SDSC HPC Training
April 12, 2019

Mary Thomas
mthomas@scsc.edu
https://www.sdsc.edu/~mthomas/

What is High Performance Computing?

High Performance Computing most generally refers to the practice of aggregating computing power in a way that delivers much higher performance than one could get out of a typical desktop computer or workstation in order to solve large problems in science, engineering, or business. (HPC Wire)

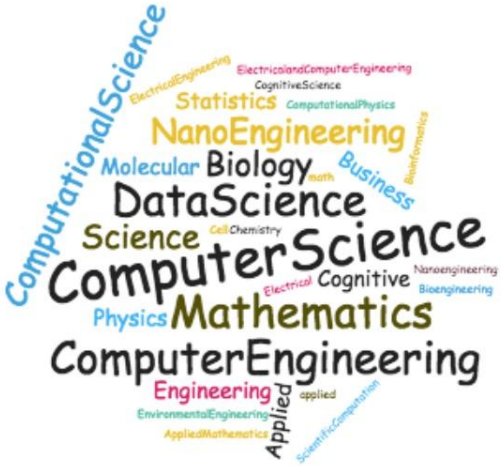
HPC-Students will get to play on Comet (and other machines)

- 2.762.6 PetaFlops (10^15)
- 27 standard racks, ~6.7 TF /rack with
  - 1944 nodes, 46,656 cores
  - 249 TB DRAM, 622 TB SSD

Agenda

- HPC Training Orientation
- About the Student Cluster Competition
- Pizza & Registration
- HPC Overview

SCC Program Orientation, Registration process & accounts; HPC overview (SDSC Team)



Goals of the SDSC HPC Students Group

http://hpc-students.sdsc.edu

- To facilitate and increase interactions between the San Diego Supercomputer Center and UCSD students:
- To educate and train students in all things HPC: parallel programming, running applications, learning hardware.
- To connect students to the wider world of HPC through events, meetings, interactions with industry, attending meetings.
- To mentor students to help train the next generation of scientists
- To collaborate with the UCSD Supercomputing club

Collaboration:

SDSC HPC Students:

- Provide space for group meetings
- Find budget and sponsors to support club activities
- Provide technical expertise for projects, tutorials, etc.
- Provide students with access to HPC resources and accounts
- Mentor projects, activities, events

http://hpc-students.sdsc.edu

UCSD Supercomputing Club:

- Approved UCSD Club
- Organize fun activities
  - Set agendas
  - Meeting times
  - Recruit members
- Find sponsors
- Organize and promote club activities and events

https://supercomputingclub.ucsd.edu

Supercomputing Student Activities

- HPC Training
- Apply to be on the student cluster competition (SCC) team - 6 students + alternates (budget permitting)
  - all travel expenses paid - travel to Denver, Co in November
  - Significant time commitment
  - Co-curriculum credit
- Mentor student volunteers at SC'19
  - https://sc19.supercomputing.org/program/studentssc/student-volunteers/
- ACM SigHPC Computing4Change competition: this is more individual but something worth supporting
  - https://www.sighpc.org/for-our-community/computing4change

About the Student Cluster Competition

SC19

Denver, CO | hpc is now.

SCC History/Background

http://www.studentclustercompetition.us/
https://sc19.supercomputing.org/

- Began in 2007 to provide HPC experience to undergraduate and high school students.
- Students design and build small clusters, learn scientific computing, run applications
- Compete in a non-stop, 48-hour challenge to complete a real-world scientific workload:
  - Operate and maintain the cluster
  - Run and test science applications
  - Reproducibility Challenge: students attempt to reproduce a science paper
  - Power outage challenge - restart system without warning

SCC: How to Compete

Site is now up: http://www.studentclustercompetition.us/2019/overview.html

ONLY 6 undergraduate students can be on the official team, and participate in the competition at SC

- Entering is competitive: we must write a formal application and submit this for review: 4/19/18
  - We must have the 6 members of the team for the application (can change later).
  - All travel and expenses will be paid for by SCC and SDSC
- To participate on the team requires a *significant* level of commitment:
  - Work during the Summer & Fall to learn HPC, applications, cluster hardware
  - Be able to travel to the meeting -- this year in Denver, Nov 17-22, 2019
  - SDSC will support a few alternates and/or graduate student mentors (budget)

SCC Team Preparation Process: 3 Steps

- Step 1: HPC Training
- Step 2: Summer HPC Cluster Buildout
- Step 3: SCC Team Finalization (August)

Step 1: HPC Training (general)

- Series of sessions:
  - SDSC Staff: Fridays from 1:00pm to 3:00pm from 4/12 to 6/7/19.
  - Supercomputing club: Select Thursday or Friday activities.
- Students will give accounts on Comet
  - Will complete several basic assignments
  - Some individual, some team
- Students who successfully complete the HPC Training program will:
  - Receive an SDSC Certificate of Completion in HPC Training.
  - Become eligible to apply to be on the SCC Core team.
  - Note: the final SCC team who will travel to SCC19 will be chosen from this group.
- *Evaluation metrics* will include:
  - participation credits - e.g. attendance, helping other students, etc.
  - Completion of tasks
  - Teamwork
  - Credit for experience (working/internships, classes taken)
  - Others TBD

Step 2: Summer HPC Buildout

- Approximately 20 students will be selected to work on the SCC core team.
  - Top group from HPC Training;
  - Students must be able to commit to the Summer/Fall buildout and possible to the SCC trip.
  - Meeting times will be chosen by the team once it is formed.
- Work throughout the Summer to:
  - Build out an HPC Cluster
  - Learn to admin the system: install software
  - Learn to run the suite of SCC science applications.
- UCSD Co-curricular credit for this activity (In progress)
  - All members of the team will receive credit.

Step 3: SCC Team Finalization

- The final 6 (plus alternates), will be chosen from the SCC core group.
- *Evaluation metrics* will include:
  - A team selection process
  - Participation
  - Ability to commit to travel (Denver, CO. from Saturday, 11/16/19 to Friday, 11/22/19)
  - Other factors TBD.
- Note: all team members will be participating in preparing for the actual meeting.

SCC Preparation Plans

- We have in-depth technical experience at hand -
  - Staff will be available to mentor and advise
  - Vendor mentor (ARM, Marvell, Hewlett Packard, AEON computing, Sandia Ntl Labs)
- We have team members who have participated at SCC in the past:
  - Student: Nigel Brown, 2018
  - Mentor: Mary Thomas, 2016 and 2017
- We have brilliant undergraduates (YOU)

SCC: Team Experience

- We have in-depth technical experience at hand -
  - Staff will be available to mentor and advise
  - Vendor mentor (ARM, Marvell, Hewlett Packard, AEON computing, Sandia Ntl Labs)
- We have team members who have participated at SCC in the past:
  - Student: Nigel Brown, 2018
  - Mentor: Mary Thomas, 2016 and 2017
- We have brilliant undergraduates (YOU)

Pizza break and Registration

Register for the Training:

https://www.eiseverywhere.com/ereg/newreg.php?eventid=419579&

SCC Sponsors & Technologies

HPC Overview

Early parallel computers….

Stone Soupercomputer (2003):

Cheapest cost/flop=$0, ~20 MFlops

https://web.archive.org/web/20031121211117/http://stonesoup.esd.ornl.gov

Comet: HPC for the “long tail of science:"

- “Long Tail” - majority of computational research is performed *at modest scale*: large number jobs that run for less than 48 hours, but can be computationally intensive and generate large amounts of data.
- Comet is an NSF-funded system available through the eXtreme Science and Engineering Discovery Environment (XSEDE) program.
- Advanced computing environment: supports science gateways, interactive computing, Jupyter notebooks, containers.

Comet System Characteristics

- **Total peak flops ~2.76 PF**
- **Dell primary integrator**
  - Intel Haswell processors w/ AVX2
  - Mellanox FDR InfiniBand
- **1944 Standard compute nodes (46,656 cores)**
  - Dual CPUs, each 12-core, 2.5 GHz
  - 128 GB DDR4 2133 MHz DRAM
  - 2*160GB GB SSDs (local disk)
- **72 GPU nodes**
  - 36 nodes same as standard nodes *plus* Two NVIDIA K80 cards, each with dual Kepler3 GPUs
  - 36 nodes with 2 14-core Intel Broadwell CPUs plus 4 NVIDIA P100 GPUs
- **4 large-memory nodes**
  - 1.5 TB DDR4 1866 MHz DRAM
  - Four Haswell processors/node
  - 64 cores/node
- **Hybrid fat-tree topology**
  - FDR (56 Gbps) InfiniBand (bisection)
  - Rack-level (72 nodes, 1,728 cores) full bisection bandwidth
  - 4:1 oversubscription cross-rack
- **Performance Storage**
  - 7.6 PB, 200 GB/s; Lustre
  - Scratch & Persistent Storage segments
- **Durable Storage**
  - 6 PB, 100 GB/s; Lustre
  - Automatic backups of critical data
- **Home directory storage**
- **Gateway hosting nodes**
- **Virtual image repository**
- **100 Gbps external connectivity to Internet2 & ESNet**

Open HPC PDF file

Task for Weeks 0 & 1

- Register for the Training:
- https://www.eiseverywhere.com/ereg/newreg.php?eventid=419579&
- Obtain your Comet account
  - Portal.xsede.org → create portal account
  - Send userID to mthomas@sdsc.edu, SUBJECT: HPC STUDENTS COMET ACCOUNT
    - Include xsede portal username
- Make sure you can logon
- Complete these exercises:
  - https://github.com/sdsc-training/webinars/tree/master/basic_linux_skills
  - https://github.com/sdsc-training/webinars/tree/master/getting_started