Advanced Computing National Collaboratory HPC Infrastructure, Kabré

Infraestructura de HPC del Colaboratorio Nacional de Computación Avanzada, Kabré

Isaac Eduardo Gómez-Sánchez[1], Jean Carlo Umaña-Jiménez[2], Melissa Arce-Montero[3]

Gómez-Sánchez, I; Umaña-Jiménez, J; Arce-Montero, M. Advanced Computing National Collaboratory HPC Infrastructure, Kabré. *Tecnología en Marcha*. Edición especial 2020. 6th Latin America High Performance Computing Conference (CARLA). Pág 112-117.

https://doi.org/10.18845/tm.v33i5.5087

1 Electrical engineering student, Costa Rica National Center for High Technology. Costa Rica. E-mail: isaac.gomez@ucr.ac.cr.
https://orcid.org/0000-0001-6991-2460
2 Computer engineer, Costa Rica National Center for High Technology. Costa Rica. E-mail: jumana@cenat.ac.cr.
https://orcid.org/0000-0003-0857-6007
3 Electrical engineering student, Costa Rica National Center for High Technology. Costa Rica. E-mail: melissa.arcemontero@ucr.ac.cr
https://orcid.org/0000-0001-7553-6369

Keywords

Cluster; high performance computing; parallel computing.

Abstract

Kabré supercomputer is a High Performance Computing (HPC) cluster that provides tools that are needed to execute specific projects. In this paper, we present the infrastructure of such supercomputer settled at Advanced Computing National Collaboratory (CNCA), how its structure is defined by user needs and some recommendations in order to enhance the use and development of an HPC cluster infrastructure for scientific research purposes.

Palabras clave

Cluster; computación de alto rendimiento; computación paralela.

Resumen

La supercomputadora Kabré es un clúster de computación de alto rendimiento (HPC) que proporciona las herramientas necesarias para ejecutar proyectos específicos. En este documento, presentamos la infraestructura de dicha supercomputadora establecida en Colaboratorio Nacional de Computación Avanzada (CNCA), cómo se define su estructura según las necesidades del usuario y algunas recomendaciones para mejorar el uso y el desarrollo de una infraestructura de clúster HPC para fines de investigación científica.

Introduction

In science, throughout history, existed many problems which meant a significant and repetitive work in order to prove a theorem, validate a method, or simply to achieve a specific result. With the birth of HPC, many research centers saw potential in parallelization as a tool that speedup repetitive, mechanical and hard to compute tasks. Examples of these are presented in [9] with the use of TensorFlow for image recognition, [8] with a comparison of three different genome assemblers (Velvet, ABySS and SOAPdenovo) implementing de Bruijin graphs, [7] with an analysis and design of an LP and tremor location application, based on amplitude decay for seismic studies, and many others. These efforts wouldn't be more than that without a computational infrastructure that allows such power. The setup of a cluster infrastructure is required in order to solve such issues. For instance [1], [5] and [11] present an infrastructure focused on solving this requirement. But, as said in [11], such a cluster is necessarily defined by the users requirements, as defining it without demand, could be an expensive way of planning.

In Costa Rica, those efforts are not left behind. In Costa Rica National Center for High Technology (CeNAT) a computational cluster is in constant development. In this paper, we show how the users requirements shaped such infrastructure, its hardware constitution to fulfill those requirements and some tools that ease the administration, development and use of this one.

Infrastructure

Kabré supercomputer, cluster of CeNAT, is an HPC platform created to provide users with tools for data analysis, modeling and visualization. It has a lot of compute servers/nodes networked together that work in parallel raising the processing speed to give high performance computing.

In table 1 can be seen the current devices that conforms Kabré. Such devices has four main areas of specialization: simulation, machine learning, big data and bioinformatics. Use examples in these areas are presented in [2], [3], [4], [6], [7] and [8]. Below is shown a brief history of the origin of each area, as well as the hardware and software infrastructure that conforms Kabré, and some of its respective trends.

Hardware infrastructure

From the beginning, a focus on simulation to be the main area was sought. Xeon Phi were used, which has a large number of cores that are not very strong, but are fast enough to facilitate the execution of parallel projects. Currently, it is the strongest area of the cluster with 32 nodes. Over time, after listening to users and knowing which projects were they working on, different aspects were polished to create an infrastructure that adapts to their needs. That's why the second area arose, machine learning. There was a need of buying equipment with graphics cards capable of executing signal analysis projects. The graphic cards used are the NVIDIA Tesla K40 and the V100. These facilitated the execution of machine learning and deep learning programs. Also new projects arose with the need to run Big Data projects, therefore Data Science appeared as a new area. It works with more data and information, so another infrastructure was built with the necessary characteristics to facilitate the execution of these projects.

Because of its current faculties, the cluster is used by many users from different science areas. Currently it has at least 266 users. The main areas in which the cluster is used are presented in table 1. Most of the users that are presented in the table come from the public universities of Costa Rica, associated to Consejo Nacional de Rectores (CONARE). With such universities as well as with other institutions like Universidad de la República in Uruguay, or University of Chicago, there are about 270 different institutions around the world that uses this cluster. Also, Kabré is connected to RedCLARA, which interconnect many institutions along the globe with a 2.5Gbps link.

Software infrastructure

All these hardware would be useless without a software platform that ease the user experience and allow the cluster infrastructure at high level. For this reason, Kabré use a software infrastructure that optimize, and enables the use of each of the aforementioned areas of specialization. The main software platform used is Slurm. As said in [10]: "Slurm is an open source, fault-tolerant, and highly scalable cluster management and job scheduling system for large and small Linux clusters". Something that experience showed is the necessity of managing the resources available in a fair and smart way. With this tool, users jobs can be scheduled using weight paradigms, setup a maximum time for each job and, of course, distribute the resources among the users.

Other tools used in Kabré are Spark and Hadoop. These two platforms are frameworks which ease the big data work done with the cluster. The Spark platform can be used with Scala, Python, Java, R programming languages. The main difference among these is the algorithms for processing data: while Hadoop use MapReduce, Spark implements its own way of processing.

Trends

New necessities appeared with time. One of them is to increase the cluster infrastructure. It's expected to buy servers for the most used and growing areas: 1 node for Bioinformatics, 1 for Big Data and 2 for Machine Learning. Technical specifications about Bioinformatics node are brought in table 1, while the extra nodes of the other 2 areas are intended to keep with similar technical specifications than originals. Also, with the desire of ease the management of all the infrastructure, it's expected to implement new software tools such as CHEF to manage server configuration details and Spack for package management. In order to improve performance and ease the cluster management, new methods, tools and equipment are constantly researched and implemented.

Recommendations

When talking about the creation of a data center, there are many variables to take into account, and many times these are ignored. For example, the conditions of the room where the equipment will be housed must have adequate space to be able to place the racks with their respective servers and leaving space to work properly. Another very important factor is the electrical connection, making sure it's enough to support the load of all the equipment to be used. In addition to this, is required to have adequate cooling equipment to prevent the equipment from overheating and reduce its life.

Also, when a data center is starting and does not have much money, is good to avoid making very risky bets such as buying some very specific technologies or betting on not so common trends. It would be best to buy something more standard that meets the goal for which the infrastructure is created. As an example, Xeon Phi technology was purchased at CNCA. However, soon after, this was discontinued, which implies a problem if in the future it's desired to invest in the infrastructure that uses this technology.

Table 1. Technical specifications.

Purpose	RAM (GB)	Storage (TB)	Processor	Nodes	Science areas of use
Simulation	92	4.256	Xeon Phi 7210 @ 1.30GHz	32	Mathematical models, Molecular dynamics, Computational physics, Quantum physics, Condensed matter, Lattice Field Theories, HPC, Climate modeling, Air pollutants modeling, Meteorology and atmospheric physics, Computational Neuroscience and Simulations, Computational Seismology, Geomatics.
	94	6	Xeon Phi 7230F @ 1.30GHz		
Big Data	64	8	Xeon E5-2650 v4 @ 2.20GHz	4	Data Mining, Big Data, Statistical Analysis.
	64	1.2	Xeon E5-2650 v4 @ 2.20GHz		
	32	1.8	Xeon E5-2620 v4 @ 2.10GHz		
	62	11.018	Xeon E5-2650 v4 @ 2.20GHz		
Machine Learning	16	1	Xeon E3-1225 v5 @ 3.30GHz	4	Machine Learning, Deep Learning, Pattern Recognition, Computational Neuroscience and Simulations.
			Graphic Card: NVIDIA Tesla K40c		
Bioinformatics (coming soon)	1024	240	Xeon Gold 6154 @ 3.00GHz	1	Bioinformatics, Biocomputing, Genetics and Genomics, Microbiology, Biotechnology, Population's genetics.

Another good practice that must be carried out is that of a good network infrastructure design. The purchase of network switches, network interfaces and cables, as well as the way in which internal connections are defined (routers, gateways, etc.) must be optimized so that the speed of the connection, both inside and outside the system, is as quick as possible. A good practice would be to have a 10Gb Base T infrastructure as a minimum, to reduce latency.

In addition, within the good practices that it is advisable to have, is the use of failure response mechanisms. The use of RAIDs, from 1 onwards; make backups and allocate equipment only for that purpose; virtualization of the administration system, since it allows you to easily create snapshots and rollback; and keep track of the installed packages, as well as their installation process, in case it's required to reinstall them or install similar packages. These are some of the practices that can be implemented as an error containment mechanism. However, it is good to constantly improve this area and seek and implement new methodologies.

Finally, two elements that have been found important to implement are the modularization of user environments and infrastructure documentation. The first one is to allow loading the packages as modules so that each user has their personalized session with the packages and their respective versions according to their needs. And the second one is nothing more than the proper handling of updated and complete documentation. It has been found that the latter is of great importance, as this can speed up the process of debugging errors, as well as the introduction of new personnel to the support area.

Conclusion

The scientific task requires the power of HPC to solve problems that require great computational power. To solve this need, a computing infrastructure with sufficient power to carry out the tasks entrusted is essential. In the CNCA there is a computer platform (Kabré) in constant development capable of solving this need through the power of HPC. It has a hardware and software infrastructure that allows its users to develop scientific works using simulation tools, machine learning, big data, bioinformatics and storage.

Defining an HPC computing platform to be used in scientific research without proper knowledge of what the user's needs will be is complicated, since it's possible to make rash decisions that result in the acquisition of equipment that quickly becomes obsolete. For this reason, it is always advisable to study the technology market and choose equipment with good support, as well as one that adapts to the user's requirements. Is also required to implement proper methods of failure containment, documentation and design of the workspace for the proper function of an HPC infrastructure purposed for scientific research.

References

[1] A. Adakin et. al., "Building a high performance computing infrastructure for novosibirsk scientific center," *Journal of Physics: Conference Series* (Vol. 331, No. 5, p. 052020), IOP Publishing, 2011.

[2] D. Alvarado, A. de la Osa & M. Frutos, "Parallelization of a Magnetohydrodynamics Modelfor Plasma Simulation," *IEEE International Work Conference on Bioinspired Intelligence (IWOBI),* 2018.

[3] S. Calderón, J. Castro & M. Zumbado, "DNLM-MA-P: A Parallelization of theDeceived Non Local Means Filter withMoving Average and Symmetric Weighting," *IEEE International Work Conference on Bioinspired Intelligence (IWOBI),* 2018.

[4] M. Calvo, D. Jiménez & E. Meneses, "Analyzing Communication Features and Community Structure of HPC Applications," *IEEE International Work Conference on Bioinspired Intelligence (IWOBI),* 2018.

[5] C. Carley, L. Sells, B. McKinney, C. Zhao & H. Neeman, "Using a shared, remote cluster for teaching HPC," *IEEE International Conference on Cluster Computing (CLUSTER)*, 1-6, 2013.

[6] J. Castro & E. Meneses, "Parallelization of a Denoising Algorithm for TonalBioacoustic Signals using OpenACC Directives," *IEEE International Work Conference on Bioinspired Intelligence (IWOBI),* 2018.

[7] G. Cornejo, L. Van der Laat, E. Meneses, J. Pacheco & M. Mora, "Using parallel computing for seismo-volcanic event location based on seismic amplitudes," *IEEE 38th Central America and Panama Convention (CONCAPAN XXXVIII),* 2018.

[8] C. Gamboa & E. Meneses, "Comparative Analysis of de Bruijn Graph Parallel Genome Assemblers," *IEEE International Work Conference on Bioinspired Intelligence (IWOBI),* 2018.

[9] G. Ramirez-Gargallo, M. Garcia-Gasulla, & F. Mantovani, "TensorFlow on State-of-the-Art HPC Clusters: A Machine Learning use Case," *2019 19th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID), 526-533,* 2019.

[10] Slurm.com, "Quick Start User Guide", 2017. [Online] Available: https://slurm.schedmd.com/quickstart.html. [Accessed: 12-Aug- 2019].

[11] S. Varrette, P. Bouvry, H. Cartiaux & F. Georgatos, "Management of an academic HPC cluster: The UL experience," *International Conference on High Performance Computing & Simulation (HPCS)*, 959-967, 2014.