Code Management and Making Contributions to the UFS-SRW Using git and GitHub

The Nuts and Bolts of Research to Operations

Michael J. Kavulich, Jr.,
Associate Scientist, NCAR JNT/RAL

September 23, 2021

This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

Outline

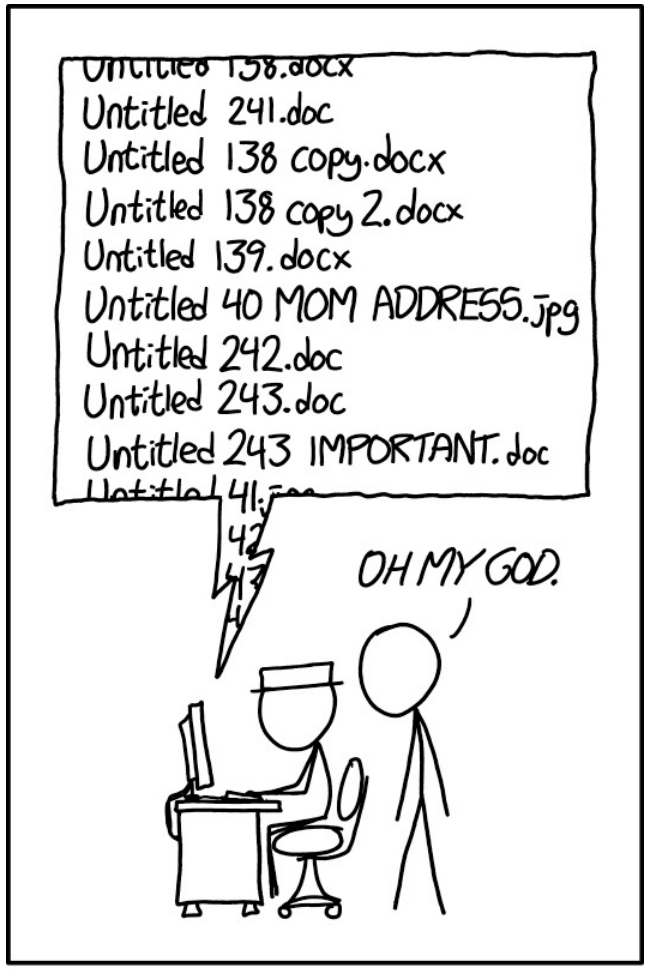
- Version control: Git and GitHub
  - Version control overview
  - Using git software on the command line
  - Using GitHub
- UFS-SRW structure, Submodules, and Manage Externals
- Making and contributing changes to the UFS code
  - Changing code for development purposes
  - Contributing code back to individual repositories
    - Testing requirements
    - Pull requests

Outline

- Version control: Git and GitHub
  - Version control overview
  - Using git software on the command line
  - Using GitHub
- UFS-SRW structure, Submodules, and Manage Externals
- Making and contributing changes to the UFS code
  - Changing code for development purposes
  - Contributing code back to individual repositories
    - Testing requirements
    - Pull requests

Why do we need version control?

- In the olden days, code development, whether for an individual or in a team setting, was a slow, divergent process with lots of potential for problems
  - Keeping track of the “official” version of the code relied on outside communication and/or naming conventions
  - Difficulty remembering when changes were made and by whom
  - Working on multiple changes simultaneously could result in frustrating conflicts and overlapping changes
  - Figuring out when and how a bug was introduced could be near impossible
- It was decided a better system was needed: version control software was developed to enable users to
  - Keep the authoritative version of the code in a central location
  - Track changes made to the code
  - Allow multiple individuals or groups to make changes to the code independently
  - Recognize and resolve when conflicting changes are made to the code



A primitive version control system

Basics of version control

- Version control software simply tracks the history of changes to files; in other words, it keeps track of different “versions” of a file or files as they are modified over time
- Three different types:
  - Linear version control
    - Simple but ubiquitous example: Microsoft Word/Google Docs
  - Centralized version control
    - Subversion
  - Distributed version control
    - git
- In this context, what we are tracking is simply plain text files
  - source code
  - run scripts
  - documentation

git version control software

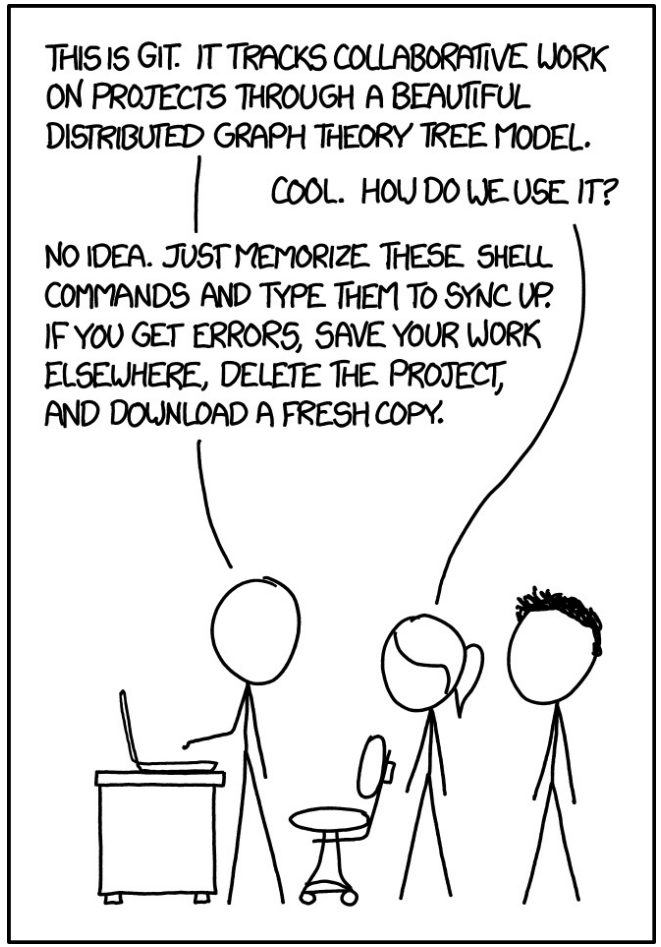
- git version control was developed for the Linux project
- Decentralized: rather than a single central copy of the code "repository" where all changes must be handled, *everyone* has an equally valid copy of the entire repository
  - This sounds complicated (and it can be), but the standard workflow ("gitflow*") used by most UFS components keeps everything organized
  - Among the many advantages to this system are
    - Simple to track local development even for minor changes
    - Internet access is not needed for active development until it is time to move the changes elsewhere
    - Inadvertent changes can usually be undone easily
- git has become by far the dominant version control system in the software community; in 2018 almost 90% of surveyed software developers preferred it as their version control software.

*my software engineer consultant/wife tells me that it is more accurately described as "gitflow branching model with forking"

How git works

- A self-contained bunch of tracked code is known as a *repository*
- git repositories can be created from scratch, but we'll focus on existing code
- Those of you in the practical sessions have already used git at least once, when you first retrieved the UFS-SRW App repository:

```
git clone -b ufs-v1.0.1 https://github.com/ufs-community/ufs-srweather-app.git
```



If all else fails…

How git works

- A self-contained bunch of tracked code is known as a *repository*
- git repositories can be created from scratch, but we'll focus on existing code
- Those of you in the practical sessions have already used git at least once, when you first retrieved the UFS-SRW App repository:

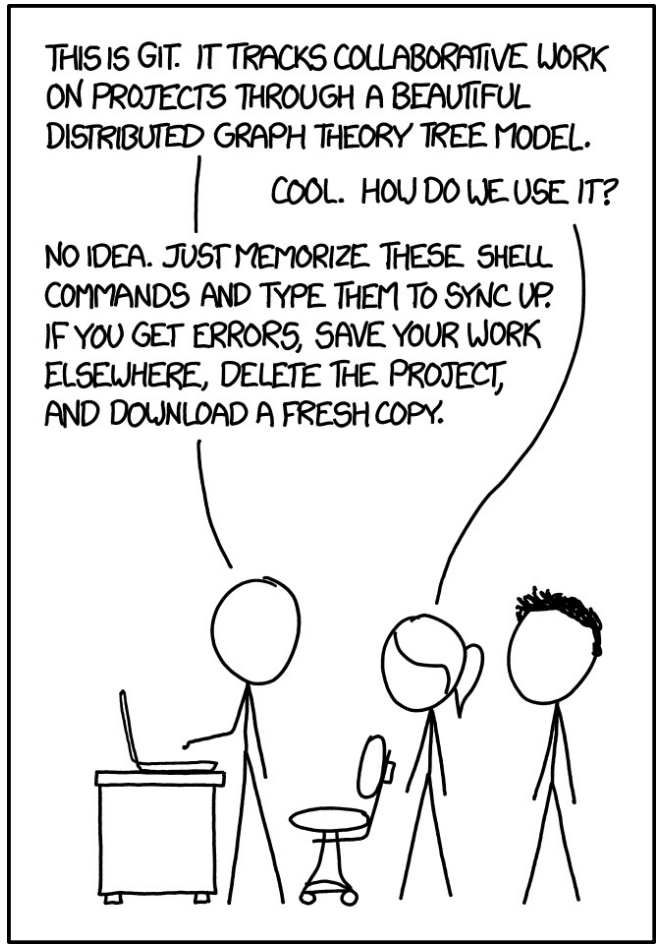
creates a *clone*, or a local copy of an existing repository

Location of the Short Range Weather App repository on GitHub

```
git clone -b ufs-v1.0.1 https://github.com/ufs-community/ufs-srweather-app.git
```

check out the branch named 'ufs-v1.0.1'

- You can try any of the commands in this presentation at home if you like, since `git clone` gives you a full copy of the repository to do whatever you want with it!



If all else fails…

Making commits with git

- A git “save point” is known as a *commit*
  - Each commit contains
    - A change to the code being tracked by git
    - A commit “message” (provided by the person who made the change)
    - A SHA-1 hash that uniquely identifies that commit, *as well as all commits that came before it*
  - The `git log` command gives a list of all commits* since the repository was created

```
commit 704d1abda3322d6eb8691bd7b9ccb39207d00778
Author: Michael Kavulich <kavulich@ucar.edu>
Date:   Thu Sep 9 12:52:40 2021 -0600

    Update regional_workflow hash (#175)

    ## DESCRIPTION OF CHANGES:
    Two problems have been addressed in the regional_workflow repository that need to be added to the app; a simple
hash update accomplishes this.
     - https://github.com/NOAA-EMC/regional_workflow/pull/590 Fixes a problem related to the EMC_post --> UPP rename
     - https://github.com/NOAA-EMC/regional_workflow/pull/592 Fixes a problem with the ./run_WE2E_tests.sh script th
at was introduced in https://github.com/NOAA-EMC/regional_workflow/pull/578
     -
    ## TESTS CONDUCTED:
    Ran set of end-to-end tests on Hera and Cheyenne; all passed

commit e55c497ebd8d20da359c0232c5dfe642165790a3
Author: Chan-Hoo.Jeon-NOAA <60152248+chan-hoo@users.noreply.github.com>
Date:   Tue Sep 7 11:18:55 2021 -0400

    Change EMC_post to UPP (#171)

commit 7c0c8c3b5fe55948d40b5bfb8c366e5c8f0afcb7
Author: Christina Holt <56881914+christinaholtNOAA@users.noreply.github.com>
```

*Technically, only the commits *in a given branch*, but we’ll explain what that means later

Making commits with git

- A git “save point” is known as a *commit*
  - You can create a commit by modifying code, “staging” that code for commit, and then committing
    - The `git status` command will show files that are different from what git has in its ledger; in this example, two files have been modified and one new one created

```
test/ufs-srweather-app> git status
On branch release/public-v1
Your branch is up-to-date with 'origin/release/public-v1'.
Changes not staged for commit:
  (use "git add <file>..." to update what will be committed)
  (use "git checkout -- <file>..." to discard changes in working directory)

        modified:   README.md
        modified:   env/build_cheyenne_gnu.env

Untracked files:
  (use "git add <file>..." to include in what will be committed)

        Newfile.txt

no changes added to commit (use "git add" and/or "git commit -a")
```

Making commits with git

- A git “save point” is known as a *commit*
  - You can create a commit by modifying code, ”staging” that code for commit, and then committing
    - The `git status` command will show files that are different from what git has in its ledger; in this example, two files have been modified and one new one created

This line shows what “branch” you are on; we’ll get to that later

These lines let us know what we may want to do next

```
test/ufs-srweather-app> git status
On branch release/public-v1
Your branch is up-to-date with 'origin/release/public-v1'.
Changes not staged for commit:
  (use "git add <file>..." to update what will be committed)
  (use "git checkout -- <file>..." to discard changes in working directory)

        modified:   README.md
        modified:   env/build_cheyenne_gnu.env

Untracked files:
  (use "git add <file>..." to include in what will be committed)

        Newfile.txt

no changes added to commit (use "git add" and/or "git commit -a")
```

git has recognized some changed files here

git also noticed there was an “untracked” file that it does not know about

Making commits with git

- A git “save point” is known as a *commit*
  - You can create a commit by modifying code, ”staging” that code for commit, and then committing
    - You can use the `git diff` command to see the changes you have made

```
diff --git a/README.md b/README.md
index 60144e2..ad1adec 100644
--- a/README.md
+++ b/README.md
@@ -9,3 +9,4 @@ https://github.com/ufs-community/ufs-srweather-app/wiki/Getting
-Started

 UFS Development Team. (2021, March 4). Unified Forecast System (UFS) Short-Ran
ge Weather (SRW) Application (Version v1.0.0). Zenodo. https://doi.org/10.5281/
zenodo.4534994

+Also, UFS is the best!
diff --git a/env/build_cheyenne_gnu.env b/env/build_cheyenne_gnu.env
index a30558c..36fc80d 100644
--- a/env/build_cheyenne_gnu.env
+++ b/env/build_cheyenne_gnu.env
@@ -14,3 +14,4 @@ export CMAKE_C_COMPILER=mpicc
 export CMAKE_CXX_COMPILER=mpicxx
 export CMAKE_Fortran_COMPILER=mpif90
 export CMAKE_Platform=cheyenne.gnu
+export AWESOME_VARIABLE='fix everything'
lines 1-18/18 (END)
```

Making commits with git

- A git “save point” is known as a *commit*
  - You can create a commit by modifying code, ”staging” that code for commit, and then committing
    - Use `git add` to stage a modified file for commit
    - For new files, you will also use `git add` to stage them for commit
    - To delete files, use `git rm`
    - You can use `git add` on full directories, or even use wildcards, but this is **strongly** discouraged
      - This makes it very easy to accidentally commit files you don’t want to commit, which can cause unintended consequences further down the line

```
>git add README.md env/build_cheyenne_gnu.env Newfile.txt
>git status
On branch release/public-v1
Your branch is up-to-date with 'origin/release/public-v1'.
Changes to be committed:
  (use "git reset HEAD <file>..." to unstage)

        new file:   Newfile.txt
        modified:   README.md
        modified:   env/build_cheyenne_gnu.env
```

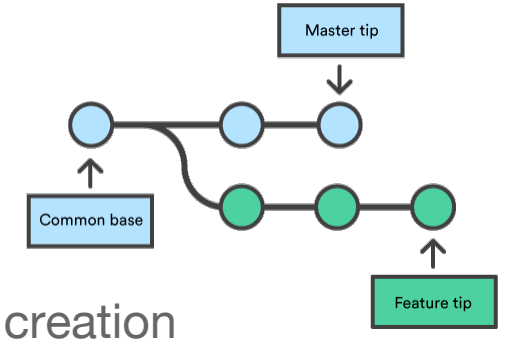
Making commits with git

- A git “save point” is known as a *commit*
  - You can create a commit by modifying code, ”staging” that code for commit, and then committing
    - Use `git commit` to commit the staged changes; this should bring up a text editor for you to enter your commit message
    - A commit message for your own changes can be as brief or as detailed as you like, but it should be enough to give a rough idea of what was changed and why.

```
Changed some files to make the UFS better. I think. It could be worse I guess.
Regardless, the real code is the friends we've made along the way.

# Please enter the commit message for your changes. Lines starting
# with '#' will be ignored, and an empty message aborts the commit.
# On branch release/public-v1
# Your branch is up-to-date with 'origin/release/public-v1'.
#
# Changes to be committed:
#       new file:   Newfile.txt
#       modified:   README.md
#       modified:   env/build_cheyenne_gnu.env
#
~
~
```

You will probably want to set the `GIT_EDITOR` *environment variable to your favorite text editor, otherwise you may end up in Emacs with no idea how to escape…*

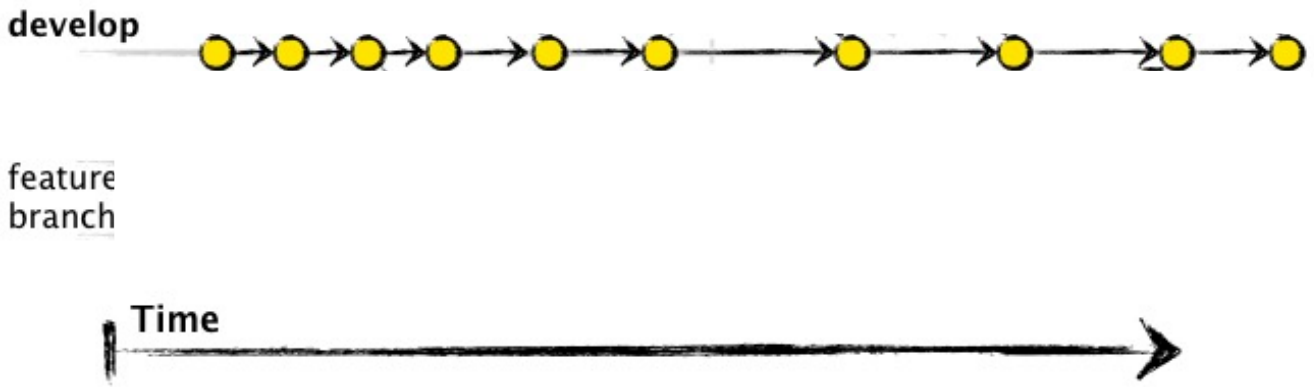
git branches



- The simplest repository will consist of a single, linear history all the way back to its creation
- However, it is useful to have the ability to work on multiple changes to a repository in parallel
- Git allows (and encourages) a “branch” functionality
  - Can be used for parallel development of different capabilities or fixes in the code
  - Can be used to separate the code undergoing active development from that being tested for a release, or being kept “stable” for some other purpose.
- If you never change anything, all commits will go on the main branch by default; this is often kept as the “authoritative” version of a project’s code
  - Most UFS components use “develop” as the main branch; for others it is “main”
  - The SRW App release that we have been using this week is on branches named “release/public-v##” depending on the release number of that component
  - The name of a branch does not typically matter; it’s just for human readability
- Use `git checkout -b your_branch_name` to create a new branch identical to the current branch; it is good practice to always create a new branch when making changes to the code that you will need to keep

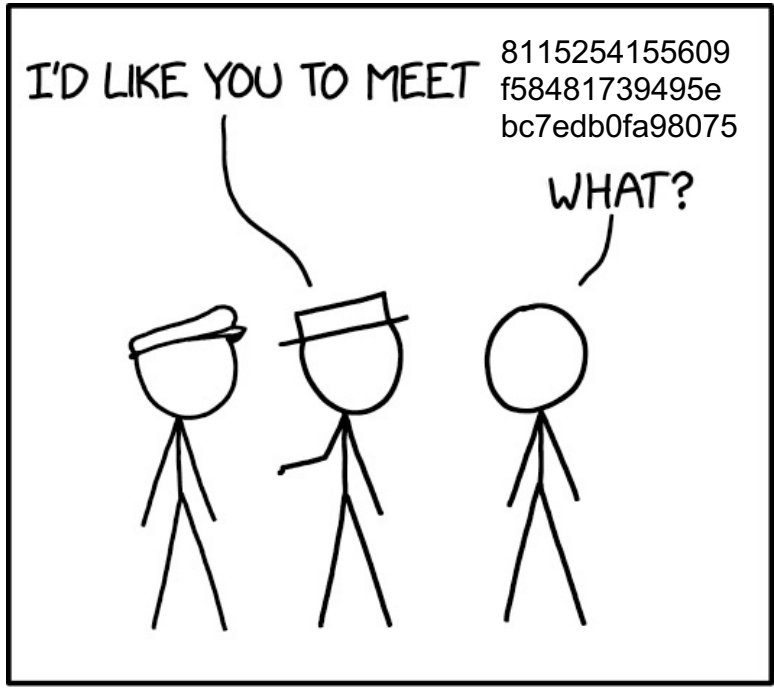
git branches

Author: Vincent Driessen
Original blog post: http://nvie.com/archives/323
License: Creative Commons



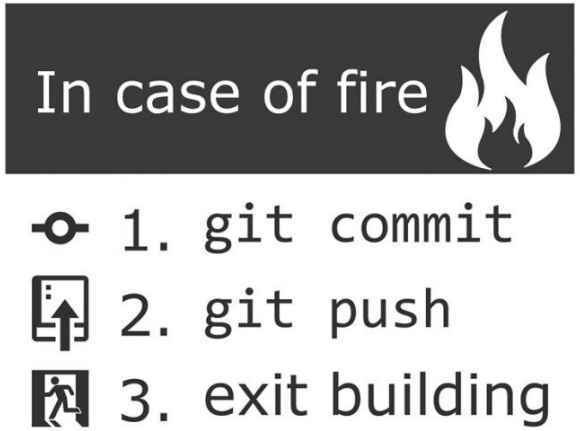
git tags

- As mentioned earlier, each git commit has a unique 40-character “hash” that identifies it
  - It is unique, but not very memorable
- Git tags allow a hash to be referenced in a human-readable way
  - Can be checked out just like branches
  - Essentially tags are an easily referenced, permanent “snapshot” of the code
- Git tags are typically created by repository managers for important events
  - An official code release, *e.g.* `v1.0.1`
  - A stable, well-tested version of the code
  - A reference to a specific event in the repository history that should be preserved



This isn't a paid advertisement for git, but…

- Even for tracking small projects on your personal machine, it's worth it
  - Just `git init` (creates new repository), `git add`, and `git commit`, and *voila*, you have a repository for your project!
  - Trust me, I wish someone had told me this in grad school
- git is an incredibly powerful tool, and we can only barely scratch the surface today. Some more very useful commands include:
  - `git diff` Can compare two files, two commits, two branches, etc.
  - `git merge` Can merge the changes from one branch to another
  - `git stash` Temporarily "stash away" your current changes without committing them
  - `git cherry-pick` Can move individual commits from one branch to another
  - `git blame` Gives a line-by-line summary of when and how each part of a file was last changed
  - `git bisect` Can help determine when a certain change occurred in the code history
- For more information on git, the official documentation is quite accessible
  - https://git-scm.com/docs/gittutorial

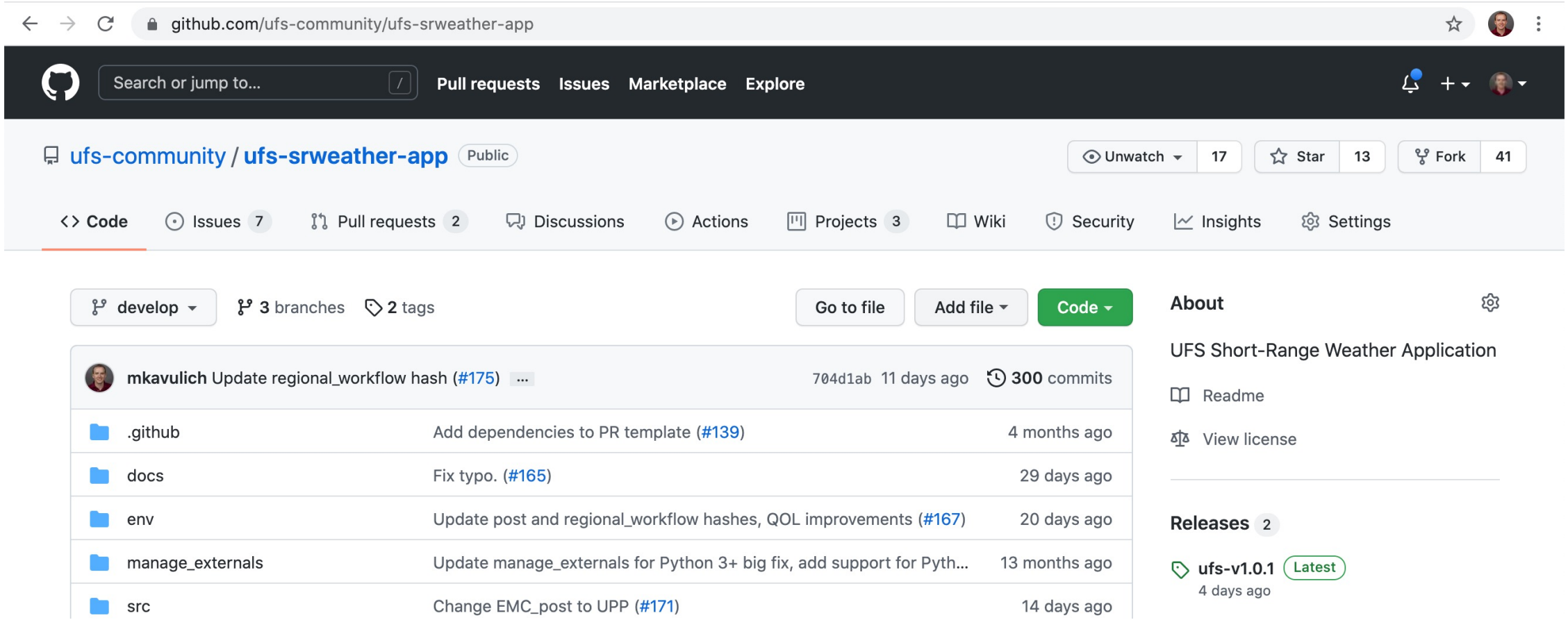


Git….Hub?

- GitHub is a website specifically for hosting and maintaining git repositories, as well as collaboration, testing, and documentation tools
- Like git itself, GitHub is widely used in the software development community
- GitHub allows for many additional capabilities on top of the built-in git functionality
  - Forks
  - Pull requests
  - Issue tracking
  - Wiki
  - etc.

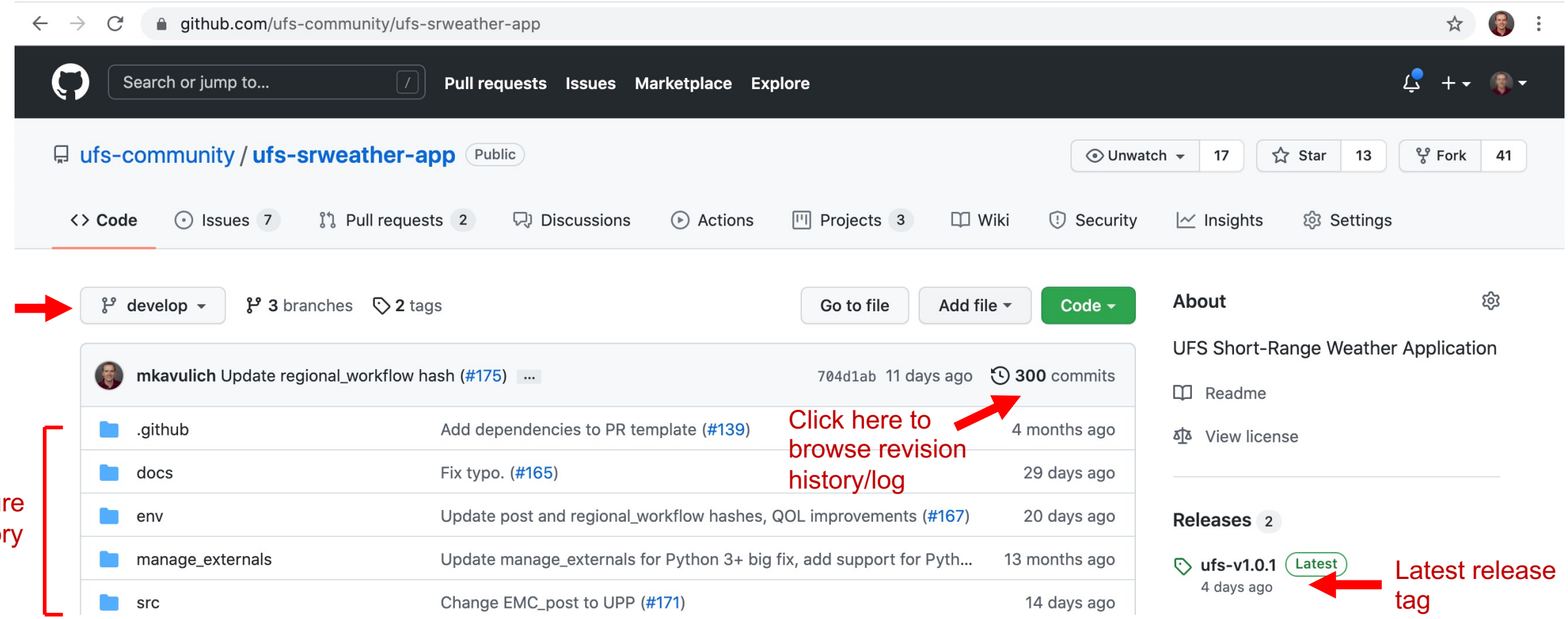
GitHub

- Example: https://github.com/ufs-community/ufs-srweather-app



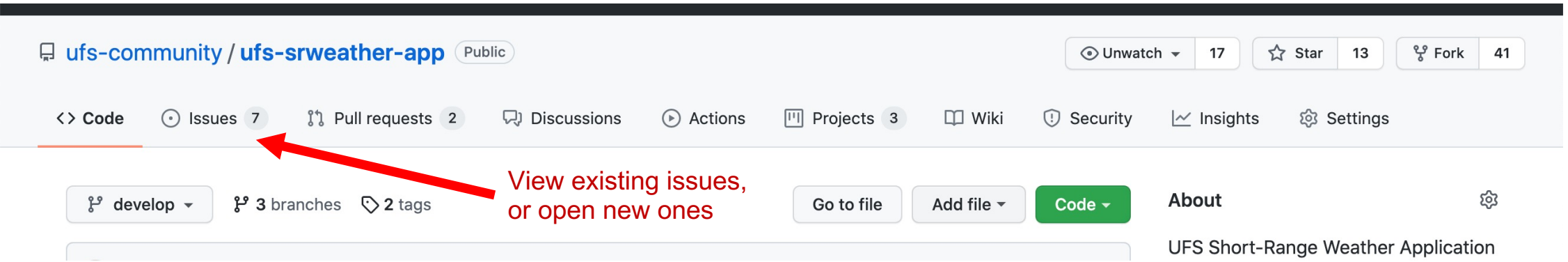
GitHub

- Example: https://github.com/ufs-community/ufs-srweather-app



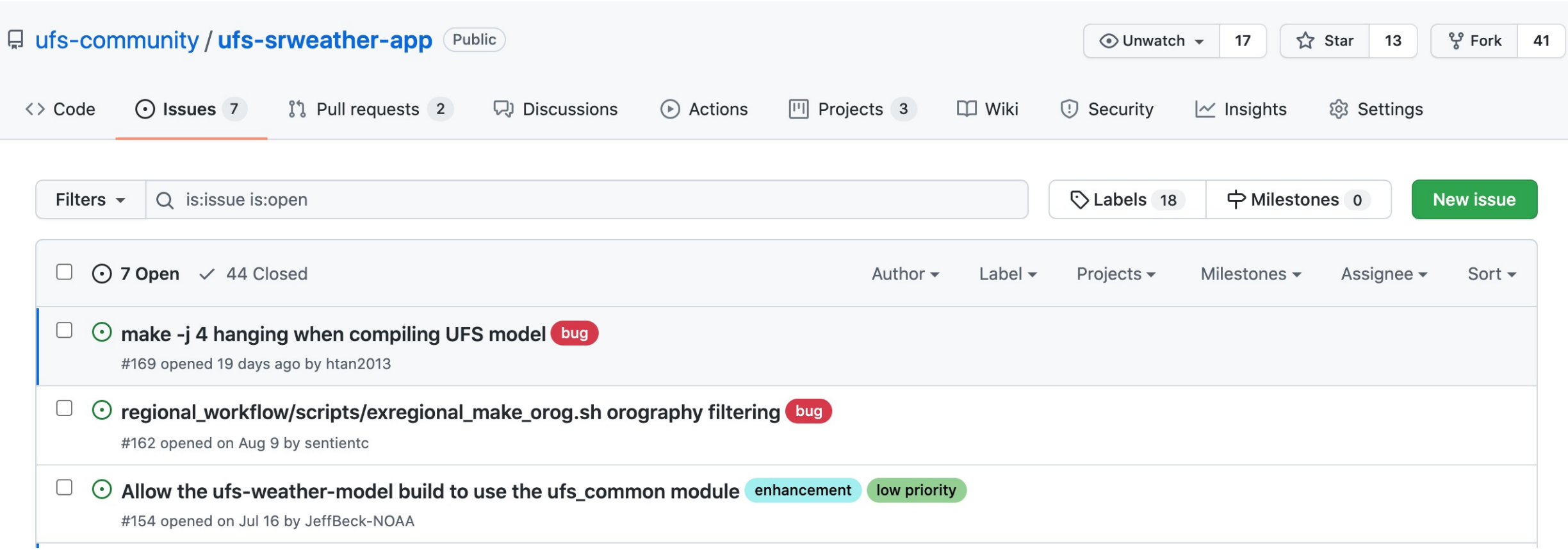
GitHub Issues

- The GitHub Issue Tracker is a great tool for communication with other collaborators on a given repository
  - Issues are simply numbered messages associated with a repository
  - Reasons for opening an issue include:
    - Pointing out a bug in the code
    - Requesting a feature
  - Typically issues consist of a title briefly describing the issue, followed by more detailed text
  - Issues can be closed (resolved) by Pull Requests



GitHub Issues

- https://github.com/ufs-community/ufs-srweather-app/issues



GitHub Issues

Fix to allow quilting with non-factors for layout #244

Open **chan-hoo** opened this issue 14 days ago · 3 comments

chan-hoo commented 14 days ago

Description:

When the layout values (lx, ly) in the input file 'input.nml' are not factors of (npx-1) and (npy-1) respectively, quilting does not work.

Solution:

Gerhard's update will be applied to the develop branch.
ESMF's decomposition flags will be added to the two files in 'ufs-weather-model/FV3/': (1) /io/module_wrt_grid_comp.F90, and (2) module_fcst_grid_comp.F90.

chan-hoo added the bug label 14 days ago

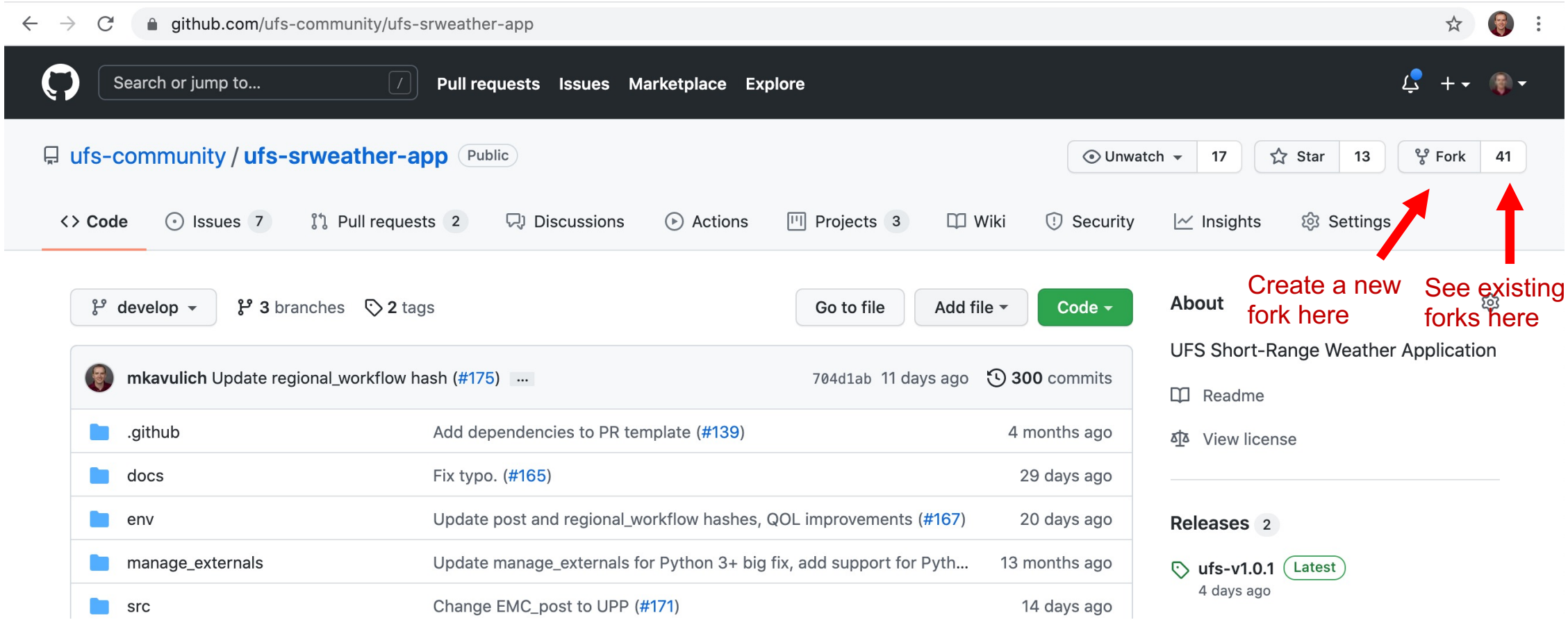
gsketefian commented 14 days ago Member

Hi all,
The code freeze for release-v1 of the UFS short-range weather app is October 30 (this Friday). It would be great to resolve this issue ASAP because it will simplify how the grid parameters are set. It would have to go into the release/public-v2 branch (but hopefully simultaneously into develop). Once this is resolved, we (the DTC-UFS-CAM team) need a couple more days to then fine-tune and test the three supported grids (RRFS_CONUS_25km, RRFS_CONUS_13km, and RRFS_CONUS_3km) and modify the workflow scripts accordingly. **@jwolff-ncar @JeffBeck-NOAA @llpcarson @mkavulich @JulieSchramm @fentoad72**

GitHub Forks

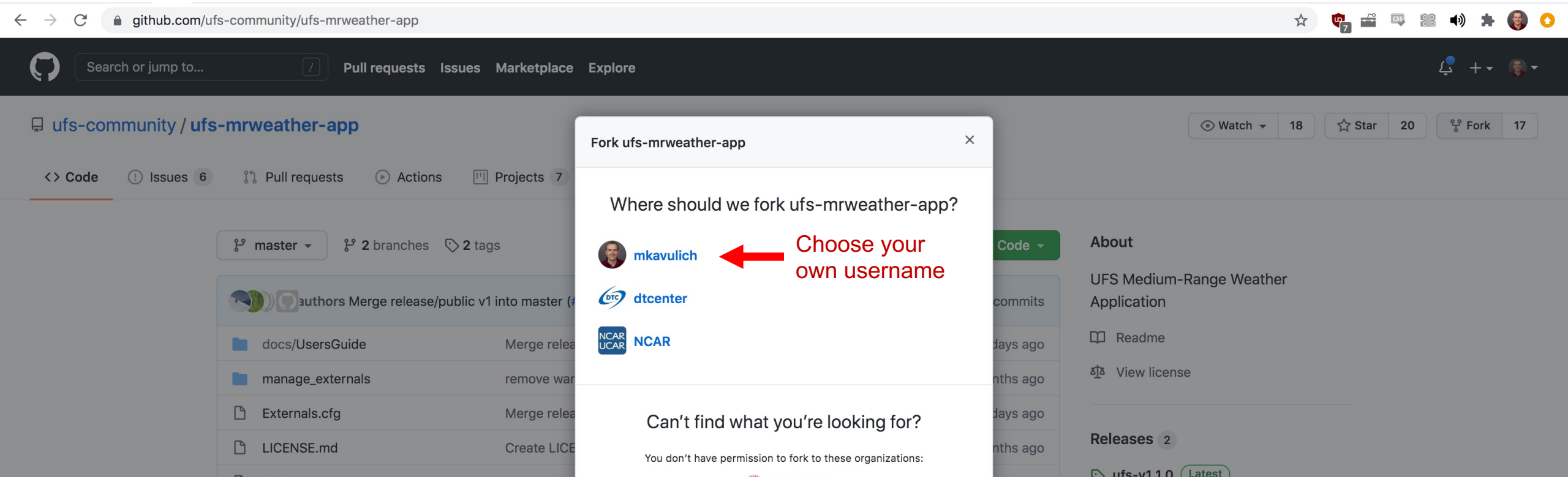
- GitHub allows individuals to keep their own copy of the “authoritative” repository; this is known as a “fork”
  - A fork, like every other git repository, is a full, stand-alone repository, containing the entire commit history, all branches and tags
  - The fork is stored under your own GitHub account, and you have full permissions to make as many changes as you want without affecting the authoritative repository

GitHub Forks



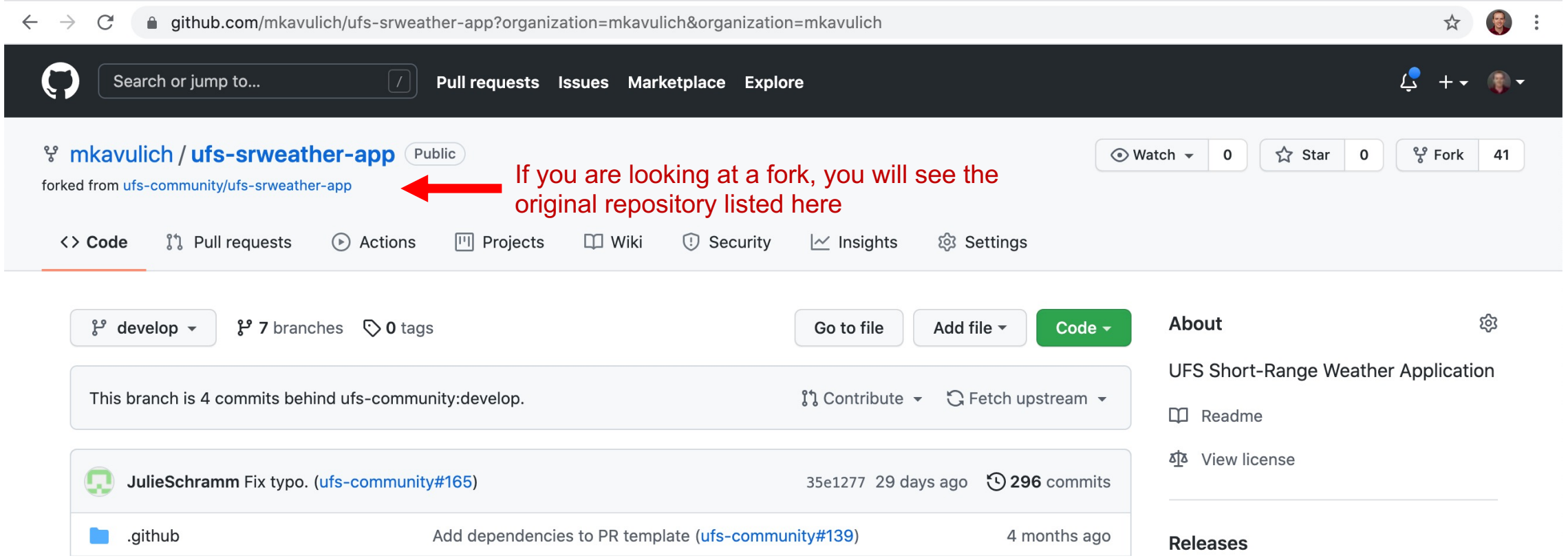
GitHub Forks

- All development and new contributions should come from a user's fork
  - You might see a box asking where the fork should be created; choose your username



GitHub Forks

- All development and new contributions should come from a user's fork
  - After forking, you should see the same code you did before, but at a different URL



GitHub Forks

- All development and new contributions should come from a user's fork
  - Once your fork is created, in order to do work with the code and make changes to the code, you will clone your fork instead of the main repository

Cloning the main repository:

```
git clone https://github.com/ufs-community/ufs-srweather-app.git
```

Cloning your fork:

```
git clone https://github.com/YOUR_GITHUB_USERNAME/ufs-srweather-app.git
```

  - Aside from the different URL, working in a clone of your fork is the same as working in a clone of the main repository
  - Don't make changes directly to the develop branch! Instead, start with develop and create a new branch for your changes

A bit of git we haven't covered yit: git push and pull

- When commits are made, they are initially only on the local clone of your repository
- In order to get your code changes back to the main repository on GitHub, you will need to "push" those commits back to the origin, using the `git push` command

```
>git push
Username for 'https://github.com': mkavulich
Password for 'https://mkavulich@github.com':
Counting objects: 6, done.
Delta compression using up to 72 threads.
Compressing objects: 100% (5/5), done.
Writing objects: 100% (6/6), 560 bytes | 0 bytes/s, done.
Total 6 (delta 4), reused 0 (delta 0)
remote: Resolving deltas: 100% (4/4), completed with 4 local objects.
remote:
remote: Create a pull request for 'test' on GitHub by visiting:
remote:      https://github.com/mkavulich/ufs-weather-model/pull/new/test
remote:
To https://github.com/mkavulich/ufs-weather-model
 * [new branch]      test -> test
>
```

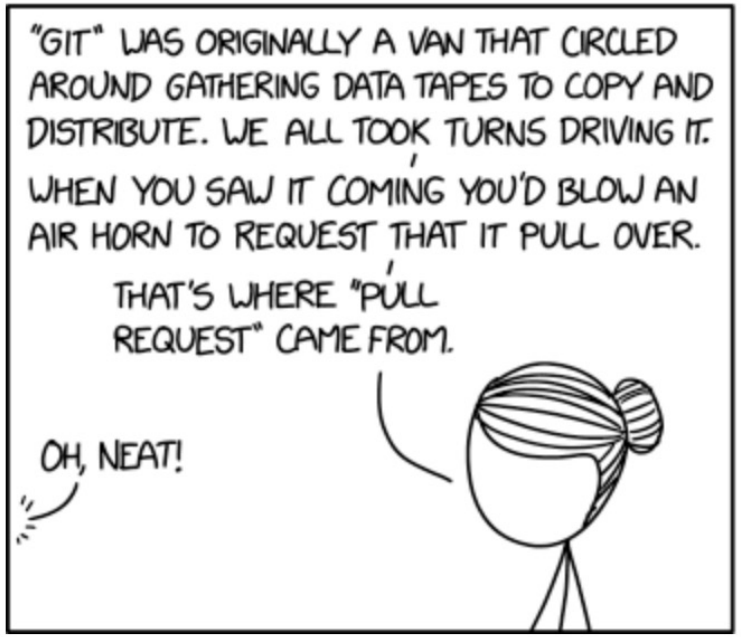
A bit of git we haven't covered yit: git push and pull

- When commits are made by others to the main repository, they do not automatically populate into your local clone
- In order to get the most up-to-date code from GitHub, you will need to "pull in" the latest commits, using the `git pull` command

```
>git pull
From https://github.com/mkavulich/ufs-weather-model
 * [new branch]      develop                -> mkavulich/develop
 * [new branch]      production/GFS.v16     -> mkavulich/production/GFS.v16
 * [new branch]      production/GFS_v15     -> mkavulich/production/GFS_v15
 * [new branch]      production/HREF.v3     -> mkavulich/production/HREF.v3
 * [new branch]      production/HREF.v3beta -> mkavulich/production/HREF.v3beta
 * [new branch]      release/public-v1      -> mkavulich/release/public-v1
 * [new branch]      release/public-v2      -> mkavulich/release/public-v2
Fetching submodule FV3
From https://github.com/NOAA-EMC/fv3atm
   6e2421c..5530a85  develop           -> origin/develop
   f5da715..6a56b8a  release/public-v2 -> origin/release/public-v2
Fetching submodule FV3/atmos_cubed_sphere
Fetching submodule FV3/ccpp/physics
From https://github.com/NCAR/ccpp-physics
   f3e6761..c95a1ae  master     -> origin/master
```

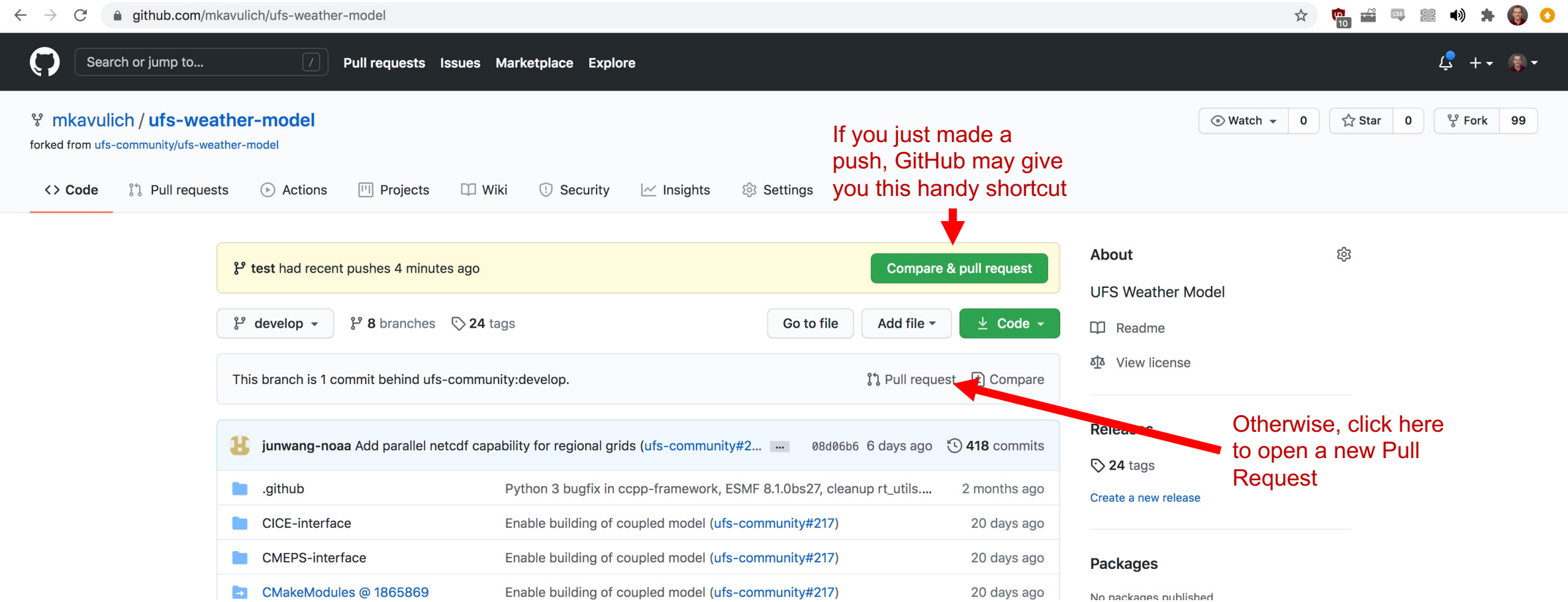
GitHub Pull Requests

- If you would like to make a change to a repository, you can do so via a “Pull Request”
  - A pull request, often abbreviated PR, is a request to have your changes “pulled” in to the official repository from your fork
  - A PR can be applied between any two branches in any repositories with a common history, but traditionally they are applied from a fork to the main repository
  - When opening a PR, it is generally expected you will provide a description of the changes, a justification for the changes (fixing a bug, adding a feature, etc.), and a summary of tests conducted
    - Different projects will have different requirements: more on that later



You are welcome to blow an air horn when opening a pull request if you wish. Just do it in your own lab.

GitHub Pull Requests

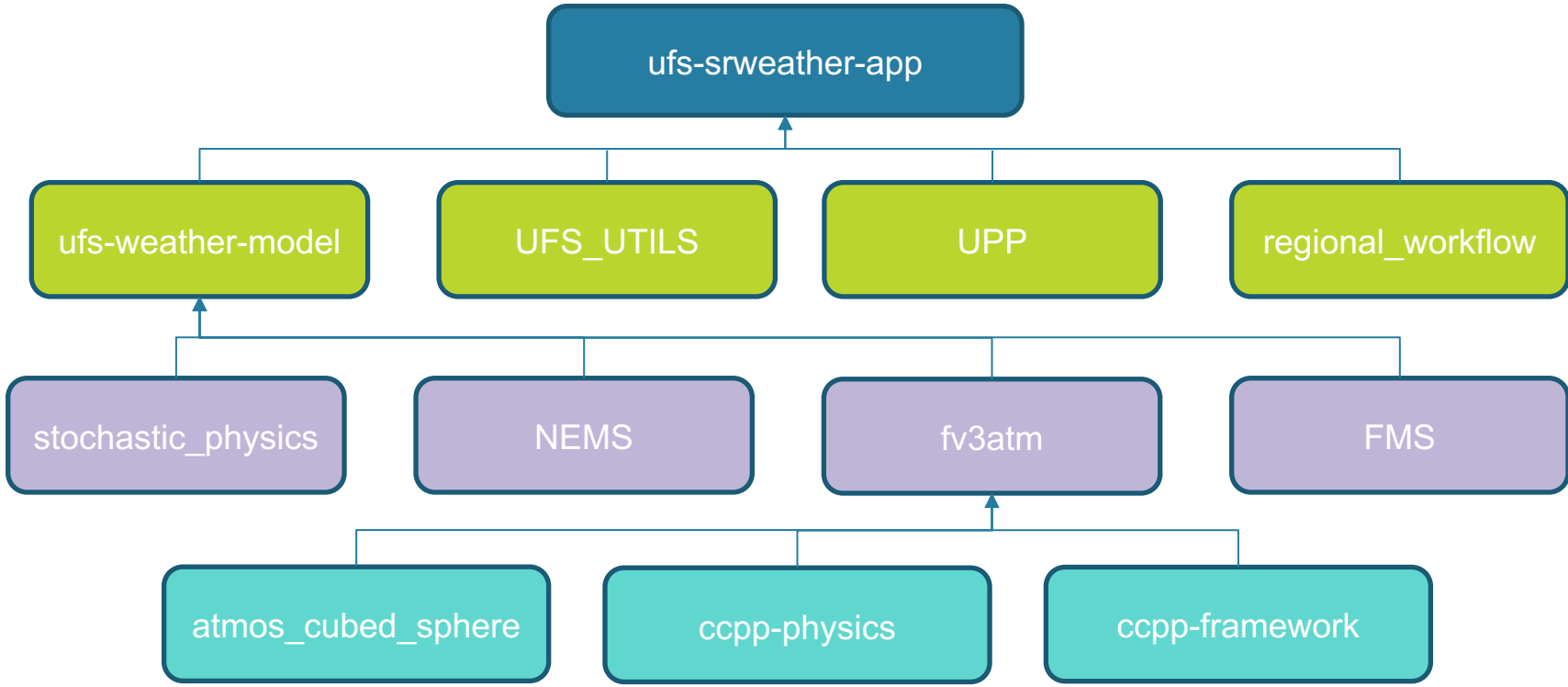


Outline

- Version control: Git and Github
  - Version control overview
  - Using git software on the command line
  - Using Github
- UFS-SRW structure, Submodules, and Manage Externals
- Making and contributing changes to the UFS code
  - Changing code for development purposes
  - Contributing code back to individual repositories
    - Testing requirements
    - Pull requests

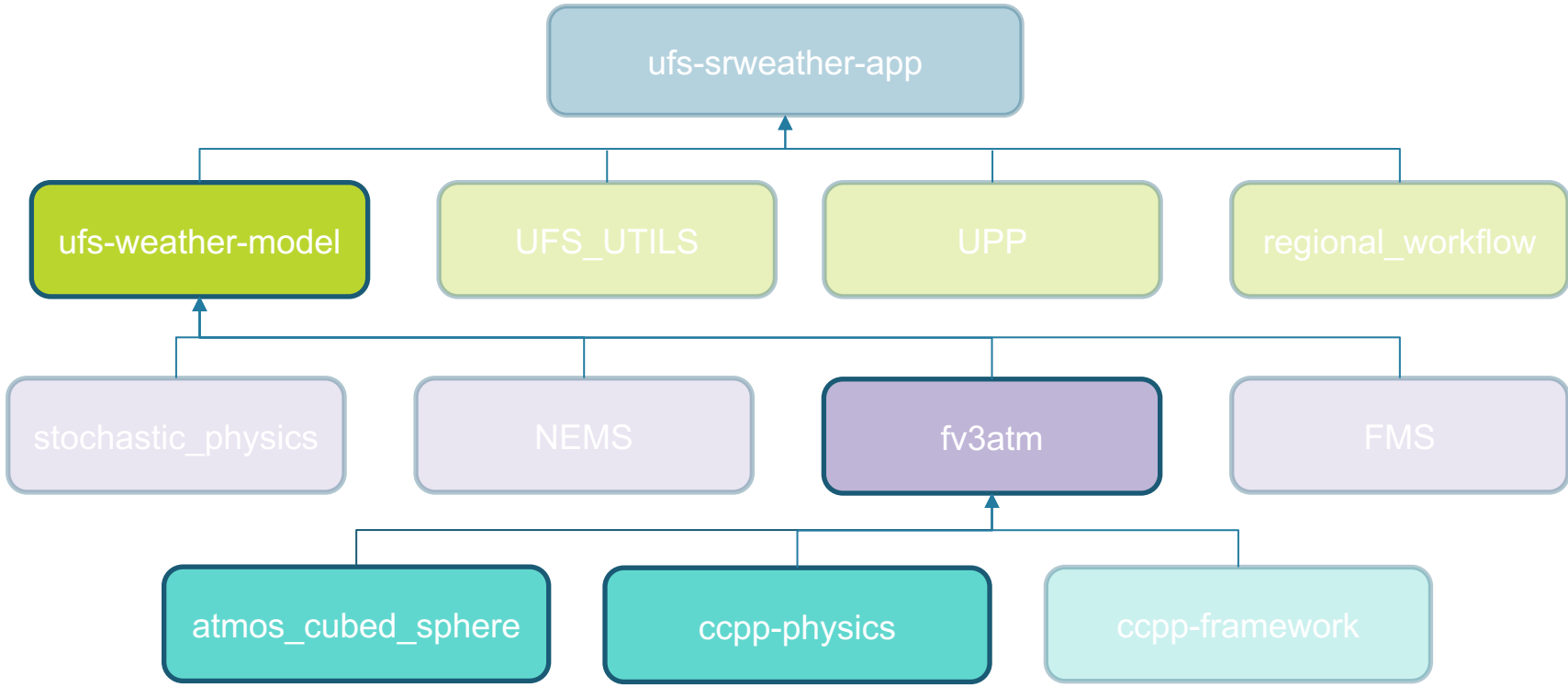
UFS-SRW Structure

- The UFS-SRW is composed of a number of individual, stand-alone codes, most of which were initially independent components
- Each of these components is in its own separate repository
- Many components, such as the weather model, have additional sub-components: not all are listed here!



UFS-SRW Structure

- ufs-weather-model
  - The main repository for the weather model and its components
- fv3atm
  - Contains the atmospheric component of the weather model
- ccpp-physics
  - Contains the GFS/RRFS physics schemes
- atmos_cubed_sphere
  - Contains the FV3 dynamical core

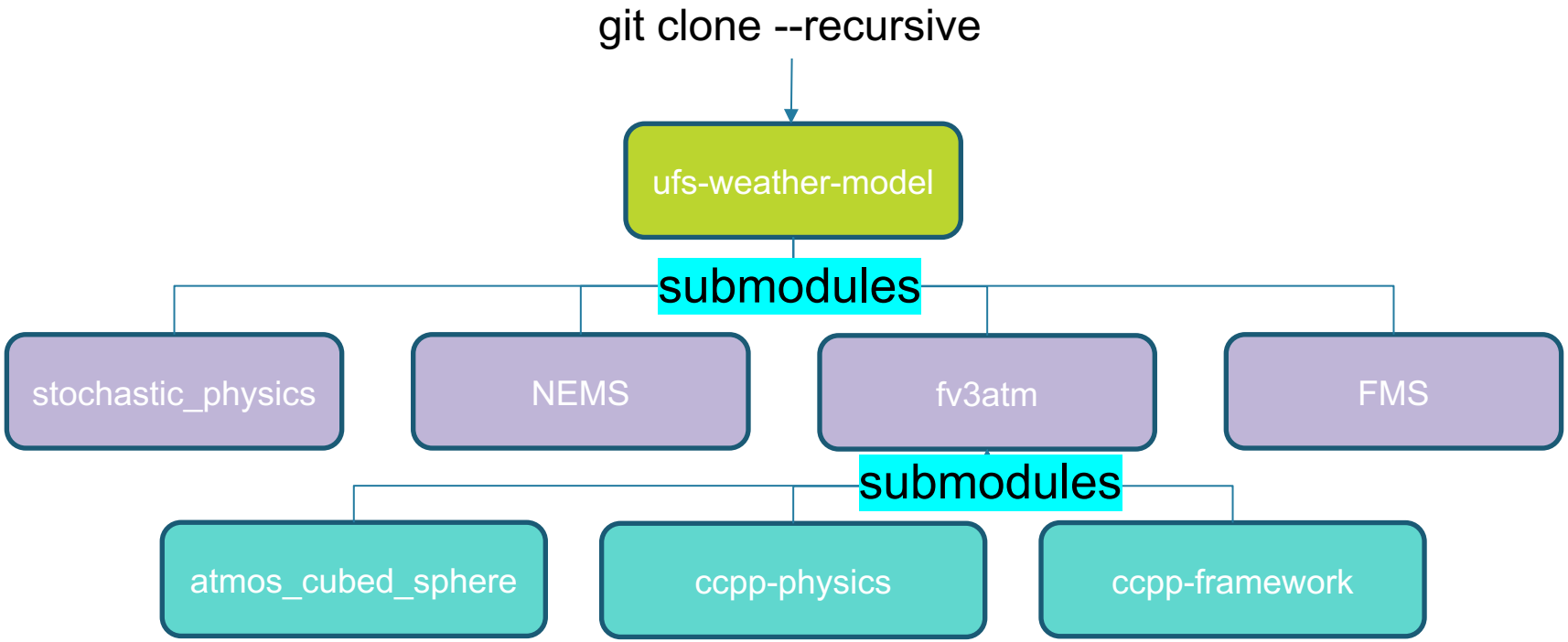


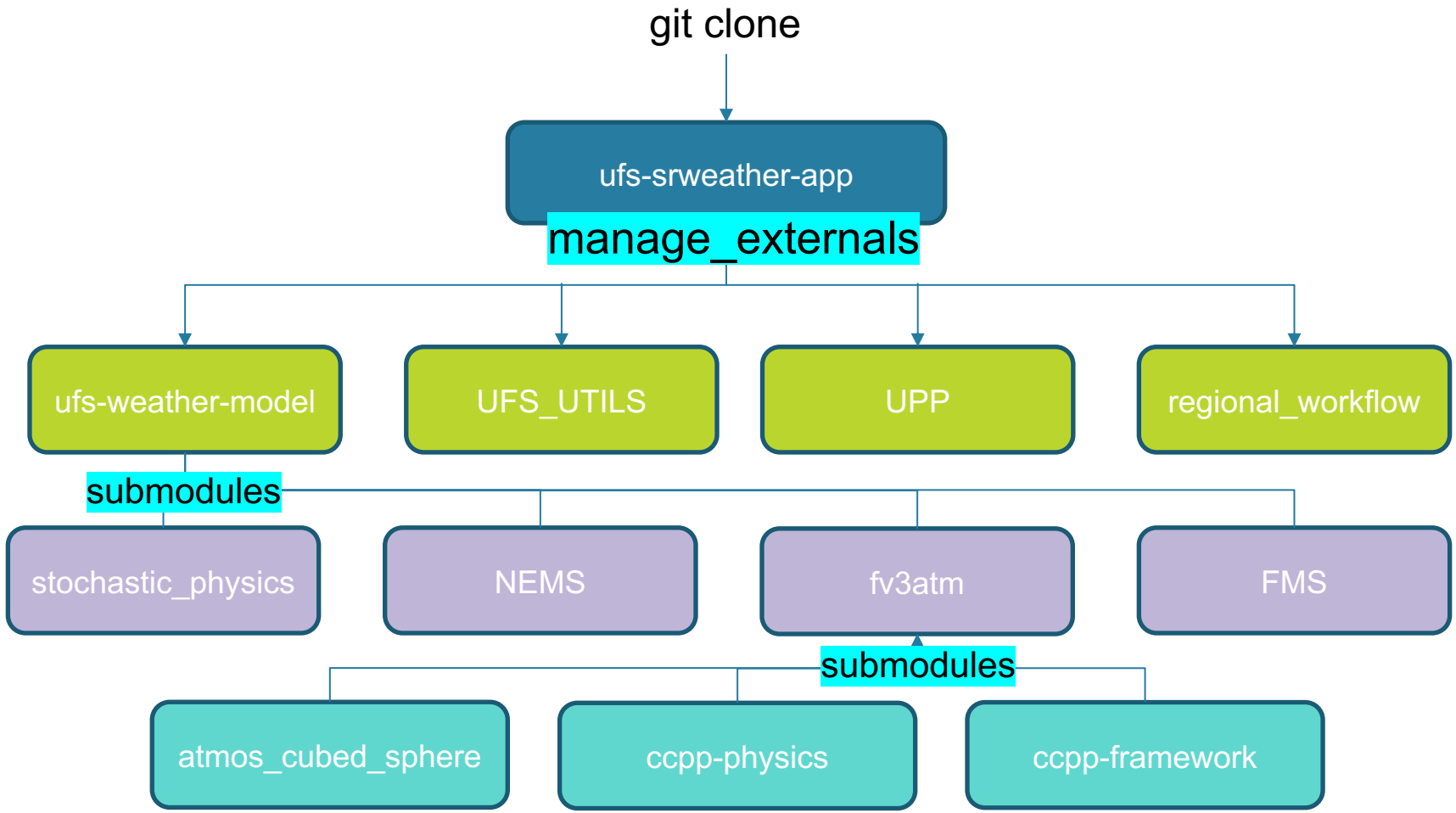
Submodules, and Manage Externals

- How are all these repositories linked together? Surely it's a nightmare to keep track of all the changes going into every repository…
- This is handled through manage_externals and submodules
  - Submodules are a native functionality of git (https://git-scm.com/book/en/v2/Git-Tools-Submodules)
    - A repository can be linked as a subdirectory as another repository
    - Submodules are tracked in a top-level ".gitmodules" file
  - manage_externals is a tool developed and maintained by a number of people at NCAR and other national labs for use with CESM (https://github.com/ESMCI/manage_externals)
    - Adds some additional functionality on top of submodules
    - External repositories are tracked in the top-level "Externals.cfg" file

```
[regional_workflow]
protocol = git
repo_url = https://github.com/NOAA-EMC/regional_workflow
tag = ufs-v1.0.1
local_path = regional_workflow
required = True

[ufs_utils]
protocol = git
repo_url = https://github.com/NOAA-EMC/UFS_UTILS
tag = ufs-v2.0.0
local_path = src/UFS_UTILS
required = True

[ufs_weather_model]
protocol = git
repo_url = https://github.com/ufs-community/ufs-weather-model
tag = ufs-v2.0.0
local_path = src/ufs_weather_model
required = True

[EMC_post]
protocol = git
repo_url = https://github.com/NOAA-EMC/EMC_post
tag = upp_v9.0.1
local_path = src/EMC_post
required = True

[externals_description]
schema_version = 1.0.0
```

Externals.cfg for the ufs-srweather-app version 1.0.1 release


git clone --recursive
ufs-weather-model
submodules
stochastic_physics
NEMS
fv3atm
FMS
submodules
atmos_cubed_sphere
ccpp-physics
ccpp-framework


git clone
ufs-srweather-app
manage_externals
ufs-weather-model
UFS_UTILS
UPP
regional_workflow
submodules
stochastic_physics
NEMS
fv3atm
FMS
submodules
atmos_cubed_sphere
ccpp-physics
ccpp-framework

Outline

- Version control: Git and Github
  - Version control overview
  - Using git software on the command line
  - Using Github
- UFS-SRW structure, Submodules, and Manage Externals
- Making and contributing changes to the UFS code
  - Changing code for development purposes
  - Contributing code back to individual repositories
    - Testing requirements
    - Pull requests

Changing code for development purposes

- The SRW App uses a `cmake`-based build system that allows for easy rebuilding of the code after making changes
- For most changes, modify the code in the ufs-srweather-app/src directory, then re-run `make`
- The `make` utility keeps track of changed files, and only rebuilds the necessary code

Caveat: If you make changes to the build system, prerequisite libraries, or other dependencies, you will need to re-build from scratch

```
>vi src/UFS_UTILS/sorc/orog_mask_tools.fd/orog.fd/mtnlm7_oclsm.f
>cd build/
>make -j 4
[  4%] Performing build step for 'UFS_UTILS'
[ 12%] Performing build step for 'EMC_post'
[ 12%] Performing build step for 'ufs_weather_model'
[  1%] Built target ccpp
[ 15%] Built target shared_lib
[  6%] Built target cpl
[ 18%] Built target shave
[ 20%] Built target fms
[ 86%] Built target nceppost
[ 21%] Built target filter_topo
[ 39%] Built target regional_esg_grid
[100%] Built target ncep_post
[ 42%] Built target global_equiv_resol
[ 28%] Built target stochastic_physics
Scanning dependencies of target orog
[ 16%] Performing install step for 'EMC_post'
[ 62%] Built target chgres_cube
[ 63%] Building Fortran object sorc/orog_mask_tools.fd/orog.fd/CMakeFiles/orog.dir/mtnlm7_oclsm.f.o
[ 80%] Built target ccppphys
[ 77%] Built target sfc_climo_gen
[ 83%] Built target vcoord_gen
[ 83%] Built target gfsphysics
[ 86%] Built target nceppost
[ 83%] Built target ipd
[ 87%] Built target make_solo_mosaic
[100%] Built target ncep_post
[ 83%] Built target ccppdriver
[ 96%] Built target make_hgrid
Install the project...
-- Install configuration: "Release"
[ 92%] Built target fv3dycore
[ 95%] Built target io
[ 96%] Built target stochastic_physics_wrapper
-- Up-to-date: /glade/scratch/kavulich/UFS_SRW/testing/test_release/ufs-srweather-app/bin/ncep_post
-- Up-to-date: /glade/scratch/kavulich/UFS_SRW/testing/test_release/ufs-srweather-app/include
```

Made changes to a single file

Only that file (and any dependencies) are re-built, making the process much faster than the initial build!

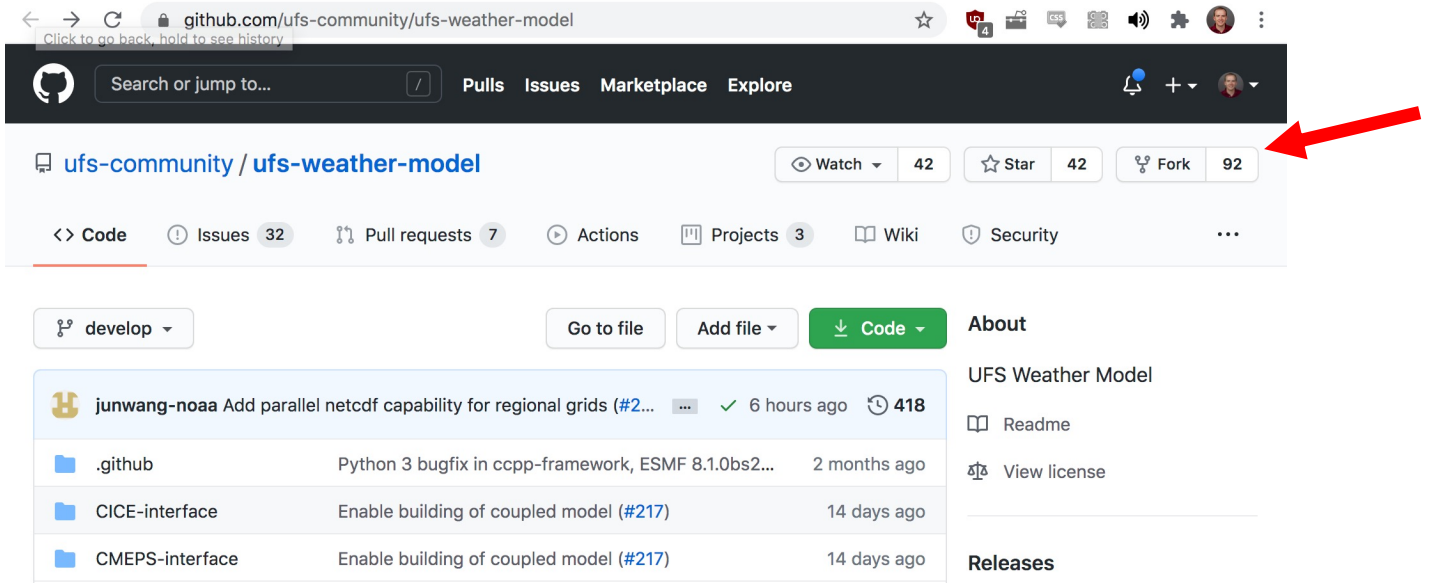
Contributing code back to the UFS

- The first thing you will need to do is switch to the develop branch of the repository
  - The released code is version 1.0.1, and is frozen aside from bug fixes since early 2021. By the time you start your development, the main branch will be far ahead of the release branch
  - Just running git clone https://github.com/ufs-community/ufs-srweather-app (without the `-b` argument) will check out the default "develop" branch
- In the develop branch, the latest well-tested hashes of each repository are cloned; these are far more recent than the tags used in the release branch
- Before making changes, you should create a "fork" of the repository/ies you will be changing

```
[regional_workflow]
protocol = git
repo_url = https://github.com/NOAA-EMC/regional_workflow
# Specify either a branch name or a hash but not both.
#branch = develop
hash = 9bac8563
local_path = regional_workflow
required = True

[ufs_utils]
protocol = git
repo_url = https://github.com/NOAA-EMC/UFS_UTILS
# Specify either a branch name or a hash but not both.
#branch = develop
hash = ea821358
local_path = src/UFS_UTILS
required = True

[ufs-weather-model]
protocol = git
repo_url = https://github.com/ufs-community/ufs-weather-model
# Specify either a branch name or a hash but not both.
#branch = develop
hash = 3b8bb78
local_path = src/ufs-weather-model
required = True

[UPP]
protocol = git
repo_url = https://github.com/NOAA-EMC/UPP
# Specify either a branch name or a hash but not both.
#branch = develop
hash = a49af05
local_path = src/UPP
required = True
```

Contributing code back to the UFS: Creating a fork

- To fork a repository, you will need to create a GitHub account (or log in to an existing one).
- Then go to the repository you are interested in; in this case the ufs-weather-model
  - https://github.com/ufs-community/ufs-weather-model
- In the top right, there is a “Fork” button that you can click to create a fork



Contributing code back to the UFS : Committing changes

- Once your fork has been created at https://github.com/YourUsername/ufs-weather-model, you can then modify Externals.cfg to check out your fork instead of the authoritative repository
  - repo_url = https://github.com/~~ufs-community~~/ufs-weather-model/ → YourUsername
- Then run `./manage_externals/checkout_externals` to check out the code as usual; this time, instead of cloning the authoritative ufs-community/ufs-weather-model repository, manage_externals will clone your fork of ufs-weather-model
- Create a new branch and commit changes to your fork as described earlier.
  - `git add newfile changed_file`
  - `git commit -m 'Added new file and changed another one...for science!'`
- Push your changes back to your fork on GitHub
  - `git push --set-upstream origin branchname`

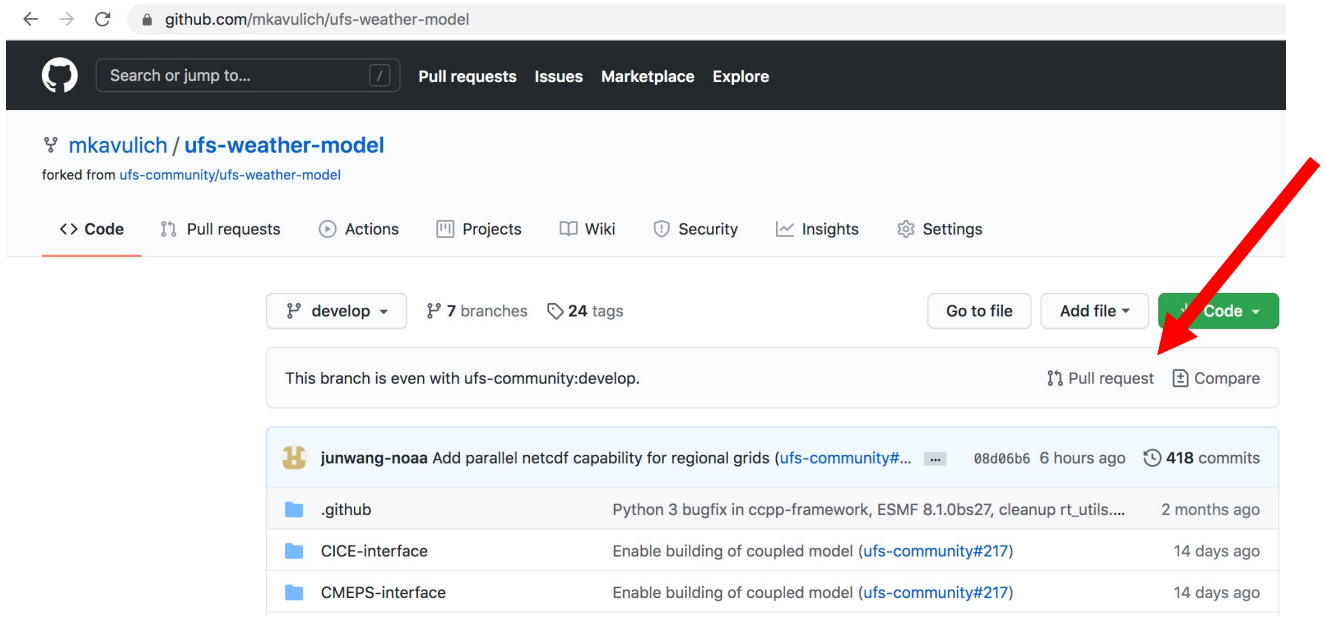
```
[regional_workflow]
protocol = git
repo_url = https://github.com/mkavulich/regional_workflow
# Specify either a branch name or a hash but not both.
branch = feature/adding_a_thing
#hash = 9bac8563
local_path = regional_workflow
required = True

[ufs_utils]
protocol = git
repo_url = https://github.com/NOAA-EMC/UFS_UTILS
# Specify either a branch name or a hash but not both.
#branch = develop
hash = ea821358
local_path = src/UFS_UTILS
required = True

[ufs-weather-model]
protocol = git
repo_url = https://github.com/mkavulich/ufs-weather-model
# Specify either a branch name or a hash but not both.
branch = feature/adding_a_related_thing
#hash = 3b8bb78
local_path = src/ufs-weather-model
required = True

[UPP]
protocol = git
repo_url = https://github.com/NOAA-EMC/UPP
# Specify either a branch name or a hash but not both.
#branch = develop
hash = a49af05
local_path = src/UPP
required = True
```

Now when I run manage_externals, it will clone my forks instead of the authoritative repositories

I haven't changed anything in UPP, so I'll leave this as-is

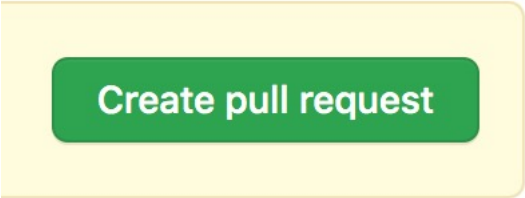
Contributing code back to the UFS: Opening a Pull Request

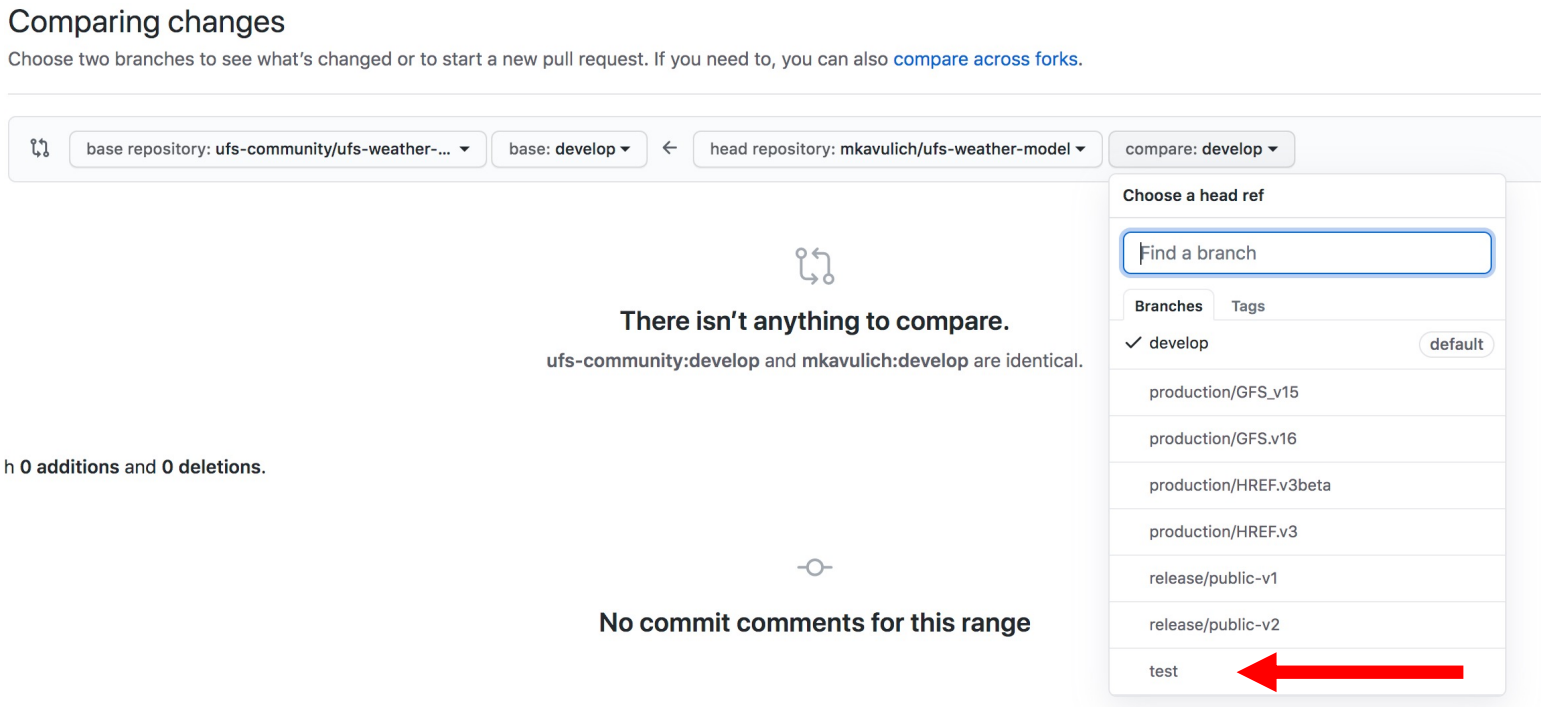
- Once your changes are working to your satisfaction, and have been pushed back to GitHub, you are now ready to open a Pull Request
- Visit your fork on GitHub via your favorite internet browser, and click “Pull Request”



Contributing code back to the UFS: Opening a Pull Request

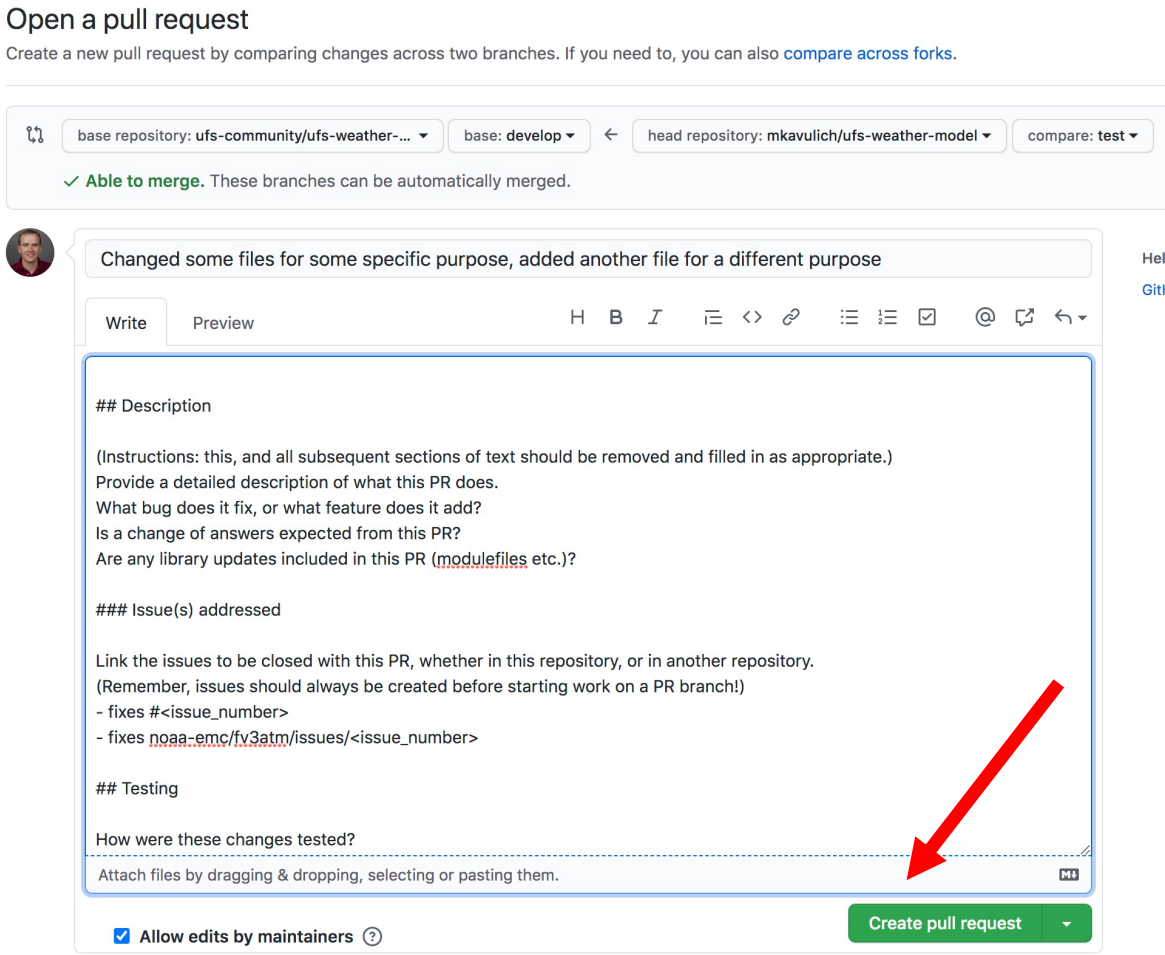
- From the dropdown menu at right, select the branch you just pushed to your fork
- After selecting the correct branch, select “Create pull request”





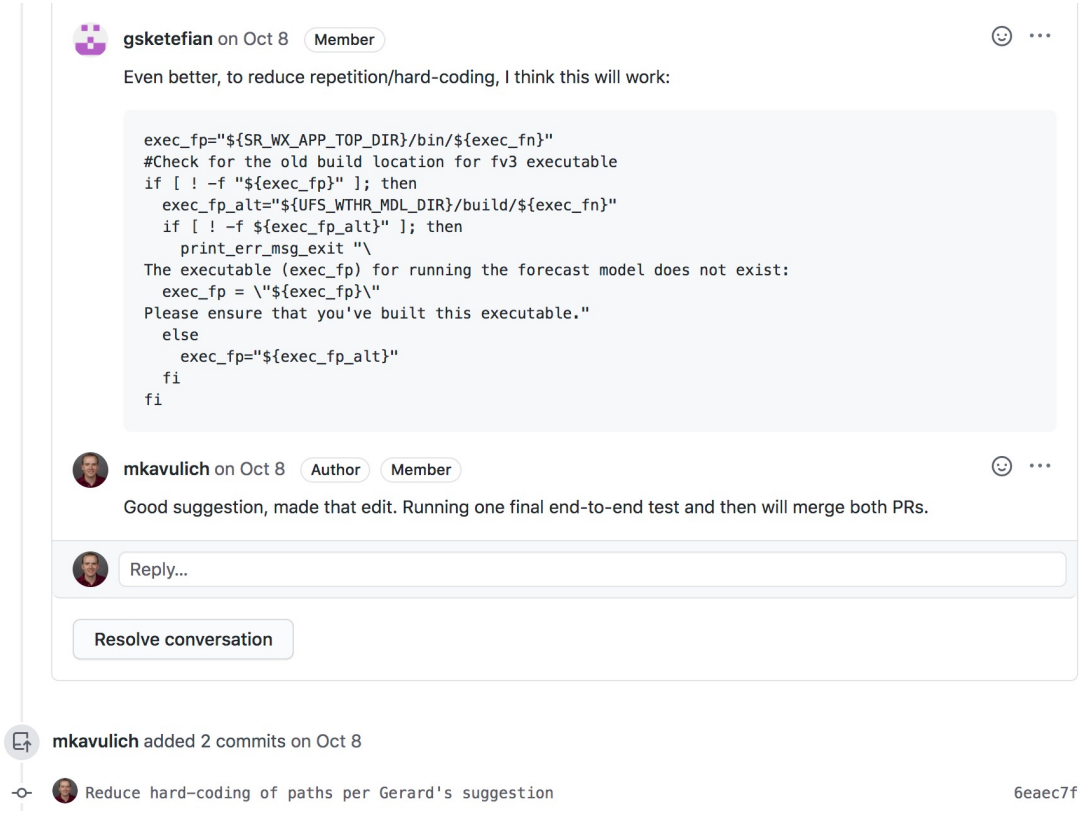
Contributing code back to the UFS: Opening a Pull Request

- And now, it's time to make your request!
  - Create a brief but descriptive title in the first box
  - Add more details about the changes and their purpose in the large box
    - For PRs that consist of many commits, this is where your own commit message history can come in handy; if you have been including descriptive commit messages all along then this step is a lot less work!
  - For some repositories, ufs-weather-model included, the message box will be filled in with a template; in that case you should follow the instructions provided
- When you are finished filling in all the details, you can hit "Create pull request" to open the PR



Contributing code back to the UFS: Continuing a Pull Request

- Be prepared to respond to questions or concerns from code managers and other community members!
- Make requested changes to the code so that your Pull Request can be approved and merged to the main branch
  - Pull requests are tied to a specific branch, so if you need to make changes, simply add new commits to that branch and push them back to GitHub

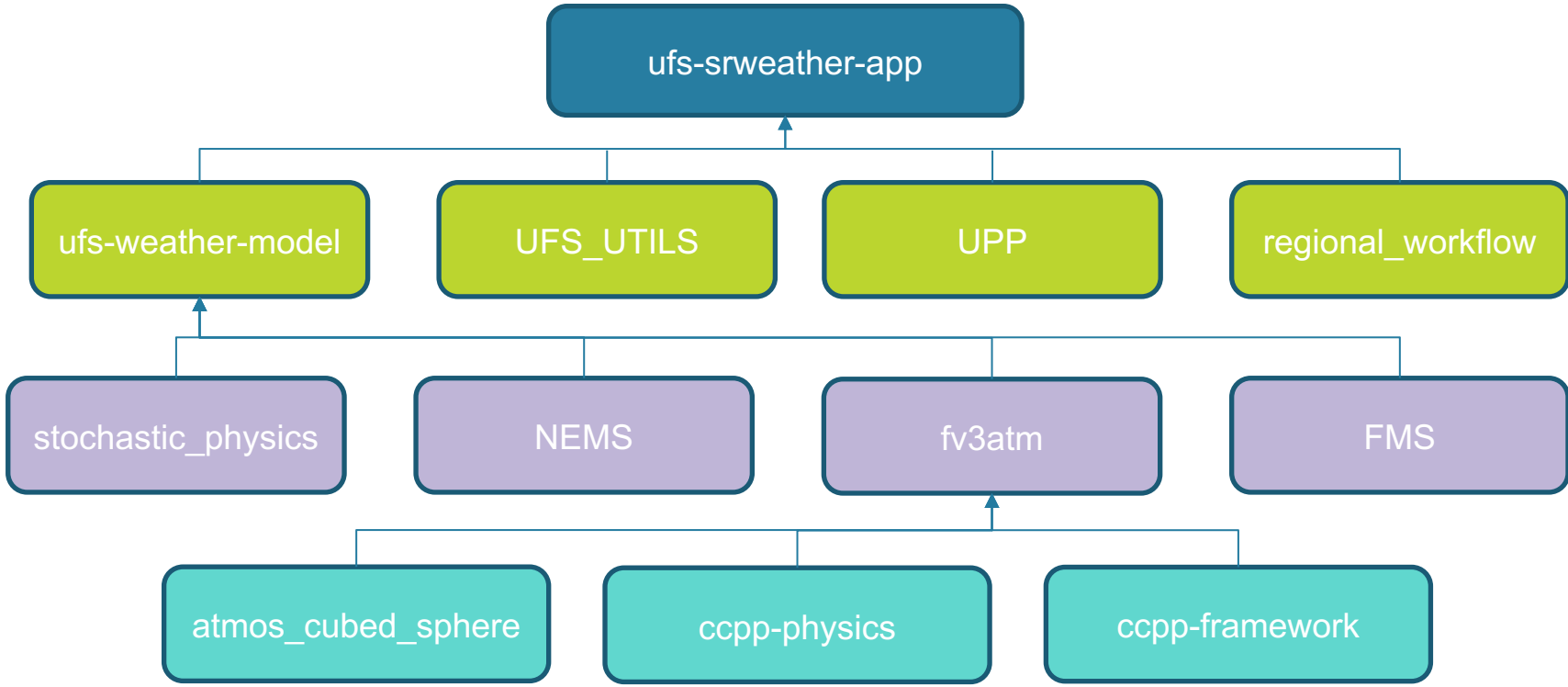


Testing requirements

- Most components of the UFS have some kind of testing system for ensuring that changes to the code are working correctly and do not break existing capabilities
  - The weather model has a fairly extensive regression testing system
    - https://github.com/ufs-community/ufs-weather-model/wiki/Running-regression-test-using-rt.sh
  - The UFS_UTILS repository has an extensive but still developing regression testing system, as well as unit testing requirements
  - UPP has a basic regression testing suite, but is growing rapidly
  - regional_workflow, and changes to the ufs-srweather-app itself, can be tested with the Workflow End-to-End (WE2E) tests
    - https://github.com/NOAA-EMC/regional_workflow/tree/develop/tests/WE2E
- These tests will need to pass before changes can be accepted into these code repositories
  - You can speed the code review process along by completing testing prior to opening your PR
  - These tests are not meant to be a burden; ask a code manager or developer for help!

Requirements for different repositories

The example I used above is for the ufs-weather-model, but different repositories have different requirements for PRs; these will be briefly detailed in the following slides



Development requirements: ufs-srweather-app and regional_workflow

- https://github.com/ufs-community/ufs-srweather-app/
- https://github.com/NOAA-EMC/regional_workflow/
- Requirements for contributing code:
  - All code changes should be submitted via pull requests from feature branches on user forks
  - End-to-end tests should be run on one or more supported platforms
- We are in the process of developing a more formal testing and code contribution process; see the relevant repository wikis for the latest and greatest information!
  - https://github.com/NOAA-EMC/regional_workflow/wiki/Code-development-and-review-in-the-regional_workflow-repository
  - https://github.com/ufs-community/ufs-srweather-app/wiki/Code-Management

Development requirements: UFS_UTILS

- https://github.com/NOAA-EMC/UFS_UTILS/
- Requirements for contributing code:
  - Requires an issue be opened prior to opening a pull request
  - All code changes must conform to NCO Implementation Standards
  - Requires regression testing on a number of platforms prior to merging
  - May require the contribution of unit tests
- Different utilities have different code managers; see the repository wiki for details

Unified Post-Processor (UPP)

- https://github.com/NOAA-EMC/UPP
- To make changes
  - Create an issue to describe the change that you will be providing
  - Open a pull request
  - Contact one of the code managers to conduct regression tests
  - Detailed instructions available on the wiki
    - https://github.com/NOAA-EMC/UPP/wiki/UPP-Code-Development

CCPP

- See Laurie's talk this afternoon!

References and further reading

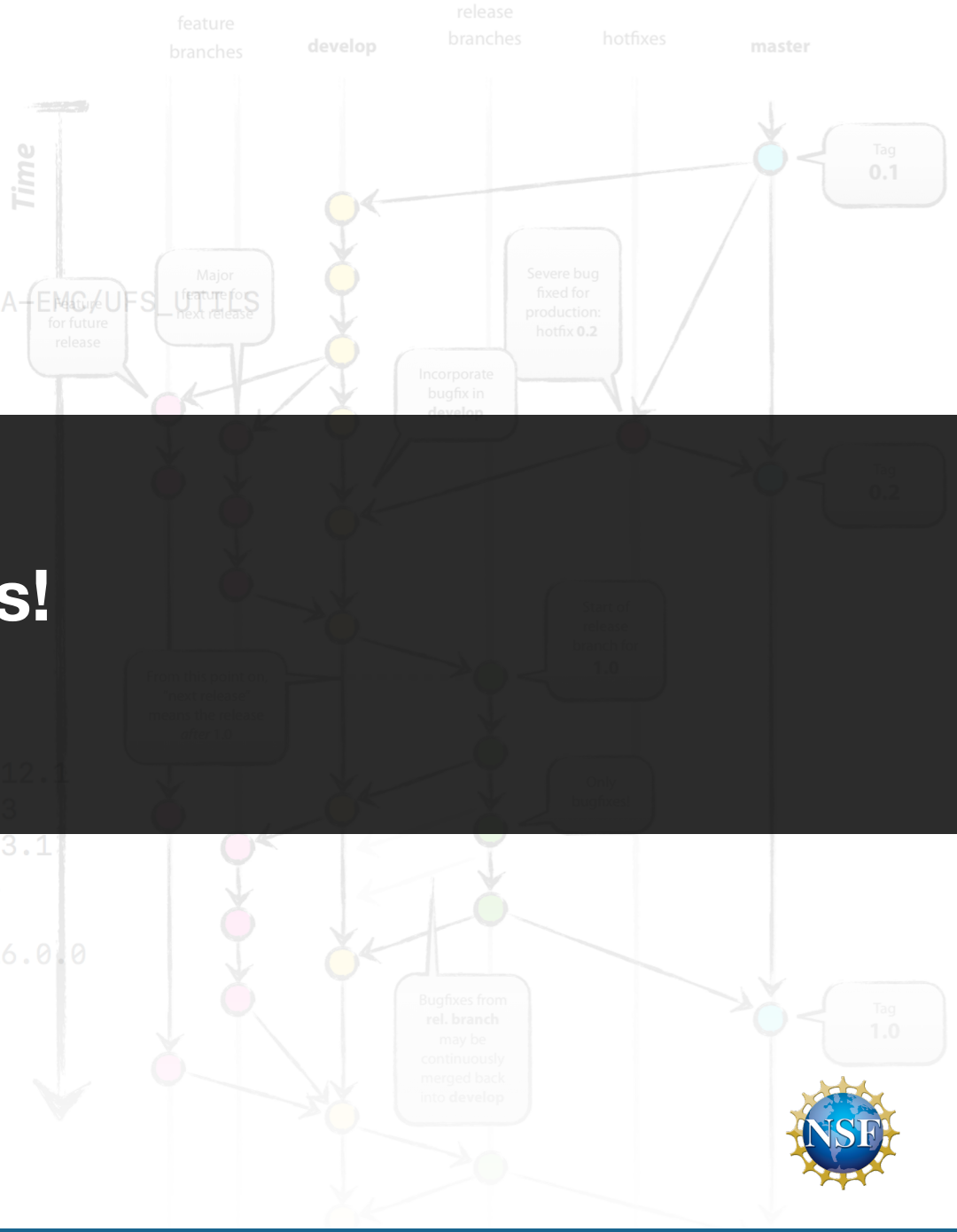
- Git documentation: https://git-scm.com/docs
  - Git visual cheat sheet: https://ndpsoftware.com/git-cheatsheet.html
- GitHub documentation: https://docs.github.com/en/free-pro-team@latest/github

  - Image credits: Randall Munroe, Simon Mutch, Vincent Driessen

Thank you for your attention!
Questions?

Messages consisting mainly of non-words are discouraged

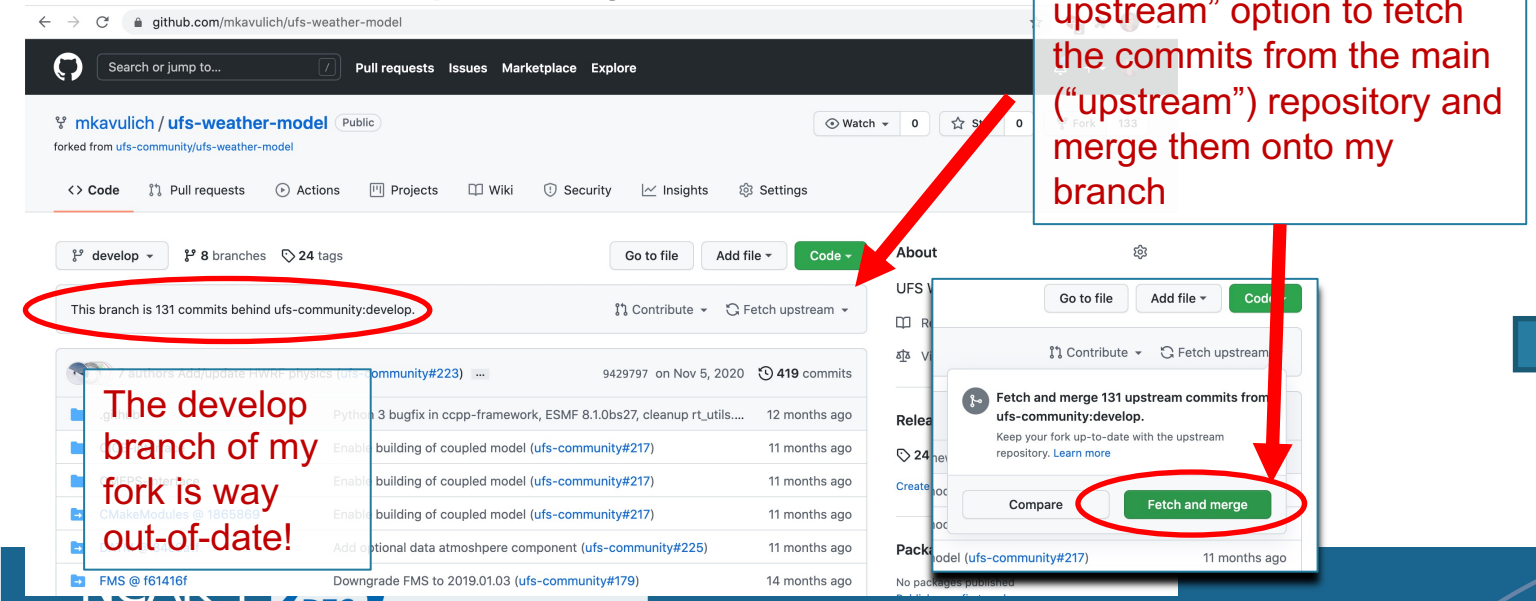
Bonus slides!



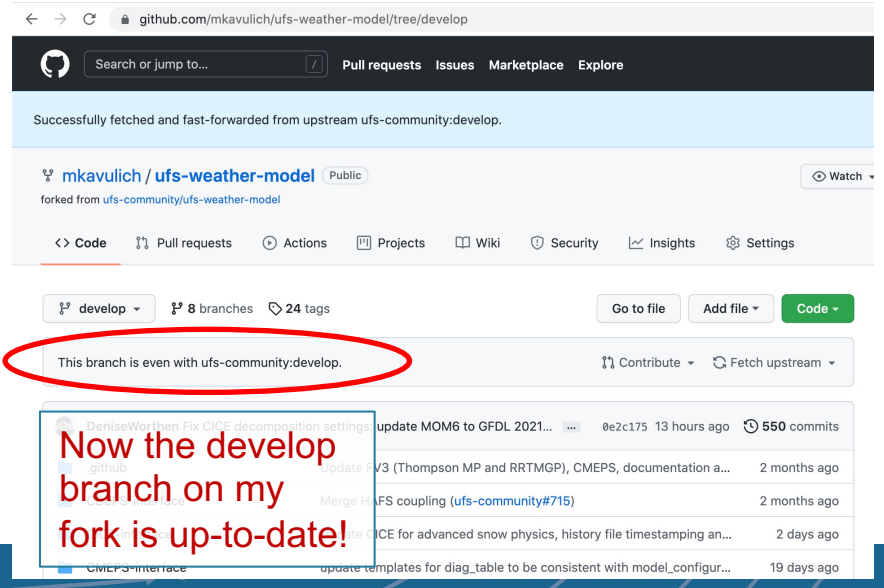
This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

GitHub Forks: keeping in sync

- Once a fork is created, it becomes its own independent repository
  - Changes in the main repository will not automatically appear in your fork!
  - You will need to keep your fork up-to-date as you continue development over time
- Luckily, keeping forks updated recently became easier: Github added a convenient "Fetch upstream" button that will update your branch with the latest changes from the main repository



Let's use the "Fetch upstream" option to fetch the commits from the main ("upstream") repository and merge them onto my branch



GitHub Forks: keeping in sync

- While forking has many advantages, it does require some additional effort in keeping your fork sync with the main repository
- This is handled through another bit of git functionality: remote repositories
- A “remote” is simply a link to another repository, either on disk or on the web
  - When you create a remote link to another repository, you can push to and pull from that repository
  - One remote is automatically created when you clone: the “origin” remote is the location where the current repository was cloned from
  - You can view remote repositories with the `git remote` command

```
>git remote -v
origin  https://github.com/ufs-community/ufs-mrweather-app.git (fetch)
origin  https://github.com/ufs-community/ufs-mrweather-app.git (push)
>
```

GitHub Forks: keeping in sync

- To keep your fork in sync with the main repository, clone fork locally, and create a remote named ”upstream” that will point to the main repository

```
> git remote add upstream https://github.com/ufs-community/ufs-weather-model
> git remote -v
origin  https://github.com/mkavulich/ufs-weather-model (fetch)
origin  https://github.com/mkavulich/ufs-weather-model (push)
upstream        https://github.com/ufs-community/ufs-weather-model (fetch)
upstream        https://github.com/ufs-community/ufs-weather-model (push)
```

- Next, use git fetch on the “upstream” repository, which will fetch the latest changes, including both new commits on each branch as well as new branches

```
> git fetch upstream
remote: Enumerating objects: 66, done.
remote: Counting objects: 100% (66/66), done.
remote: Total 96 (delta 66), reused 66 (delta 66), pack-reused 30
Unpacking objects: 100% (96/96), done.
From https://github.com/ufs-community/ufs-weather-model
 * [new branch]      develop                -> upstream/develop
 * [new branch]      production/GFS.v16     -> upstream/production/GFS.v16
 * [new branch]      production/GFS_v15     -> upstream/production/GFS_v15
 * [new branch]      production/HREF.v3     -> upstream/production/HREF.v3
 * [new branch]      production/HREF.v3beta -> upstream/production/HREF.v3beta
 * [new branch]      release/public-v1      -> upstream/release/public-v1
 * [new branch]      release/public-v2      -> upstream/release/public-v2
```

GitHub Forks: keeping in sync

- Next, perform a merge on the main branch for the repository. Including the `--ff` flag is recommended to avoid potential problems

```
> git merge --ff upstream/develop
Updating 08d06b6..9429797
Fast-forward
 FV3                                                        |   2 +-
 tests/RegressionTests_cheyenne.gnu.log                     | 266 ++++++++++++++----
 tests/RegressionTests_cheyenne.intel.log                   | 514 +++++++++++++++++++++--------------
 tests/RegressionTests_hera.gnu.log                         | 266 ++++++++++++++----
 tests/RegressionTests_hera.intel.log                       | 721 ++++++++++++++++++++++++++++---------------------
 tests/RegressionTests_orion.intel.log                      | 578 ++++++++++++++++++++++++---------------
 tests/RegressionTests_wcoss_cray.log                       | 502 ++++++++++++++++++++--------------
 tests/RegressionTests_wcoss_dell_p3.log                    | 626 ++++++++++++++++++++++++------------------
 tests/default_vars.sh                                      |   6 +
 tests/fv3_conf/ccpp_c96_HAFS_v0_hwrf_run.IN                |  85 ++++++
```

- Finally, push the synced branch to your fork on GitHub; assuming everything went well, the main branch on your fork is now synced with the main brain in the authoritative repository!

```
> git push origin
Username for 'https://github.com': mkavulich
Password for 'https://mkavulich@github.com':
Total 0 (delta 0), reused 0 (delta 0)
To https://github.com/mkavulich/ufs-weather-model
   08d06b6..9429797  develop -> develop
```

Bonus slides: changing NCEPLIBS code

- You will need to build your own version of NCEPLIBS, rather than using a pre-installed version
- Example: https://github.com/NOAA-EMC/NCEPLIBS-external/blob/release/public-v1/doc/README_cheyenne_intel.txt
  - It is not necessary to re-build the NCEPLIBS-external package
    - You will need to set the `-DCMAKE_INSTALL_PREFIX` flag when running cmake to install NCEPLIBS in a directory of your choosing
  - Before building a new copy of NCEPLIBS, you will need to point to the branch where you have made your modifications
    - The release/public-v1 branch of NCEPLIBS uses git submodules directly rather than manage_externals

Development requirements: FV3 dynamical core

- https://github.com/NOAA-EMC/GFDL_atmos_cubed_sphere/
  - Main development branch is dev/emc
  - Fork of https://github.com/NOAA-GFDL/GFDL_atmos_cubed_sphere
  - Development at GFDL takes place here: https://gitlab.gfdl.noaa.gov
  - No regression testing suite yet, but making changes to the dynamical core should not be taken lightly!
    - Thorough justification for the changes should be provided (referencing an existing Issue may suffice)
    - Testing should be done to ensure results will not change
    - If results *will* change, you should be prepared with scientific justification for the differences

Bonus slides: changing NCEPLIBS, chgres_cube, or post code

> `git clone -b develop git@github.com:NOAA-EMC/NCEPLIBS`

> `cd NCEPLIBS/`

> `vi .gitmodules`

- Edit the NCEPLIBS-post and/or UFS_UTILS url and branch to point to your fork and branch

> `git submodule update --init --recursive`

- Since we did not clone with the `--recursive` tag, this step is needed clone all of the submodules prior to building
- We now see our fork is being cloned, rather than the main repository

```
[submodule "NCEPLIBS-post"]
        path = NCEPLIBS-post
        url = https://github.com/NOAA-EMC/EMC_post
        branch = release/public-v1
[submodule "UFS_UTILS"]
        path = UFS_UTILS
        url = https://github.com/NOAA-EMC/UFS_UTILS
        branch = release/public-v1
".gitmodules" 76 lines --85%--
```

```
Submodule path 'NCEPLIBS-landsfcutil': checked out 'fa0368c8915d702
Submodule path 'NCEPLIBS-nemsio': checked out 'a1cb631cc2fcd43390b0
Submodule path 'NCEPLIBS-nemsiogfs': checked out 'c8a7bc2336c13e7f0
remote: Enumerating objects: 9, done.
remote: Counting objects: 100% (9/9), done.
remote: Total 28 (delta 9), reused 9 (delta 9), pack-reused 19
Unpacking objects: 100% (28/28), done.
From https://github.com/mkavulich/EMC_post
 * branch            bc6074f08d9b380bd9638ac269bf33c6ec52a641 -> FE
Submodule path 'NCEPLIBS-post': checked out 'bc6074f08d9b380bd9638a
Submodule 'cmake' (https://github.com/NOAA-EMC/CMakeModules) regist
Cloning into '/glade/scratch/kavulich/tmp/NCEPLIBS/NCEPLIBS-post/cm
```