Seminary

High-Performance Computing (HPC) and its applications

Master in Space Applications for Emergency Early Warning and Response

Author: Zader, Pablo Javier
Professor: Dr. Marcelo Scavuzzo
Córdoba, Argentina
October 2013
# Contents

Introduction........................................................................................................................................... 4  
Definitions............................................................................................................................................... 4  
Background............................................................................................................................................... 7  
Performance ranking of supercomputers on the world............................................................................. 8  
Energy-Efficient ranking of supercomputers on the world....................................................................... 11  
Ranking of 50 fastest Supercomputers in Latin America....................................................................... 13  
  Argentina in LARTop50....................................................................................................................... 13  
Areas of applications............................................................................................................................... 14  
Use case .................................................................................................................................................. 14  
  Optimizing grid computing configuration and scheduling for geospatial analysis: An example  
  with interpolating DEM....................................................................................................................... 14  
Reference.................................................................................................................................................. 18  

Abstract

Supercomputing also called HPC (acronym for, High-Performance Computing) is required for treatment with data and/or computations intensive for a given problem.

Today's supercomputers are being used to address some of the major problems of the scientific world but also for industry world. This problems include computer simulations that enable scientists to discover new phenomena and test hypotheses require massive amounts of performance.

The aim is then to disaggregate the problems into tasks that are sent to one or more computer clusters to be executed. In this context, appears also the concepts of distributed and parallel computing.

The objective is to bibliographic search to learn about of HPC today in the world, several supercomputer in the world and which applications are run on HPC environments.

Keywords: HPC, grid computing, parallel computing, distributed computing, supercomputing, image processing, cluster
Introduction

Definitions

Cluster
A group of interconnected, whole computers working together as a unified computing resource that can create the illusion of being one machine. Each computer in a cluster is typically referred to as a node (William Stallings, 2010). Lists four benefits that can be achieved with clustering:

- Scalability
- High availability
- High-performance
- Load Balancing

Cluster Configurations (William Stallings, 2010)

CPU and/or core
CPU: (Central Processing Unit) is the hardware inside the computer running the basic arithmetic operations and logical. Core: each CPU integrated into a single processor.

Distributed computing
A distributed system is a software system in which components located on networked computers communicate and coordinate their actions by passing messages\(^1\). In this type computing, the processors have its private memory; the information between processors is exchanged via passing message.

Distributed Computing

\(^{1}\)
http://en.wikipedia.org/wiki/Distributed_computing
**Grid computing**
A group of cluster and/or computers acting as one. Are generally not located in the same place. Use the Internet to communicate unlike supercomputers where the core are connected by high-speed bus (i.e. Infiniband).

Example

SETI@home Project
One of the more famous examples of grid computing projects (and distributed computing) is run by SETI, also known as Search for Extra-Terrestrial Intelligence. The application looks for radio signals or other forms of communication in space, in an effort to prove the existence of extra-terrestrial intelligence.

**High-Performance Computing (HPC)**
A research area dealing with supercomputers and the software that runs on supercomputers. The emphasis is on scientific applications, which may involve heavy use of vector and matrix computation, and parallel algorithms (William Stallings, 2010).

**FLOPS**
The computer performance is measured in FLoating-point Operations Per Second (# of operations performed each second).

**Mainframes**
A mainframe computer is a large, powerful computer that handles the processing for many users simultaneously (up to several hundred users). The name mainframe originated after minicomputers appeared in the 1960’s to distinguish the larger systems from the smaller minicomputers (Beach, Thomas E, 2012).
The term mainframe is used for the larger, most powerful computers other than supercomputers. Typical characteristics of a mainframe are that it supports a large database, has elaborate I/O hardware, and is used in a central data processing facility.

**MPI, OpenMPI**
MPI stands for the Message Passing Interface. Written by the MPI Forum (a large committee comprising of a cross-section between industry and research representatives), MPI is a standardized API typically used for parallel and/or distributed computing.
Open MPI is an open source, freely available implementation of the MPI documents. The Open MPI software achieves high performance; the Open MPI project is quite receptive to community input.

**Parallel computing**
Is the art of dividing a program into several smaller parts in order to run them simultaneously. There are two common types of parallelization:

- Data parallelism is to subdivide the input data set of the program and run as many instances of the program (each on one processor or core) as subdivisions of data we have.
- Task parallelism is to subdivide the main task into subtasks disjoint, which runs on the same or a different set of data. Each subtask is sent to a processor or core.

Parallel computers can be roughly classified according to the level at which the hardware

2 [http://setiathome.berkeley.edu/](http://setiathome.berkeley.edu/)
3 [http://www.open-mpi.org/faq/?category=general](http://www.open-mpi.org/faq/?category=general)
supports parallelism, with multi-core and multi-processor computers having multiple processing elements within a single machine, while clusters, MPPs, and grids use multiple computers to work on the same task\(^4\). In this type computing, the processors shared the memory; information is exchanged between processors to the read the memory.

**Supercomputer**
Computer equipment is able to calculus much more than desktop computers and are used for specific purposes.
Example
The Cray-1 was a supercomputer designed, manufactured and marketed by Cray Research. The first Cray-1 system was installed at Los Alamos National Laboratory in 1976 and it went on to become one of the best known and most successful supercomputers in history. The Cray-1's architect was Seymour Cray, the chief engineer was Cray Research co-founder Lester Davis.

**Speedup**
Is the measure that indicates how much a program in parallel is faster than the same program in sequential and is defined by the following formula:

\[
speedup = \frac{\text{time to execute program on a single processor}}{\text{time to execute program on N parallel processors}}
\]

\(^4\) [http://en.wikipedia.org/wiki/Parallel_computing](http://en.wikipedia.org/wiki/Parallel_computing)
Background

The first supercomputer was created in the 60s, the CDC 6600, manufactured by Control Data Corporation with performance of about 1 MFLOPS and 10 parallel unit processing, used for research in nuclear physics was the world's fastest computer from 1964 to 1969.

In the 70's the CDC 7600 supercomputer used only a few processors and achieves 10MFLOPS on hand-compiled code, with a peak of 36 MFLOPS. Then in the 90's machines with thousands of processors appeared, parallel supercomputers with tens of thousands of “off-the-shelf” processors were the norm.

Today, commodity PC’s which you can purchase have more than one core i.e. dual core, quad core processors, for example Intel core i7, AMD Opteron, what was a supercomputer in those years, nowadays we have in our home through a PC, notebook or a Tablet also.

The growth of this technology has been marked by Moore's Law, which was established by Gordon Moore in 1965 and states that the number of transistors on integrated circuits doubles approximately every two years. The next figure show, plot of CPU transistor counts against dates of introduction. Nota: logarithmic vertical scale; the line corresponds to exponential growth with transistor count doubling every two years. (Figure 1).


Performance ranking of supercomputers on the world

The Performance is measured in FLoating point Operations Per Second (FLOPS or flop/s) . In the personal computers (PC) is in the range of 2 – 13 Gflop/s⁶ (Giga FLOPS) in contrast to the Tianhe-2, a supercomputer developed by China’s National University of Defense Technology, is the world’s new No. 1 system with a performance of 33.86 petaflop/s on the Linpack[8] benchmark⁷.

Resume of Tianhe-2:

<table>
<thead>
<tr>
<th>Location</th>
<th>Guangzhou, China</th>
</tr>
</thead>
<tbody>
<tr>
<td>Architecture</td>
<td>Intel Xeon E5, Xeon Phi</td>
</tr>
<tr>
<td>Power</td>
<td>17.6 MW (24 MW with cooling)</td>
</tr>
<tr>
<td>Operating system</td>
<td>Kylin Linux</td>
</tr>
<tr>
<td>Memory</td>
<td>1,375 TiB (1,000 TiB CPU and 375 TiB Coprocessor)</td>
</tr>
<tr>
<td>Storage</td>
<td>12.4 PB</td>
</tr>
<tr>
<td>Speed</td>
<td>33.86 PFLOPS</td>
</tr>
<tr>
<td>Cost</td>
<td>2.4 billion Yuan (390 million USD)</td>
</tr>
<tr>
<td>Purpose</td>
<td>Research and education</td>
</tr>
<tr>
<td>Dimensions</td>
<td>720 m²</td>
</tr>
</tbody>
</table>


The TOP500⁸ project ranks and details the 500 most powerful known supercomputers in the world, published every six months.

The main objective of the TOP500 list is to provide a ranked list of general purpose systems that are in common use for high end applications. The authors of the Top500 reserve the right to independently verify submitted LINPACK results, and exclude systems from the list which are not valid or not general purpose in nature. By general purpose system we mean that the computer system must be able to be used to solve a range of scientific problems. Any system designed specifically to solve the LINPACK benchmark problem or have as its major purpose the goal of a high Top500 ranking will be disqualified[3].

The following figure (figure 2) can display the first 8 supercomputers in the world (the reader can see the complete list by entering the web site) which is obtened from top500 website as a result of a query to the database that it provides and ordered bye value Rmax. The list corresponds to June, 2013.

---

⁶ Most microprocessors today can do 4 FLOPs per clock cycle. Therefore a 2.5-GHz processor has a theoretical performance of 10 billion FLOPs = 10 GFLOPs. Calculo de GFLOPs = cores × clock × FLOPs /cycle

⁷ Reflect the performance of a dedicated system for solving a dense system of linear equations.

⁸ [http://www.top500.org/](http://www.top500.org/)
The meaning of the columns we describe below:

- Rank - Position within the TOP500 ranking
- System - Type indicated by manufacturer or vendor
- Site - Location and country
- Cores - Number of processors (Cores)
- Rmax - Maximal LINPACK performance achieved
- Rpeak - Theoretical peak performance
- Power(kW) - maximum energy

Other statistic we can see is Performance Development (in Gflop/s) (Figure 3), where we see three curves: the orange curve represents the top of the supercomputers between 1993 and 2013, the yellow curve gives the position 500 of the supercomputer between 1993 and 2013, and the green curve gives the sum of the performance of supercomputers of the positions 1 and 500 between 1993 and 2013. The most current data up to June 2013, where the top supercomputer yields about 33,862,700 Gflop/s (33 PetaFlop/s) and the last of the list gives 96,619 Gflop/s.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Site</th>
<th>System</th>
<th>Cores</th>
<th>Rmax (TFlop/s)</th>
<th>Rpeak (TFlop/s)</th>
<th>Power (kW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>National University of Defense Technology, China</td>
<td>Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.20GHz, TH Express-2, Intel Xeon Phi 3151P</td>
<td>3,120,000</td>
<td>33,862.7</td>
<td>54,902.4</td>
<td>17.806</td>
</tr>
<tr>
<td>2</td>
<td>DOE/SC/Oak Ridge National Laboratory, United States</td>
<td>Titan - Cray XK7, Opteron 6274 16C 2.20GHz, Cray Gemini Interconnect, NVIDIA K20x</td>
<td>560,640</td>
<td>17,590.0</td>
<td>27,112.5</td>
<td>8,200</td>
</tr>
<tr>
<td>3</td>
<td>DOE/ANSA/LLNL, United States</td>
<td>Sequoia - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM</td>
<td>1,572,684</td>
<td>17,173.2</td>
<td>20,132.7</td>
<td>7,989</td>
</tr>
<tr>
<td>4</td>
<td>RIKEN Advanced Institute for Computational Science (AICS), Japan</td>
<td>K computer, SPARC64 VIIIfx 2.0GHz, Tofu Interconnect</td>
<td>705,024</td>
<td>10,510.0</td>
<td>11,200.4</td>
<td>12,660</td>
</tr>
<tr>
<td>5</td>
<td>DOE/SC/Argonne National Laboratory, United States</td>
<td>Mira - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM</td>
<td>788,432</td>
<td>8,568.3</td>
<td>10,066.3</td>
<td>3,945</td>
</tr>
<tr>
<td>6</td>
<td>Texas Advanced Computing Center/Univ. of Texas, United States</td>
<td>Stampede - PowerEdge C8220, Xeon E5-2680v2 8C 2.70GHz, Infiniband FDR, Intel Xeon Phi 5E10P</td>
<td>482,462</td>
<td>5,168.1</td>
<td>6,520.1</td>
<td>4,510</td>
</tr>
<tr>
<td>7</td>
<td>Forschungszentrum Juelich (FZJ), Germany</td>
<td>JUQUEEN - BlueGene/Q, Power BQC 16C 1.60GHz, Custom Interconnect</td>
<td>456,752</td>
<td>5,008.9</td>
<td>5,872.0</td>
<td>2,301</td>
</tr>
<tr>
<td>8</td>
<td>DOE/ANSA/LLNL, United States</td>
<td>Vulcan - BlueGene/Q, Power BQC 16C 1.60GHz, Custom Interconnect IBM</td>
<td>393,216</td>
<td>4,293.3</td>
<td>5,033.2</td>
<td>1,972</td>
</tr>
</tbody>
</table>

Figure 2. Top supercomputers http://www.top500.org/lists/2013/06/
and finally we can see the Projected Performance Development (figure 3).
Energy-Efficient ranking of supercomputers on the world

The Green500[4] list, ranks the top 500 supercomputers in the world by energy efficiency. Supercomputers are compared by performance-per-watt. The following figure (figure 4) can display the first 10 supercomputers in the world (the reader can see the complete list by entering the website) which is obtained from green500 website as a result of a query to the database that it provides and ordered by value MFLOPS/W. The list corresponds to June, 2013.

![Figure 4. Ranking supercomputer by MFLOPS/W](image)

The first supercomputer of the list EURORA has the following characteristic:

**Architecture:** 1 rack  
**Model:** Eurora prototype  
**Processor Type:**  
- Intel Xeon CPU E5-2658 @ 2.10GHz  
- Intel Xeon CPU E5-2687W @ 3.10GHz  
**Accelerator Type:**  
- Nvidia Tesla K20s  
- Intel Xeon-Phi 5120D  
**Computing Cores:** 1024  
**Computing Nodes:** 64  
**RAM:** 16 GByte DDR3 1600MHz per node(5 nodes with 32 GByte)  
**Internal Network:**  
- 1 FPGA (Altera Stratix V) per node  
- IB QDR interconnect  
- 3D Torus interconnect  
**Disk Space:** 90 GByte SSD per node  
**Sustained Performance:** 3,150 MFlop/w
The Graph500[5] list, ranks the top 500 supercomputers in the world by data intensive load. Supercomputers are compared by GTEPS (10^9 Traversed edges per second). The following figure (figure 5) can display the first 7 supercomputers in the world (the reader can see the complete list by entering the web site) which is obtained from graph500 website as a result of a query to the database that it provides and ordered by value GTEPS. The list corresponds to June, 2013.

![Figure 5. Ranking supercomputer by GTEPS](image)

The first supercomputer of the list Sequoia has the following more important characteristic:\(^9\):

**IBM Sequoia**

- **Operators**: LLNL (Lawrence Livermore National Laboratory)
- **Location**: Livermore, Alameda County, SFBA, Northern California, United States
- **Power**: 7.9 MW
- **Operating system**: CNK operating system
- **Space**: 3,000 square feet (280 m2)
- **Memory**: 1.5 PiB
- **Speed**: 16.32 PFLOPS
- **Purpose**: NW&UN, astronomy, energy, human genome, and climate change

Ranking of 50 fastest Supercomputers in Latin America

The Top 50 for Latin America project [LARTop50] was born in 2011 at the National University of San Luis, Argentina and is integrated by researchers from several countries in Latin America. LARTop50 aims to create and maintain the list with the statistics of the 50 supercomputing systems with better computational performance in Latin America. Its purpose is to provide reliable basis for tracking and detecting trends in high-performance computing[6].

The following figure (figure 6) can display the first 5 supercomputers in Latin America (the reader can see the complete list by entering the web site) which is obtained from LARTop50 website as a result of a query to the database that it provides and ordered by value Rmax in GFLOPS. The list corresponds to July, 2013.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Name</th>
<th>Site</th>
<th>Country</th>
<th>Vendor</th>
<th>Model</th>
<th>Processor</th>
<th>Nodes</th>
<th>Cores</th>
<th>Rmax(GFlop)</th>
<th>Rpeak</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MIZTU</td>
<td>Universidad Nacional Autónoma de México - UNAM</td>
<td>Mexico</td>
<td>Hewlett Packard</td>
<td>HP Cluster Platform 3000SL</td>
<td>Intel E5-2670</td>
<td>332</td>
<td>5280</td>
<td>80429</td>
<td>12058</td>
</tr>
<tr>
<td>2</td>
<td>CENAPAD-SP</td>
<td>IBM P750 Centro Nacional Desempeño - CENAPAD-SP</td>
<td>Brazil</td>
<td>IBM P750</td>
<td>PowerPC_POWER7</td>
<td>40</td>
<td>1260</td>
<td>27020</td>
<td>3700</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>LEVQUE</td>
<td>National Laboratory for High Performance Computing - NHPE</td>
<td>Chile</td>
<td>IBM</td>
<td>iDataPlex</td>
<td>Xeon Nehalem X5550</td>
<td>66</td>
<td>5280</td>
<td>4847</td>
<td>5639</td>
</tr>
<tr>
<td>4</td>
<td>MEDUSA</td>
<td>Centro de Investigaciones en Optica, A.C. - CIO</td>
<td>Mexico</td>
<td>Lufac</td>
<td>LCI-4C0</td>
<td>Xeon X5675</td>
<td>38</td>
<td>432</td>
<td>4842</td>
<td>5305</td>
</tr>
<tr>
<td>5</td>
<td>ISAC</td>
<td>Comisión Nacional de Energía Atómica - GTIC/DCAP/DCAP</td>
<td>Argentina</td>
<td>SIASA (INTEL)</td>
<td>Heterogénero</td>
<td>Xeon X2220/Xeon E5-2670</td>
<td>162</td>
<td>644</td>
<td>2922</td>
<td>5810</td>
</tr>
</tbody>
</table>

Figure 6. Top 50 Latin American Supercomputer

Argentina in LARTop50

In the table below we can see the positions occupied by Argentina in the top 50 supercomputers in America.

<table>
<thead>
<tr>
<th>RANK</th>
<th>Name</th>
<th>Segmen t</th>
<th>Name</th>
<th>Vendor</th>
<th>Year</th>
<th>Number of nodes</th>
<th>Processor</th>
<th>Operating system</th>
<th>Number of cores</th>
<th>Rmax</th>
<th>Rpeak</th>
<th>MPI library</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>CNEA-GTIC-DCAP</td>
<td>Governmental</td>
<td>ISAAC</td>
<td>SIASA (INTEL)</td>
<td>2010</td>
<td>162</td>
<td>XEON X3220/ES420</td>
<td>2.618-274.18.16l</td>
<td>5.centos.plu s (5.1)</td>
<td>644</td>
<td>2922,00</td>
<td>5610,00</td>
</tr>
<tr>
<td>10</td>
<td>Rosario</td>
<td>Investigati on</td>
<td>PILUSO</td>
<td>DELL-AMD</td>
<td>2013</td>
<td>22</td>
<td>AMD OPTERON 6382 SE</td>
<td>CENTOS 6.3</td>
<td>704</td>
<td>1080,00</td>
<td>7321,60</td>
<td>Intel MPI</td>
</tr>
<tr>
<td>12</td>
<td>ICB-ITIC</td>
<td>Academic</td>
<td>ICB-ITIC</td>
<td>Vertex</td>
<td>2012</td>
<td>2</td>
<td>AMD OPTERON 6372</td>
<td>Linux</td>
<td>128</td>
<td>225,60</td>
<td>1074,00</td>
<td>openmpi</td>
</tr>
<tr>
<td>13</td>
<td>Unidad de Cálculo en Física y Química Teórica - UNICAPYOT - INIFTA</td>
<td>Academic</td>
<td>CALCULO</td>
<td>Intel</td>
<td>2010</td>
<td>30</td>
<td>Intel Core 2 Duo</td>
<td>Linux Rocks 6.0</td>
<td>128</td>
<td>78,60</td>
<td>1361,00</td>
<td>OpenMPI</td>
</tr>
</tbody>
</table>
Areas of applications

• Full simulation of engineering systems
• Full simulation of biological systems
• Astrophysics
• Materials science
• Bio-informatics, proteomics, pharmaco-genetics
• Scientifically accurate 3D functional models of the human body
• Biodiversity and biocomplexity
• Climate and Atmospheric Research
• Energy
• Digital libraries for science and engineering

Use case

Optimizing grid computing configuration and scheduling for geospatial analysis: An example with interpolating DEM (Qunying Huang, Chaowei Yang, 2010)

Problems

• Many geographic analyses are very time-consuming and do not scale well when large datasets are involved
• The interpolation of DEMs for large geographic areas could become a problem in practical application, especially for web applications such as terrain visualization

Solutions

• HPC and parallel computing, such as grid computing
• Configuring and scheduling a high performance grid computing platform to improve the performance of geographic analyses through a systematic study on how the number of cores, processors, grid nodes, different network connections and concurrent request impact the speedup of geospatial analyses.
• Condor, a grid middleware, is used to schedule the DEM interpolation tasks for different grid configurations

Methodology

• Use different grid computing configurations to support the process of DEM interpolation
• DEM domain is decomposed into subdomains for balances the workload on the grid computing platform
• Uniform grid decomposition of the domain is used to partition an entire DEM into several subdomains with each having the same size.
• The Condor middleware is utilized to schedule the DEM interpolation tasks
• Introduces grid computing architecture and methodologies used for fast interpolation of high resolution DEMs.
• comparative performance of grid computing environments with different number of CPU cores and CPUs, different network connections of computers, and concurrent requests.

Experiment design
The next graphic show a grid computing environment where is made of test:

Joint Center of Intelligent Spatial Computing (CISC) at George Mason University
Among the most important tests that are performed to determine the optimal configuration of the grid, below mentioned three:

- Scalability analysis of parallel processing systematic:
  We considerate the executing of algorithm DEM interpolate on cluster HPC homogeneous and getting the resulting time of execution in 1, 2, 4, ..., 180 and 200 cores.
we conclude that the choice of between 60 and 70 cores are needed to significantly reduce processing time by 1084 seconds to 50 seconds.

- **Homogeneous grid vs. heterogeneous grid:**
  Time of process sequential DEM interpolate algorithms is 1084 s. After running same algorithms in 2, 4, 8, 12, 16 and 20 cores in two different grids and getting the result “speedup” we can compare the results as in the fig 7.

![Fig. 7. Speedup of DEM interpolation by different number of CPUs relative to one CPU on both homogeneous and heterogeneous grid platform.](image)

In the previous graph, we note that the best gain in acceleration is achieved with 12 cores for a heterogeneous grid while with 20 cores on a homogeneous grid.

- **Multi-core and multiprocessors performance analysis**
  Decomposed the whole DEM into 4, 16, 36, 60 and 96 subdomains and executed on the five groups of grid nodes of the homogeneous grid pool.
  It can be seen that the group of 4 cores on different processors (G44) has the best performance when running the IDW algorithm, while 4 cores on one processor (G11) have the worst performance. This result illustrates the issue of memory contention of multi-cores. For one core, it uses all L2 cache but for four cores, they share a single L2 cache so that every core in the four-core system has been allocated only 1/4 available cache compared to the one core system.

![Fig. 9. Execution time for all jobs by five groups of cores in HPC pool (execution time is the time HPC pool uses to dispatch, run and integrate the interpolating core performance and G44 has the best performance).](image)

In the previous graph, we note that the best run time is achieved with 4 processors and 1 core per processor.
Conclusions

HPC may be the only way to achieve computational goals in a given amount of time.
Size: Many problems that are interesting to scientists and engineers cannot fit on a PC usually because they need more than a few GB of RAM, or more than a few hundred GB of disk.
Speed: Many problems that are interesting to scientists and engineers would take a very long time to run on a PC: months or even years; but a problem that would take a month on a PC might only take a few hours on a Supercomputer.
Expensive solution: it is better to join forces.
Reference

    DESIGNING FOR PERFORMANCE EIGHTH EDITION
    November 17, 2012.
    for geospatial analysis: An example with interpolating DEM