

Simplifying the contribution process for both contributors & maintainers

A case study of the  integration in  easybuild

GitHub



FOSDEM 2018

Testing and Automation devroom

Feb 3rd 2018, Brussels (Belgium)

Kenneth Hoste

kenneth.hoste@ugent.be

GitHub: @boegel

Twitter: @kehoste



*GitHub is a code hosting service
for git repositories*

<https://github.com>

- launched in April 2008
- currently most popular platform for collaborating on FOSS
 - over 24M users, 67M repositories
 - over 100M *merged* pull requests
- also provides issue tracker, code review facilities, wiki, ...



*Git is a free and open source (GPLv2)
distributed version control system*

<https://git-scm.com>

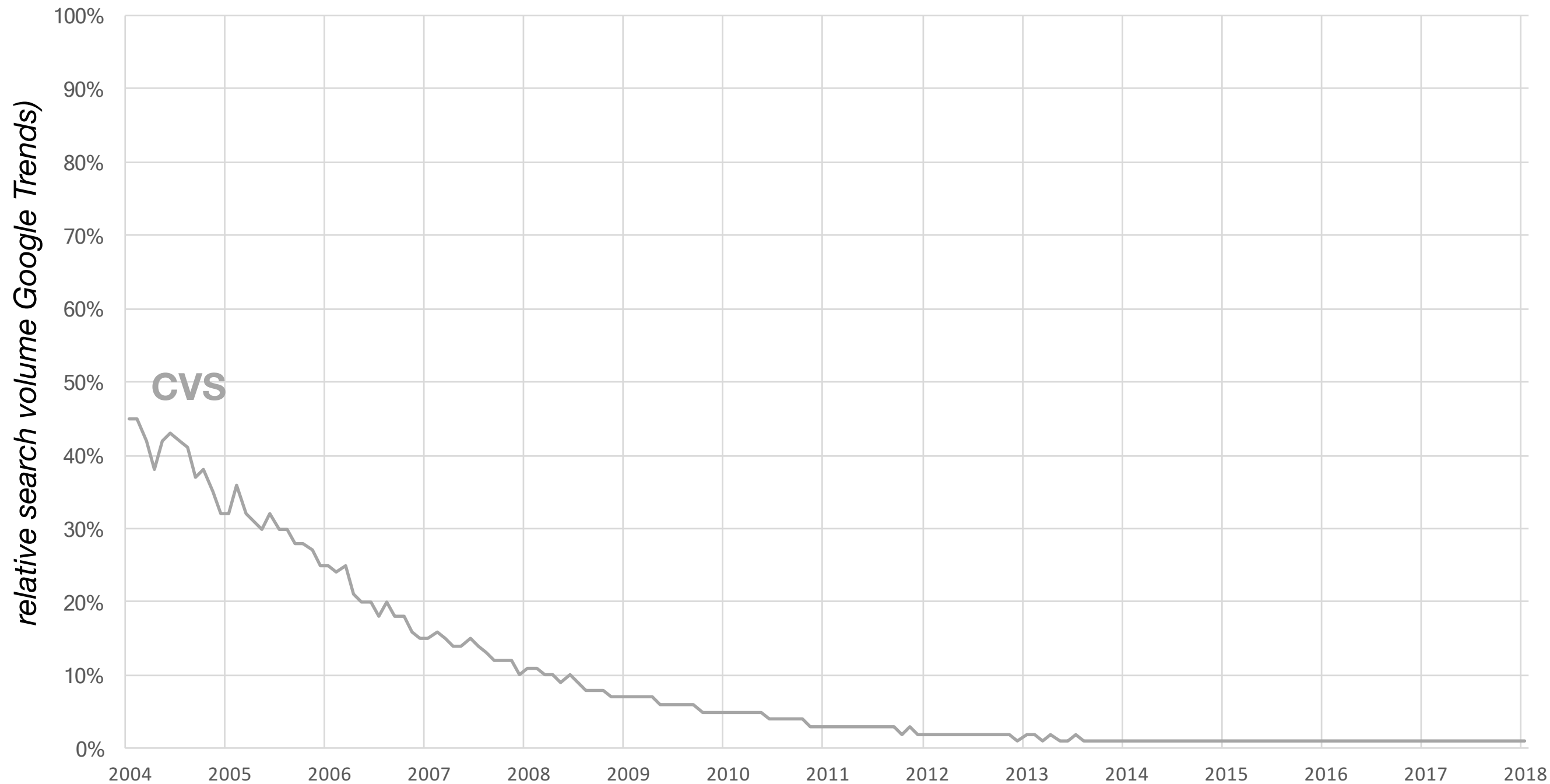
- created in 2005 by Linus Torvalds
- currently (by far) most popular version control system (VCS)
- design goals:
 - good performance, even for large projects (Linux kernel)
 - support for non-linear development & distributed workflow
 - to not be like CVS (Concurrent Versions System)
- internals similar to modern *blockchain* technologies (Merkle trees)

Git workflow: branch, stage, commit, push

```
# edit file to contribute  
# (you can use emacs too, if you insist)  
$ vim file  
  
# create branch, and check it out  
$ git checkout -b branch_name  
  
# stage file  
$ git add file  
  
# commit changes with sensible commit message  
$ git commit -m "this is a sensible commit message"  
  
# push branch to remote repository  
$ git push remote_repo branch_name
```

A little bit of VCS history

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)



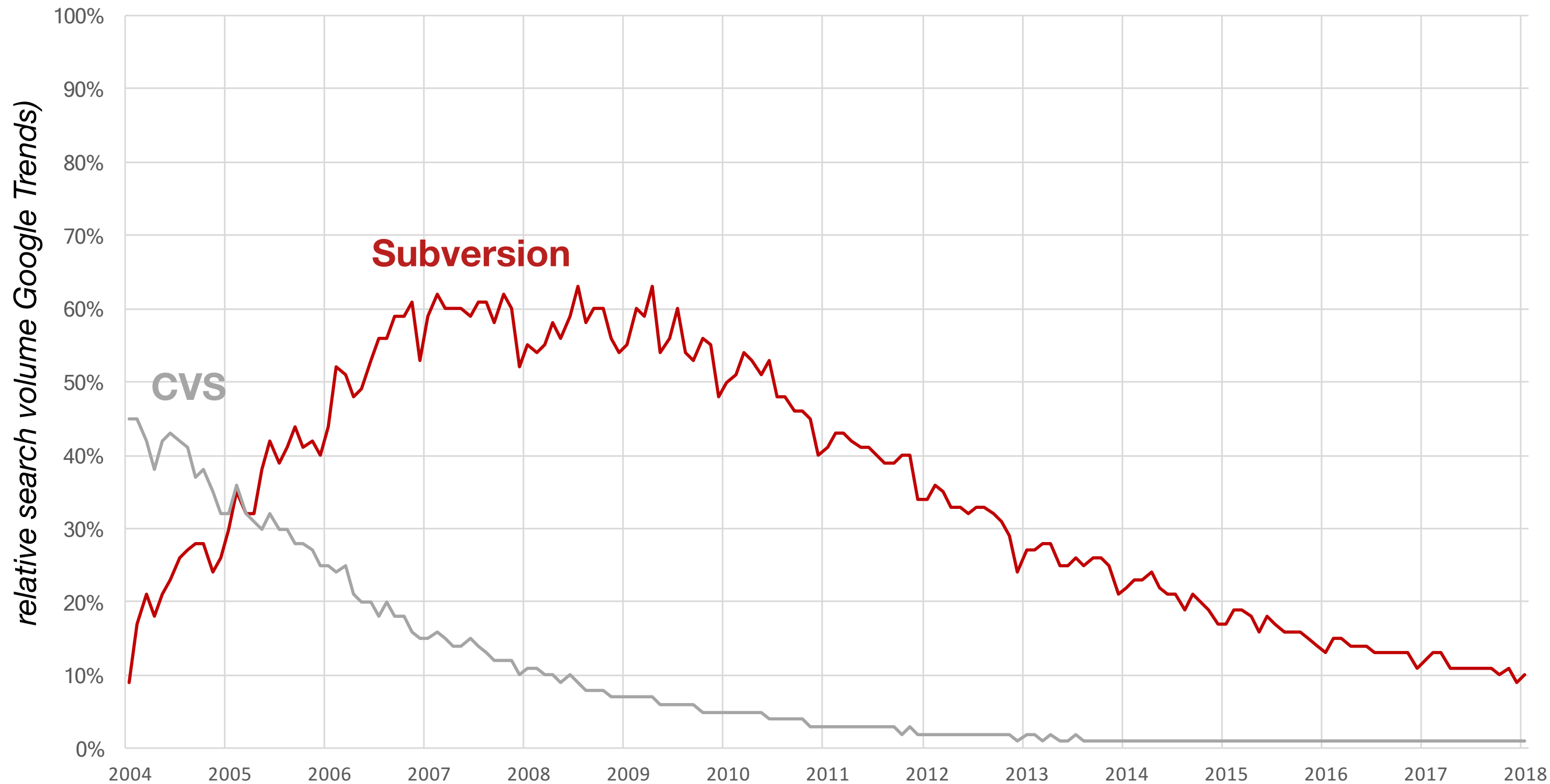
Concurrent Versions System (CVS)

<http://www.nongnu.org/cvs>

initial release: Nov 1990, last release: May 2008

A little bit of VCS history

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)



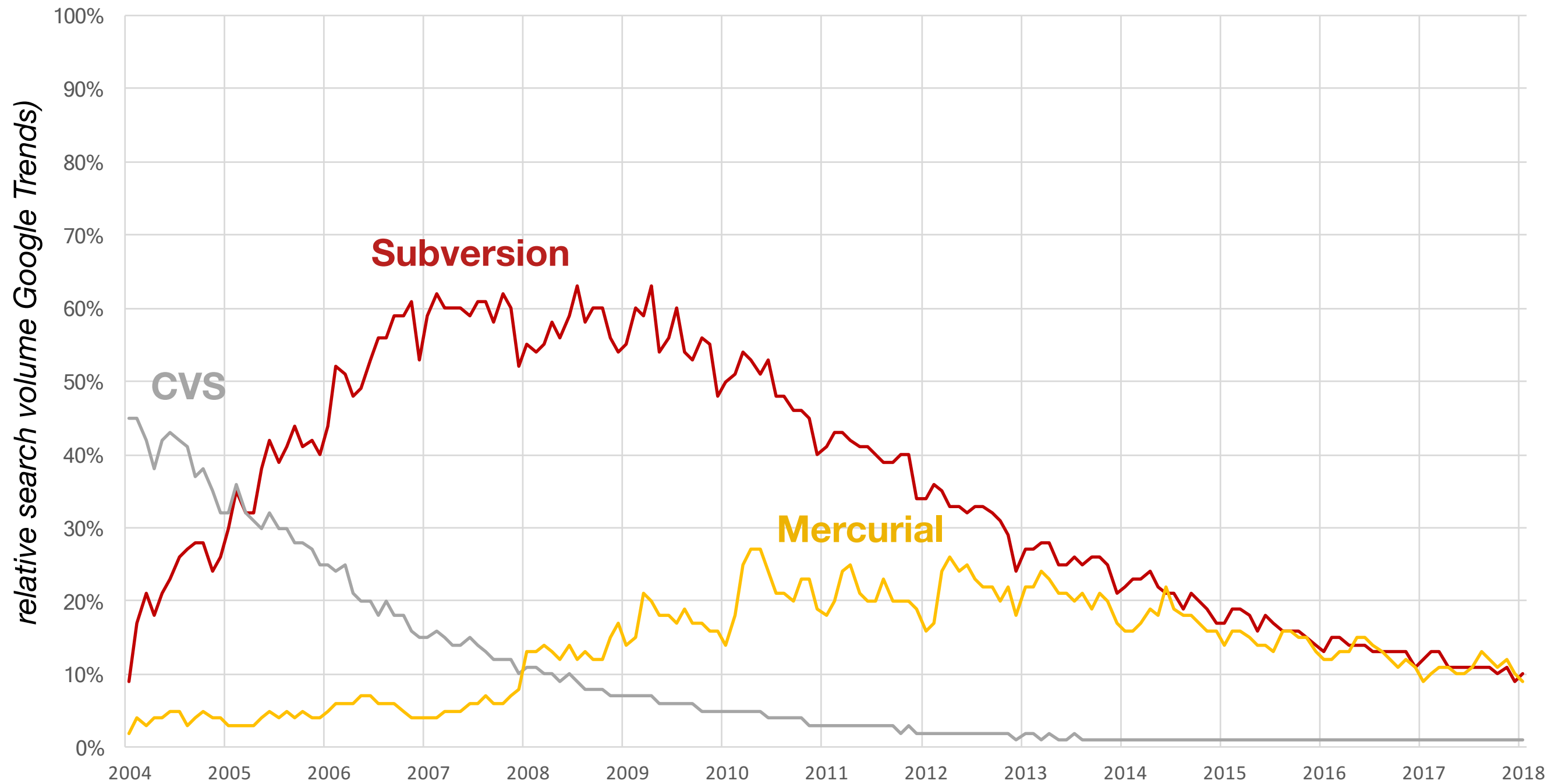
Apache Subversion (SVN)

<https://subversion.apache.org>

initial release: Oct 2000, last release: Aug 2017

A little bit of VCS history

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)



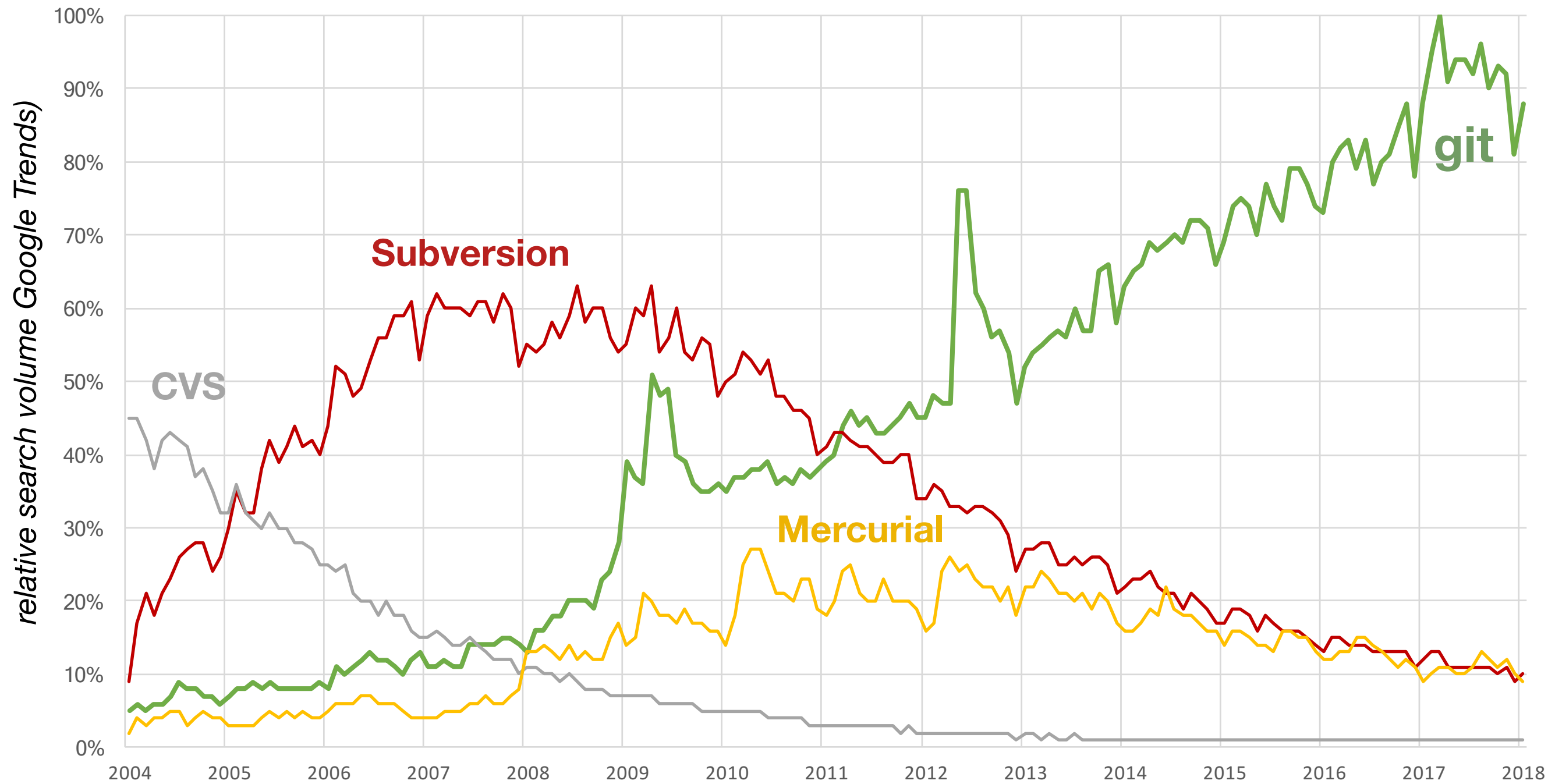
Mercurial

<https://www.mercurial-scm.org>

initial release: April 2005, last release: Dec 2017

A little bit of VCS history

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)



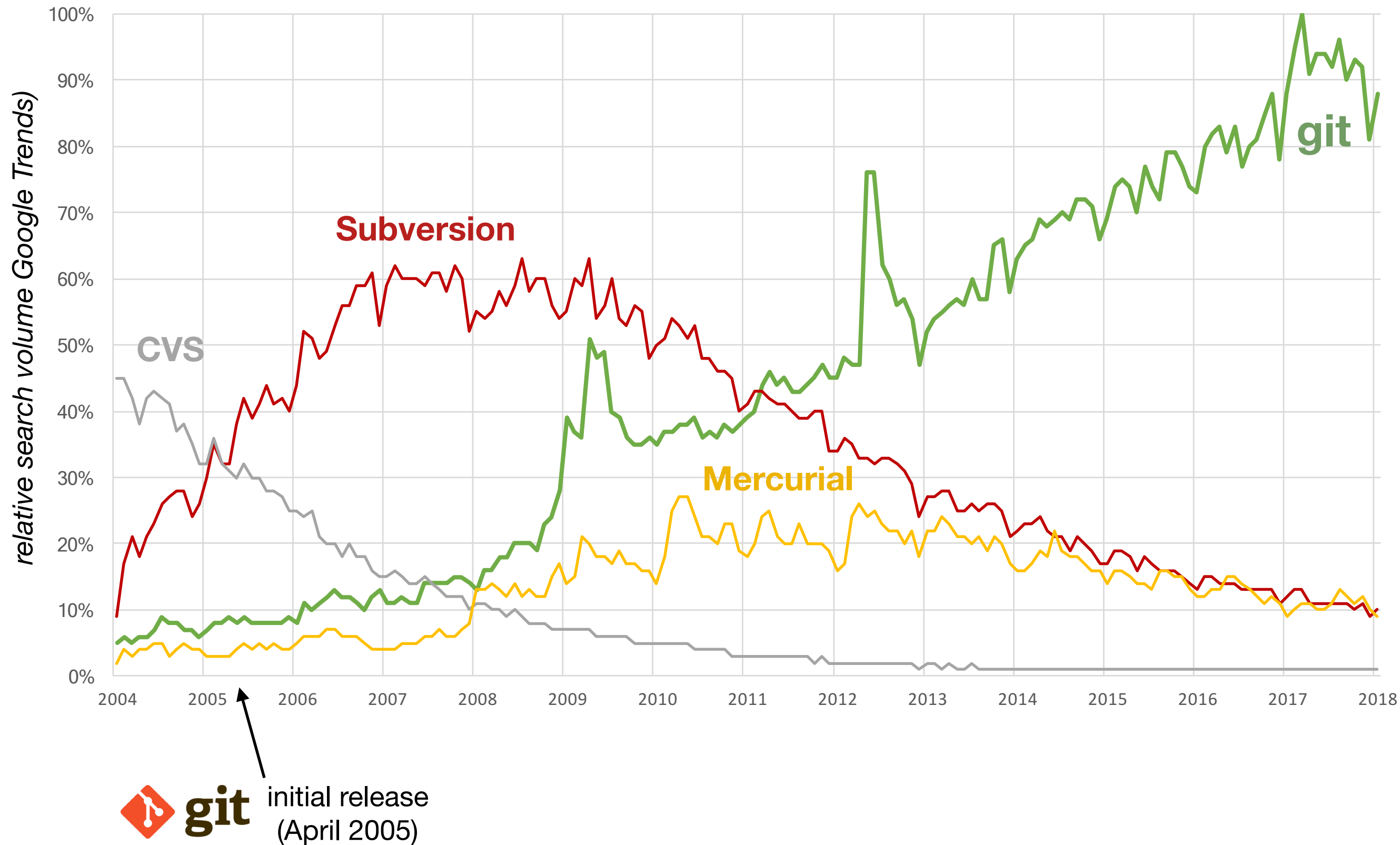
git

<https://git-scm.com>

initial release: April 2005, last release: Jan 2018

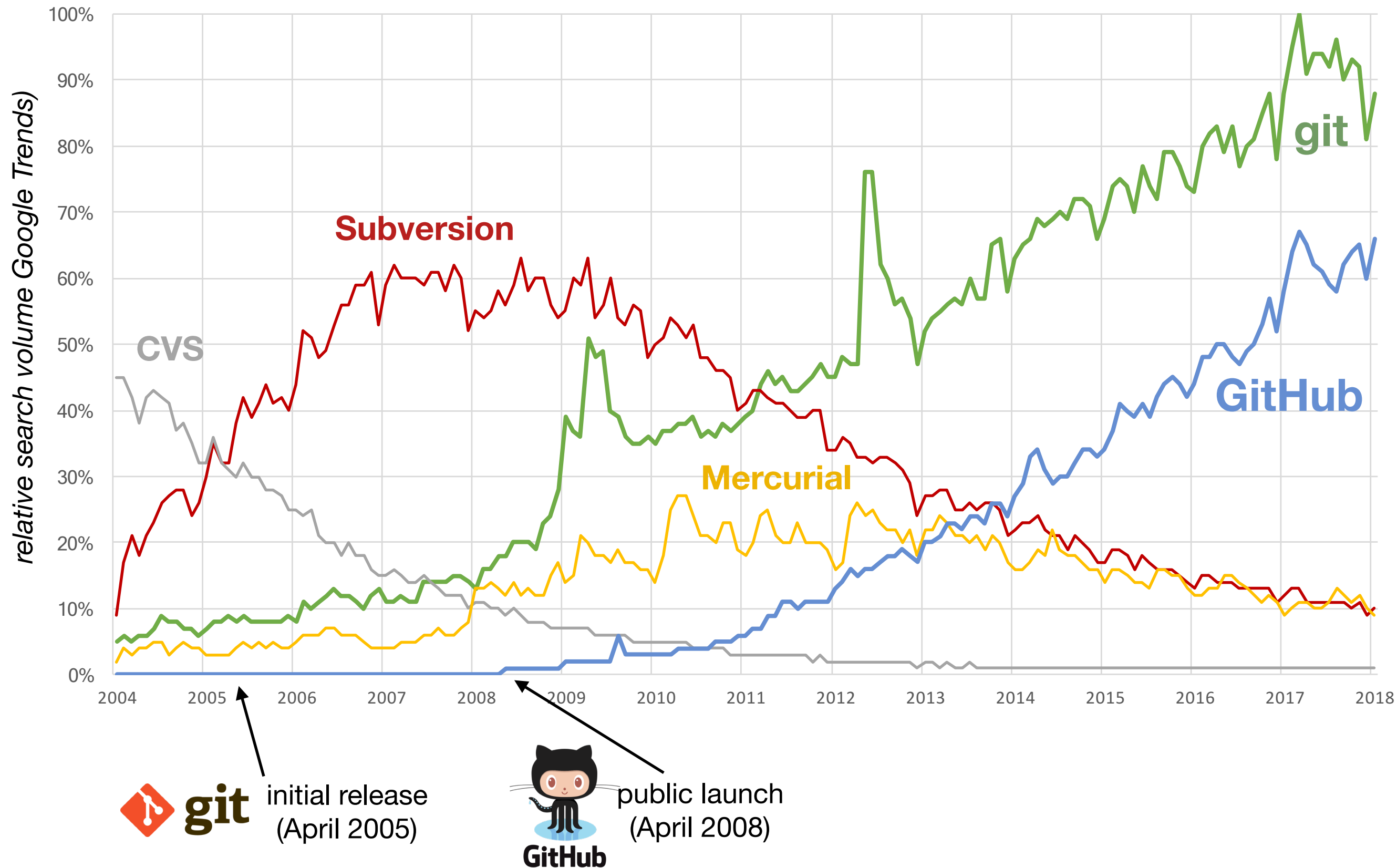
Why did **git** become so popular?

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)



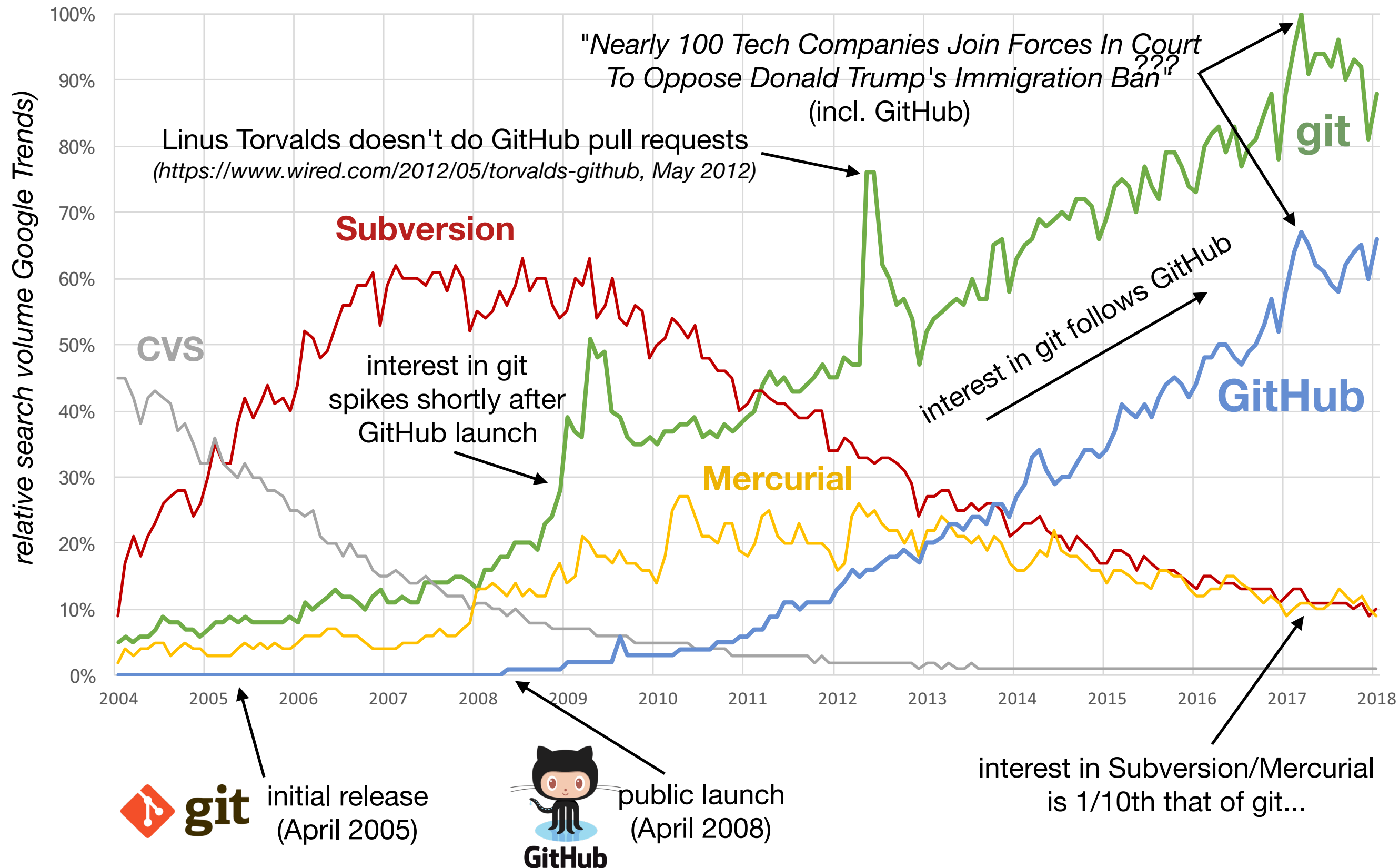
Is **git** so popular thanks to **GitHub**?

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)






Is git so popular thanks to GitHub?

(inspired by: <https://blog.gitprime.com/git-didnt-beat-svn-github-did>, May 2016)



Is **git** so popular thanks to **GitHub**?



- Hard to be sure, but it's likely  has had a major impact.
GitHub
- Maybe there even was some sort of feedback loop...
(**GitHub** increases interest in  **git**, which in turns makes **GitHub** more popular?)
- ... or maybe it's just because **GitHub** has 'git' in its name,
and I can't get Google Trends to behave...
- Whatever the reason,  **git** is the most popular (D)VCS today.
- You may need it in order to contribute to a FOSS project.
- *Which means you will need to learn how to use it...*

git has a steep learning curve

<https://stackoverflow.com/questions/tagged/git>

16362

votes

75

answers

5.9m
views

How to undo the most recent commits in Git?

I accidentally committed wrong files to Git, but I haven't pushed the commit to the server yet. How can I undo those commits from Git?

git

git-commit

git-reset

git-revert

community wiki

61 revs, 40 users 14%

Peter Mortensen

12440

votes

40

answers

5.1m
views

How do I delete a Git branch both locally and remotely?

I want to delete a branch both locally and on my remote project fork on GitHub. Failed Attempts to Delete Remote Branch \$ git branch -d remotes/origin/bugfix error: branch 'remotes/origin/bugfix' ...

git

git-branch

git-remote

asked Jan 5 '10 at 1:12



Matthew Rankin

227k ● 33 ● 100 ● 139

9236

votes

44

answers

2.1m
views

What is the difference between 'git pull' and 'git fetch'?

What are the differences between git pull and git fetch?

git

git-pull

git-fetch

asked Nov 15 '08 at 9:51



Pablo

94.9k ● 96 ● 258 ● 387

(in comparison: the infamous "How to exit Vim" has 'only' 1.3M views...)

git has a steep learning curve

<https://twitter.com/msccccc/status/954098595623391233>



Mike Coutermarsh

@msccccc

Just googled "undo last commit but keep changes"

And I work for GitHub.

This stuff is hard!

6:10 PM - 18 Jan 2018

321 Retweets **1,010** Likes



git has a steep learning curve

(from <https://www.quora.com/Why-is-Git-so-hard-to-learn>)



Terry Lambert, former Sr. Software Engineer at Google (2011-2012)

Answered Nov 10, 2016

The primary problem with git is that there are too many ways to do the same few, conceptually simple things.

***It's the perl of source code control systems.
This is not a compliment.***



William Estoque, Ruby, Ruby on Rails, Javascript, Engineer, Startups

Answered Nov 4, 2016

Git is hard to learn because people try to learn everything at once.



Ori Shalev, a software engineer

Answered Jan 2, 2017

Because probably like most people, you skipped the basics. .



Jesse McGrew, 15+ years as a professional and open source developer

Answered Jan 8, 2017

Git is hard to learn because its developers gave up on making it easy.

git is not for everybody


- main target audience is *software developers*
- in particular *teams* of people working together on software
- a lot of (computer) people don't need it for their *daily work*
- some feel they lack the *time* to learn it
- others *try* to learn, but can't seem to get their head around it...
- not everybody knows an experienced git user to *help* them
- some people don't **want** to learn it (and that's fine)

git can be a major hurdle for contributors

- Potential contributors may not be familiar (yet) with `git`...
- Do they ***have*** to learn `git` first before they can contribute?



The burden for contributors

- be sufficiently familiar with the project they want to contribute to
 - programming language(s) used
 - overall structure of the project's codebase
 - project-specific policies & expectations,
incl. code style rules, legal aspects, (git) workflow, testing, ...
 - recommended approach to implement a bug fix/enhancement
- know enough about  **git** & **GitHub** to open a '*correct*' pull request
- initial effort for newcomers is often (heavily) underestimated
- documentation is very useful, but not sufficient for everyone

A little bit about easybuild ...

<http://easybuilders.github.io/easybuild>

- **framework for installing (scientific) software on HPC systems**
- uses easyconfig files as 'recipes'
 - key-value definitions for easyconfig parameters (Python syntax)
 - specify software name, version, compiler toolchain, dependencies, ...
- common contributions to EasyBuild:
 - add easyconfig file for a new software version, or for other toolchain
 - very often based on an existing easyconfig file (copy & modify)
 - statement that "this particular (set of) recipe(s) works for me"

"high-performance computing"

Would you like to know more?

Installing software for scientists on a multi-user HPC system

A comparison between conda, EasyBuild, Guix, Nix & Spack

 **CONDA**

 **easybuild**

 **Guix**

 **Nix**

 **Spack**

Sunday February 4th 2018 - 9am, room H.1302 (Depage)

HPC, Big Data & Data Science devroom @  FOSDEM'18

https://fosdem.org/2018/schedule/event/installing_software_for_scientists

Example usage of EasyBuild

```
$ eb sklearn.eb
== temporary log file in case of crash /tmp/eb-Wqpzyc/easybuild-5ll606.log
== processing EasyBuild easyconfig /home/boegel/sklearn.eb
== building and installing scikit-learn/0.19.1-intel-2017b-Python-3.6.3...
== fetching files...
== creating build dir, resetting environment...
== unpacking...
== patching...
== preparing...
== configuring...
== building...
== testing...
== installing...
== taking care of extensions...
== postprocessing...
== sanity checking...
== cleaning up...
== creating module...
== permissions...
== packaging...
== COMPLETED: Installation ended successfully
== Results of the build can be found in the log file(s) /example/software/scikit-learn/0.19.1-
intel-2017b-Python-3.6.3/easybuild/easybuild-scikit-learn-0.19.1-20180125.210511.log
== Build succeeded for 1 out of 1
== Temporary log file(s) /tmp/eb-Wqpzyc/easybuild-5ll606.log* have been removed.
== Temporary directory /tmp/eb-Wqpzyc has been removed.
```

*EasyBuild consumes **easyconfig files**,
which specify what to install*

Example easyconfig file

```
easyblock = 'PythonPackage'

name = 'scikit-learn'
version = '0.19.1'
versionsuffix = '-Python-%(pyver)s'
```

easy to tweak: **software version**,
compiler toolchain (version),
dependencies, configure options, ...

```
homepage = 'http://scikit-learn.org/stable/index.html'
description = "Scikit-learn is a Python library for machine learning algorithms."
```

```
toolchain = {'name': 'intel', 'version': '2017b'}
```

```
source_urls = [PYPI_SOURCE]
sources = [SOURCE_TAR_GZ]
checksums = ['5ca0ad32ee04abe0d4ba02c8d89d501b4e5e0304bdf4d45c2e9875a735b323a0']
```

```
dependencies = [
    ('Python', '3.6.3'),
    ('matplotlib', '2.1.0', versionsuffix),
]
```

```
options = {'modulename': 'sklearn'}
```

```
sanity_check_paths = {
    'files': [],
    'dirs': ['lib/python%(pyshortver)s/site-packages/sklearn'],
}
```

```
moduleclass = 'data'
```


So you want to contribute to easybuild ...

To contribute an easyconfig file,
you first need to prepare:

I shall ~~say~~ do zis only once!



1) create a  account via
GitHub

[Sign up for GitHub](#)

2)  **Fork** <https://github.com/easybuilders/easybuild-easyconfigs>

3) `git clone git@github.com:YOU/easybuild-easyconfigs.git`

(you *may* want to change this bit)

Branch, stage, commit, push

Once you have cloned the easyconfigs repository, you can really get to work...

```
# create and check out a new branch
```

```
$ git checkout -b example
```

```
# stage the new/changed easyconfig file(s)
```

```
$ git add easybuild/easyconfigs/s/scikit-learn/*.eb
```

```
# commit the changes with a sensible commit message
```

```
$ git commit -m "update scikit-learn to 0.19.1"
```

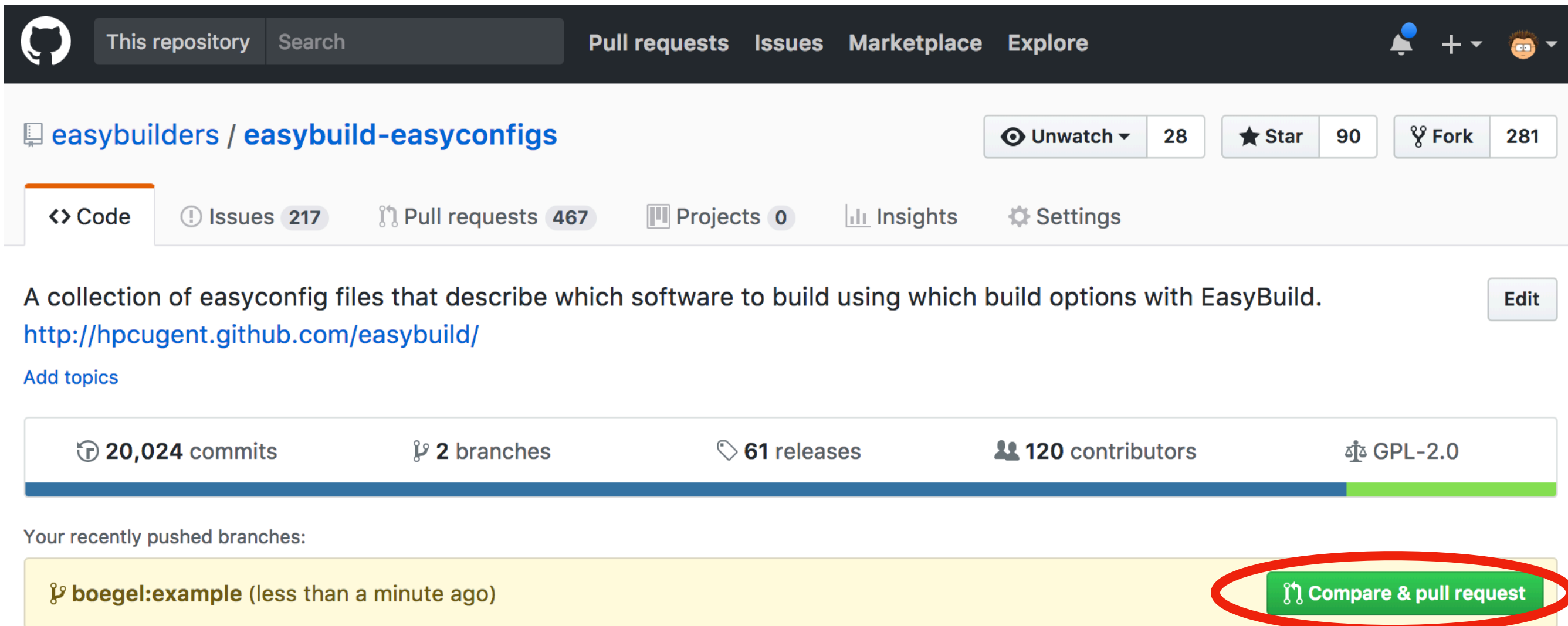
```
# push your branch to your fork on GitHub
```

```
$ git push origin example
```

(this may be different in your setup, see `git remote -v`)

Clickety, clickety, click.

Next, log on to  and open a pull request...
GitHub



The screenshot shows the GitHub interface for the repository `easybuilders / easybuild-easyconfigs`. The repository has 28 watchers, 90 stars, and 281 forks. It contains 217 issues, 467 pull requests, 0 projects, and 120 contributors. The license is GPL-2.0. The description states: "A collection of easyconfig files that describe which software to build using which build options with EasyBuild." and provides the URL <http://hpcugent.github.com/easybuild/>. The repository has 20,024 commits, 2 branches, 61 releases, and 120 contributors. The license is GPL-2.0. The "Your recently pushed branches:" section shows a branch named `boegel:example` pushed less than a minute ago. A red circle highlights the `Compare & pull request` button next to this branch.

You can use this useful button , if you're quick enough!



Almost there...

Open a pull request

Create a new pull request by comparing changes across two branches. If you need to, you can also [compare across forks](#).

base fork: easybuilders/easybuild-eas...
base: **develop** ▼

head fork: boegel/easybuild-easyconfigs ▼

compare: example ▼

✓ **Able to merge.** These branches can be automatically merged.

{data}[intel/2017b] scikit-learn 0.19.1 (Python 3.6.3)

Write Preview

AA ▼ B i “ <> 🔗 ⋮ 1/2/3 ✓

This is my first contribution to EasyBuild, please be gentle...

Attach files by dragging & dropping, [selecting them](#), or pasting from the clipboard.

☒ Allow edits from maintainers. [Learn more](#)

Create pull request


1 commit 1 file changed 0 commit comments 1 contributor

Not too quick...

- target **develop**
- tweak PR title
- (brief description)
- double-check changed files

If you're sure, click the shiny green button...

Are we done yet?

- once your pull request is open, your changes will be:
 - quickly tested by  Travis CI (automagically)
 - reviewed by a maintainer
 - thoroughly tested by at least one person
- additional changes may be requested...
- which means you may have to go back to `git`...
- maybe even multiple times... (spread over multiple hours/days)



Finetuning your contribution

Do you still remember the funny branch name you used for that PR?

```
# check out your branch again
```

```
$ git checkout example
```

```
# edit some files, as requested by the friendly reviewer
```

```
$ vim ...
```

```
# commit (all) your changes
```

```
$ git commit -am "fixed suggestions by reviewer"
```

```
# push your updated branch to GitHub to update the PR
```

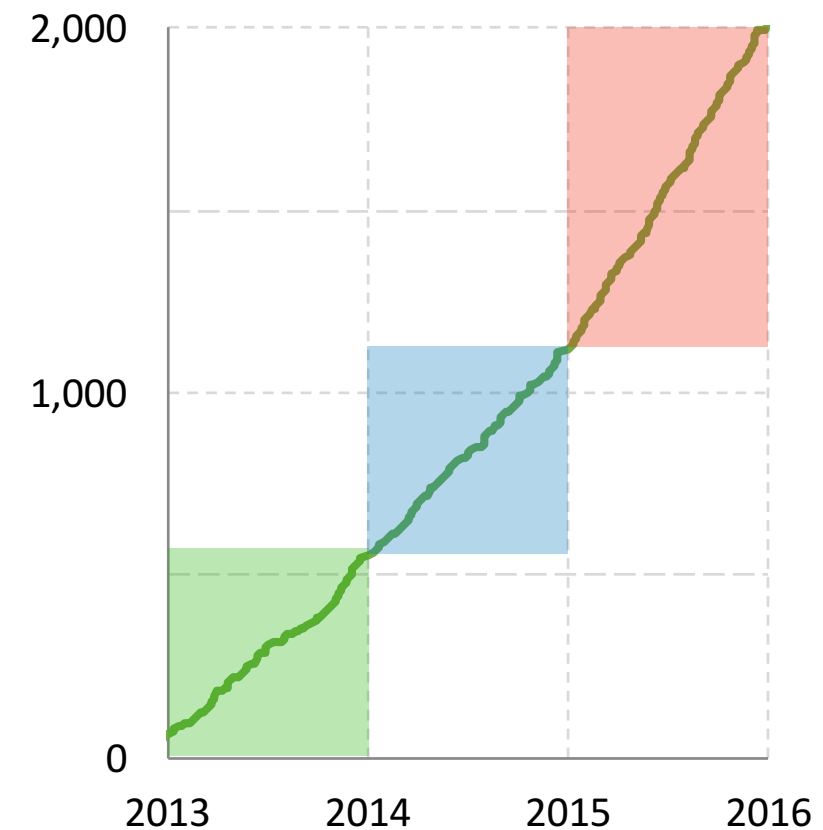
```
$ git push origin example
```

Too much overhead, manual steps...

- *"Isn't there an easier/more efficient way for doing this?!"*
 - need to come up with a sensible funny branch name first...
 - 4 git commands to prepare the contribution
 - at least 4 well aimed mouse clicks to open the pull request
 - if requested, N x 3 more git commands for finetuning the PR
 - + a bit of cleanup afterwards
(delete branch both in local repository clone and on GitHub)
- All of this assuming you:
 - take into account the project policies
 - don't mess up with git at any time...


Contributions to

- more incoming contributions in 2015
 - ~500 easyconfig PRs in 2013 & 2014
 - almost double that in 2015...
- several people were struggling with the contribution procedure
 - mainly because they were not sufficiently familiar with `git`
 - most candidate contributors were not software developers
 - resulted in a lot of frustration & wasted time
 - some people even gave up on contributing :(
- quite annoying for a project that aims to *automate* tedious software installation procedures...



GitHub integration in easybuild

http://easybuild.readthedocs.io/en/latest/Integration_with_GitHub.html

- goals:
 - automate contribution workflow
 - avoid direct interaction with `git` and  GitHub
- initial implementation added support for:
 - opening new PRs & updating existing PRs **`eb --new-pr`**
`eb --update-pr`
 - downloading & using easyconfigs from a PR **`eb --from-pr`**
 - uploading test reports to a PR **`eb --upload-test-report`**
- limited to easyconfig files, for now...
 - but implemented such that it is easy to extend to easyblocks, and even the EasyBuild framework itself

Under the covers

- **GitPython** (<https://pypi.python.org/pypi/GitPython>)
 - Python library to interact with  **git**
-  **REST API** (<https://developer.github.com/v3>)
GitHub
 - to automate interaction with  (less clickety, clickety)
GitHub
- **keyring** Python package (<https://pypi.python.org/pypi/keyring>)
 - Python library that provides access to system keyring
 - required to securely store  tokens
GitHub

(outlined approach is specific to Python, but it should be easy to implement something similar in other programming languages)

Opening new pull requests

eb --new-pr

```
$ eb --new-pr sklearn.eb
```

```
== temporary log file in case of crash /tmp/eb-Iv07fs/easybuild-inS8Ny.log
== fetching branch 'develop' from
    https://github.com/easybuilders/easybuild-easyconfigs.git...
== copying easyconfigs to
    /tmp/eb-Iv07fs/git-working-dirhYsAA_/easybuild-easyconfigs...
```

Opening pull request

```
* target: easybuilders/easybuild-easyconfigs:develop
* from: boegel/easybuild-easyconfigs:20180125211924_new_pr_scikit-learn0191
* title: "{data}[intel/2017b] scikit-learn v0.19.1"
* description:
"""
```

```
(created using `eb --new-pr`)
```

```
"""
```

```
* overview of changes:
```

```
.../s/scikit-learn/scikit-learn-0.19.1-intel-2017b-Python-3.6.3.eb | 31 ++++++++
1 file changed, 31 insertions(+)
```

```
Opened pull request: https://github.com/easybuilders/easybuild-easyconfigs/pull/12345
== Temporary log file(s) /tmp/eb-Iv07fs/easybuild-inS8Ny.log* have been removed.
== Temporary directory /tmp/eb-Iv07fs has been removed.
```

one single 'eb' command
no 'git' commands, no clickety on GitHub

Opening new pull requests

eb --new-pr

```
$ eb --new-pr sklearn.eb
== temporary log file in case of crash /tmp/eb-Iv07fs/easybuild-inS8Ny.log
== fetching branch 'develop' from
https://github.com/easybuilders/easybuild-easyconfigs.git...
== copying easyconfigs to
/tmp/eb-Iv07fs/git-working-dirhYsAA_/easybuild-easyconfigs...
```

Opening pull request

```
* target: easybuilders/easybuild-easyconfigs:develop
* from: boegel/easybuild-easyconfigs:20180125211924_new_pr_scikit-learn0191
* title: "{data}[intel/2017b] scikit-learn v0.19.1"
* description:
"""
```

```
(created using `eb --new-pr`)
```

```
"""
```

```
* overview of changes:
```

```
.../s/scikit-learn/scikit-learn-0.19.1-intel-2017b-Python-3.6.3.eb | 31 ++++++
1 file changed, 31 insertions(+)
```

repository is **cloned** to temporary directory,
develop branch is **checked out** as a base

```
Opened pull request: https://github.com/easybuilders/easybuild-easyconfigs/pull/12345
== Temporary log file(s) /tmp/eb-Iv07fs/easybuild-inS8Ny.log* have been removed.
== Temporary directory /tmp/eb-Iv07fs has been removed.
```

Opening new pull requests

eb --new-pr

```
$ eb --new-pr sklearn.eb
== temporary log file in case of crash /tmp/eb-Iv07fs/easybuild-inS8Ny.log
== fetching branch 'develop' from
    https://github.com/easybuilders/easybuild-easyconfigs.git...
== copying easyconfigs to
    /tmp/eb-Iv07fs/git-working-dirhYsAA_/easybuild-easyconfigs...
```

*auto-derived
branch name*

Opening pull request

```
* target: easybuilders/easybuild-easyconfigs:develop
* from: boegel/easybuild-easyconfigs 20180125211924_new_pr_scikit-learn0191
* title: "{data}[intel/2017b] scikit-learn v0.19.1"
* description:
"""
```

(created using `eb --new-pr`)

"""

```
* overview of changes:
```

```
.../s/scikit-learn/scikit-learn-0.19.1-intel-2017b-Python-3.6.3.eb | 31 ++++++++
1 file changed, 31 insertions(+)
```

```
Opened pull request: https://github.com/easybuilders/easybuild-easyconfigs/pull/12345
== Temporary log file(s) /tmp/eb-Iv07fs/easybuild-inS8Ny.log* have been removed.
== Temporary directory /tmp/eb-Iv07fs has been removed.
```

a new branch is created & pushed
(you can specify the branch name yourself)

Opening new pull requests

eb --new-pr

```
$ eb --new-pr sklearn.eb
== temporary log file in case of crash /tmp/eb-Iv07fs/easybuild-inS8Ny.log
== fetching branch 'develop' from
    https://github.com/easybuilders/easybuild-easyconfigs.git...
== copying easyconfigs to
    /tmp/eb-Iv07fs/git-working-dirhYsAA_/easybuild-easyconfigs...
```

Opening pull request

```
* target: easybuilders/easybuild-easyconfigs:develop
* from: boegel/easybuild-easyconfigs:20180125211924_new_pr_scikit-learn0191
* title: "{data}[intel/2017b] scikit-learn v0.19.1"
* description:
"""
```

```
(created using `eb --new-pr`)
```

```
"""
```

```
* overview of changes:
```

```
.../s/scikit-learn/scikit-learn-0.19.1-intel-2017b-Python-3.6.3.eb | 31 ++++++
1 file changed, 31 insertions(+)
```

the easyconfig file is *renamed*
and *moved to the right location*

```
Opened pull request: https://github.com/easybuilders/easybuild-easyconfigs/pull/12345
== Temporary log file(s) /tmp/eb-Iv07fs/easybuild-inS8Ny.log* have been removed.
== Temporary directory /tmp/eb-Iv07fs has been removed.
```

Opening new pull requests

eb --new-pr

```
$ eb --new-pr sklearn.eb
== temporary log file in case of crash /tmp/eb-Iv07fs/easybuild-inS8Ny.log
== fetching branch 'develop' from
    https://github.com/easybuilders/easybuild-easyconfigs.git...
== copying easyconfigs to
    /tmp/eb-Iv07fs/git-working-dirhYsAA_/easybuild-easyconfigs...
```

Opening pull request

```
* target: easybuilders/easybuild-easyconfigs:develop
* from: boegel/easybuild-easyconfigs:20180125211924_new_pr_scikit-learn0191
* title: "{data}[intel/2017b] scikit-learn v0.19.1"
```

```
* description:
"""
(created using `eb --new-pr`)
"""
```

the pull request is prepared:
target develop branch, *title & description*

```
* overview of changes:
.../s/scikit-learn/scikit-learn-0.19.1-intel-2017b-Python-3.6.3.eb | 31 ++++++
1 file changed, 31 insertions(+)
```

```
Opened pull request: https://github.com/easybuilders/easybuild-easyconfigs/pull/12345
== Temporary log file(s) /tmp/eb-Iv07fs/easybuild-inS8Ny.log* have been removed.
== Temporary directory /tmp/eb-Iv07fs has been removed.
```

Opening new pull requests

eb --new-pr

```
$ eb --new-pr sklearn.eb
== temporary log file in case of crash /tmp/eb-Iv07fs/easybuild-inS8Ny.log
== fetching branch 'develop' from
    https://github.com/easybuilders/easybuild-easyconfigs.git...
== copying easyconfigs to
    /tmp/eb-Iv07fs/git-working-dirhYsAA_/easybuild-easyconfigs...
```

Opening pull request

```
* target: easybuilders/easybuild-easyconfigs:develop
* from: boegel/easybuild-easyconfigs:20180125211924_new_pr_scikit-learn0191
* title: "{data}[intel/2017b] scikit-learn v0.19.1"
* description:
    """
    (created using `eb --new-pr`)
```

a high-level *overview of the changes* is shown (similar to output of 'git stat')

```
"""
* overview of changes:
.../s/scikit-learn/scikit-learn-0.19.1-intel-2017b-Python-3.6.3.eb | 31 ++++++++
1 file changed, 31 insertions(+)
```

```
Opened pull request: https://github.com/easybuilders/easybuild-easyconfigs/pull/12345
== Temporary log file(s) /tmp/eb-Iv07fs/easybuild-inS8Ny.log* have been removed.
== Temporary directory /tmp/eb-Iv07fs has been removed.
```

Opening new pull requests

eb --new-pr

```
$ eb --new-pr sklearn.eb
== temporary log file in case of crash /tmp/eb-Iv07fs/easybuild-inS8Ny.log
== fetching branch 'develop' from
    https://github.com/easybuilders/easybuild-easyconfigs.git...
== copying easyconfigs to
    /tmp/eb-Iv07fs/git-working-dirhYsAA_/easybuild-easyconfigs...
```

Opening pull request

```
* target: easybuilders/easybuild-easyconfigs:develop
* from: boegel/easybuild-easyconfigs:20180125211924_new_pr_scikit-learn0191
* title: "{data}[intel/2017b] scikit-learn v0.19.1"
* description:
    ""
```

```
(created using `eb --new-pr`)
```

```
""
```

```
* overview of changes:
```

```
.../s/scikit-learn/scikit-learn-0.19.1-intel-2017b-Python-3.6.3.eb | 31 ++++++++
1 file changed, 31 insertions(+)
```

the *pull request is opened automatically*,
using the GitHub API & token

```
Opened pull request: https://github.com/easybuilders/easybuild-easyconfigs/pull/12345
```

```
== Temporary log file(s) /tmp/eb-Iv07fs/easybuild-inS8Ny.log* have been removed.
```

```
== Temporary directory /tmp/eb-Iv07fs has been removed.
```


Opening a pull request is now... easy!

```
$ mv sklearn.eb scikit-learn-0.19.1-intel-2017b-Python-3.6.3.eb
$ mv scikit*.eb easybuild/easyconfigs/s/scikit-learn
$ git checkout develop && git pull upstream develop
$ git checkout -b scikit_learn_0191_intel_2017b
$ git add easybuild/easyconfigs/s/scikit-learn
$ git commit -m "{data}[intel/2017b] scikit-learn v0.19.1"
$ git push origin scikit_learn_0191_intel_2017b
```

+ log into GitHub to actually open the pull request (clickety, clickety...)

one single eb command

no git commands

no  interaction
GitHub

metadata is automatically
derived from easyconfig

saves a lot of time!

eb --new-pr sklearn.eb

Updating existing pull requests

eb --update-pr

to update a PR: provide PR#, updated files and commit message

```
$ eb --update-pr 12345 sklearn.eb --pr-commit-msg "minor style fixes"
== temporary log file in case of crash /tmp/eb-EKclWF/easybuild-wlYcat.log
== Determined branch name corresponding to easybuilders/easybuild-
easyconfigs PR #12345: 20180125211924_new_pr_scikit-learn0191
== fetching branch '20180125211924_new_pr_scikit-learn0191' from https://
github.com/boegel/easybuild-easyconfigs.git...
== copying easyconfigs to /tmp/eb-EKclWF/git-working-dirbIzGHn/easybuild-
easyconfigs...
Overview of changes:
.../s/scikit-learn/scikit-learn-0.19.1-intel-2017b-Python-3.6.3.eb | 2 +-
1 file changed, 1 insertion(+), 1 deletion(-)

Updated easybuilders/easybuild-easyconfigs PR #12345 by pushing to branch
boegel/20180125211924_new_pr_scikit-learn0191
== Temporary log file(s) /tmp/eb-EKclWF/easybuild-wlYcat.log* have been
removed.
== Temporary directory /tmp/eb-EKclWF has been removed.
```

Updating existing pull requests

eb --update-pr

branch that corresponds with PR is determined *automatically*

```
$ eb --update-pr 12345 sklearn.eb --pr-commit-msg "minor style fixes"
== temporary log file in case of crash /tmp/eb-EKclWF/easybuild-w1Ycat.log
== Determined branch name corresponding to easybuilders/easybuild-
easyconfigs PR #12345: 20180125211924_new_pr_scikit-learn0191
== fetching branch '20180125211924_new_pr_scikit-learn0191' from https://
github.com/boegel/easybuild-easyconfigs.git...
== copying easyconfigs to /tmp/eb-EKclWF/git-working-dirbIzGHn/easybuild-
easyconfigs...
Overview of changes:
.../s/scikit-learn/scikit-learn-0.19.1-intel-2017b-Python-3.6.3.eb | 2 +-
1 file changed, 1 insertion(+), 1 deletion(-)

Updated easybuilders/easybuild-easyconfigs PR #12345 by pushing to branch
boegel/20180125211924_new_pr_scikit-learn0191
== Temporary log file(s) /tmp/eb-EKclWF/easybuild-w1Ycat.log* have been
removed.
== Temporary directory /tmp/eb-EKclWF has been removed.
```

Updating existing pull requests

eb --update-pr

PR branch is downloaded to temporary directory,
updated files are copied

```
$ eb --update-pr 12345 sklearn.eb --pr-commit-msg "minor style fixes"
== temporary log file in case of crash /tmp/eb-EKclWF/easybuild-w1Ycat.log
== Determined branch name corresponding to easybuilders/easybuild-
easyconfigs PR #12345: 20180125211924_new_pr_scikit-learn0191
== fetching branch '20180125211924_new_pr_scikit-learn0191' from https://
github.com/boegel/easybuild-easyconfigs.git...
== copying easyconfigs to /tmp/eb-EKclWF/git-working-dirbIzGHn/easybuild-
easyconfigs...
Overview of changes:
.../s/scikit-learn/scikit-learn-0.19.1-intel-2017b-Python-3.6.3.eb | 2 +-
1 file changed, 1 insertion(+), 1 deletion(-)

Updated easybuilders/easybuild-easyconfigs PR #12345 by pushing to branch
boegel/20180125211924_new_pr_scikit-learn0191
== Temporary log file(s) /tmp/eb-EKclWF/easybuild-w1Ycat.log* have been
removed.
== Temporary directory /tmp/eb-EKclWF has been removed.
```

Updating existing pull requests

eb --update-pr

updated files are renamed and moved to right location

```
$ eb --update-pr 12345 sklearn.eb --pr-commit-msg "minor style fixes"
== temporary log file in case of crash /tmp/eb-EKclWF/easybuild-w1Ycat.log
== Determined branch name corresponding to easybuilders/easybuild-
easyconfigs PR #12345: 20180125211924_new_pr_scikit-learn0191
== fetching branch '20180125211924_new_pr_scikit-learn0191' from https://
github.com/boegel/easybuild-easyconfigs.git...
== copying easyconfigs to /tmp/eb-EKclWF/git-working-dirbIzGHn/easybuild-
easyconfigs...
Overview of changes:
.../s/scikit-learn/scikit-learn-0.19.1-intel-2017b-Python-3.6.3.eb | 2 +-
1 file changed, 1 insertion(+), 1 deletion(-)
```

```
Updated easybuilders/easybuild-easyconfigs PR #12345 by pushing to branch
boegel/20180125211924_new_pr_scikit-learn0191
== Temporary log file(s) /tmp/eb-EKclWF/easybuild-w1Ycat.log* have been
removed.
== Temporary directory /tmp/eb-EKclWF has been removed.
```

Updating existing pull requests

eb --update-pr

changes are committed, high-level overview is shown (cfr. 'git stat')

```
$ eb --update-pr 12345 sklearn.eb --pr-commit-msg "minor style fixes"
== temporary log file in case of crash /tmp/eb-EKclWF/easybuild-w1Ycat.log
== Determined branch name corresponding to easybuilders/easybuild-
easyconfigs PR #12345: 20180125211924_new_pr_scikit-learn0191
== fetching branch '20180125211924_new_pr_scikit-learn0191' from https://
github.com/boegel/easybuild-easyconfigs.git...
== copying easyconfigs to /tmp/eb-EKclWF/git-working-dirbIzGHn/easybuild-
easyconfigs...
```

Overview of changes:

```
.../s/scikit-learn/scikit-learn-0.19.1-intel-2017b-Python-3.6.3.eb | 2 +-
1 file changed, 1 insertion(+), 1 deletion(-)
```

```
Updated easybuilders/easybuild-easyconfigs PR #12345 by pushing to branch
boegel/20180125211924_new_pr_scikit-learn0191
```

```
== Temporary log file(s) /tmp/eb-EKclWF/easybuild-w1Ycat.log* have been
removed.
```

```
== Temporary directory /tmp/eb-EKclWF has been removed.
```

Updating existing pull requests

eb --update-pr

modified branch is pushed to GitHub to update the PR

```
$ eb --update-pr 12345 sklearn.eb --pr-commit-msg "minor style fixes"
== temporary log file in case of crash /tmp/eb-EKclWF/easybuild-wlYcat.log
== Determined branch name corresponding to easybuilders/easybuild-
easyconfigs PR #12345: 20180125211924_new_pr_scikit-learn0191
== fetching branch '20180125211924_new_pr_scikit-learn0191' from https://
github.com/boegel/easybuild-easyconfigs.git...
== copying easyconfigs to /tmp/eb-EKclWF/git-working-dirbIzGHn/easybuild-
easyconfigs...
Overview of changes:
.../s/scikit-learn/scikit-learn-0.19.1-intel-2017b-Python-3.6.3.eb | 2 +-
1 file changed, 1 insertion(+), 1 deletion(-)

Updated easybuilders/easybuild-easyconfigs PR #12345 by pushing to branch
boegel/20180125211924_new_pr_scikit-learn0191
== Temporary log file(s) /tmp/eb-EKclWF/easybuild-wlYcat.log* have been
removed.
== Temporary directory /tmp/eb-EKclWF has been removed.
```


Updating a PR in 1, ~~2~~, ~~3~~

First, figure out the (funny) branch name corresponding to the PR...

```
$ git checkout scikit_learn_0191_intel_2017b
$ cd easybuild/easyconfigs/s/scikit-learn
$ vim scikit-learn-0.19.1-intel-2017b-Python-3.6.3.eb
$ git add scikit*.eb
$ git commit -m "fix typo"
$ git push origin scikit_learn_0191_intel_2017b
```

one single eb command

no git commands

no  interaction
GitHub

no local branch to clean up
once PR is merged

saves a lot of time!

```
eb --update-pr 12345 sklearn.eb --pr-commit-msg "fix typo"
```


Using easyconfigs from a PR

eb --from-pr

- no need to manually download easyconfigs to feed them to eb
- PR is (virtually) merged with current develop branch in /tmp
- very useful when testing contributions!

```
$ eb --from-pr 12345
```

```
== temporary log file in case of crash /tmp/eb-kQRw7W/easybuild-HU4hce.log
== processing EasyBuild easyconfig /tmp/eb-kQRw7W/files_pr12345/s/scikit-learn/scikit-learn-0.19.1-intel-2017b-Python-3.6.3.eb
== building and installing scikit-learn/0.19.1-intel-2017b-Python-3.6.3...
== fetching files...
== creating build dir, resetting environment...
== unpacking...
== patching...
== preparing...
== configuring...
== building...
== testing...
== installing...
```


Uploading test reports to a PR

eb --upload-test-report

- combined with **--from-pr** to test contribution and report results
- test report with info on host & environment is uploaded as a gist
- a comment is added in the PR to trigger a notification by



```
$ eb --from-pr 12345 --upload-test-report
```

```
...  
== building and installing scikit-learn/0.19.1-intel-2017b-Python-3.6.3...  
...  
== COMPLETED: Installation ended successfully
```

```
Adding comment to easybuild-easyconfigs issue #12345: 'Test report by @boegel  
**SUCCESS**  
Build succeeded for 1 out of 1 (1 easyconfigs in this PR)  
<hostname> - Linux CentOS 7.5, Intel Xeon CPU E5-2680 v3 @ 2.50GHz, Python 2.7.5  
See https://gist.github.com/3c...58 for a full test report.'  
== Test report uploaded to https://gist.github.com/3c...58 and mentioned in a  
comment in easyconfigs PR#12345  
== Build succeeded for 1 out of 1
```

Uploading test reports to a PR

```
eb --upload-test-report
```



boegel commented on 21 Nov 2017

Owner



Test report by @boegel

SUCCESS

Build succeeded for 1 out of 1 (1 easyconfigs in this PR)

node2403.golett.os - Linux centos linux 7.4.1708, Intel(R) Xeon(R) CPU E5-2680 v3 @ 2.50GHz,
Python 2.7.5

See <https://gist.github.com/2ce867c657e05a8717978f3a73b9e019> for a full test report.



verdurin commented 3 days ago

Member



Test report by @verdurin

FAILED

Build succeeded for 0 out of 1 (1 easyconfigs in this PR)

rescomp2 - Linux centos 6.8, Intel(R) Xeon(R) CPU E5-2650 v2 @ 2.60GHz, Python 2.7.11

See <https://gist.github.com/0dd5b4a3338a1350fe3b92b247bf84bb> for a full test report.

Uploading test reports to a PR

eb --upload-test-report

EasyBuild test report for easyconfigs PR #5375

 [easybuild_test_report_easyconfigs_pr5375_20175921-UTC-10-59-58.md](#)

Test report for <https://github.com/easybuilders/easybuild-easyconfigs/pull/5375>

Test result

Build succeeded for 1 out of 1 (1 easyconfigs in this PR)

Overview of tested easyconfigs (in order)

- **SUCCESS** *scikit-learn-0.19.1-intel-2017b-Python-3.6.3.eb*

Time info

- start: Tue, 21 Nov 2017 10:57:11 +0000 (UTC)
- end: Tue, 21 Nov 2017 10:59:58 +0000 (UTC)

Uploading test reports to a PR

```
eb --upload-test-report
```

EasyBuild test report for easyconfigs PR #5719

 [easybuild_test_report_easyconfigs_pr5719_20180124-UTC-18-01-27.md](#)

Test report for <https://github.com/easybuilders/easybuild-easyconfigs/pull/5719>

Test result

Build succeeded for 0 out of 1 (1 easyconfigs in this PR)

Overview of tested easyconfigs (in order)

- **FAIL (build issue)** *RSEM-1.3.0-intel-2017b.eb* (partial log available at <https://gist.github.com/7d9e48b400232b5fa0924e2bf28d240a>)

Time info

- start: Wed, 24 Jan 2018 18:01:18 +0000 (UTC)

 **a separate gist provides partial build log**

More **GitHub** integration, MORE !!!1!

GitHub integration was extended gradually over time:

- checking whether requirements are in place `eb --check-github`
- verifying & installing a GitHub token `eb --install-github-token`
- dry run mode for `--new-pr` and `--update-pr` `eb --dry-run`
- downloading and using *partial* PRs `eb --from-pr 123 example.eb`
- (p)reviewing PRs `eb --preview-pr` `eb --review-pr`
- merging PRs, after checking they are eligible `eb --merge-pr`
- improve performance by avoiding to clone from scratch

Checking requirements for **GitHub** integration

eb --check-github

```
$ eb --check-github
```

```
Checking status of GitHub integration...
```

```
Making sure we're online...OK
```

various **checks** are performed
to verify whether required
stuff is in place

```
* GitHub user... boegel => OK
* GitHub token... e87..3b1 (len: 40) => OK (validated)
* git command... OK ("git version 1.8.3.1")
* GitPython module... OK (GitPython version 2.1.1)
* push access to boegel/easybuild-easyconfigs repo @ GitHub... OK
* creating gists... OK
* location to Git working dirs... OK (/home/example)
```

```
All checks PASSED!
```

```
Status of GitHub integration:
```

```
* --from-pr: OK
* --new-pr: OK
* --review-pr: OK
* --update-pr: OK
* --upload-test-report: OK
```

Checking configuration for **GitHub** integration

eb --check-github

```
$ eb --check-github
```

```
Checking status of GitHub integration...
```

```
Making sure we're online...OK
```

```
* GitHub user... boegel => OK
* GitHub token... e87..3b1 (len: 40) => OK (validated)
* git command... OK ("git version 1.8.3.1")
* GitPython module... OK (GitPython version 2.1.1)
* push access to boegel/easybuild-easyconfigs repo @ GitHub... OK
* creating gists... OK
* location to Git working dirs... OK (/home/example)
```

```
All checks PASSED!
```

```
Status of GitHub integration:
```

```
* --from-pr: OK
* --new-pr: OK
* --review-pr: OK
* --update-pr: OK
* --upload-test-report: OK
```

overview of **what is supported**
is produced based on
result of checks

Reviewing pull requests

eb --review-pr

- reviewer should ideally also compare new easyconfig files with existing ones for the same software
 - just a version bump or different toolchain?
 - also other changes?
 - to check whether fixes that are already in place in develop branch are not overlooked
- quite tedious to do manually...
- what are the most relevant easyconfigs to compare with?
 - same software (major) version, same toolchain (version), ...

Previewing & reviewing pull requests

eb --preview-pr

eb --review-pr

- (multi)diff between contributed easyconfig and existing one(s)
- can be used both to *preview* a contribution and to *review* a PR

```
$ eb --review-pr 5736
```

removed

added

```
Comparing BamTools-2.5.1-intel-2017b.eb with
```

```
BamTools-2.5.0-intel-2017b.eb
```

```
=====
```

```
2 - version = '2.5.0' (1/1)
```

```
2 + version = '2.5.1' (1/1)
```

comparison is done with (most) relevant
easyconfigs in current **develop** branch

```
-----
```

```
11 - checksums = [ 'dd4185bdba6e3adf2c24b7f93a57233d' ] (1/1)
```

```
11 + checksums = [ '98e90632058f85bd5eed6088b3ff912e' ] (1/1)
```

```
-----
```

Merging pull requests

eb --merge-pr

verifies whether PR is **eligible** to be merged, and merges it if so

```
$ eb --merge-pr 5722
```

```
easybuilders/easybuild-easyconfigs PR #5722 was submitted by vanzod,  
you are using GitHub account 'boegel'
```

```
Checking eligibility of easybuilders/easybuild-easyconfigs PR #5722  
for merging...
```

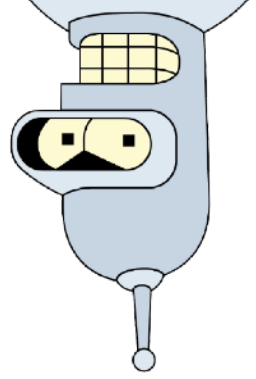
```
* targets develop branch: OK  
* test suite passes: OK  
* last test report is successful: OK  
* approved review: OK (by boegel)  
* milestone is set: OK (3.6.0)
```

merging pull requests can be
done **from the command line**
(after approved visual review)

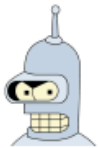
```
Review OK, merging pull request!
```

```
Adding comment to easybuild-easyconfigs issue #5722:  
'Going in, thanks @vanzod!'
```

boegelbot



- bot that creates comments in PRs if tests fail in Travis
- mainly to trigger **GitHub** to send notifications
- see `pr_check.py` in <https://github.com/boegel/eb-scripts>



boegelbot commented on 8 Sep 2017



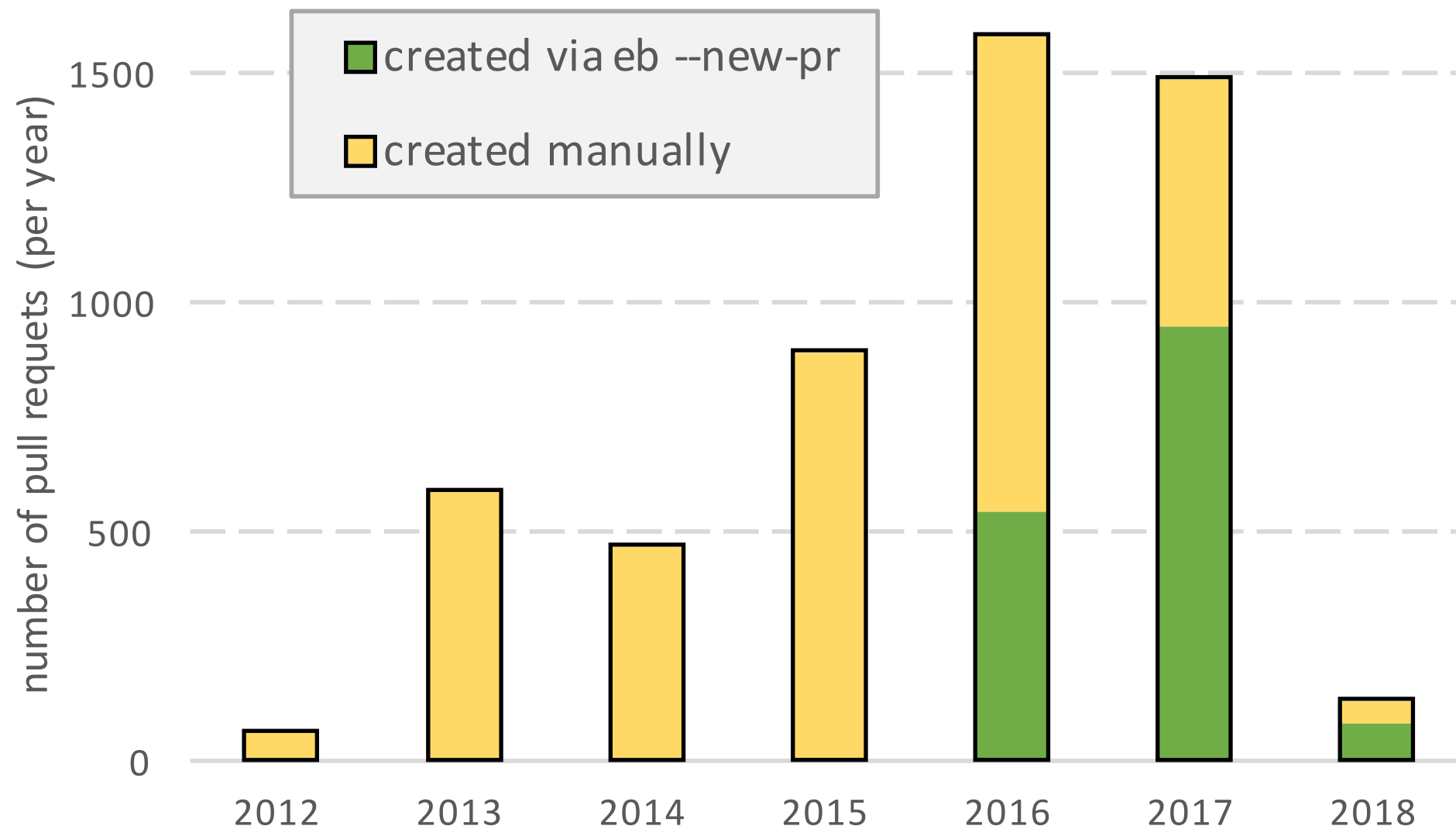
Travis test report: 8/8 runs failed - see <https://travis-ci.org/easybuilders/easybuild-easyconfigs/builds/273198653>

Only showing partial log for 1st failed test suite run 5974.1;
full log at <https://travis-ci.org/easybuilders/easybuild-easyconfigs/jobs/273198655>

```
...
ERROR: Test for parsing of easyconfig Stata-15.eb
-----
Traceback (most recent call last):
  File "<string>", line 1, in innertest
  File "/home/travis/build/easybuilders/easybuild-easyconfigs/test/easyconfigs/easyconfigs.py", line
    ecs = process_easyconfig(spec)
```

Impact on contributions

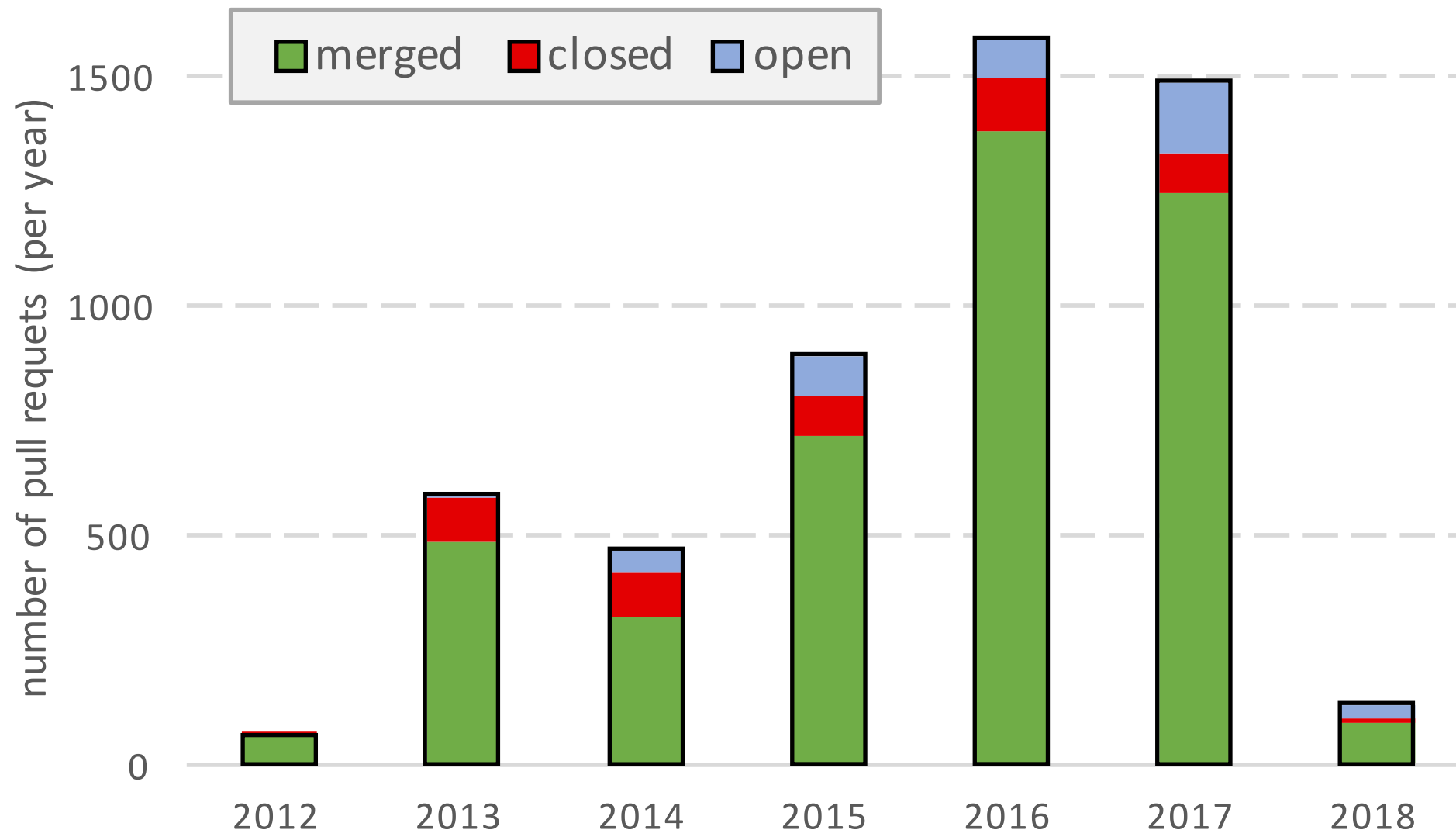
<https://github.com/easybuilders/easybuild-easyconfigs/pulls>



- > 500 PRs opened using eb --new-pr in 2016 (34%)
- almost 1000 PRs in 2017 (63%)
- 80% more PRs in 2016/2017 compared to 2015

Impact on contributions

<https://github.com/easybuilders/easybuild-easyconfigs/pulls>



- **GitHub** integration has helped (a lot) to keep pace with incoming PRs
- ~85% of PRs gets merged (usually a good reason for the ones that are not)
- would have been difficult without automation due to amount of PRs


Conclusions

-  **git** can be a (major) hurdle for potential contributors


- **contribution procedure can be automated**

via `GitPython` (or equivalent library) +  REST API
GitHub

- contributors don't *need* to use  **git** (directly) anymore!

-  **GitHub** integration in EasyBuild has resulted in:
 - more contributions
 - being able to keep pace with incoming contributions
 - **less "wasted" time for both contributors & maintainers**

Simplifying the contribution process for both contributors & maintainers

A case study of the  integration in  easybuild

GitHub



FOSDEM 2018

Testing and Automation devroom

Feb 3rd 2018, Brussels (Belgium)

Kenneth Hoste

kenneth.hoste@ugent.be

GitHub: @boegel

Twitter: @kehoste