Heterogeneous Grid Computing: Issues and Early Benchmarks

Eamonn Kenny[1], Brian Coghlan[1], George Tsouloupas[2],
Marios Dikaiakos[2], John Walsh[1], Stephen Childs[1], David O'Callaghan[1],
and Geoff Quigley[1]

[1] Department of Computer Science, Trinity College Dublin, Ireland
{ekenny, coghlan, walshj1, childss, ocallwd, gquigle}@cs.tcd.ie
[2] Dept. of Computer Science, University of Cyprus, 1678, Nicosia, Cyprus
{georget, mdd}@ucy.ac.cy

Abstract. A heterogeneous implementation of the current LCG2/EGEE grid computing software is supported in the Grid-Ireland infrastructure. The porting and testing of the current software version of LCG2 is presented for different flavours of Linux, namely Red Hat 7.3, Red Hat 9 and Fedora Core 2. The GridBench micro-benchmarks developed in CrossGrid are used to compare the different platforms.

1   Introduction

Grid-Ireland uses the LCG[1] and EGEE[2] grid middleware, which originally assumed reference ports to Red Hat 7.3 and Microsoft Windows, but subsequently this has been revised to include Scientific Linux 3 (SL3). Unfortunately this is a very restrictive situation, counter to the original heterogeneous ethos of grid computing. As a result of our interest in heterogeneity, we at Trinity College Dublin began porting to non-reference platforms in October 2003. Subsequently EGEE have almost finished porting the current LCG2 grid implementation to Scientific Linux on 32-bit and 64-bit architectures.

2   Porting for Heterogeneity

The LCG2/EGEE software components are shown in the form of a dependency graph in Figure 1. Grid-Ireland wished, in the first instance, that the porting of the LCG2 software to other platforms would focus on the ability to execute Globus and EDG jobs on worker nodes, and that replica management, R-GMA and VOMS would be supported.

To avail of the base functionality requires Globus and various EDG support packages. Since Globus 2.4.3 is known to have many bugs, the University of Wisconsin-Madison corrects these and packages all the necessary components as part of the Virtual Data Toolkit (VDT) [3]. We have assisted Maarten Litmaath

in CERN to port VDT-1.1.14 to IRIX and Fedora Core 2. A Red Hat 9 port is already provided by VDT. A port exists for Globus to Mac OS X and AIX but the VDT version must be ported to both of these platforms.

Grid-Ireland also wished that MPI, replica management and the OpenPBS client be provided on each worker node. In some cases Torque might be required since newer versions of operating systems are not always provided for in OpenPBS. Also the R-GMA information system producer and consumer APIs and the VOMS client were required.
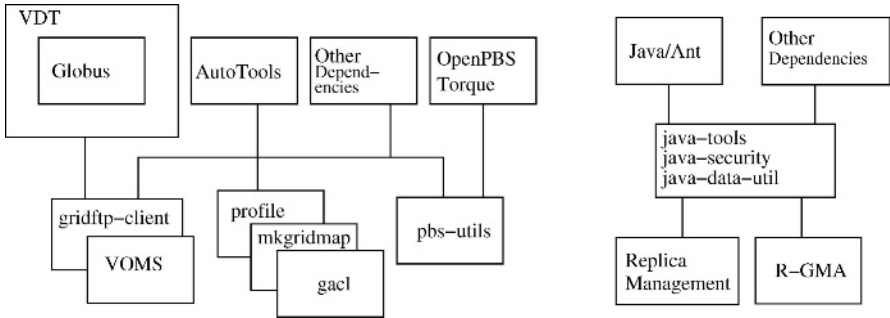

Fig. 1. LCG2/EGEE software components

At the moment there is no requirement by Grid-Ireland for the workload management system (WMS) but it appears that there is logging of WMS events from worker nodes to the resource broker. This logging activity can be disabled but we would prefer to retain this desirable feature. WMS consists of many modules, but it might be able to be refactored to delineate those specific to the event logging, so that just this functionality needs to be ported. It should be noted that if the whole of WMS were ported successfully then almost everything will have been ported because it depends on so many other packages.

There are a number of on-going issues, but we have successfully ported the functionality for job submission to Fedora Core 2, IRIX 6.5.14 and 6.5.17m, AIX 5.2L and Red Hat 9. We also plan to do this for Mac OS X v10.3 very soon, and a number of other platforms if the need arises within Grid-Ireland.

3   Micro-benchmarking Results

Presently EDG job submission is possible for the following non-reference platforms: SGI IRIX 6.5, Fedora Core 2, Red Hat 7.3 and 9 as described in [4], where preliminary results using a fast fourier transform (FFT) were used to show differences in computational speed between different architectures. A routine such as a FFT cannot be independently used as a benchmark since it gives no explicit information about I/O, CPU, caching, floating point or disk write speed.

The CrossGrid GridBench benchmarks[5] developed at the University of Cyprus
provides more precise benchmarking.

GridBench is a tool for benchmarking Grids. It consists of *i)* the GridBench
Framework and *ii)* the GridBench Suite. The GridBench framework provides
mechanisms for defining and executing benchmarks as well as archiving and

Table 1. Micro-Benchmark Results

OS Type	Version	Bonnie (kB/s)	HPL (GFlop/s)	b_eff_io (MB/s)	epdhrystones (dhrystones)	Whetstone (MIPS)	epflops-4 (MFlops)
Fedora Core	2	193536	0.648350	46.712	3858611	707.86	825
Red Hat	7.3	159431	0.696549	52.515	4201680	742.71	898
Red Hat	9	154663	0.729079	45.318	4255319	750.35	899

Table 2. Machine Specifications

OS Type	Version	CPU Speed (GHz)	memsize/Total (MB)	epstream (add)	epstream (triad)
Fedora Core	2	2.8	951/2044	1600	960
Red Hat	7.3	2.8	979/2044	1600	960
Red Hat	9	2.8	872/2044	1600	960
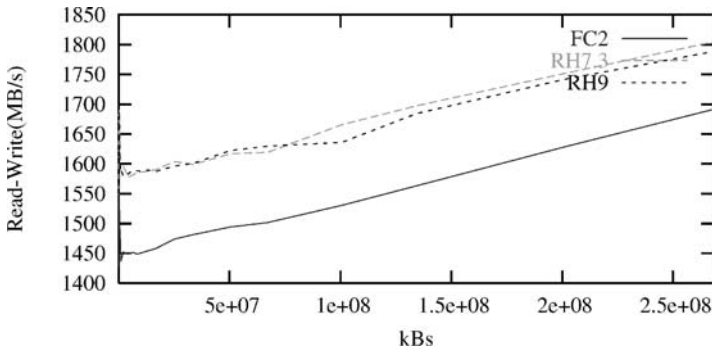
Fig. 2. Cachebench for Fedora Core 2, Red Hat 7.3 and Red Hat 9

managing the results. The Gridbench Framework also provides a user-friendly
graphic user interface serving as a "virtual workbench" for conducting Grid
benchmarks and tests. The GridBench Suite provides a set of benchmarks aim-
ing to characterize Grid resources at different levels (e.g. the worker-node level
or the cluster level). The Gridbench Framework and Suite are described in [5, 6].
Gridbench is currently being enhanced for fully automating the process of bench-
marking heterogeneous resources.

A collaboration between TCD and UCY has resulted in the use of 11 of the GridBench benchmarks from the GridBench suite. Each of these modules, along with a suitable job description language (JDL) file, are required for the job submission to a single workernode to obtain the benchmark results. Initially all 11 modules were sent simultaneously as a batch job, but it was found that the results obtained were inconsistent, with fluctuating wall clock speeds; this was due to the load tailing off as the next submitted job arrived on the worker node. To achieve consistent load between submissions, each job submission must be staggered in time. The results in Table 1 and Figure 2 were computed on 3 worker nodes with identical hardware specifications (see Table 2).

4   Conclusions

The base worker node port of the LCG2/EGEE grid software for Globus and EDG job submission is now completed for Fedora Core 2, Red Hat 9, IRIX and AIX. The results presented in Table 1 and 2 show conclusively that Fedora Core 2 is fastest for write access to the disk (see bonnie in Table 1), while being slower for computational performance (see cachebench in Fig. 2, epflops in Table 1 and HPL in Table 1). In the case of integer computations it is markedly slower (see epdhrystones in Table 1). Results were also obtained for blasbench and the gridbench file transfer benchmark (gbftb) with no noticable differences.

At first the Red Hat 7.3 RPMs were used to perform all the benchmarks on each Linux platform. To ensure that we could be confident of the validity of our results it was necessary to build mpich-1.2.6 and GridBench under Fedora Core 2 using the gcc-3.3.3 compiler. Compiling Gridbench under Fedora Core 2 with identical optimization options to that of Red Hat 7.3 gave no performance gains.

Acknowledgements

References

1. LHC: Large hadron collider computing grid project. http://lcg.web.cern.ch/LCG/ (2004)
2. EGEE: Enabling grids for e-science in europe. http://www.eu-egee.org/ (2004)
3. VDT: Virtual data toolkit. http://www.cs.wisc.edu/vdt/ (2004)
4. Kenny, E., Coghlan, B., Walsh, J., Childs, S., O'Callaghan, D., Quigley, G.: Hetero-geneity of Computing Nodes for Grid Computing. Submitted to EGC 2005 (2004)

5. Tsouloupas, G., Dikaiakos, M.D.: Gridbench: A tool for benchmarking grids. In: Proceedings of the 4th International Workshop on GridComputing (GRID2003), Phoenix, AZ, IEEE (2003) 60–67
6. Tsouloupas, G., Dikaiakos, M.D.: Characterization of computational grid resources using low-level benchmarks. Technical Report TR-2004-5, Dept. of Computer Science, University of Cyprus (2004)