



# MASUD RAHMAN

(a.k.a., Mohammad Masudur Rahman)

PhD Student

**Software Research Lab**

Department of Computer Science

University of Saskatchewan, Canada

Email: [masud.rahman@usask.ca](mailto:masud.rahman@usask.ca)

# RECAP

- **SDLC Phases (6)**
- *Requirement Gathering & Analysis*
  - UML
- *System Design*
  - Database
- *Implementation & Coding*
  - Source code documentation
  - **Version control systems**
- *Software Testing*
  - Unit & Integration testing
- *System Deployment*
  - Build tools
- *Software Maintenance & Evolution*
  - Clone detection, **Concept Location**



# Git & version control

# What is version control?

- A way to:
  - Keep a log of all the changes made to your scripts & data
  - Synchronize all the copies of your scripts in all your computers
  - Be able to go back to previous working versions

# Version control - example

- A project on which I am working on at the moment (Genotype Space)
- Started working on April 2012
- ~2,000 revisions

The screenshot shows a Bitbucket repository named "genotype\_space" with 2,000 commits. The commits are listed in a table with columns for Author, Commit, and Message. The commits are dated from April 2012 to the present. The messages describe various changes such as launching jobs, implementing network properties, and refactoring code.

Author	Commit	Message
Giovanni Ma...	95667ff	W19: launched job for chromosomes 1-21, window size 19
Giovanni Ma...	cdde9f5	SIMS: network properties for the two neutral + incomplete sweep simulations
Giovanni Ma...	fa58639	INCOMPLETE: launching background calculation for neutral and pos selection
Giovanni Ma...	7aae9ae	SAMPLE_SIZE: finished running reports for w11 and w3
Giovanni Ma...	969eeb2	BOXPLOT: generated boxplots again, including av_closeness and
Giovanni Ma...	ec0dded	SAMPLE_SIZE: WINDOW_SIZE: plotting the av_w_degree_inv instead of the
Giovanni Ma...	79453d5	BOXPLOT: refactored the way the --simplify option is applied
Giovanni Ma...	32dcda2c	PLOTS: WINDOW_SIZE: fixed title

# Example log

```
changeset: 1902:70f5d3dfecc44
tag:      tip
utente:   Giovanni Marco Dall'Olio <dalloliogm@gmail.com>
data:     Mon Mar 25 18:31:13 2013 +0100
file:    results/tables/multiple_windows/w11/background_by_gene_w11_s492.cs
descrizione:
WINDOW_SIZE: ASW: removing the old results based on sample_size 492.

changeset: 1901:473ced3a7262
utente:   Giovanni Marco Dall'Olio <dalloliogm@gmail.com>
data:     Mon Mar 25 18:16:48 2013 +0100
file:    .hgtags
descrizione:
TAG: Version of the repo used for the talk in Ferrara

changeset: 1900:5c7be5ed9d89
utente:   Giovanni Marco Dall'Olio <dalloliogm@gmail.com>
data:     Mon Mar 25 18:13:28 2013 +0100
file:    logs/test_logs/test_logs_latest.log
descrizione:
TEST_LOGS: on henikoff lab computer
```

-- MOST: \*stdin\* (1,1) 0%

Press 'Q' to quit, 'H' for help, and SPACE to scroll.

gioby@trellos: ~/workspace/genotype\_space | gioby@trellos: ~/workspace/genotype\_space | gioby@trellos: ~/workspace/genotype



This repository

Search

Pull requests Issues Marketplace Explore



masud-technope / CMPT370Public

Unwatch ▾

2

Star

1

Fork

0

Code

Issues 0

Pull requests 0

Projects 0

Wiki

Insights

Settings

Branch: master ▾

Commits on Oct 4, 2017

**Tutorial Oct 03 uploaded**

masud-technope committed 13 days ago



7dd2fe1



Commits on Sep 28, 2017

**UML items uploaded**

masud-technope committed 18 days ago



ab67dd2



Commits on Sep 21, 2017

**Tutorial materials for Sep 19 class**

masud-technope committed 26 days ago



f58a1a0



Commits on Sep 19, 2017

**Adding the insert queries**

masud-technope committed 28 days ago



3c30cae

**Merge branch 'master' of https://github.com/masud-technope/CMPT370Public**

...

masud-technope committed 28 days ago



5215962

**Added the toy database table creation SQL**

masud-technope committed 28 days ago



1e0d5ce



Commits on Sep 13, 2017

**Update README.txt**

masud-technope committed on Sep 13



3430d62

**Merge branch 'master' of https://github.com/masud-technope/CMPT370Public**

...

masud-technope committed on Sep 13



ba4d2eb



[This repository](#) [Search](#)[Pull requests](#) [Issues](#) [Marketplace](#) [Explore](#)[masud-technope / CMPT370Public](#)[Unwatch](#) [2](#)[Star](#) [1](#)[Fork](#) [0](#)[Code](#) [Issues 0](#) [Pull requests 0](#) [Projects 0](#) [Wiki](#) [Insights](#) [Settings](#)

Resources for CMPT 370 tutorials

[Edit](#)[uml](#) [database](#) [software-documentation](#) [doxygen](#) [astyle](#) [mysql-database](#) [git](#) [svn](#) [junit](#) [Manage topics](#)

13 commits

1 branch

0 releases

1 contributor

Branch: master ▾

[New pull request](#)[Create new file](#)[Upload files](#)[Find file](#)[Clone or download](#) ▾

	masud-technope Tutorial Oct 03 uploaded	Latest commit 7dd2fe1 13 days ago
	Tutorial-Oct03	Tutorial Oct 03 uploaded
	Tutorial-Sep12	Tutorial materials for Sep 19 class
	Tutorial-Sep19	Tutorial materials for Sep 19 class
	Tutorial-Sep26	UML items uploaded
	README.md	Update README.md
	README.md	

# CMPT370Public

Resources for CMPT 370 tutorials

# Version control – how does it work?

- Every time I make significant changes to a script, or get new results, I save a new “revision”
- Each revision is associated to a description, and a date

# Advantages of version control

- What was the version of my scripts when I generated the results included in the paper, 2 years ago?

# Advantages of version control

- What was the version of my scripts when I generated the results included in the paper, 2 years ago?
- I've accidentally removed a portion of my code. Can I go back?

# Advantages of version control

- What was the version of my scripts when I generated the results included in the paper, 2 years ago?
- I've accidentally removed a portion of my code. Can I go back?
- I've found a folder containing scripts and results. Are the scripts updated to the latest version?

# Git and Version control

- A popular software for version control is “git”
- Used for the Linux kernel
- Adopted by <http://github.com>, popular hosting website for open source projects

# How to get git?

- <http://git-scm.com/>
- Linux: apt-get install git



# How to get git?

- <http://git-scm.com/>
- Linux: apt-get install git
- Windows, Mac, Linux: you can install “tortoise-hg”, an
  - <http://tortoisehg.bitbucket.org/>



# Definition: the repository

- The repository is the folder where all the scripts (and eventually results) of a project are stored

# How to create a repository (on your computer)

- Connect to evopcompute
- Create a folder
- “cd” to it
- Type `git init`

# git init

- The command `git init` creates a new repository in the current folder
- All the files and subfolder will be included in the repo (although you will have to add them manually)

# git clone

- Alternatively, you can “clone” an existing repository
- `git clone <url-to-repository>`

# git add

- Let's add a new file to the repository
- First, let's create a file:
  - gedit README.txt
- Now, let's tell git to track it:
  - git add README.txt

# Effects of git add

- The command `git add` tells git that you want to keep track of all the changes of that file

basics

stages

branch

merge

remotes

advanced

working directory

staging area

repository

basics

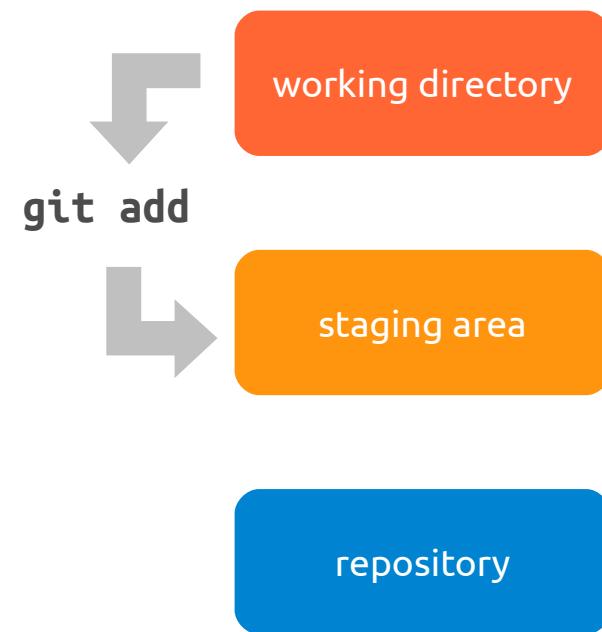
stages

branch

merge

remotes

advanced



basics

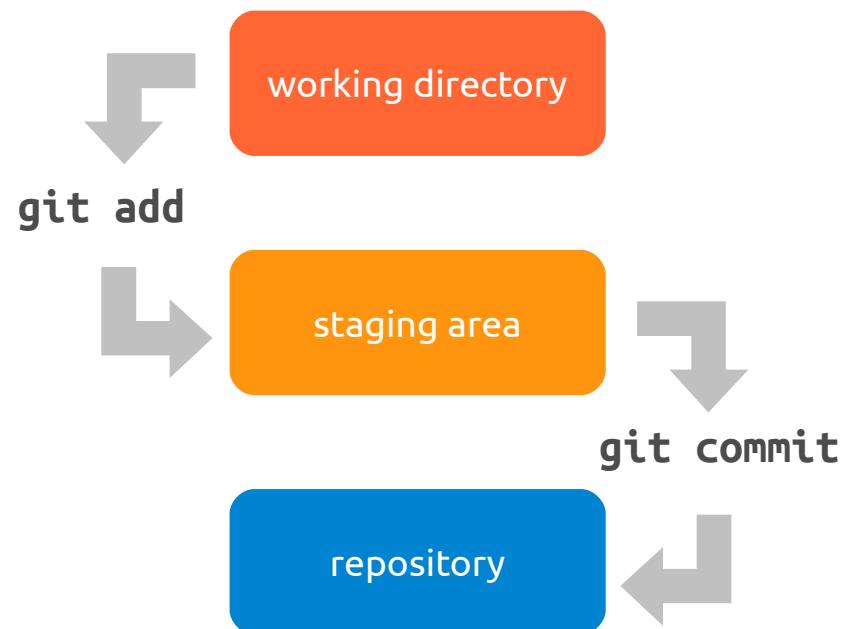
stages

branch

merge

remotes

advanced



basics

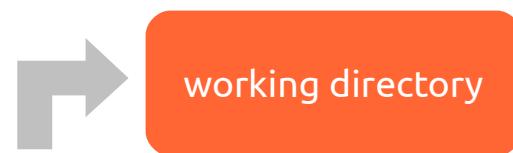
stages

branch

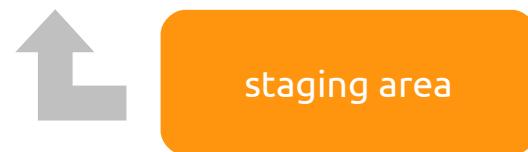
merge

remotes

advanced



`git reset HEAD <file>`



basics

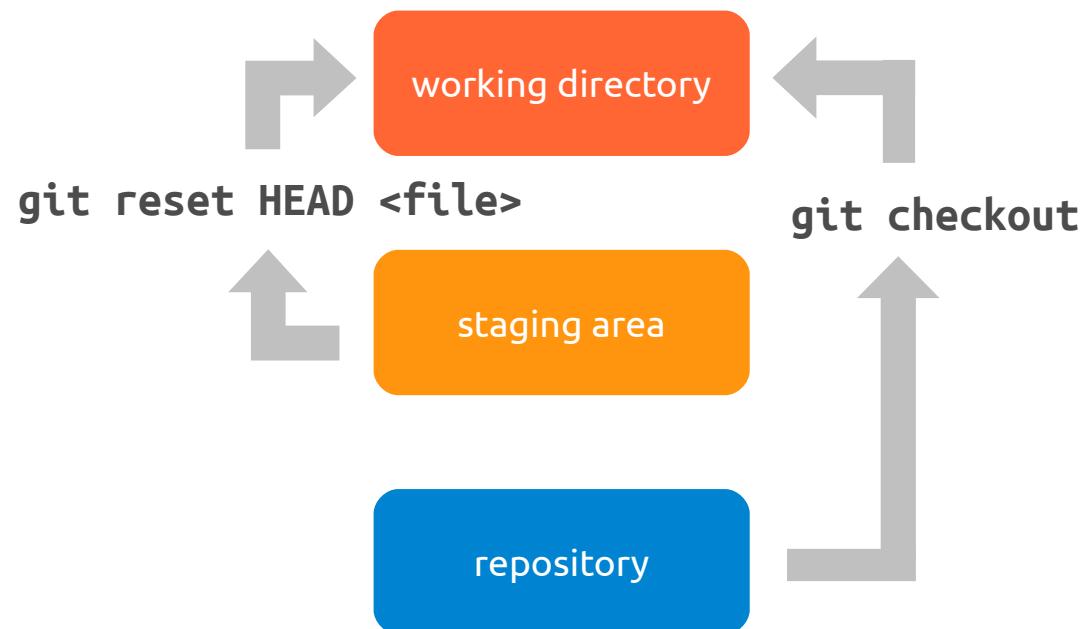
stages

branch

merge

remotes

advanced



# git commit

- Let's save our first revision:
  - `git commit README.txt -m "DOCS: added a readme"`

# git commit

- Let's save our first revision:
  - `git commit README.txt -m "DOCS: added a readme"`
- Explanation:
  - `git commit` → create a new commit
  - `README.txt` → the new commit refers only to the file `README.txt`
  - `-m "...."` → description of the commit

basics

stages

branch

merge

remotes

advanced

## Writing \*bad\* commit messages

"fix"

":(:("

"changes"

"it works!"

"final commit"

"Testing in progress ;-)"

"TODO: write meaningful commit message"

"Your commit is writing checks your merge can't cash"

basics

stages

branch

merge

remotes

advanced

## Writing good commit messages

Write commit messages as if you're giving *commands* to the codebase

“Add DAO interfaces for Entities”

“Implement basic version of AddressDAO”

“Fix bug in delete method of UserDAO”

“Move package security to at.ac.tuwien.service”

“Add generated fxml files for UI”

# git log

- Let's see our new commit:
  - `git log`

# Let's modify our file: git diff

- The command `git diff` tells you all the changes made since the last commit
- Let's modify our `README.txt` file, and see how the changes are tracked
  - `echo "new line after first commit" > README.txt`
  - `git diff`

# Let's make more commits

- echo “new line after first commit” > README.txt
- git diff
- git commit -m “DOCS: added a new line”
- echo “another new line after first commit” > README.txt
- git commit -m “DOCS: added another new line”
- git log

# git revert

- If we don't like some new changes, we can revert them:
  - `git log`
  - `git revert <id of the revision>`

# Syncronyzing a repository with github / bitbucket

- A common practice is to have multiple copies of your code, stored in different places
- Github and bitbucket are two free hosting services for code
- Bitbucket offers unlimited space hosting, so you can keep a backup of your code there

# How to synchronize your repo with github

- Create an account at <http://github.com>
- Follow the instructions on setting up a ssh key
  - <https://help.github.com/articles/set-up-git>
  - <https://help.github.com/articles/generating-ssh-keys>
- Create a new repository from the web interface

# git pull & push

- **git clone** → download a repository from github/bitbucket, for the first time
- **git pull** → download the latest changes from your repository (e.g. github/bitbucket), and update the code on your computer
- **git push** → upload all the latest changes to the remote repo

# STORE GIT CREDENTIALS

- \$ **git config credential.helper store**
- \$ git push http://example.com/repo.git  
    Username: <type your username>  
    Password: <type your password>
- [several days later]
- \$ git push http://example.com/repo.git [your credentials are used automatically]



# Git pull & push – best practices

- Always download the latest changes (`git pull`) before uploading new ones (`git push`)
- Otherwise, you will have some version conflicts
- Never force a `git push`

# How to collaborate on github/bitbucket

- Let's imagine we want to contribute to biopython:

The screenshot shows the GitHub repository page for 'biopython / biopython'. The top navigation bar includes 'Search or type a command', 'Explore', 'Gist', 'Blog', and 'Help' buttons. The repository header shows it's 'PUBLIC' and has a profile picture for 'dalolloiogn'. It features a 'Watch' button (345), an 'Unstar' button, a 'Fork' button (173), and a 'Code' tab which is currently selected. Below the tabs are buttons for 'ZIP', 'HTTP', 'SSH', and 'Git Read-Only', along with the URL 'https://github.com/biopython/biopython.git'. A 'Read-Only access' link is also present. The main content area displays a list of recent commits from the 'master' branch, with a total of '1000+ commits'. Each commit is shown with the author's name (e.g., 'peterjc'), the date, and a brief description. For example, the first commit was made an hour ago and fixed BWA doctests.

Author	Date	Description
peterjc	an hour ago	Forgot the Python 2.5 with_statement import
peterjc	a day ago	Fixed BWA doctests [peterjc]
tiagoantao	2 months ago	BioSQL support in Jython [tiagoantao]
peterjc	21 hours ago	Use filenames with SeqIO.read in Tutorial [peterjc]
cbrueffer	4 months ago	Scripts: Some more PEP8 whitespace cleanup. [cbrueffer]
peterjc	an hour ago	Forgot the Python 2.5 with_statement import [peterjc]
peterjc	8 days ago	Add biopython.egg-info to git ignore list [peterjc]
peterjc	3 days ago	Run 2to3 separately from setup.py to try and appease TravisCI limits. [peterjc]
peterjc	22 days ago	Thank Andrea Rizzi for his contributions [peterjc]
peterjc	25 days ago	Mention mmcif changes in NEWS and DEPRECATED files [peterjc]

basics

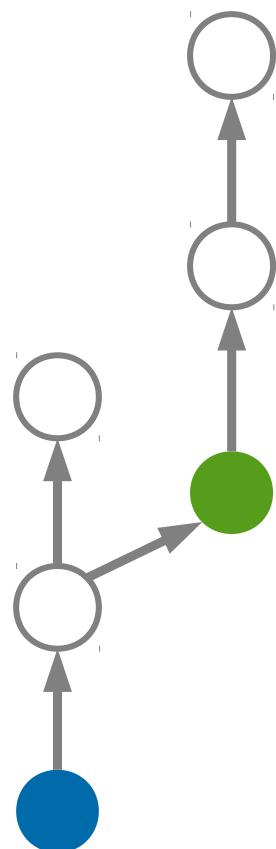
stages

branch

merge

remotes

advanced



basics

stages

branch

merge

remotes

advanced

## Manage branches

Create a new branch from the one you have currently checked out

```
$ git branch <branch>
```

Rename a branch

```
$ git branch -m <oldbranch> <newbranch>
```

Delete a branch

```
$ git branch -D <branch>
```

basics

stages

branch

merge

remotes

advanced

## List branches

```
$ git branch
  develop
* master
  new-feature
```

basics

stages

branch

merge

remotes

advanced

## List branches

```
$ git branch
  develop
* master
  new-feature
```

Branch you have  
currently  
checked out

basics

stages

branch

merge

remotes

advanced

## Change into a branch

```
$ git checkout develop  
Switched to branch 'develop'
```

basics

stages

branch

merge

remotes

advanced

# basics

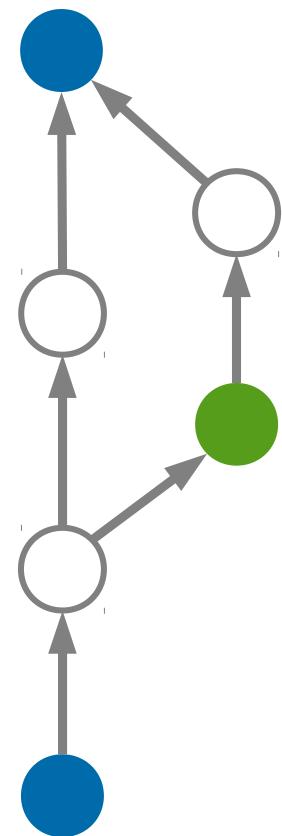
# stages

# branch

# merge

## remotes

## advanced



basics

stages

branch

merge

remotes

advanced

## Merge branches

Merge the specified branch into the current branch (the one you have checked out)

```
$ git checkout master  
$ git merge topic
```

Merges topic  
into master

basics

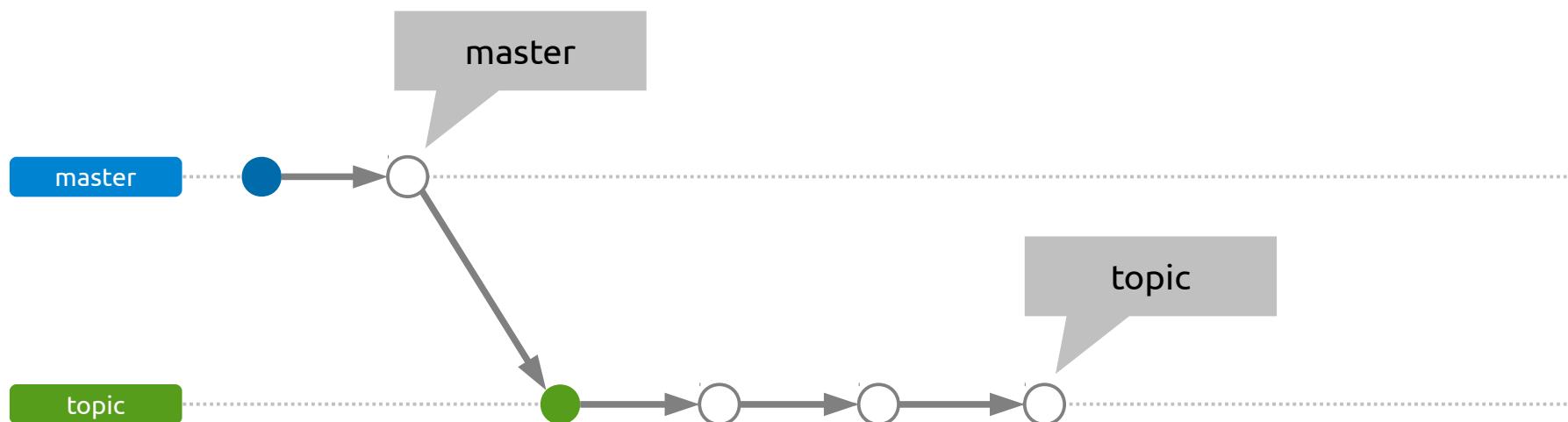
stages

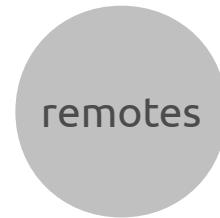
branch

merge

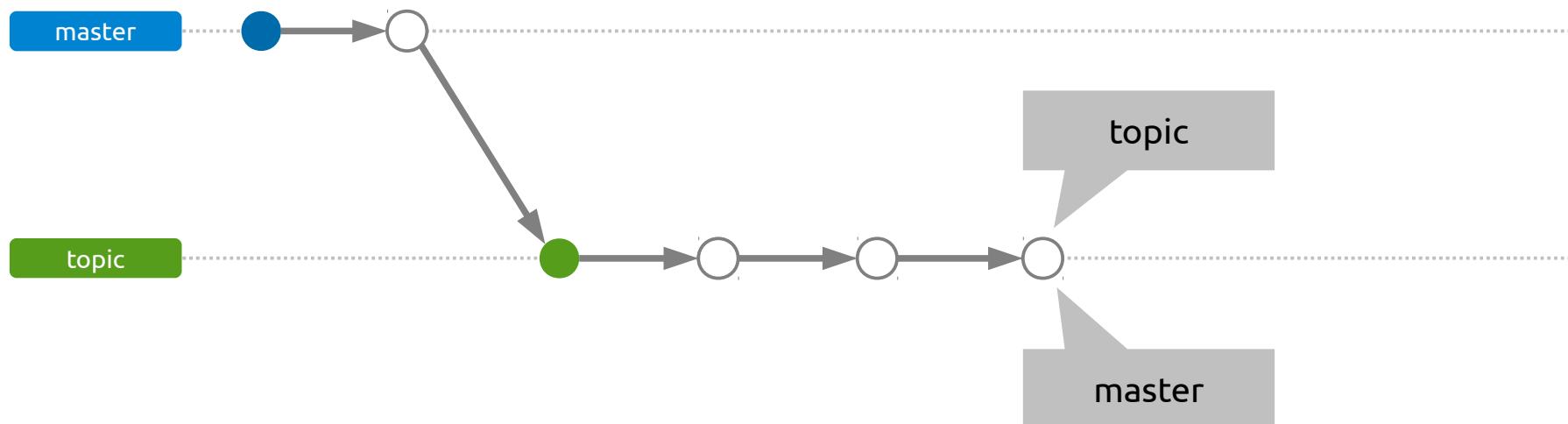
remotes

advanced





## Fast-forward



basics

stages

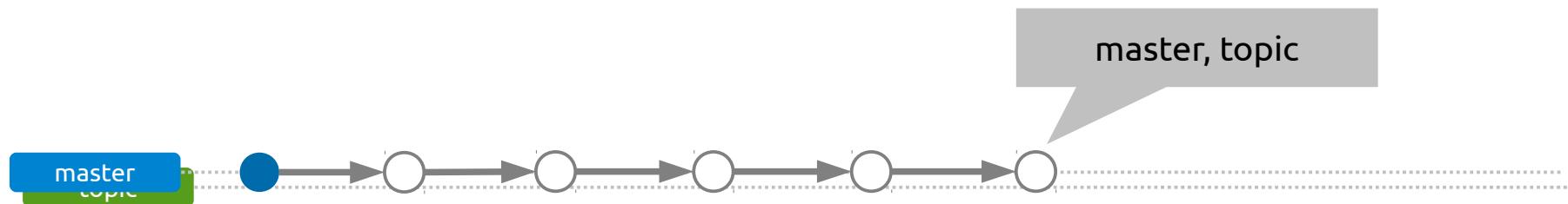
branch

merge

remotes

advanced

## Fast-forward



basics

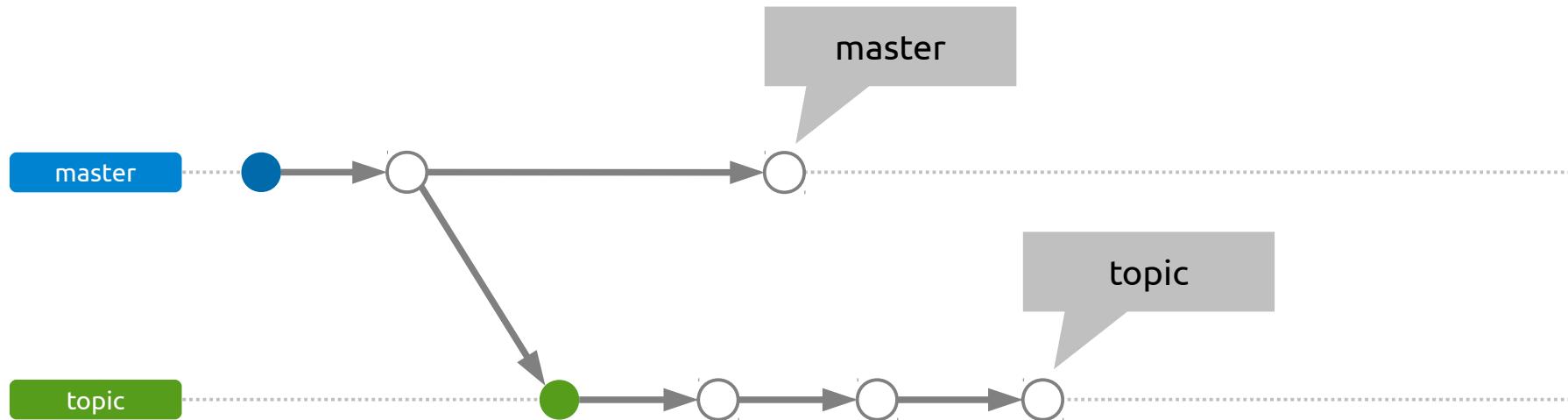
stages

branch

merge

remotes

advanced



basics

stages

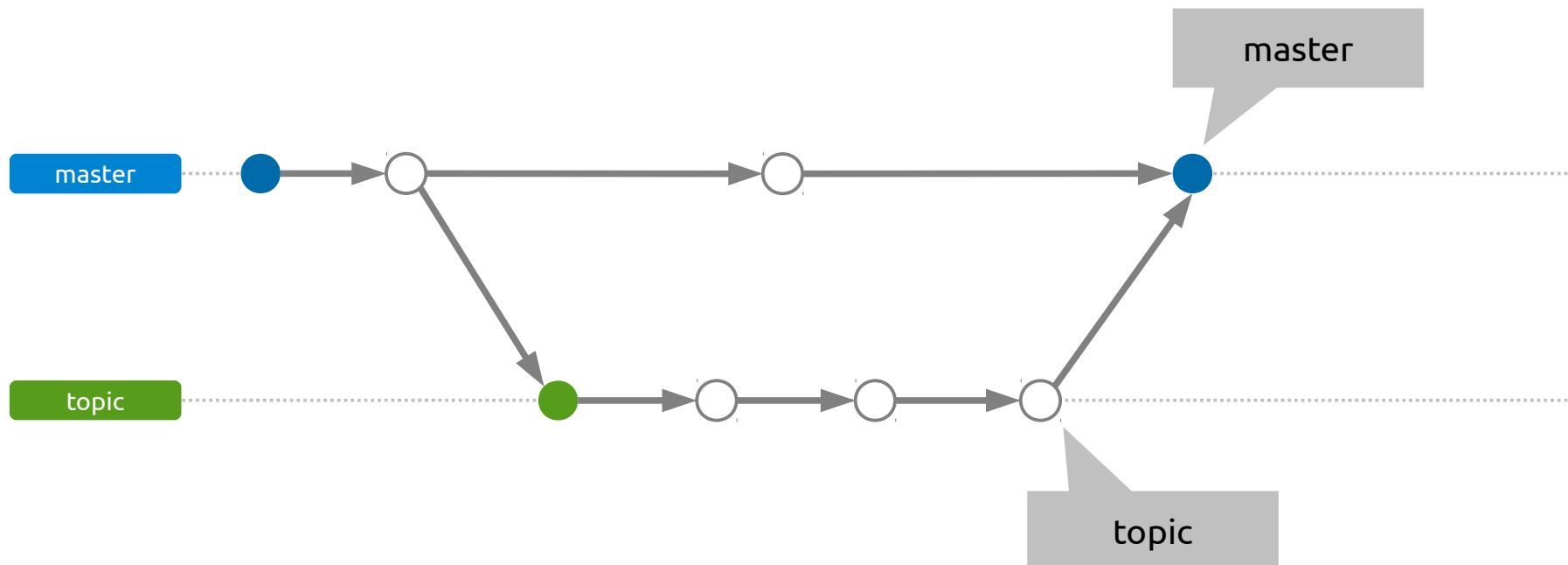
branch

merge

remotes

advanced

## Recursive three-way



basics

stages

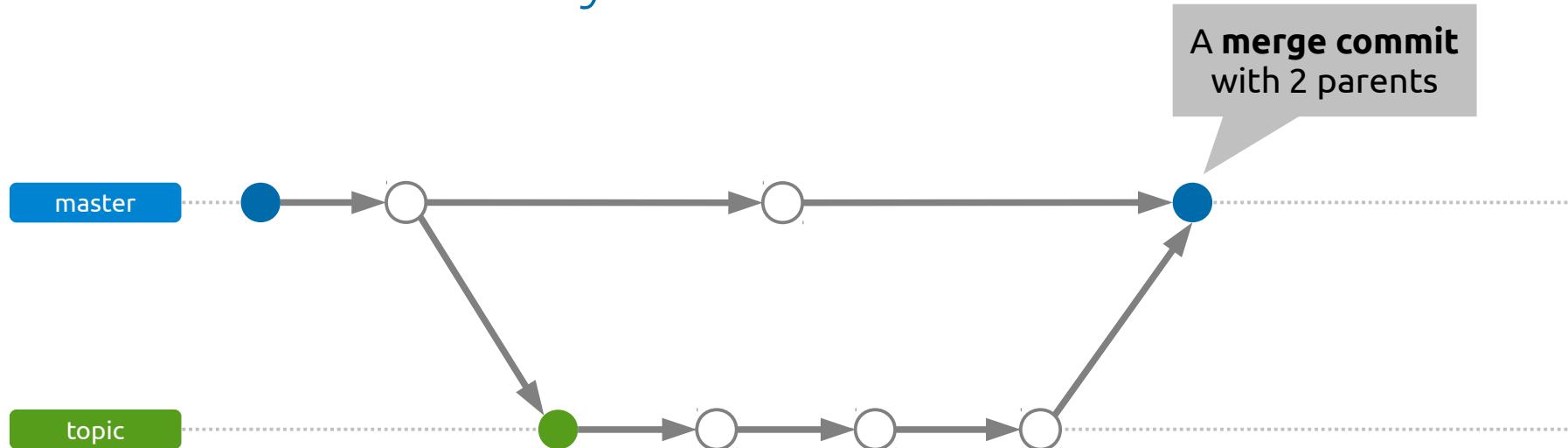
branch

merge

remotes

advanced

## Recursive three-way



basics

stages

branch

merge

remotes

advanced

# Conflicts

basics

stages

branch

merge

remotes

advanced

master

basics

stages

branch

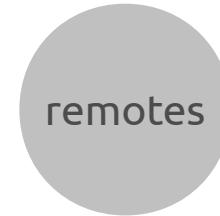
merge

remotes

advanced

master

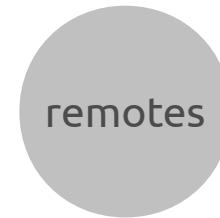
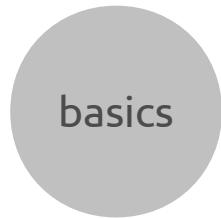
```
class Main {  
    public static void main(String... args) {  
    }  
}
```



master

topic

```
class Main {  
    public static void main(String... args) {  
    }  
}
```

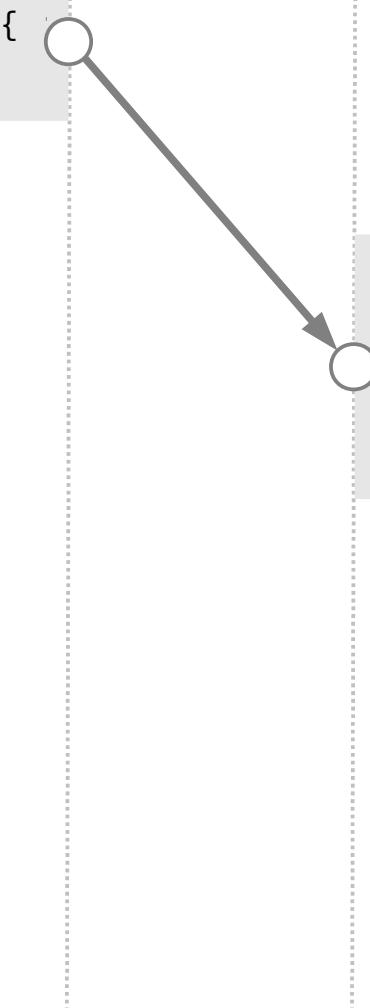


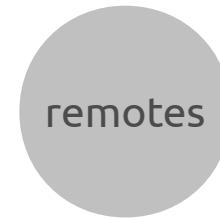
master

topic

```
class Main {  
    public static void main(String... args) {  
    }  
}
```

```
class Main {  
    static int status = 0;  
  
    public static void main(String... args) {  
        System.exit(status);  
    }  
}
```





master

topic

```
class Main {  
    public static void main(String... args) {  
    }  
}
```

```
class Main {  
    public static void main(String... args) {  
        System.exit(0);  
    }  
}
```

```
class Main {  
    static int status = 0;  
  
    public static void main(String... args) {  
        System.exit(status);  
    }  
}
```



master      topic

```
class Main {  
    public static void main(String... args) {  
    }  
}
```

```
class Main {  
    public static void main(String... args) {  
        System.exit(0);  
    }  
}
```

```
class Main {  
    static int status = 0;  
  
    public static void main(String... args) {  
        System.exit(status);  
    }  
}
```

```
class Main {  
    static int status = 0;  
  
    public static void main(String... args) {  
        <<<<< HEAD  
        System.exit(0);  
=====  
        System.exit(status);  
=====>>>>> topic  
    }  
}
```

# Conflict!

```
└ ~/C/G/n/git-exercise-1 ➔ git merge 97751003          22:04:41 10/12/2014
Auto-merging introduction.txt
CONFLICT (content): Merge conflict in introduction.txt
Automatic merge failed; fix conflicts and then commit the result.
```

# Fix conflict

- Find the conflict part and fix them:

<<<<< HEAD

Original contents

=====

New contents

>>>>> test\_merge

- Add updated files
- Commit

# Push failed

```
! [rejected]          master -> master (fetch first)
error: failed to push some refs to 'https://github.com/nccumath/git-exercise-1.git'
hint: Updates were rejected because the remote contains work that you do
hint: not have locally. This is usually caused by another repository pushing
hint: to the same ref. You may want to first integrate the remote changes
hint: (e.g., 'git pull ...') before pushing again.
hint: See the 'Note about fast-forwards' in 'git push --help' for details.
```

Try: git pull

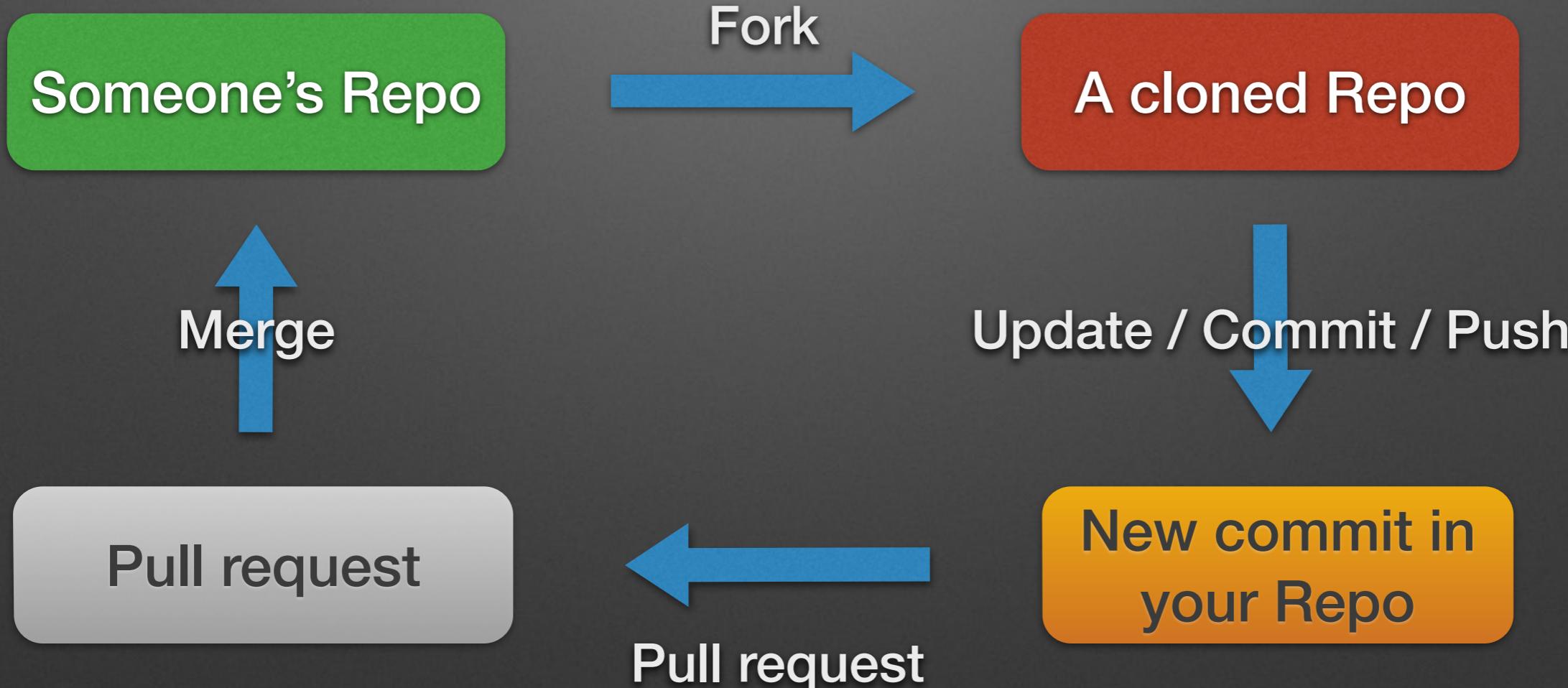
# git log --graph

```
* commit 3582729a8adf89170fe05cd53ddbe537222062ec
|\ Merge: 6aeb606 100619a
| | Author: Jim Yeh <lemonlatte@gmail.com>
| | Date: Mon Oct 6 00:59:16 2014 +0800
| |
| |     Merge branch 'test_merge'
| |
| |     Conflicts:
| |         introduction.txt
| |
| * commit 100619a3037fa62ac9a68cd83b915ff07d8f547c
| | Author: Jim Yeh <lemonlatte@gmail.com>
| | Date: Mon Oct 6 00:54:08 2014 +0800
| |
| |     test merge
| |
| * commit 6aeb606139d07dc7d781194fd3e71449ca78fae0
|/| Author: Jim Yeh <lemonlatte@gmail.com>
| | Date: Sun Oct 5 22:41:53 2014 +0800
| |
| |     Add a sample text file.
```

# How to collaborate on github/bitbucket

- To contribute to a open source repository on github/bitbucket, we will have to:
  - “fork” the repository using the web interface
  - Make changes to our forked repository
  - Send a “pull request” to the original repository once ready

# Fork / Pull Request



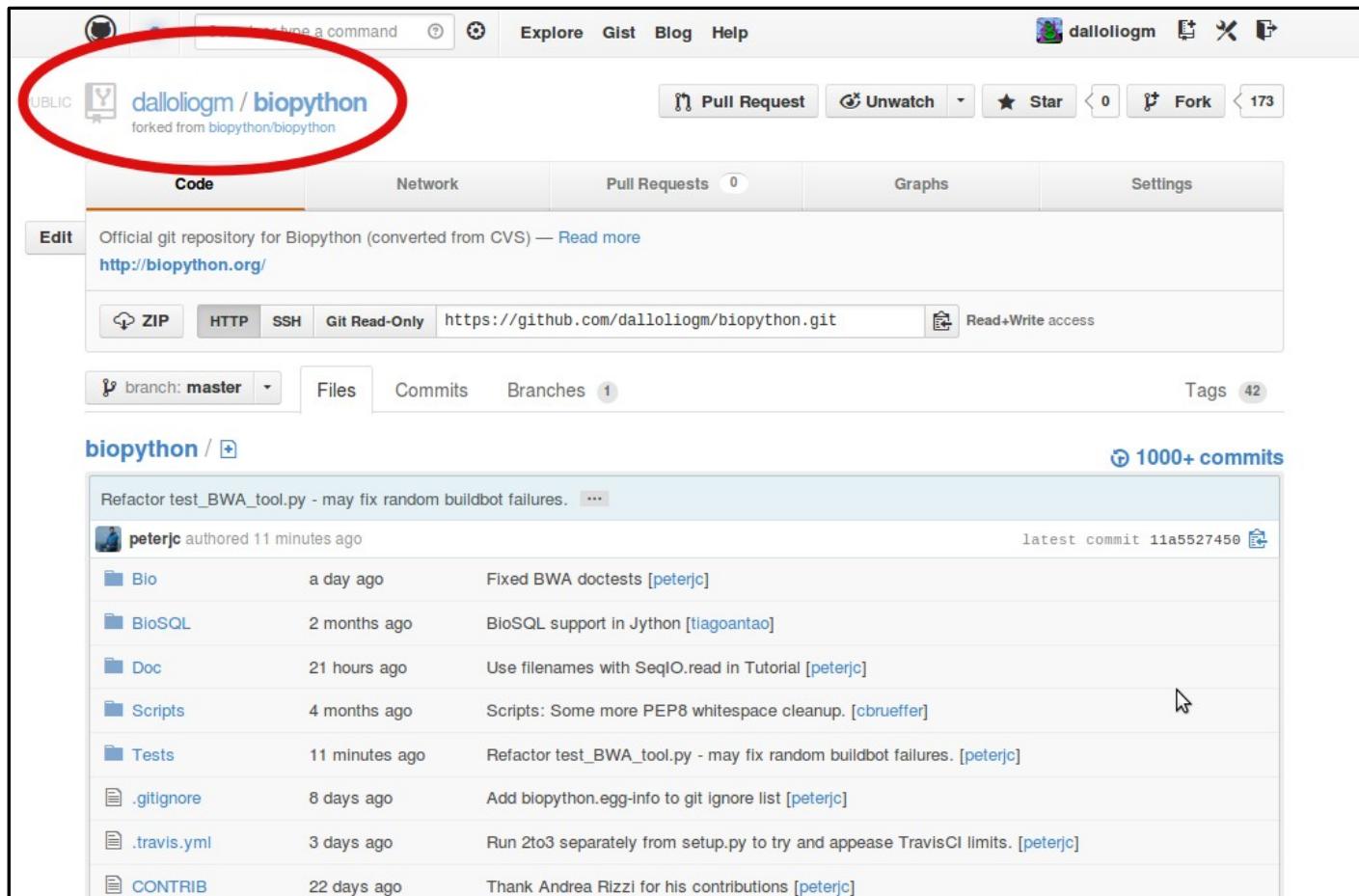
# Fork repository

The screenshot shows a GitHub repository page for 'biopython / biopython'. At the top right, there is a 'Fork' button, which is highlighted with a large red circle. Below the header, there are tabs for 'Code', 'Network', and 'Pull Requests'. The 'Code' tab is selected. Underneath, there is a brief description of the repository: 'Official git repository for Biopython (converted from CVS) — Read more http://biopython.org/'. Below this, there are download options: 'ZIP', 'HTTP', 'SSH', and 'Git Read-Only', along with the URL 'https://github.com/biopython/biopython.git'. A 'Read-Only access' button is also present. At the bottom of the main content area, there is a summary of recent commits by user 'peterjc':

File	Time Ago	Commit Message
Bio	a day ago	Fixed BWA doctests [peterjc]
BioSQL	2 months ago	BioSQL support in Jython [tiagoantao]
Doc	21 hours ago	Use filenames with SeqIO.read in Tutorial [peterjc]
Scripts	4 months ago	Scripts: Some more PEP8 whitespace cleanup. [cbrueffer]
Tests	an hour ago	Forgot the Python 2.5 with_statement import [peterjc]
.gitignore	8 days ago	Add biopython.egg-info to git ignore list [peterjc]
.travis.yml	3 days ago	Run 2to3 separately from setup.py to try and appease TravisCI limits. [peterjc]
CONTRIB	22 days ago	Thank Andrea Rizzi for his contributions [peterjc]

On the right side of the commit list, it says '1000+ commits'.

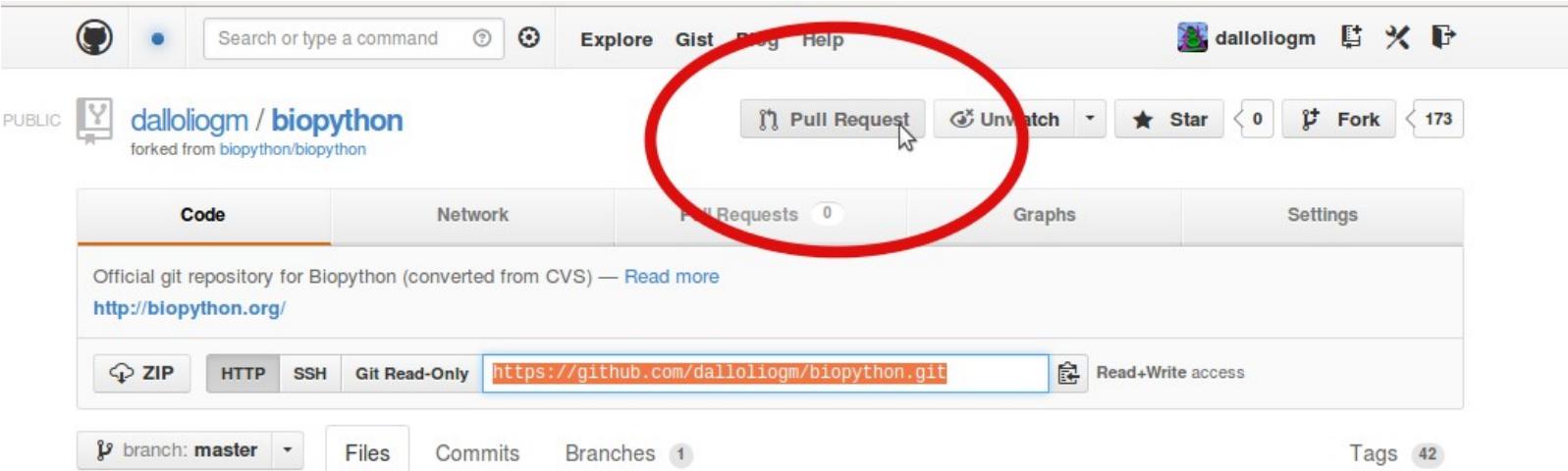
# Forked Repository



The screenshot shows a GitHub repository page for `dalloliogm / biopython`. A red circle highlights the text "forked from biopython/biopython" located just below the repository name. The page includes standard GitHub navigation like Explore, Gist, Blog, Help, and a user profile for `dalloliogm`. It also shows basic repository statistics: 0 pull requests, 173 forks, and 1000+ commits. The main content area displays a list of recent commits by various contributors, such as `peterjc`, `tiagoantao`, and `cbrueffer`.

File	Author	Date	Commit Message
Bio	peterjc	a day ago	Fixed BWA doctests [peterjc]
BioSQL		2 months ago	BioSQL support in Jython [tiagoantao]
Doc		21 hours ago	Use filenames with SeqIO.read in Tutorial [peterjc]
Scripts		4 months ago	Scripts: Some more PEP8 whitespace cleanup. [cbrueffer]
Tests	peterjc	11 minutes ago	Refactor test_BWA_tool.py - may fix random buildbot failures. [peterjc]
.gitignore		8 days ago	Add biopython.egg-info to git ignore list [peterjc]
.travis.yml		3 days ago	Run 2to3 separately from setup.py to try and appease TravisCI limits. [peterjc]
CONTRIB		22 days ago	Thank Andrea Rizzi for his contributions [peterjc]

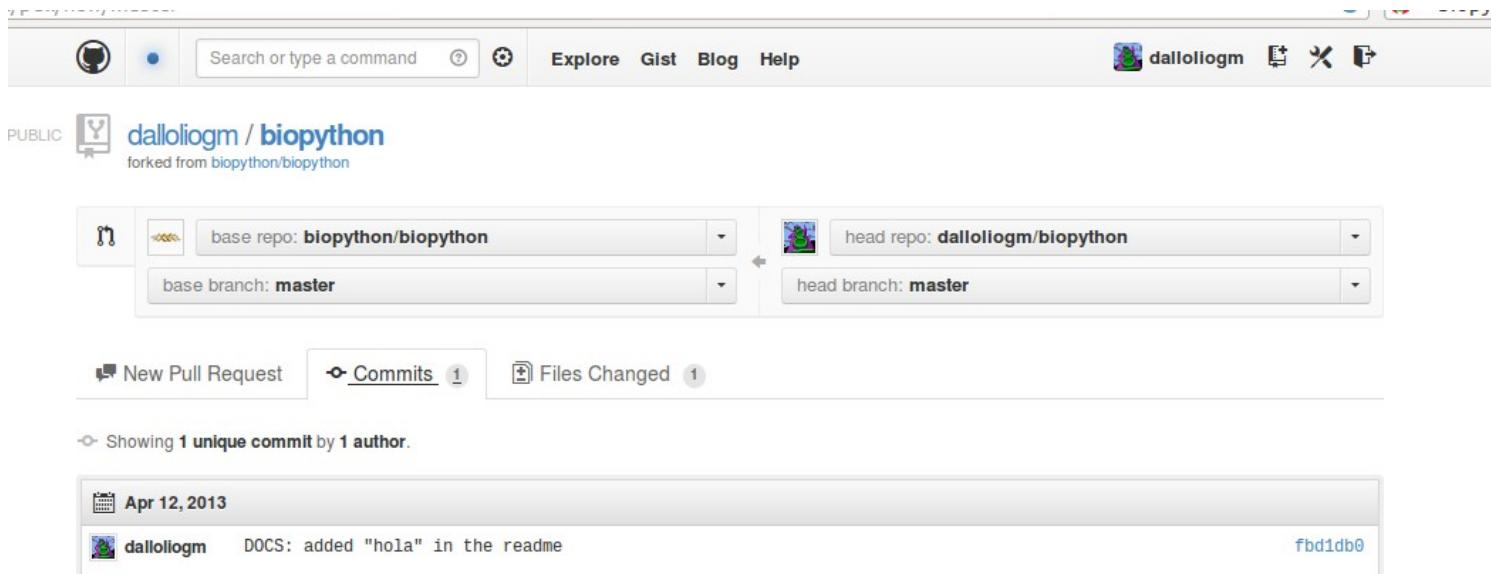
# Pull Request



A screenshot of a GitHub repository page for `dalloiogm / biopython`. The page shows the repository's main interface with tabs for Code, Network, Pull Requests (which is highlighted with a red circle), Graphs, and Settings. Below the tabs, there is a brief description of the repository and its URL (`http://biopython.org/`). The Pull Requests section shows 0 open pull requests. At the bottom, there are links for ZIP, HTTP, SSH, and Git Read-Only, along with a copy icon and a 'Read+Write access' button. The commit history section shows 1000+ commits by user `peterjc`, with the latest commit being `11a5527450`.

File	Time Ago	Commit Message
Bio	a day ago	Fixed BWA doctests [peterjc]
BioSQL	2 months ago	BioSQL support in Jython [tiagoantao]
Doc	21 hours ago	Use filenames with SeqIO.read in Tutorial [peterjc]
Scripts	4 months ago	Scripts: Some more PEP8 whitespace cleanup. [cbrueffer]
Tests	18 minutes ago	Refactor test_BWA_tool.py - may fix random buildbot failures. [peterjc]
.gitignore	8 days ago	Add biopython.egg-info to git ignore list [peterjc]
.travis.yml	3 days ago	Run 2to3 separately from setup.py to try and appease TravisCI limits. [peterjc]
CONTRIB	22 days ago	Thank Andrea Rizzi for his contributions [peterjc]

# Pull Request



# SUMMARY

- Version control System (Git)
- Repository creation/cloning
- Commit, merge, pull, push
- Branching
- Pull requests.



## ACKNOWLEDGEMENTS

- Thomas Rausch
- Jim Yeh
- Oscar Nierstrasz
- <https://www.slideshare.net/>



# THANK YOU!! QUESTIONS?

Masud Rahman

[masud.rahman@usask.ca](mailto:masud.rahman@usask.ca)

