Galaxy Deployments at Indiana University

Carrie Ganote
National Center for Genome Analysis Support

April 16, 2015
About IU and NCGAS, and me

A bit about Indiana University:
Very culturally diverse campuses
Known for our Jacob’s School of Music and Kelley School of Business

A bit about NCGAS:
National Center for Genome Analysis Support
NSF/NIH/IU funded group offering bioinformatics support/Linux help and lab experience to researchers

A bit about me:
NCGAS Staff since Oct 2012
First year PhD student, woohoo!
An unhealthy obsession with turkeys

National Center for Genome Analysis Support: http://ncgas.org
National Center for Genome Analysis Support: http://ncgas.org
Fin

Thanks for watching!
Questions and comments:
Email help@ncgas.org
IU Galaxy

- Running since 2012 on Mason cluster at IU
- Deployed on Quarry, Big Red2 and soon to be Karst
- Apache proxy + Kerberos Authentication
- Lustre Filesystem
- Most tools, with emphasis on large memory applications
- University-wide, plus collaborators
- Still using hg

Fancy javascript tool menus!

National Center for Genome Analysis Support: http://ncgas.org
BLAST is an embarrassingly parallel program – it’s our only tool that uses the grid.
The Game Plan

Galaxy Host - VM
   Galaxy Framework
      Job
         PROFIT
            Pulsar Runner (localhost)
               Job script submits to OSG then submits to Mason

OSG
   Highly distributed grid
      Splits input files
      Creates Condor jobs
      Collects all results
      Compress and tar

Mason
   Large memory cluster
      Unpacks results
      Merges results in memory
      Returns final result

Photo by Soichi Hayashi. National Center for Genome Analysis Support: http://ncgas.org
Trinity Galaxy

Welcome to the Trinity Galaxy instance.

This instance has been configured such that only users who are logged in may use it. If you need help, send us a question or read the FAQs.

If you would like to register for an account, please click here.

Login

Email address:
cganote@gmail.com

Password:
******

Forgot password? Reset here

National Center for Genome Analysis Support: [http://ncgas.org](http://ncgas.org)
Trinity Galaxy

Trinity Galaxy Registration Form

Please enter the required information below.
Questions? Need help? Feel free to email the NCGAS team at help@ncgas.org.

Information about You

Full Name:
* must provide value

Institution name:
* must provide value

Institution location, city:
* must provide value

Institution location, state:
* must provide value

Institution location, country:
* must provide value

National Center for Genome Analysis Support: [http://ncgas.org](http://ncgas.org)
Trinity Galaxy

- One tool/pipeline emphasis
- Ftp server
- Local users

lib
galaxy
  config.py
  managers
    context.py
  tools
    parameters
      grouping.py
  api
    remote_files.py

National Center for Genome Analysis Support: http://ncgas.org
Challenges

Job Runners:
• PBS – still breaks with Torque 4x
  • Seems to have trouble interacting with the dynamic runner
• Drmaa – one destination possible?
• Dev servers are always somehow different than production!

Future Goals
• Trinity needs some more guinea pigs!
  Sign up now!
• IU Galaxy still needs to be moved to Github
• OSG-BLAST is still in beta
Fin (for real)

Thanks for watching!
Questions and comments:
Email help@ncgas.org

National Center for Genome Analysis Support: http://ncgas.org