

TGAC Galaxy server and training

Nicola Soranzo

nicola.soranzo@tgac.ac.uk



- Introduction
- Status of the new Galaxy server
- Ongoing projects
- Future work

- **Installation and maintenance** of a new **Galaxy platform**, hosted at TGAC, for internal bioinformaticians and external researchers
- Exploit TGAC's hardware and software expertise and resources, promoting the Centre as a National Capability for UK bioscience
- **Training** on Galaxy for BBSRC institutes
- Network with the Galaxy community

What is Galaxy?

- Web-based platform for data-intensive biomedical research
 - Open source (Academic Free License)
 - Developed by the A. Nekrutenko lab at Penn State and the J. Taylor lab at Johns Hopkins University
- >2500 citations <http://www.citeulike.org/group/16008/>
- >70 *public* Galaxy servers
 - None in the UK!

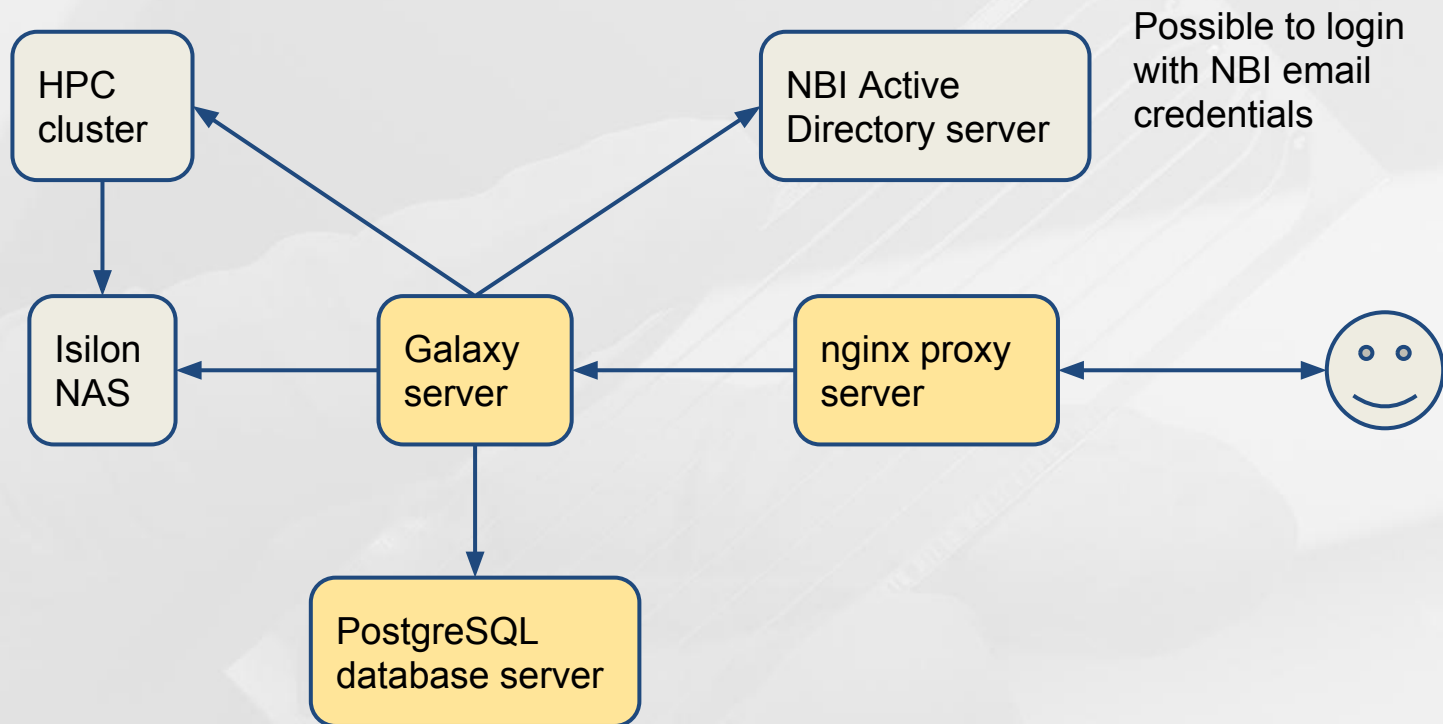


- Uniform access to bioinformatics tools
 - every software can be embedded
- Tool Shed: free tool store <https://toolshed.g2.bx.psu.edu/>
- Workflow: editing / extracting; reuse on different inputs; automatic scheduling of jobs
- Data visualisations
- Reproducibility: every analysis step is tracked
- Sharing of datasets, workflows and histories
- Scalability: can use cluster or cloud resources

New TGAC server architecture

<https://galaxy.tgac.ac.uk/>

Reuse software
modules



- NGS mapping: BWA, Bowtie, Bowtie2
- RNA-Seq: TopHat, Cufflinks suite, SortMeRNA, StringTie
- De novo assembly: ABySS, SPAdes
- Phylogenetics: ClustalW, T-Coffee, MAFFT
- Variant analysis: FreeBayes
- FastQC, Samtools, bedtools, Picard, NCBI BLAST+, EMBOSS

- Presently accessible only inside the NRP
 - We will open the server to the other BBSRC institutes...
 - ...and then to the rest of the world
 - Sustainability
- More tools and reference genomes/indexes
- Training!

- Presently one-to-one
- Survey to be sent to NRP to investigate the training needs in Galaxy
 - expectations, personal projects/data
- First TGAC course to be held early 2016
 - Second NRP event mid 2016
 - Focused events: RNA-Seq, metagenomics...
- Train the trainer for Galaxy?

Acknowledgments



Rob Davey



Tim Stitt, Chris
Bridson and CiS
team @ NBI



Vicky and Emily
@ 360° Division



Galaxy Team and community