Pushing the limits of job flexibility on HPC

Carrie Ganote

Guest starring: Asma Bankapur, Cicada Dennis, Tom Doak, Sheri Sanders, Bhavya Nalagampalli Papudeshi, Brian Haas and Phil Blood
The backstory

GALAXY SERVER

FILE SYSTEM
(LUSTRE)

COMPUTE NODE

RAMDISK
LOGICAL DISK
LOCATED IN RAM

CPU

LOCAL DISK /TMP
Rerun interface

Job

Wrapper script

Try #1

Try #2

Try #3

Time
Rerun interface

Job #1

Wrapper script
Create a directory using username and user-specified tag

User reruns job

Job #2

Wrapper script
Move to custom directory and pick up where the job left off
Rerun interface

Trinity De novo assembly of RNA-Seq data using Trinity on the Karst cluster (Galaxy Tool Version 2.4.0)

**Paired or Single-end data?**

*Paired* dropdown

**Left/Forward strand reads**

- 230: SRR073754_1.fastq

**Right/Reverse strand reads**

- 231: SRR073754_2.fastq

**Allow Job Rerun**

To make a job rerunnable, you will need to specify a unique tag to label the job, with no spaces or weird characters.

*test2*

**Execute** button
Less hacky ways of doing this

- Utilize the job resubmit-on-fail logic implemented by the Galaxy team
- Implement a more general mechanism for specifying the job working directory in job destinations
- Remove the requirement for user input to rerun checkpoint-friendly software
- Move these mechanisms out of the tool wrapper altogether
Our mad plans

IU Galaxy

Pulsar

National Center for Genome Analysis Support - ncgas.org
Fin

Thanks to:
Asma Bankapur
Cicada Dennis
Brian Haas
Phil Blood
Tom Doak
Sheri Sanders
Bhavya NP

Thank you for being awesome!