PeepCode press

$9

Git Internals

Source code control and beyond

by Scott Chacon

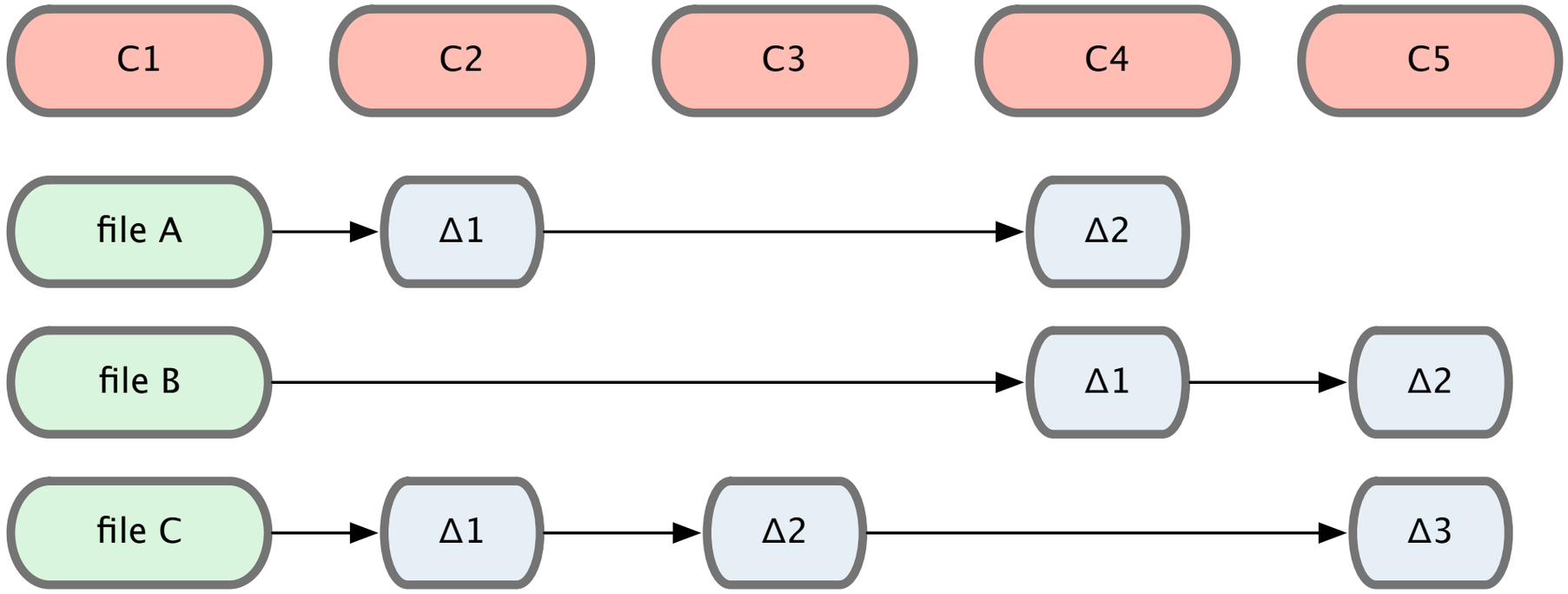
What is Git?

git is

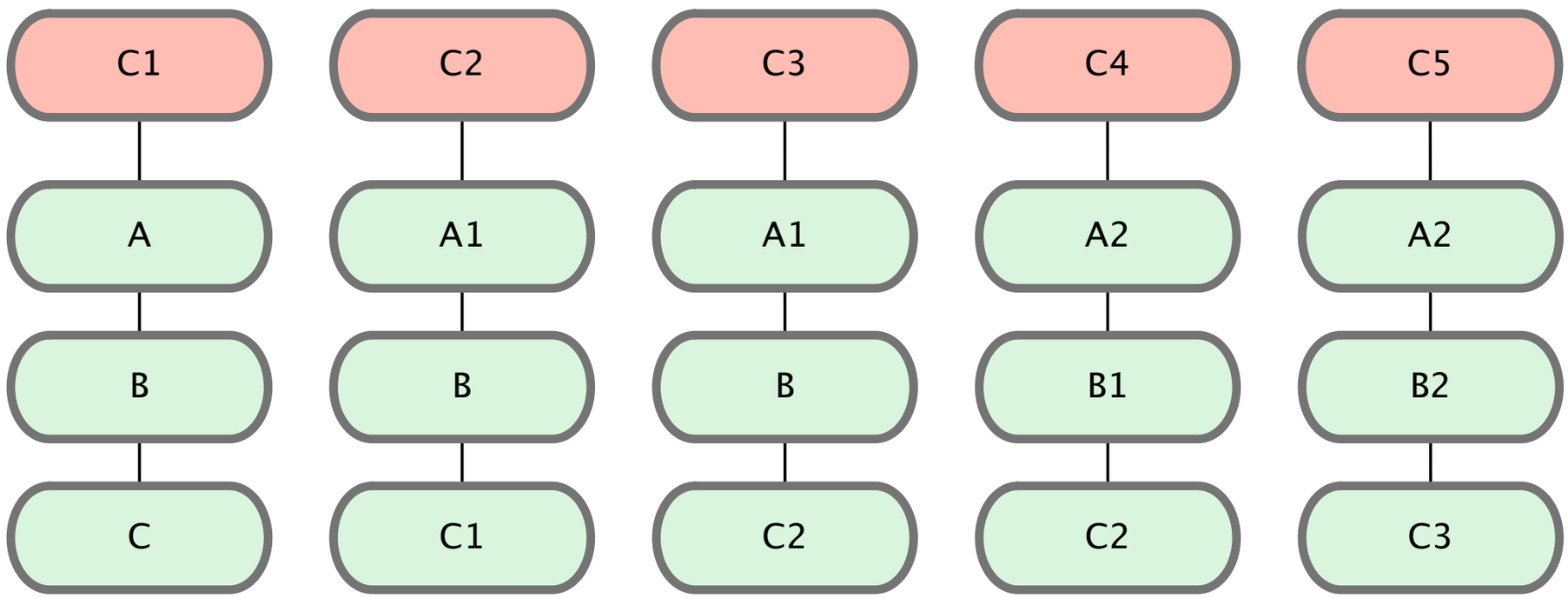
directory content management system

source control taxonomy

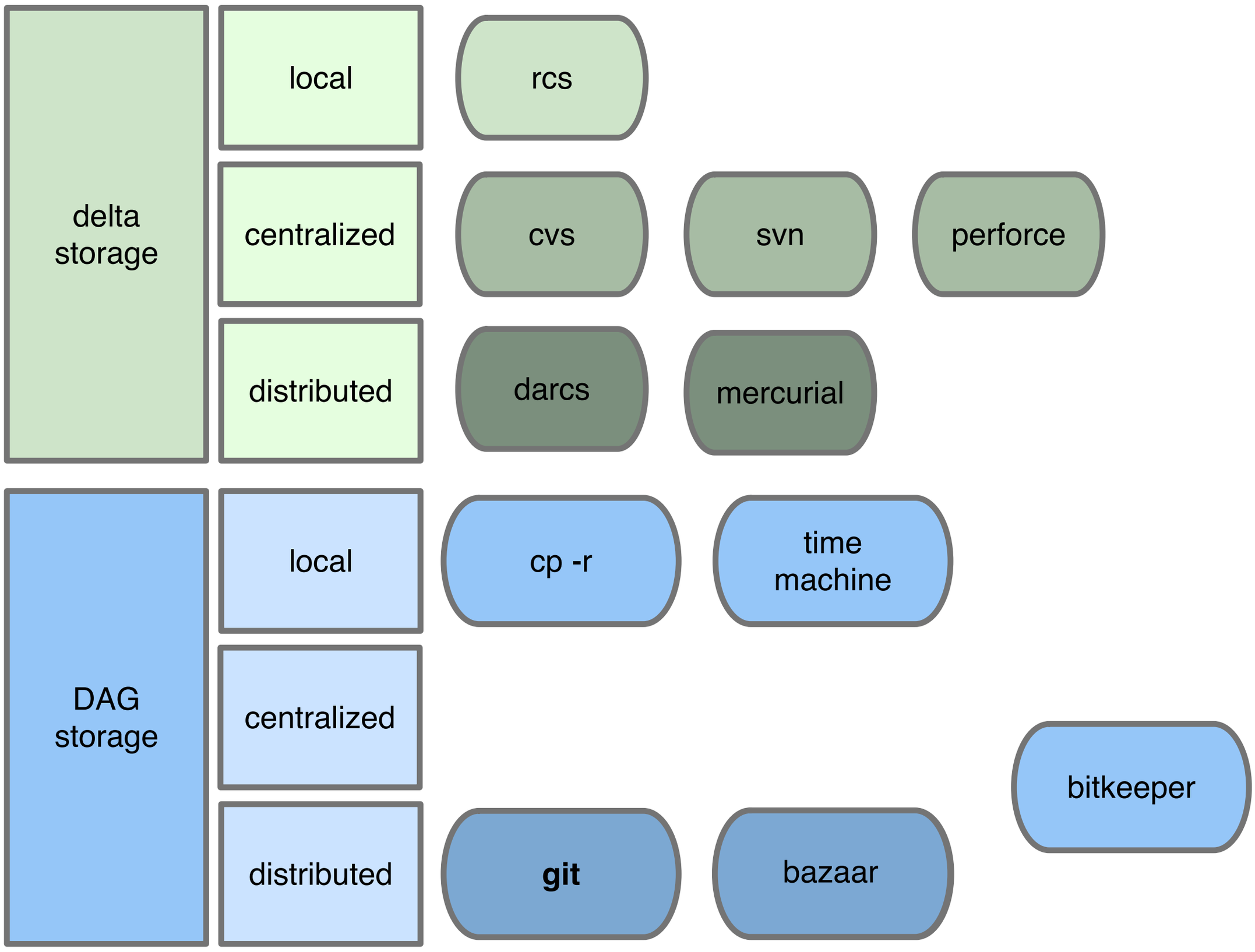
delta storage



DAG storage



source control taxonomy



How Does Git Work?

Git Directory

config file

hooks

index

object database

references

object database

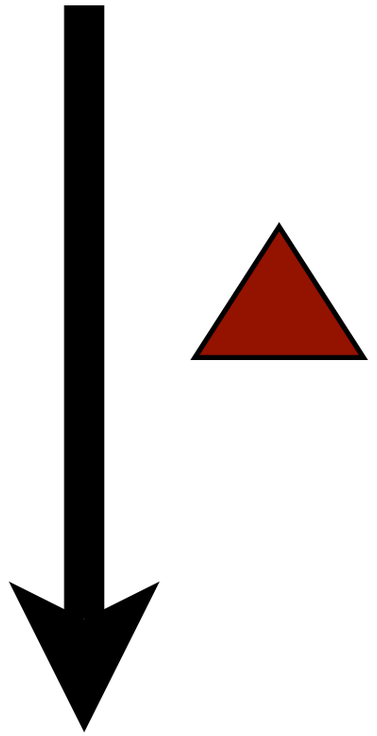
```
content

new_content = type + ' ' + content.size + \0
                + content

sha = Digest::SHA1.hexdigest(new_content)

"824aed035c0aa75d64c..."

compressed = zlib::deflate(new_content)

path = ".git/objects/82/4aed035c0aa75d64c..."

File.open(path, 'w') {|f| f.write(compressed)}
```

object database

`git gc`

same file with minor differences

```
.git/objects/82/4aed035c0aa75d64c...
.git/objects/1d/c9cbcb76cbb80fce1...
.git/objects/63/874f37013c1740acd...
.git/objects/04/fb8aee105e6e445e8...
.git/objects/45/b983be36b73c0788d...
.git/objects/f1/032eed02413a1145c...
```



```
.git/objects/pack/pack-999727..9f600.pack
.git/objects/pack/pack-999727..9f600.idx
```

object database

```
new_content = type + ' ' + content.size + \0
                   + content
```

4 types of git objects

object database

blob

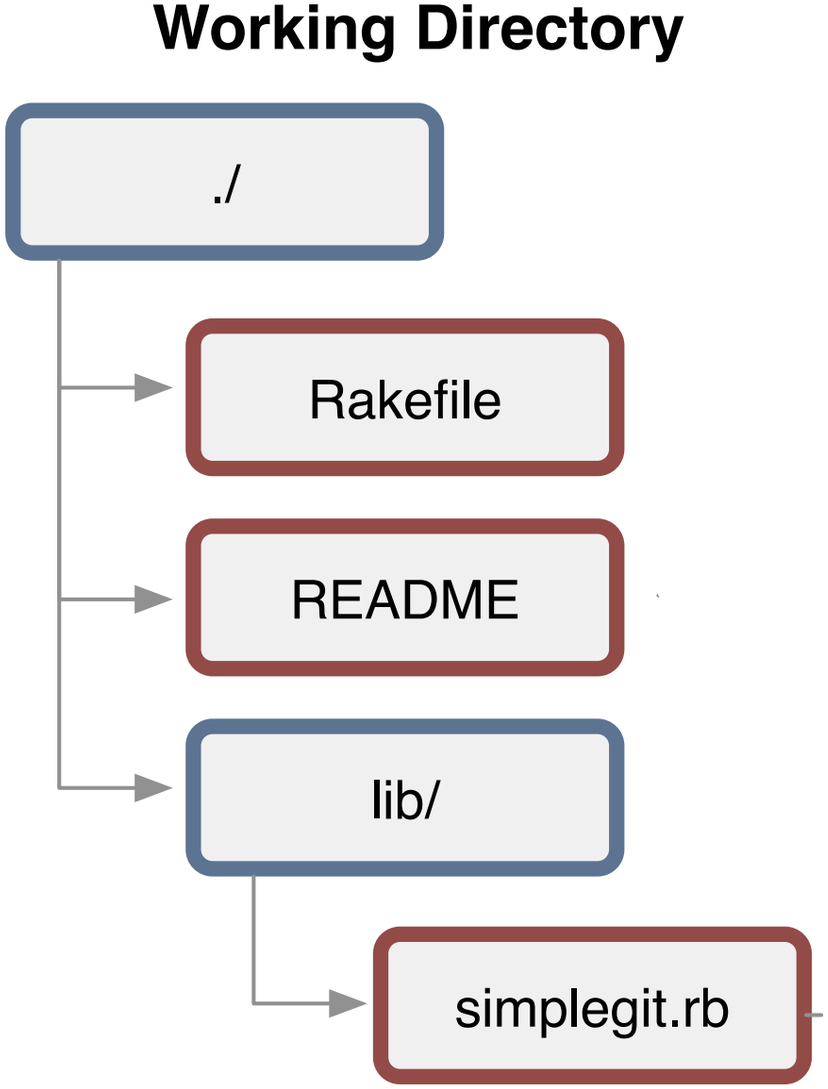
tree

commit

tag

object database

blob



object database

object

compress

header

content

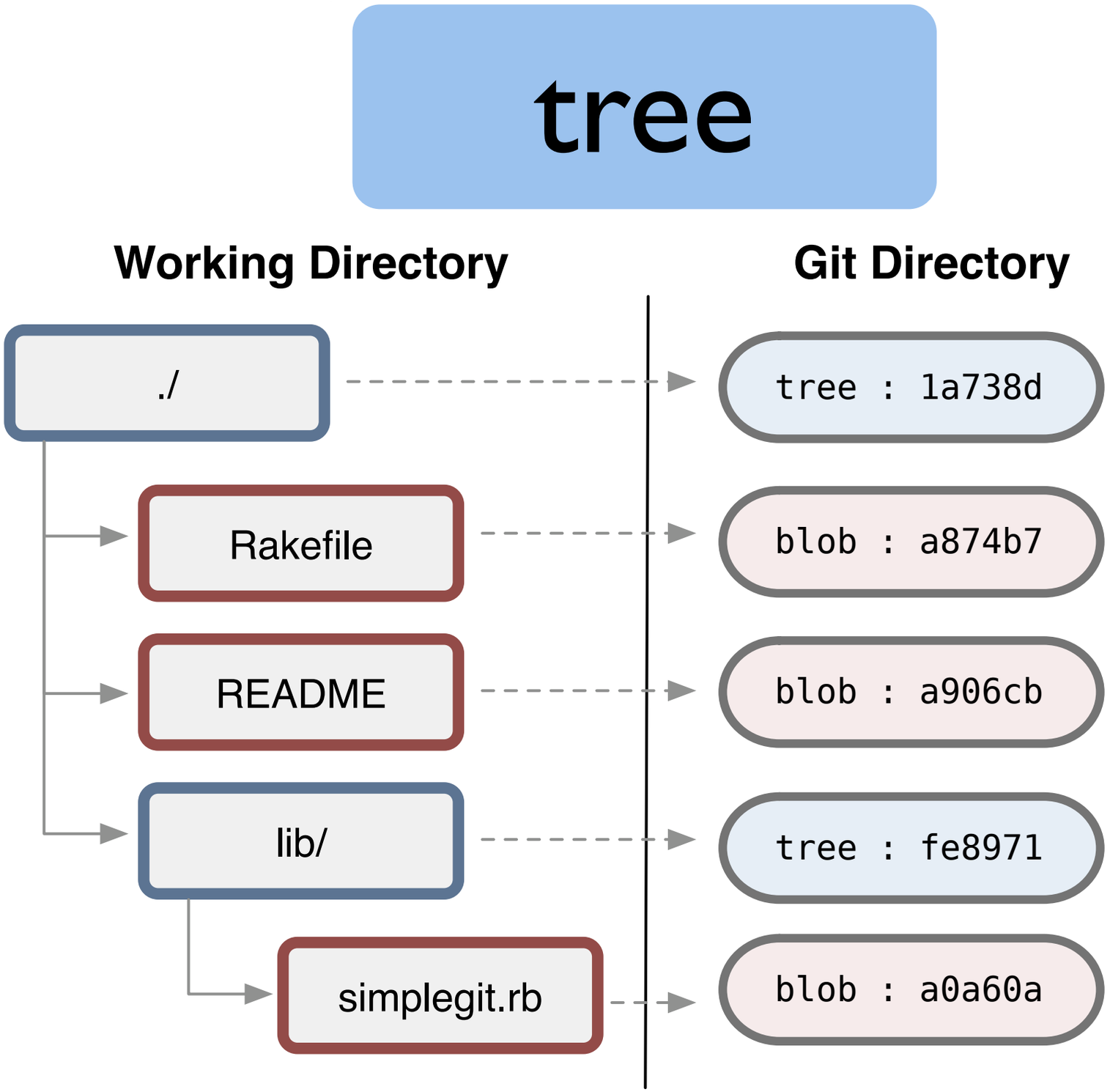
```
blob : a906cb
```

zlib::deflate

```
blob [content size]\0
```

```
SimpleGit Ruby Library
======================

This library calls git commands and
returns the output.

Author : Scott Chacon
```

object database



object database

```
tree : 1a738d
```

zlib::deflate

```
tree [content size]\0

100644 blob a906cb    README
100644 blob a874b7    Rakefile
040000 tree fe8971    lib
```

“inode” info

filename

object database

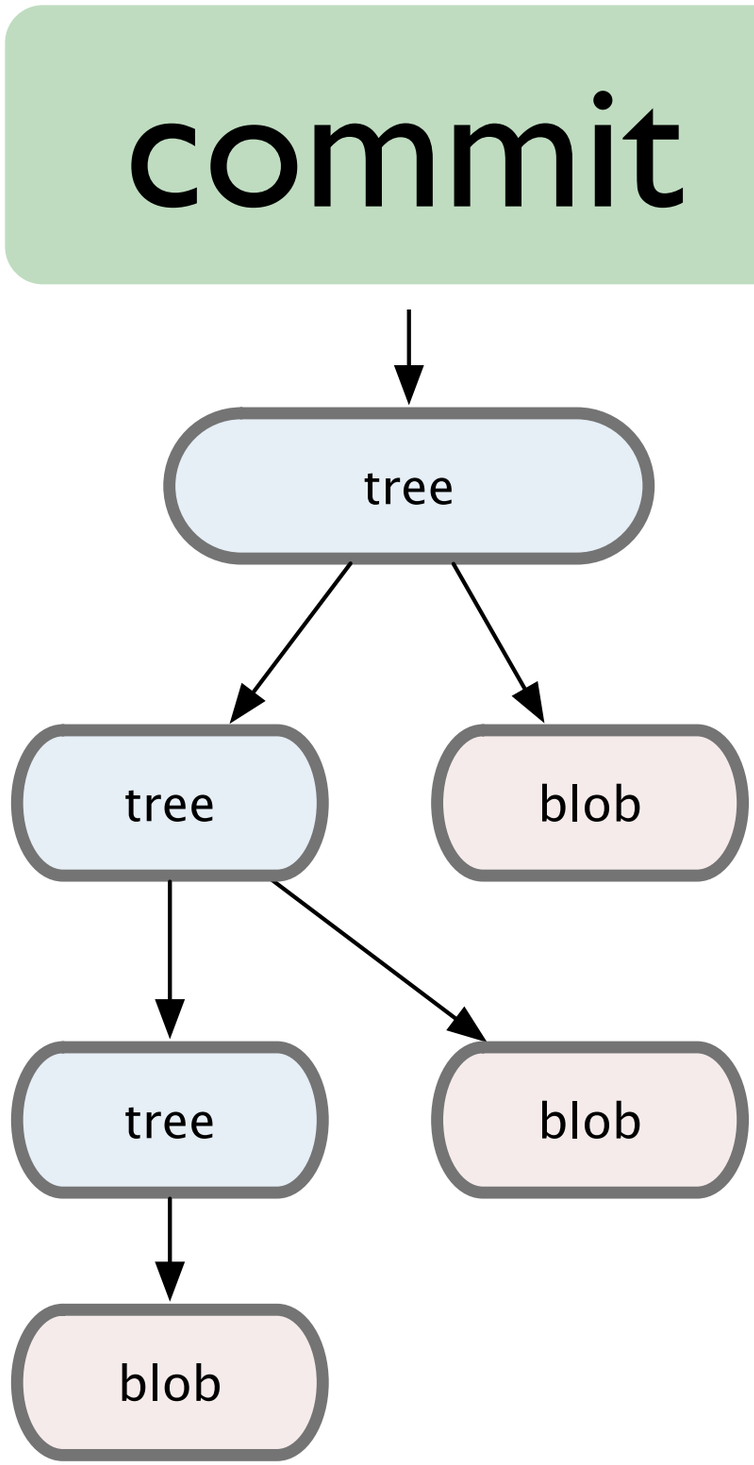
```
tree : 1a738d
```

zlib::deflate

```
tree [content size]\0

100644 blob a906cb    README
100644 blob a874b7    Rakefile
040000 tree fe8971    lib
```

mode type “block pointer”

object database



object database

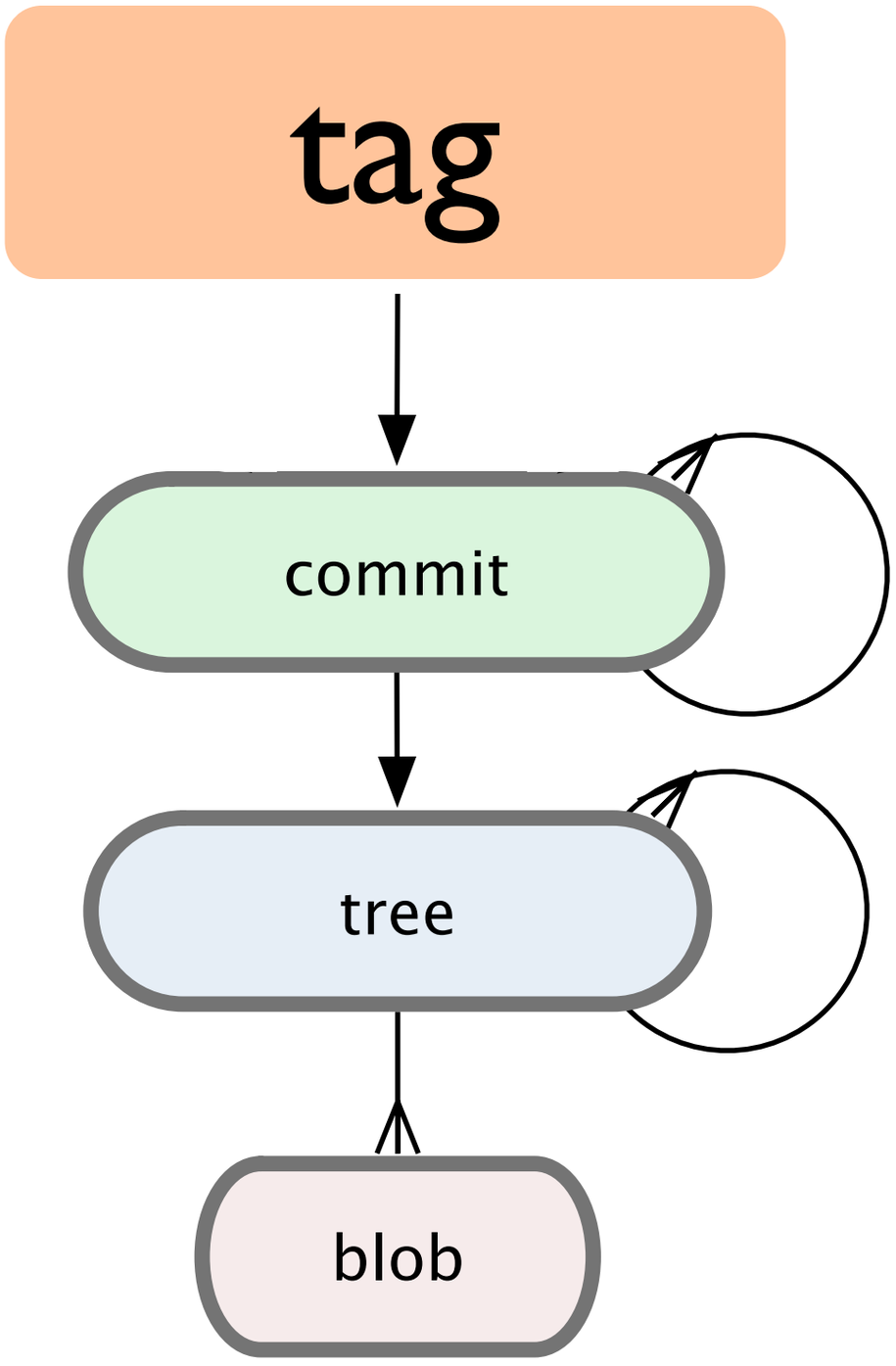
commit : e1b3ec

zlib::deflate

```
commit [content size]\0
```

```
tree e1b3ec
parent a11bef
author Scott Chacon
    <schacon@gmail.com> 1205624433
committer Scott Chacon
    <schacon@gmail.com> 1205624433

my second commit, which is better than
the first
```

object database



object database

tag : 0c819c

zlib::deflate

```
tag [content size]\0
```

```
object 0576fa
type commit
tag v0.1
tagger Scott Chacon
          <schacon@gmail.com> 1205624655

this is my v0.1 tag
```

object database

blob

tree

immutable

commit

tag

Git Directory

config file

hooks

index

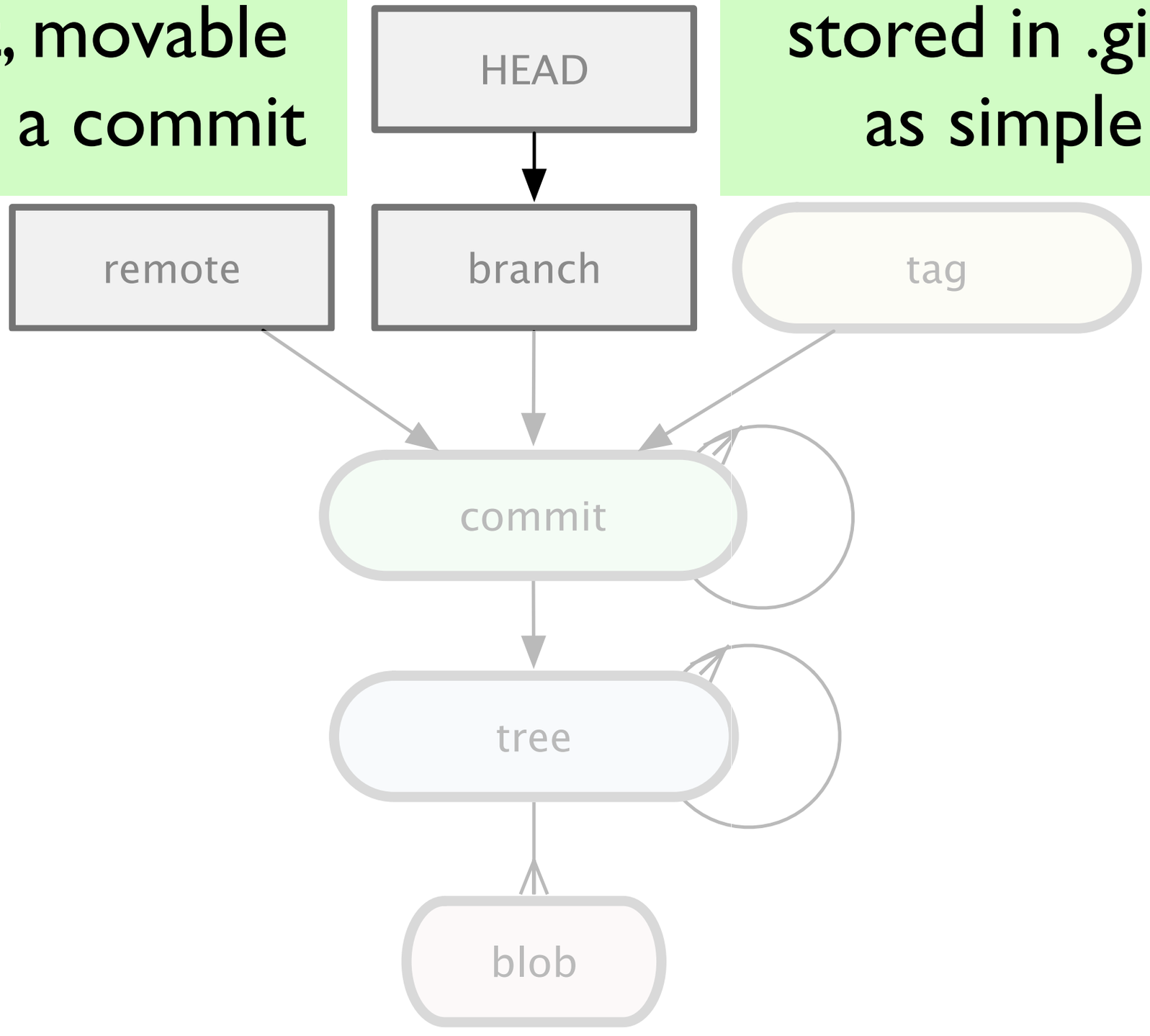
object database

references

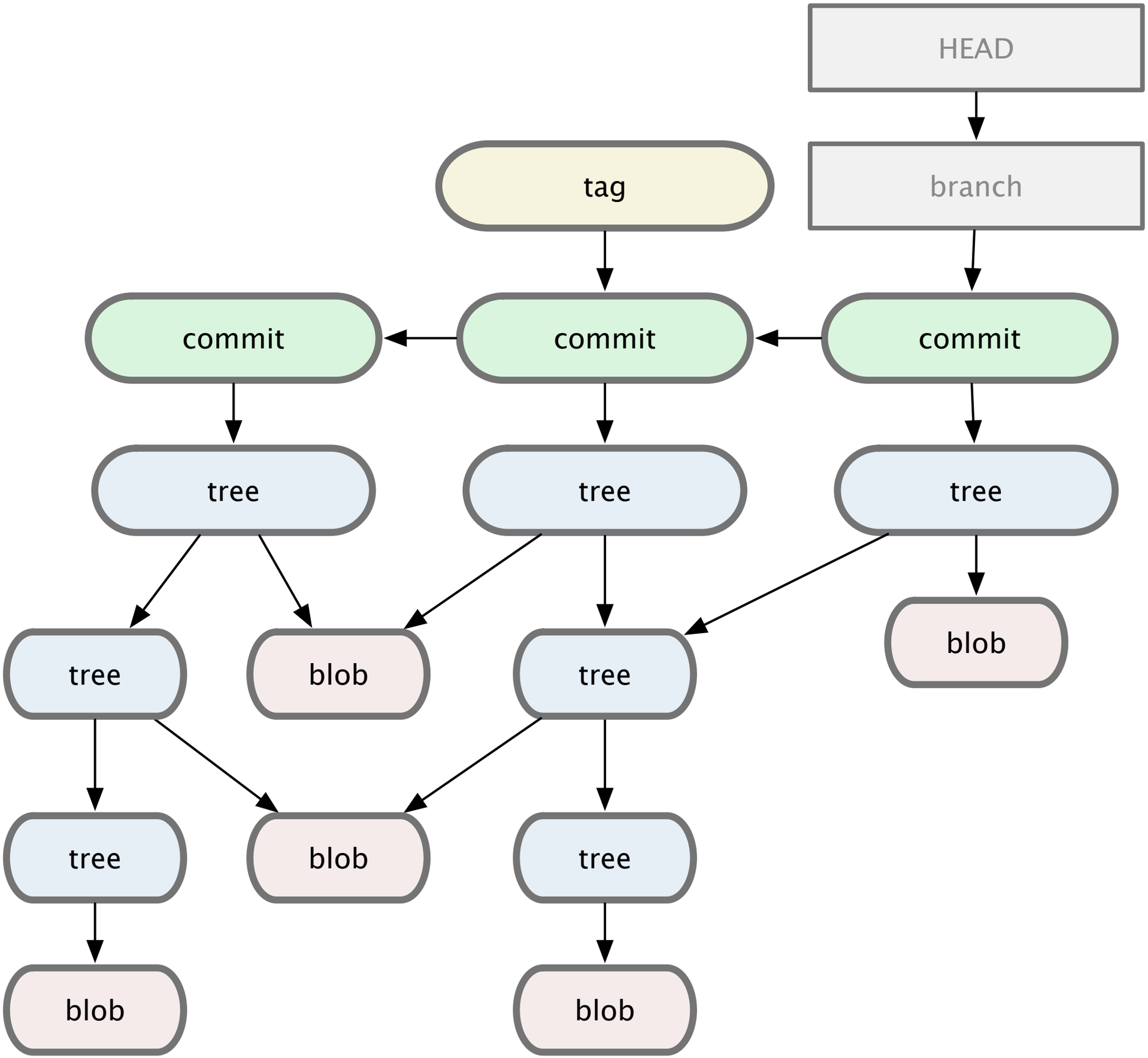
references

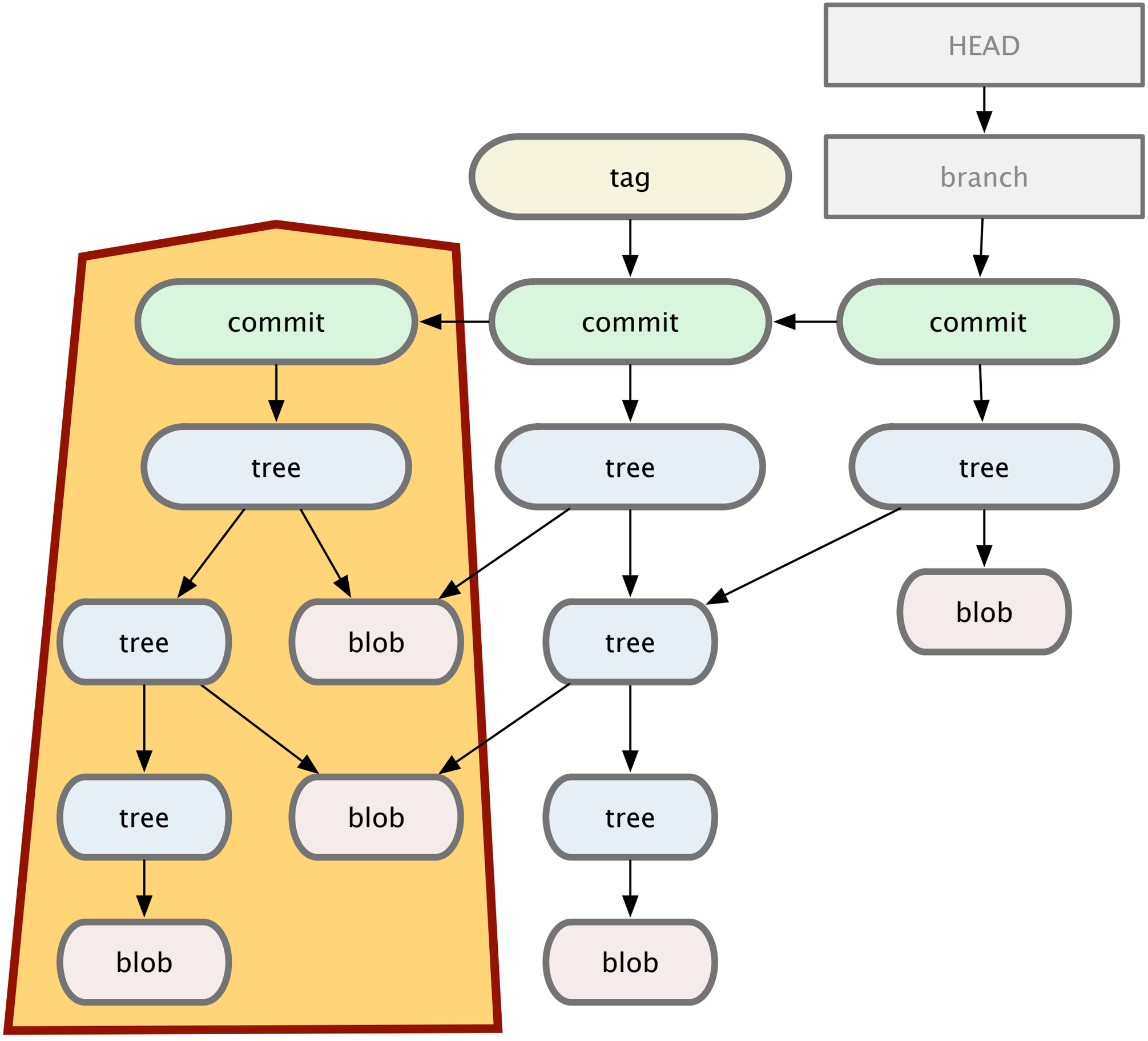
lightweight, movable pointers to a commit

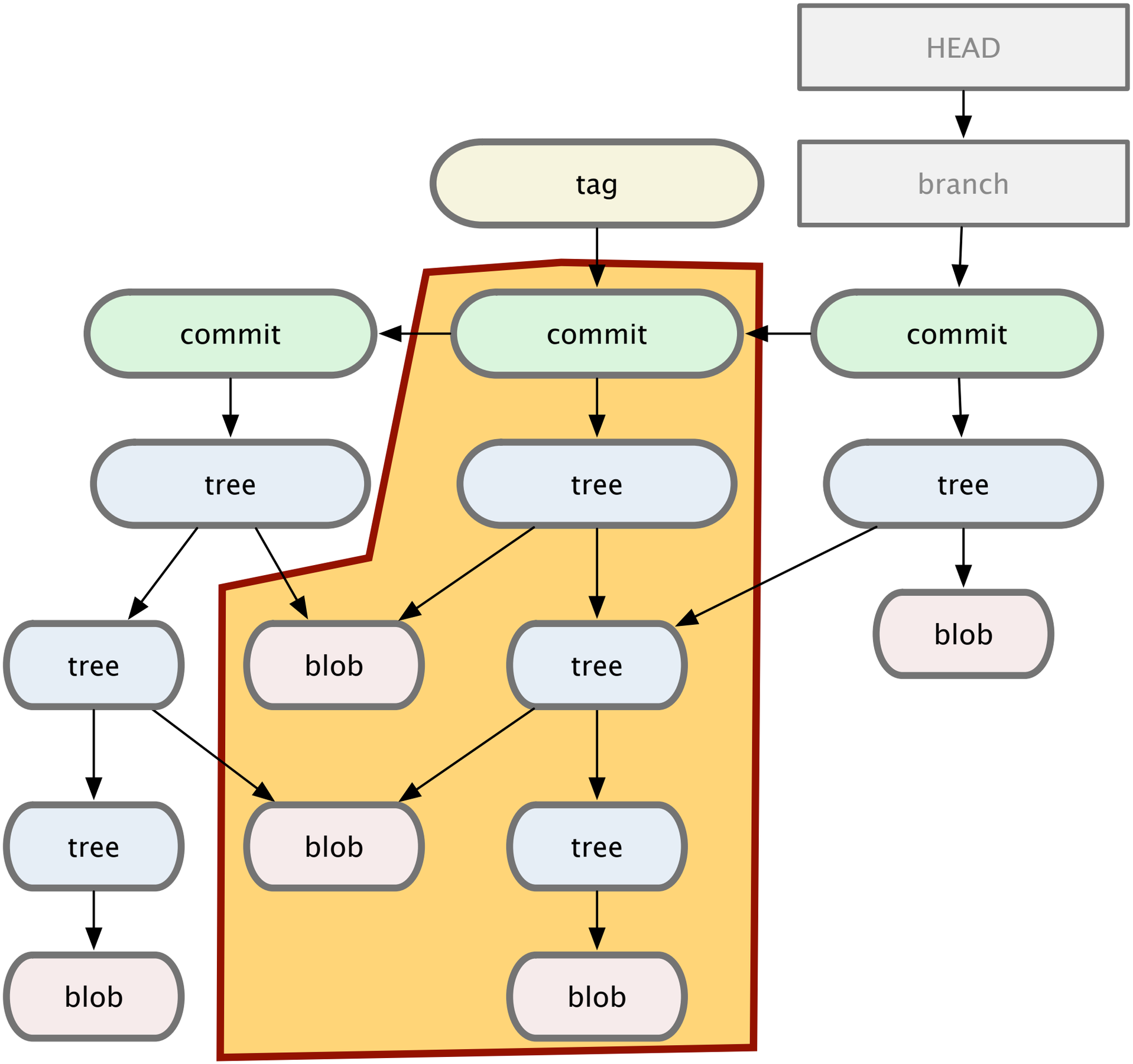
stored in .git/refs/* as simple files

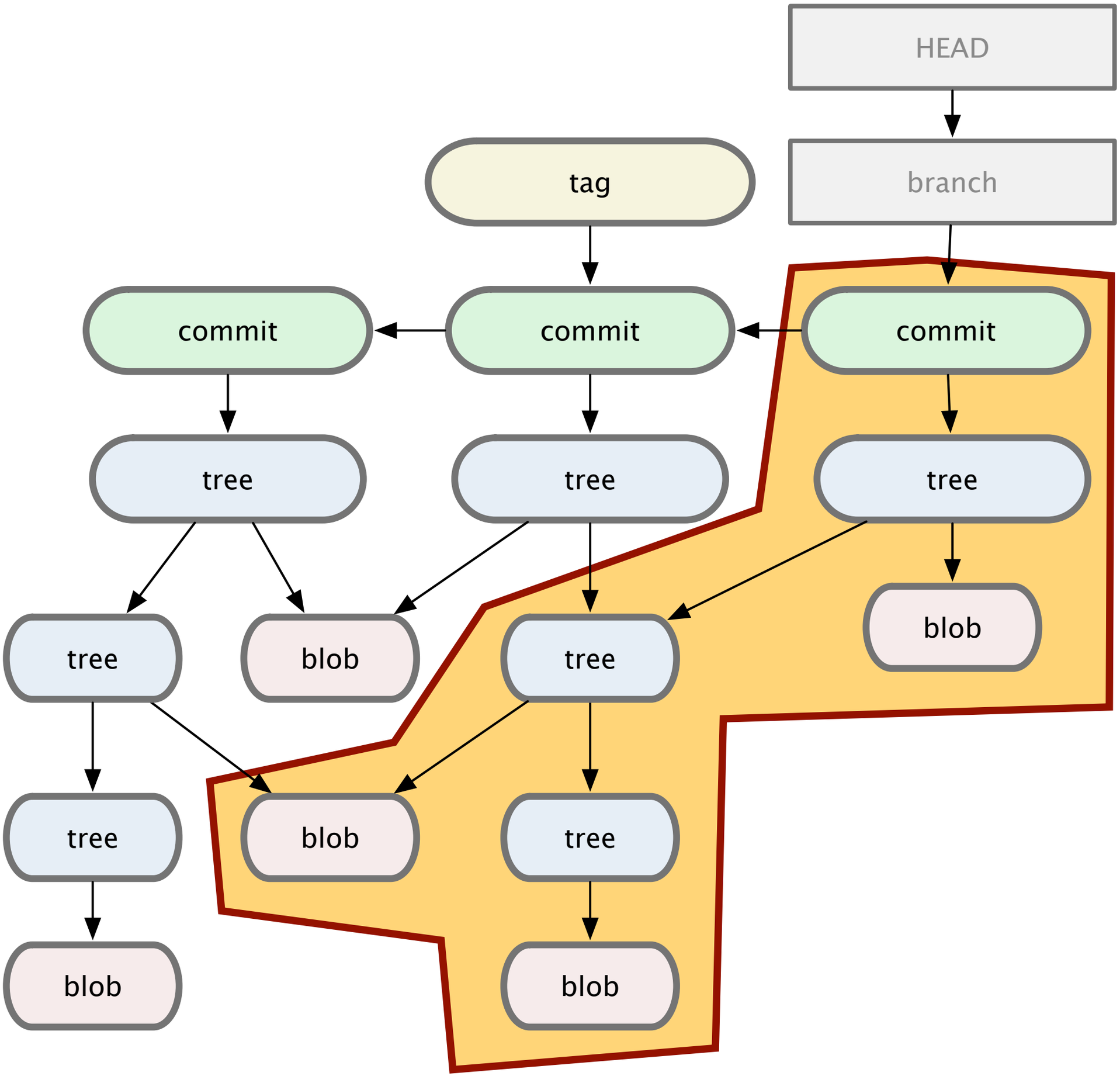


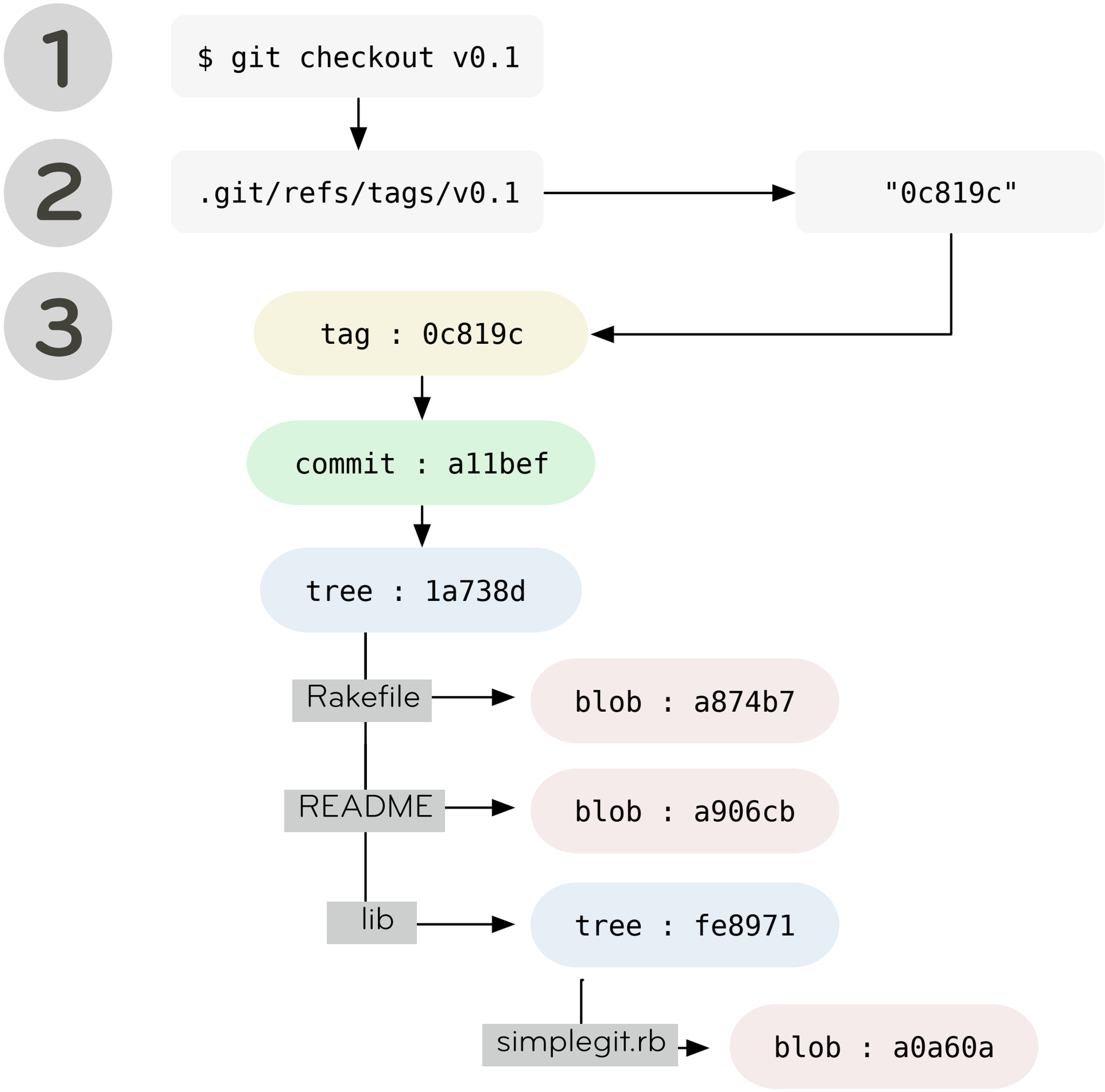
the object model


HEAD
branch
tag
commit
commit
commit
tree
tree
tree
tree
blob
tree
blob
tree
blob
blob
blob

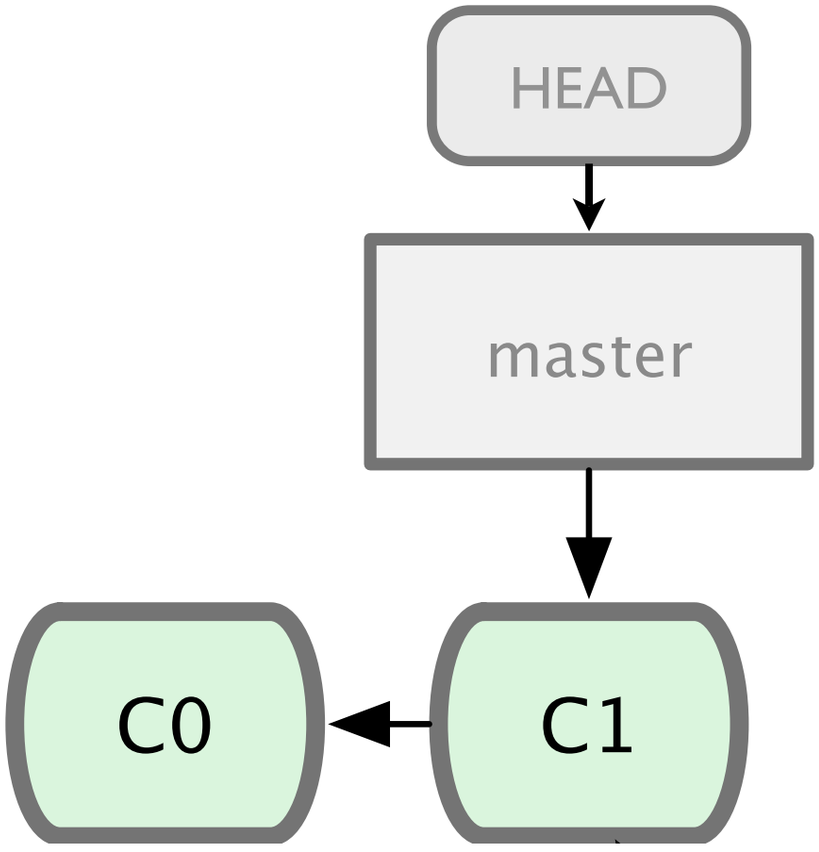

HEAD
branch
tag
commit
commit
commit
tree
tree
tree
tree
blob
tree
blob
tree
blob
tree
blob
blob

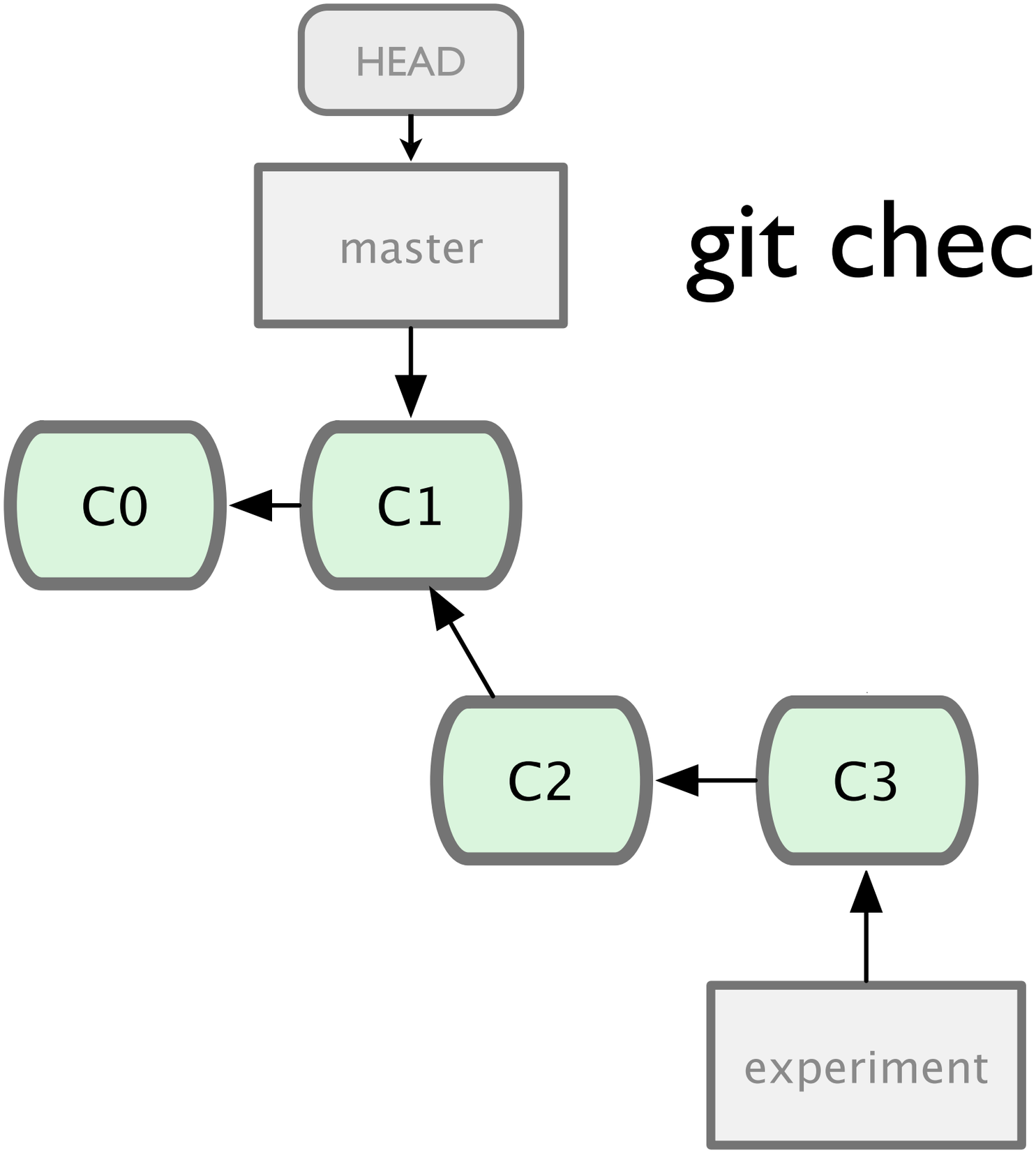

HEAD
branch
tag
commit
commit
commit
tree
tree
tree
tree
blob
tree
blob
tree
blob
blob
tree
blob


HEAD
branch
tag
commit
commit
commit
tree
tree
tree
tree
blob
blob
tree
tree
blob
tree
blob
blob

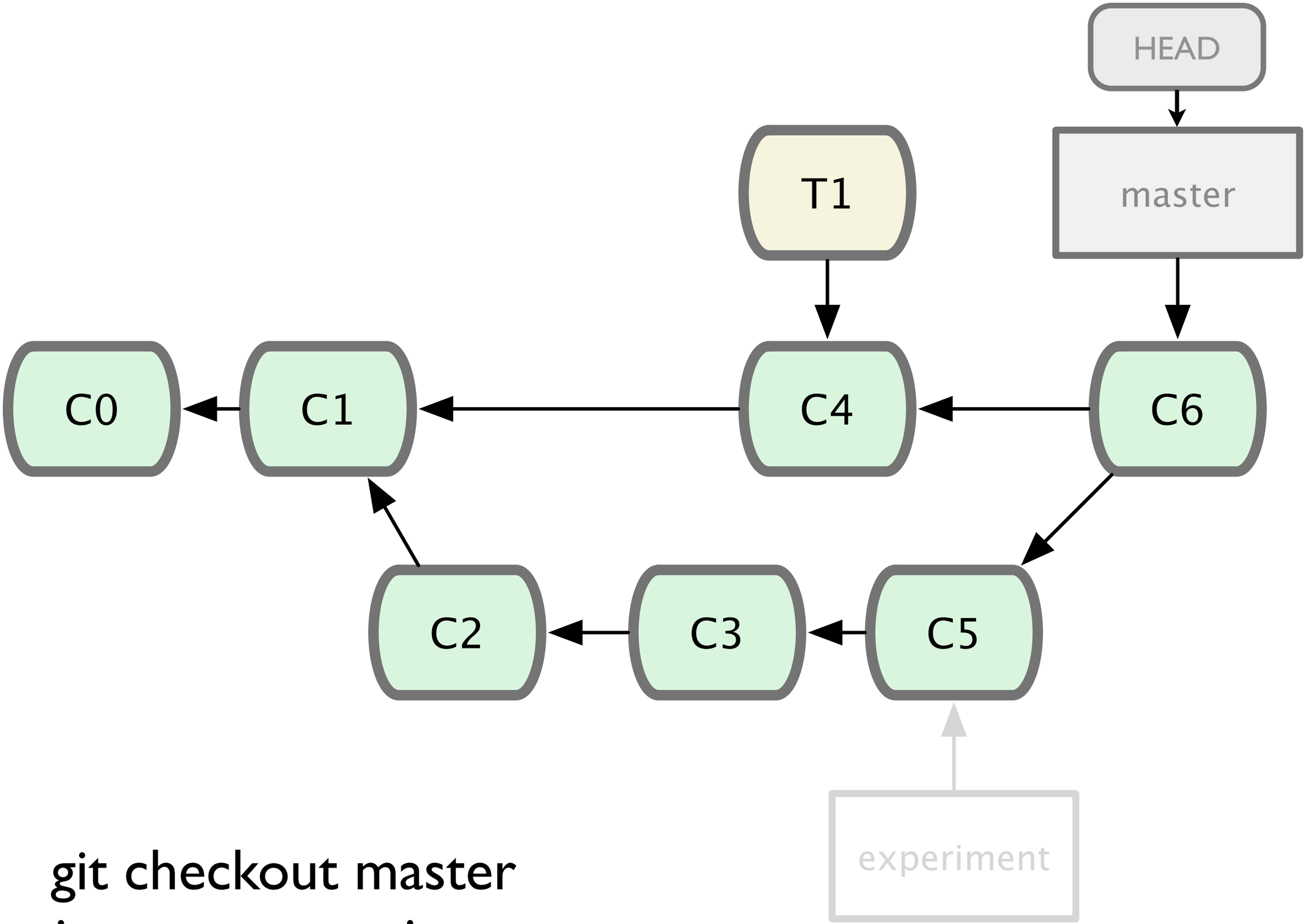

1
$ git checkout v0.1
2
.git/refs/tags/v0.1
"0c819c"
3
tag : 0c819c
commit : a11bef
tree : 1a738d
Rakefile
blob : a874b7
README
blob : a906cb
lib
tree : fe8971
simplegit.rb
blob : a0a60a

branching and merging


HEAD
master
C0
C1



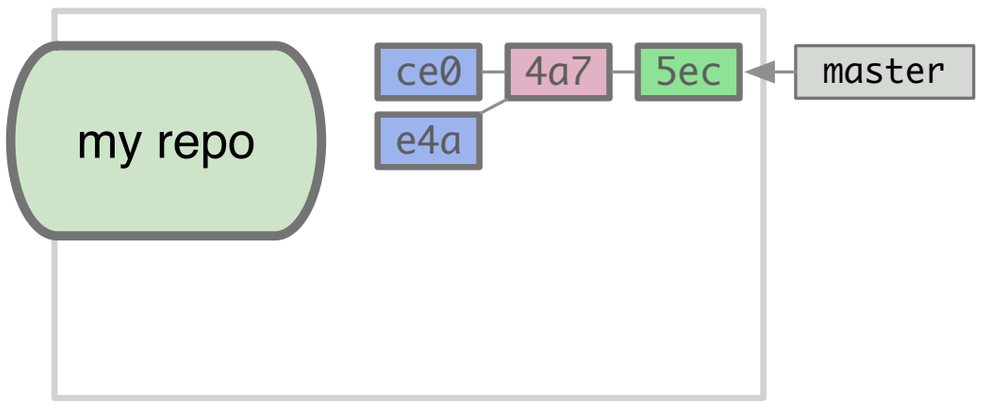
git checkout master



git checkout master
git merge experiment

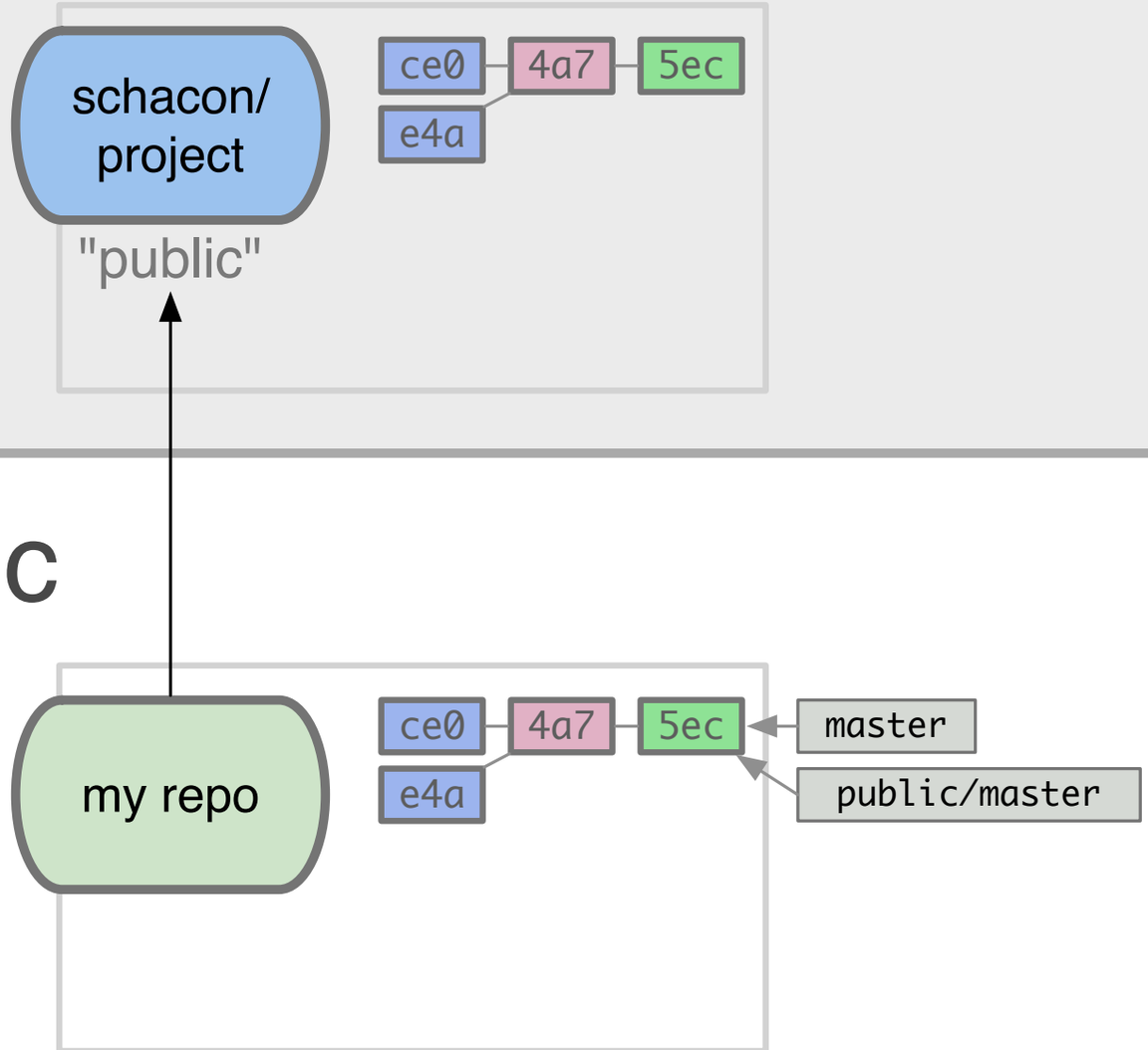
developer
nick

developer
jessica

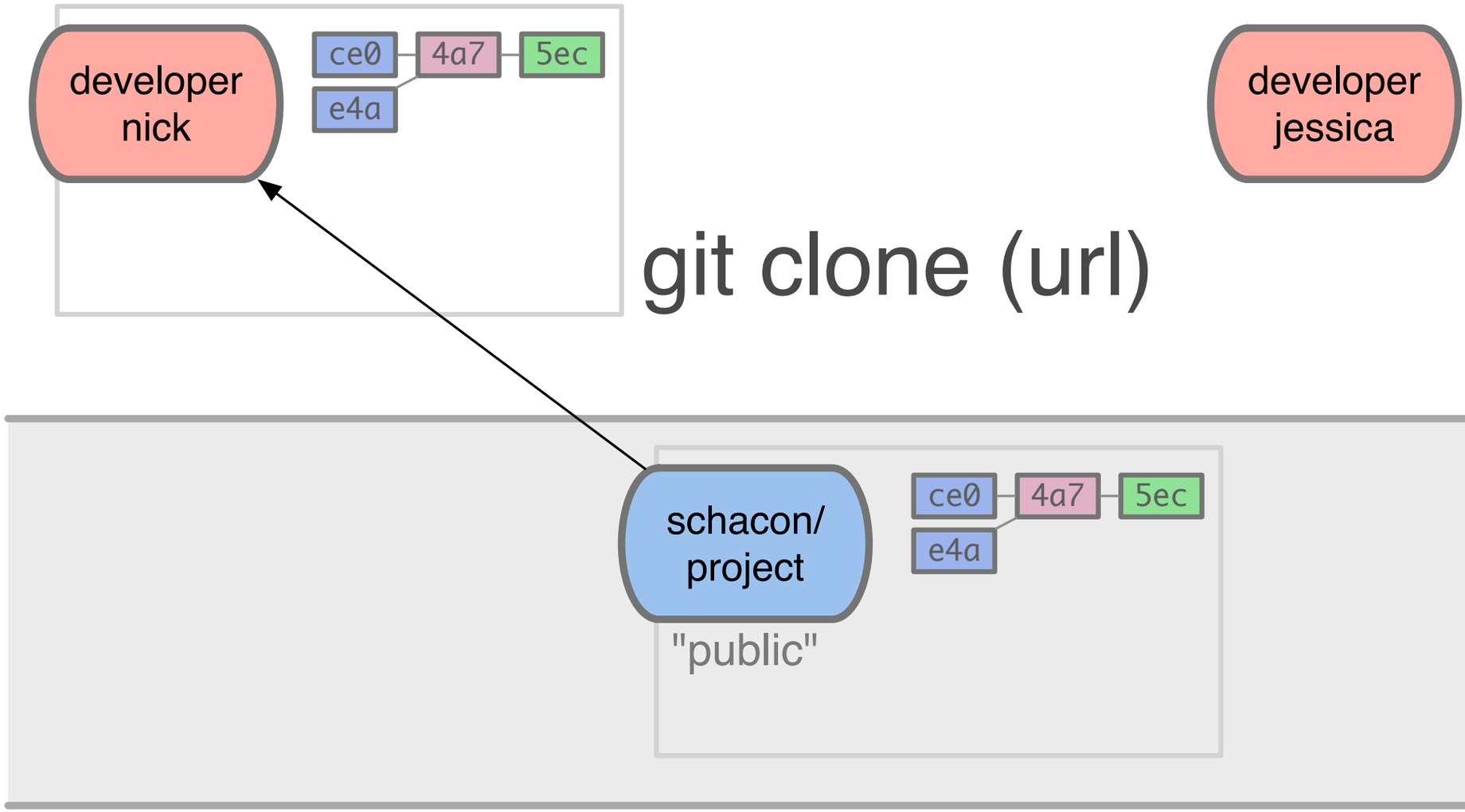


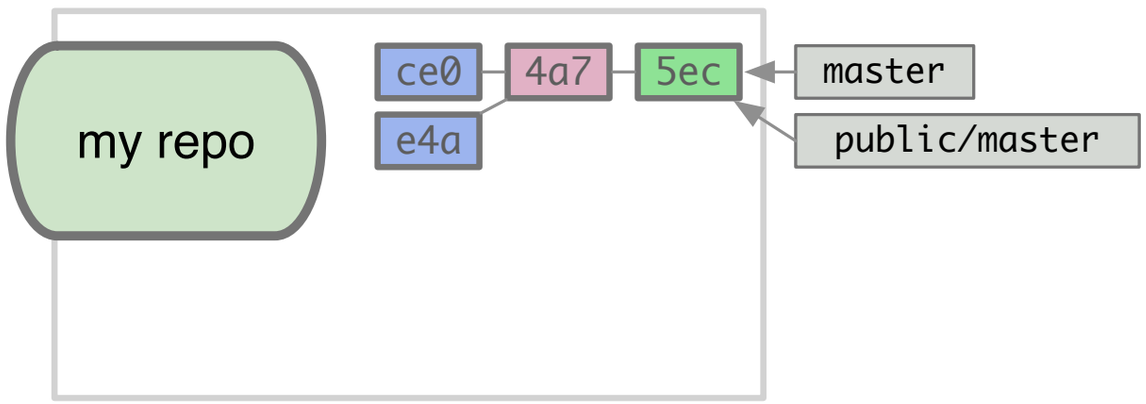
developer nick

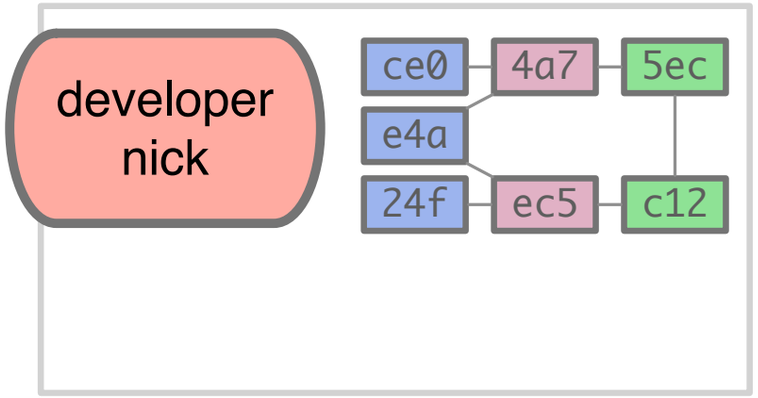
developer jessica



git push public

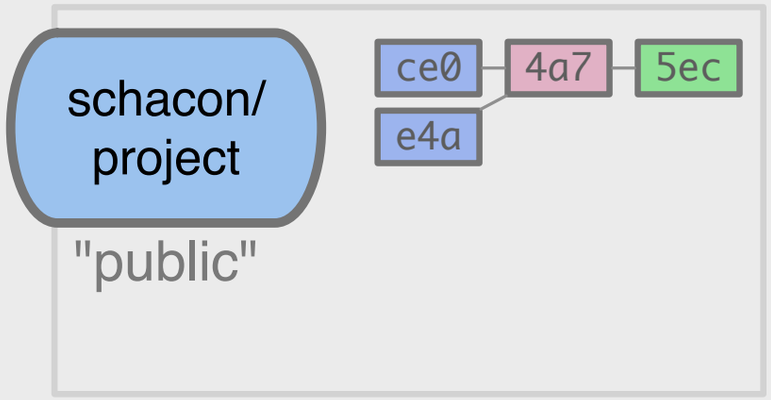

developer
nick
ce0
4a7
5ec
e4a
developer
jessica
git clone (url)
schacon/
project
ce0
4a7
5ec
e4a
"public"

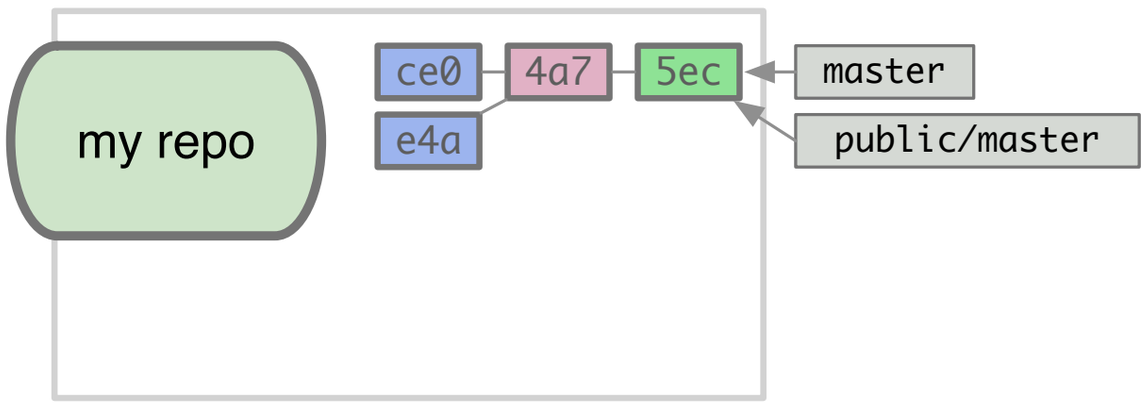

my repo
ce0
4a7
5ec
e4a
master
public/master

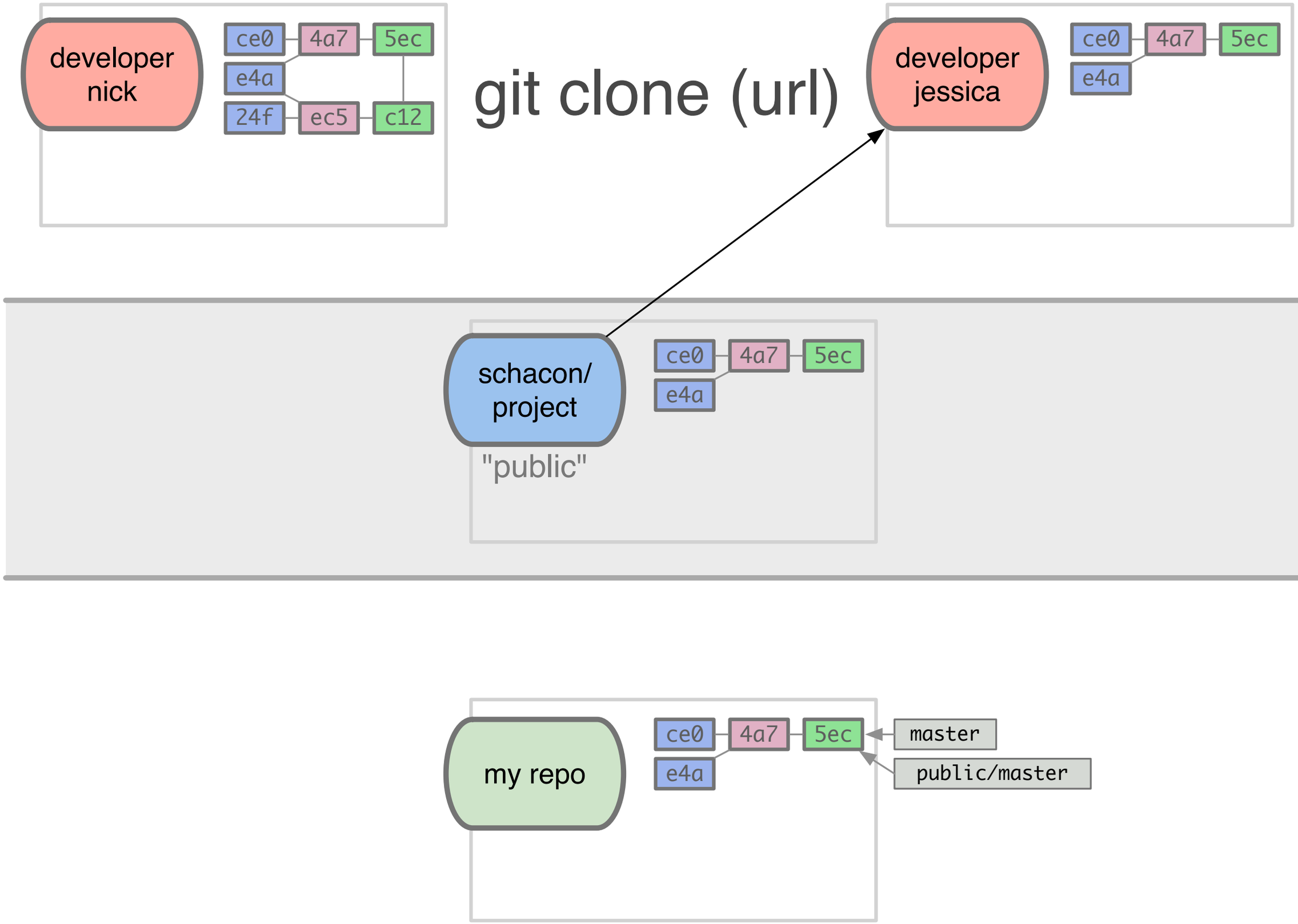

developer
nick
ce0
4a7
5ec
e4a
24f
ec5
c12

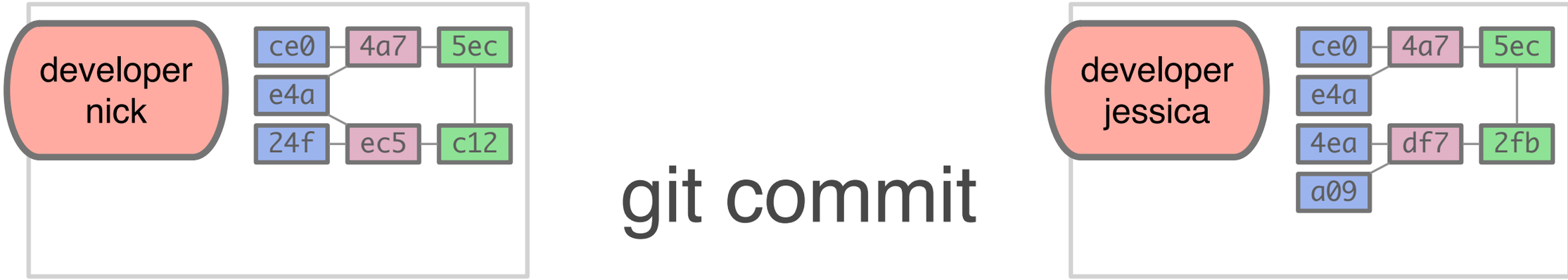
developer
jessica

git commit


schacon/
project
"public"
ce0
4a7
5ec
e4a

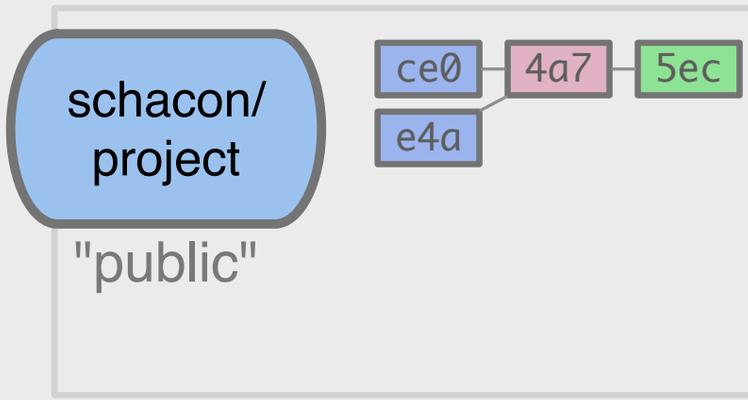

my repo
ce0
4a7
5ec
e4a
master
public/master

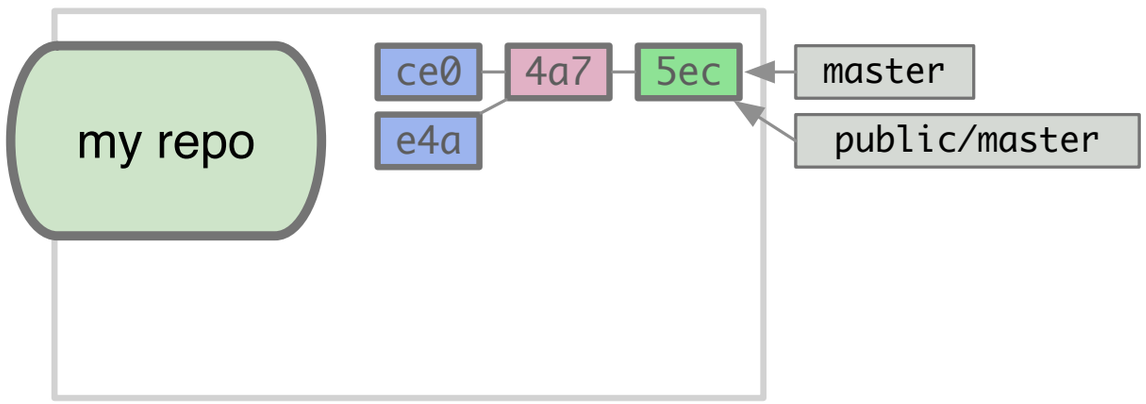

developer
nick
ce0
4a7
5ec
e4a
24f
ec5
c12
git clone (url)
developer
jessica
ce0
4a7
5ec
e4a
schacon/
project
ce0
4a7
5ec
e4a
"public"
my repo
ce0
4a7
5ec
e4a
master
public/master

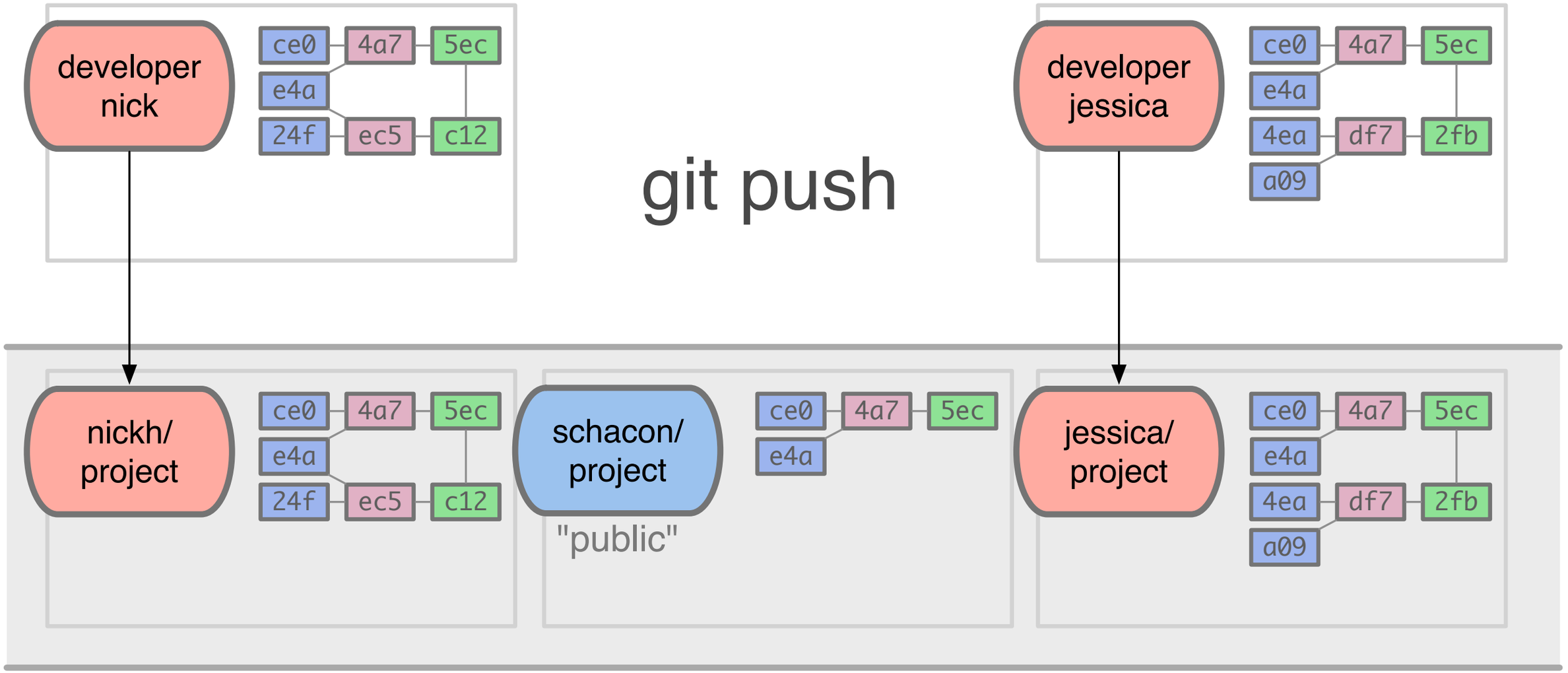

developer
nick
ce0
4a7
5ec
e4a
24f
ec5
c12

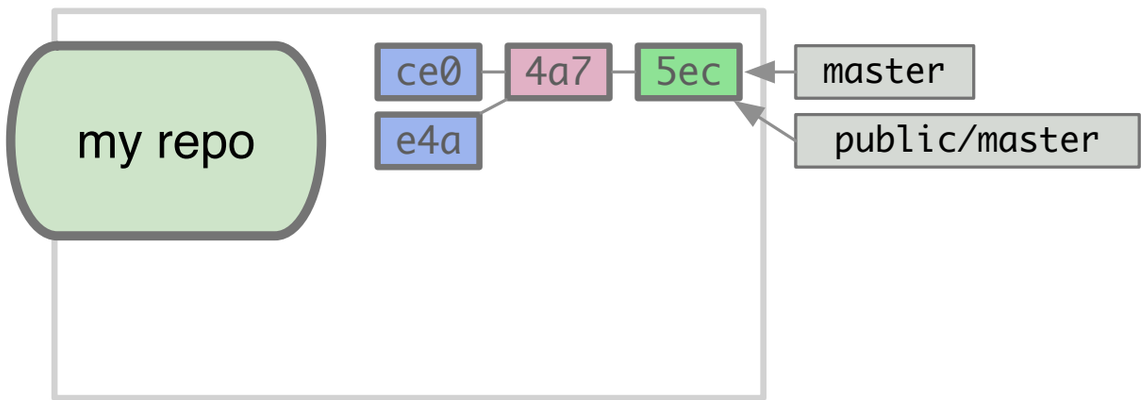
git commit

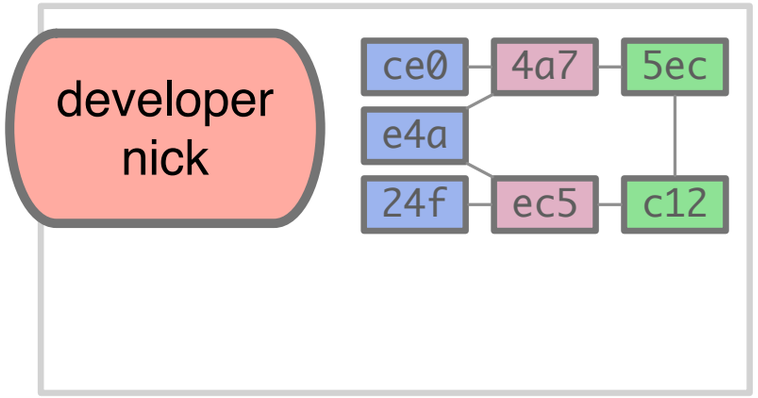
developer
jessica
ce0
4a7
5ec
e4a
4ea
df7
2fb
a09

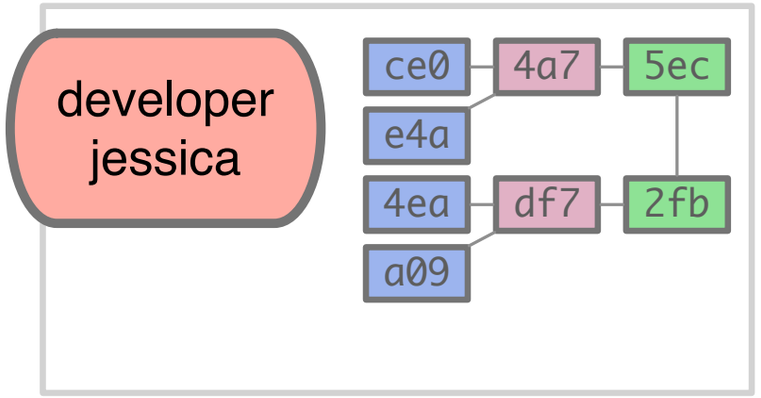

schacon/
project
"public"
ce0
4a7
5ec
e4a

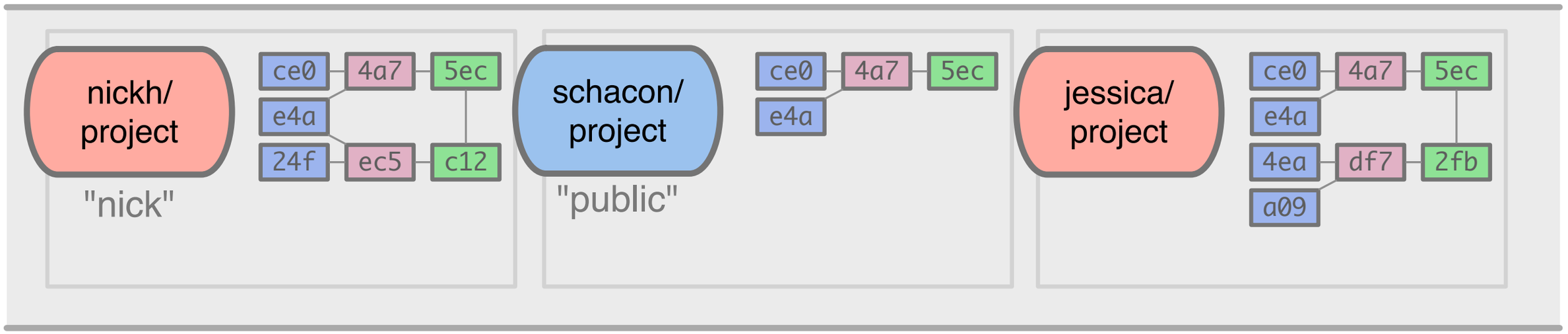

my repo
ce0
4a7
5ec
e4a
master
public/master

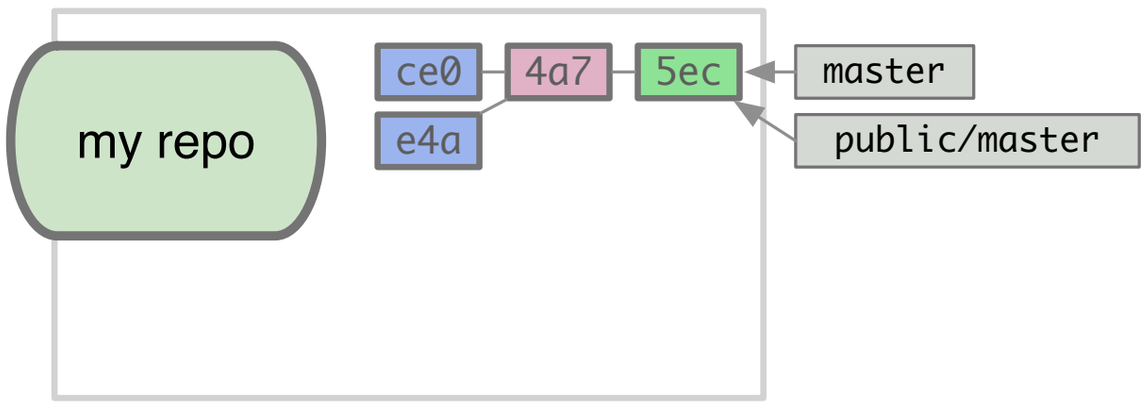

developer
nick
ce0
4a7
5ec
e4a
24f
ec5
c12
git push
developer
jessica
ce0
4a7
5ec
e4a
4ea
df7
2fb
a09
nickh/
project
ce0
4a7
5ec
e4a
24f
ec5
c12
schacon/
project
"public"
ce0
4a7
5ec
e4a
jessica/
project
ce0
4a7
5ec
e4a
4ea
df7
2fb
a09


my repo
ce0
4a7
5ec
e4a
master
public/master

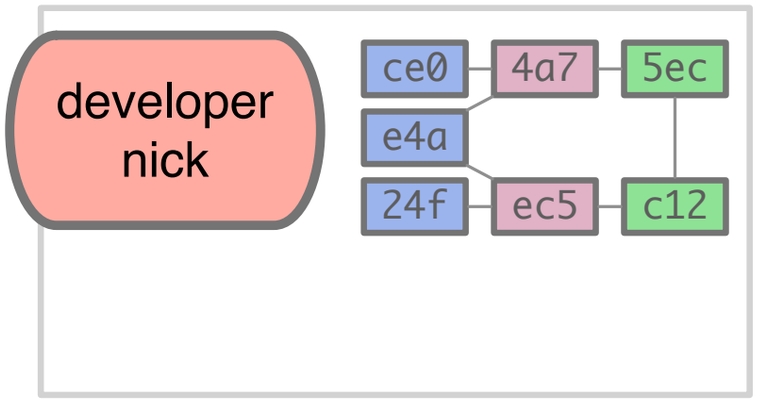


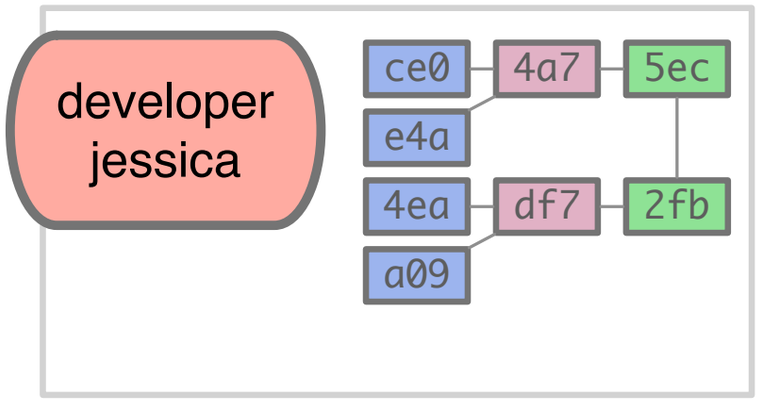


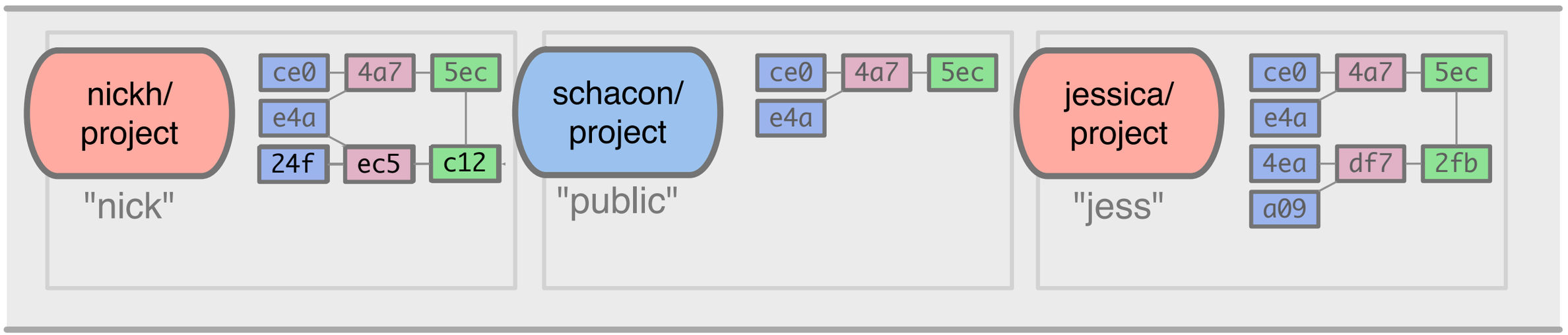


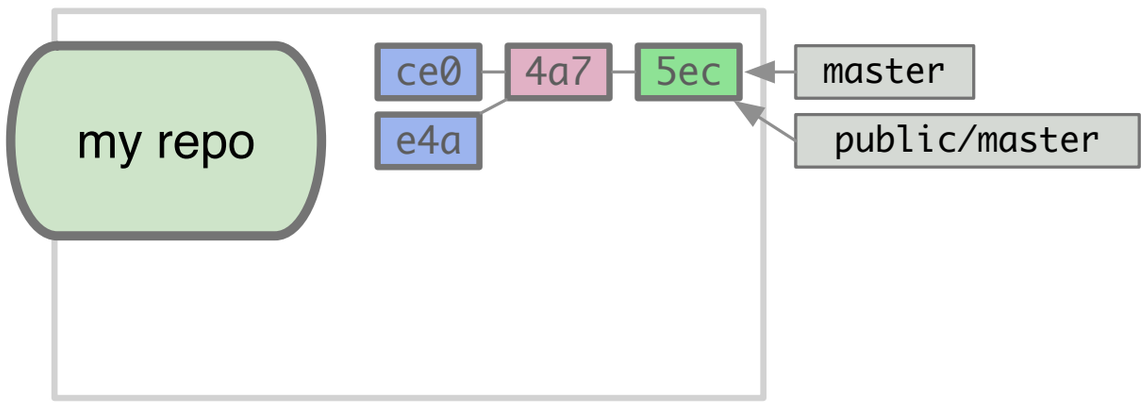


git remote add nick git://github.com/nickh/project.git

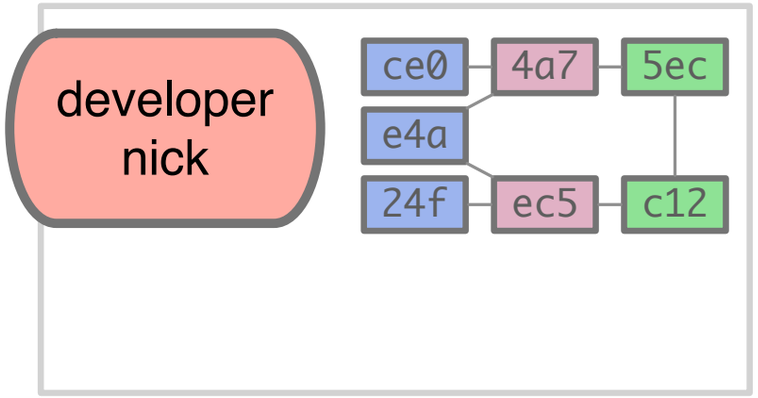


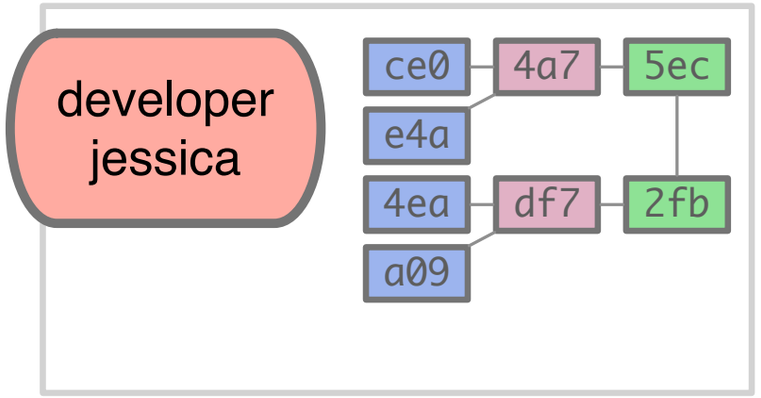


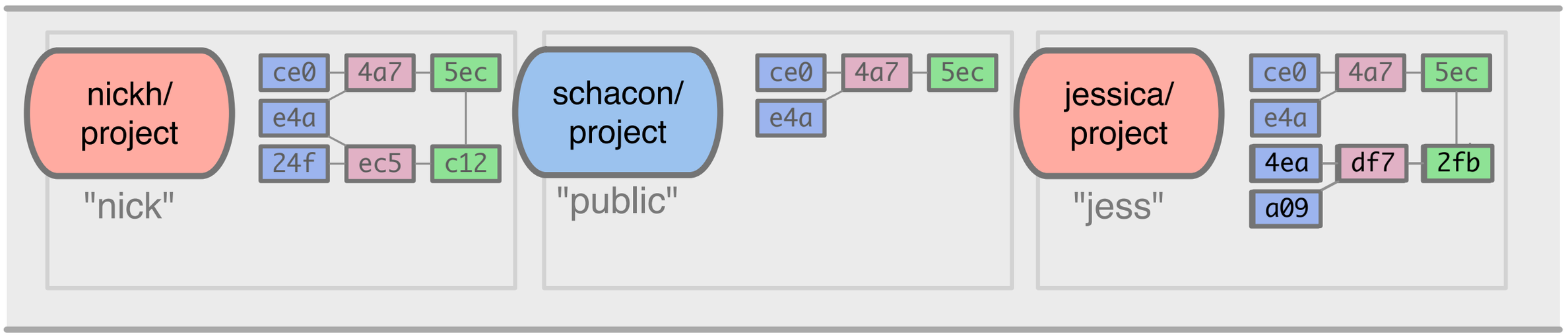


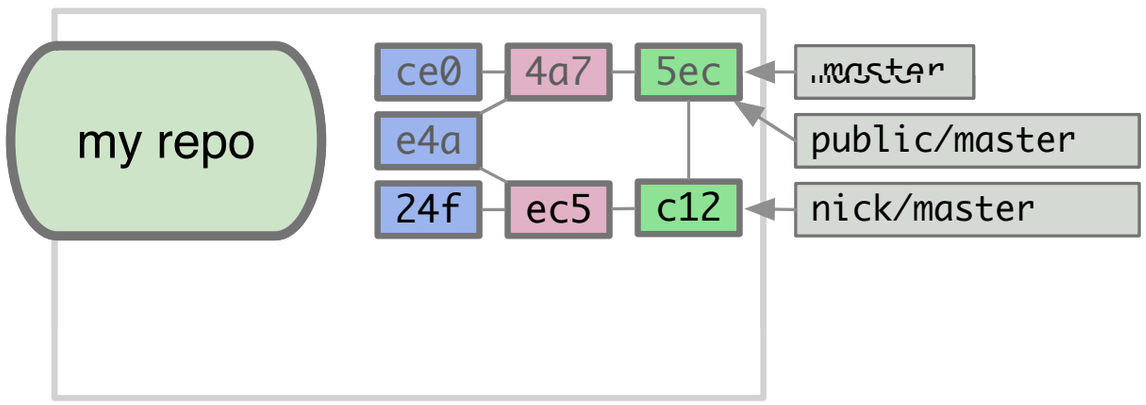


git fetch nick

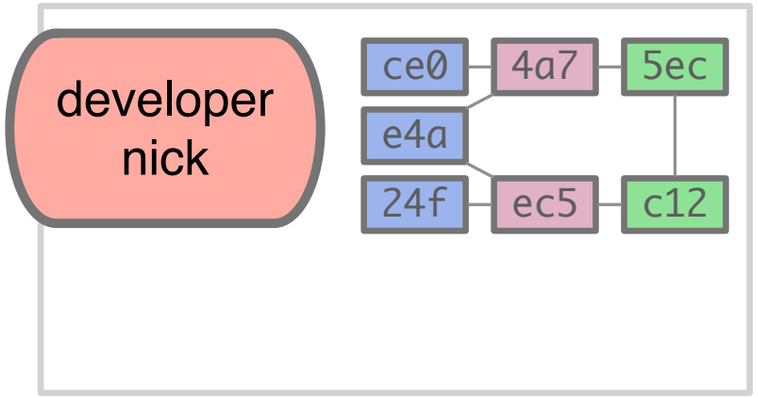


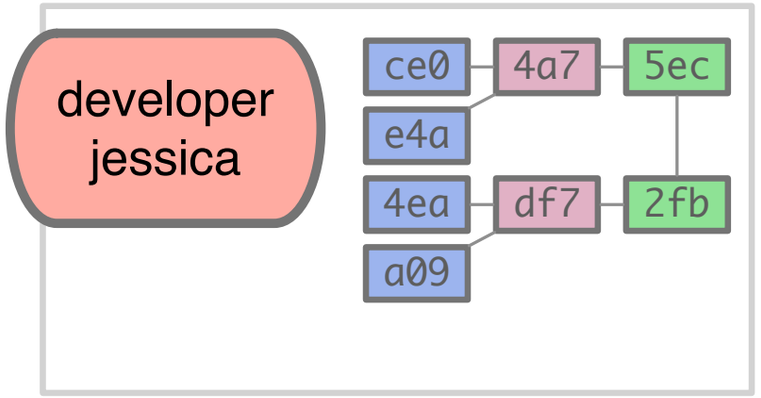


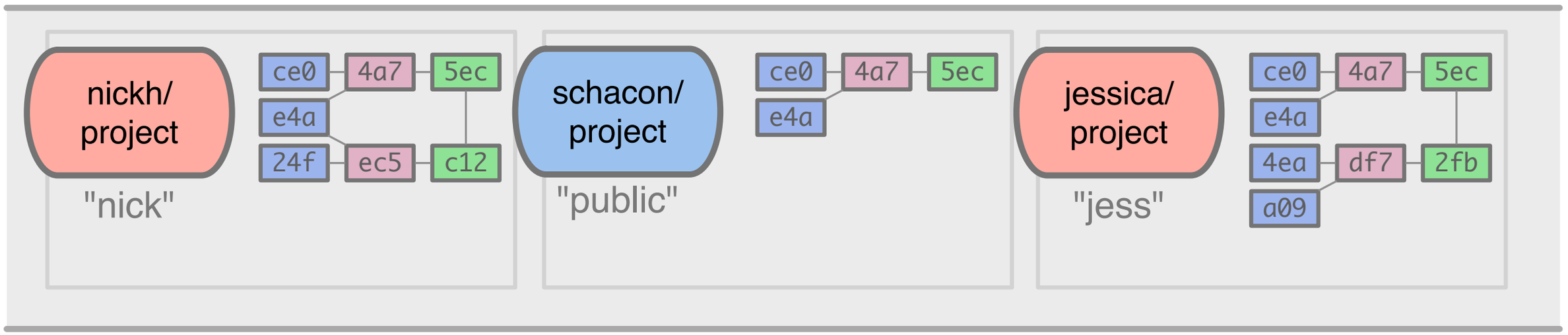


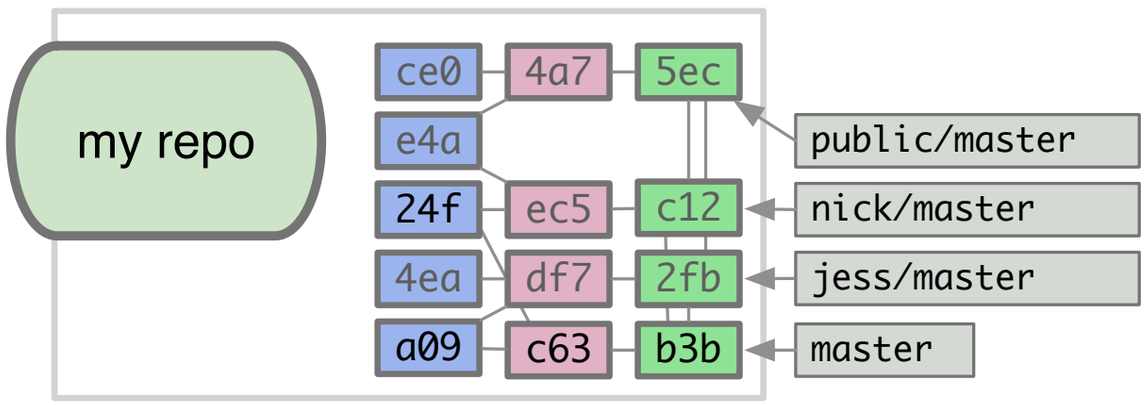


git fetch jess

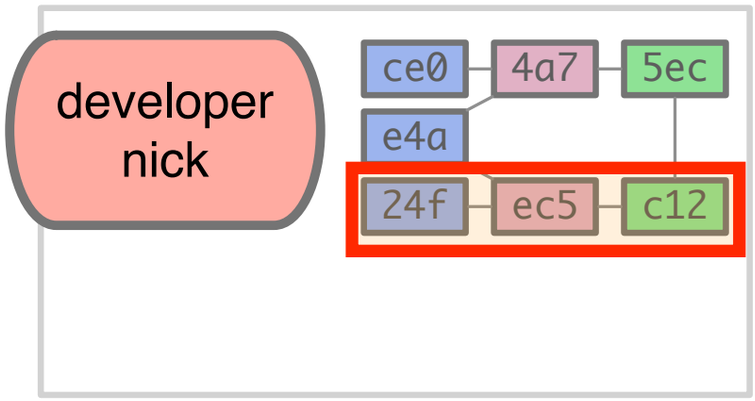


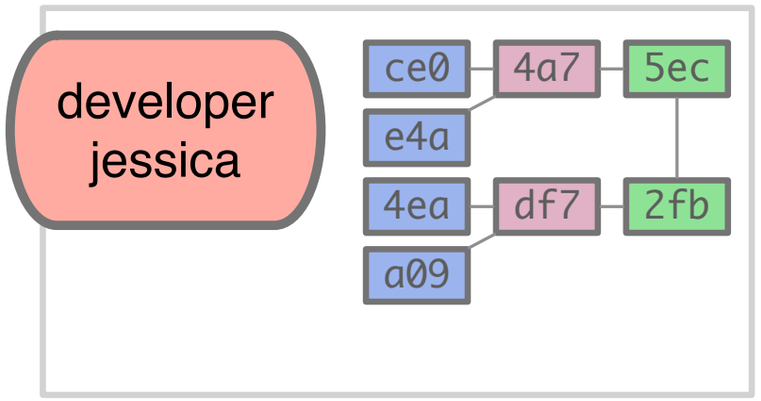


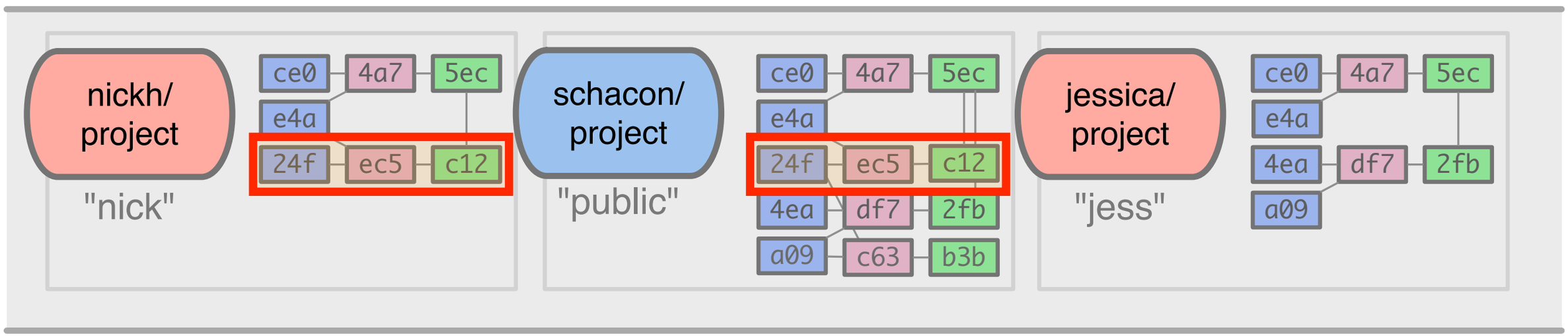


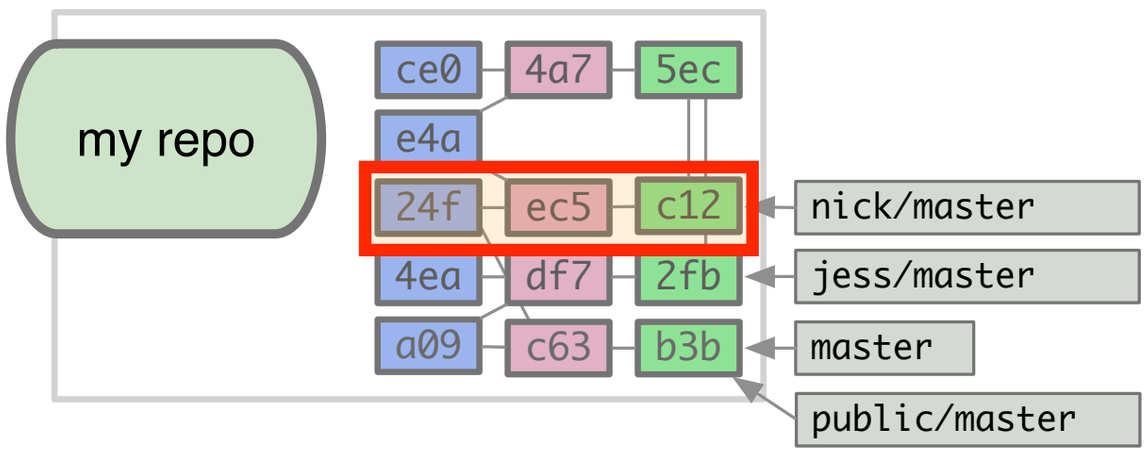


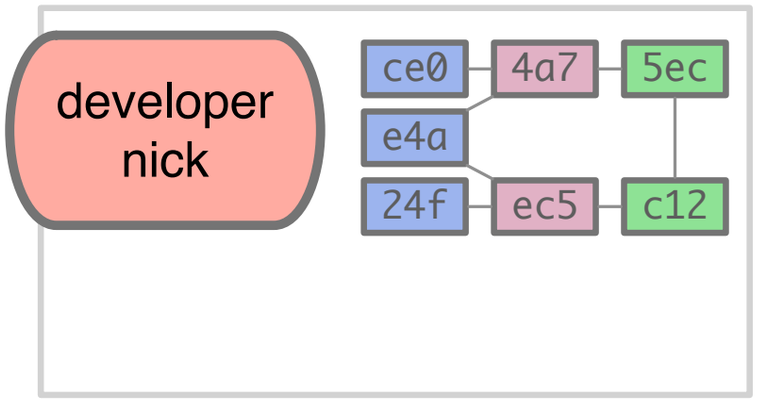
git merge nick jess

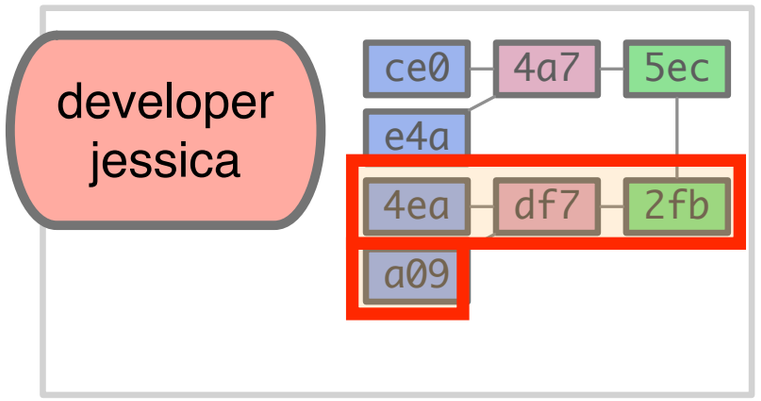

developer
nick
ce0
4a7
5ec
e4a
24f
ec5
c12

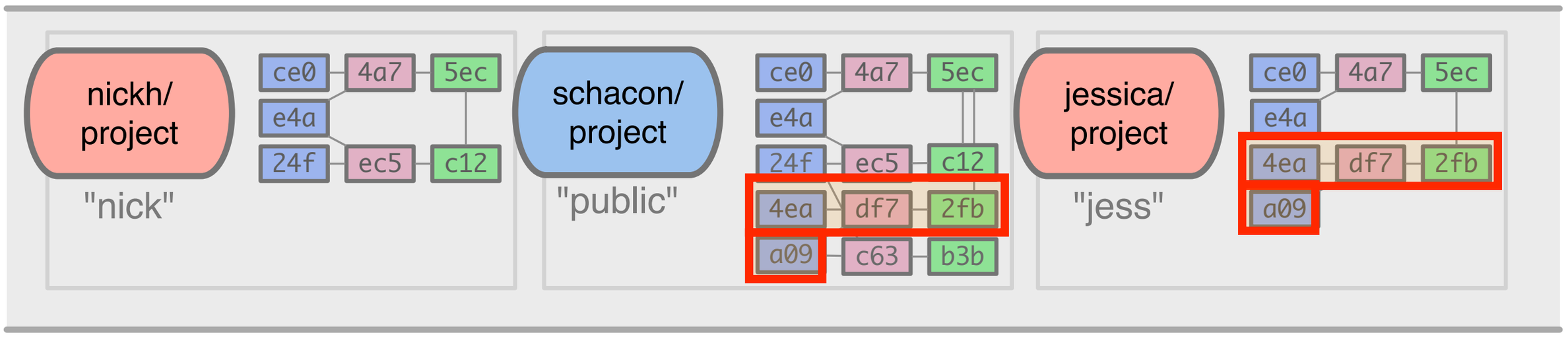

developer
jessica
ce0
4a7
5ec
e4a
4ea
df7
2fb
a09

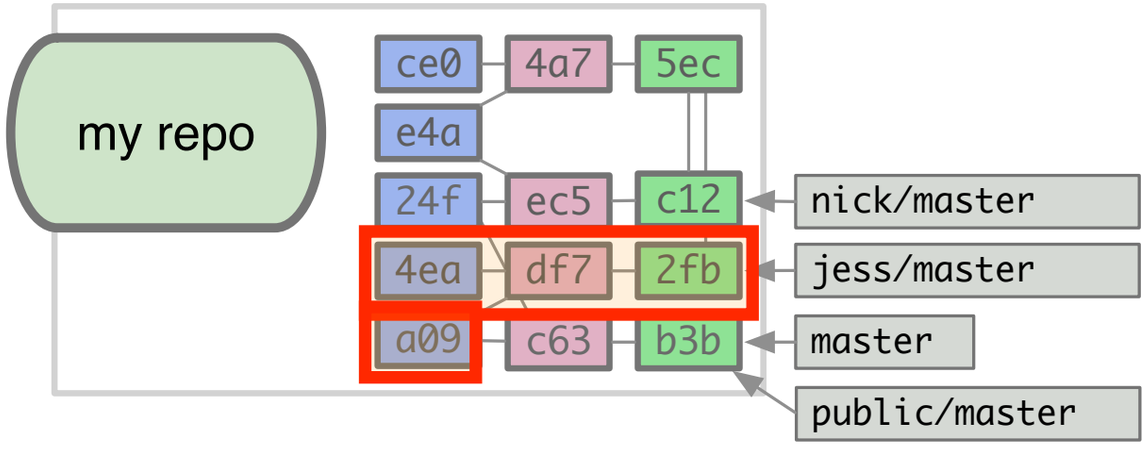

nickh/
project
"nick"
ce0
4a7
5ec
e4a
24f
ec5
c12
schacon/
project
"public"
ce0
4a7
5ec
e4a
24f
ec5
c12
4ea
df7
2fb
a09
c63
b3b
jessica/
project
"jess"
ce0
4a7
5ec
e4a
4ea
df7
2fb
a09

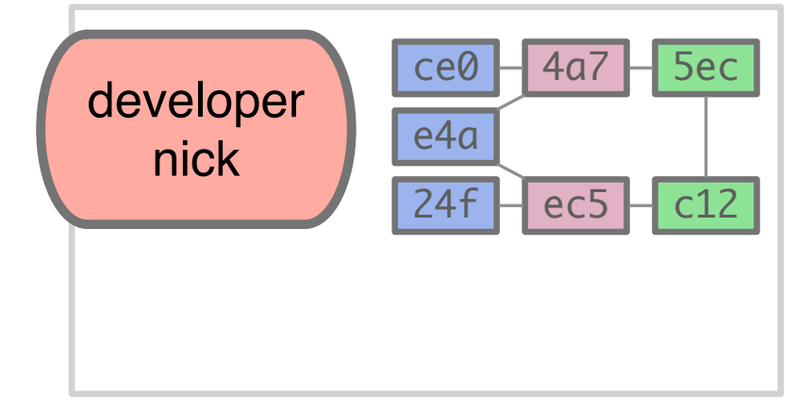

my repo
ce0
4a7
5ec
e4a
24f
ec5
c12
4ea
df7
2fb
a09
c63
b3b
nick/master
jess/master
master
public/master

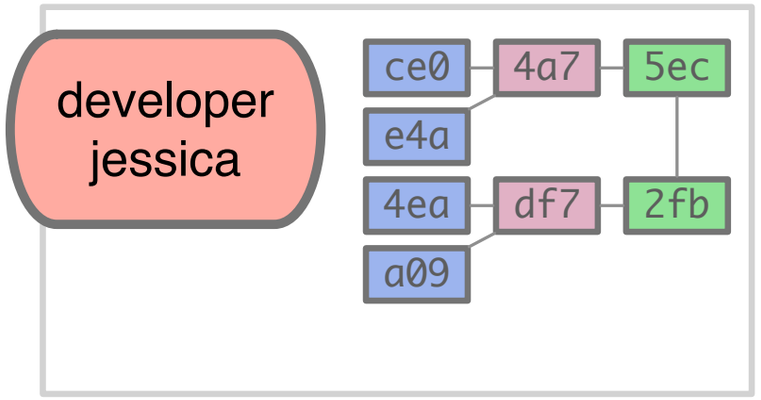

developer
nick
ce0
4a7
5ec
e4a
24f
ec5
c12

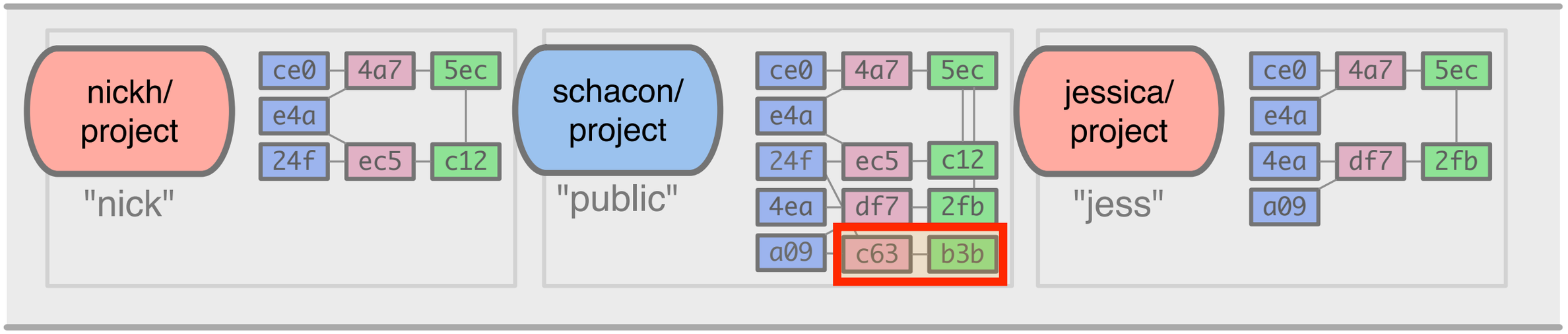

developer
jessica
ce0
4a7
5ec
e4a
4ea
df7
2fb
a09

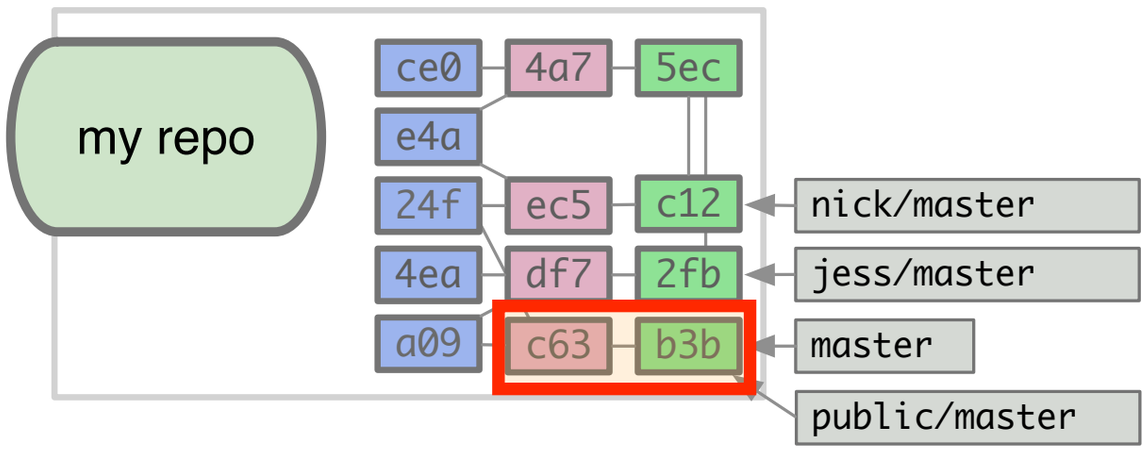

nickh/
project
"nick"
ce0
4a7
5ec
e4a
24f
ec5
c12
schacon/
project
"public"
ce0
4a7
5ec
e4a
24f
ec5
c12
4ea
df7
2fb
a09
c63
b3b
jessica/
project
"jess"
ce0
4a7
5ec
e4a
4ea
df7
2fb
a09


my repo
ce0
4a7
5ec
e4a
24f
ec5
c12
4ea
df7
2fb
a09
c63
b3b
nick/master
jess/master
master
public/master

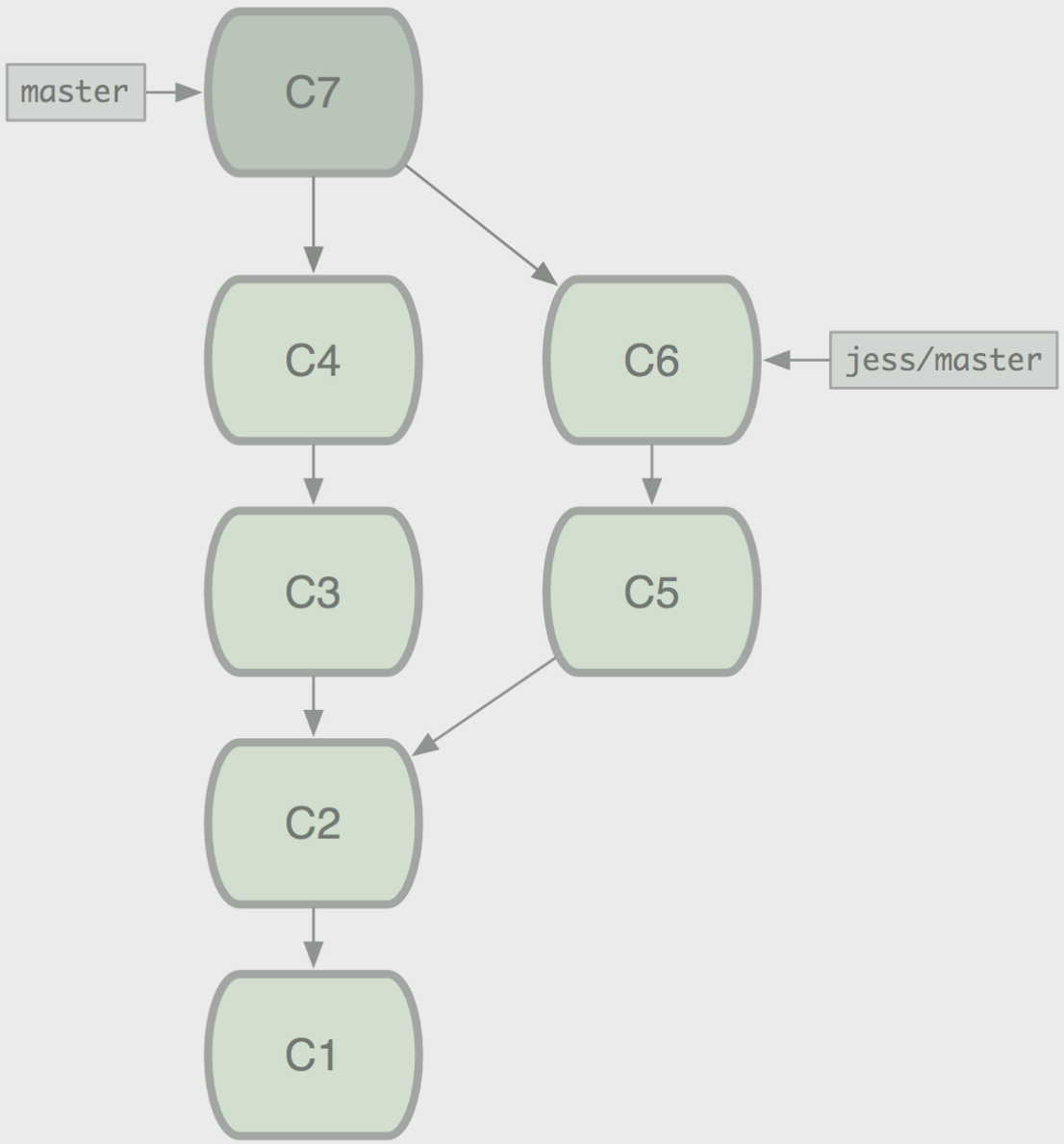

developer
nick
ce0
4a7
5ec
e4a
24f
ec5
c12


developer
jessica
ce0
4a7
5ec
e4a
4ea
df7
2fb
a09

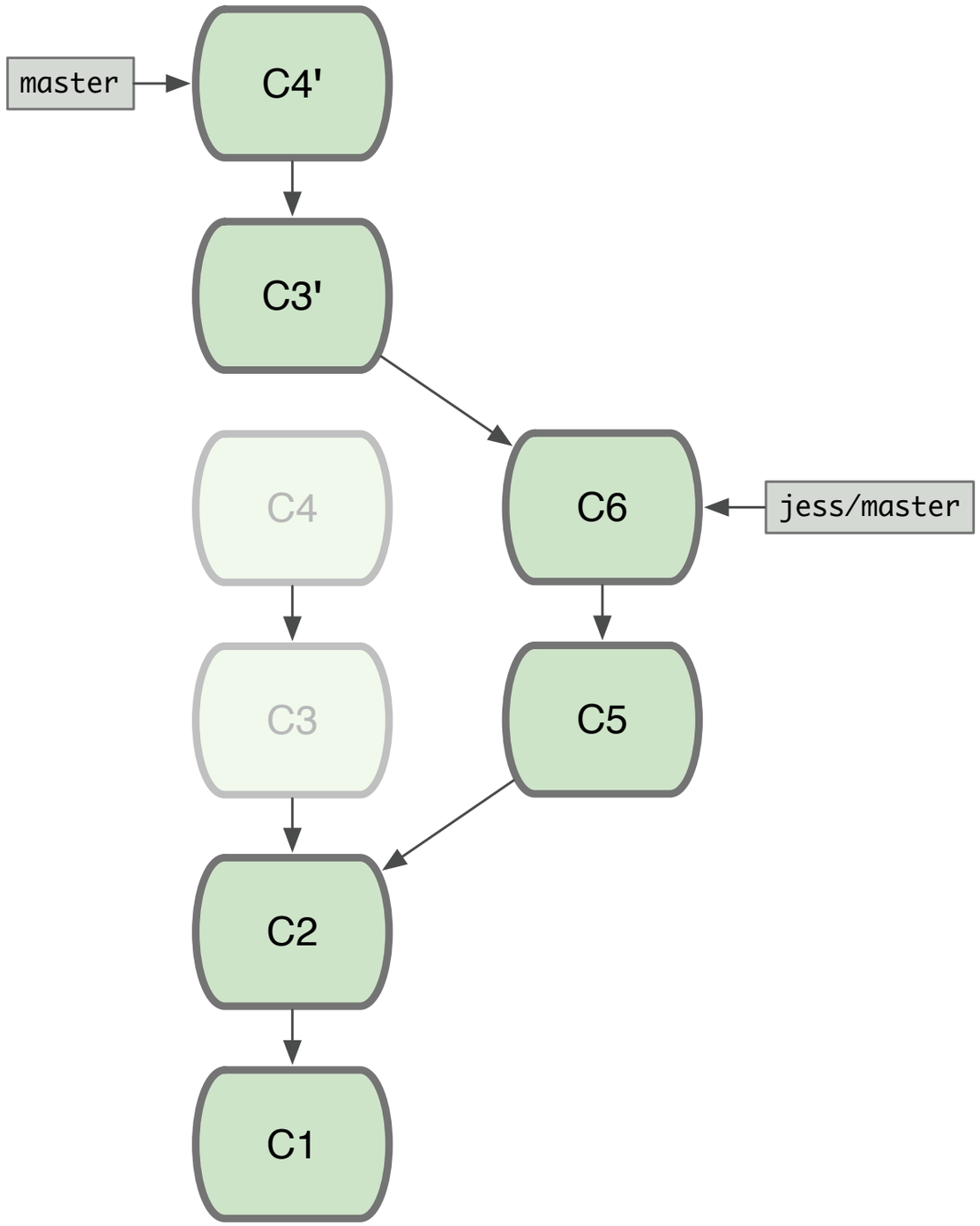

nickh/
project
"nick"
ce0
4a7
5ec
e4a
24f
ec5
c12
schacon/
project
"public"
ce0
4a7
5ec
e4a
24f
ec5
c12
4ea
df7
2fb
a09
c63
b3b
jessica/
project
"jess"
ce0
4a7
5ec
e4a
4ea
df7
2fb
a09


my repo
ce0
4a7
5ec
e4a
24f
ec5
c12
4ea
df7
2fb
a09
c63
b3b
nick/master
jess/master
master
public/master

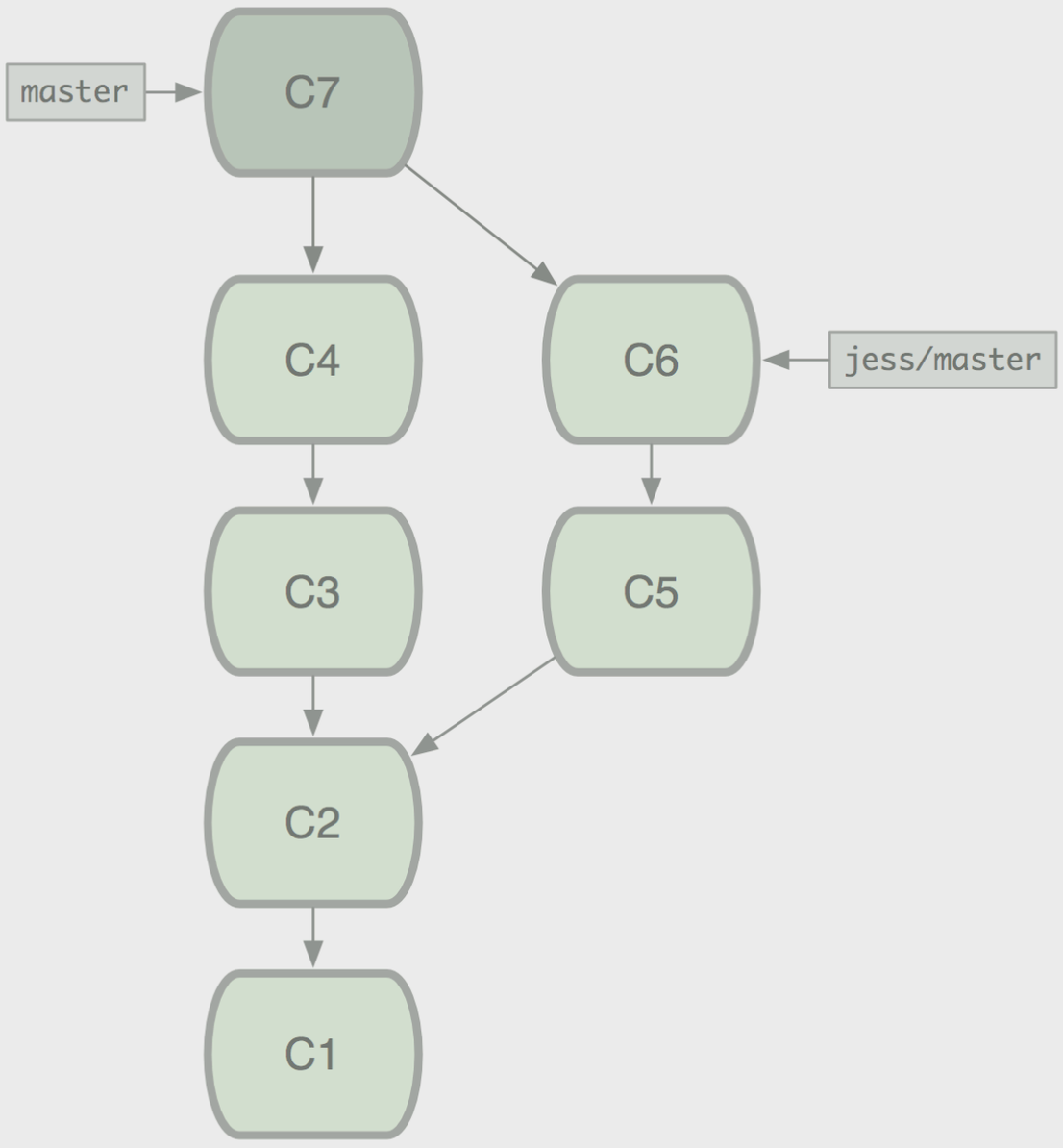
git merge



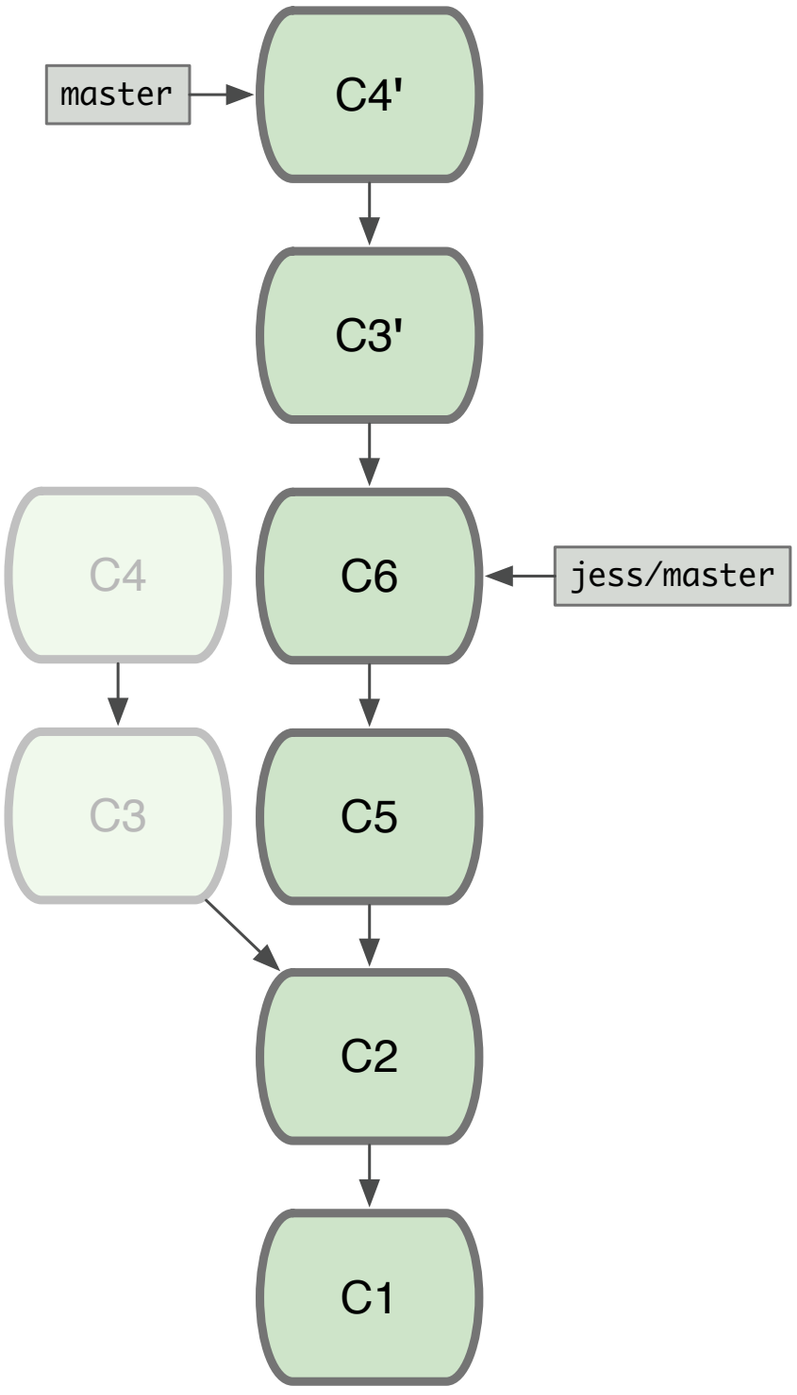
git rebase

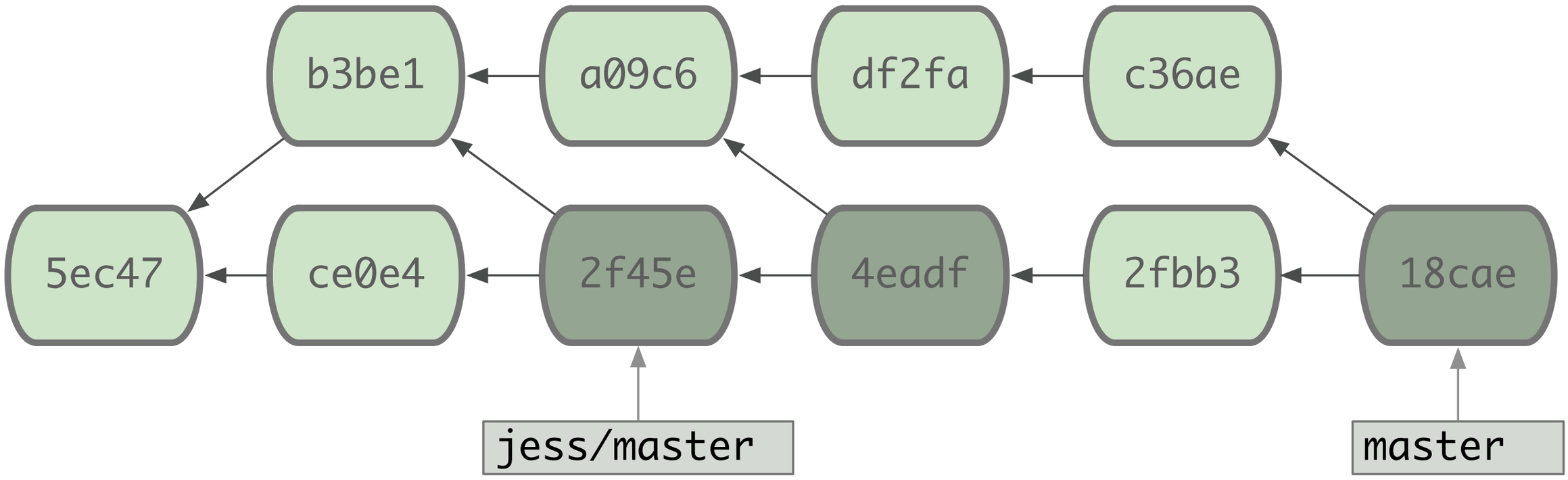


git merge

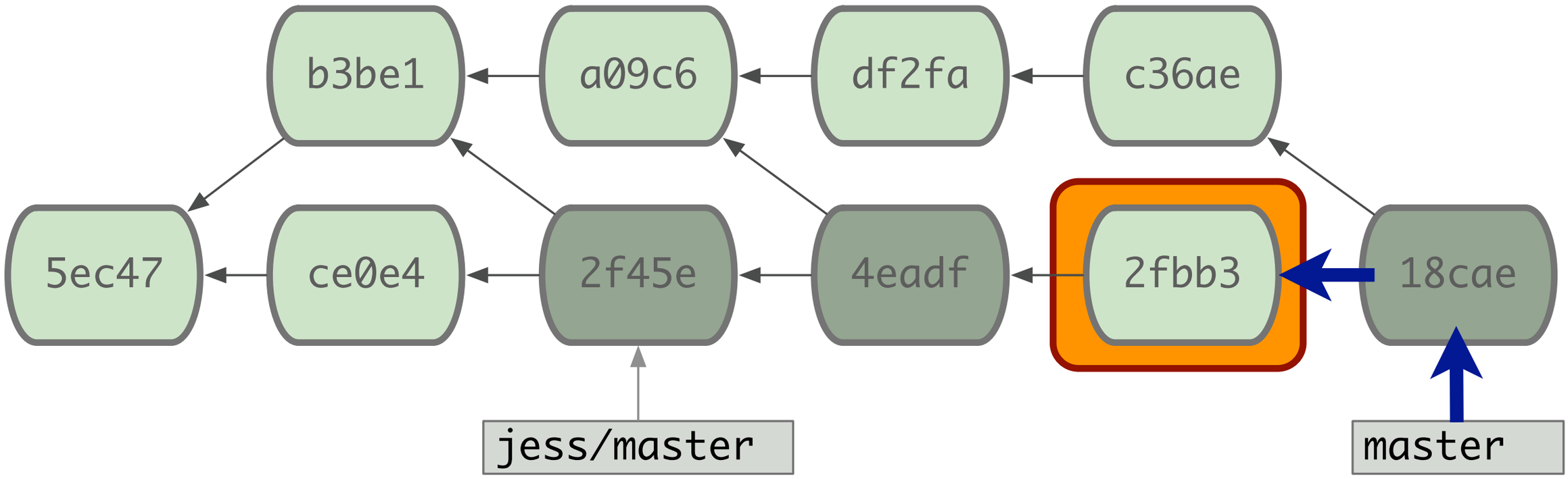


git rebase

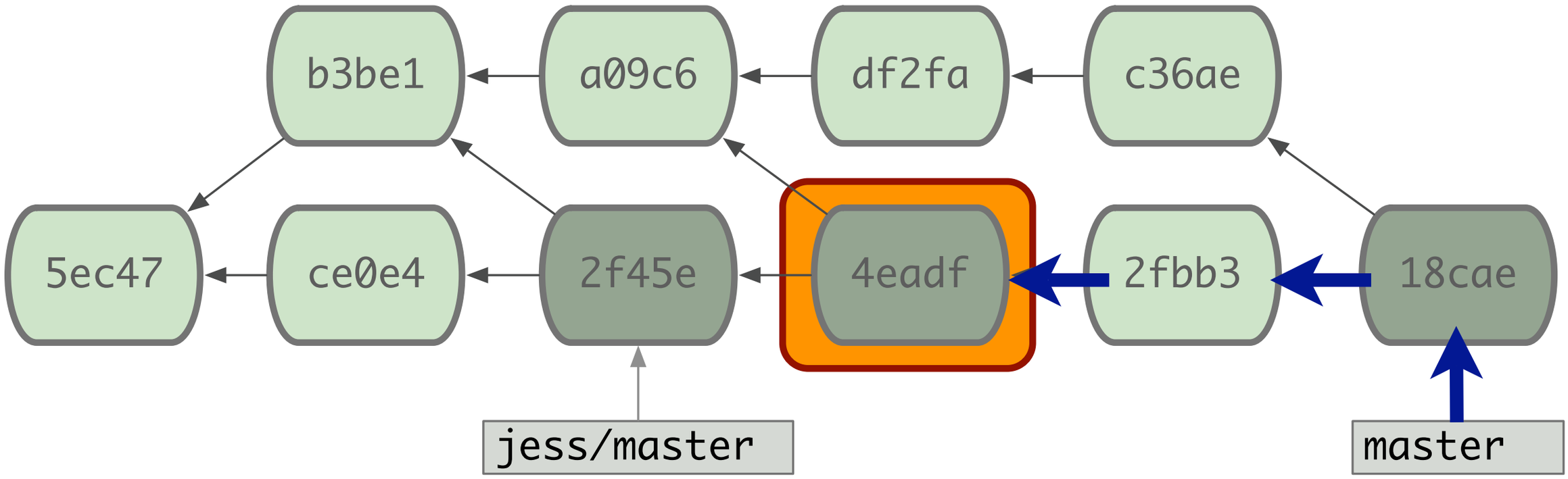



b3be1
a09c6
df2fa
c36ae
5ec47
ce0e4
2f45e
4eadf
2fbb3
18cae
jess/master
master

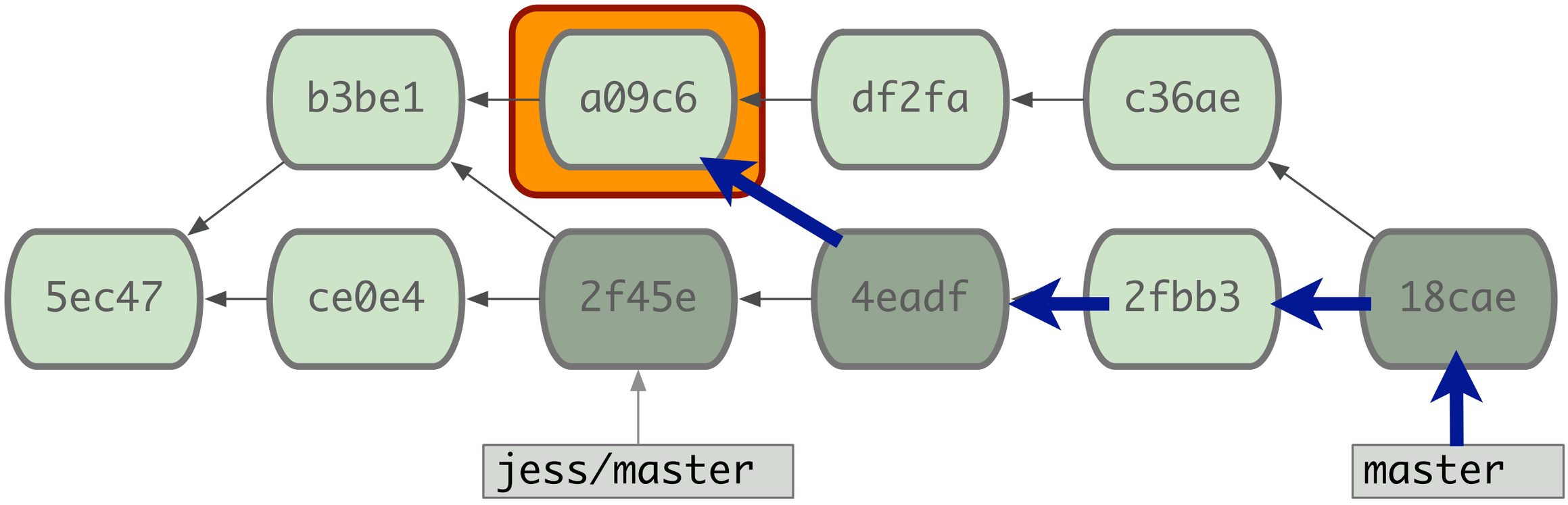
master^


b3be1
a09c6
df2fa
c36ae
5ec47
ce0e4
2f45e
4eadf
2fbb3
18cae
jess/master
master

master~2


b3be1
a09c6
df2fa
c36ae
5ec47
ce0e4
2f45e
4eadf
2fbb3
18cae
jess/master
master

master^^^2
master~2^2



setup

```
git config --global user.name
"Scott Chacon"
```

```
git config --global user.email
"schacon@gmail.com"
```

```
$>cat ~/.gitconfig
[user]
        name = Scott Chacon
        email = schacon@gmail.com
[color]
        status = auto
        branch = auto
        interactive = auto
        diff = auto
```

git clone

```
$>git clone git://github.com/schacon/ticgit.git
```

Git Protocols

ssh:// push

http[s]://

git://

file:///

~~rsync://~~

Git Protocols

pull

ssh://

http[s]://

git://

file:///

~~rsync://~~

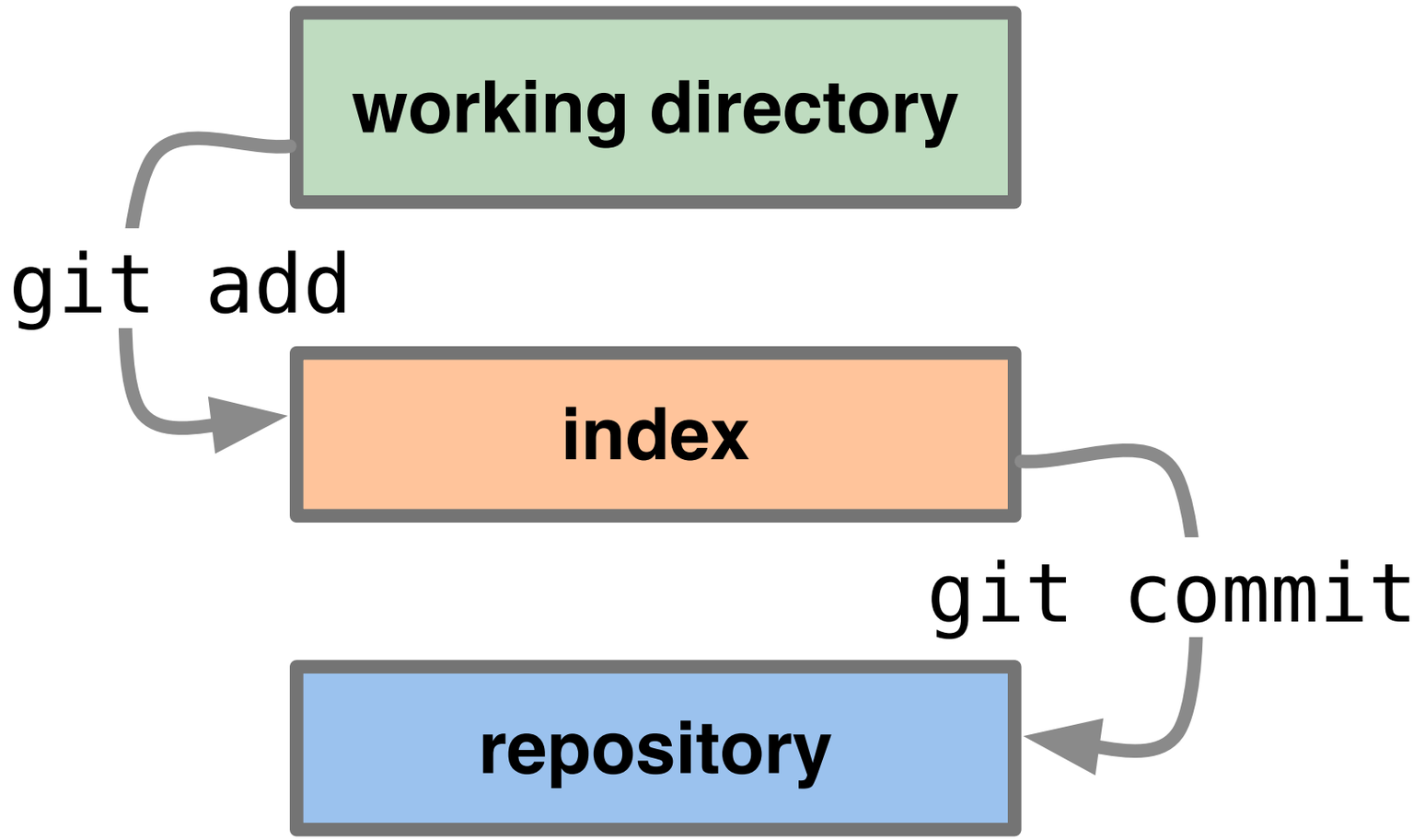
git clone

```
$>git clone git://github.com/schacon/ticgit.git
Initialized empty Git repository in /Users/schacon/examples
/ticgit/.git/
remote: Generating pack...
remote: Done counting 483 objects.
remote: Deltifying 483 objects...
remote:  100% (483/483) done
remote: Total 483 (delta 198), reused 483 (delta 198)
Receiving objects: 100% (483/483), 60.15 KiB, done.
Resolving deltas: 100% (198/198), done.
$>_
```

```
$>git init
Initialized empty Git repository in .git/
$>git add .
$>git commit -m 'my first commit'
Created initial commit bfe09f9: my first commit
 3 files changed, 3 insertions(+), 0 deletions(-)
 create mode 100644 README
 create mode 100644 Rakefile
 create mode 100644 lib/mylib.rb
$>_
```

```
$>cat .gitignore
tmp/**/*
log/*
config/database.yml
config/environments/production.rb
.DS_Store
```

```
$>git status
# On branch master
# Changes to be committed:
#   (use "git reset HEAD
#
#       modified:   READM
#
# Changed but not updated
#   (use "git add <file>.
#
#       modified:   TODO
#
```




```
$>git commit -m 'updated
Created commit c4f8fd6: u
 1 files changed, 1 inser
$>git status
# On branch master
# Changed but not updated:
#   (use "git add <file>..." to update what will be committed)
#
#       modified:   TODO
#
no changes added to commit (use "git add" and/or "git commit -a")
$>_
```

```
$>git log
commit 310154e3c7db47d8bac935c2c43aee6afac11aae
Author: Scott Chacon <schacon@gmail.com>
Date:   Sun Apr 13 10:49:15 2008 -0700

    updated README formatting and added blame

commit f7f3f6dd8fd3fa40f052427c32785a0fa01aaa5f
Author: Magnus Chacon <mchacon@gmail.com>
Date:   Sun Apr 13 10:45:01 2008 -0700

    changed my name a bit

commit 710f0f8d2cdf5af87033b9ec08859a505f9a6af5
Author: Magnus Chacon <mchacon@gmail.com>
Date:   Sun Apr 13 10:34:16 2008 -0700

    added ls-files

commit c110d7ff8cfb86fd5cce9a8aee462678dbb4ef9b
Author: Scott Chacon <schacon@gmail.com>
Date:   Sun Apr 6 12:13:36 2008 -0700

    made the ls-tree function recursive and list trees
```

```
$>git log --pretty=oneline
3e625e2139a71bfbaa85b055c139bceafab7e2c0 updated rakefile
4fefbaafd51ff3e79b84e9ee1a0dac2848e21a98 commit the todo
c4f8fd68ae3897d1d10456ed2b93d87cdc53d6e8 updated README
a5f4a0daa9b13bb85283461ddaba6e589e34ccb5 added cat-file
310154e3c7db47d8bac935c2c43aee6afac11aae updated README formatting
f7f3f6dd8fd3fa40f052427c32785a0fa01aaa5f changed my name a bit
710f0f8d2cdf5af87033b9ec08859a505f9a6af5 added ls-files
c110d7ff8cfb86fd5cce9a8aee462678dbb4ef9b made the ls-tree function
ce9b0d5551762048735dd67917046b44176317e0 limiting log to 30
e22e01e39c39521edc8ccd3aa5df228209e7ad2d -added todo options
2c0d4d7e3d36a8cd5da1eab3e0c1e83faa3bfd9f added limit to log functio
47c668a9e8f93f65ff9019c587f3b2204bd6c98d fixed conflict
ca28a28e64192af4527ea63665dceaa9bdba6a4b new version 0.1.2
f30ba8480d9aba88580696651e80d9aabe8e15a2 new version 0.2.0
1edee6b1d61823a2de3b09c160d7080b8d1b3a40 added a new function
4b0780c5b7a4ae1217f4da742a172e7d787ca3eb rakefile and todo file add
0d57a4eb03d35e1e1041857d1cb01199eeb23cc4 Revert "rakefile and todo
f65a297eb0ace63e5b302e849064f37ded909c0b rakefile and todo file add
c035c611b559537d725d875cfa10c05c17d48e4f Merge branch 'newfunc'
d6fad7dd6f98481cab874b192aaa79d03490d4a0 added more description
1a8c32ea9725886aaadc35dbf4d68468d3af0abe added lstree function
40e5c10f67fa01c01da5687d08bd3bc21bb6a44b added email to readme
cf25cc3bfb0ece7dc3609b8dc0cc4a1e19ffbcd4 committing all changes
0c8a9ec46029a4e92a428cb98c9693f09f69a3ff changed the verison number
```

```
$>git log --pretty=format:"%h %an %ar - %s"
3e625e2 Magnus Chacon 47 minutes ago - updated rakefile
4fefbaa Magnus Chacon 56 minutes ago - commit the todo
c4f8fd6 Magnus Chacon 60 minutes ago - updated README
a5f4a0d Magnus Chacon 2 weeks ago - added cat-file
310154e Scott Chacon 2 weeks ago - updated README formatting and a
f7f3f6d Magnus Chacon 2 weeks ago - changed my name a bit
710f0f8 Magnus Chacon 2 weeks ago - added ls-files
c110d7f Scott Chacon 3 weeks ago - made the ls-tree function recur
ce9b0d5 Scott Chacon 3 weeks ago - limiting log to 30
e22e01e Scott Chacon 3 weeks ago - -added todo options
2c0d4d7 Scott Chacon 3 weeks ago - added limit to log function
47c668a Scott Chacon 3 weeks ago - fixed conflict
ca28a28 Scott Chacon 3 weeks ago - new version 0.1.2
f30ba84 Scott Chacon 3 weeks ago - new version 0.2.0
1edee6b Scott Chacon 3 weeks ago - added a new function
4b0780c Scott Chacon 3 weeks ago - rakefile and todo file added
0d57a4e Scott Chacon 3 weeks ago - Revert "rakefile and todo file
f65a297 Scott Chacon 3 weeks ago - rakefile and todo file added
c035c61 Scott Chacon 3 weeks ago - Merge branch 'newfunc'
d6fad7d Scott Chacon 3 weeks ago - added more description
1a8c32e Scott Chacon 3 weeks ago - added lstree function
40e5c10 Scott Chacon 3 weeks ago - added email to readme
cf25cc3 Scott Chacon 6 weeks ago - committing all changes
0c8a9ec Scott Chacon 6 weeks ago - changed the verison number
0576fac Scott Chacon 6 weeks ago - my second commit, which is bett
```

git diff (treeish1) (treeish2)

```
diff --git a/.gitignore b/.gitignore
index 224f138..7e56ec7 100644
--- a/.gitignore
+++ b/.gitignore
@@ -1,4 +1,3 @@
 bin/*.log
 pkg
 coverage
-ticgit*gem
diff --git a/Rakefile b/Rakefile
index b30e89c..bd7b91a 100644
--- a/Rakefile
+++ b/Rakefile
@@ -31,7 +31,6 @@ desc "Clean out the coverage and pkg directories"
 task :clean do
   rm_rf 'coverage'
   rm_rf 'pkg'
-  rm Dir.glob('ticgit*gem')
 end

 task :default => "pkg/#{spec.name}-#{spec.version}.gem" do
diff --git a/ticgit.gemspec b/ticgit.gemspec
index c0e96b7..775eaf1 100644
--- a/ticgit.gemspec
+++ b/ticgit.gemspec
@@ -6,7 +6,7 @@ Gem::Specification.new do |s|
     s.author    =   "Scott Chacon"
     s.email     =   "schacon@gmail.com"
     s.summary   =   "A distributed ticketing system for Git projects."
-    s.files     =   ["lib/ticgit/base.rb", "lib/ticgit/cli.rb", "lib/ticgit/comment.rb"
+    s.files     =   ["lib/ticgit", "lib/ticgit/base.rb", "lib/ticgit/cli.rb", "lib/ticg

     s.bindir = 'bin'
     s.executables << "ti"
```

branching workflow

normal development

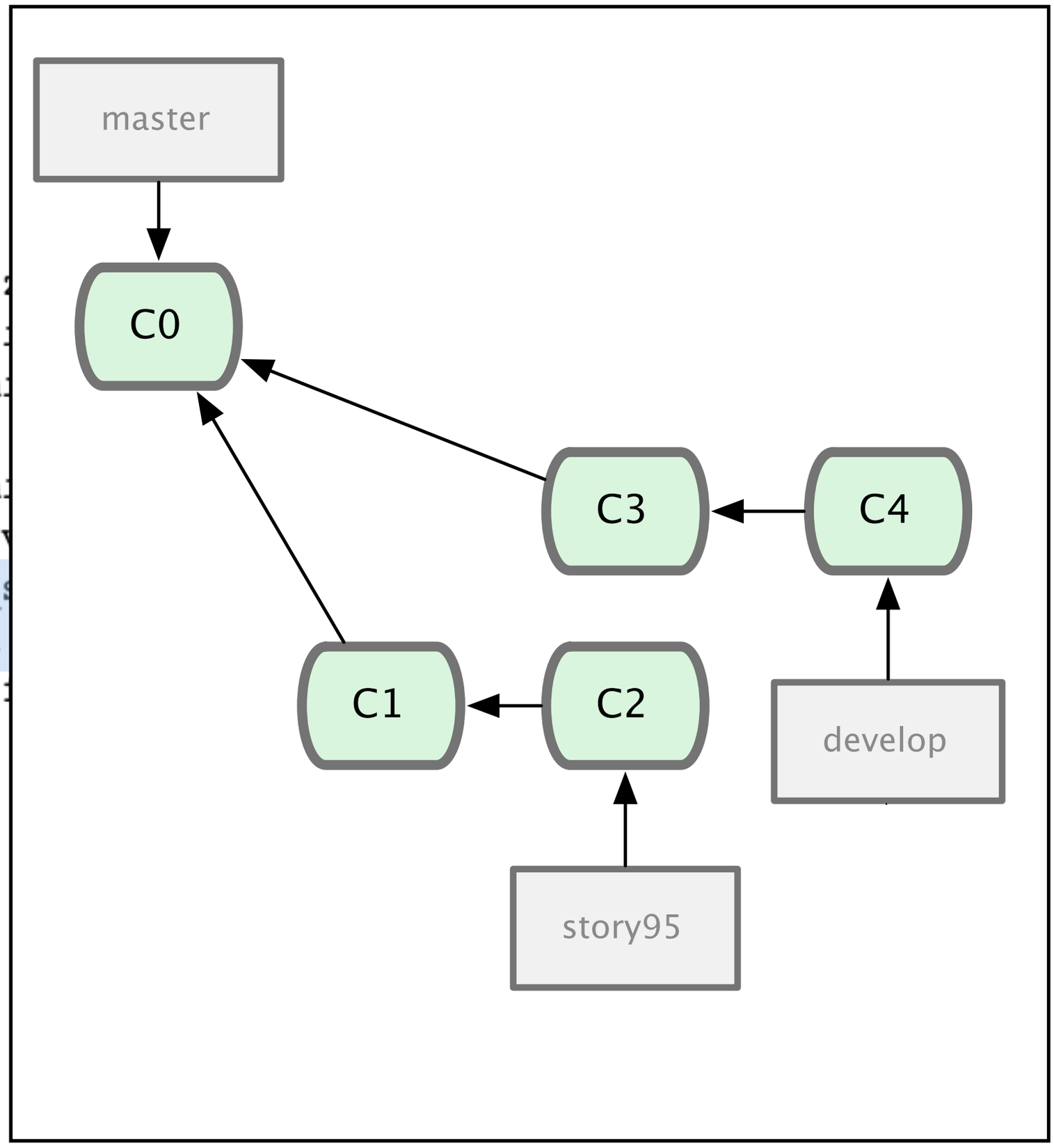
want to try an idea

have to do a hotfix

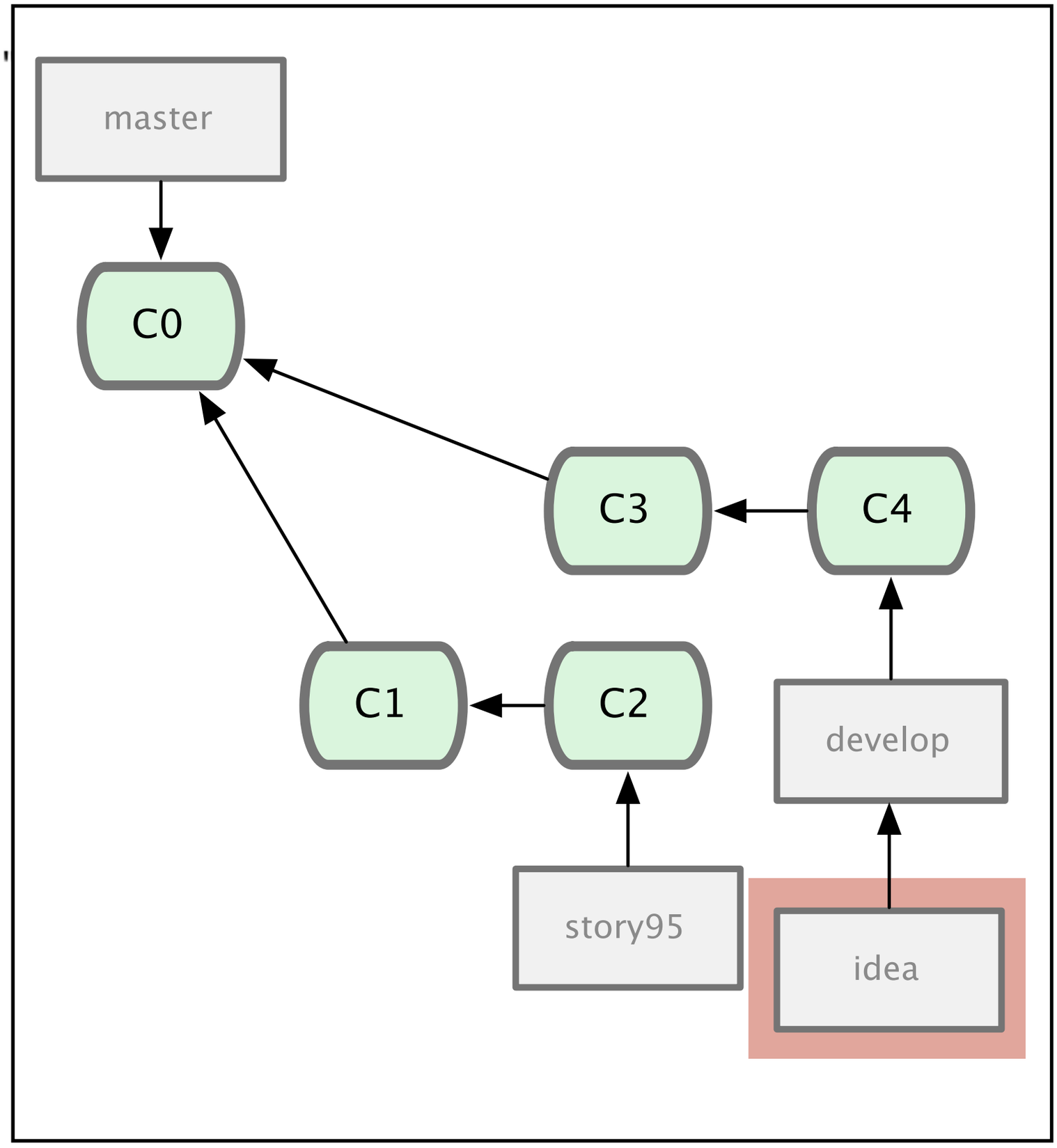
merge idea into development

push to production

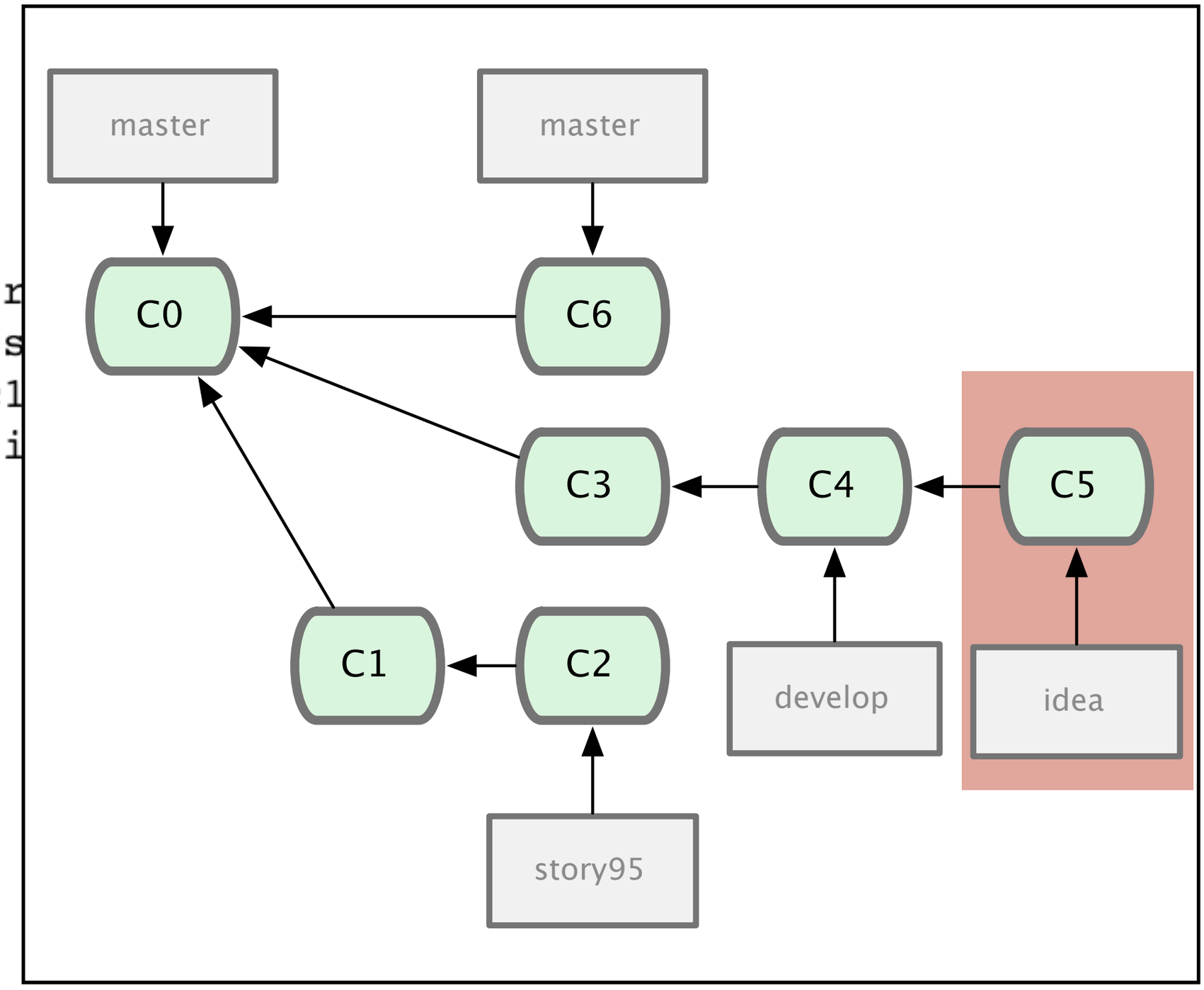
```
$>git branch
* develop
  master
  story95
$>git show-branch
* [develop] new file_si
 ! [master] added cat-fi
  ! [story95] show-commi
---
  + [story95] show-commi
  + [story95^] added rev
*    [develop] new file_s
*    [develop^] modified
*++ [master] added cat-f
$>_
```



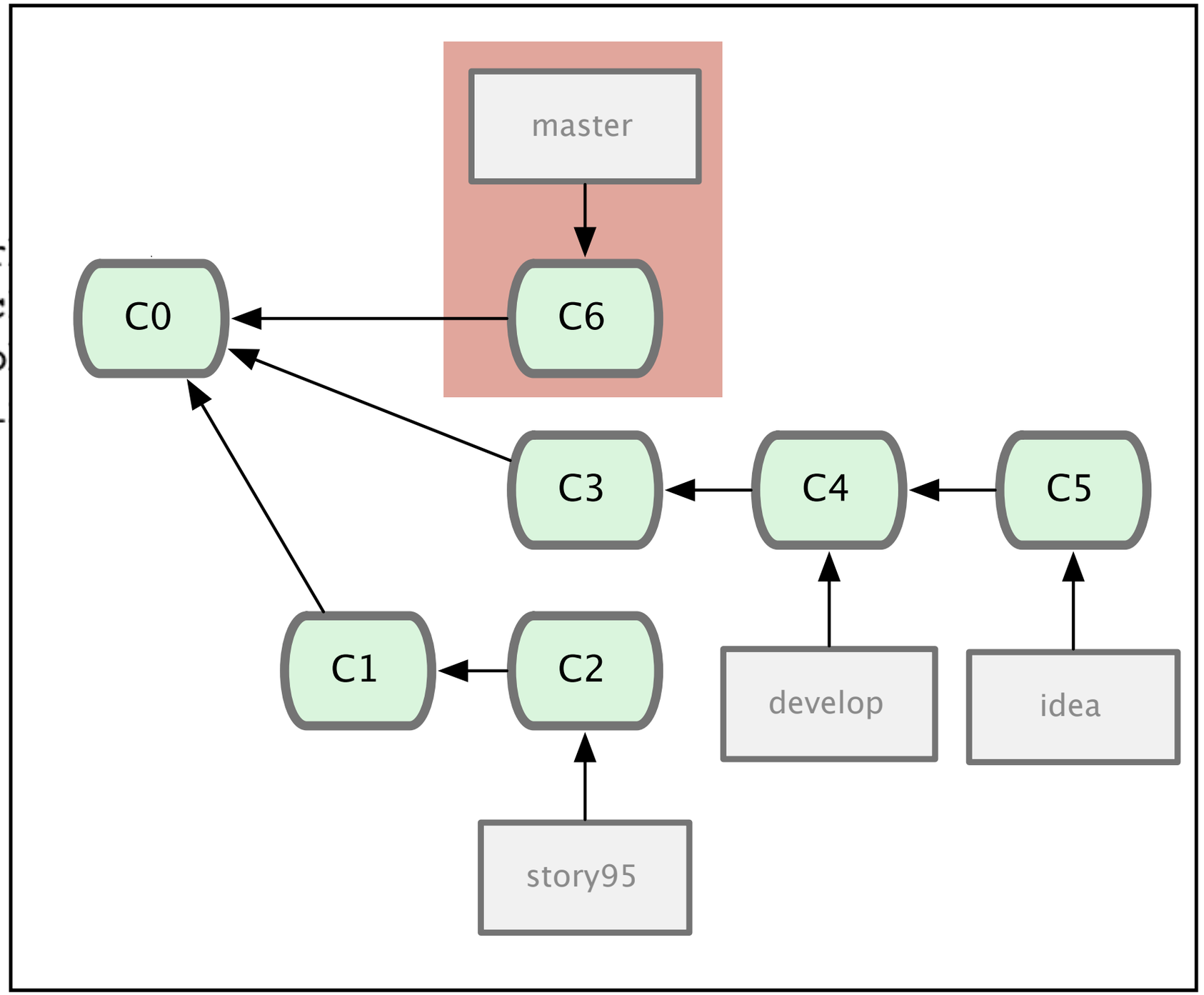
```
$>git checkout -b idea
Switched to a new branch '
$>git branch
  develop
* idea
  master
  story95
$>
```



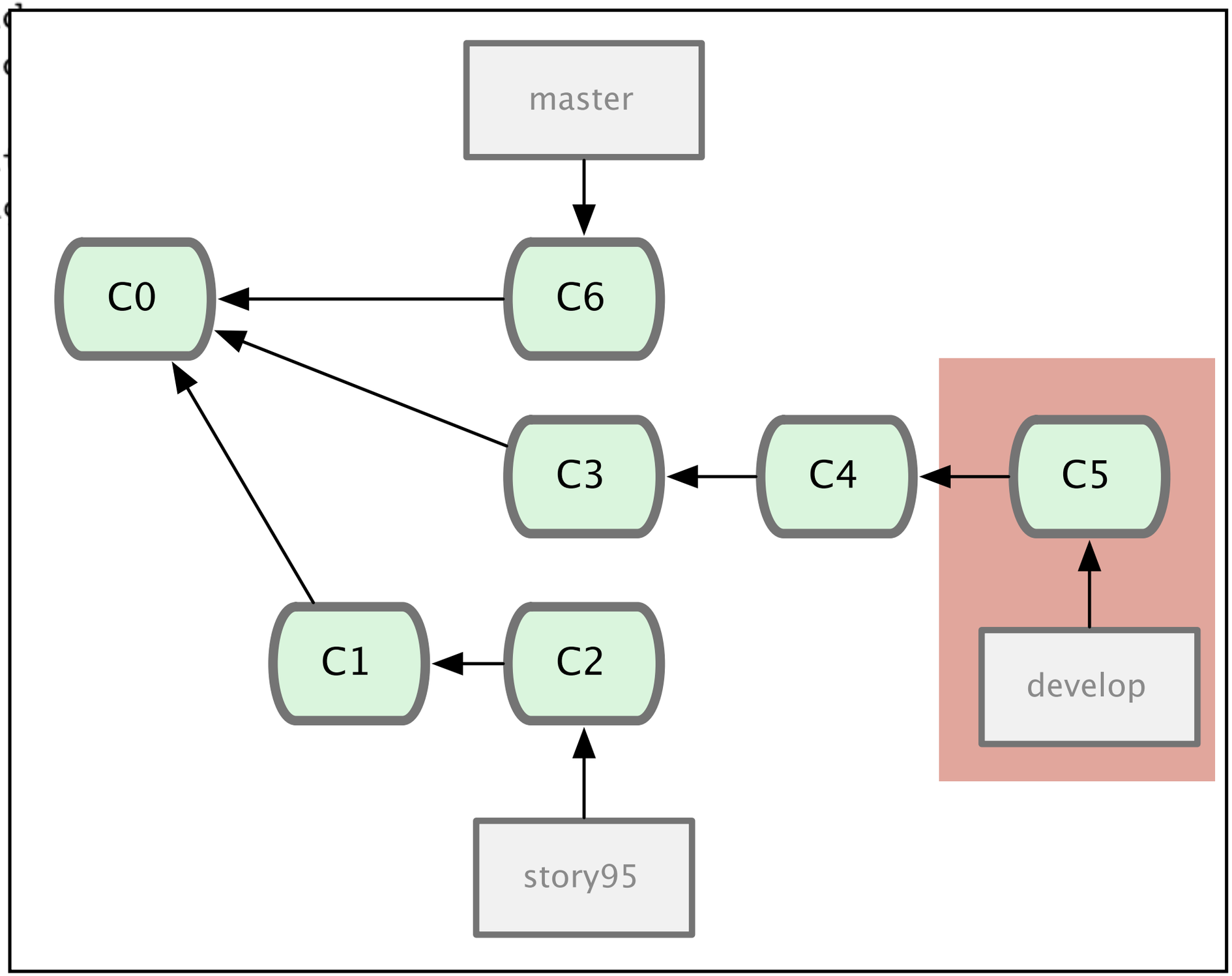
```
$>git checkout -b idea
Switched to a new branch "idea"
$>git branch
  develop
* idea
  master
  story95
$>vim lib/simplegit.r
$>git commit -a -m 's
Created commit 8268c1
 1 files changed, 4 i
```



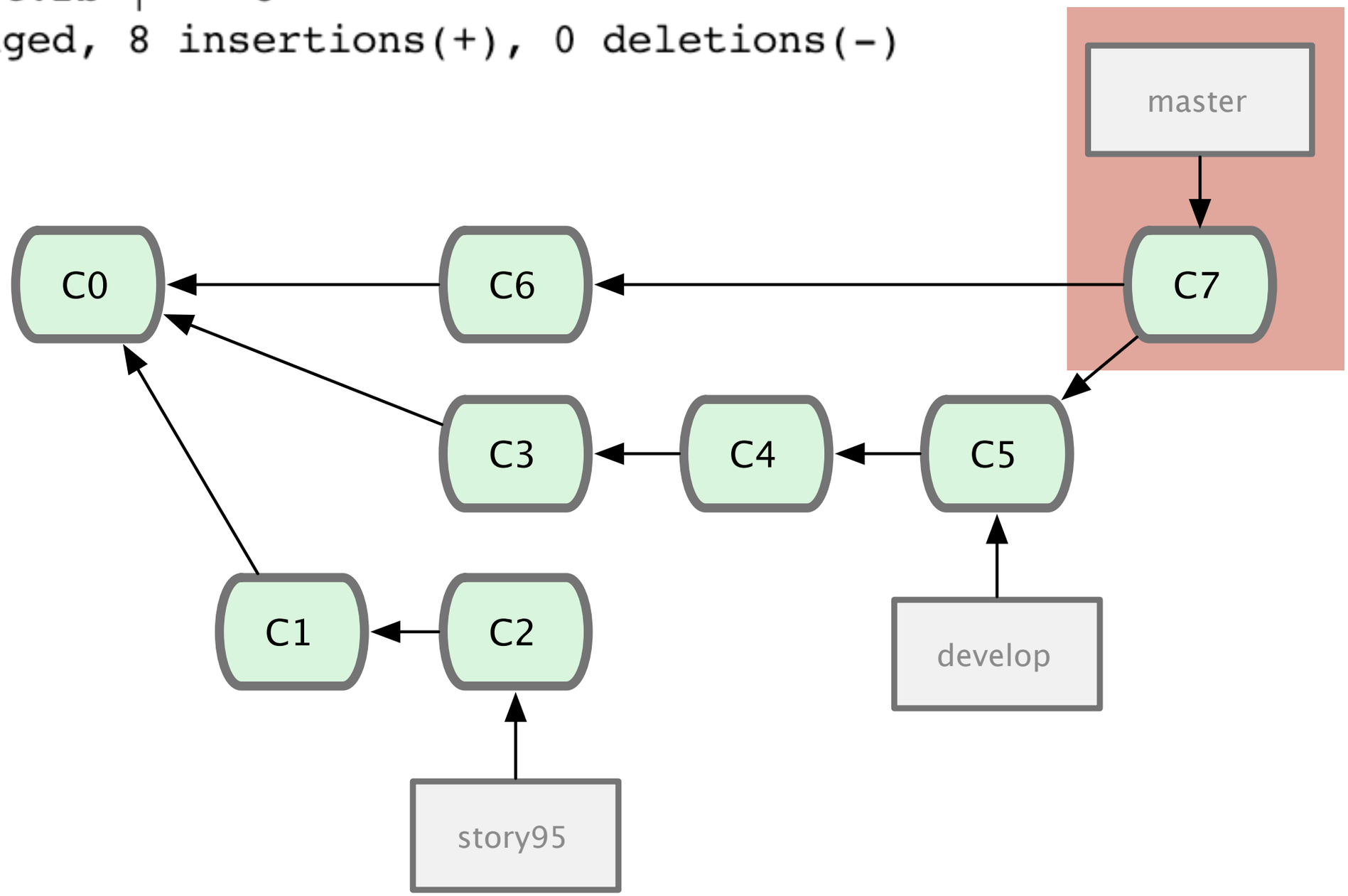
```
$>git checkout master
Switched to branch "master"
$>git branch
  develop
  idea
* master
  story95
$>vim lib/simplegit.r
$>git commit -a -m 'c
Created commit 749bcb
 1 files changed, 1 i
$>git tag -a 'v1.3'
```



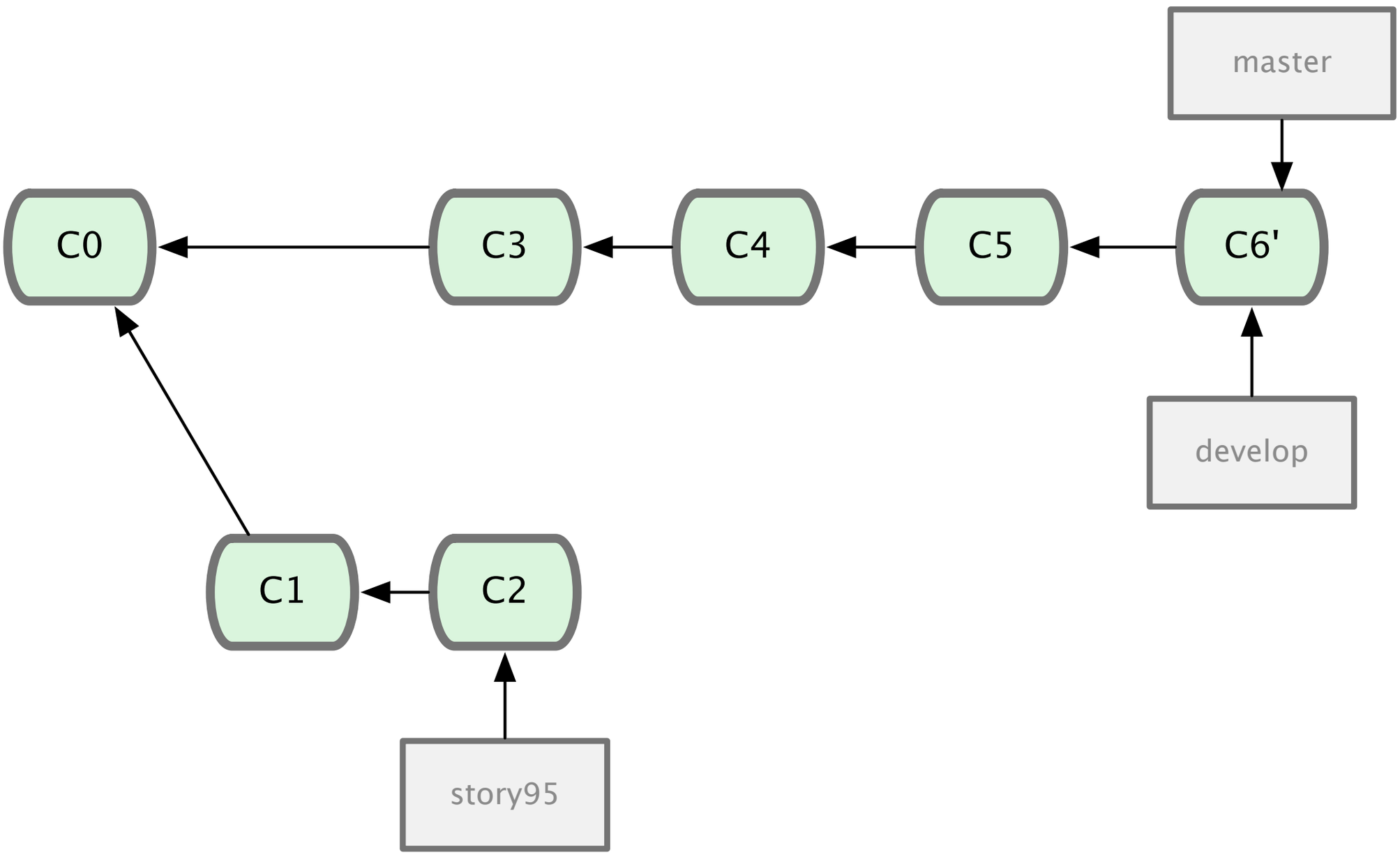
```
$>git checkout develop
Switched to branch "develop"
$>git merge id
Updating 8f35
Fast forward
 lib/simplegi
 1 files chan
```



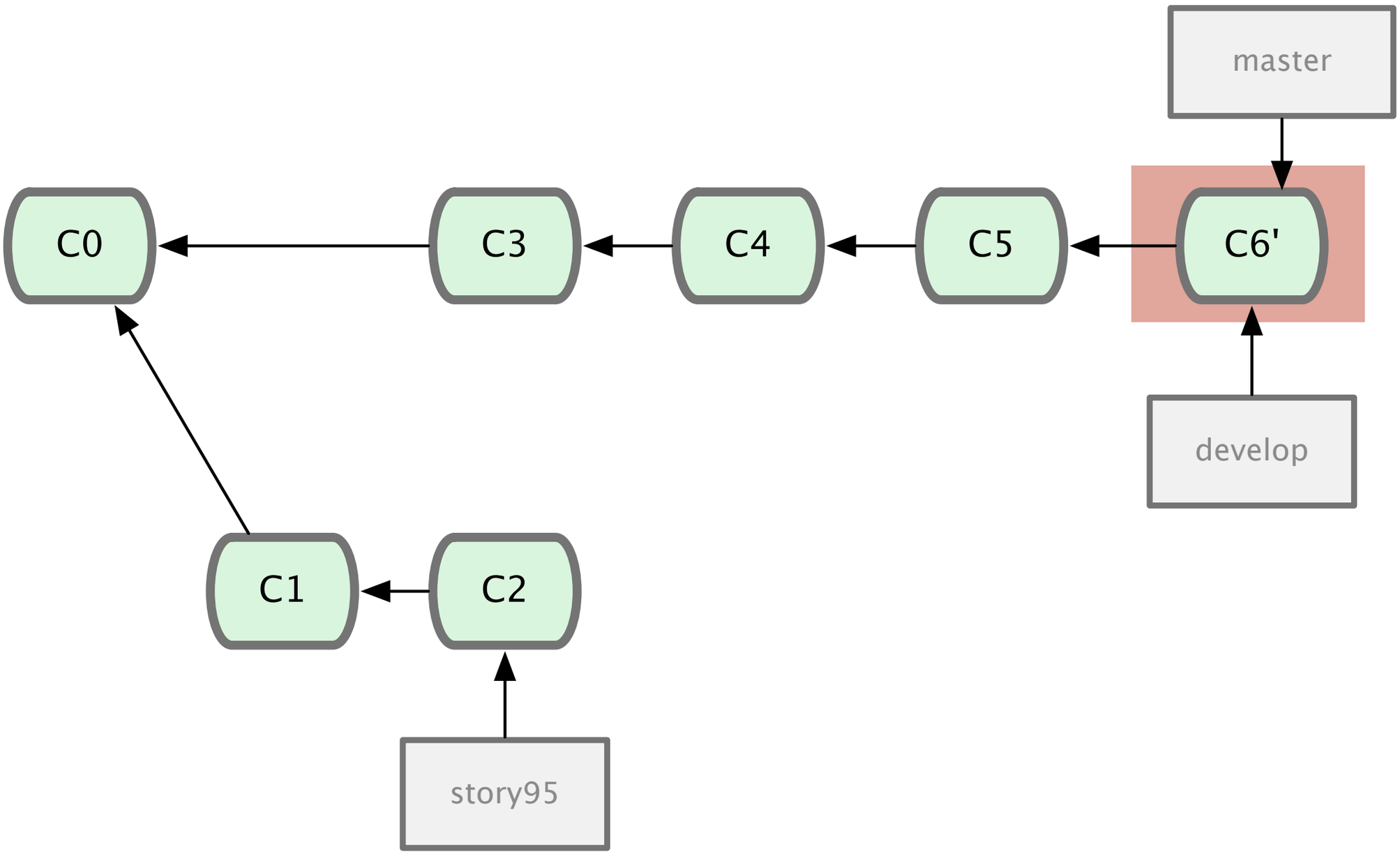
```
$>git checkout master
Switched to branch "master"
$>git merge develop
Auto-merged lib/simplegit.rb
Merge made by recursive.
 lib/simplegit.rb |     8 ++++++++
 1 files changed, 8 insertions(+), 0 deletions(-)
$>
```



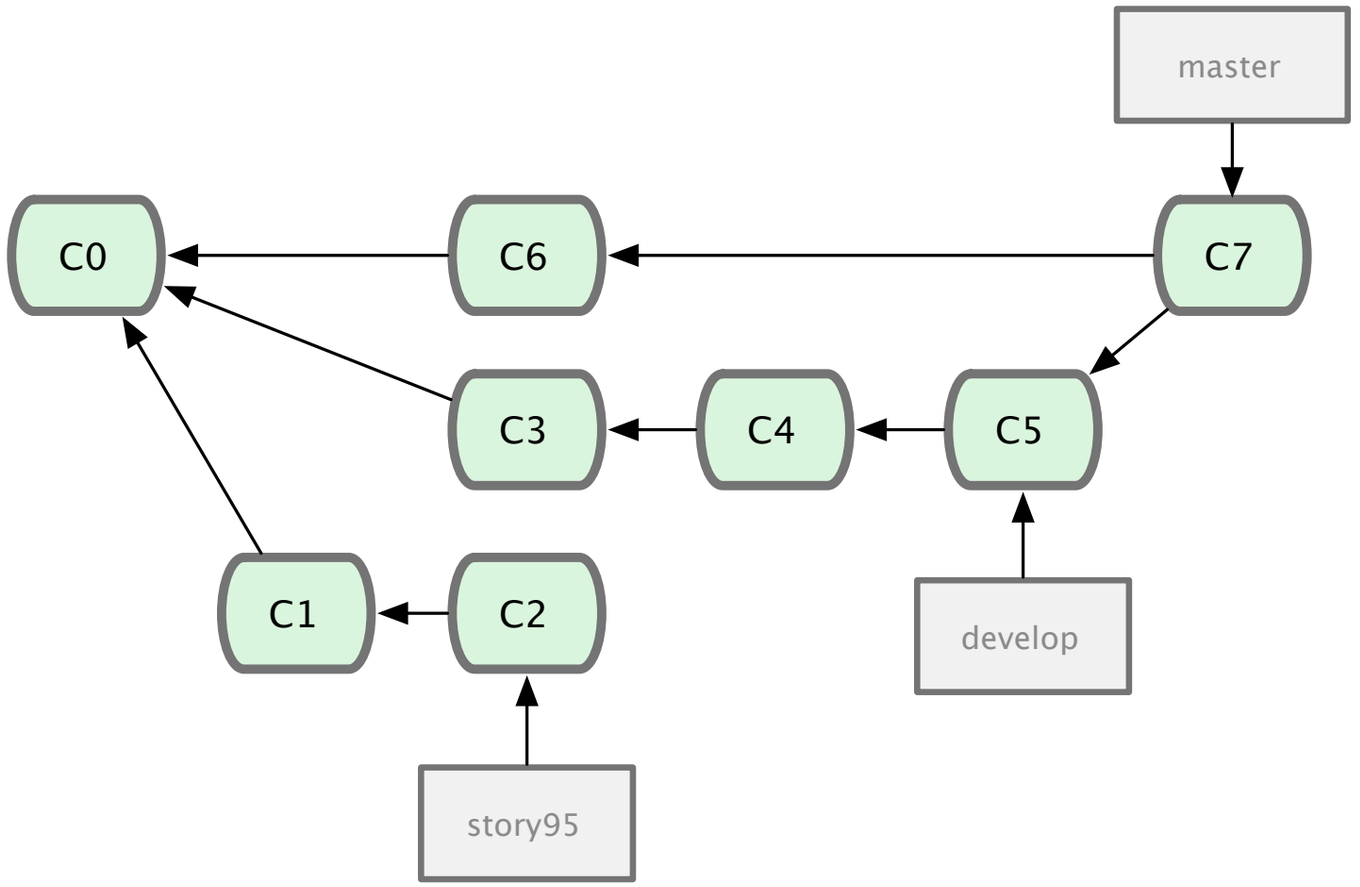
```
$>git checkout master
Switched to branch "master"
$>git rebase develop
First, rewinding head to replay your work on top of it...
HEAD is now at 8268c1a split show command into raw and normal
Applying changed log output to 30 from 25
$>
```



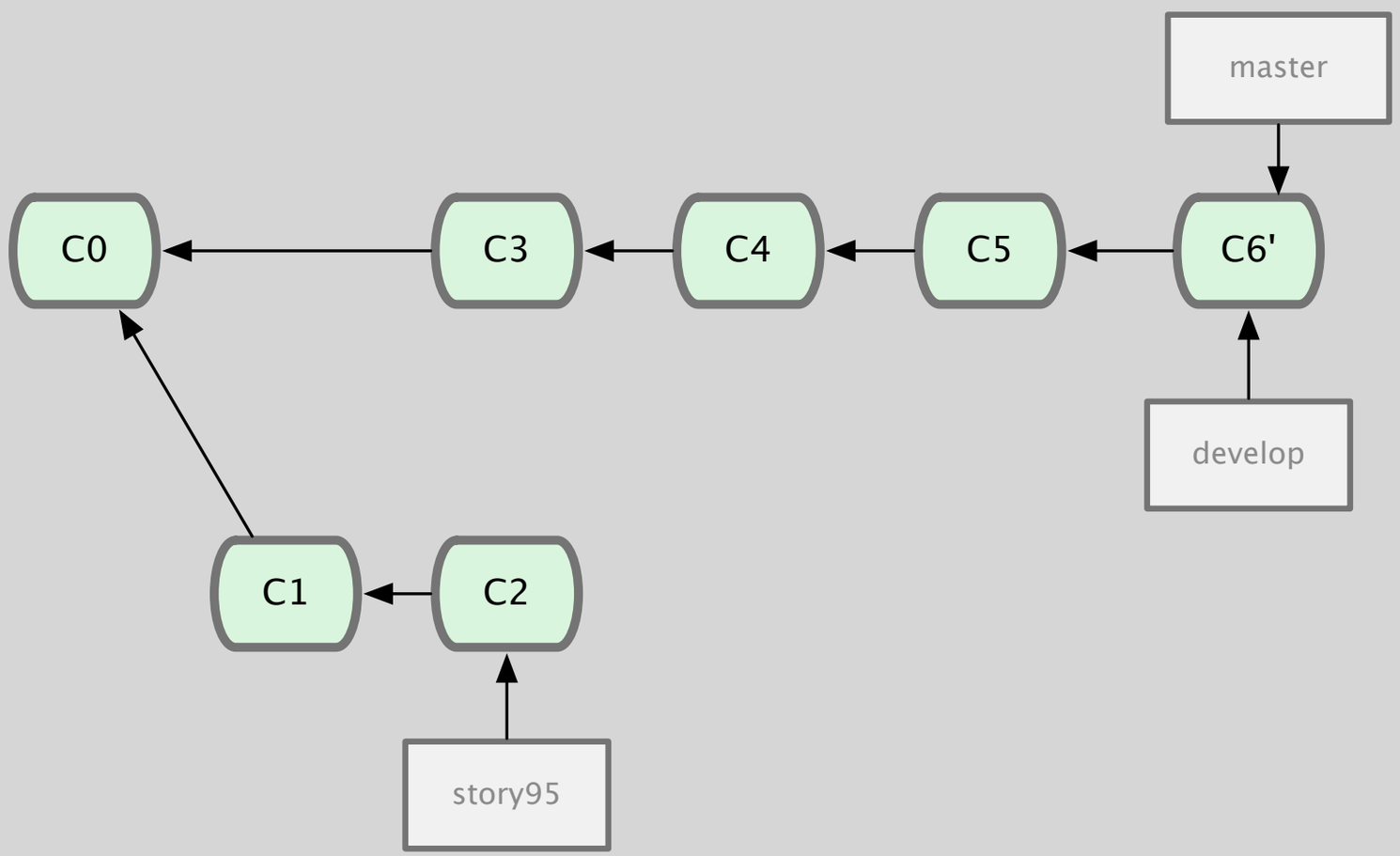
```
$>git checkout master
Switched to branch "master"
$>git rebase develop
First, rewinding head to replay your work on top of it...
HEAD is now at 8268c1a split show command into raw and normal
Applying changed log output to 30 from 25
$>
```



merge



rebase



sharing git

github
git repository hosting

schacon

0

account | profile | guides | log out

repositories: all | search

News Feed

for you | from you

Magnus Chacon committed to simplegit 4 minutes ago

fc62e5543b195f18391886b9f663d5a7eca38e84

added file_size

Magnus Chacon committed to simplegit 27 minutes ago

f40831985e265eb8cca2d0e66c11ac48ccfe170d

changes

Magnus Chacon committed to simplegit 27 minutes ago

25ce6394a94d8c3c30240960b386f79f4ebdd1da

changes

Magnus Chacon committed to simplegit 27 minutes ago

3e625e2139a71bfbaa85b055c139bceafab7e2c0

updated rakefile

Magnus Chacon committed to simplegit 27 minutes ago

4fefbaafd51ff3e79b84e9ee1a0dac2848e21a98

commit the todo

Your Repositories (create a new one)

all | public | private | sources | forks

- capistrano
- dst
- facebox
- git-ruby
- git-source
- git-wiki
- grit
- munger
- rest-client
- ruby-git
- simplegit
- ticgit

Watched Repositories

github
git repository hosting

schacon

account | profile | guides | log out

0

repositories: all | search

News Feed

for you | from you

Magnus Chacon committed to simplegit 4 minutes ago

fc62e5543b195f18391886b9f663d5a7eca38e84

added file_size

Magnus Chacon committed to simplegit 27 minutes ago

f40831985e265eb8cca2d0e66c11ac48ccfe170d

changes

Magnus Chacon committed to simplegit 27 minutes ago

25ce6394a94d8c3c30240960b386f79f4ebdd1da

changes

Magnus Chacon committed to simplegit 27 minutes ago

3e625e2139a71bfbaa85b055c139bceafab7e2c0

updated rakefile

Magnus Chacon committed to simplegit 27 minutes ago

4fefbaafd51ff3e79b84e9ee1a0dac2848e21a98

commit the todo

Your Repositories (create a new one)

all | public | private | sources | forks

- capistrano
- dst
- facebox
- git-ruby
- git-source
- git-wiki
- grit
- munger
- rest-client
- ruby-git
- simplegit
- ticgit

Watched Repositories

Create a New Repository

Create a new empty repository into which you can push your local git repo.

NOTE: If you intend to push a copy of a repository that is already hosted on GitHub, then you should fork it instead.

Project Name

simplegit2

Description

another repo for simplegit for demo puposes

Homepage URL

Who has access to this repository? (You can change this later)

- (•) Anyone (learn more about **public** repos)
- ( ) Only the people I specify (learn more about **private** repos)

Submit

github
git repository hosting

schacon | account | profile | guides | log out

0 | repositories: all | search

schacon / simplegit2 edit

this repo is viewable by everyone

Description:	another repo for simplegit for demo puposes edit
Homepage:	Click to edit edit
Public Clone URL:	git://github.com/schacon/simplegit2.git
Your Clone URL:	git@github.com:schacon/simplegit2.git
Push URL:	git@github.com:schacon/simplegit2.git

github
git repository hosting

schacon | account | profile | guides | log out

0 | repositories: all | search

schacon / simplegit2 edit

this repo is viewable by everyone

Description:	another repo for simplegit for demo puposes edit
Homepage:	Click to edit edit
Public Clone URL:	git://github.com/schacon/simplegit2.git
Your Clone URL:	git@github.com:schacon/simplegit2.git
Push URL:	git@github.com:schacon/simplegit2.git

```
$>git remote
$>git remote add public git@github.com:schacon/simplegit2.git
$>git push public master
Counting objects: 115, done.
Compressing objects: 100% (73/73), done.
Writing objects: 100% (115/115), 11.20 KiB, done.
Total 115 (delta 42), reused 65 (delta 24)
refs/heads/master: 0000000000000000000000000000000000000000 -> fc6
5543b195f18391886b9f663d5a7eca38e84
To git@github.com:schacon/simplegit2.git
 * [new branch]      master -> master
$>git remote
public
$>git remote show public
* remote public
  URL: git@github.com:schacon/simplegit2.git
  Tracked remote branches
    master
$>_
```

github
git repository hosting

schacon
0

account | profile | guides | log out
repositories: all | search

Source Browser | Commits | Wiki | Network (1) | Admin

master | all branches | all tags

schacon / simplegit2 edit | pull request | unwatch | download

this repo is viewable by everyone

Description: another repo for simplegit for demo puposes edit
Homepage: Click to edit edit
Public Clone URL: git://github.com/schacon/simplegit2.git
Your Clone URL: git@github.com:schacon/simplegit2.git
Push URL: git@github.com:schacon/simplegit2.git

added file_size

Magnus Chacon (author)
17 minutes ago

commit fc62e5543b195f18391886b9f663d5a7eca38e84
tree 3ec275327fe6f8474fd4d044ebb1adfe04ca38db
parent f40831985e265eb8cca2d0e66c11ac48ccfe170d

simplegit2 /

name	age	message	history
README	about 14 hours ago	updated README [Magnus Chacon]	
Rakefile	39 minutes ago	changes [Magnus Chacon]	
TODO	39 minutes ago	changes [Magnus Chacon]	
lib/	16 minutes ago	added file_size [Magnus Chacon]	

distributed workflow

fetch, pull and push

```
$>git remote
github
$>git remote show github
* remote github
  URL: git://github.com/schacon/simplegit.git
  New remote branches (next fetch will store in remotes/github)
    master
$>git fetch github
remote: Generating pack...
remote: Done counting 106 objects.
remote: Result has 101 objects.
remote: Deltifying 101 objects...
remote:  100% (101/101) done
remote: Total 101 (delta 35), reused 101 (delta 35)
Receiving objects: 100% (101/101), 9.49 KiB, done.
Resolving deltas: 100% (35/35), done.
From git://github.com/schacon/simplegit
 * [new branch]      master     -> github/master
$>git branch -a
  develop
  master
* story95
  github/master
$>_
```

```
$>git remote
public
$>git remote show public
* remote public
  URL: git@github.com:schacon/simplegit.git
  Tracked remote branches
    master
$>git push public master
Counting objects: 7, done.
Compressing objects: 100% (3/3), done.
Writing objects: 100% (4/4), 365 bytes, done.
Total 4 (delta 2), reused 0 (delta 0)
refs/heads/master: f40831985e265eb8cca2d0e66c11ac48ccfe170d -> fc62
5543b195f18391886b9f663d5a7eca38e84
To git@github.com:schacon/simplegit.git
   f408319..fc62e55  master -> master
$>_
```

- **Local Commands**
  - git config
  - git init
  - git add
  - git commit
  - git status
  - git tag
  - git log
- **Branchy Commands**
  - git checkout
  - git branch
  - git merge
  - git rebase
- **Remotey Commands**
  - git remote
  - git fetch
  - git pull
  - git clone
  - git push
- **Patchy Commands**
  - git diff
  - git apply
  - git format-patch
  - git am