

TEAM INTRODUCTION

This year's Clemson CyberTigers, advised by Dr. Jon Calhoun, represent diversity in gender, ethnicity, and perspective. Students are a mix of Computer Engineering and Computer Science majors and bring different skills from their unique backgrounds.



Kristen Guernsey

Application Lead: NAMD
Bio: Kristen has researched HPC data compression and fault tolerance and presented a paper at SC23. She interned at TACC as a system administrator and is interested in parallel architectures.



Benjamin Schlueter

Application Lead: ICON
Bio: Benjamin has done two HPC internships at LANL in high-speed networking, parallel file-systems, and data storage. He researches HPC fault tolerance at Clemson and presented a paper on posits and floating point at FTXS at SC23.



Samuel Quan

Benchmark Lead: Mystery
Bio: Sam has researched AI interpretability/privacy, lossy compression, and improving compilers for use in physics simulations. He is continuing his HPC research during the year.
Minor: Physics



Chloe Crozier

Benchmark Lead: MLPerf
Bio: Chloe interned as a cloud engineer at Deloitte and an automation developer for the Navy. She is working on a Capstone project with NVIDIA to develop a full-stack RAG application.
Minor: Mathematics & Economics



Ainara Garcia

Application Lead: Reproducibility
Bio: Ainara has done research in solving linear systems, data structures and algorithms, and complete/partial compression techniques. She was a SULI intern at Argonne National Labs and is presenting a poster at SC24.
Minor: Spanish Studies



Marcus McAbee

Benchmark Lead: HPL
Bio: Marcus has researched ML, signal denoising, and data compression. He is continuing his REU work on compression and has interests in computer architecture and hardware acceleration.
Minor: Mathematics

HARDWARE AND SOFTWARE

- The team is familiar with Clemson's Palmetto cluster, so we are configuring our cluster software stack to be nearly identical to it
- Our cluster contains a **head node** and **5 compute nodes** (3 CPU, 2 GPU)
- InfiniBand HDR** with **200 Gbps** (2x 100 Gbps) is used to ensure low latency node to node communication
- DDR5/4800 RAM** increases memory bandwidth to maximize throughput on memory constrained problems
- NVIDIA A100** GPUs offload work from CPUs for higher computational throughput

Hardware	Specs	Software	Details
Head Node	Model: PowerEdge R7515 CPU: AMD EPYC 16C/32T Memory: 128 GB	System	OS: Rocky Linux (8.10) Kernel: 4.18.0-372.32.1
GPU Nodes	Model: PowerEdge XE8545 GPU: NVIDIA A100 40G GPU TDP: 250 W CPU: 2x AMD EPYC 96C/192T Memory: 768GB DDR5	Parallelism	OpenMPI (4.1.4) OpenMP (5.2)
CPU Nodes	Model: R7625 CPU: 2x AMD EPYC 96C/192T Memory: 768GB DDR5	Compiler	gcc (12.1.0)
Storage	3.84 TB SSD	GPU Acceleration	CUDA (12.0.1)
NIC	InfiniBand HDR (200 Gbps) Mellanox ConnectX-6	Network Stack	Mellanox (OFED 23.04-1.1.3.0)
		Scheduler	Slurm (24.05)
		Package Managers	DNF (4.7.0) Spack (0.22)

BENCHMARKS

HPL – Marcus McAbee

- HPL evaluates the performance of computers, particularly HPC clusters, by solving **dense linear systems** using **double-precision arithmetic**
- Our A100 GPUs, based on the Ampere architecture, are packed with **high memory bandwidth**, significant **CUDA cores**, and are designed for HPC workloads
- We will leverage **NVIDIA-optimized HPC benchmark libraries** such as BLAS
- Team members developed HPL expertise by optimizing runs on our Raspberry Pi clusters



MLPerf – Chloe Crozier

- MLPerf is a benchmark for evaluating the performance of **ML/AI systems**, measuring how effectively they process inputs and generate results
- It tests inference performance on nine different models, including **LLMs**, text-to-image, **NLP**, computer vision, and medical image segmentation
- It covers both datacenter and edge case scenarios to capture a **full spectrum of performance**
- We will maximize our benchmark results by **balancing CPU-GPU workloads** and optimizing **network configurations**



Mystery – Samuel Quan

- To prepare for the mystery benchmark revealed at the competition, we will **survey and run previous SC and common HPC benchmarks**
- We will install **common libraries** and learn techniques used in well-known benchmarks
- Team members have experience running benchmarks such as **HPCG**, **STREAM**, and **OSU** from classes and SCC23
- We will test our **final hardware configuration** using any additional benchmarks we practiced while preparing for the mystery benchmark



REPRODUCIBILITY

Data Lifecycles – Ainara Garcia and Chloe Crozier

To successfully reproduce the results from the paper "*Data Flow Lifecycles for Optimizing Workflow Coordination*", we must:

- Pay attention to distributed and scalable measurement techniques
- Focus on how **DFL critical paths** and **caterpillar trees** are used to find bottlenecks and opportunities for optimizing workflows, particularly in task assignment and data movement
- Analyze the techniques proposed for workflow performance improvement, including **distributed caching**, **task-data locality**, and **anticipatory data movement**
- Pay attention to how workflows are represented using **DAGs** enriched with data flow lifecycles, which include tasks, data objects, and their interactions

POWER & TIME MANAGEMENT

- Tools and strategies such as **DVFS** and **under-volting** will increase our system's efficiency
- We will utilize **NVIDIA SMI** for precise GPU monitoring and management
- Our team will devise a detailed **competition schedule** to balance workloads and remain within the power budget
- We will create optimized **CPU-only** and **CPU+GPU builds** for certain applications to aid in load balancing

APPLICATIONS

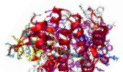
ICON – Benjamin Schlueter

- ICON is a flexible and scalable model for **weather and climate prediction**
- It uses **mesh refinement** with local zoom, simulating atmosphere and ocean with different parameters
- 192 CPU cores per node allows for high single node throughput
- Acceleration via **OpenACC** allows offloading of computation from CPUs to A100 GPUs
- We will use **nvfort** compiler optimized NVIDIA GPUs



NAMD – Kristen Guernsey

- NAMD is a parallel **molecular dynamics** application for simulating large biomolecular systems
- With the **Charm++** library, computation is partitioned across processors and nodes
- For peak performance, we will:
 - Restrict I/O operations** through parameter tuning
 - Find the desired **check-pointing frequency**
 - Determine the ideal **number of patches** for parallel efficiency
 - Run the optimal version for our cluster's hardware



Mystery – Marcus McAbee

- To prepare for the mystery application, we will practice running applications used in past SC competitions
- We have also compiled a **list of well-documented, scientific and mathematic applications** for HPC with example inputs and expected outputs
- Weekly meetings ensure that **everyone participates in troubleshooting and optimizing** the applications we use for practice



PREPAREDNESS

Why Will We Win?

- Team members **meet twice per week** to discuss our strategies and progress on application optimization
- We **attend webinars** for benchmarks and applications
- All team members have built **Raspberry Pi clusters**
- We are enrolled in classes based on **CUDA**, **parallel computing**, **networking**, and **OS**

