

UTSouthwestern

Medical Center

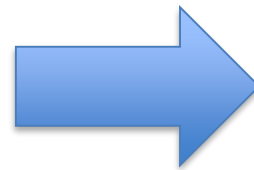
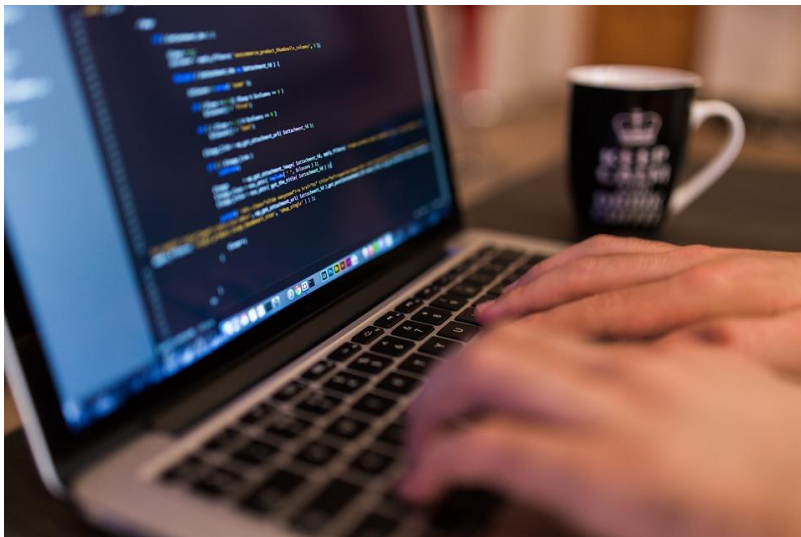
Lyda Hill Department of Bioinformatics

BioHPC

Managing Software Projects in a Team

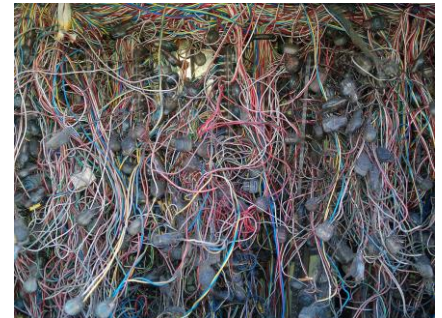
Why think about how to manage software projects?

- Most scientific software starts out written by 1 or 2 people
- Usually written for a specific project
- Emphasis is on getting things done and published



Why think about how to manage software projects?

- Maybe the area of research becomes the focus of the lab
- The software becomes a critical tool that everyone needs
- More features please! Fix this bug! – Complexity increases
- New lab members start to contribute
- The original author moves on to a new institution



Why think about managing software projects?

- Without management of the project, problems happen... and get worse over time



- Planning

Know (roughly) what needs to be done, who will do it.

- Version Control

Track *all* changes, allow multiple developers to work together easily.

- Testing

Ensure bugs are fixed, and developers don't break each others code.

BioHPC's Git service (using GitLab) is a powerful platform to manage team software projects. *Not just version control!*

The left screenshot shows the 'Issues' tab of the BioHPC GitLab interface. It displays a list of issues with checkboxes, search filters, and a 'New Merge Request' button. The issues listed include:

- add gpfs fileset quota: backup archive root (opened 6 days ago by Wei Guo, high priority)
- Live LDAP check fixes (opened 2 months ago by David Trudgian)
- Quota fix merge (opened 3 months ago by David Trudgian)
- Upgrade to Django CMS3.4 (opened 3 months ago by David Trudgian)
- Correction to required fields for UserRegistration admin. Fixes #47 (opened 3 months ago by David Trudgian)
- Nicer password checks, fixes #65 fixes #37 (opened 3 months ago by David Trudgian)
- Update usage contributions for Q1 2017 (opened 3 months ago by David Trudgian)

The right screenshot shows the 'Repository' tab of the BioHPC GitLab interface. It displays a commit graph for the 'master' branch, showing a timeline of commits from June to July. The commits listed include:

- [Live Change] 2017 Q3 accounting
- Merge remote-tracking branch 'origin/live'
- [Live change] Make menu entries shorter so can see during training session
- [Live Change] Add Nikon Elements to DCV image
- Close #70 - confusing start time/date for webGUI
- Merge remote-tracking branch 'origin/live'
- Closes #69 - strip default from module names
- Merge branch 'quota-newfix' into 'master'
- add gpfs fileset quota: backup archive root
- New DCV KVM Image
- LIVE FIX - fix webGUI after param_runner addition
- Merge remote-tracking branch 'origin/master' into live

- Planning - Know (roughly) what needs to be done, who will do it.

GitLab Issues, Milestones, Boards

- Version Control – Track changes, allow multiple developers to work together easily.

GitLab Repositories, Merge Requests

- Testing – Ensure bugs are fixed, and developers don't break each others code.

GitLab CI

Without some planning & tracking of projects, everyone tends to drift in their own area of expertise or enjoyment.

Often a good thing in research, we want to explore new ideas – so heavyweight project management not so applicable as in industry.

.. but ...

Too much mission creep can be bad – nothing gets finished.

Issues – Record, assign, track bugs and new features

The screenshot displays the 'UT Issues' interface. At the top, there's a search bar and navigation icons. Below the header, a summary shows 'Open 14', 'Closed 44', and 'All 58' issues. A 'New Issue' button is visible. The main section contains a list of issues with filters for Author (David Trudgian), Milestone, and Labels. The issues listed are:

- Select all for import files is confusing** (bug) - biohpc/astrocyte#25 - opened 3 months ago by David Trudgian - updated 3 months ago
- CRI Request - Prefix imported files that have the same name** (enhancement, feature) - biohpc/astrocyte#24 - opened 4 months ago by David Trudgian - updated 4 months ago
- identifier for the job** - biohpc/param_runner#1 - opened 6 months ago by David Trudgian - updated 6 months ago
- uploaded file name should not contain special characters** (bug) - biohpc/biohpc_portal#64 - opened 6 months ago by Wei Guo - updated 6 months ago
- BUG - missing docs files at import cause error on workflow page** (bug) - biohpc/astrocyte#22 - opened 8 months ago by David Trudgian - updated 8 months ago

Each project has an issue tracker – users can see issues across all their projects

Record things to do as issues – bugs to fix, new features to add, ideas to explore

Can be assigned to users, labels, milestones and discussed by members of project

Milestones – Group issues into larger goals

New Milestone

Title

1.3


Description

Write

Preview

B **I** **Q** **</>** **☰** **☷** **☑** **✕**

Release version 1.3

Markdown is supported  Attach a file

Start Date

2017-04-14

[Clear start date](#)

Due Date

Select due date

◀ April 2017 ▶

sun	mon	tue	wed	thu	fri	sat
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30						

Create milestone

Cancel

Can be very useful to focus efforts on most important tasks.

People really like to shoot for a good score (% completion of milestone).

Milestones – By topic

[List](#) [Board](#) [Labels](#) [Milestones](#)

[Open 3](#) [Closed 1](#) [All 4](#)

Red Hat 7 Cluster
8 Issues · 0 Merge Requests
Jun 30, 2017–Aug 30, 2017

[Edit](#) [Close Milestone](#) [Delete](#)

UI / usability enhancements
3 Issues · 0 Merge Requests
expires on Dec 31, 2017

[Edit](#) [Close Milestone](#) [Delete](#)

UTSW LDAP Integration
3 Issues · 0 Merge Requests

[Edit](#) [Close Milestone](#) [Delete](#)

Issues grouped by a common theme of work

Good for projects that don't have a highly structured release cycle

Milestones – By version

List Board Labels **Milestones**

Open 4 Closed 1 All 5

Future
0 Issues · 0 Merge Requests

Security - Sensitive data analysis
2 Issues · 0 Merge Requests

v1.0.0
1 Issue · 0 Merge Requests

v1.1.0
9 Issues · 0 Merge Requests

Edit Close Milestone Delete

Edit Close Milestone Delete

Edit Close Milestone Delete

Edit Close Milestone Delete

Issues grouped by a target release version
Good for projects that have a structured release cycle

More speculative work can be kept in non-versioned milestones

Boards – An overview of work ongoing / pending

The screenshot displays a Kanban board interface with the following components:

- Navigation Bar:** Project, Repository, **Issues 24**, Merge Requests 0, Wiki, Settings.
- Board Tabs:** List, **Board**, Labels, Milestones.
- Filters:** Author, Assignee, Milestone, Labels, Filter by name..., Add list, Add issues.
- Columns:**
 - bug (5 items):**
 - #64: uploaded file name should not contain special characters
 - #60: web job resubmitted by refreshing success page
 - #58: Remove unsuccessful web desktop start attempts (medium)
 - #45: Non-ascii characters in registration form trigger server error (LDAP)
 - #43: Consider impact of disabled users on storage accounting - update disabled user quota info? (medium, quota)
 - enhancement (2 items):**
 - #40: Make biohpc_utils admin view nice (quota)
 - #17: Web job submission should get partition list dynamically (RHEL7)
 - RHEL7 (8 items):**
 - #71: Build RHEL7 test VM (Infrastructure)
 - #72: Build RHEL7 production VM (Infrastructure)
 - #73: Switch vagrant dev environment to CentOS 7. (vagrant)
 - #74: Quota testing on RHEL7 (testing)
 - #75: Migrate vagrant to SLURM machine RHEL7 cluster (vagrant)
 - #76: Upgrade to guacamole / tomcat 7 (Infrastructure, ssh, vagrant)
 - #77: Update yum for SLURM 17.x
 - Done (15 items):**
 - #66: Date Time picker is confusing (Avatar)
 - #59: sorted department and groups for quota accounting page
 - #57: web job submission result page error.
 - #48: Add ability to import quota information from other systems via REST JSON calls. (Avatar, feature, high priority)
 - #68: webgui script and web visualization session script can only be used with user
 - #37: Password re-use / history considerations (security)

*Snapshot of work going on, or waiting to be done.
Can be used to visually arrange and prioritize work via labels.*

In the Git on BioHPC training we talked about how to use git

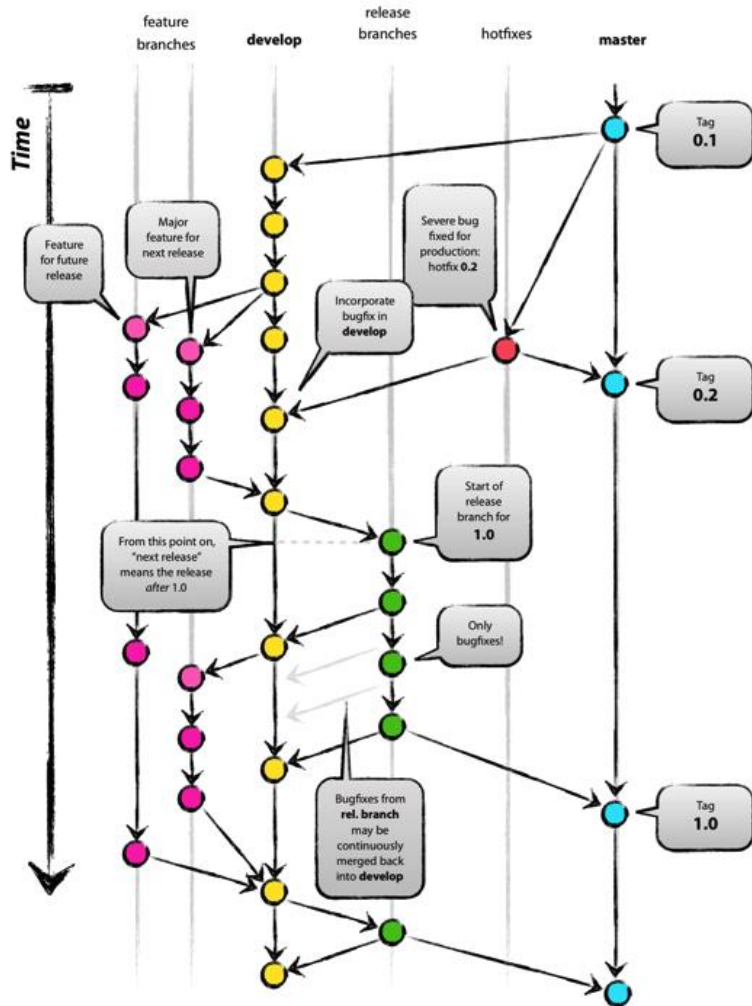
Push, pull, branch merge

Git is very flexible – *too flexible!*

Everyone using git in different ways on the same project = big mess!

We need some guidance to establish a git workflow for a project

Git Flow - Too complex for most projects

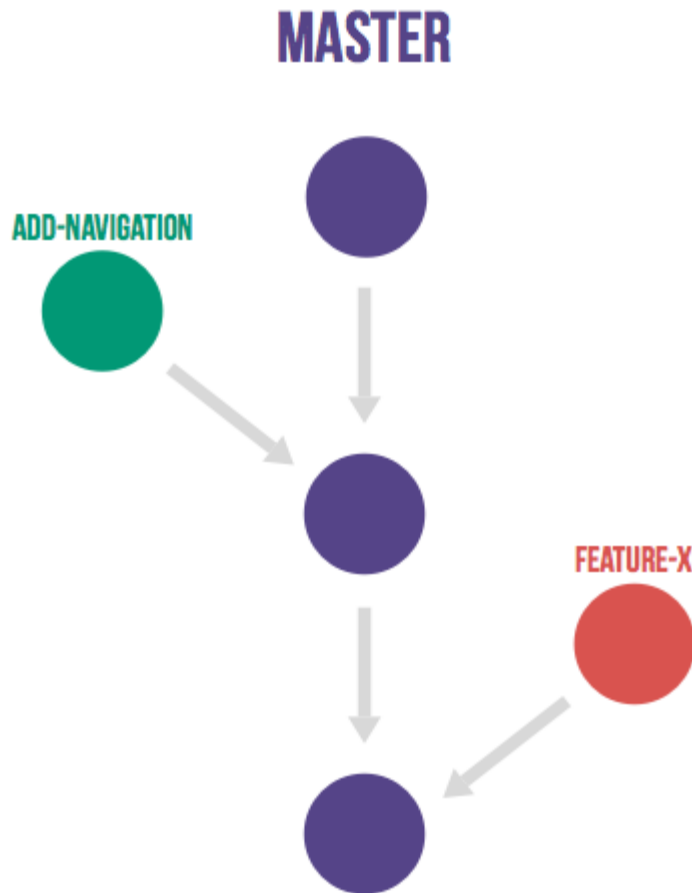


Often seen in web documentation

Designed around needs of industry developing and maintaining versions of software

Overkill for most scientific projects

GitLab Flow – A nice compromise



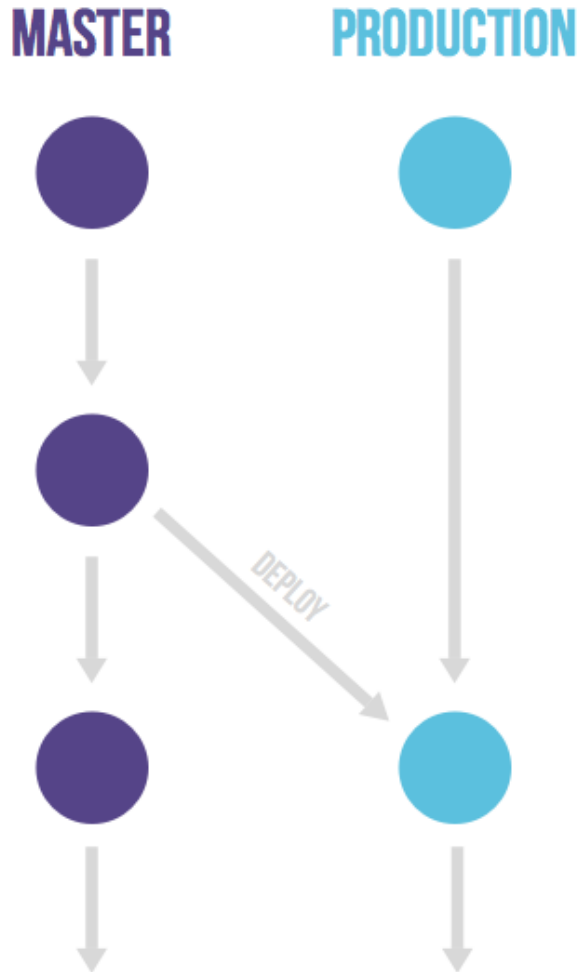
Simple to explain

Any new piece of work (bug-fix, feature) is a new feature branch

Feature branches are merged into master branch

Master branch is the current state of the project

Keeping a Production/Live Branch



Variant – master + production

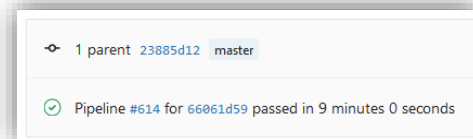
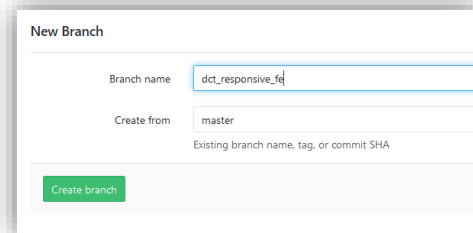
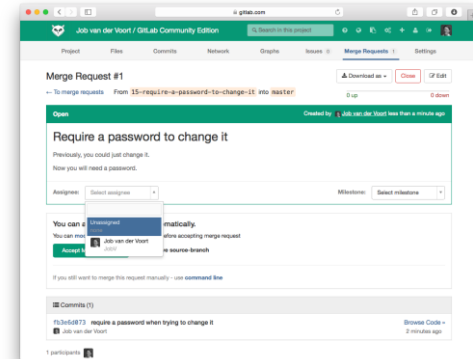
Lots of scientific work needs a stable place for analysis projects, and an exploratory place for new developments.

Work merged into master branch

Add a production branch for stability

Feature branches are where work gets done

- Create or pick issue(s) to work on
- Create a new branch from master
- Code and test



1. Use feature branches, no direct commits on master.
4. Perform testing, code reviews before merges into master, not afterwards.
9. Everyone starts from master, and targets master.



<https://about.gitlab.com/2016/07/27/the-11-rules-of-gitlab-flow/>

Managing Merges

In small projects, everyone can merge directly into master, without much trouble.

In large projects things need to be reviewed first...

- Not everyone is aware of how their changes could affect others

- May be some coding rules to enforce

Merge requests propose merging changes, allowing them to be reviewed and accepted/rejected/modified

Informal – anyone can accept a request

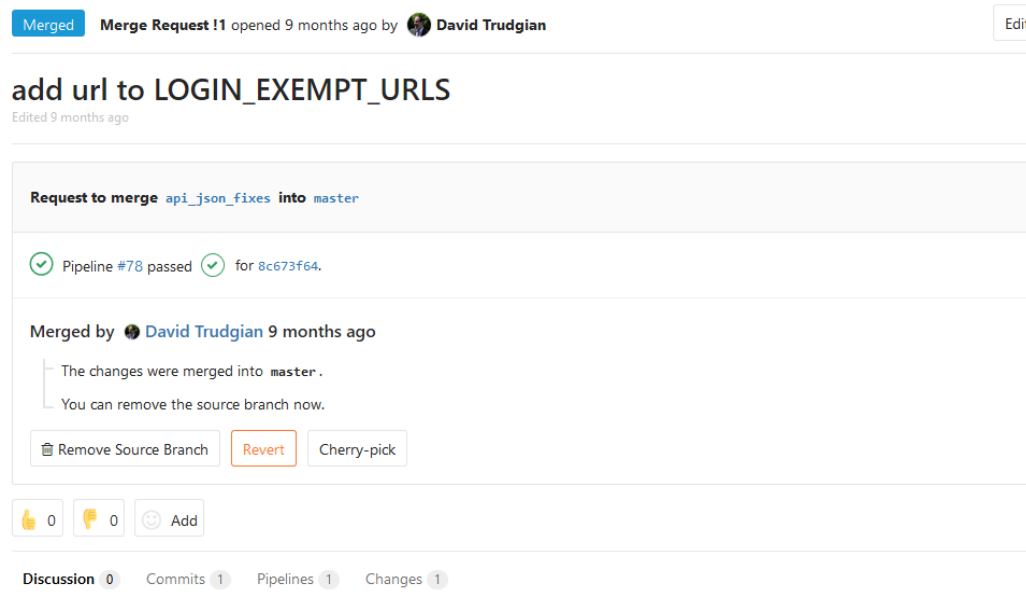
Formal – project gatekeeper has to OK



Merge Requests

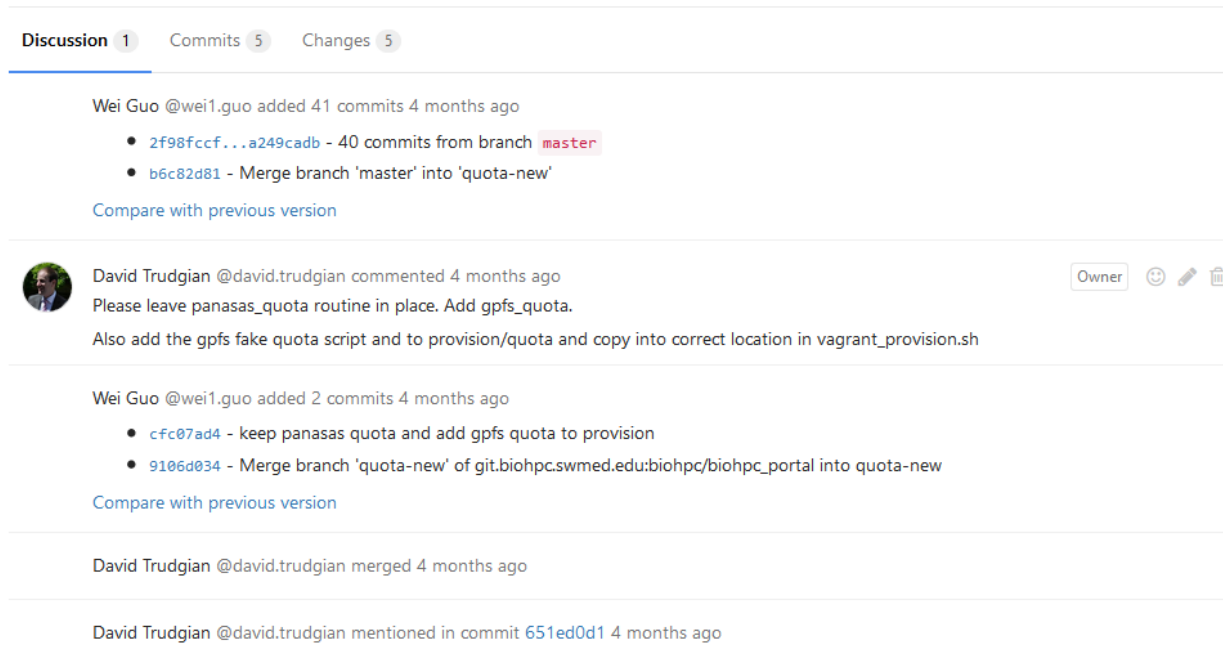
Merge changes into the master branch using merge requests

Merge requests allow easy review and comment before changes that might affect others become part of the main code branch



Make programming social – use mentions, ask advice, learn by reviewing

You can look at changes in a merge request and comment on the request.



The screenshot displays a GitHub merge request interface. At the top, there are tabs for 'Discussion' (1), 'Commits' (5), and 'Changes' (5). The 'Discussion' tab is selected. Below the tabs, a user named Wei Guo (@wei1.guo) is noted as having added 41 commits 4 months ago. A list of commits is shown: '2f98fccf...a249cadb' (40 commits from branch 'master') and 'b6c82d81' (Merge branch 'master' into 'quota-new'). A link to 'Compare with previous version' is provided. Below this, a comment by David Trudgian (@david.trudgian) is shown, dated 4 months ago. The comment text is: 'Please leave panasas_quota routine in place. Add gpfs_quota. Also add the gpfs fake quota script and to provision/quota and copy into correct location in vagrant_provision.sh'. To the right of the comment are icons for 'Owner', a smiley face, a pencil, and a trash can. Below the comment, another entry shows Wei Guo adding 2 commits 4 months ago, with a list of commits: 'cfc07ad4' (keep panasas quota and add gpfs quota to provision) and '9106d034' (Merge branch 'quota-new' of git.biohpc.swmed.edu:biohpc/biohpc_portal into quota-new). A 'Compare with previous version' link is also present. At the bottom, two more entries are visible: 'David Trudgian merged 4 months ago' and 'David Trudgian mentioned in commit 651ed0d1 4 months ago'.

Discussion 1 Commits 5 Changes 5

Wei Guo @wei1.guo added 41 commits 4 months ago

- 2f98fccf...a249cadb - 40 commits from branch master
- b6c82d81 - Merge branch 'master' into 'quota-new'

[Compare with previous version](#)

David Trudgian @david.trudgian commented 4 months ago Owner 😊 ✎ 🗑

Please leave panasas_quota routine in place. Add gpfs_quota.

Also add the gpfs fake quota script and to provision/quota and copy into correct location in vagrant_provision.sh

Wei Guo @wei1.guo added 2 commits 4 months ago

- cfc07ad4 - keep panasas quota and add gpfs quota to provision
- 9106d034 - Merge branch 'quota-new' of git.biohpc.swmed.edu:biohpc/biohpc_portal into quota-new

[Compare with previous version](#)

David Trudgian @david.trudgian merged 4 months ago

David Trudgian @david.trudgian mentioned in commit 651ed0d1 4 months ago

The team works on the request together, to ensure it's successful.

Code Review

View all the code changes – comment on individual lines, files, sections of code

52

lib/galaxy/auth/__init__.py

☒ Show comments

View

▼

@@ -20,19 +20,24 @@

20

20

21

21

<auth>

22

22

<authenticator>

23

-#

<type>activedirectory</type>

24

-#

<filter>'[username]'.endswith('@students.latrobe.edu.au')</filter>

23

+#

<type>ldap_ad</type>

nsoranzo on Apr 7, 2015

Member

Presently the type in this example should be `ldap`.

dannon on Apr 7, 2015

Owner

Nice catch!

More eyes = more chances to catch bugs before they are merged

































Automated Testing / CI Pipelines

Testing is tedious – *unless you automate it*

Testing is required to keep projects working as they become more complex

There are tools to write tests in every popular language

GitLab CI (Pipelines) can run tests on every code change for you

	#385 by 	 live  717c2b1f SLURM limits for executor by default		 00:06:32  4 months ago
	#384 by 	 master  d897d355 Merge branch 'master' of https://gi...		 00:06:23  4 months ago
	#79 by 	 master  0a38132a Merge branch 'api_json_fixes' into '...		 9 months ago
	#78 by  latest	 api_json_fixes  8c673f64 add url to LOGIN_EXEMPT_URLS		 9 months ago
	#71 by 	 live  5d6424b0 Merge remote branch 'origin/mast...		 9 months ago

 .gitlab-ci.yml 136 Bytes

```
1  before_script:
2    - sh provision/setup_venv.sh
3
4  test:
5    script:
6      - source .venv/bin/activate
7      - module add nextflow/0.20.1
8      - py.test
```

A simple YAML file that defines steps in a testing/CI pipeline
Each step can take defined actions
GitLab runs for all new commits, and tracks results

The CI tasks are run by a program called 'gitlab-runner'

This is available as a module. You can run it on a workstation / thin-client, or in an interactive session or batch job.

```
module add gitlab runner

# Register your project to the runner using the CI token in
# the project settings
gitlab-runner register

# Start the runner, so it can retrieve and run builds
gitlab-runner run
```

[Project](#) [Repository](#) [Issues](#) 12 [Merge Requests](#) 0 [Pipelines](#) [Wiki](#) **[Settings](#)**

[General](#) [Members](#) [Integrations](#) [Repository](#) **[Pipelines](#)** [Pages](#)

Specific Runners

How to setup a specific Runner for a new project

1. Install a Runner compatible with GitLab CI (checkout the [GitLab Runner section](#) for information on how to install it).
2. Specify the following URL during the Runner setup: `https://git.biohpc.swmed.edu/`
3. Use the following registration token during setup: `oAzYySyUjHjhLcHG7hwQ`
4. Start the Runner!

Runners activated for this project

● 5d0d8fe2 

Disable for this project

biohpcws064.biohpc.swmed.edu

#1

We've covered some modern practices for managing software projects.

How can BioHPC help your group with your software development challenges?

- Do you have large projects, that are becoming difficult to control?
- Do you have difficulty managing git version control in your group?
- Do you want to use automated testing / CI?
- Is there anything else that BioHPC could provide that would be useful?

Chat with us, or contact biohpc-help@utsouthwestern.edu for an appointment.