



Pegasus Users Group

MEETING



USC



Thank you for joining us. The presentation will begin shortly.

Pegasus is funded by the National Science Foundation under grant #1664162



Pegasus Office Hours!



Questions from Registration

- The new Python3 API and its use cases.
 - The new Python API allows users to compose workflows, create various catalogs Pegasus relies on. In addition, allows you to submit/monitor/debug your workflows
 - <https://pegasus.isi.edu/documentation/reference-guide/api-reference.html>



Questions from Registration

- Developing data-driven workflows, e.g., that listen to available new data sets.
 - Ensemble Manager has notion of triggers that can be used to spawn new workflows.
 - file pattern based cron trigger
<https://pegasus.isi.edu/documentation/reference-guide/pegasus-service.html#file-pattern-timed-interval-based-workflow-trigger>
 - Where does your data reside?
 - remote database
 - file server
 - cloud storage



Questions from Registration

- Computational chemistry, structural chemistry, molecular dynamics and Monte Carlo simulations
 - There are Pegasus workflows developed in these areas.
 - One of which is Nanodiamond refinement workflow provided by scientists at the Spallation Neutron Source (Oak Ridge National Laboratory)
(<https://github.com/papajim/SNS-Nanodiamond-Workflow>)
 - Regarding Parameter sweeping and Monte Carlo simulations, in the past we have explored interactions of Dakota (<https://dakota.sandia.gov/>) and MCViNE (<https://mcvine.ornl.gov/>) with Pegasus
 - Pegasus Hub examples



Questions from Registration

- Big Data, Data Management
 - Pegasus strives to provide complete data management for your workflows from
 - discovering locations of input data
 - staging input data/application containers efficiently to your execution environment
 - staging out outputs to user specified location
 - cleaning up data that is no longer required
 - <https://pegasus.isi.edu/documentation/reference-guide/data-management.html>



Questions from Registration

- Approaches and tricks that can be applied to various problems to receive high throughput.
 - The user guide has a list of tips and tricks
 - <https://pegasus.isi.edu/documentation/user-guide/optimization.html>



Questions from Registration

- Using containers (Docker and Singularity) in workflows.
 - Pegasus has well defined support for both docker and singularity containers. Users in their transformation catalog can refer to the container they want their jobs to execute in
 - <https://pegasus.isi.edu/documentation/user-guide/containers.htm>
!



Questions from Registration

- Batch/parallel processing jobs
 - <https://pegasus.isi.edu/documentation/user-guide/execution-environments.html#hpc-clusters>



Questions from Registration

- In situ analytics in Pegasus-based workflow
 - Pegasus has preliminary support for decaf as a means to run in situ applications
 - We have an ongoing research effort studying how Pegasus could benefit from in situ frameworks such as Decaf or DataSpaces





Thank You!

