Contribution ID: **58** Type: **not specified**

Addressing the Challenges of Executing Massive Computational Clusters in the Cloud

Tuesday, 26 March 2019 11:45 (25 minutes)

This talk will discuss how we worked with Dr. Amy Apon, Brandon Posey, AWS and the Clemson DICE lab team dynamically provisioned a large scale computational cluster of more than one million cores utilizing Amazon Web Services (AWS). We discuss the trade-offs, challenges, and solutions associated with creating such a large scale cluster with commercial cloud resources. We utilize our large scale cluster to study a parameter sweep workflow composed of message-passing parallel topic modeling jobs on multiple datasets.

At peak, we achieve a simultaneous core count of 1,119,196 vCPUs across nearly 50,000 instances, and are able to execute almost half a million jobs within two hours utilizing AWS Spot Instances in a single AWS region.

Additionally we will discuss a follow on project that the DICE Lab is currently working on in the Google Cloud Platform (GCP) that will enable a Computer Vision analytics system to concurrently processes hundreds of thousands of hours of highway traffic video providing statistics on congestions, vehicle trajectories and neural net pre-annotation. We will discuss how this project will differ from the previous one and how additional boundaries are being pushed.

Relevant Papers:
https://ieeexplore.ieee.org/abstract/document/8411029

https://tigerprints.clemson.edu/computing_pubs/38/

Primary authors: Dr HERZOG, Alexander (Clemson University); Dr APON, Amy (Clemson University); WILSON, Boyd (Omnibond); Dr POSEY, Brandon (BMW); GROPP, Christopher (Clemson University)

Presenter: WILSON, Boyd (Omnibond)

Session Classification: Grid, Cloud and Virtualization

Track Classification: Grid, Cloud & Virtualisation