ETH zürich



Version control with Git

Best Practices in Programming

Piotr Kupczyk, Swen Vermeul, John Hennig July 5, 2023







What is a Version Control System (VCS)?

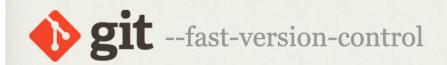
"... is a class of systems responsible for **managing changes** to **computer programs**, documents, large web sites, or other collections of information..."

[Wikipedia: Version control]





What is Git?



Git is a free and open source distributed version control system designed to handle everything from small to very large projects with speed and efficiency.

[git-scm.com]





Version Control System (VCS)

History of files and folders





Tracks source code changes over time

(Who? What? When? Why?)





Version: 1

Author: Paul Date: 1.01.2021

Comment: Initial version





2

Version: 2

Author: Kate Date: 12.01.2021

Comment: Add feature X

Version: 1

Author: Paul Date: 1.01.2021

Comment: Initial version





Version: 2

Author: Kate Date: 12.01.2021

Comment: Add feature X

Version: 1

Author: Paul Date: 1.01.2021

Comment: Initial version

Version: 3

Author: Kate Date: 13.01.2021

Comment: Refactor class C





Version: 2

Author: Kate Date: 12.01.2021

Comment: Add feature X

Version: 4
Author: Paul

Date: 21.01.2021

Comment: Remove unused code

Version: 1

Author: Paul Date: 1.01.2021

Comment: Initial version

Version: 3

Author: Kate Date: 13.01.2021

Comment: Refactor class C





Version: 2

Author: Kate
Date: 12.01.2021

Comment: Add feature X

Version: 4
Author: Paul

Date: 21.01.2021

Comment: Remove unused code

Version: 1

Author: Paul Date: 1.01.2021

Comment: Initial version

Version: 3

Author: Kate Date: 13.01.2021

Comment: Refactor class C

Version: 5

Author: Paul Date: 23.01.2021

Comment: Fix bug reported in issue #123





Good performance

Version: 2

Author: Kate Date: 12.01.2021 Comment: Add feature X

> Version: 4 Author: Paul

Date: 21.01.2021 Comment: Remove unused code Version: 1

Author: Paul Date: 1.01.2021

Comment: Initial version

Version: 3

Author: Kate Date: 13.01.2021

Comment: Refactor class C

Version: 5

Author: Paul Date: 23.01.2021

Comment: Fix bug reported in issue #123





Good performance

Version: 2

Author: Paul
Date: 1.01.2021
Comment: Initial version

Author: Kate Date: 12.01.2021

Comment: Add feature X

Author: Kate Date: 13.01.2021

Comment: Refactor class C

Version: 3

Version: 1

Version: 4
Author: Paul
Date: 21.01.2021

Comment: Remove unused code

Version: 5
Author: Paul

Date: 23.01.2021

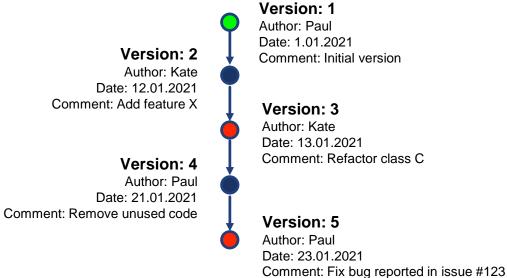
Comment: Fix bug reported in issue #123

Bad performance





Good performance

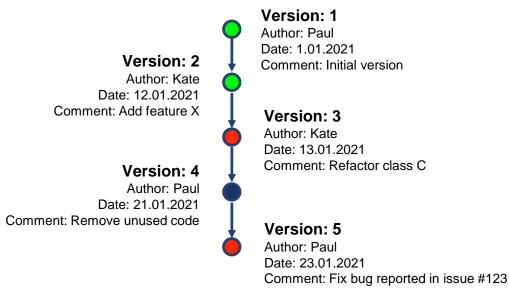


Bad performance





Good performance



Bad performance





Branches





Branches

How to work independently of others?





Version: 1

Author: Paul Date: 1.01.2021

Comment: Initial version





Version: 2

Author: Kate Date: 12.01.2021

Comment: Add feature X

Version: 1

Author: Paul Date: 1.01.2021

Comment: Initial version





Version: 2

Author: Kate Date: 12.01.2021

Comment: Add feature X

Version: 1

Author: Paul Date: 1.01.2021

Comment: Initial version

Version: 3

Author: Paul Date: 13.01.2021

Comment: Start feature Y





Version: 2

Author: Kate Date: 12.01.2021

Comment: Add feature X

Version: 4
Author: Kate

Date: 21.01.2021

Comment: How am I supposed to work?!

Version: 1

Author: Paul Date: 1.01.2021

Comment: Initial version

Version: 3

Author: Paul Date: 13.01.2021

Comment: Start feature Y





Version: 2

Author: Kate Date: 12.01.2021 Comment: Add feature X

Version: 4

Author: Kate Date: 21.01.2021

Comment: How am I supposed to work?!

Version: 1

Author: Paul Date: 1.01.2021

Comment: Initial version

Version: 3

Author: Paul Date: 13.01.2021

Comment: Start feature Y

Version: 5

Author: Paul Date: 23.01.2021

Comment: Still working on feature Y. Sorry Kate...





Version: 2

Author: Kate

Date: 12.01.2021

Comment: Add feature X

Version: 4
Author: Kate

Date: 21.01.2021

Comment: How am I supposed to work?!

Version: 6

Author: Paul Date: 24.01.2021

Comment: Finished feature Y.

Version: 1

Author: Paul Date: 1.01.2021

Comment: Initial version

Version: 3

Author: Paul Date: 13.01.2021

Comment: Start feature Y

Version: 5

Author: Paul Date: 23.01.2021

Comment: Still working on feature Y. Sorry Kate...





Version: 1

Author: Paul Date: 1.01.2021

Comment: Initial version





Version: 1

Author: Paul Date: 1.01.2021

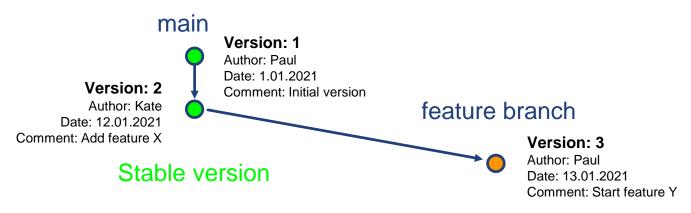
Comment: Initial version

Version: 2 Author: Kate

Date: 12.01.2021 Comment: Add feature X

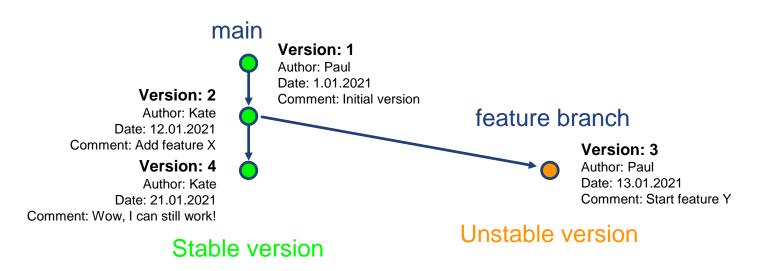






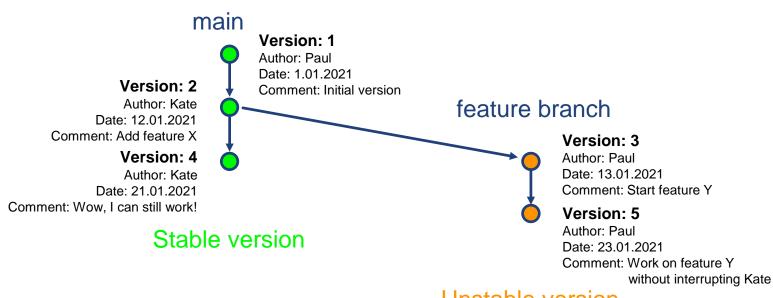






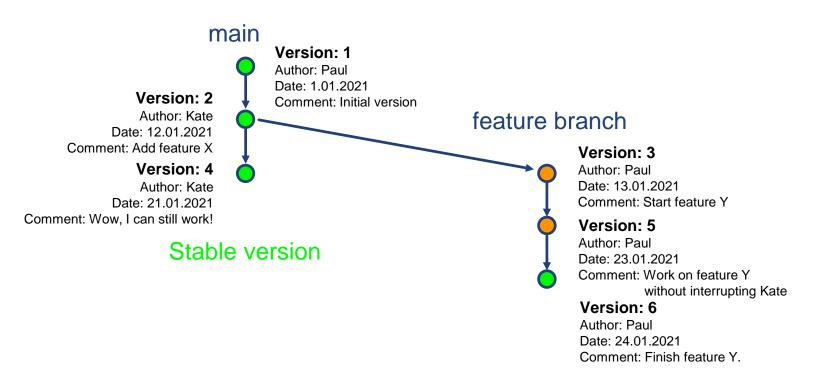






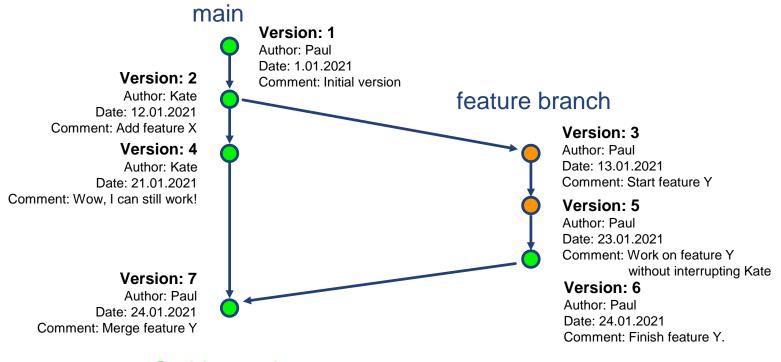








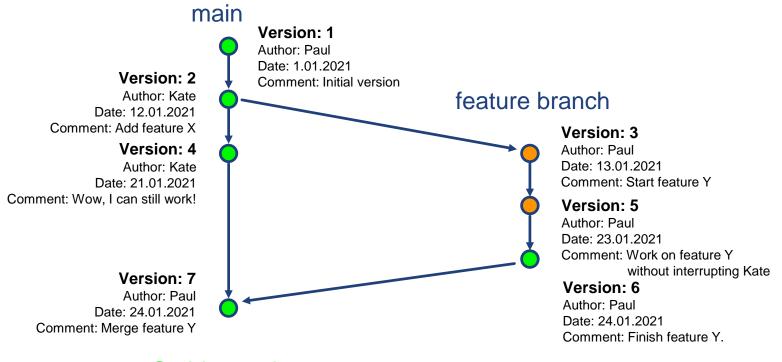




Stable version







Stable version







Git is a free and open source distributed version control system designed to handle everything from small to very large projects with speed and efficiency.

Collaboration

How to share changes with others?



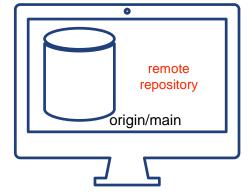








e.g. <u>GitHub.com</u>





Paul's machine

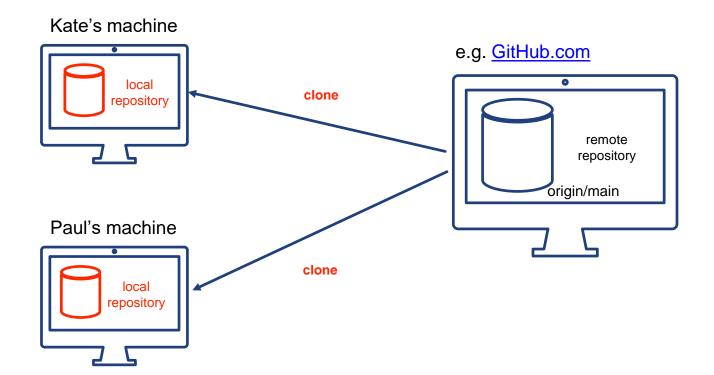






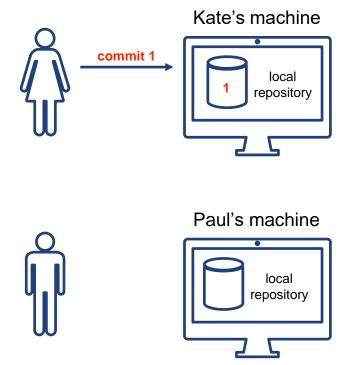


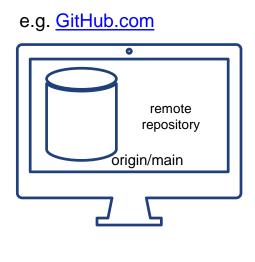






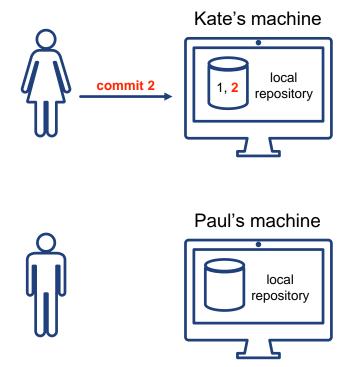


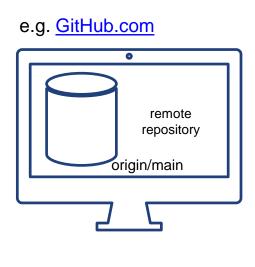






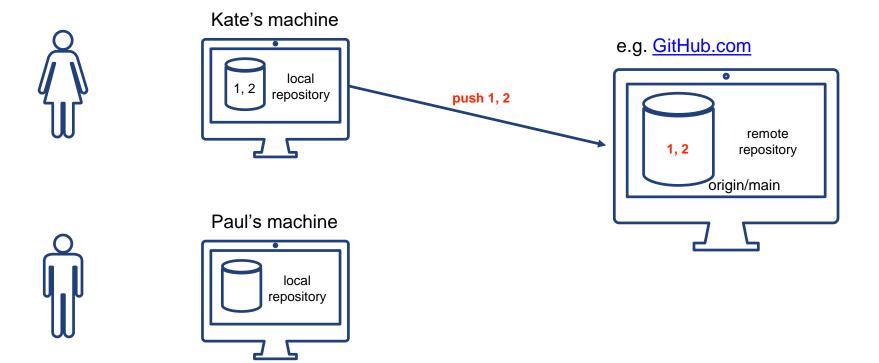






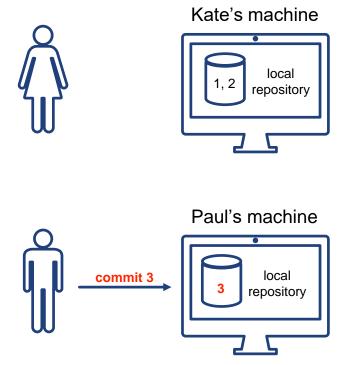


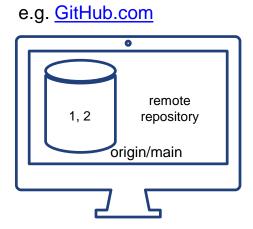






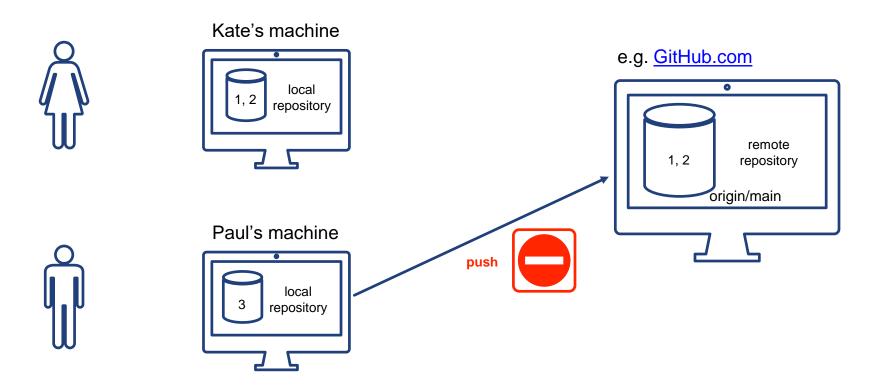






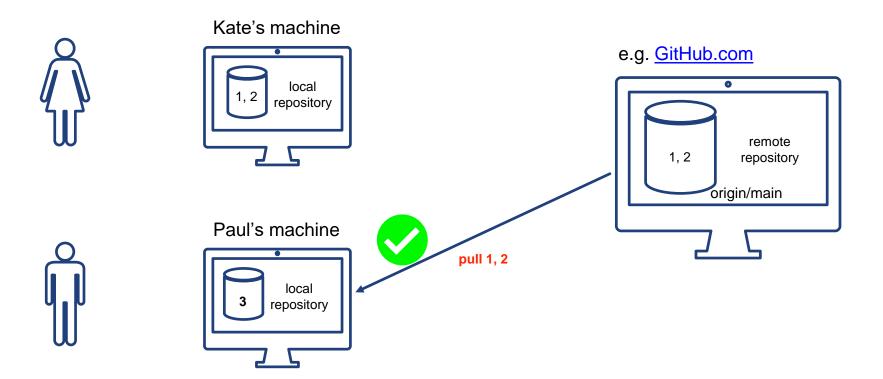








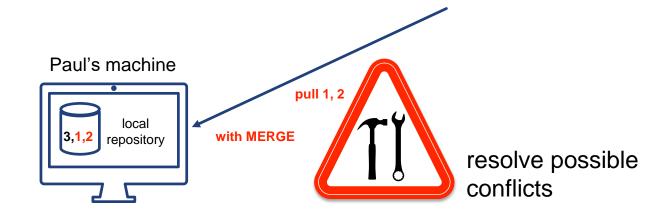








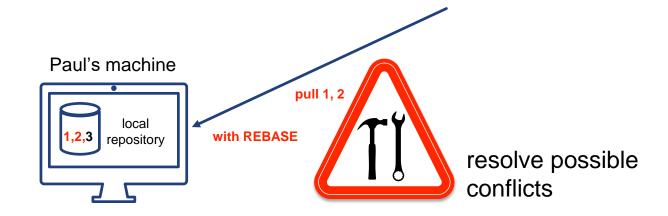






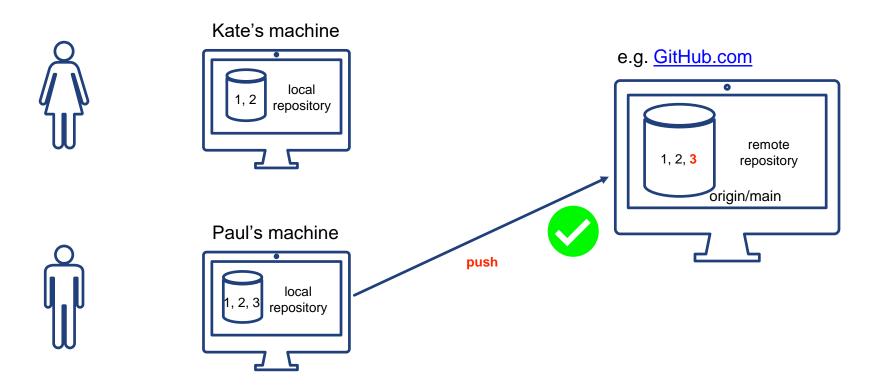






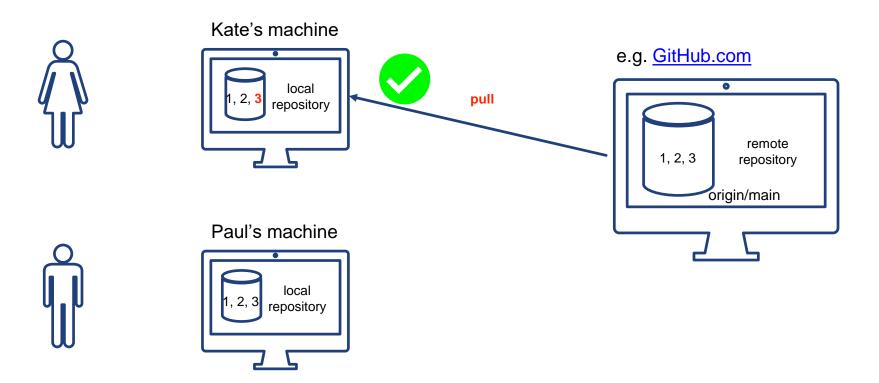
















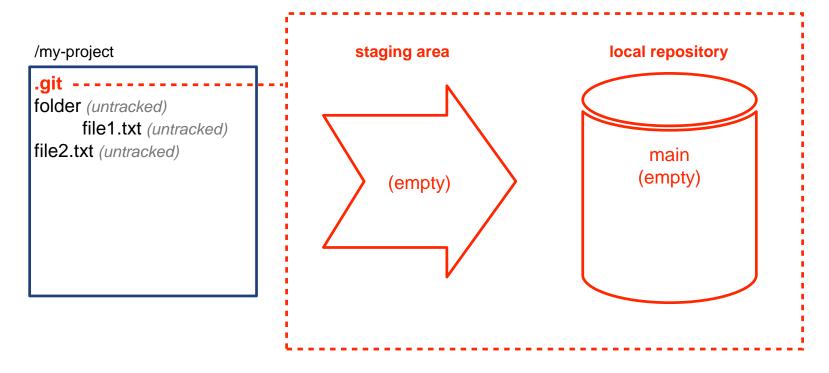
/my-project

folder file1.txt file2.txt





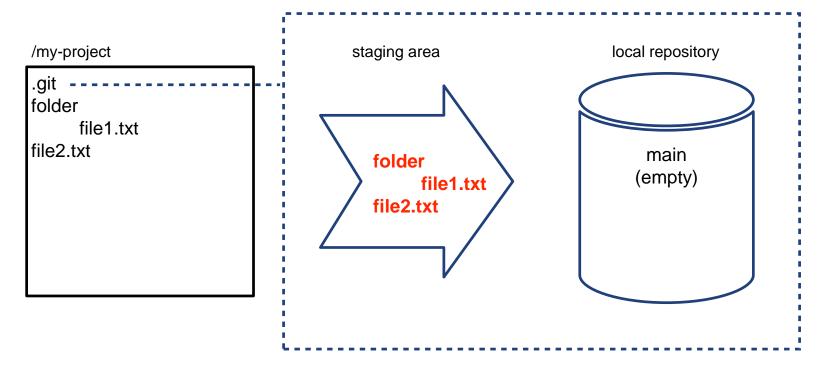
1. Initialize local Git repository: git init







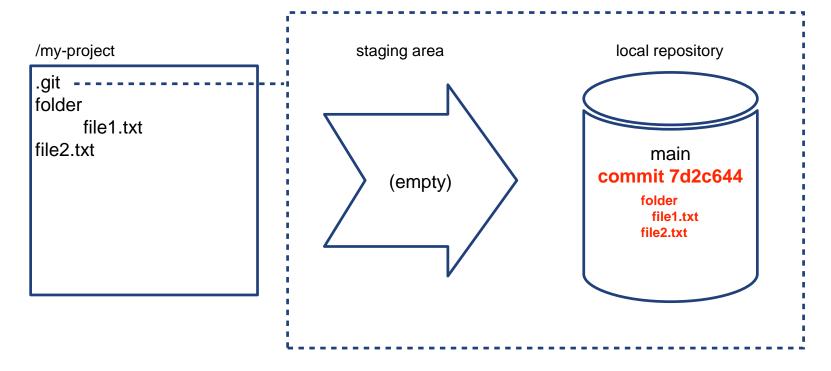
2. Tell Git to track all files: git add .







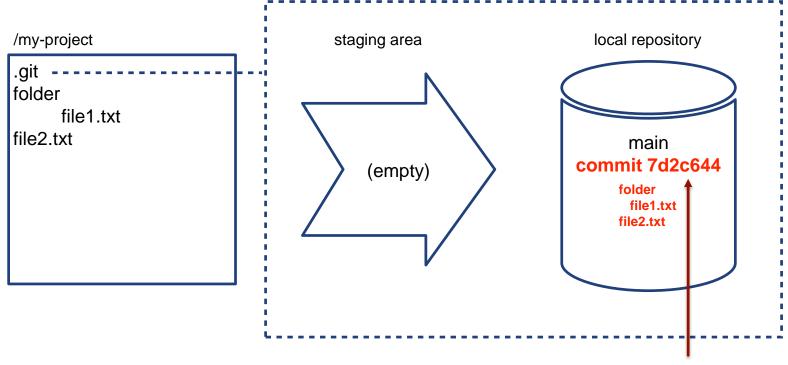
3. Commit staged files: git commit -m "New!"







3. Commit staged files: git commit -m "New!"

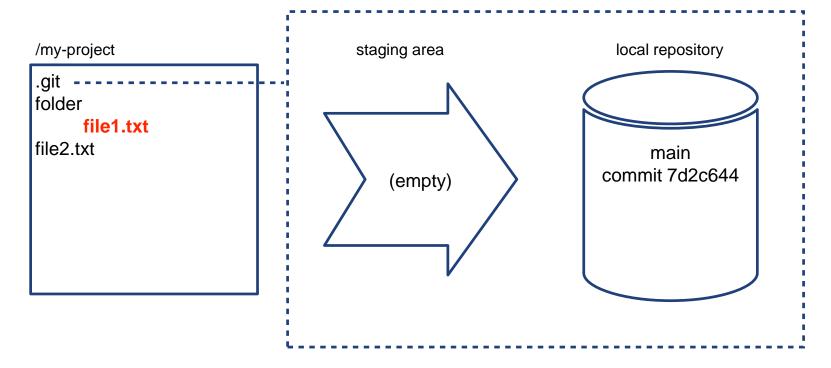


unique identifier (SHA-1 hash)





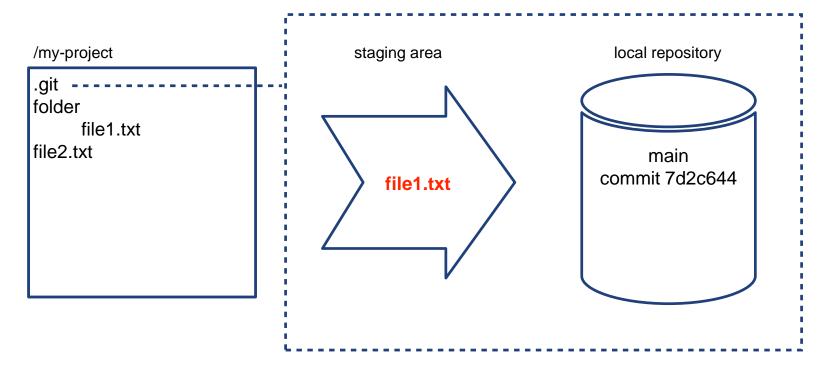
4. Keep changing project files







5. Stage changes of tracked files: git add --update

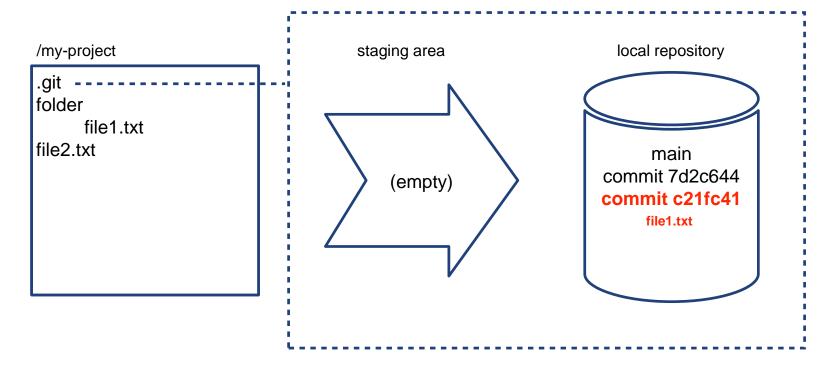


Or git add -u for short.





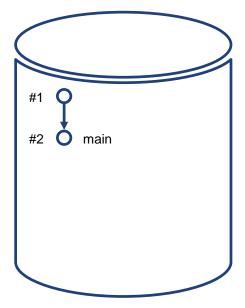
6. Commit staged files: git commit -m "Newer!"







remote repository (origin)

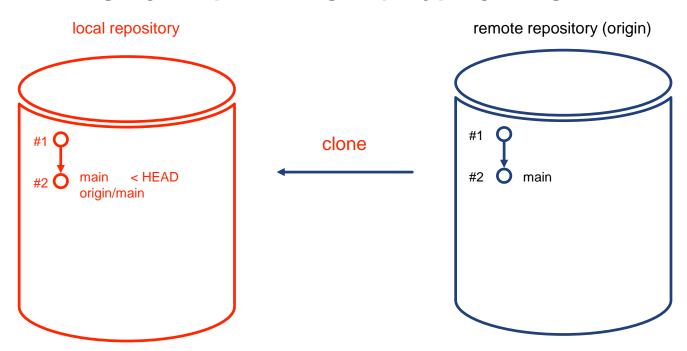






1. Clone remote repository:

git clone git@example.com:group/myproject.git

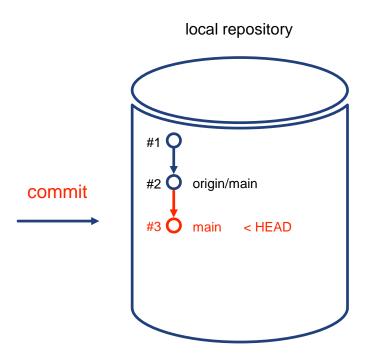


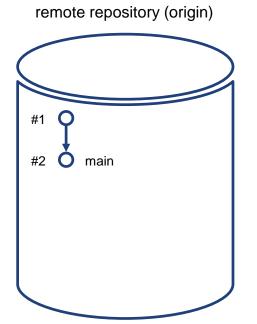
\$ git status
On branch main
Your branch is up to date with 'origin/main'.





2. Commit changes locally: git commit



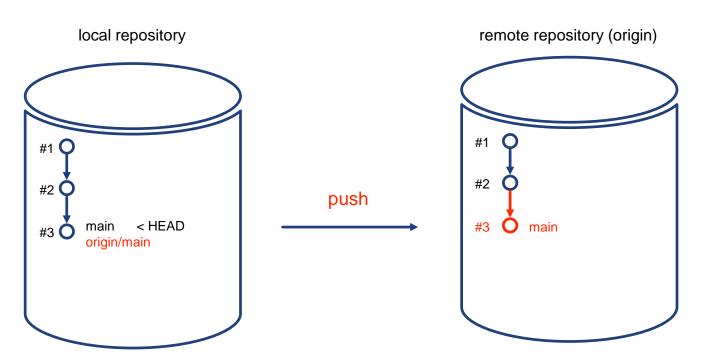


\$ git status
On branch main
Your branch is ahead of 'origin/main' by 1 commit.





3. Push changes to remote repository: git push

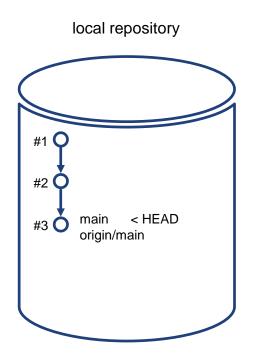


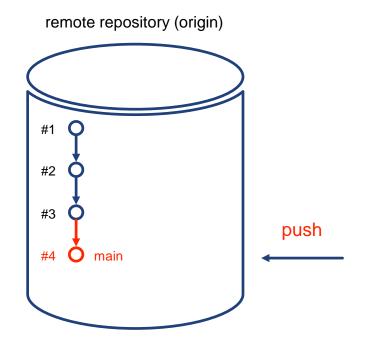
\$ git status
On branch main
Your branch is up to date with 'origin/main'.





4. Someone pushes changes to the remote repository

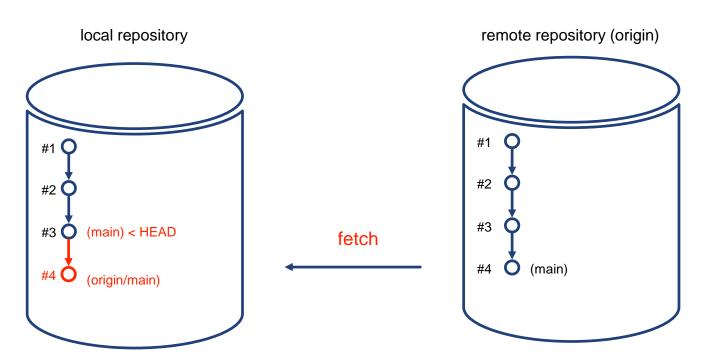








5. get latest from repo (without pull/merge): git fetch



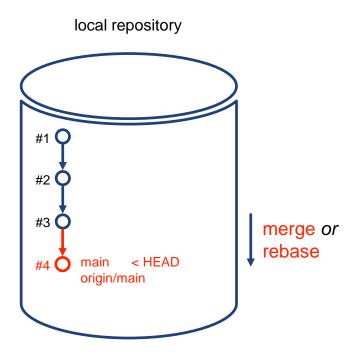
\$ git status

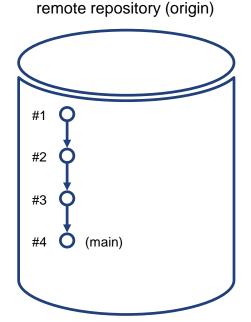
Your branch is behind 'origin/main' by 1 commit, and can be fast-forwarded.





6. Merge latest changes: git merge origin/main [--ff-only]





\$ git status
On branch main
Your branch is up to date with 'origin/main'.



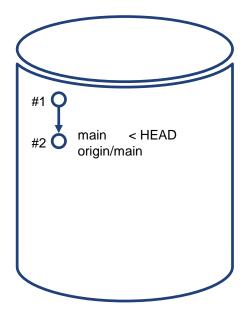


Working with branches



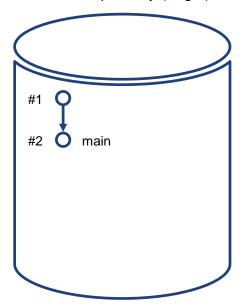






\$ git status
On branch main
Your branch is up to date with 'origin/main'.

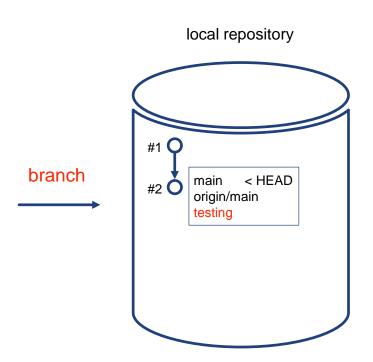
remote repository (origin)

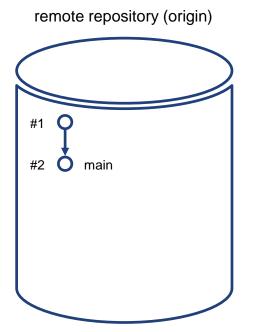






1. Create a branch: git branch testing

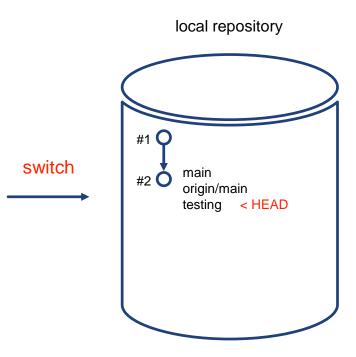




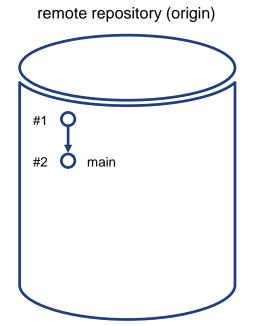




2. Switch to the branch: git switch testing



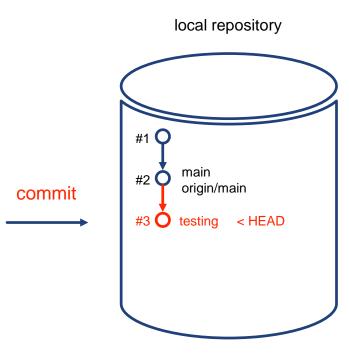
\$ git status
On branch testing



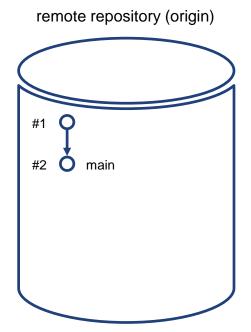




3. Commit changes to the branch: git commit



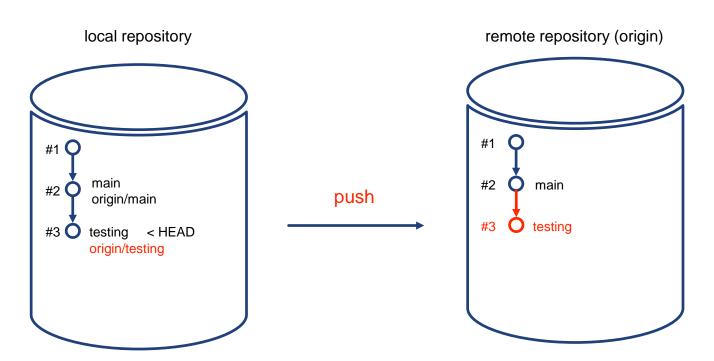
\$ git status
On branch testing







4. Push to remote repository: git push

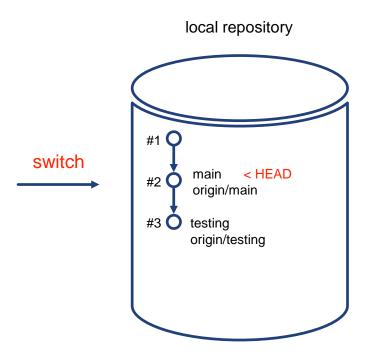


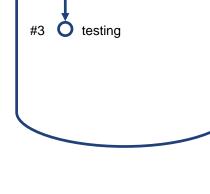
\$ git status
On branch testing
Your branch is up to date with 'origin/testing'.





5. Switch back to main: git switch main





main

#1

#2

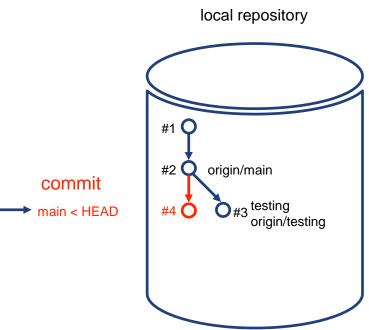
remote repository (origin)

\$ git status
On branch main
Your branch is up to date with 'origin/main'.

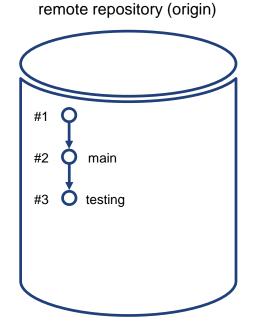




6. Commit changes to main: git commit



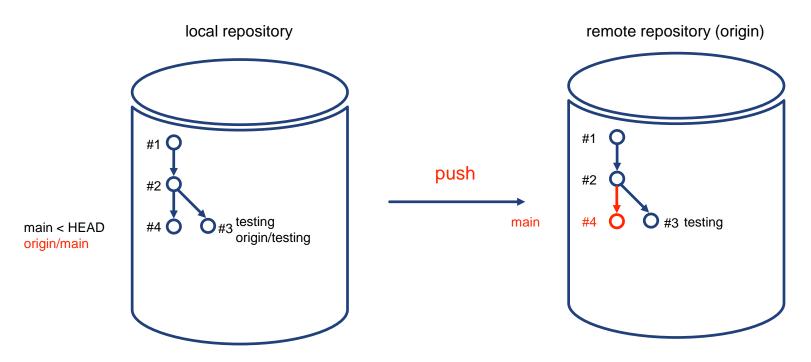
\$ git status
On branch main
Your branch is ahead of 'origin/main' by 1 commit.







7. Push main to remote: git push



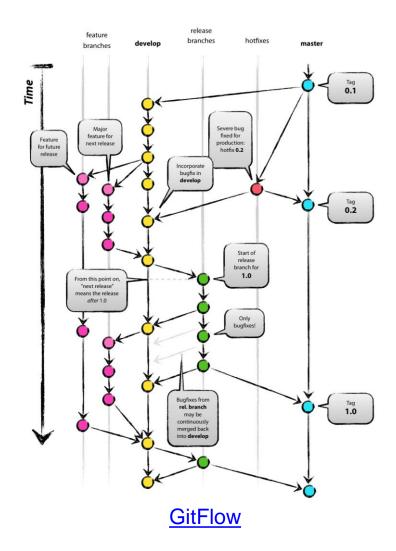
\$ git status
On branch main
Your branch is up to date with 'origin/testing'.





Branching strategies

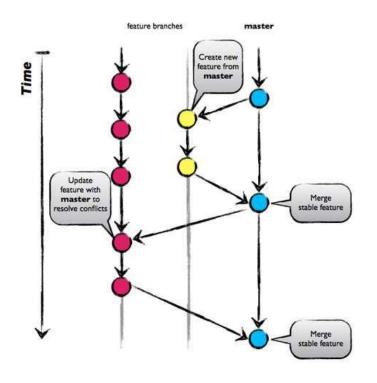
As your project grows in complexity, you may need a proper branching strategy.



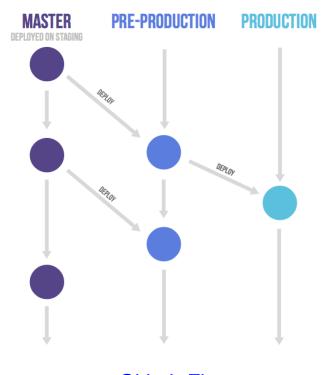




Branching strategies



But don't make it more complicated than needed.



GitHub Flow

GitLab Flow





Why Git as VCS?

- Distributed
 - Local repository: speed + freedom
 - Remote repository: collaboration + backup
- Fast
 - Most operations are local.
 - Operations are lightweight, e.g. a branch is just a "moving label".
- Industry standard





Best practices

- Atomic commits
 - Make scope-limited changes that are easy to grasp (and hence summarize).
 - Don't "also" reformat/refactor code you need not change right now.
 - Keep the diff-s small. No distractions because of things like white-space.
- Document the change
 - Start commit message with one-line summary. (Easy with atomic commits.)
 - Elaborate the reasoning in a follow-up paragraph, if needed.
 - Think of your future self: In which commit did I make that change?
- Branching
 - Branch often, you can always merge (or rebase).
 - One branch for each issue/task. Break it down into commits as you see fit.
 - In the main branch, a (simple) "linear commit history" is often desirable.
- Keep out large files
 - Use extension like Git-LFS or Git-annex for large binary or data files.





Git remote servers with web UI

Git web hosts:

ETH: <u>gitlab.ethz.ch</u>

SWITCH: <u>gitlab.switch.ch</u>

GitLab: <u>gitlab.com</u>

GitHub: <u>github.com</u>

BitBucket: <u>bitbucket.org</u>

- Benefits:
 - code browser
 - merge/pull requests
 - issue tracker
 - documentation hosting (wiki, web site)
 - continuous integration, continuous deployment (CI/CD)





Git hands-on session

Let's dig in!

https://siscourses.ethz.ch/git-handson/hands_on_git.html