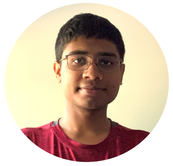
Husky Cows

SC20 Northeastern Williams

Meet the Team

Chris Bunn
Team Lead & MemXCT
Northeastern University
5th year, CE & Econ

Peter Zhao
Sys Admin & MemXCT
Williams College
4th year, CS & Math



Ritvik Rao
CESM
Northeastern University
3rd year, CE/CS & Econ

Karmen Liang
HPCG & IO500
Williams College
4th year, CS & Art

Carolina Binns
LINPACK & IO500
Northeastern University
5th year, CE

Andrew Jacob
GROMACS
Northeastern University
3rd year, CE/CS

Cloud Configuration

Software

- Each workload will be ran in its own designated Cyclecloud cluster
- Every cluster will run Ubuntu 18.04 and use SLURM for workload management
- A persistent BeeGFS cluster will be used to store files for all workloads
- Applications will be compiled using GNU compilers and nvcc (for GPU-based applications)
- For MPI-based applications, we will use OpenMPI where possible.

Hardware

Generally, our cloud strategy is to utilize AMD instances for CPU-only workloads because they have better core per dollar performance compared to Intel-based instances. We believe that the instances below optimize on key design features associated with each workload.

LINPACK / HPCG
ND40rs v2 as these utilize 8 NVLink interconnected V100 GPUs, and having high GPU performance is important

IO500
L80s v2 will optimize IO throughput and latency by using low-latency NVMe SSDs for fast data read/writes.

CESM
HB120rs v2 due to its 120 CPU cores, 480GB of RAM, and HDR InfiniBand support. CESM has been shown to scale well with the number of cores.

GROMACS
ND40rs v2, as our research on the Discovery Cluster shows that this workload can exploit the efficient inter-GPU communication offered with NVLINK.

MEMXCT
ND40rs v2 provides the most power and enough memory relative to what was tested in the publication.

MYSTERY APP
If GPU-intensive: ND40rs v2.
If CPU-intensive: H-series, perhaps HB120rs v2 due to its powerful 120 AMD EPYC 7002-series CPU cores.

Cloud Strategy & Budgeting

Our team will first divide the budget evenly amongst our applications for the first training period. Each team member will use their initial budget to begin testing their workload on the cloud and to determine their budget expectations for the final competition.

We will reallocate our budget accordingly for the second training session. Changes or observations in the second testing phase may result in fine tuning of the budget prior to the final competition.

At the start of the competition, we will run all of our applications (CESM, MemXCT, Gromacs, and the mystery application) simultaneously. Once the output of these applications has been verified, we will then utilize the remaining budget on our benchmarks. This approach ensures that we maximize our cloud budget.

Before and during the competition, we plan on using a Grafana dashboard to track our cloud spending and implement a hard shut off for all applications regardless of completion to ensure we stay within budget.

Our Strengths

Our team combines the skill sets and backgrounds of students from both Northeastern University and Williams College. We are excited to bring members from Williams on board, bringing a fresh perspective to our competition approach from previous years. Combined with our thorough optimization approach and hardware choices, we have a strong path to victory.

Diverse Backgrounds
We represent a range of experience, with rising 3rd, 4th, and 4th years majoring in Computer Science, Computer Engineering, Mathematics, Economics, and Art History

Industry & Research Experience
We have hands-on experience at companies such as Intel, Microsoft, Alteryx, Acacia Communications, Analogic, JPMorgan, and MongoDB, as well as research organizations such as MIT Lincoln Labs and MITRE.

Leadership & Community
We are involved in various campus clubs and have leadership roles in the Society of Women Engineers and Underrepresented Identities in CS, volunteered to increase accessibility to STEM education, and built community through Questbridge and Out in STEM.

We want to thank our advisors and sponsors:

Northeastern Information Technology Services

NCAR

David Kaeli
Professor, ECE
Northeastern University

Kelly Shaw
Associate Professor, CS
Williams College

Optimization Strategies

Standard

In preparing for the competition, our team laid out a general strategy to apply to all of our benchmarks and applications. Though not all of these steps are applicable to every workload, they provided us with a starting point for optimization. With each of the following steps, team members tracked changes in performance:

- Run application on the Discovery Cluster with a basic workload and record baseline measurements
- Enable multi-core support
- Enable single- or multi-GPU support
- Enable multi-node support
- Tune application specific parameters
- Graph data collected on performance improvements and scalability to make decisions about the ideal configuration
- Transfer over workload files to Azure Cyclecloud and recompile using optimal workload parameters

To minimize cloud spending, we will be using the Discovery Cluster to tune our workload parameters before transferring to CycleCloud. We will use hardware on Discovery that is as similar to our planned CycleCloud SKUs as possible.

Advanced

Beyond this general strategy, team members utilized several more advanced techniques, depending on their application. These efforts include swapping out core libraries, unraveling nested loops, and hand selecting compiler optimizations. While the gain from these techniques may be less significant than our standard strategy, we believe they will help set us apart from the competition.