LANEY COLLEGE SUPERCOMPUTING

ABOUT US

In relation to our background, our team is extremely diverse. In fact, we come from: USA, China, Vietnam, Russia, Australia and Italy. The Bay Area is a melting pot of cultures, here we have the opportunity to interact with people from different background all the time. When forming the team for the supercomputing competition, we welcomed everybody. Thanks to that, we were able to form a very diverse team. We came from very different academic and career backgrounds into a similar space. With such diversity, we were able to fill well all of the engineering areas. When a problem would arise, we were collectively able to provide a valuable input.

Steve

Steve Leak has worked in HPC for about 20 years and is now a User Support Consultant at NERSC. He moved from Australia to the US five years ago for fun and adventure, and aside from HPC, data and software, he's into traveling, paragliding and watching the antics of his dog!

Nigel

Originally from Charleston, South Carolina, Nigel moved to the Bay Area to work in music. Since then he has pivoted to a career in technology. He transferred from Laney College to the University of California, San Diego to study Computer Engineering full time while working research internships for IBM and the San Diego Supercomputing Center.

Simon

Simon was born and raised in Eritrea, now he is US citizen. He is planning to transfer to UC universities. He has prepared for his major by excelling in a wide range of math and science courses along with other extracurricular activities like supercomputing competition and tutoring. He is also working on various programming projects to build a strong foundation for his future endeavors. He will be graduating in Math at BCC next summer. His passions are soccer, cycling, and traveling.

Dave

Dongyao Liang immigrated from China to the US when he was 18. Spent the first year to study English and then went to Laney College to learn computer science. He'll transfer to a 4-year university in 2019. So far, he's been keen on HPC and software designs. He's also into pool hall, watching meme videos and other indoor activities.

Noel

Born and raised in Italy, Noel moved to the States in 2016 to pursue his studies in Computer Science. He will be graduating in computer science at Laney next summer. He is a member of Phi Theta Kappa with 4.0 GPA planning to transfer at UC Berkley. Along with the supercomputing competition, Noel is doing an internship for as a Backend Engineer. He is trilingual in Italian, Spanish and English. His passions are soccer and traveling.

Khuong

My Name's Khuong Dinh. I am from Vietnam and an international student here in the States, I just finished an Associate degree in computer science and look forward to gain more knowledge, skills and opportunities. I am in the middle of a process to immigrate to the US so thing's abit hassled, but I thanks that this competition welcomed me and let me experienced such practical and team work environment.(btw I'm not a spy... OR AM I? lol kidding)

Jameelah

Jameelah is born and raised in San Francisco. She plans on transferring into the UC System to study Electrical Engineering and Computer Science. She has a back ground building and calibrating particle mass sensors at Lawrence Berkeley National Laboratory. She's currently doing research on energy harvesters at Berkeley Sensor & Actuator Center. In her free time she enjoys cycling, reading, and cooking.

Daniel

Daniel Gens is a Site Reliability Engineer at National Energy Research Scientific Computing Center (NERSC) located at Lawrence Berkeley National Laboratory. At LBNL, Daniel supports the operations of two of the world's fastest supercomputers, focusing on system administration, Python software development and writing Web-based & API integrations. As a trans guy and an aspiring activist, he is working to promote inclusion and recruitment for trans and gender non-conforming people at LBNL and in STEM through his work as the co-chair of LBNL's LGBTQIA+ Employee Resource Group, Lambda Alliance. Dan also works closely with local high schools and community colleges and has co-authored a computer science program for minority high school students.

Pr. J Williams

Johnnie is a faculty member at Laney College in Oakland, where he teaches courses in Python, Java, C++, Data Structures & Algorithms. Prior to working at Peralta, Johnnie worked as a programmer for start-up companies that were pioneering innovative technologies such as biometric verification and smart card technology. He has also worked as a programmer for a leading video game company. Johnnie was born and raised in Oakland and graduated with a degree in Cognitive Science from Occidental College in Los Angeles.
Johnnie was previously the Apprenticeship Director at #YesWeCode, where he developed apprenticeship programs and sourced engineering talent for numerous companies in the Bay Area

Darrell

Raised in Berkeley, California, Darrell is currently a policy analyst and software engineer studying Computer Science at Peralta Community College. He's passionate about civic programming, being the integration of software innovation into government policy and greater good.

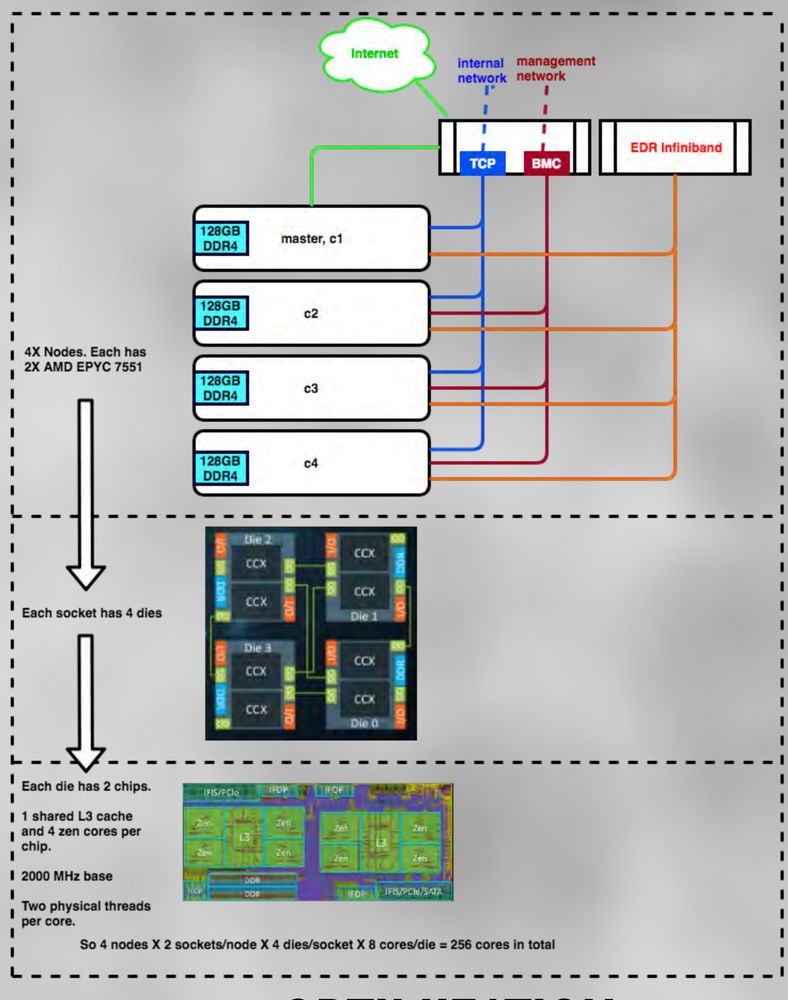
ARCHITECTURE

Hardware provided by Cray, based on AMD EPYC CPUs & Mellanox Infiniband networking
OS: Centos 7.5 on all nodes
Warewulf as the provisioning tool (over internal network)
Resource Manager: Slurm
Package Manager: Spack
Mellanox drivers for Infiniband
Parallel File System: OrangeFS over Infiniband
Base Compiler Family: GNU 7
AMD Optimized C/C++ compiler (AOCC) and other AMD libs
OpenMPI 3

Each Node:

- 2 x AMD EPYC 7551: 32 cores, 64 threads, 2GHz base clock
- 128 GB DDR4
- 250GB flash drive
- 10GbE Ethernet
- Mellanox ConnectX-5 VPI adapter card for EDR Infiniband
- Baseboard management controller

NVidia generously provided four V100 GPUs, but the form factor of our nodes and the GPUs proved an insurmountable mismatch, so we are unable to incorporate these into our hardware



THE APPLICATIONS

- HPL: or Top 500, benchmark [1] is the most widely recognized and discussed metric for ranking high performance computing systems.
  We first looked for a couple of good block sizes wth a fixed problem size then increased the problem size to use most of our memory then configured other parameters. Finally we tried out AOCC and AMDlibs optimized libraries to optimize it.

- HPCG: is composed of computations and data access patterns more commonly found in applications. After configuring with different local dimension sizes, we found out that the 128x128x128 symmetric problem size is the largest size that fits into our memory. We also observed using 120x224x64 nonsymmetric problem size we got good performance

- OpenMC: is an open source Monte Carlo particle transport code focused on reactor modeling and reactor physics methods research. We spent some time working to get a parallel HDF5 installed. Then building OpenMC with GCC was smooth.

- Horovod: an open source library that improves on both obstructions to scaling:it employs efficient inter-GPU communication via ring reduction and requires only a few lines of modification to user code, enabling faster, easier distributed machine learning in TensorFlow.

- SeisSol: is a software package for simulating wave propagation and dynamic rupture based on the arbitrary high-order accurate derivative discontinuous Galerkin method. SeisSol is specifically built for Intel CPUs, so to run on AMD we built for Intel Haswell, since the EPYC instruction set includes all of the Haswell instruction sets.

PREPARATION

As a team completely new to High Performance Computing, most of our preparation has been centered around learning the field. Since the start of the year we have:

- Looked at computer architecture, and how nodes and cores and threads affect MPI and OpenMP
- Built a practice cluster from scavenged hardware at NERSC
- Assembled and installed our Cray CS500 nodes with Centos 7.5 and OpenHPC
- Solved Centos install problems, configured networking with internal and external networks and NAT to network across them, and learned how to secure a system
- Had a cloud computing introduction with IBM

- During preparation we ran applications on both Cori at NERSC and COMET at SDSC, as well as on our own cluster

OPTIMIZATION

We experimented heavily with different compilers and implementations of MPI to decide our final software stack for each benchmark.We used Arm Performance Reports to find whether HPL and HPCG configurations were compute or MPI bound

OUR STRENGTHS

This configuration gives us more CPU cores and nodes than most recent teams while staying comfortably within the power budget
Use of flash-only storage and EDR Infiniband for MPI will minimize the network and I/O overhead of applications, and the management network allows us to power nodes up and down from the master node

Sponsored by: