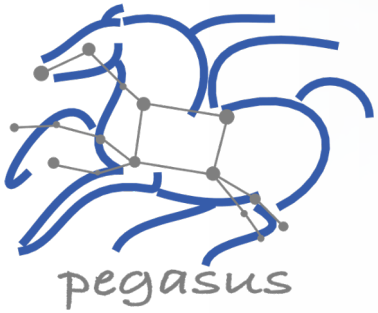


Pegasus 5.0 Tutorial

Workflow Management System

Karan Vahi

University of Southern California, School of Engineering
Information Sciences Institute
vahi@isi.edu

Setup

It is the same (but hosted) as the self-guided tutorial available in the Pegasus documentation: https://pegasus.isi.edu/documentation/user-guide/tutorial.html

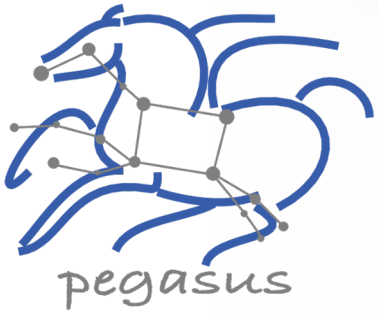
Please claim an instance by putting you name next to an unused instance in: shorturl.at/oEP12 (see Zoom chat for clickable link!)

Follow the link next to your name.

Docker Container / Jupyter Notebook

Container is for tutorial purposes - most production workflows have dedicated submit hosts

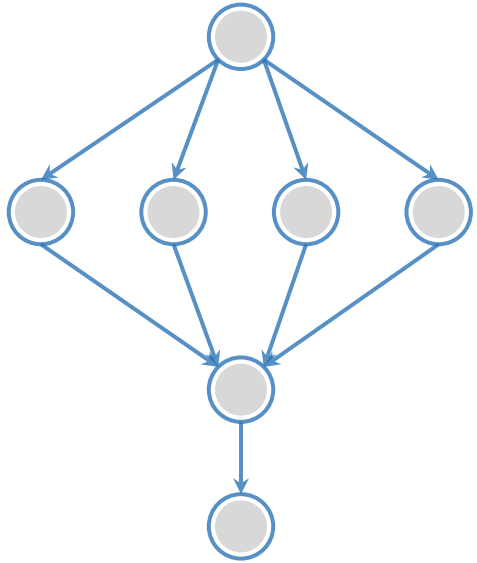
Jupyter is optional. You can choose to use just the workflow abstraction API, the full workflow management API, inside or outside Jupyter.

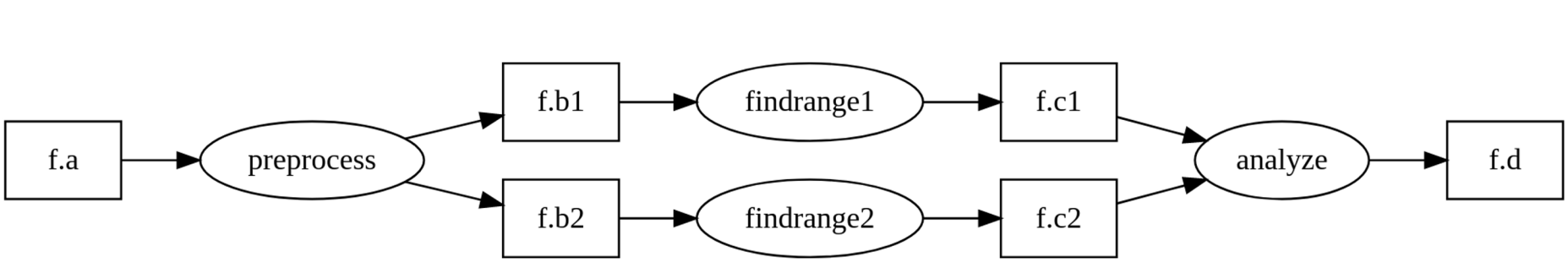


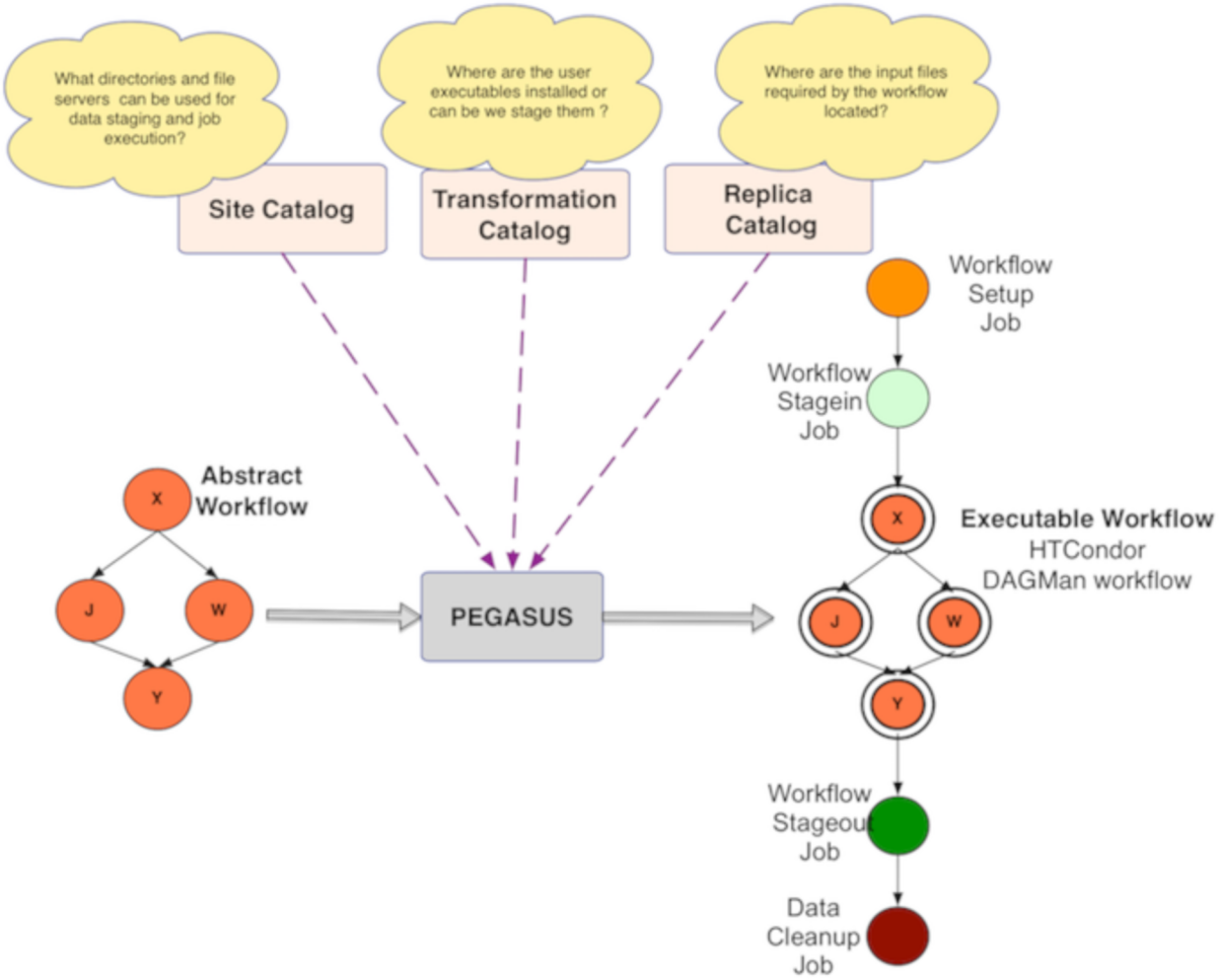
API

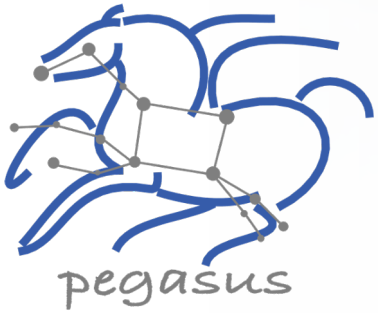
Key Pegasus Concepts

- **Pegasus WMS == Pegasus planner (mapper) + DAGMan workflow engine + HTCondor scheduler/broker**
  - Pegasus maps workflows to infrastructure
  - DAGMan manages dependencies and reliability
  - HTCondor is used as a broker to interface with different schedulers
- **Workflows are DAGs**
  - Nodes: jobs, edges: dependencies
  - No while loops, no conditional branches
  - Jobs are standalone executables
- **Planning occurs ahead of execution**
- **Planning converts an abstract workflow into a concrete, executable workflow**
  - Planner is like a compiler

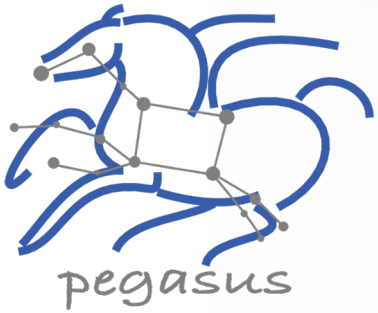



f.a
preprocess
f.b1
findrange1
f.c1
f.b2
findrange2
f.c2
analyze
f.d

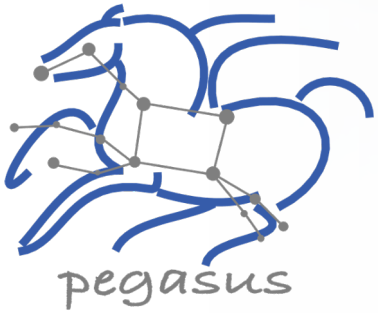

What directories and file servers can be used for data staging and job execution?
Where are the user executables installed or can be we stage them ?
Where are the input files required by the workflow located?
Site Catalog
Transformation Catalog
Replica Catalog
Workflow Setup Job
Workflow Stagein Job
Abstract Workflow
X
J
W
Y
PEGASUS
Executable Workflow
HTCondor
DAGMan workflow
Workflow Stageout Job
Data Cleanup Job



Debugging



Command Line Tools



Summary

Pegasus

est. 2001

Automate, recover, and debug scientific computations.

Get Started

Pegasus Website
https://pegasus.isi.edu

Users Mailing List
pegasus-users@isi.edu

Support
pegasus-support@isi.edu

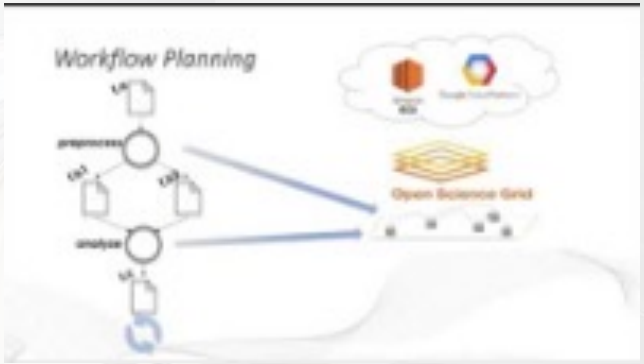
Slack
Ask for an invite by trying to join pegasus-users.slack.com in the Slack app

Pegasus Online Office Hours
https://pegasus.isi.edu/blog/online-pegasus-office-hours/

Bi-monthly basis on second Friday of the month, where we address user questions and also apprise the community of new developments

YouTube Channel

https://www.youtube.com/channel/UCwJQln1CqBvTJqiNr9X9F1Q/featured



Pegasus in 5 Minutes