HPC Rankings Based on Real Applications

NOLAN BAKER, The University of Delaware
AARON JARMUSCH, The University of Delaware
SUNITA CHANDRASEKARAN, The University of Delaware
RUDOLF EIGENMANN, The University of Delaware

CCS Concepts: • **Computer systems organization → Architectures**; *Parallel architectures*; Multicore architectures.

Additional Key Words and Phrases: performance, benchmarking, HPC

ACM Reference Format:
Nolan Baker, Aaron Jarmusch, Sunita Chandrasekaran, and Rudolf Eigenmann. 2020. HPC Rankings Based on Real Applications . 1, 1 (October 2020), 3 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

High performance computing (HPC) systems has become increasingly prevalent. There are several systems available across the world, including several listed under Top500 [3]. Often, potential users are looking for the best systems to target, and to that end the Top500 ranks the supercomputers based on a number of criteria. This serves as a motivation for our project. In our project, we explore the idea of ranking of several clusters using the (Standard Performance Evaluation Corporation) HPG (High Performance Group) ACCEL results [1]. SPEC is a non-profit organization that builds benchmark suites encompassing several programming models including MPI, OpenMP and OpenACC. The purpose of the suite is to stress test both the hardware characteristics and the software stack.

SPEC ACCEL HPG utilizes various science applications to create a “base” and a “peak” score. The “base” score and other result details will be displayed on a developed website in a ranking style. This public interface will allow users to filter specialized information while utilizing popup windows to display individual supercomputer details.

2 DEVELOPMENT OF WEBSITE

The SPEC HPG ranking website will use ACCEL results from the SPEC.org site. Our website uses a web-scraper that is be on an automatic timer to capture and input new results into a MySQL database directly. The new website will access all information on the database and display the information for viewers in an easy to understand format.

The creation of the SPEC Ranking website utilizes Indiana University and University of Texas Austin’s ’Jetstream’ system [6][2]. Jetstream utilizes cloud resources and hardware to create instant Virtual Machines (VM) for scientific, engineering and educational purposes. We use one such virtual machine. The web-scraper, database, and website are all hosted on a Jetstream VM. Jetstream

Authors’ addresses: Nolan Baker, The University of Delaware; Aaron Jarmusch, The University of Delaware; Sunita Chandrasekaran, The University of Delaware; Rudolf Eigenmann, The University of Delaware.

XXXX-XXXX/2020/10-ART $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

allowed for co-operation for every step of the web creation process during a physical contact-free time.

2.1 Web-Scraping with Python

The SPEC ACCEL data is updated as new results are uploaded to the SPEC.org website. Utilizing a fully automated system, a cron job (a time-based job scheduler) is established on a Jetstream VM. Periodically, this timer will run our web-scraper program too input any newly updated results from SPEC.org into our database.

Data, displayed on the SPEC website as files, has a direct link. These files can be quickly read and parsed through BeautifulSoup, a Python package for parsing documents[5]. Then, a python data organizer, Pandas, allows for easy integration into the database[4].

2.2 Database

The database of choice is MySQL, which contains one table to hold the organized list of each benchmark entry. Once the system captures more results, the system can be evaluated for each quarter of a year. Each entry in the table consists of an individual system and its specifications. In order to distinguish between systems, a name key is used, designed by SPEC but re-purposed for MySQL organization.

2.3 Web-development

We utilize 'PHPMyAdmin' (PHP) to visualize the data, including each column and entry and automate the process using a script.

We use CSS, HTML, PHP and Javascript to display the benchmarks and configure the overall display of the website. A 'bootstrap template' is initially used for simplicity, and further configured to our needs. To make the website user-friendly, here are the key features of our design:

- A search function at the top of the home page, allowing users to search for terms in or typed of data.
- A filter function to display more (and all) info per system, including, but not limited to memory speed, power consumption, operating system, etc.
- A selection function per column to list each system out by that column, either numerically or alphabetically.
- An 'About' page, which discusses the origin of the scores
- A 'Contact' page to receive feedback on the website.

3 CONCLUSION

Researchers must continually evaluate their methods for performance accuracy in real world systems settings to advance science. In the year 2020, for example, many individuals are researching herd immunity models, which require large data sets and complex formulas to process. To that end, demand for HPC systems is growing exponentially. Information regarding performance, energy among other factors of a system made publicly available, in an organized manner, will be very beneficial to the end users.

ACKNOWLEDGMENTS

To Winona Snapp-Childs, for mentoring us through the REU Jetstream program. To Sunita and Rudy for being true mentors with this project. Helping us complete this project to the best of our ability.

REFERENCES

[1] Sunita Chandrasekaran Barbara Chapman Shuai Che Mathew Colgrove Huiyu Feng Alexander Grund Robert Henschel Wen-Mei W. Hwu Huian Li Matthias S. Müller Wolfgang E. Nagel Maxim Perminov Pavel Shelepugin Kevin Skadron John Stratton Alexey Titov Ke Wang Matthijs van Waveren Brian Whitney Sandra Wienke Rengan Xu Kalyan Kumaran Guido Juckel, William Brantley. 2015. *SPEC ACCEL: A Standard Application Suite for Measuring Hardware Accelerator Performance.* Vol. 8966. Lecture Notes in Computer Science.

[2] Maytal Dahan Ian Foster Kelly Gaither Andrew Grimshaw Victor Hazlewood Scott Lathrop Dave Lifka Gregory D. Peterson Ralph Roskies J. Ray Scott Nancy Wilkins-Diehr John Towns, Timothy Cockerill. Sept.-Oct. 2014. *XSEDE: Accelerating Scientific Discovery.* Vol. 16, no. 5, pp. 62-74. Computing in Science Engineering.

[3] Hans Werner Meuer, Erich Strohmaier, Jack Dongarra, and Horst D. Simon. 2014. *The TOP500: History, Trends, and Future Directions in High Performance Computing* (1st ed.). Chapman Hall/CRC.

[4] The pandas development team. 2020. *pandas-dev/pandas: Pandas.* https://doi.org/10.5281/zenodo.3509134

[5] Leonard Richardson. 2007. Beautiful soup documentation. *April* (2007).

[6] Cockerill T.M. Foster I. Hancock D. Merchant N. Skidmore E. Stanzione D. Taylor J. Tuecke S. Turner G. Vaughn M. Stewart, C.A. and N.I. Gaffney. 2015. *Jetstream: a self-provisioned, scalable science and engineering cloud environment.* Vol. 200. Proceedings of the 2015 XSEDE Conference: Scientific Advancements Enabled by Enhanced Cyberinfrastructure.