High Performance Computing (HPC) in Geomatics Engineering

Subash Ghimire

Assistant Professor and Head of Department
Department of Geomatics Engineering, Kathmandu University, Nepal

Abstract: *With the technological developments, the amount of data produced by many Geoinformation communities on a daily basis has increased to terabyte levels. Remotely sensed data acquired by means of various sensing techniques, causes great difficulties in storing, transferring, and analyzing with conventional methods. Therefore, High Performance Computing is necessary and important for all who deals with large geospatial data such as, students, faculties, researchers, government and private sectors to solve the problems that will not fit in a normal computer. The main of aim of this paper is to analyze the application of High-Performance Computing (HPC) techniques in processing of geospatial data and to provide an overview of HPC application courses at Department of Geomatics Engineering, Kathmandu University. This study is based on literature review and secondary data. The main application of HPC is found in well-known problems in GIS such as georeferencing, generating digital surface model, map matching, Image fusion etc. A slope calculations and computations based on huge digital elevation model data can be processed on Graphics Processing Unit using Compute Unified Device Architecture that enable hundreds of threads to run concurrently employing multiprocessors on a graphics card. The various Geomatics engineering related research work employing HPC has been presented in this paper.*

Keywords: *High Performance Computing, Geomatics Engineering, Kathmandu University*

1. Introduction

Geomatics is a very active and advanced field of engineering that deals with multi-resolution geospatial and spatio-temporal data for all kinds of engineering applications. HPC have brought about fundamental changes in the handling and processing of geospatial data (Blais & Esche, 2014). HPC brings various technologies such as computer algorithm, architecture and system software under one roof to solve very complex problems effectively, efficiently and quickly (Bhojne et al., 2013). The various computing-intensive applications require large processing power. Generally, HPC is familiar to the practice of collecting (grouping) computing power in such a way, that delivers more higher performance comparing to typical desktop computer to solve complex problems in Geomatics Engineering (Stojanovic & Stojanovic, 2013). It focuses on efficient processing of geospatial data in the supercomputing domain for the benefit of spatial analytics and spatial data handling. It is important for the students, faculties, researchers, governments, private sectors etc to solve the problems that will not fit in a simple computer. HPC cluster has hundreds or thousands of computer servers which are networked together. The nodes in each cluster work in parallel with each other, boosting processing speed to deliver high-performance computing. It is a use of parallel processing technique for executing advanced applications reliably, efficiently and quickly.

HPC is an application of supercomputers for computational Geomatics problems that are either too large for standard computers. A supercomputer is types of computers that are capable of handling big data or do a great amount of computation simultaneously. It consists of many processors that are able to perform billions and trillions of calculations or computations per second. Many modern software applications that are developed to model real world process are performed in HPC. The parallel processing significantly supports Geographic Information System (GIS) that includes handling, processing and analytics of large spatial data sets (Stojanovic & Stojanovic, 2013).

2. Application of HPC in Geomatics Engineering

The high increase in the amount of geospatial data produced daily by many geoinformation communities' reveals big challenges in data storage and extracting information timely. Many geospatial sensors designed in today's technologies are used in geoinformation science and earth observation. Real-time geospatial data must be processed and analyzed in a short time in order to be able to provide time-critical decision support in time-critical applications (Pektürk & Ünal, 2018).

The application of HPC principles in Geomatics engineering began from middle of 1990s (Armstrong, 1995). Currently, with the increasing volume of geospatial data required for computational and data-intensive problem solving in different GIS application domains, have emerged as a prominent research area (Shekhar, 2010). The HPC is used to evaluate two parallel/distributed architectures and programming models. The distributed application for map-matching computation considering moving points and a road network was implemented following MPI. The slope computation was carried on GPU using CUDA (Stojanovic & Stojanovic, 2013). Similarly, python a high performance computer language has

NumPy and Numba both of which allow very fast array and matrix computations.

Experimental evaluations show improvement in performance and indicate feasibility of using NoW and multiprocessors on a graphic card for HPC in GIS. GPU has come out as a powerful co-processor for general-purpose calculation. Comparing with CPUs, GPU has higher memory bandwidth and higher computation power. HPC has a broad range of applications that demand accelerating and performance improvements such as database searching, web search engines, hydrological and environmental modeling, and climate change modeling etc. (Patel & Hwu, 2008).

Currently, available advanced geoinformation collection, management and dissemination and mobile positioning technologies supports management of huge quantities of spatial data at low cost. Therefore, HPC has an application for processing and storing large spatial database size used in various geoinformation applications such as environmental monitoring, climate change observation, traffic management and optimization and Location-Based Services (LBS). HPC may be used to overcome the computational intractability of large and complex spatio-temporal data sets.

Similarly, HPC are successfully used in processing both large size raster geospatial data produced by satellites in terabytes (Shekhar, 2010; Zhang, 2010). It is also used in spatial query processing and overlay computation over large amount of spatial vector data (Park et al., 2010). An overlay calculations and computations and spatial queries are very important capabilities of geographical analysis in various GIS applications and required HPC**.** Both, spatial join query processing and spatial overlay computations are time-consuming in handling large-scale and voluminous spatio-temporal data, and thus are excellent for application of HPC. Shi (2010) discusses basic research challenges in application of HPC techniques in a service-oriented GIS. HPC may be used to archive data of various land and properties, which is of national interest and also require high confidentiality and security.

3. Geomatics Engineering Related Research work employing HPC

Various Geomatics Engineering related research work employing HPC is tabulated as follows.

S No	Author	Contribution	Results
1	Zhang (2010)	Considered a new HPC framework for processing geospatial data in a personal computing environment.	It argued that modern personal computers equipped with multicore CPU and many-core GPU provide excellent support for spatial data processing
2	(Bhojne et al., 2013)	Review on recent development in HPC technology for satellite data processing and analyzing.	Parallel processing- most image fusion algorithms and hyper spectral image processing based on neural architectures and morphological model. Distributed processing- A distributed network of inexpensive PCs can be designed that is optimal to deal with the type of computationally intensive problems encountered in processing remotely sensed images. Cluster computing- Geo-referencing, image transformation, image mosaicking, etc.
3	Van der Merwe & Meyer (2009)	Automatic Digital Surface Models (DSM) generation from satellite imagery using GPU was performed.	According to preliminary results, GPU algorithm decreased processing time by 900%.
4	Beutel et al. (2010)	Considered construction of a grid Digital Elevation Model (DEM) from 3D point clouds generated by the LiDAR equipment.	The obtained results showed that using of a GPU for this type of GIS application can significantly speed up the computation
5	Akhter et al. (2010)	Develop the methodology which enables GRASS GIS software to run on HPC systems.	Different implementations for parallel/distributed GRASS modules are presented on three different programming platforms.

4. HPC at Kathmandu University

Kathmandu University has installed the supercomputer at Information Technology (IT) Park, Banepa. The supercomputer which consists of high performance computing (HPC) servers was donated by CERN, computing centers based in Switzerland. This supercomputer has 184 CPU servers, 16 disk servers and 12 network switches, with a total processor count of over 2,500 and 8 TB of memory. This will contribute towards a new high-performance computing facility for research and educational purposes in Nepal. Similarly, Research & Development projects in the schools of science and engineering have to borrow their computing time abroad, either through online data transfer, which is inevitably slow for large data transfers.

5. Major Courses at Department of Geomatics Engineering employing HPC

Several recent teaching/ learning and research works employ high-performance processing techniques on geospatial data at Department of Geomatics

Engineering, Kathmandu University and are mentioned on following subsections.

5.1 BE in Geomatics Engineering

The following major courses in BE in Geomatics Engineering needs HPC

- GIS and Remote sensing
- Computational Geomatics
- Python programming
- Hydrological modeling
- Environmental modeling
- Photogrammetry
- Geomatics engineering students project

5.2 ME/MS in Geoinformatics and Master in Land Administration

The following major core/elective courses in graduate programs employ HPC

- Remote Sensing
- Geo-visualization & Advanced Cartography
- Geo-statistics
- Spatial analysis and Modeling
- Digital Image Processing
- Location Based Services (Mobile and Web)
- Hyper Spectral Remote Sensing
- Spatial Data Infrastructures / OpenGIS
- Scientific Geo-computing
- GIS and Remote sensing
- Urban Planning and Management
- Spatial Engineering in Agriculture and Forestry
- 3D Geoinformatics
- Geospatial Data Mining
- Application of Unmanned Aerial Systems
- Problem Assessment Project
- Master thesis

5.3 PhD in Geomatics Engineering

PhD scholars at Kathmandu University may employ HPC for Database searching and web search engines. The students, faculties and researchers may use the supercomputer as a backup data centre.

6. Major stakeholders of HPC

The collaborations with national and international organizations is necessary to share the experience, skills and knowledge on High performance computing. The related stakeholders for HPC are

- Ministry of Land Management, Co-operatives and Poverty alleviation
- Survey Department
- ICIMOD
- Federal and Provincial Ministries & Departments
- Metropolitan City
- Nepal Army/ Nepal Police/Armed Police Force
- Hydropower Companies
- Nepal Electricity Authority
- Private sector in Infrastructure development activities/Consultancy
- NGOs/INGOs
- National and International Universities, Private and public companies

7. Challenges

One of the methods for achieving performance improvements is to decrease the size of components used to manufacture computers. The technological development has made possibility to place billions of transistors on a single chip. But, raising the clock frequency, as a major approach of computer performance advance, has limited due to energy consumption and heat-dissipation. This is also the challenges faced by KU supercomputer at IT Park.

8. Future work

The Department of Geomatics Engineering at Kathmandu University will consider the comprehensive usage of other HPC techniques and platforms in Geomatics Engineering such as cloud computing and the adaptation of various GIS algorithms to cloud infrastructure. To achieve this, the department is committed for collaboration with other departments at Kathmandu University and national and international HPC communities.

9. Conclusion

In recent years, massive spatio temporal data has been generated with advances in GIS, remote sensing, mobile positioning, LiDAR and cloud computing and sensor networks. HPC promotes in processing and analysis of such Big Data. This paper shows that application of HPC in Geomatics Engineering. It reveals that using parallelization and HPC in GIS represents a promising research and development field. Kathmandu University has installed the supercomputer at Information Technology (IT) Park, Banepa. This supercomputer has 184 CPU servers, 16 disk servers and 12 network switches, with a total processor count of over 2,500 and 8 TB of memory. This will contribute towards a new high-performance computing facility for research and educational purposes in Nepal. There are many courses which emplys HPC at Department of Geomatics Engineering The collaboration with national and international organization is necessary for joint research projects and development of supercomputer laboratories to fulfill the need of hardware and software at Department of Geomatics Engineering

References

Akhter, S., Aida, K. & Chemin, Y. (2010).GRASS GIS on high performance computing with MPI, OpenMP and Ninf-G programming framework. *Proceeding of ISPRS 2010*, Japan.

Armstrong, M.P. (1995). Is there a role for high performance computing in GIS? *Journal of the Urban and Regional Information Systems Association*, Vol. 7, No. 1, pp.7–10.

Beutel, A., Мшlhave, T. & Agarwal, P.K. (2010). Natural neighbor interpolation based grid DEM construction using a GPU. *Proceedings of 18th ACM SIGSPATIAL International Symposium on Advances in Geographic Information Systems*, pp.172–181.

Bhojne, M., Chakravarti, A., Pallav, A. & V, S. (2013). High Performance Computing for Satellite Image Processing and Analyzing – A Review. *International Journal of Computer Applications Technology and Research,* Vol 2, No. 4, 424 – 430.

Blais, R. & Esche, H. (2014). Geomatics and the New Cyber-Infrastructure, *Geomatica,* Vol. 62, No. 1, 2008, pp. 431- 443.

Patel, S. & Hwu, W.W. (2008). Accelerator architectures, *IEEE Micro*, Vol. 28, No. 4, pp.4–12.

Pektürk,M.K. & Ünal, M.,(2018). Performance-Aware High-Performance Computing for Remote Sensing Big Data Analytics. *Open access peer-reviewed chapter, DOI: 10.5772/intechopen.75934.*

Shekhar, S. (2010) 'High performance computing with spatial datasets', *Invited Talk, HPDGIS 2010 Workshop: International Workshop on High Performance and Distributed Geographic Information Systems*, San Jose, California.

Shi, X. (2010). High performance computing: fundamental research challenges in service oriented GIS, *ACM SIGSPATIAL – HPDGIS 2010 Workshop: International Workshop on High Performance and Distributed Geographic Information Systems*, San Jose, California, pp.31–34.

Stojanovic, N. & Stojanovic, D. (2013).High-performance computing in GIS: techniques and applications, *Int. J. Reasoning-based Intelligent Systems*, Vol. 5, No. 1, pp.42–49.

Van der Merwe, D.& Meyer, J. (2009). Automatic digital surface modeling using a graphics processing unit, *Proceedings of the Southern Africa Telecommunication Networks and Applications Conference (SATNAC)* 2009.

Zhang, J. (2010). Towards personal high-performance geospatial computing (HPC-G): perspectives and a case study, *ACM SIGSPATIAL – HPDGIS 2010 Workshop: International Workshop on High Performance and Distributed Geographic Information Systems*, San Jose, California, pp.3–10.

URL 1**:** https://kathmandupost.com/money/2019/02/13/nepal-finally-gets-its-first-supercomputer, retrieved on 24 March 2020.

URL 2**:** http://geom.ku.edu.np/, retrieved on 24 March 2020.

URL 3**:** https://www.usgs.gov/core-science-systems/ngp/cegis/high-performance-computing?qt-science_support_page_related_con=1#qt-science_support_page_related_con,retrieved on 28 March 2020.