CS 434: Parallel and Distributed Computing

Percy Brown: 40382022

February 25, 2021

Lab Assignment 1

1 5 Top HPC Machines in the World

High - Performance computing (HPC) may generally refer to as the practice of combining computing power in a way that delivers much higher performance. Therefore, HPC Machines help us to process data and perform calculations at higher speeds other than a typical desktop computer. Below are the top 5 HPC Machines in the world as of November 2020.

1.1 Supercomputer Fugaku - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D

The Supercomputer Fugaku development plan was initiated by the Ministry of Education, Culture, Sports, Science and Technology of Japan in 2014, giving birth to the development of Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, which is currently ranked first amongst the top 500 high - performance computing machines in the whole world as of November 2020. The machine was manufactured by Fujitsu, a Japanese multinational information technology equipment and services company headquartered in Tokyo. It has a memory capacity of 5,087,232 GB, with 7,630,848 cores. At the heart of the Fugaku supercomputer is its Arm - based processor, A64FX, with 48 core CPUs running at a clock speed of 2.2GHz. For its system interconnect, it uses Tofu interconnect D, which provides Increased resources for high-density node configuration. Using the LINPACK Benchmark as a performance yardstick, Fugaku is now capable of 442,010 TFlop/s, meaning that it can perform 442,010 floating operations per second, making it a powerful beast, whiles its theoretical peak performance is rated at 537,212 TFlop/s. The supercomputer is pretty power - efficient, pulling 29,899.23 kW of power. It uses the Red Hat Enterprise Linux operating system which supports and powers software and technologies for automation, cloud, containers, middleware, storage, application development, microservices, virtualization, management and more. It consists of a total of 158,976 nodes.It consists of a total of 158,976 nodes. For more information about the Supercomputer Fugaku, visit `http://www.sharelatex.com`

1.2 Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband

The US Summit, an IBM-built supercomputer now running at the Department of Energy's (DOE) Oak Ridge National Laboratory (ORNL), and is currently ranked second amongst the top 500 high - performance computing machines in the whole world as of November 2020. The machine was manufactured by International Business Machines (IBM), an American multinational technology company headquartered in New York. It has a memory capacity of 2,801,664 GB, with 2,414,592 cores. At the heart of the US Summit is its IBM POWER9 architecture, with 22 core CPUs running at a clock speed of 3.07GHz. For its system interconnect, it uses Dual-rail Mellanox EDR Infiniband, for storage and inter-process communications. Using the LINPACK Benchmark as a performance yardstick, The Summit's maximal achieved performance is rated at 148,600 TFlop/s, meaning that it can perform 148,600 floating operations per second, whiles its theoretical peak performance is rated at 200,795 TFlop/s. The supercomputer consumes 10,096.00 kW of power. It uses the Red Hat Enterprise Linux 7.4 version as its operating system which introduces support for Network Bound Disk Encryption (NBDE). It consists of a total of 4608 nodes. The Summit supercomputer provides scientists and researchers the opportunity to solve complex tasks in the fields of energy, artificial intelligence, human health and other research areas. For more information about the Summit, visit `https://www.olcf.ornl.gov/olcf-resources/compute-systems/summit/`

1.3 Sierra - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband

Sierra, Livermore's latest advanced technology high performance computing system, joined Lawrence Livermore National Laboratory's lineup of supercomputers in 2018, and is currently ranked third amongst the top 500 high - performance computing machines in the whole world as of November 2020. Seirra was manufactured in the United States by International Business Machines (IBM), an American multinational technology company headquartered in New York, NVIDA, an American multinational technology company incorporated in Delaware and based in Santa Clara, California and Melanox, an Israeli-American multinational supplier of computer networking products based on InfiniBand and Ethernet technology. It has a memory capacity of 1,382,400 GB, with 1,572,480 cores. At the heart of Sierra is its IBM POWER9 architecture, with 22 core CPUs running at a clock speed of 3.1GHz. For its system interconnect, it uses Dual-rail Mellanox EDR Infiniband for storage and inter-process communications. Using the LINPACK Benchmark as a performance yardstick, Sierra's maximal achieved performance is rated at 94,640 TFlop/s, meaning that it can perform 94,640 floating operations per second, whiles its theoretical peak performance is rated at 125,712 TFlop/s. The Sierra supercomputer consumes 7,438.28 kW of power. It uses the Red Hat Enterprise Linux operating system which supports and powers software and technologies for automation, cloud, containers, middleware, storage, application development, microservices, virtualization, management and more. It consists of a total of 4,474 nodes. The Sierra supercomputer provides computational resources that are essential for nuclear weapon scientists to fulfill the National Nuclear Security Administration's stockpile stewardship mission through simulation in lieu of underground testing. For more information about Sierra, visit `https://hpc.llnl.gov/hardware/platforms/sierra`

1.4 Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway

The Sunway TaihuLight is a Chinese supercomputer which is currently ranked fourth amongst the top 500 high - performance computing machines in the whole world as of November 2020. It was designed by the National Research Center of Parallel Computer Engineering Technology (NRCPC) and is located at the National Supercomputing Center the city of Wuxi, in Jiangsu province, China. It has a memory capacity of 1,310,720 GB, with 10,649,600 cores. At the heart of Sunway TaihuLight is its Sunway SW26010, a 260-core many-core processor running at a clock speed of 1.45GHz. It uses its custom Sunway interconnect, based on PCIe 3.0 and delivers 16 GB/sec of peak bandwidth between nodes, with 1ms latency. Using the LINPACK Benchmark as a performance yardstick, Sunway TaihuLight's maximal achieved performance is rated at 93,014.6 TFlop/s, meaning that it can perform 93,014.6 floating operations per second, whiles its theoretical peak performance is rated at 125,436 TFlop/s. Sunway TaihuLight supercomputer consumes 15,371.00 kW of power. The system runs on its own operating system, Sunway RaiseOS 2.0.5, which is based on Linux. The system is already utilized in many tasks related to both nature and society. For more information about Sunway TaihuLight, visit `https://steemit.com/technology/@asmolokalo/sunway-taihulight`

1.5 Selene - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband

The world's fastest commercial machine, Selene was named the world's fifth-fastest supercomputer in the world on November 2020's closely watched list of TOP500 supercomputers. Selene, manufactured in United States by NVIDIA, an American multinational technology company incorporated in Delaware and based in California, is at the center of some of its most ambitious technology efforts. It has a memory capacity of 1,120,000 GB, with 555,520 cores. At the heart of Selene is its architecture (Extreme Performance Yield Computing architecture (AMD EPYC 7742) with 64 core CPUs running at a clock speed of 2.25GHz. For its system interconnect, it uses Mellanox HDR Infiniband for storage and inter-process communications. Using the LINPACK Benchmark as a performance yardstick, Selene's maximal achieved performance is rated at 63,460 TFlop/s, meaning that it can perform 63,460 floating operations per second, whiles its theoretical peak performance is rated at 79,215 TFlop/s. Selene supercomputer consumes 2,646.00 kW of power. The system runs on Ubuntu 20.04.1 LTS operating system, which is based on Linux. What makes Selene remarkable is how closely it is wired into the day-to-day work of some of NVIDIA's top researchers. For more information about Selene, visit `https://www.nvidia.com/DGXSuperPOD`

2 Top HPC Machine in Africa

The top HPC Machine in Africa is the Lengau Supercomputer. Lengau is located in South Africa and was manufactured by Dell EMC in collaboration with the Centre for High Performance Computing (CHPC) at the Council for Scientific and Industrial Research (CSIR). It has a memory capacity of 175,232 GB, with 32,856 cores. At the heart of Lengau is Intel Xeon (R) E5-2690V3 processor with 12 core CPUs running at a clock speed of 2.6GHz. For its system interconnect, it uses FDR Infiniband Network which features very high throughput and low latency for storage and inter-process communications. Using the LINPACK Benchmark as a performance yardstick, Lengau's maximal achieved performance is rated at 1,029.32 TFlop/s, meaning that it can perform 1,029.32 floating operations per second, whiles its theoretical peak performance is rated at 1,366.81 TFlop/s. The Lengau supercomputer consumes 685.00 kW of power. The system runs on CentOS operating system, which is based on Linux. The supercomputer will provide research institutions and private industries the required processing power for modelling, big data processing and research. It consists of 1008 nodes. For more information about Lengau, visit `http://www.chpc.ac.za/`

3 Applications of HPC Machines in Ghana

3.1 Heathcare

One major application of HPC Machines in Ghana could be in the field of healthcare. High-performance computing has advanced medical research in many ways, thus applying it in Ghana's health system can improve medical research and healthcare activities. For example, it has led to the creation of a noninvasive robotic arm controlled by the brain—a potentially life-changing breakthrough for people living with paralyzed limbs. Medicine and computing are as intimately intertwined as DNA's double helix. Computers already store confidential patient information, track vital signs and analyze drug efficacy. The rise of HPC will allow Ghana's medical professionals to digitize even more complex processes, too, like genome sequencing and drug testing. Also, HPC is very has been vital to analysis of cancer genomics data, both for providing the necessary computational power and the security needed for handling sensitive patient genomic datasets. The potential of HPC to transform healthcare is clear, and thus having HPC machines in the healthcare sector of Ghana will greatly improve our healthcare services and not be over-reliant on foreign healthcare interventions.

3.2 Urban planning

Smart people are book smart and street smart. A smart city is data smart. Major metropolises across the Ghana have begun collecting sensor data on weather, traffic patterns and noise levels, all of which allow officials to make data-driven decisions about everything from when to issue smog warnings to how often trains in some parts of the country should run. It also lets them quantify longer-term issues like climate change and infrastructure decay. Because smart city sensor networks collect so much data, they need HPC to parse it all, hence the need for HPC machines in the country to aid in urban planning.

3.3 Finance

The financial marketplace in Ghana today is dominated by growing regulatory requirements, constantly changing liquidity conditions and increasingly complex asset-class strategies by firms across the spectrum. In such an environment, it's no longer only top-tier banks that need vast amounts of high performance computing (HPC) capabilities to handle the huge number of calculations needed for risk management, trade idea generation, compliance or a host of other functions that sell-side and buy-side organisations carry out. Ghana's financial sector is known for its compute-intensive needs, although those needs do vary significantly depending on the type of firm and its market profile. Banks and brokers can typically use HPC and grid computing for tasks such as real-time risk management, as they handle thousands of transactions for clients around the world at any given moment. Proprietary trading outfits, meanwhile, often need HPC to run sophisticated models. The potential of HPC to transform improve activities in our financial sector is clear.

4 Challenges of using HPC Machines in Ghana

4.1 Energy consumption

Power and energy consumption are seen of one of the most critical design factor for any next generation large-scale HPC system. The price centers have to pay for energy is shifting the budgets from investment to operating costs, leading to scenarios in which the sizes of systems will be determined by their power needs, rather by the initial hardware cost. As a consequence, virtually all funding agencies for HPC projects in Ghana must have set aggressive goals for peak power requirements in the machines. With the current energy supply state in Ghana, it is a very pressing challenge to use hogh-performance computing machines in the country, especially when we cannot aford to provide constant power to all parts of the country as it stands.

4.2 Financing and Maintenance

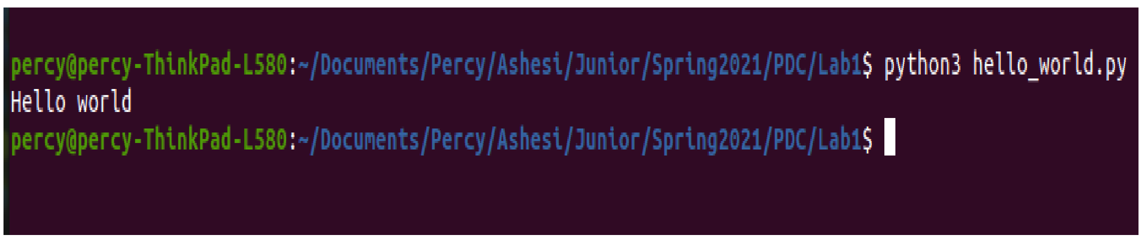
Although having high-performance computing machines in Ghana can improve upon our lives and activities in areas such as healthcare and finance among others, Ghana's challenge will be the financing and maintenance in having these supercomputers in various fields in the country. We cannot use HPC machines in just one sector whiles others are still lagging behind in terms of efficient services and operations. Therefore, it will be very challenging for Ghana, especially as a developing country to finance and maintain HPC machines in the country.

5 Part 2 - Familiarization and understanding of Unix programming environment

Since my main environment is windows 10, I installed Ubuntu 20.04.2 for dual booting. I am also currently using Visual Studio Code as my editor in ubuntu.

5.1 Python

I wrote a simple python program to print "hello world" and run it on the unix command line. Find a screenshot after running the program below.



5.2 OpenMP Library

Using the gcc compiler, I run a sample program successfully that uses the OpenMP library, i.e. adding "-lopenmp -lgomp." Find a screenshot after compilation and running the program below.

5.3 PThread Library

Using the gcc/g++ compiler, I run a sample program successfully that uses the PThread library, i.e. adding "-lpthread". Find a screenshot after compilation and running the program below.

```
percy@percy-ThinkPad-L580:~/Documents/Percy/Ashesi/Junior/Spring2021/PDC/Lab1$ gcc sample_pthreads.c -o sample_pt -lpthread
sample_pthreads.c: In function ‘main’:
sample_pthreads.c:14:13: warning: implicit declaration of function ‘pthread_create’ [-Wimplicit-function-declaration]
   14 |     iret1 = pthread_create( &thread1, NULL, print_message_function, (void*) message1);
      |             ^~~~~~~~~~~~~~
sample_pthreads.c:18:5: warning: implicit declaration of function ‘pthread_join’ [-Wimplicit-function-declaration]
   18 |     pthread_join( thread1, NULL);
      |     ^~~~~~~~~~~~
percy@percy-ThinkPad-L580:~/Documents/Percy/Ashesi/Junior/Spring2021/PDC/Lab1$ ./sample_pt
Thread 1
Thread 2
Thread 1 returns: 0
Thread 2 returns: 0
percy@percy-ThinkPad-L580:~/Documents/Percy/Ashesi/Junior/Spring2021/PDC/Lab1$
```

5.4 GitHub

The source code files will be located under my github repository name **Parallel and Dsitributed Computing** under the folder named **lab1**.