SOONER SUPERCOMPUTING TEAM

The Sooner Team

Meet the Team

Braden Starcher

Braden is a sophomore in Computer Engineering at the University of Oklahoma. He voyaged to SC13 and was hooked on HPC ever since. Braden is interested in applying his HPC experiences to solving big data problems in the future.

Chase Gough

Chase is a multi-disciplinary studies senior. His interest in HPC began when meeting the OSCER sys admins while working for OU IT. Chase's interest in the technology driving HPC has only grown since then, and he aims to support that growth through a career in the industry.

KatyFranks

Katy is a senior in Geographic Information Science at the University of Oklahoma. She became interested in HPC when she was a meteorology student and became fascinated with all the possibilities HPC has to offer. She hopes to pursue a career in HPC and GIS.

Mark Glose

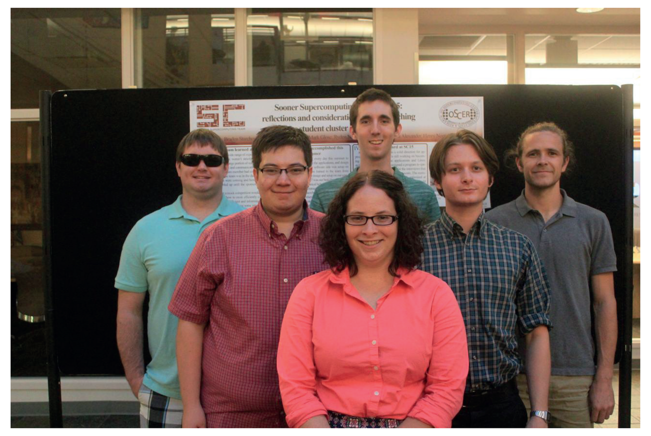
Mark is a junior in Computer Engineering at the University of Oklahoma. Mark became interested in HPC when he first attended the Sooner Supercomputing Club meeting. Mark wants to use HPC to explore advances in machine learning and big data.

Nico Scalzo

Nico is a junior at the University of Oklahoma pursuing a bachelors in Computer Engineering and a Masters in Electrical Engineering. Nico became interested in HPC when he took a tour of OU's cluster Boomer. Nico wants to use the skills he learned in efficiently running hardware to push the boundaries of the performance of consumer hardware.

Ryden Neal

Ryden is a junior at the University of Oklahoma pursuing a bachelors in Computer Science. Ryden became interested in HPC after learning of the club and competition. Ryden is interested in how HPC will further his CS education.



Back row, from left to right: Ryden Neal, Nicholas Scalzo, Chase Gough
Middle row, from left to right: Mark Glose, Braden Starcher
Front row: Kathryn Franks

Thanks to Our Sponsors

DELL intel Adaptive COMPUTING SanDisk OSCER INFORMATION TECHNOLOGY The UNIVERSITY of OKLAHOMA Mellanox TECHNOLOGIES

Supercomputing at OU & OSCER

Interest in supercomputing at OU began when three students attended SC13. The three decided to compete at SC14 and quickly acquired team members who were enticed by the tales of SC13. The OU Supercomputing Team and Club was officially initiated in the spring semester of 2015 after OU returned from SC14. Last year's team has stayed mostly involved with the Supercomputing club and has assisted the team this year as well. The Supercomputing Club reaches out to students by hosting general meetings with speakers from the field and professors with current or past HPC experience.

Founded in 2001, the OU Supercomputing Center for Education & Research (OSCER), a division of the University of Oklahoma Information Technology, helps undergraduates, grad students, faculty and staff to learn and use advanced computing in their science and engineering research and education.

This is accomplished via Cyberinfrastructure (CI) education, expertise and resources: For education, OSCER has been sponsoring "Supercomputing in Plain English," a series of presentations on core CI concepts that targets researchers in domain STEM disciplines; for expertise, OSCER works directly with researchers to advance the computing- and data-intensive aspects of their research; for resources, OSCER has the fastest academic supercomputer and the largest academic research archive in the state, as well as an internal research cloud.

How We've Prepared for SC15

The team is comprised of veteran and newbie members. Before the applications were announced, the veterans spent time assisting potential newbie club members get acquainted with working in a Linux environment and how to run code on OSCER's cluster. Once there applications were released, whoever was interested claimed an application and began to run jobs and attempt to find ways to optimize performance. We analyzed performance bottlenecks and used this information to choose an appropriate cluster architecture. We then worked with our sponsors advice and options to come out with our current cluster. Our team evolved the following responsibilities:
Braden – MILC, Sysadmin, Chase – HPL, Sysadmin, Ryden – Trinity
Katy – Team Captain WRF, Mark – Repast, Nico – Main Sysadmin

We have an advantage in that we aimed for a cluster design that would fit in as many nodes without adding complexity while keeping high speed interconnect. We chose to set up four SAS SSDs in the master node for faster read write and configured a RAID 0 for higher performance NFS. The slave nodes all have 1TB HDD. While looking into the nature of the applications during the summer we saw little need for accelerators. Some codes had ports for offloading to GPUs, but the added complexity. extra points of failure and power draw were not as worthwhile as adding an entire compute node. On the software side, we went with Centos 6.7. There was some issues with HPC Repast that didn't play well with Centos 7. We're using Adaptive Computing's Moab with Torque for scheduling. We used Allinea for reports and profiling. We believe we've come with a system that will not only just work but give us solid scores all around.

Applications, Optimization, Trials, Tribulations

HPL

Our cluster
2.3 GHZ x 16 Flops/Hz x 24 Cores/Node x 11 Nodes
≈ 9715.2 GFlops

With Turbo Turned on
3.1 GHZ x 16 Flops/Hz x 24 Cores/Node x 11 Nodes
≈ 13094.4 GFlops

WRF

Weather Research Forecasting is a mesoscale numerical weather forecasting and research model. It is used for operational weather forecasting as well as for research simulations. There are two different simulations WRF can run. Atmospheric simulations with real time (weather observations) or idealized case/hypothetical situation. Over the summer I spoke with many professionals working at the National Weather Center who utilize WRF on a day to day basis. They gave me tips as to what compilers are best to use for the program and other useful compilation flags.

MILC

MILC is a Multiple Instruction Multiple Data (MIMD) Lattice Computation (MILC) code that performs large scale quantum chromodynamics (QCD) simulations. QCD is the theory of strong interaction of subatomic physics. I began by compiling with GNU compilers--everything worked in accordance to the instructions mostly. Being sure of accuracy has been an issue, following the instructions and comparing by hand seems accurate, but the accuracy check executable says otherwise. Compiling the SCIDAC packages with Intel compilers also pointed to errors in the code that gcc was most likely forgiving. NERSC.gov and HPC Advisory Council had a resources for telling me the nature of the code that helped decide our cluster architecture.

Trinity

Trinity is an RNA sequencing application that had no issues compiling at first. When I compiled with GNU everything worked perfectly. For optimization I looked at compiling it with intel compilers since that was our processor. In order to compile I had to make several changes to some of the makefiles to replace some of the optimization flags that weren't working properly. After doing so Trinity compiled correctly. The use of Intel compilers provided a great boost to performance with the same data run in Trinity. An interesting thing I have found is that the application has been worked on by numerous people since all of the coding files contained different formatting and naming conventions.

Cluster Specs

11 Nodes
Dell Power Edge 430
Intel E5-2670 v3 Xeon 2.3 GHz
64 GB RAM
1 TB SATA HDD/Slave Node
4 SAS SSDs in Master Node
Mellanox FDR10 Infiniband
Mellanox FDR 10 Switch
Dell X 1018 Ethernet Switch

HPC Repast

Over the summer I worked on getting HPC Repast 2.0 and Repast 2.1 up and running. I then learned there would be a version of HPC Repast unique to this competition and we quickly started work on that. Adding HDF5 to the mix seemed to be a problem for me as I could get Repast compiled but we had some trouble getting it to run correctly. With the 2.1 version of Repast we did some tests and found that in non-square grid designs, it's more efficient to give the y axis more cores. What I found interesting about Repast is that there are non-HPC versions.