intel

cloudycluster by Omnibond

Google Cloud

SGCI

globus online

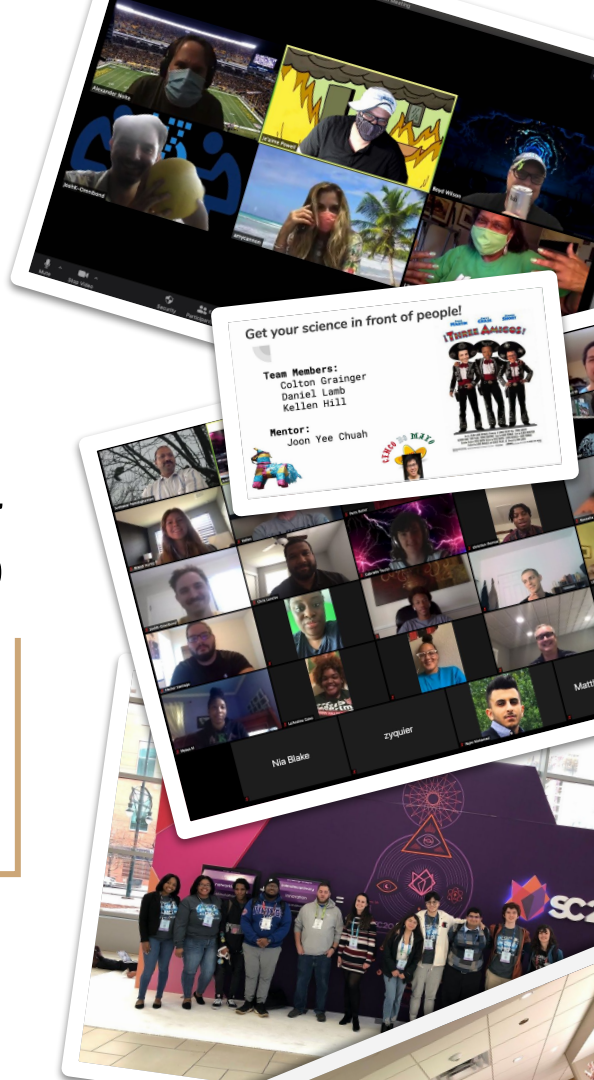
TACC
TEXAS ADVANCED COMPUTING CENTER

XSEDE

sighpc

HPC in the City
CloudyCluster Training

SC20
Everywhere we are | more than hpc.



Agenda

- Introductions
- Hackathon Objective
- Deliverables and Resources
- General Information
- Google Cloud Overview

Google Cloud Platform

Presenter: Je'aime Powell

Organizers

Alex Nolte - *University of Tartu*
alexander.nolte@ut.ee

Boyd Wilson - *Omnibond*
boyd@omnibond.com

Amy Cannon - *Omnibond*
amycannon@omnibond.com

Je'aime Powell - *TACC*
jpowell@tacc.utexas.edu

Linda Hayden - *ECSU*
haydenl@mindspring.com

The Objective of HPC in the City

The hackathon aims to harness the resources, skills, and knowledge found in the HPC community in an effort to provide applied exposure towards the conference host city's local students from 2-4 year post-secondary educational institutions. In short, the hackathon will provide HPC skills and training while targeting problems that directly affect the participants.

- Develop an understanding of an Atlanta based issue through application of data analysis/presentation or management.

What you should expect to gain:

- Increased familiarity with data science in the cloud
- Experience collaborative software engineering
- Develop professional communication skills

Team Deliverables and Resources

Deliverables:

- **Source code Including Comments**
- **PDF of presentation**
  - Team members with pictures
  - Use of HPC technology in the project
  - Regional (Atlanta) implications of the project
- **Github Link**
  - README.md project description

Resources:

- Mentors/Specialists
- Slack (Ad-Hoc Communication)
- Google Cloud (Provided Credits)
- Cloudy Cluster
- Most Commonly Used:
  - Python
  - Jupyter Notebooks
  - Node.Js (JavaScript)
  - HTML
- Datasets

General Information (the 3 T's)

- **Teams**
  - 4-5 Students
  - 1 Primary Mentor
  - 1 Specialist/Staff
- **Time (*Draft*)**
  - November 5th - 9th
    - 11/5@~6pm ET Event Start
      - Team formation
    - 11/[6-9] @ 11 ET & 6pm ET- Checkins
    - 11/9@6pm ET-Final Presentations

- **Topic Examples**
  - Data Analysis of COVID 19
  - Economic disparities and their effects on college participation
  - Genomics, Molecular Dynamics, or Weather Modeling in the Cloud.
  - Social Justice
  - Presidential Election
  - Public Data Management
  - Graduation Rates
  - Broadband Access
  - Insurance vs. Public Health Resilience

Presenter: Boyd Wilson

CloudyCluster and HPC/HTC Overview

Google Cloud Platform

High * Computing (H*C)

- **High Performance Computing(HPC)** [Generic]
  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.

- **High Throughput Computing(HTC)**
  European Grid Infrastructure defines HTC as *"a computing paradigm that focuses on the efficient execution of a large number of loosely-coupled tasks"*,[2] HTC systems are independent, sequential jobs that can be individually scheduled on many different computing resources across multiple administrative boundaries.

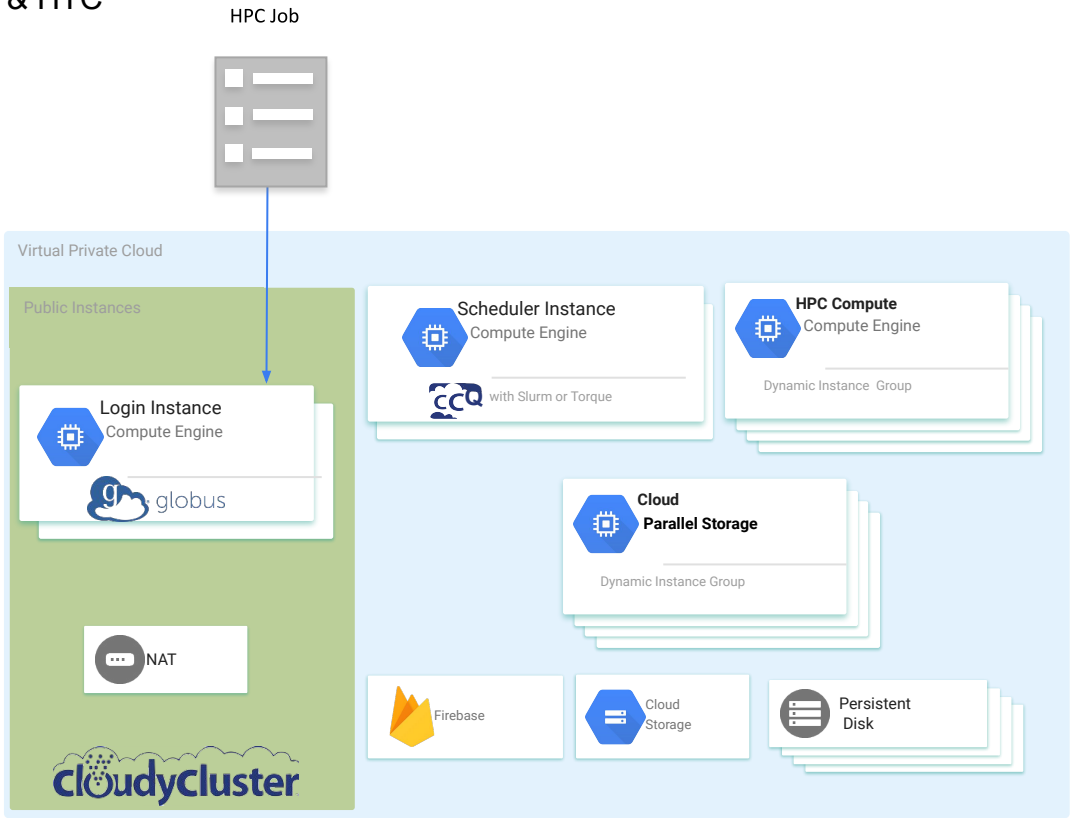
- **High Performance Computing(HPC)**
  HPC systems tend to focus on tightly coupled parallel jobs, and as such they must execute within a particular site with low-latency interconnects. Conversely,

- **Message Passing Interface (MPI)**
  Nodes need to interact throughout the job (dependant on data from other nodes)

- **Pleasingly Parallel**
  No dependencies between nodes, data can be processed in any order.

Architecture

Self-Service Elastic HPC & HTC



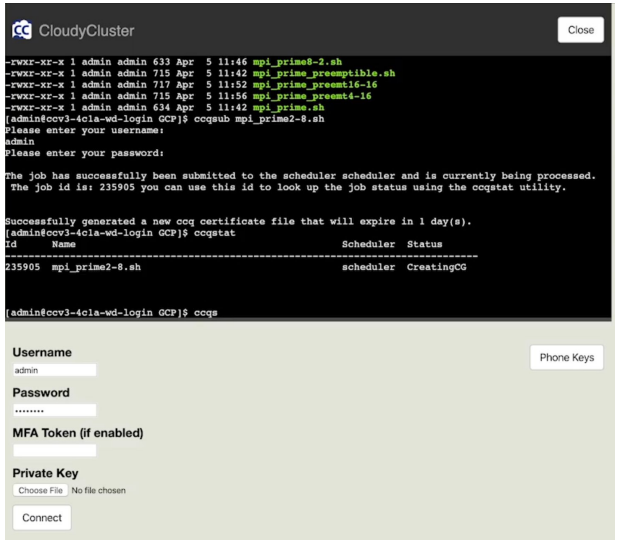
Google Cloud

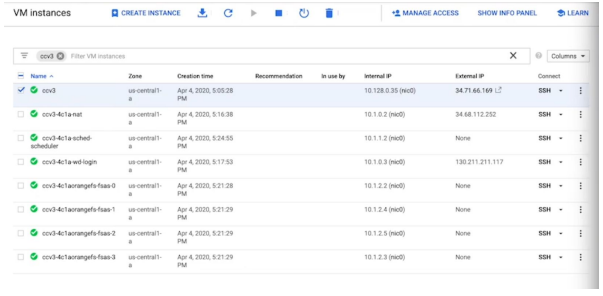
The Job (HPC not Italian)

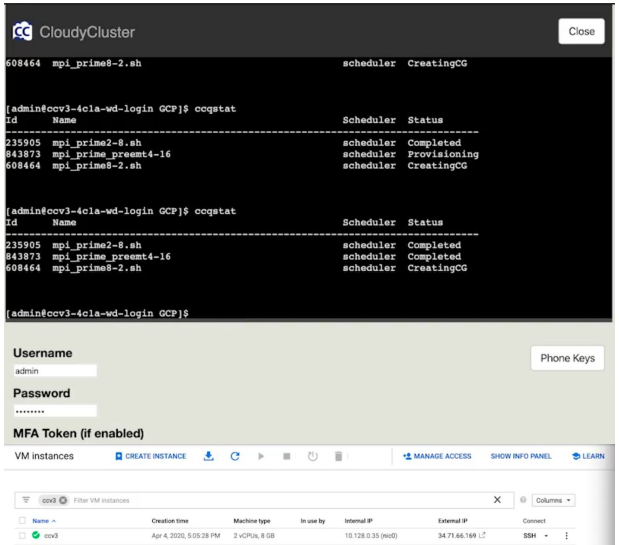
- HPC jobs are shell scripts with specific commented directives to tell schedulers what to do.

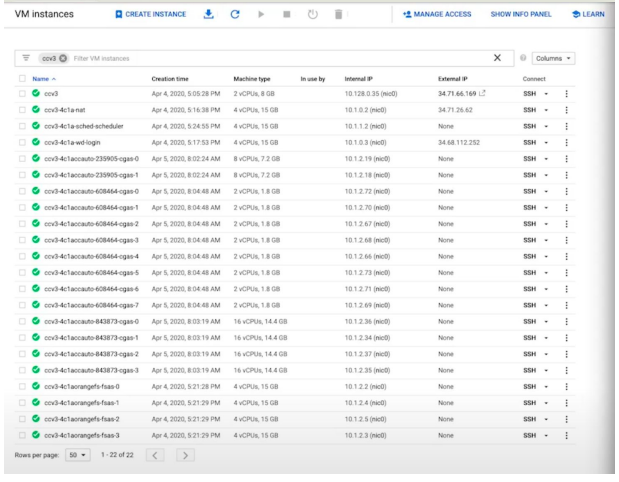
```
#!/bin/bash
# Indicates that CCQ should launch the instances in
preemptible mode
#CC -gcpup
#Uncomment this section for use with Torque/Maui
##PBS -l nodes=4:ppn=2
#Uncomment this section for use with Slurm
#SBATCH -N 4
#SBATCH --ntasks-per-node=2
#Need to change the location of the shared FS to the
name you specified in the CloudyCluster creation
wizard when launching the clus
Ter

export SHARED FS NAME=/mnt/orangefs
#Uncomment this section for use with openMPI
module add openmpi/3.0.0
#Uncomment this section for use with mpich
#module add mpich/3.2
cd $SHARED FS NAME/samplejobs/mpi
mpirun -np 8 $SHARED_FS_NAME/samplejobs/mpi/mpi_prime
```

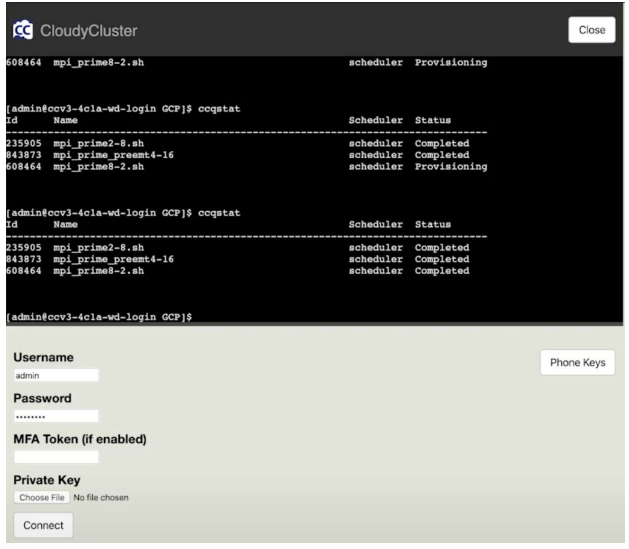
CCQ Job Execution

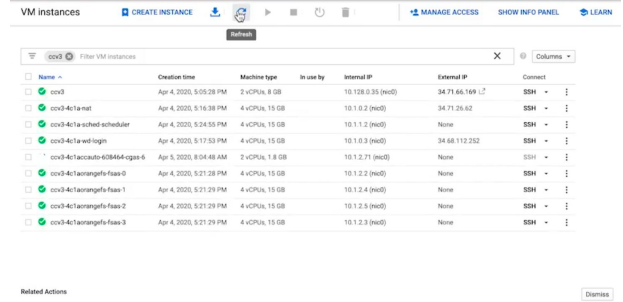












Live stuff!

General Information (the 3 T's)

- **Teams**
  - 4-5 Students
  - 1 Primary Mentor
  - 1 Specialist/Staff
- **Time (*Draft*)**
  - November 5th - 9th
    - 11/5@~6pm ET Event Start
      - Team formation
    - 11/[6-9] @ 11 ET & 6pm ET- Checkins
    - 11/9@6pm ET-Final Presentations

- **Topic Examples**
  - Data Analysis of COVID 19
  - Economic disparities and their effects on college participation
  - Genomics, Molecular Dynamics, or Weather Modeling in the Cloud.
  - Social Justice
  - Presidential Election
  - Public Data Management
  - Graduation Rates
  - Broadband Access
  - Insurance vs. Public Health Resilience

Questions and Concerns

Contact Information:

Boyd Wilson
(*HPC in the City Organizing Committee Member*)

Email: boyd@omnibond.com

Twitter: @boydwilson

HPC in the City Event Site: http://hackhpc.org/hpc/