CHEP 2016

INTEGRATING HPC, CLOUD, AND CONTAINERS FOR DATA-INTENSIVE COMPUTING PLATFORMS

DAN MURPHY-OLSON,
RYAN AYDELOTT

Highlights

- ARGONNE PROVIDES A BROAD PORTFOLIO OF COMPUTING RESOURCES TO RESEARCHERS. SINCE 2011 WE HAVE BEEN PROVIDING A CLOUD COMPUTING RESOURCE TO RESEARCHERS, PRIMARILY USING OPENSTACK. OVER THE LAST YEAR WE'VE BEEN WORKING TO BETTER SUPPORT CONTAINERS IN THE CONTEXT OF HPC. SEVERAL OF OUR OPERATING ENVIRONMENTS NOW LEVERAGE A COMBINATION OF THE THREE TECHNOLOGIES WHICH PROVIDES INFRASTRUCTURE TAILORED TO THE NEEDS OF THE SPECIFIC WORKLOAD.

- AS PART OF THIS WORK WE HAVE INTEGRATED TWO SYSTEMS AT ARGONNE, ONE PROVIDING PRIMARILY HPC RESOURCES, AND THE OTHER PROVIDING CLOUD COMPUTING RESOURCES.

- THE PRIMARY INTEGRATION POINT IS THE GPFS FILESYSTEM, WHICH IS MOUNTED ACROSS THE CLUSTER AND SHARED OVER NFS TO THE OPENSTACK TENANTS.

- HPC PROJECTS ARE MAPPED TO OPENSTACK TENANTS. USERS WITH ACCESS TO THE OPENSTACK TENANT HAVE FULL ROOT ACCESS TO THE GPFS PROJECT FILESET.

- USING THIS INTEGRATION USERS CAN EASILY USE THEIR OWN CUSTOM COMPUTE IMAGES TO ACCESS DATA STORED ON THE HPC FILESYSTEM. BECAUSE THE FILESYSTEM IS SHARED USING NFS TO THE TENANT NETWORK COMPLEX CLIENT CONFIGURATION IS AVOIDED.

- FUTURE WORK INCLUDES PERFORMANCE TUNING BETWEEN OPENSTACK CLIENTS AND THE GPFS NFS SERVER. ALSO, A LINK BETWEEN THE APPLICATION GATEWAY AND THE OPENSTACK SCHEDULER.