Contribution ID: **62** Type: **not specified**

Large-Scale Genomics Experiments Enabled by OSG Connect Resources

Tuesday, 7 March 2017 17:00 (20 minutes)

Several petabytes of raw DNA sequencing data have been deposited into public databases in recent years, introducing novel opportunities for mining useful biological information. The Open Science Grid (OSG) provides hardware and software infrastructure that have enabled us to address complex biological questions at a larger scale than previously possible with our local HPC resources at Clemson. With the help of the OSG support staff, we have developed two functional Pegasus workflows for processing and interpreting large genomic datasets. From software compilation to workflow optimization, we have encountered technical challenges that were quickly eased by the dedicated support of the OSG staff. The successes and challenges that we have encountered throughout this process will be discussed.

Presenters: Prof. FICKLIN, Stephen (Washington State University); Mr POEHLMAN, William (Clemson University)

Session Classification: Life Sciences