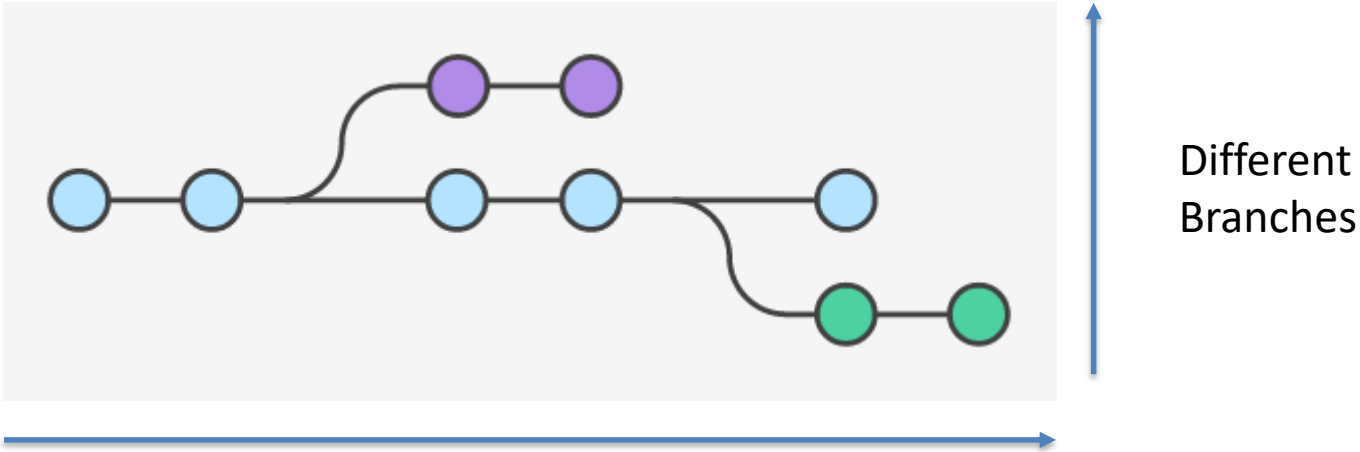
UT Southwestern Medical Center | BioHPC

Git Version Control

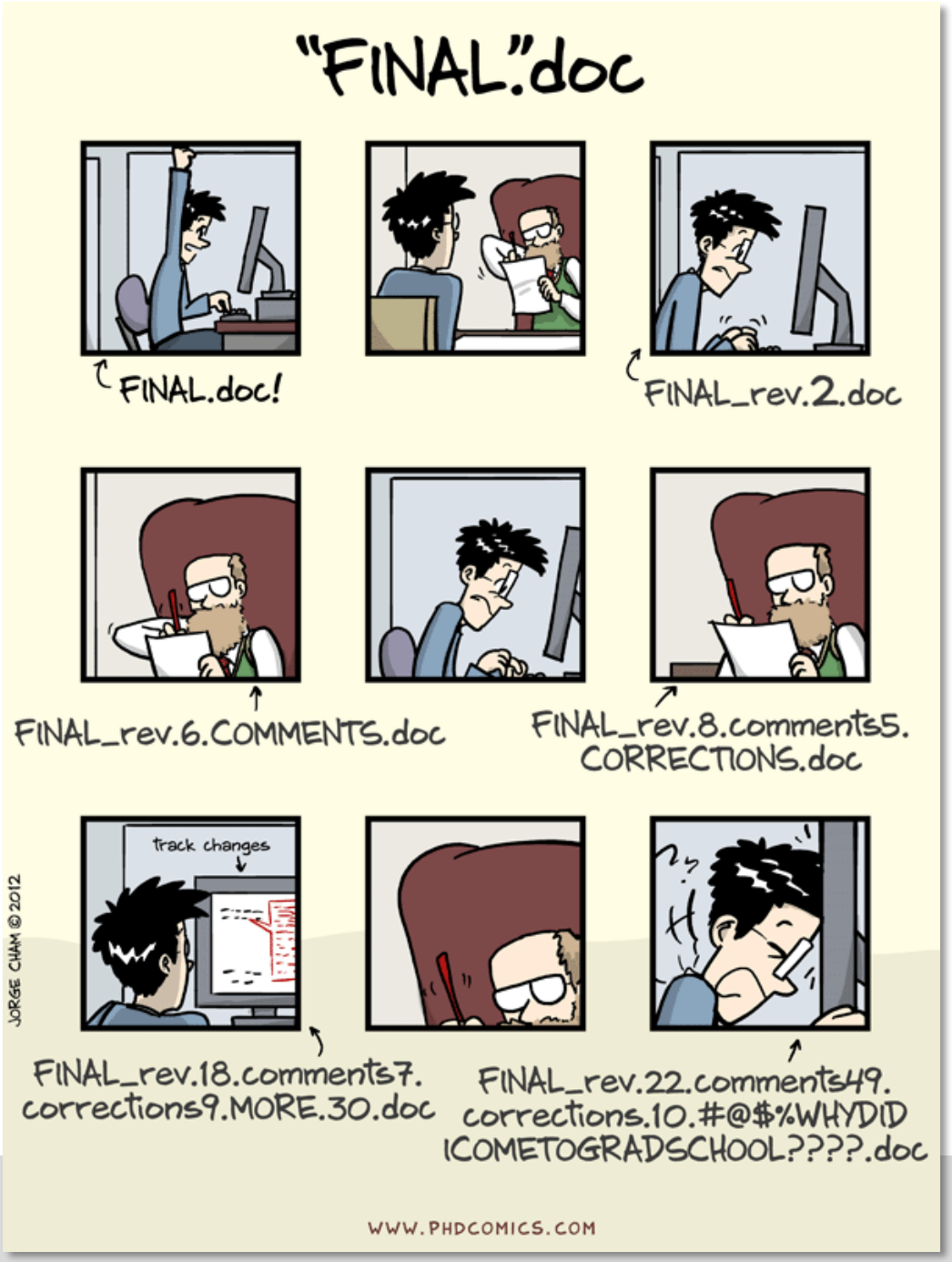
Updated for 2017-04-12

What is Version Control?

Version control is a system that records changes to a file or set of files over time so that you can recall specific versions later.
It is all about managing the *history* of and multiple versions (*branches*) of source code, documents, web sites, etc.



Acknowledgment – Diagrams taken from Atlassian Git Tutorials (https://www.atlassian.com/git/tutorials/)



Bad enough for a single manuscript!

What if you have 100s of files of code?

"Piled Higher and Deeper" by Jorge Cham
www.phdcomics.com

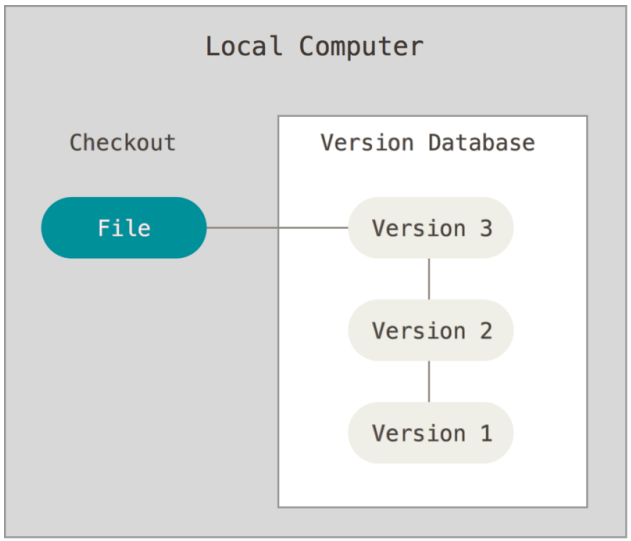
Why Use Version Control?

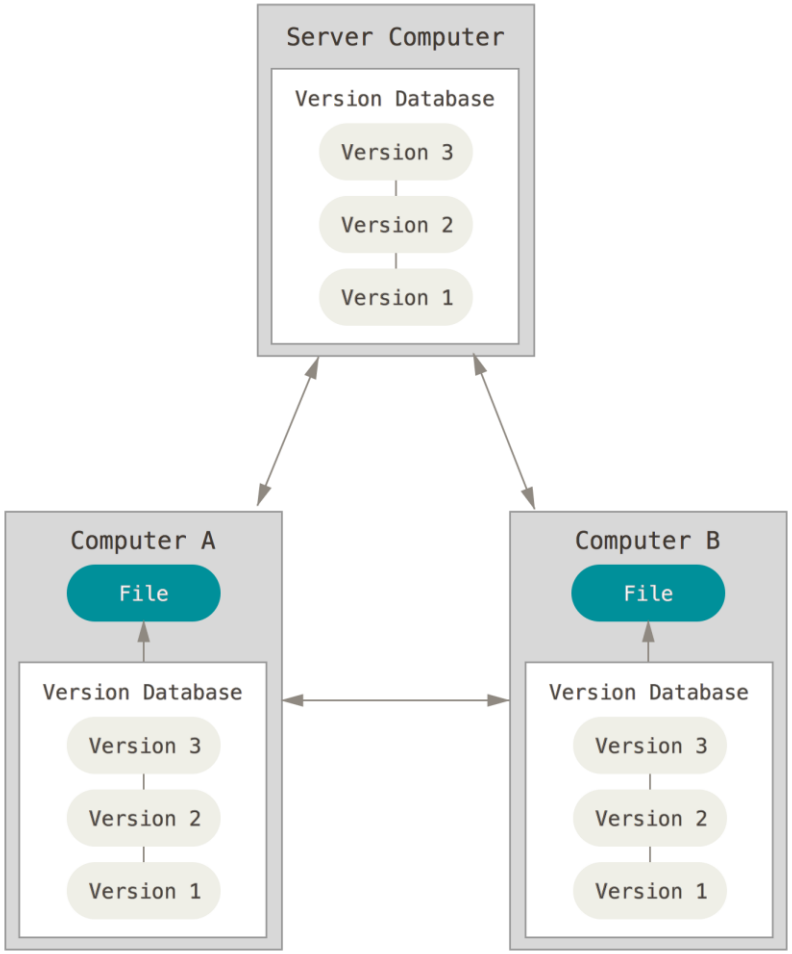
Working by yourself

- Gives you a “time machine” for going back to earlier versions

Working with others

- Greatly simplifies concurrent work, merging changes

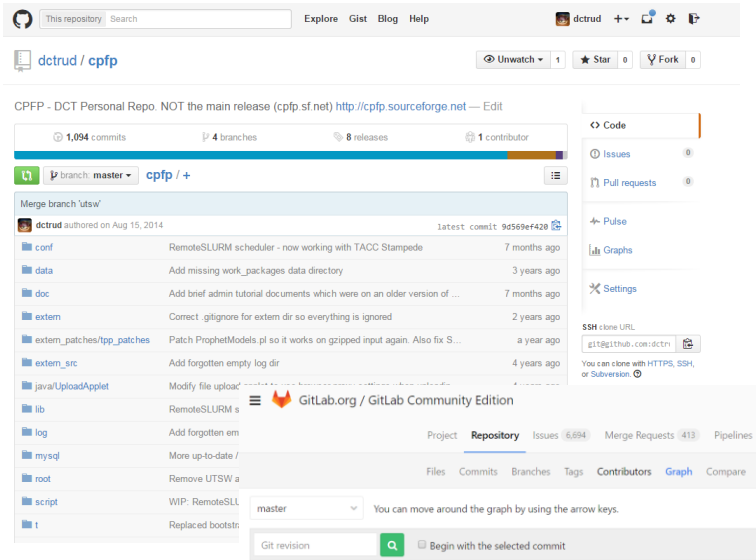


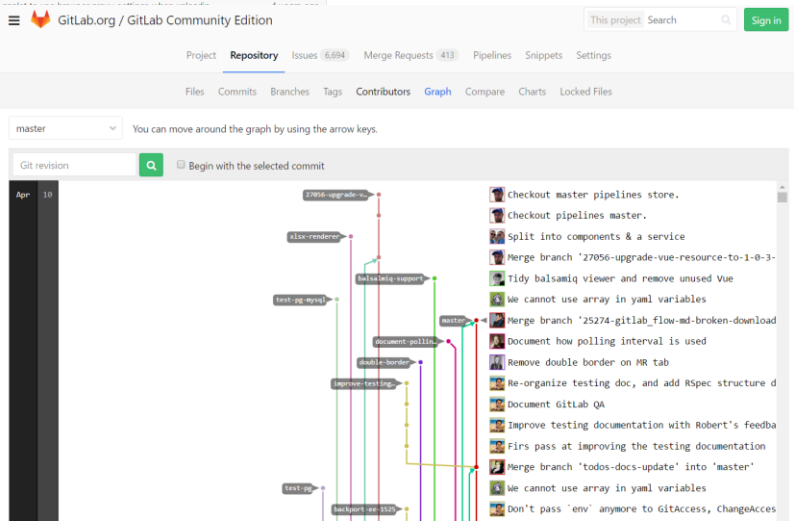


Why Git?

Most used tool by far!

GitHub

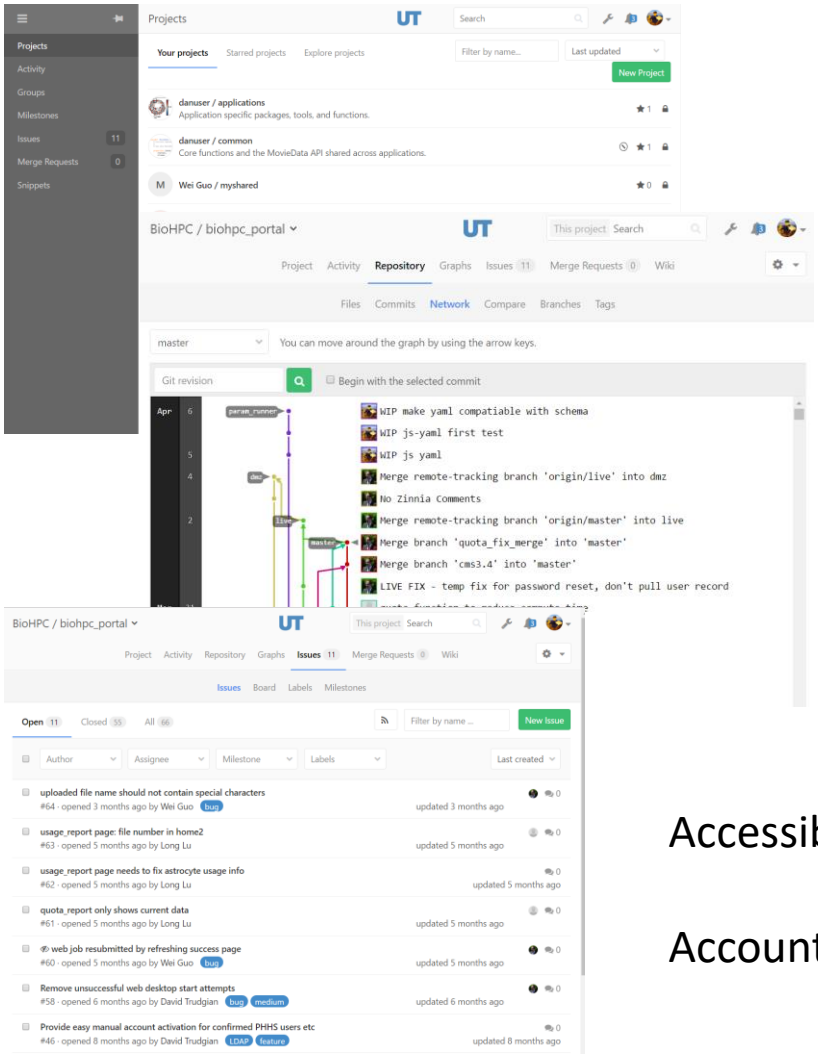




GitLab

We have https://git.biohpc.swmed.edu – based on GitLab, it's a lot like GitHub

https://git.biohpc.swmed.edu



Browse files and history

Manage access rights

Create branches / forks

Perform *basic* editing online

Track issues/bugs, merge requests

Accessible from internet

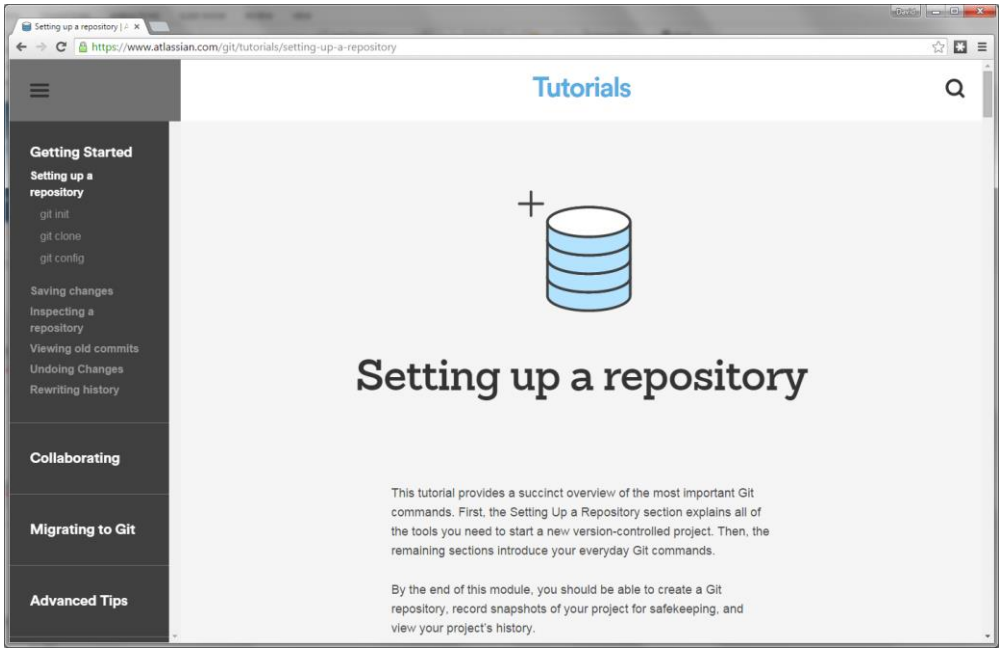
Accounts for non-UTSW collaborators available

A Comprehensive Tutorial

We have time to run through the most common features you might use.

An *excellent* tutorial on almost everything about git can be found at:

https://www.atlassian.com/git/tutorials

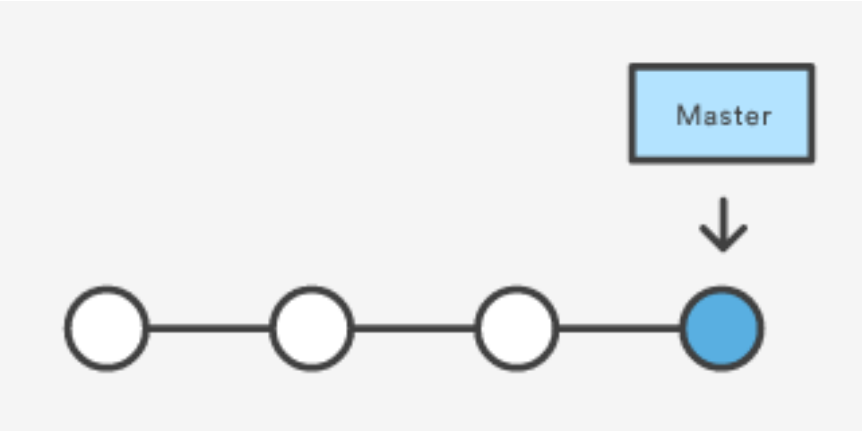


An Example Scenario….

John has some code. It's very important so he wants to keep track of his work.

John creates a git *repository* containing his code, *local* to his machine.

The repository stores sets of changes John makes, called *commits.*



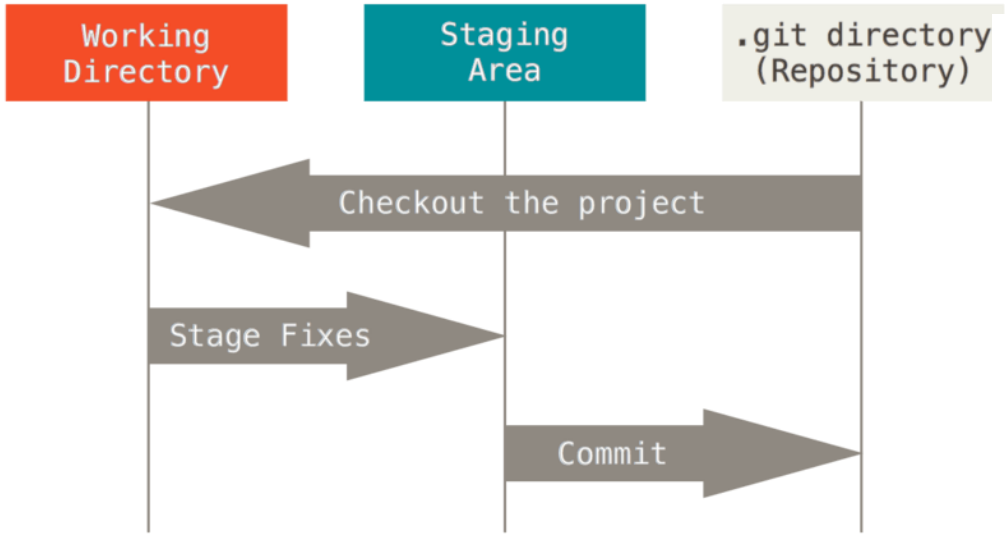
There is a single *master branch*. The newest commit is the *HEAD* of the branch.

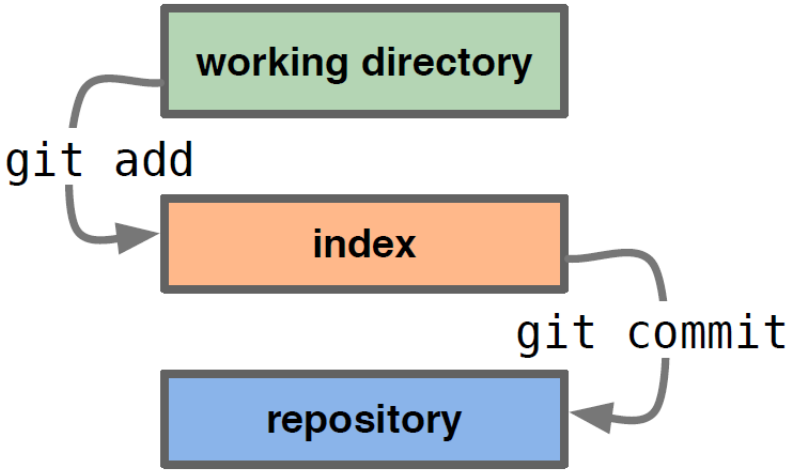
The Three States

Basic git workflow:

- Modify files at the working directory
- Stage files by adding snapshots to staging area
- Commit the changes and update snapshots in the git directory

The index is where we can *stage* changes that will become a *commit*



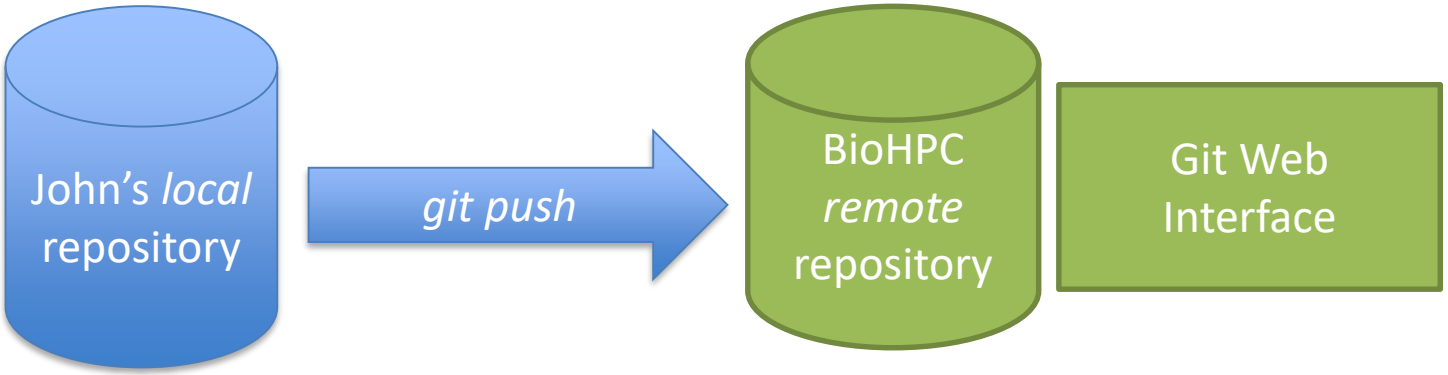


An Example Scenario….

John wants to make his code available to everyone else in his lab, to view or use.

John creates a *project* on the BioHPC git service.

John *pushes* his *repository* to the *remote* BioHPC git service.



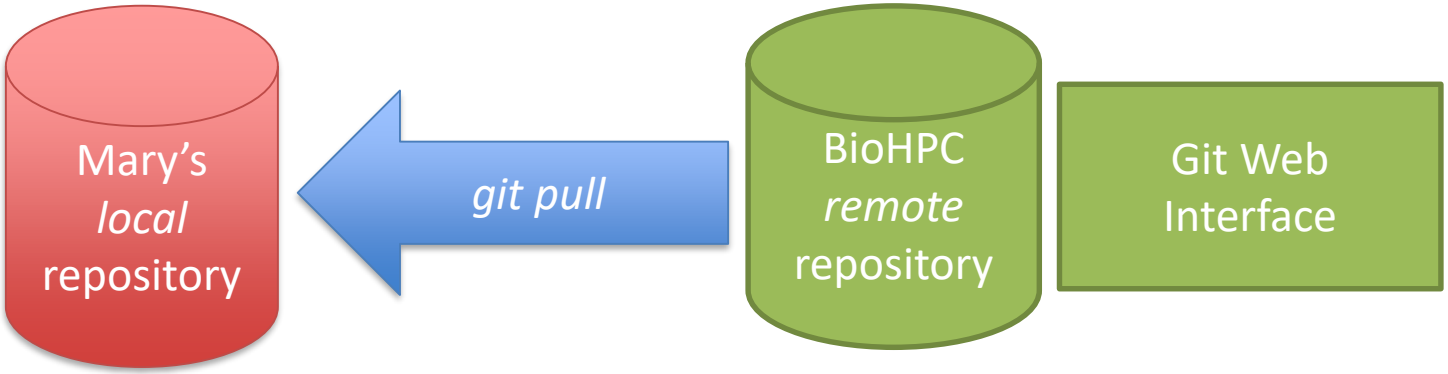
Now everyone can see John's code as he works on it, using the web interface.

An Example Scenario….

Mary wants to use John's code.

Mary *clones* John's code from the central BioHPC remote repository, to a *local* repository on her machine

Mary can *pull* any change John makes, into her local repository.

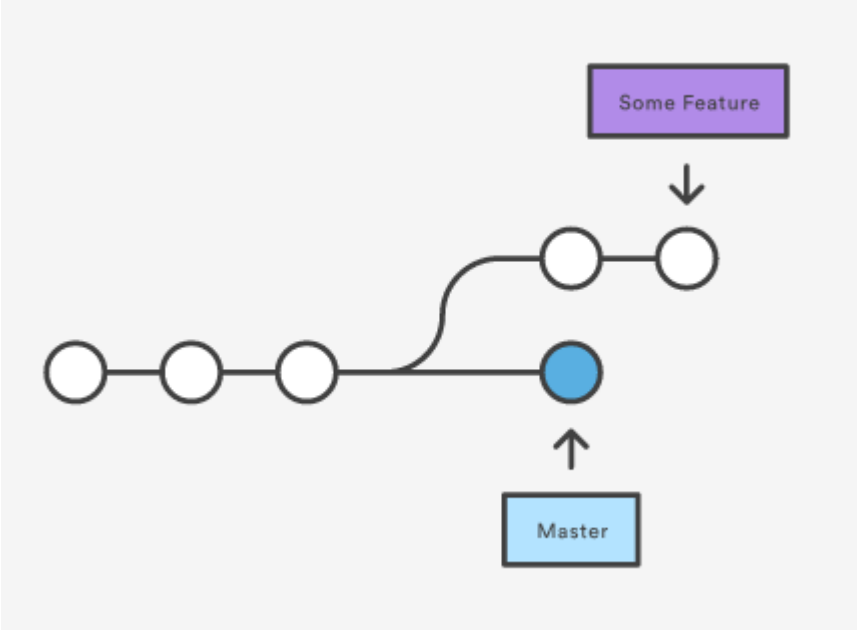


An Example Scenario….

Mary wants to add her new algorithm to John's code.

Mary creates a *branch* in her local repository and modifies the code in that branch to use her new algorithm

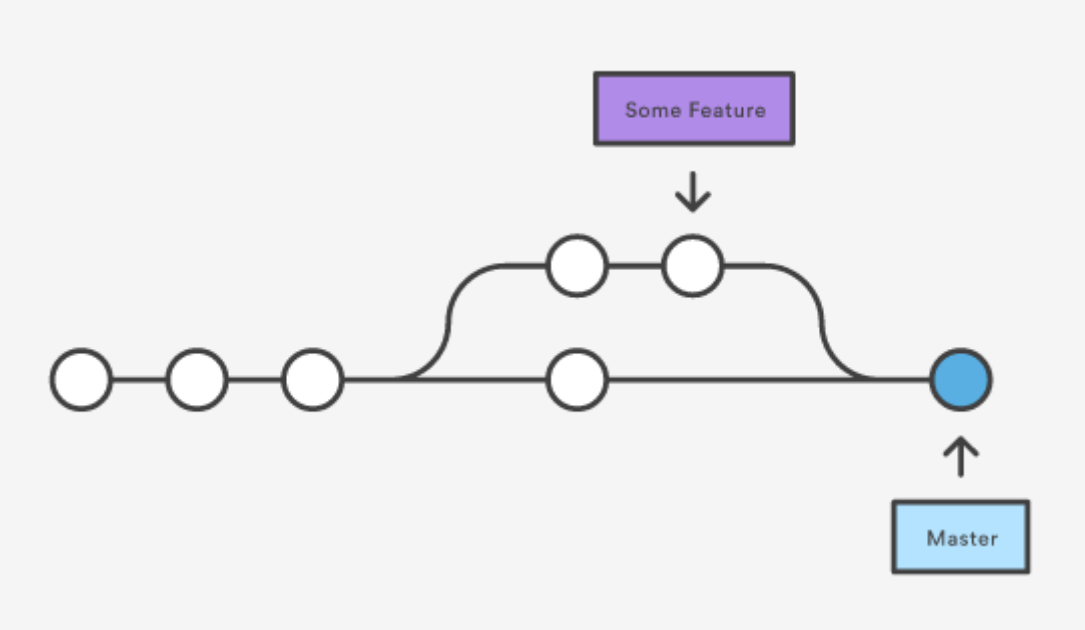
Mary can *push* the branch to the BioHPC remote repository so that others can use it.



An Example Scenario….

Mary's algorithm is great, so John wants to bring her changes into his *master* branch.

John *fetches* Mary's *feature branch* from the *remote,* and *merges* her *commits* into his *master* branch. If they have both changed the same portions of code John must *resolve* any *conflicts.*



Check Version

Before you begin, first make sure Git is installed and which version is installed.
On BioHPC cluster, use command:

```
$ module add git
$ git --version
```

Git 1.7.9 and earlier may experience problems with the clone command and https – so use ssh!

Download and Install

Windows: http://git-scm.com/download/win

Mac: http://git-scm.com/download/mac

Linux: yum install git (CentOS, RHEL) or apt-get install git (Ubuntu)

Setting Things Up – Name and Email

Git assigns all commits to a user name and email address.

You have to tell git your name and email address, once on each system you use:

Only once across all BioHPC workstations/clients & Nucleus

```
$ git config --global user.name "John Doe"
$ git config --global user.email johndoe@example.com
```

If you don't do this your history will be very messy!

Setting Things Up - SSH Keys

Git connects to remote repositories using either https or ssh

ssh works more reliably, but you need to create an ssh key-pair, and let the BioHPC git service know your public key.

On BioHPC you already have a public key at ~/.ssh/id_dsa.pub

On Mac use ssh-keygen –t rsa in a terminal window

On Windows use ssh-keygen –t rsa in a 'Git Bash' session

Setting Things Up - SSH Keys

Now copy the contents of the *public key* id_dsa.pub / id_rsa.pub into your git.biohpc.swmed.edu user profile.

- Login to git.biohpc.swmed.edu
- Go to 'profile settings' (person icon in top right)
- Go to 'ssh keys'
- Add your ssh key – paste content of file in the 'Key' box, and give it a meaningful title.

Check to see if your SSH key is working properly:

```
$ ssh -T git@git.biohpc.swmed.edu
Welcome to GitLab, Wei Guo!
```

Demo

Demo of:

Creating a project on the BioHPC git service

Creating a repository

Making changes, committing, pushing

Branching

Merging, conflicts

See the accompanying notes for reference material.

Git Advice

- Not good for storing huge binary files, data
  e.g. keep large test images out of your git repository
  `git help gitignore` or search web for gitignore

- Decide on a structure for your project earlier recommended that master branch is stable, tested code
  Pull before push

- Create branches for ongoing work, commit & push things often surprising how often a complete history is useful.

Git Advice

There are *many* ways of accomplishing the same thing in git

The same operation can often be performed using shortcuts:

```
git branch feature
git checkout feature
```

is the same as

```
git checkout -b feature
```

In web guides, many things are equivalent but written in a different way

Resources

Notes Specific to the BioHPC git server

https://portal.biohpc.swmed.edu/content/guides/using-biohpc-git/

Tutorials

https://www.atlassian.com/git/tutorials/

Videos

https://www.youtube.com/watch?v=r63f51ce84A