Porting a parallel program from the NASA Center for Climate Simulation (NCCS) Discoverer supercomputer to desktops for validation of the Multi-sensor Aerosol Products Sampling System (MAPSS)

Charles Sleeper
Southwestern Oklahoma State University

Jeremy Evert
Southwestern Oklahoma State University, jeremy.evert@swosu.edu

Abstract

The National Aeronautics and Space Administration (NASA) produces nearly two gigabytes of data per second. NASA researchers leverage millions of dollars of computing hardware to analyze this data. NASA shares this data with the rest of the world. Advances in computer technology have provided modern desktop computers more powerful than the fastest supercomputers in the world from two and three decades ago. This provides many possibilities for greater use... Read More

Follow this and additional works at: http://dc.swosu.edu/cpgs_EDSBT_BCS_student

Recommended Citation
Sleeper, Charles and Evert, Jeremy, "Porting a parallel program from the NASA Center for Climate Simulation (NCCS) Discoverer supercomputer to desktops for validation of the Multi-sensor Aerosol Products Sampling System (MAPSS)" (2017). Student Research. 5.
http://dc.swosu.edu/cpgs_EDSBT_BCS_student/5

This Poster is brought to you for free and open access by the Business & Computer Science at SWOSU Digital Commons. It has been accepted for inclusion in Student Research by an authorized administrator of SWOSU Digital Commons. An ADA compliant document is available upon request. For more information, please contact phillip.fitzsimmons@swosu.edu.
The National Aeronautics and Space Administration (NASA) produces nearly two gigabytes of data per second. NASA researchers leverage millions of dollars of computer hardware to analyze this data. NASA shares this data with the rest of the world. Advances in computer technology have provided supercomputers and software engineers to desktops and undergrads. Undergraduate student researchers studied the MAPSS software system, created for the NASA Goddard Space Flight Center supercomputers, to conduct a validation study of NASA Earth Atmosphere aerosol data. Undergrads rewrote parts of the software allowing it to run on an Intel i7 processor running a Linux system. The students completed translation of four of the seven satellite sensors, and developed automation software allowing MAPSS to be portable between individual computers. The students provided documentation of this process allowing future students to complete the translation of the remaining sensor systems and the validation study. This should provide greater use of the data that streams from NASA every day.

### Research Questions

**Do NASA satellites sensor platforms produce similar readings to ground-based sensors?**

Can NASA research be verified by undergraduate researchers? What level of Cyber infrastructure resources are necessary to examine a data set this size?

### Atmospheric Aerosol Data Collection

Space-based sensor took measurements looking down to Earth, while ground-based sensors took measurements looking up. Statistics for each of the aerosol products involved extracting measurements that fall within 2.5 km of the chosen locations, and within 30 minutes of each satellite fly over.

### Results

- May 2016: SWOSU students received an Oklahoma NASA EPSCoR Research Initiation Grant. Students formed a summer research team. The students selected LaTeX for documentation of all publications, GitHub as their knowledge management tool. All students focused on their areas of expertise. Some began a literature review of the NASA MAPSS project. Other students developed a Python translation of the functionality of the MAPSS software. SWOSU student Jack Gillory took the lead on translating the MAPSS source code to run on a Linux distribution and standard desktop computer. Charles Sleeper began developing a Raspberry Pi Cluster to support future research partnerships with area High Schools.
- May 2016: The SWOSU HPC team participated in the first Oklahoma HPC competition. Students deployed a Raspberry Pi cluster. Students also developed code to find the determinant of a four by four matrix that is used in the MAPSS software. The MAPSS software continues to be in use after two decades of development. The team of software engineers at the NASA Goddard Space Flight Center had tuned the software for optimal performance on several supercomputers, including the Discoverer 12. Jack Gillory was able to deliver functionality with four of the nine modules within the software.
- August 2016: SWOSU students participated in a NSF C* Grant proposal. The SWOSU team coordinated a Cyber Infrastructure proposal for three universities. During the fall 2017 Oklahoma Supercomputing Symposium, this plan was praised by members of the One OCI team as being one of the better written plans the state had seen in recent years.
- September 2016: SWOSU meets with Southeastern Oklahoma State University research faculty to begin collaboration on future NASA OK EPSCoR Grants.
- December 2016: Oklahoma State University provided SWOSU with eight servers from a retired cluster. This cluster was deployed using CentOS with the SLURM job scheduler. Students installed the operating system, set up the administrator accounts, and worked with university IT staff to enable access to the server from various classrooms.
- December 2016: The SWOSU HPC team reached out to Dr. Toshihisa Matsui for a copy of the NU-WRF model. The SWOSU COMSC 3133 Software Engineering students will begin analysis of the software. This software was used as one of the three components of the Student Cluster Competition at The 2016 International Conference for High Performance Computing, Networking, Storage and Analysis (SC16)

### References


---

**Timeline**

- May 2016: SWOSU students received an Oklahoma NASA EPSCoR Research Initiation Grant. Students formed a summer research team. The students selected LaTeX for documentation of all publications, GitHub as their knowledge management tool. All students focused on their areas of expertise. Some began a literature review of the NASA MAPSS project. Other students developed a Python translation of the functionality of the MAPSS software. SWOSU student Jack Gillory took the lead on translating the MAPSS source code to run on a Linux distribution and standard desktop computer. Charles Sleeper began developing a Raspberry Pi Cluster to support future research partnerships with area High Schools.
- May 2016: The SWOSU HPC team participated in the first Oklahoma HPC competition. Students deployed a Raspberry Pi cluster. Students also developed code to find the determinant of a four by four matrix that is used in the MAPSS software. The MAPSS software continues to be in use after two decades of development. The team of software engineers at the NASA Goddard Space Flight Center had tuned the software for optimal performance on several supercomputers, including the Discoverer 12. Jack Gillory was able to deliver functionality with four of the nine modules within the software.
- August 2016: SWOSU students participated in a NSF C* Grant proposal. The SWOSU team coordinated a Cyber Infrastructure proposal for three universities. During the fall 2017 Oklahoma Supercomputing Symposium, this plan was praised by members of the One OCI team as being one of the better written plans the state had seen in recent years.
- September 2016: SWOSU meets with Southeastern Oklahoma State University research faculty to begin collaboration on future NASA OK EPSCoR Grants.
- December 2016: Oklahoma State University provided SWOSU with eight servers from a retired cluster. This cluster was deployed using CentOS with the SLURM job scheduler. Students installed the operating system, set up the administrator accounts, and worked with university IT staff to enable access to the server from various classrooms.
- December 2016: The SWOSU HPC team reached out to Dr. Toshihisa Matsui for a copy of the NU-WRF model. The SWOSU COMSC 3133 Software Engineering students will begin analysis of the software. This software was used as one of the three components of the Student Cluster Competition at The 2016 International Conference for High Performance Computing, Networking, Storage and Analysis (SC16)

---

**Research Challenges**

- The MAPSS software contained hundreds of source files written in several programming languages over two decades by multiple authors.
- Some files for the software are automatically generated by the NASA supercomputer and are not automatically generated within a standard installation of Linux Mint. These files were manually recreated from error messages.