SGX3's FACULTY HACK @ GATEWAYS' 24

Training for Faculty Mentors

9/12/24 18:00 - 20:00 ET

Agenda

- Welcome from Dr. Sandra Geising, *SGX3 PI*
- Overview of the Faculty Hackathon Logistics
- Deliverables (GitHub, Poster Template, Blog Post)
- First Steps when meeting your faculty by Elijah Maccarthy
- Lessons Learned from past Faculty Hacks by John Holmen
- What Mentors are asked to share with the faculty
- Mentor's contribution by Dr. Bernadette Boscoe
- Mentor's contribution by Dr. K. Munene
- Faculty/Mentor teams for 2024

Overview

Event Information

The hack will begin with a series of virtual sessions and will conclude with an in person poster presentation of team findings at the Gateways 2024 conference.

Schedule:

September 16th - 20th, 2024 (6pm - 8pm ET) - Virtual Sessions [Two optional sessions Sept. 16th and 19th)

October 8th - 10th, 2024 - Poster Session [Gateways 2024 will take place in Bozeman, MT.]

Overview

The FacultyHack@Gateway2024 will involve 10 Computer Science or science discipline area faculty. Faculty teams will adapt High-Performance Computing (HPC) tools for use in their courses. They will leave with "ready-to-go" course outlines, supporting data, and identified resources. Each team will be assigned a technical mentor to help with this process. Teams completing all four (4) challenges receive a $1000 honorarium.

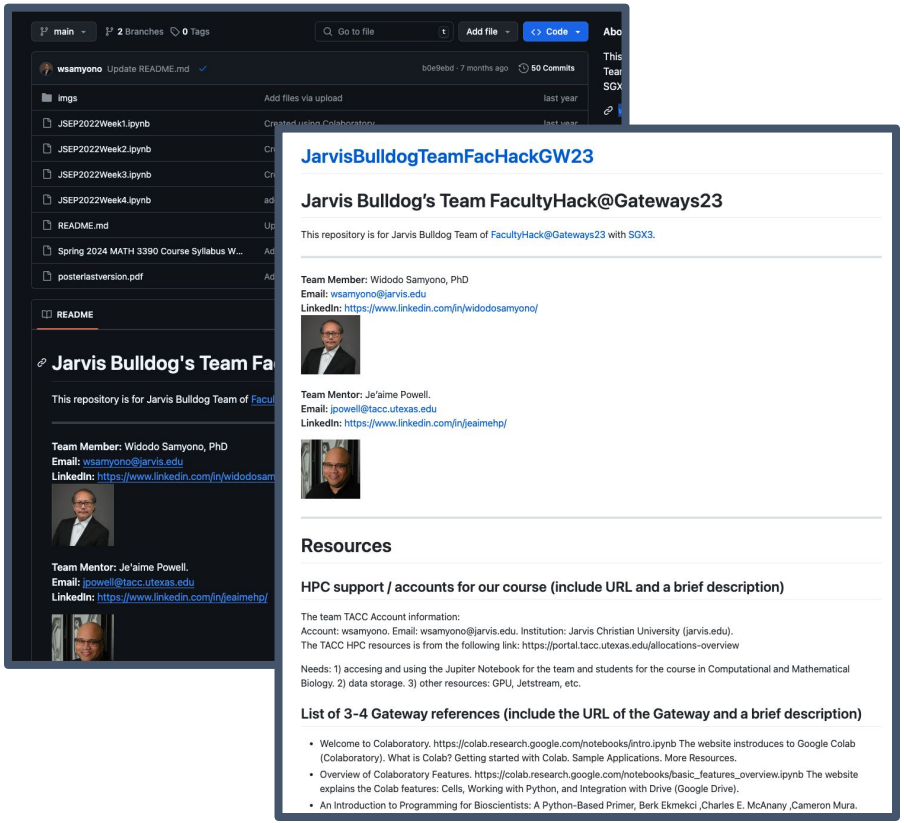
Challenges:
Attend all HPC training sessions Attend the Gateways 2024 conference in Bozeman, MT (Travel support is provided); Make a poster presentation of revised courses at Gateways 2024; Produce a Blog Post on your SGX3 Curriculum project which will be uploaded to sciencegateways.org/networking-community/blogs; GitHub repository with poster, README.md, description and code/datasetws

Outcomes:
A completely revised course description with implementation schedule. Assignment of a Gateways community mentor to provide use cases, resources and next step suggestions. Robust access to HPC resources for research and instruction. Opportunities to collaborate with other HPC educators and technical personnel. Enhanced computer science courses with HPC content at the home institutions.

Deliverables

- GitHub Repository
  - README.md
    - Project description
    - Blog post
    - Code/Datasets
  - Poster
  - Curriculum
- Poster presented at Gateways24

GitHub Example



GitHub:
https://github.com/wsamyono/JarvisBulldogTeamFacHackGW23

GitPages:
https://wsamyono.github.io/JarvisBulldogTeamFacHackGW23

Poster Template 2024

SGX3's FACULTY HACK @ GATEWAYS' 24

Title: Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam.

SGX3 award # 2231406

Revised Course Description

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Implementation Schedule

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Sample HPC/Gateways Exercise

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Resource Needs/List

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Gateway Community Mentor Syllabus Suggestions

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Resources / Science Gateways

- Resource
- Resource
- Resource
- Resource
- Resource
- Resource

Use Cases

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Datasets

- Dataset
- Dataset
- Dataset
- Dataset
- Dataset
- Dataset

Possible Expansions

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Authors

Author Photo

Author 1 Name
Affiliation
Email

Author 2 Photo

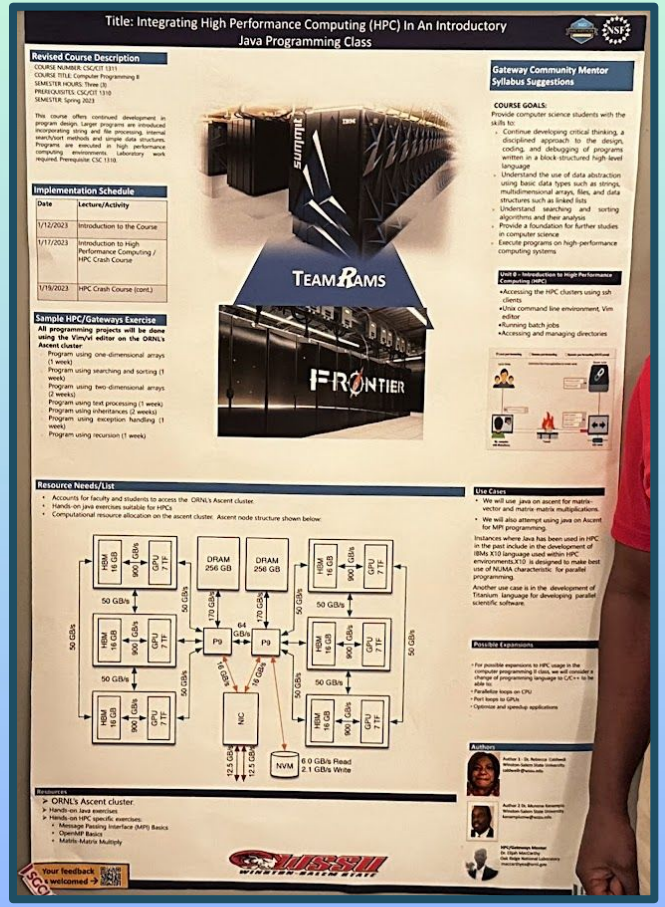
Author 2 Name
Affiliation
Email

Author 2 Photo

HPC/Gateways Mentor
Name
Affiliation
Email

MORE INFORMATION → https://hackhpc.github.io/facultyhack-gateways23

Sample Poster from 2022



Best Poster Award Gateways 2022

https://hackhpc.github.io/facultyhack-gateways24

Sample Poster from 2023

SGX3's FACULTY HACK @ GATEWAYS' 24

Scaling Up: Incorporating HPC experience into an undergraduate Introduction to Data Science course using Gateways with Jupyter notebooks and GPU-enabled instances

Revised Course Description

The **revised course description** will add a new module, ***Data Science using a GPU-enabled HPC*** at the end of the quarter-length course.

CS 356: Special Topics- Introduction to Data Science
The goals of this course are to learn how to acquire, clean, analyze, and visualize data using Python, libraries including Pandas, and Jupyter notebooks.

In the revised description, the added module will have students explore an HPC platform, detect GPUs in their Jupyter notebooks, and load a large dataset to build a machine learning model. Given code to train a model, students will learn how to run jobs on the HPC cluster, and store results for analysis.

Learning outcomes:
Students will

- Familiarize themselves with HPC environments and workflows to do analytic tasks
- Move data and code into the compute space, build a machine learning model, save it, and test it
- Utilize GPUs for training and explore the exciting realm of HPCs for scientific research

Implementation Schedule

Fall 2023:
Explore Gateways resources, using ACCESS-CI obtain accounts for instructor and ensure availability for ~25 students to have their own spawned containers

Winter 2024:
Develop Jupyter notebooks, load data and notebooks into an instance for testing, check for compute-credits needed to run the module, budget as needed

Spring 2024:
Begin course, add and test student accounts, implement module

Summer 2024:
Refine/refactor material, prepare for second offering Fall 2024/Winter 2025

Your feedback is welcome!

Sample HPC/Gateways Exercise

Image classification using machine learning is an effective way to introduce HPCs and the necessity of GPUs to newcomers.

We explore Cropnet classifier[1], a Tensorflow model that takes images of cassava leaves as input and detects various diseases if present. Cassava root is a major source of food across the world.

Previous to this exercise, students will have already studied the cassava project and its aims, and tested the existing trained model and code. Now, the focus is on retraining the model using GPUs on a cluster, in a platform. In this exercise, students write code to 1) detect GPUs, load the training, test, and validation sets, and train the model and save it.

```
[ ] plot(examples, predictions)
```

[1] https://github.com/google/cropnet/blob/master/cassava_disease_V1/2

Resource Needs/List

- HPC Platform that allows for ~25 student accounts GPU access, ideally GitHub authentication to access containers
- Jupyter notebooks in HPC environment, each student with JupyterHub spawned Kubernetes cluster
- Way to load and download data
- Compute credits to train ~25 ML models

Gateway Community Mentor Syllabus Suggestions

Mentors suggested to show the utility of using HPCs at scale, and to experience what happens when local computers cannot compute certain tasks requiring a GPU.

Proposed activity: Have students attempt to train the model on their laptops or desktops. It will likely fail, or take a considerable amount of time.

Next, have students try to train the model on the free version of Colab. Here, too, it will likely fail, timeout, or take a considerable amount of time to run.

Next, have students train the model on an HPC instance. Students will see the necessity and efficacy of using HPC instances to train machine learning models requiring GPUs.

Resources / Science Gateways

- JupyterHub: for notebooks
- TACC: for clusters and jobs
- Jetstream2: for clusters and jobs
- Exosphere: for interfaces
- SciServer: for domain science resources

Use Cases

- Students will access TACC Jto learn how to familiarize themselves with cloud-browser-based platforms, log in to an instance, and understand how to run jobs
- Students will use Jetstream2 to have their own container spawned for them, see pre-loaded configurations and then add libraries as needed
- Students will use exosphere to explore GUI interfaces to Jetstream2
- Students will use SciServer for domain science examples and compute environments

Special Thanks

Charlie Dey, Texas Advanced Computing Center
Je'aime Powell, Texas Advanced Computing Center
Linda Hayden, Elizabeth City State University

Datasets

- Cassava leaf image dataset: https://www.tensorflow.org/datasets/catalog/cassava
- Kaggle competition: https://www.kaggle.com/c/cassava-disease/overview

Possible Expansions

- From the Cassava disease detection model, students can create a TF-lite app to deploy and run on mobile phones to detect plant disease in the field

-Deploy a Cassava disease detection model to run on an endpoint and provide a method to upload images and detect them in batch or real-time

-With more cassava plant disease data or even different plant disease data, build a transfer learning model and test its efficacy and deploy it via an endpoint or a mobile app

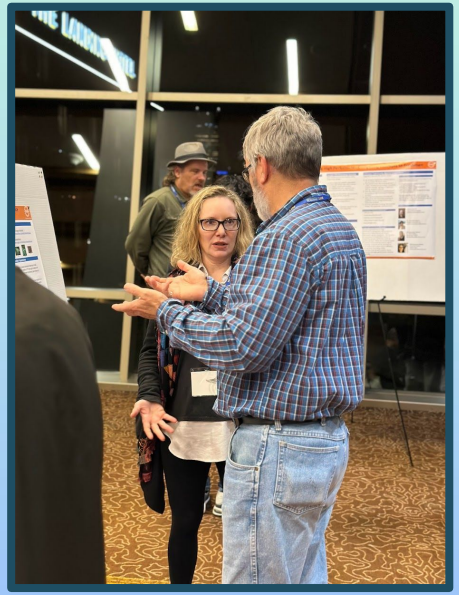
Authors

Bernie Boscoe
Assistant Professor, Computer Science
Southern Oregon University
boscoeb@sou.edu

HPC/Gateways Mentor
Mohamed Elbakary, PhD
Associate Professor of Electrical and Computer Engineering
Elizabeth City State University
melbakary@ecsu.edu

HPC/Gateways Mentor
Veronica Vergara
Oak Ridge National Laboratory
vergarav@ornl.gov

MORE INFORMATION → https://hackhpc.github.io/facultyhack-gateways23



https://hackhpc.github.io/facultyhack-gateways24

First Steps
when meeting your faculty

- Create action items for mentors and faculty
- Create slide for next check-in
- Overview and goals of the course from faculty
- How to collaborate between check-ins.

SGX3's FACULTY HACK @ GATEWAYS' 24

Lessons Learned from past Faculty Hacks:

John Holmen
(*Oak Ridge National Labs*)

Mentors share with Faculty

IDENTIFY HPC RESOURCES FOR CLASS ACCOUNTS

IDENTIFY DATA SETS

BE A VIRTUAL LECTURER FOR THE CLASS TO DISCUSS YOUR PROJECTS AND YOUR PROFESSIONAL JOURNEY

MAKE SUGGESTIONS FOR THE FACULTY'S NEXT STEPS FOR THE GATEWAY POSTER

IDENTIFY HOMEWORK ASSIGNMENTS, TEAM PROJECTS AND TEACHING EXAMPLE OPPORTUNITIES

HELP FACULTY SET UP THE ACCOUNTS AND DOCUMENT THE PROCESS FOR STUDENTS USAGE

https://hackhpc.github.io/facultyhack-gateways24

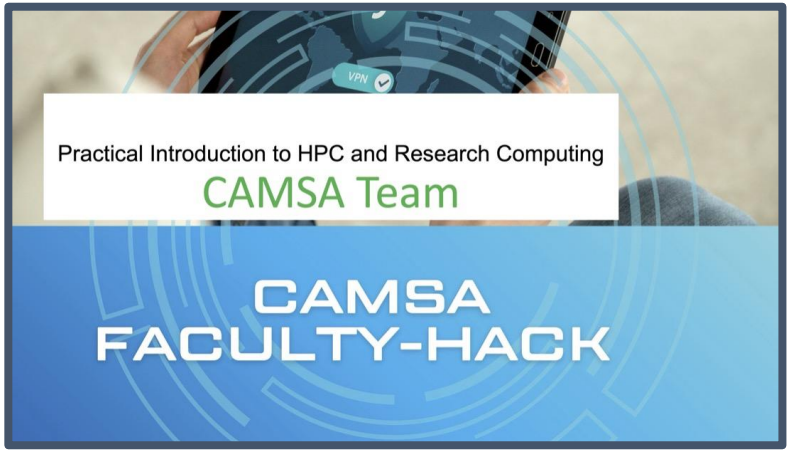
SGX3's FACULTY HACK @ GATEWAYS' 24

Mentors Contribution by Dr. Bernadette Boscoe (*South Oregon University*)

SDSC's FACULTY HACK @ GATEWAYS' 24

Mentors Contribution by Dr. K Munene (*Winston-Salem State University*)

https://hackhpc.github.io/facultyhack-gateways24



4: Syllabus

- This course provides exposure to advanced topics in computer networks including recent research findings in this field. The topics include: internetworking, Internet concept, Client-server model for applications, Network and internet management. Also, this course covers recently emerging protocols and technologies such as: Virtualization and Software Defined Networks (SDNs), IPv6, wireless networks, Secure Socket Layer, and Transport Layer Security.
- The course integrates also hands on labs about the usage of High-Performance Computing (HPC) in computer networks and other computing Disciplines. The goal is to allow students to use such resources in their other courses or future research or experiments.

SGX3's FACULTY HACK @ GATEWAYS' 24

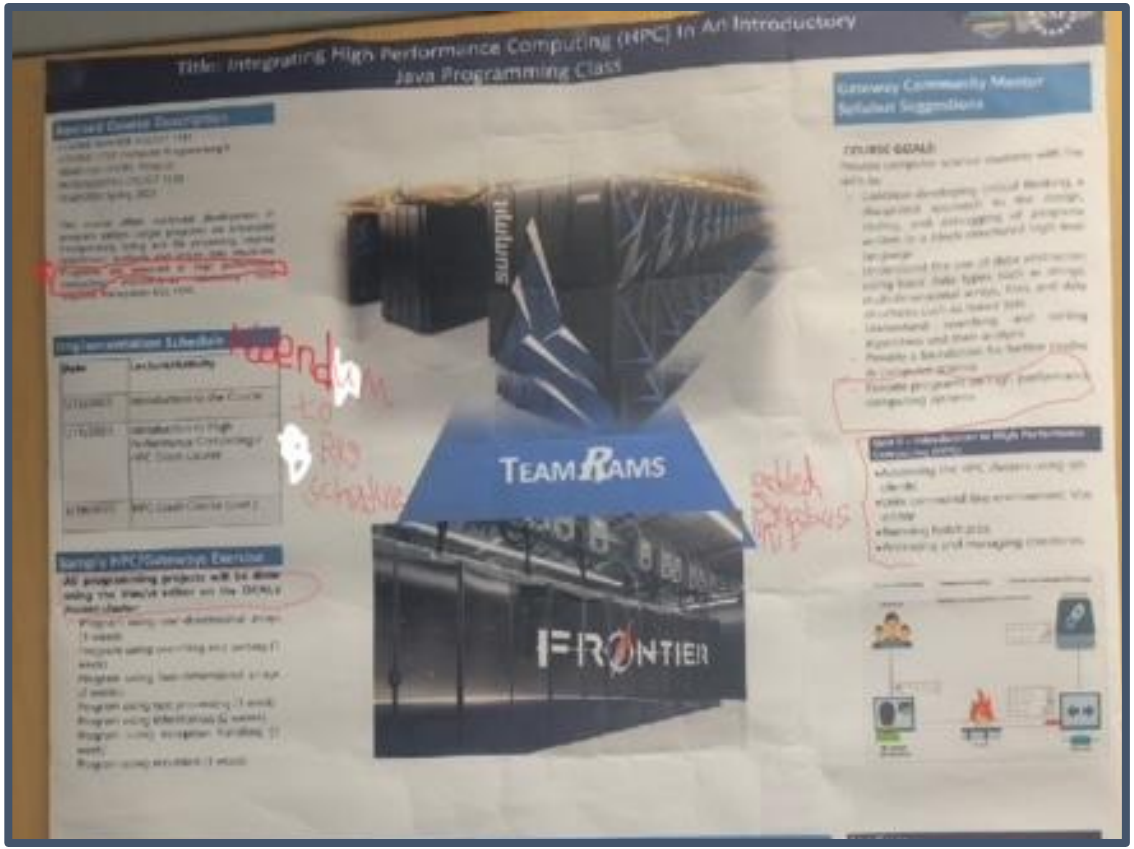
5: Sample Exercises

- In the past I used existing public materials in testbeds such as Geni.net, XSEDE, Deterlab, etc. Additionally, I provide samples of my own experience implementing those experiments and my own feedback to previous students.
- There is one particular example I used and like this year is ChameleonCloud shared experiments portal: (https://www.chameleoncloud.org/experiment/share/)
- The portal focuses on an idea that I like (reproducibility), and allows users to submit their feedback and also their own experiments.
- I am planning also to utilize the new TAMU cluster ACES (https://portal-aces.hprc.tamu.edu/pun/sys/dashboard)
- I am a member of SWEETER grant with TAMU with the aim of enabling the spread of usage of HPC resources across A&M system and the region.

6: Supporting Gateways

- https://access-ci.org/
- https://hprc.tamu.edu/

SC'23's FACULTY HACK @ GATEWAYS' 24



https://hackhpc.github.io/facultyhack-gateways24

2024 Faculty/Mentor Matches

Faculty	Mentor 1	Mentor 2
Ahmad Al-Omari	*Izzat Alsmadi*	*Charlie Dey ?*
Sabrina Perry	*Izzat Alsmadi*	*Fernando Posada*
Sungbum Hong	*Fernando Posada*	*Sam Fagbemi*
Shrikant Pawar	*Chard Kyle ?*	*LaTasha Roberts*
Nikhil Shrangare	*John Holmen*	*Sam Fagbemi*
Olabisi Ojo	*Hector Corzo*	*Boyd Wilson ?*
Lloyd Mitchell	*Charlie Dey ?*	*Sheryl Bradford*
Mohammed Elmellouki	*John Holmen*	*Sheryl Bradford*
Olamide Tawose	*Hector Corzo*	*Boyd Wilson ?*
Wanjun HU	*Chard Kyle ?*	*LaTasha Roberts*

https://hackhpc.github.io/facultyhack-gateways24

SGX3 Extend. Expand. Exemplify. | US RSE | Oak Ridge National Laboratory | Omnibond Engineering ▪ Trust ▪ Identity | TACC Texas Advanced Computing Center

SGX3's FACULTY HACK @ GATEWAYS' 24

Next Session:

Monday Sept. 16, 2024

Kickoff!