tool provides a *ranking module* that allows the user to interactively build a *ranking model*. A *ranking model* consists of *filtering*, *aggregation* and *ranking functions* (Figure 2).

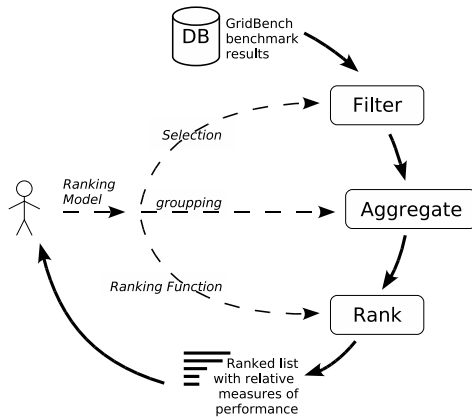


Fig. 2. The ranking process.

**Filtering** refers to a user selection regarding which results will be included or excluded in the ranking process. The user selection can be based on any of the following types of filtering: (i) *attribute filtering*, (ii) *monitoring-based filtering* and (iii) *filtering based on metric quality*,

**Aggregation** allows the user to specify grouping of the measurements and produces a normalized set of statistics for each metric: *mean*, *standard-deviation*, *min*, *max*, *average-deviation* and *count*.

Finally, the user can interactively construct a **Ranking Function** using the GUI that includes metric statistics as well as metric quality indicators, and values from Grid Information Services e.g.  $R = C_1 \cdot \alpha_{mean} + C_2 \cdot \beta_{mean} + C_3 \cdot \gamma_{avdev} + C_4 \cdot \delta_{min} + \dots$

We have performed several experiments for ranking the resources of EGEE's South Eastern Europe region based on

different preferences. In Figure 3 we show the result of a *ranking function* that involves three normalized metrics (Mflops/s, dhrystones/s and memory bandwidth).

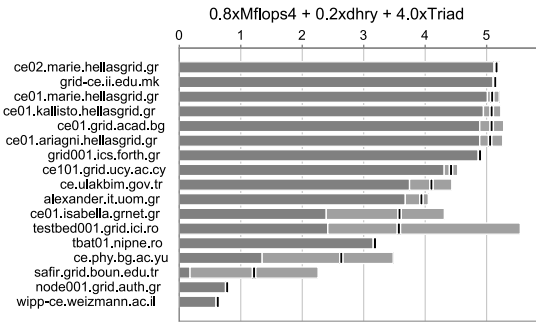


Fig. 3. Ranking of SEE resources based on a *ranking function* involving several metrics.

Figure 4 shows the result of a ranking function that also factors in the number of CPU's at a given site to produce an *aggregate* metric. The number of CPU's just like other interesting values from an information system can be part of the ranking function.

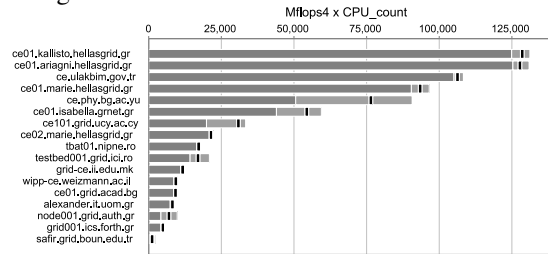


Fig. 4. Ranking based on aggregate floating-point performance.

### IV. CONCLUSIONS AND FUTURE WORK

The process of picking resources on the Grid can be simplified using a tool that allows easy and effective user interaction to run benchmarks and construct rankings based on knowledge of their application's characteristics. In on-going work we are working on (i) preserving the user preferences and actions that resulted in the user-driven rankings in a way that can be reproducible as part of a resource brokering service, and (ii) establish mechanisms that can help the user understand an application's characteristics in terms of the low-level metrics included in the tool.

### REFERENCES

- [1] EGEE: Enabling Grids for eScience in Europe. <http://www.eu-egee.org>, (accessed April 2004).
- [2] Marios D. Dikaiakos. Grid benchmarking: Vision, challenges, and current status. *Concurrency and Computation: Practice and Experience Journal*, (accepted for publication).
- [3] G. Tsouloupas and M. D. Dikaiakos. GridBench: A Tool for Benchmarking Grids. In *Proceedings of the 4th International Workshop on Grid Computing (Grid2003)*, pages 60–67. IEEE Computer Society, November 2003.
- [4] George Tsouloupas and Marios D. Dikaiakos. Characterization of Computational Grid Resources Using Low-level Benchmarks. Technical Report TR-2004-5, Dept. of Computer Science, University of Cyprus, December 2004.
- [5] George Tsouloupas and Marios D. Dikaiakos. Gridbench: A workbench for grid benchmarking. In *EGC*, pages 211–225, 2005.