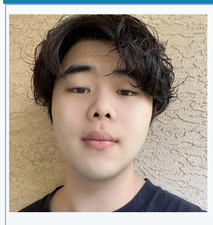
TEAM SUPERSCALAR UC San Diego

Jacob (Xiaochen) Li, Max Apodaca, Arunav Gupta, Zihao Kong, Hongyi Pan, Hongyu Zhou (Students)
Mary Thomas[1], Lewis Carroll[2], Marty Kandes[1], Mahidhar Tatineni[1] , Paul Yu[3] (Advisors)

[1] San Diego Supercomputer Center, [2] AMD, [3] Microsoft Azure

https://creazilla.com/nodes/55654-flamingo-emoji-clipart

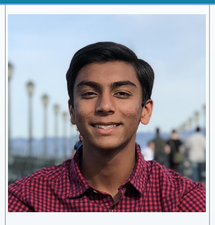
About the Team



Jacob Xiaochen Li
Computer Science, Math

Skills: Cloud Administration, PID Control, RL.

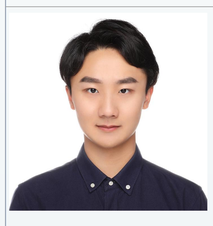
Role: Team Leader, Team Lead for Reproducibility Challenge (MemXCT)



Arunav Gupta
Data Science and Math

Skills : Machine Learning, Cloud Computing, Statistics

Role : Team Lead for HPCG Benchmark, 2nd Lead for MemXCT



Zihao Kong
Computer Engineering

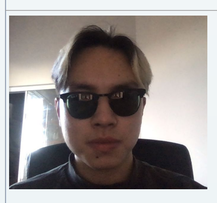
Skills: RTL design, FPGA, Embedded programming, CUDA programming, Cloud computing

Role: Team Lead for Gromacs

Max Apodaca
Computer Engineering

Skills : Cloud Computing, Containerization, Checkpoint/Restore

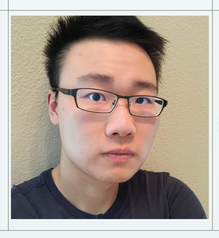
Role : Team Lead for IO500, CycleCloud Templates



Hongyi Pan
Cognitive Science

Skills : Machine learning, Natural Language Processing

Role : Team Lead for CESM



Hongyu Zou
Computer Science

Skills : Systems, Databases, and Networks

Role : Team Lead for HPL

Our team is composed of students from different majors, years, skills, and different parts of the world and reflects the diverse interests of UCSD undergraduates in HPC. However, it does not reflect diversity of the campus as well as we had hoped. Due to Covid, the opportunities to build a new program and to reach out to a broad range of organizations was limited. Our original team included one member from an underrepresented group, but unfortunately, this student had to drop out of the program. We'll start reaching out to students earlier next year.

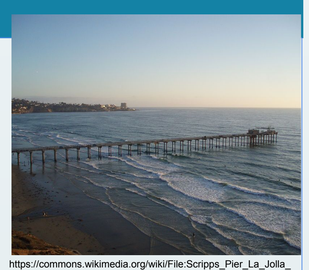
About UC San Diego & SDSC

We are next to the beach and all other facts are secondary.

UC San Diego has a large diverse campus located in the border city of San Diego. The nearby beach is a popular attraction and so are the various sculptures located around campus.

The San Diego Supercomputing Center is associated with UC San Diego and hosts the UCSD Supercomputing undergraduate association.

https://commons.wikimedia.org/wiki/File:Fallen_Star_by_Do_Ho_Suh_-_2.jpg



https://commons.wikimedia.org/wiki/File:Scripps_Pier_La_Jolla_California.jpg

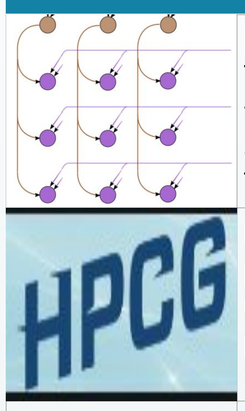
Team Preparation

- We started work on Comet, a local supercomputer, to learn how to compile and run applications
- Moved to custom Azure instance
- Ended with CycleCloud
- We have selected one image to keep it simple
- We selected different machine types to get best performance for each specific application
- Templates install commonly used components (Slurm + BeeGFS)
- We ran benchmarks on different SKUs for each applications letting the Azure documentation guide us while also calculating performance / $. These two metrics did not always line up in which case we chose performance / $.
- We plan to dry run the full competition (excluding the mystery app)
- To address availability issues we have backup SKUs for each application.
- We have practiced using Discord to communicate allowing us to adapt to the virtual nature of the competition.

Acknowledgements

THANK YOU TO OUR SPONSORS!

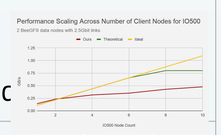
Application & Optimization Strategies

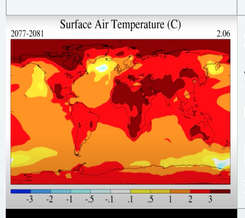


HPL - The LINPACK benchmark solves a dense system of linear equations. The model requires a VM that can provide dense computational power, so the **HC-series** VMs ($1.2/Tflops, 3.6 Tflops) would be a good fit. To accelerate the computation, we will use the **NCv3-series** ($0.44/Tflops, 7 Tflops * 4 GPU).

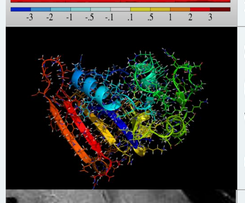
HPCG - This application is quite memory-bound, as the FLOPs are directly proportional to memory access rate. In order to optimize it, we will be using compute instances with high memory bandwidth (HB CPU), and possibly an NC-24 GPU accelerated model. The HB instances come with 50.51 Gb/s/$/hr, while NC is at 7.43 Gb/s/$/hr.

IO/500 - is a file IO throughput and latency benchmark.
The write speed scaling is sub linear with BeeGFS.
HB120rs has the best bandwidth / $, at 50.51 Gbps/$/hr.
We used a BeeGFS backend as it is simple, adaptable and c

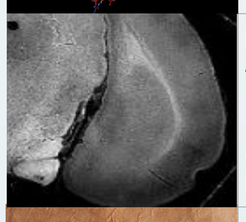




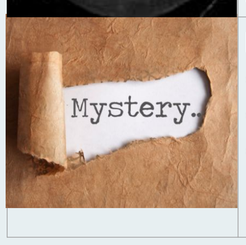
CESM - We used the HPC cluster ***Comet*** @UCSD as a testbed to learn how to Spack and modules to install and configure the model. For Azure CycleCloud we will use the CPU-based HB-series with BeeGFS for good IO performance. The most difficult part was configuring the obtuse system settings.



GROMACS - This application can be accelerated using a mixture of CPU (for the PME calculations) and GPU (for particle to particle calculations). We are choosing the **HC-series** for CPU and **N-series** for GPU. The major challenge for GROMACS has been installing software dependencies.



Reproducibility Challenge - MemXCT reconstructs X-ray CT images through implementation of *SpMV,* t*wo-level pseudo-Hilbert ordering*, and *multi-stage input buffering.* Memory bandwidth and latency results in significant performance bottleneck. Memory optimized VMs like **HB-series** and **E-series** will be a good fit.



Mystery App - We assume it will need some basic elements:

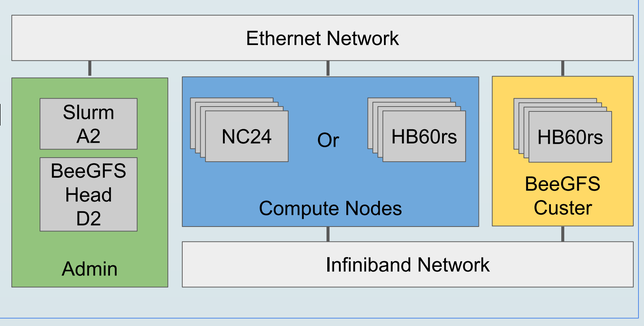
- Scheduler with Slurm
- Storage backend with BeeGFS
- Flexibility on type of VM (Computer, Mem Network)

Cluster Management & Cost Management Strategies

Key takeaways for staying under budget:

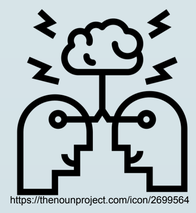
- Deploy faster instances to obtain better scaling for a similar cost / core.
  - Limit total speed but use fast instances.
  - Use HB60rs – single node scales better for throughput than multiple cheaper nodes.
  - Use the latest GCC (10.X) to get the best optimizations for the Zen 2 uArch while maintaining cross compatibility for applications like MemXCT.
- Obtain Cyclecloud cost reports for each cluster; use the CycleCloud CLI to monitor usage in real time.
- Develop HPC Software stack optimized for CentOS image, with a standard template that installs BeeGFS and SLURM.
- Customize images for our clusters which contain drivers for each of our SKUs, a common software stack and tools. This will save time when firing up new VMs.
- Profile the startup time and performance for each VM and application to predict costs more accurately

An example cluster setup for application deployment. The exact compute node would depend on the application but will almost certainly be Infiniband equipped unless it very memory intensive. BeeGFS will be the backend and could be scaled to cheaper nodes if storage is not a priority.



Why we will win

While our members come from diverse backgrounds we share a common passion for exploration. We are lucky enough to have a supercomputer in our "backyard" and every member plays on it daily. Once we had the opportunity to join the SCC competition, we jumped on it as we were excited and confident to learn more about HPC. Our super bi-weekly meetings have shown that everyone on our eclectic team is willing to give it their all. With that, our expert mentors, and attention to detail we know we will emerge victorious.



https://thenounproject.com/icon/2699564