Galaxy in Production at the University of Minnesota

Galaxy Admin Meeting, June 2016 James Johnson and Evan Bollig

MN Supercomputing Institute for Advanced Computational Research



MSI By the Numbers

- Full Time Staff: 42
- Student Assistants: 6
- Active User Groups: 594
- Active Users: 4,063
- Institutions: UMN, other EDUs in MN, Commercial
- Services:
 - 1. HPC; 2. Interactive HPC; 3. Data storage; 4. Webs/DBs; 5. Consulting
- Compute Clusters:
 - Mesabi, 675 TFlops (peak), >18,000 Cores, 67 TB RAM
 - Itasca, 100 TFlops (peak), 8,744 Cores, 31.3 TB RAM
 - Big (Hadoop): 40 Nodes, 10TB RAM, 1.2 PB HDFS
 - Stratus (OpenStack): 20 Nodes; 5TB RAM
- Storage Systems:
 - Primary: 4.1 PB usable, Panasas
 - Secondary: 3.1 PB usable, CEPH/HP&Supermicro
 - Archive: 3.5 PB Tape (BlackPearl)
- Active Projects: 100+





MN Supercomputing Institute for Advanced Computational Research





MN Supercomputing Institute for Advanced Computational Research



Who Uses MSI?

Outline

- Three Production Use-cases
- Site-specific Configurations for Tier I Service
- Clinical Pipeline Details
- Improving the Tier I Experience



Galaxy.msi.umn.edu

- 2 Production servers available to UMN researchers
 - Genomics galaxy.msi.umn.edu
 - 855 active users
 - Data: 767.7 TB generated with 267.61 current, and 500.09 purged
 - Proteomics galaxyp.msi.umn.edu
 - 118 active users
 - Data: 13.38 TB generated with 3.18 current, and 10.20 purged
- Tier I service
 - Most jobs scheduled on HPC resources (PBS-Torque/Moab)
 - CentOS 6
 - Use HPC Panasas storage
 - (Anyone using Ceph for storage?)





NSF grant to adapt galaxy for proteomics MassSpec Proteomics Requirements:

- Workflows need to operate on collections of data files, the number of inputs may vary with each sample
- Some vendor software is only available for Windows
- Datatypes for proteomics
- Proteomics tools



Galaxy-P

Galaxy-P was developed by John Chilton

- Initially a fork of Galaxy
- Multiple file datasets permit workflows to operate with variable numbers of inputs
- LWR (Light Weight Runner) plugin staged files to another host to run applications, including Windows
- Proteomics datatypes and tools developed with notable contributions from Ira Cooke, Australia



Galaxy-P

Galaxy-P functionality merged into Galaxy framework

- John Chilton developed Dataset Collections as a more general purpose replacement for Multiple file datasets
- Pulsar is the current evolution of LWR
- Galaxy-P is now just a standard Galaxy server with "a particular set of skills": i.e. installed proteomics tools and workflows

Using galaxy enables integrated genomics and proteomics sample analysis on a common platform



Galaxy Tools for Windows

Example tool for Windows (msconvert_win):

- <u>https://github.com/galaxyproteomics/tools-galaxyp/tree/master/tools/msconvert/msconvert_win</u>
- Python wrapper runs on Windows execing application

^oPulsar Setup and configuration for Windows:

https://pulsar.readthedocs.io/en/latest/readme.html#quickstart

- No automatic installation of tool dependencies
- job_conf.xml configuration for Windows requires bypassing automatic tool_script.sh generation:
 - https://github.com/galaxyproject/galaxy/pull/2492
 - o <plugin id="pulsar_legacy" type="runner" load="galaxy.jobs.runners.pulsar: PulsarLegacyJobRunner" shell="none">
 - o <destination id="destination_blue_pulsar_rest" runner="plugin_pulsar_rest" shell="none">



Toolshed

- Internal tool development toolshed
- MSI specific tools, e.g. copying out galaxy datasets to MSI
- Initial development of tools and dependencies
- Use for tool development education
- A place where one can fail less publicly



Site-specific Configurations

 \bullet \bullet \bullet

MN Supercomputing Institute for Advanced Computational Research



MSI vs Galaxy Accounting

- MSI organizes disk space by PI researcher:

 /home/<group>/ (with group assigned to PI)
 PI charged for disk usage in that path
- Galaxy accounting and quotas by user:
 - Galaxy owns all data in a linux sense
 - can't assign by file path
 - Users uniquely identified by email address
 - Users may work on behalf of multiple PIs



MSI vs Galaxy Accounting

- Users have 2-step login to Galaxy:
 - WebPage1: UMN authentication
 - WebPage2: Select from user's linux groups
 - Galaxy REMOTE_USER: <UMN_ID>+<GROUP>
 - Galaxy User email: <UMN_ID>+<GROUP>@msi. umn.edu
- Accounting tallies all Galaxy disk_usage by group
 - Galaxy API used to set quota for each galaxy user to the used amount of the group allocation

MN Supercomputing Institute for Advanced Computational Research



UMGC Data

- UMN Genomics Center Sequencers:
 - HiSeq 2500
 - HiSeq 2000
 - o MiSeq
- Data Management:
 - Incoming data auto-populated within read-only /home/group/data_release
 - Galaxy Shared Data library created and populated with symlinked datasets, library permissions restricted to PI
 - Allows access from both command line and Galaxy



User Import of Shared Data

- PI moves data files for a Galaxy Shared Data library under: /home/<group>/galaxy/
- PI submits help request to <u>help@msi.umn.edu</u>
- Help Desk chowns files to galaxy, adds data files as symlinked datasets in a Galaxy Shared Data library, and assigns library/folder/dataset permissions to PI and user.
- Galaxy file ownership prevents a command line user from removing a file without notifying a Galaxy admin to remove the Shared Data library dataset



Running Jobs on Mesabi

- Use IP-over-IB network for DRMAA traffic (internal route)
- PBS-DRMAA
 - HPC clusters configured in config/job_conf.xml
 - Most tools are assigned by dynamic_job_runner
 - Assignments read from JSON file
 - Multi-level, tutorial accounts are assigned minimum resources
- Account for SUs (-A group)
 - Galaxy has its own SUs (CPU usage trivial in total HPC)
 - Can assign CPU usage to user if warranted



Clinical Pipelines

 \bullet \bullet \bullet

MN Supercomputing Institute for Advanced Computational Research



Project Summary

- Clinical Laboratory Improvements Amendments (CLIA)
 - Quality laboratory testing (requires locked stack)
 - May include PHI data (this project does NOT)
- In partnership with BMGC, MDL and Fairview Hospital
 - Personalized treatment plan based on patient genome.
- NGSDP -- Next-gen sequencing panel for phenotypes
 - TruSight One (Hiseq) panel detects 4813 genes (~2400 medically relevant)
- Submitted to OTC for path to commercialization



Timeline to Production

- NGSDP v1 (TrueSight One panel; HiSeq)
 - Two phase workflow
 - 567 genes \bigcirc
 - Designed by John Chilton and Jesse Erdmann
- NGSDP v2
 - Expanded to 4813 genes

- ScanIndel • In production
- Whole Exome • Beta testing
- CNV
 - Beta testing



MN Supercomputing Institute for Advanced Computational Research

UNIVERSITY OF MINNESOTA Driven to Discoversm

NGSDP Pipeline (User perspective)

4813 Gene Panel

108 Workflow Steps

- 1. User logs into MSI
- 2. User runs launcher on sample sheet ... Launcher submits job to MOAB queue
- 3. MOAB runs job
 - a. Job orchestrates
 - VM provisioning and data xfer
 - b. Job executes pipeline

MN Supercomputing Institute for Advanced Computational Research



Tier II /

Galaxy Integration

- Forked 2014.06.02 version of Galaxy
- Added:
 - Workflows:
 - Pipeline for 1-4 lanes, and 4+2 libraries
 - Shutdown (*)
 - Tools
 - Data transfer
 - Instance termination
- All other tools come from Galaxy Toolshed
 - All stock except bwa_wrappers



AWS Details

- AMI snapshot of Galaxy instance
 - All tools, workflows, etc. are installed
 - Reference data uploaded and indexed, then put into S3.
 - Users and API Keys configured
- At sample run time:
 - VM runs m2.xlarge (2 vCPU, 17G Mem, 420G storage)
 - Pull reference data from S3
 - Put sample data from MSI
 - Start Galaxy
 - Launch workflow (monitor.py)



Impact of Gene Expansion



• Avg monthly bill: \$1300

From Project Update (5/22/14)

MN Supercomputing Institute for Advanced Computational Research

UNIVERSITY OF MINNESOTA **Driven to Discover**SM

Need for automation

Data Analysis Pipeline Concerns

 Increase Time and Effort: Have to modify 5 files formats before we can launch (vs. 2-3 in the old workflow).

> 3-4 min x 10 samples = 30 – 40 minutes 5 files x 10 samples = 50 potential error points

Resolution: ? Batch Launching.

We had talked about this even with the old workflow but with the increased analysis duration, could be priority.



Low Redo Rate



University of Minnesota Genomics Center (UMGC) Driven to Discover[™]

Automating Workflow Launch

- Monitor.py uses Galaxy WebAPI
 - Launches workflow in new history
 - Email user if job enters error state, continue monitoring
- Regexp match Workflow inputs
 - o "tool_state": "{\"name\": \".*L(0+)?1_R(0+)?1_.*.fastq\"}",
 - Import only the matched inputs into history
- Maintains queue of queued, running jobs left in workflow
- Only launch CLIA_workflow_complete.ga if no errors were encountered.



New Workflow by Component



https://drive.google.com/file/d/0BwqwiTk1X4kkUGFmTkRHNGJ1X3c/view?usp=sharing

MN Supercomputing Institute for Advanced Computational Research



NGSDP Wins

Jobs

- Fully automated workflow
 - Single phase
 - Regular expression inputs
 - MOAB queue integration
- 1 hour job launch
- Total time per sample: 30 hours (~6x faster)
- Cost per sample: O(\$10)



GermLine Jobs Submitted by Month (Avg: 100 / Month)

MN Supercomputing Institute for Advanced Computational Research



Integrating the PBS Scheduler

Integrates into MOAB

- "slot" == "license" (see FlexLM integration example)
- Job resource: -lsoftware=ngsdp___m1.ngsdp+1
 - TenantName__Flavor.Name+SlotsRequested

Reports available (ARES) and total (CRES) tenant slots:

GLOBAL UPDATETIME=1411161328 STATE=idle

ARES=ngsdp__m1.tiny:8, ngsdp__m1.webserver:4, ngsdp__m1.small:8, ngsdp__m1.medium:8, ngsdp__m1.large:4, ngsdp__m1.xlarge:2, ngsdp__m1.venti:0, ngsdp__m1.grande:0, ngsdp__m1.ngsdp:2 CRES=ngsdp__m1.tiny:10, ngsdp__m1.webserver:5, ngsdp__m1.small:10, ngsdp__m1.medium:10, ngsdp__m1.large: 5, ngsdp__m1.xlarge:2, ngsdp__m1.venti:0, ngsdp__m1.grande:1, ngsdp__m1.ngsdp:4

MN Supercomputing Institute for Advanced Computational Research



Evolving Needs: Improving the Tier I Experience

 \bullet \bullet \bullet

MN Supercomputing Institute for Advanced Computational Research





MN Supercomputing Institute for Advanced Computational Research



Age of Files by Group



MN Supercomputing Institute for Advanced Computational Research

UNIVERSITY OF MINNESOTA Driven to Discover

Toward Sustainable Storage

- MSI is shared resource; demand for storage is high
 - Focus on relatively small amount of "hot" data.
 - Discard intermediate/scratch
 - Develop tiers of storage.
- Galaxy must abide by the same restrictions
- Tier II: Spinning disk for short term data (CEPH)
- Archive/Tier III: Tape, yes, Tape. Long term storage



Storage Tiers

High Performance Storage

- 3.1 PB Usable
- 48GB/s read/write
- Available on HPC resources

Tier-2 Storage

- 3.1 PB Usable
- Available via an S3 interface
- Available anywhere in the world
- Archive Storage
 - > 3.5 PB tape-based storage
 - Available via an S3 Interface
 - offline storage



MN Supercomputing Institute for Advanced Computational Research



Tier Two Storage

- 2015 Developed In-house expertise with CEPH (Object File Storage)
- 2016 Deployed 3.1 PB of CEPH Storage with an S3 interface.
 - Deployed node "bricks" of 60 x drive systems with 12 SSDs System
 - Multiple 10 GbE per "brick"
 - Software-based Erasure Encoding 4+2 (RAID-6)
- Supercomputing 2015: MSI-led BoF on CEPH shows the HPC community warming



Ceph Storage at MSI



Date

MN Supercomputing Institute for Advanced Computational Research

Data Size (TB)

UNIVERSITY OF MINNESOTA Driven to Discover

Tier Three Storage

- Piloting program currently.
- Media costs is as low as \$0.60/TB/year (assuming a 14 year tape lifetime).
- Significant investment in the library and drives, but media can be taken off-line and stored elsewhere.
- Blackpearl device from Spectralogic to allow for selfservice use of the tape archive. (Up to 1 GB/sec transfer rates.)



Globus

- MSI spearheading new Gopher Science Net (100 Gbps)
- UMN-wide Globus license
- Need Galaxy-Globus
 Integration
 - Aware of work by Ravi Madduri (2011); have not tested
 - Wishlist: auto-migrate data to/from Tier II and Archive



© 2010–2016 The University of Chicago legal

MN Supercomputing Institute for Advanced Computational Research

UNIVERSITY OF MINNESOTA Driven to Discover

Stratus.msi.umn.edu

- OpenStack Research Computing Cluster
 - 20 compute nodes (400 cores)
 - 5TB Memory
 - 200 TB CEPH RBD
- Scheduled tenant resources Hybrid HPC-Cloud Pipelines
- Cheaper than Amazon



Galaxy Multi-Omics

3yr NSF Grant started 2015

- Enhance the Galaxy environment with new interactive visualization tools and data exchange functionalities necessary for effective multi-omic data analysis
- Extend the Galaxy environment to analyze and process diverse metabolomics data and support workflows for metabolic activity profiling
- Extend the Galaxy environment for integrative genomicproteomic data analysis supporting proteogenomic and metaproteomic applications

MN Supercomputing Institute for Advanced Computational Research



Galaxy Multi-Omics

4yr NIH Grant starting 2016

- Integrated proteomics-genomics analysis supporting proteogenomic workflows for characterizing cancerassociated protein variants and proteotypes
- Integrated metaproteomic analysis environment and workflows for characterizing microbiome-derived cancerassociated proteins
- Integrated metabolomic analysis environment and workflows for characterizing metabolic activity profiles in cancer





 \bullet \bullet \bullet

MN Supercomputing Institute for Advanced Computational Research



Galaxy Storage History



MN Supercomputing Institute for Advanced Computational Research

