Computational Fluid Dynamics in Solid Earth Sciences–a HPC challenge

Dinámica de Fluidos Computacional en Ciencias de la Tierra Sólida-un reto para el HPC

Vlad Constantin Manea*, **, Marina Manea*, **, Mihai Pomeran**, Lucian Besutiu**, Luminita Zlagnean**

ABSTRACT

Presently, the Solid Earth Sciences started to move towards implementing High Performance Computational (HPC) research facilities. One of the key tenants of HPC is performance, which strongly depends on the interaction between software and hardware. In this paper, they are presented benchmark results from two HPC systems. Testing a Computational Fluid Dynamics (CFD) code specific for Solid Earth Sciences, the HPC system Horus, based on Gigabit Ethernet, performed reasonably well compared with its counterpart CyberDyn, based on Infiniband QDR fabric. However, the HPCC CyberDyn based on low-latency high-speed QDR network dedicated to MPI traffic outperformed the HPCC Horus. Due to the high-resolution simulations involved in geodynamic research studies, HPC facilities used in Earth Sciences should benefit from larger up-front investment in future systems that are based on high-speed interconnects.

RESUMEN

Actualmente, las Ciencias de la Tierra Sólida comenzaron a avanzar hacia la implementación de las infraestructuras de Cómputo de Alto Rendimiento (HPC, por sus siglas en inglés). Una de las características principales del HPC es el rendimiento, que depende fuertemente de la interacción software-hardware. En este trabajo se presentan los resultados de una serie de pruebas realizadas en dos sistemas HPC distintos. Probando un código de Dinámica de Fluidos Computacional (CFD) para las Ciencias de la Tierra Sólida, el sistema HPCC Horus, basado en Ethernet Gigabit, dio resultados excelentes comparándolo con el HPCC CyberDyn, basado en Infiniband QDR. De todos modos, HPCC CyberDyn, basado en una red QDR de alta velocidad y baja latencia dedicada al trafico MPI, supera al HPC Horus. Debido a la necesidad de simulaciones de alta-resolución para geodinámica, las HPC utilizadas en Ciencias de la Tierra deben beneficiar inversiones más grandes en sistemas con interconexiones de alta velocidad.

INTRODUCTION

In recent years, modeling and computation have come to play a central key role in modern Earth Sciences [1-3], and one of the reasons is due to their dependence on fine spatial grids and small time steps for integration used in order to solve numerically systems of equations that express mathematically a physical process [4-6]. Presently, the Solid Earth Sciences started to shift towards implementing high performance computational research facilities, as it can be seen in many universities and research centers: Argonne National Laboratory, California Institute of Technology, Johns Hopkins University, Purdue University, Los Alamos National Laboratory, University of California Berkeley, University of California San Diego, Woods hole Oceanographic Institution, Australian National University, Cardiff University, Geological Survey of Norway or Monash University (just to name some of them). When optimized for the unique and particular needs of the Solid Earth community, such high performance-computing infrastructure certainly provides a key tool enabling

Recibido: 26 de marzo de 2012
Aceptado: 21 de agosto de 2012

Keywords:
High performance computing cluster; numerical modeling; Computational Fluid Dynamics.
Palabras clave:
Clúster de alto rendimiento; modelado numérico; Dinámica de Fluidos Computacional.

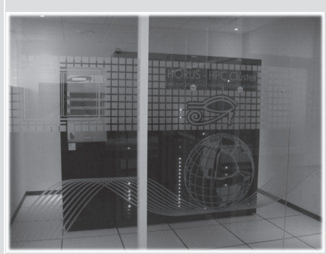
*Computational Geodynamics Laboratory. Centro de Geociencias. Universidad Nacional Autónoma de México. Mexico. Telephone number: (442) 238-1104, ext. 133; fax: (442) 5623-4129. E-mail: vlad@geociencias.unam.mx
**Solid Earth Dynamics Department. Institute of Geodynamics of the Romanian Academy. Romania.

rapid major advances in this exciting area of research (Computational Infrastructure for geodynamics or CIG, as it is presented in http://www.geodynamics.org). Numerical methods have now progressed to the point that numeric simulations have become a central part of modern Earth Sciences, in particular for geodynamics [7-11]. Such computational systems are structured specifically for the Solid Earth community's simulation needs, which include large number of computing cores, fast and reliable storage capacity and considerable amount of memory -everything configured in a system designed for long run-times. One of the key tenants of HPC is performance [12-15], and designing a High Performance Computational (HPC) solution tailored to a specific research field as the Solid Earth is often a challenge. The HPC system performance strongly depends on the software-hardware interaction, therefore, prior knowledge on how well specific parallelized software performs on different HPC architectures can weight significantly on choosing the final configuration [15, 16].

HPCC system configurations

In this paper are presented a series of performance test results from two different HPC systems: one low-end HPCC (Horus) with 300 cores (table 1) and 1,6 TFlops theoretical peak performance and one high-end HPCC (CyberDyn) with 1 344 cores and 11,7 TFlops theoretical peak performance (table 1). Horus uses CentOS 5,2 and as operating environment the open source Rocks Cluster Distribution-5.1 (http://www.rocksclusters.org). CyberDyn employs Bright Cluster Management (http://www.brightcomputing.com) and Scientific Linux (www.scientificlinux.org).

Table 1.
Comparison between HPCC Horus and HPCC Cyberdyn's configurations.

	HPCC Horus	HPCC CyberDyn
Master node		
Server model:	1 × Dell PE2970	2 × Dell R715 (failover configured)
Processor type:	AMD Opteron 2 × Dual core 2,8 GHz	AMD Opteron 2 × Twelve core 2,1 GHz
Memory:	8 GB RAM	68 GB RAM
Cards:	PERC 5i Integrated Raid Controller	PERC H200 Integrated Raid Controller Broadcom NetXtreme Dual port SFP + Direct Attach 10 GbE NIC
Computing nodes		
Server model:	38 × DELL Sc1435/PE1950	28 × DELL R815
Processor type:	AMD 2 × Quad core 2,6 GHz Intel 2 × Quad core 2,6 GHz	AMD 4 × Twelve core 2,1 GHz
Memory:	16 GB-32 GB/node	96 GB-128 GB/node
Network fabric:	1000T Ethernet (1 Gbit/sec)	InfiniBand QDR (40 Gbit/sec)
Network		
MPI trafic	HP Procurve 48G–2900 1000T-48 Port unmanaged high performance switch, 10 GbE and stacking capable	Qlogic 12300-BS01 36 port InfiniBand Quad Data rate
Management/IPMI	It is used the same switch for MPI traffic and cluster management	2 × Dell PowerConnect 6248–48 Port managed Layer 3, 10 GbE and stacking capable
Storage		
	2 × Direct Attached Storage arrays connected to the master node. Each one has a 15 TB RAID 5 volume	1 × Panasas 8 Series with 40 TB storage
Precision cooling (temperature 21 °C±2 °C, humidity 50 %±5 %)		
	Liebert 8 t	Liebert 2 t × 7 t
Inside view		
		

Apart from the number of computing cores, the main difference between the two HPC systems is the interconnect architecture. The HPC Horus system uses a centralized Gigabit Ethernet network for administrative traffic, data sharing (Network File System –NFS- or other protocols) and Message Passing Interface (MPI) or applications processing traffic. The second and larger HPC system CyberDyn uses two internal networks. The first (a Gigabit Ethernet network) is used for scheduling, node maintenance and basic logins, while the second internal network is QDR InfiniBand and is dedicated exclusively to computational parallel-MPI traffic. In order to benchmark the Earth's mantle flow in detail, large HPC facilities and parallelized codes are required. To benchmark the two HPC systems, it was used package CitcomS, which is widely used in the Solid Earth scientific community (www.geodynamics.org). CitcomS software is a Computational Fluid Dynamics (CFD) code based on finite element code and designed to solve thermal convection problems relevant to Earth's mantle [17, 18]. Written in C, the code runs on a variety of parallel processing computers, including shared and distributed memory platforms and is based on domain decomposition. This parallelized numeric code requires specific libraries -which implements the MPI standard. In order to accurately perform the comparative benchmark, on both HPCC systems, CitcomS is compiled using OpenMPI version 1.4.2. Also, on both HPC systems, Sun Grid Engine (SGE) as job scheduler was used. For these specific benchmarks, it was selected a series of different Earth's mantle convection problems, from the simple purely thermal convection to the more complex thermo-chemical convection problem. All numeric simulations were performed within full spherical and regional shell domains. They were used different mesh resolutions and computing cores, with the finer the mesh the higher the number of cores used. Each test was performed several times to ensure that consistent, repeatable and accurate results were obtained.

RESULTS

Below it is presented in details the benchmark results obtained on both HPCC systems. Due to the limitation at 300 cores for HPCC Horus, the results are comparable to only 192 computing cores for full-spherical models and 256 computing cores for regional models. On both HPC systems, it was performed a series of benchmarks using three different FEM simulations: two simple thermal convection problems (one as regional model and the other one as full-spherical model) and a more complex thermo-chemical simulation as full-spherical model.

In the case of purely thermal convection simulations, on both HPC systems, it was obtained similar performance for grid sizes limited to 129 nodes × 129 nodes × 65 nodes (figure 1-2). Increasing the mesh size and the number of computing cores, the HPCC CyberDyn starts outperforming the HPCC Horus because of the low-latency high-speed Quad Data Rate (QDR) network dedicated to MPI traffic.

A thermo-chemical simulation performed in a full-spherical shell represents the third comparative test between HPCC Horus and HPCC CyberDyn. Compared with the previous tests, these simulations involve a large number of particles (or tracers) in order to track the thermo-chemical changes inside the model. The tracers are generated pseudo-randomly, with a total number equal to tracers per element × total number of finite elements. In the simulations, they were used a number of 20 tracers per finite element, therefore, the total number of tracers varied from 1 million to 500 millions -depending on the model resolution. The benchmark results show that for complex numeric simulations HPCC CyberDyn performs better that Horus, because of the fast low-latency InfiniBand QDR network fabric and the high-performance Panasas storage. It was also found that, for specially high-resolution models, the maximum number of computing cores that offers the minimum wall time is around 384 (figure 3). Although Horus is slower than CyberDyn for these high-end FEM simulations, it is observable a continuous decreasing in wall time for almost all model resolutions and number of computing cores. This result demonstrates that HPCC Horus still has a real potential to expand to probably over 500 computing cores in the future.

DISCUSSION AND CONCLUSIONS

The high-speed Infiniband interconnect offers the possibility to exploit the full potential of large clusters and represents a key component that positively influences both scalability and performance on large HPC systems. Testing a CFD code specific for Earth Sciences, the HPC system Horus based on Gigabit Ethernet performed remarkably well compared with its counterpart CyberDyn -which is based on InfiniBand QDR fabric. However, HPC systems based on Gigabit Ethernet is still a quite popular cost-effective choice, but suitable for small (eventually medium-size) high-performance clusters running CFD codes specific for Earth Sciences. On the other hand, for medium and large HPC systems running Earth Sciences CFD codes, low-latency high-bandwidth as Infiniband fabric is highly recommended. Since presently the authors are moving towards high-resolution simulations for geodynamic predictions that require the same scale as observations (from several to thousands of kilometers), HPC facilities used in Earth Sciences should benefit from larger up-front investment in future systems that are based on high-speed interconnects.

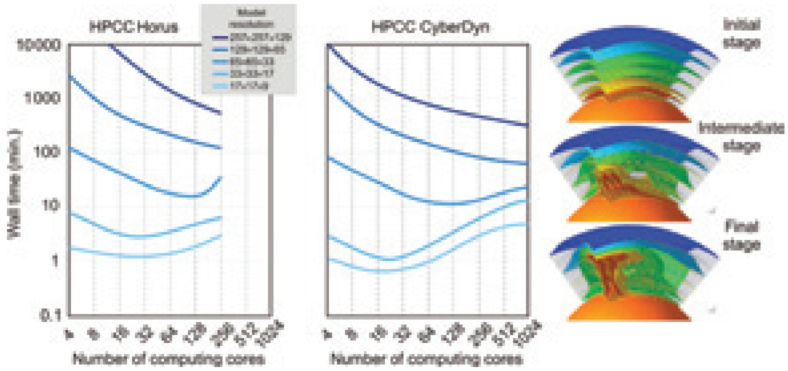


Figure 1. Comparison between benchmark results on HPCC Horus and HPCC CyberDyn for a purely thermal convection FEM simulation in a regional model, which shows the influence of mesh size on wall time as a function of number of processors. To the right is shown, as temperature isosurfaces, three different evolutionary stages of the thermal convection simulation (visualization performed with open source software OpenDX). Warm colors correspond to high temperature and cold colors represent low temperature inside the Earth's mantle. Orange sphere at the initial stage represents the Earth's iron core.

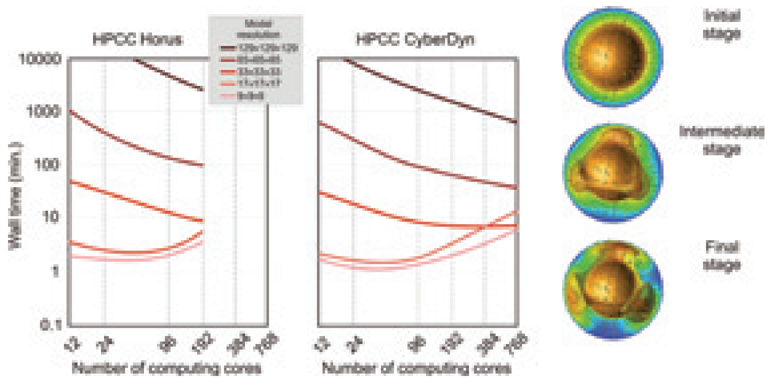


Figure 2. Comparison between benchmark results on HPCC Horus and HPCC CyberDyn for a purely thermal convection FEM simulation in a full-spherical model, which shows the influence of mesh size on wall time as a function of number of processors. To the right is shown, as temperature isosurfaces, three different evolutionary stages of the thermal convection simulation (visualization performed with OpenDX). Warm colors correspond to high temperature and cold colors represent low temperature inside the Earth's mantle. Orange sphere at the initial stage represents the Earth's iron core.

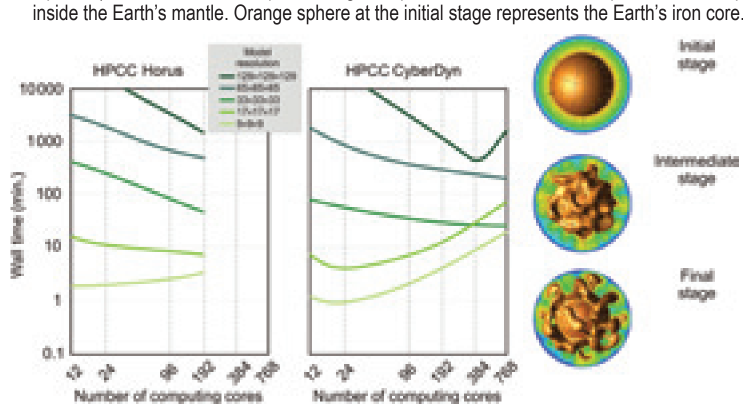


Figure 3. Comparison between benchmark results on HPCC Horus and HPCC CyberDyn for a thermo-chemical convection FEM simulation in a full-spherical model, which shows the influence of mesh size on wall time as a function of number of processors. To the right is shown, as temperature isosurfaces, three different evolutionary stages of the thermo-chemical convection simulation (visualization performed with OpenDX). Warm colors correspond to high temperature and cold colors represent low temperature inside the Earth's mantle. Orange sphere at the initial stage represents the Earth's iron core.

Although the powerful CyberDyn supercomputer has the performance delivered on the network side using high-speed InfiniBand connectivity, for Solid Earth scientists striving toward reduced fabric costs and simplified use -InfiniBand could pose several challenges. However, with the arrival of the recent 10 Gigabit Ethernet (10GbE) network adapters, HPC cluster users can easily overcome fabric challenges and continue to push the Solid Earth Science envelope. Moreover, using Ethernet connectivity as a unified fabric for cluster interconnects and storage can help lowering the total cost of ownership, by significantly reducing the number of switches and cables required. Using Ethernet connectivity for HPCC it can now be provided high performance, efficiency and scalability, enabling Earth Sciences research laboratories to obtain the low latency that HPC clusters need for CFD ultra-high resolution numeric simulations. In the near future, in order to acquire full advantage of the new HPC architectures, it will be required more than simple adaptation of existing algorithms to new hardware. Actually, it will require the design a new set of parallelized numerical algorithms that enables fast and efficient implementation.

While high performance computing has entered recently in a period of rapid change with the emergence of multi-core and many-core architectures, there is a certain amount of uncertainty on which, among all the possible modern HPC architectures, best fits the computational needs of Earth Sciences. In this paper, it is hoped to shed some light on all these uncertainties and to provide a solid base for choosing the appropriate HPC architecture for future and modern Solid Earth Sciences research labs.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the use of the HPCC CyberDyn for all numeric simulations at the Institute of Geodynamic of the Romanian Academy. This research has been conducted through the CYBERDYN project (POS CCE O212_ID 593).

REFERENCES

[1] Gerya, T. (2009). *Introduction to Numerical Geodynamic Modelling*. Cambridge University Press. Cambridge, UK.

[2] Ismail-Zadeh, A. and Tackley, P. (2010). *Computational Methods for geodynamics.* Cambridge University Press.

[3] Kaus, B. J. P., Schmid, D. W. and Schmalholz, S. M. (2011). Numerical modelling in structural geology. *Journal of Structural Geology*. Available online: http://www.gfd.geophys.ethz.ch/˜kausb/kss10_Review_MS_JSG.pdf. Consulted 2011.

[4] Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (1993). Numerical Recipes in C: *The Art of Scientific Computing*. 2nd edition. Cambridge University Press. Cambridge, UK.

[5] Hughes, T. J. R. (2000). *The finite element method*. Dover Publications.

[6] Spiegelman, M. (2004). *Myths and Methods in Modeling.* Columbia University Course lecture notes. Available online: http://www.ldeo.columbia.edu/~mspieg/mmm/course.pdf. Consulted 2011.

[7] Moresi, L. N., Dufour, F. and Muhlhaus, H.-B. (2003). A Lagrangian integration point finite element method for large deformation modeling of viscoelastic geomaterials. *Journal of Computational Physics 184*: pp. 476-497.

[8] Choi, E., Gurnis, M., Kientz, S. and Stark, C. (2010). *SNAC, User Manual, Version 1.2*. Available online: www.geodynamics.org/cig/software/snac/snac.pdf. Consulted 2011.

[9] Choi, E., Lavier, L. and Gurnis, M. (2008). Thermomechanics of mid-ocean ridge segmentation. *Physics of the Earth and Planetary Interiors 171*: pp. 374-386.

[10] Landry, W., Hodkinson, L. and Kientz, S. (2011). *GALE, User Manual, Version 2.0.0*. Available online: www.geodynamics.org/cig/software/gale/gale.pdf. Consulted 2011.

[11] Tan, E., Gurnis, M., Armendariz, L., Strand, L. and Kientz, S. (2012). *CitcomS, User Manual, Version 3.2.0*. Available online: www.geodynamics.org/cig/software/citcoms/citcoms.pdf. Consulted 2012.

[12] Terry, P., Shan, A. and Huttunen, P. (2004). Improving Application Performance on HPC Systems with Process Synchronization. *Linux Journal 127*: pp. 68, 70, 72-74.

[13] Borill, J., Oliker, L., Shalf, J., Shan, H. and Uselton, A. (2009). HPC global file system performance analysis using a scientific application derived benchmark. *Parallel Computing 35*(6): pp. 358-373.

[14] Yi, H. (2011). Performance evaluation of GAIA supercomputer using NPB multi-zone benchmark. *Computer Physics Communications 182*(1): pp. 263-265.

[15] Jin, H., Jespersen, D., Mehrotra, P., Biswas, R., Huang, L. and Chapman, B. (2011). High performance computing using MPI and OpenMP on multi-core parallel systems. *Parallel Computing 37*(9): pp. 562-575.

[16] Gray, A., Bethune, I., Kenway, R., Smith, L., Guest, M., Kitchen, C., Calleja, P., Korzynski, A., Rankin, S., Ashworth, M., Porter, A., Todorov, I., Plummer, M., Jones, E., Steenman-Clark, L., Ralston, B. and Laughton, C. (2012). Mapping application performance to HPC architecture. *Computer Physics Communications 183*(3): pp. 520-529.

[17] Zhong, S., Zuber, M. T., Moresi, L. N. and Gurnis, M. (2000). The role of temperature-dependent viscosity and surface plates in spherical shell models of mantle convection. *Journal of Geophysical Research 105*: pp. 11 063-11 082.

[18] Tan, E., Choi, E., Thoutireddy, P., Gurnis, M. and Aivazis, M. (2006). GeoFramework: Coupling multiple models of mantle convection within a computational framework. Geochemistry, *Geophysics, Geosystems* 7: p. Q06001. doi:10.1029/2005GC001155