

Git

Stefano Entatti

1 Cos'è un sistema di controllo di versione

Nella realizzazione di un progetto software potresti esserti accorto di voler tenere traccia dei tuoi progressi, magari registrandoli in modo da sapere cosa hai fatto e quando, o di voler tornare al momento esattamente prima a quella modifica che non fa funzionare più nulla.

Potresti provare, ogni volta che apporti modifiche importanti, a creare una nuova copia del progetto, numerarla e magari annotarti quali cambiamenti ha portato. Man mano che il progetto avanza, le copie possono diventare moltissime e serve comunque confrontare i file riga per riga per capire cosa cambia da una copia all'altra.

Se al progetto lavorano più persone, (e quindi carichi il progetto su drive) non è comunque possibile che più persone modifichino il codice nello stesso momento, e se ognuno modifica una copia diversa, queste vanno poi unite sempre a mano con dei copia e incolla.

Un sistema di controllo di versione tiene traccia di tutte le modifiche effettuate su un insieme di file (che compongono il progetto), ovvero di chi le ha fatte, quando, quali righe sono state tolte e quali aggiunte. Ogni registrazione di modifica riporta l'autore, un titolo e una descrizione che spiega in dettaglio il perchè e stata necessaria o qualche peculiarità della modifica stessa in modo da facilitare la comprensione ad altri sviluppatore o ad una rilettura in futuro. Inoltre in questo modo è possibile navigare e riportarsi in qualsiasi punto della storia del progetto.

Per permettere a più persone di lavorare nello stesso momento ognuno copia in locale il progetto da un punto di riferimento (server comune ad esempio), e dopo aver apportato le modifiche necessarie è possibile unirle tenendo traccia di chi ha modificato cosa ed in quale ordine.

Esistono molti software di controllo versione, tra i più famosi ci sono Git, CVS, Subversion, Mercurial, Bazaar e BitKeeper.

2 Il sistema di controllo versione Git

Git è il sistema di controllo versione più utilizzato, ed è anche uno dei più semplici da usare. È nato nel 2005 ed è stato ideato da Linus Torvalds (il creatore di Linux) per facilitare lo sviluppo del kernel di Linux, uno dei progetti opensource più grandi, a cui lavorano moltissime persone. Anche git è opensource.

Git è distribuito, ciò significa che l'intero **repository**, ovvero tutti i file che compongono un progetto, tra cui sorgenti, readme e file creati da git per tenere traccia dello storico delle modifiche può essere memorizzato su un server a cui possono accedere tutti i membri del progetto. Allo stesso tempo ognuno può tenere una copia completa del repository in locale e utilizzare tutte le funzionalità di git anche **offline**. L'interazione col server è limitata a quando lo sviluppatore desidera scaricare le modifiche effettuate da altri o vuole caricare le proprie sul server.

Git è nato come un programma a riga di comando, e anche se viene considerato più efficiente se utilizzato in questo modo, oggi esistono molti client grafici e estensioni che permettono di usufruire della maggior parte delle funzionalità di git direttamente dall'IDE.

3 Branching

Git è in grado di gestire contemporaneamente più sviluppi all'interno della stessa copia del progetto, questi si chiamano **branch**. Con più sviluppi si intendono più tipologie di modifiche differenti anche all'interno dello stesso file. Si può saltare da uno sviluppo all'altro senza perdere nessuna modifica evitando di avere più copie del progetto.

Ogni repository git possiede almeno un branch, chiamato **master**. Possono essere creati infiniti branch, ognuno dei quali può essere utilizzato per la realizzazione di una particolare funzionalità, in modo che modifiche su parti molto diverse del codice non possano causare conflitti. Spesso un singolo sviluppatore lavora su un proprio branch, in modo da non dover gestire anche i cambiamenti fatti da altri mentre svolge il suo compito.

Oppure il branch master può essere quello che viene distribuito agli utilizzatori del software perchè considerato stabile, mentre gli altri possono essere sperimentali e dunque contenere bug o modifiche che possono essere scartate senza compromettere tutto il resto.

A volte vari branch servono a separare versioni con funzionalità leggermente diverse dello stesso software.

Un branch viene creato a partire da un altro. Dopo la creazione di un branch “figlio” le modifiche che nel tempo vengono apportate su questo rimangono completamente separate da quelle che eventualmente continuano ad essere apportate sul “padre”. Se vengono creati più branch, almeno uno di questi discende da master.

Ovviamente, due branch possono essere uniti tramite un **merge**. Gli algoritmi di merge confrontano due branch riga per riga, mantenendo quella appartenente al branch in cui è stata modificata. Se però la stessa riga è stata cambiata in entrambi i branch si crea un conflitto e git richiede all’utente di risolverlo a mano.

Si può passare in qualsiasi momento da un branch all’altro.

4 Comandi principali

Git offre moltissimi comandi, ognuno dei quali svolge un piccolo insieme di funzionalità. Ogni comando (in questa sezione tratteremo i principali) è sempre preceduto da “git”. Ognuno può accettare un diverso numero di argomenti e di opzioni (tratteremo le principali per ogni comando), spesso precedute da un singolo o da un doppio meno.

4.1 init, clone

Per creare un nuovo repository locale si utilizza init (stando all’interno della cartella del progetto del quale si vuole tenere traccia):

```
$ git init prova
Inizializzato repository Git vuoto in /home/user/prova/.git/
```

Per copiare un intero repository remoto si utilizza git clone. Viene scaricato solo il branch master, ma una volta clonato si può accedere a tutti gli altri

```
$ git clone https://github.com/torvalds/linux.git
```

Per utilizzare tutti gli altri comandi, occorre posizionarsi all’ interno della cartella del repository, che viene creata in automatico da clone e init.

4.2 Commit

Ogni volta che si fa una certa quantità di cambiamenti è utile fare un commit, ovvero segnare un punto nello sviluppo a cui sarà sempre possibile tornare, quindi registrare e descrivere queste modifiche.

```
$ git commit
```

Questo comando aprirà l’editor di default di git. Nella prima riga va scritto il titolo del commit. È una buona norma che il titolo mantenga una lunghezza massima di 72 caratteri e contenga solo nomi e verbi al presente. La seconda si lascia sempre vuota e dalla terza inizia la descrizione, che può essere molto lunga e dettagliata per spiegare in modo discorsivo cosa si è fatto, perché e se eventualmente ha causato dei problemi.

se non si necessita di una descrizione si può utilizzare l’opzione -m (“message”):

```
$ git commit -m "add options page"
```

Ci sono varie teorie sulla lunghezza e il contenuto dei messaggi e delle descrizioni dei commit e su *ogni quanto* si debba committare. In genere un commit deve essere relativo ad un solo argomento e non comprendere modifiche totalmente indipendenti tra di loro. I messaggi di commit non devono essere generici (come “fix crash”) altrimenti col passare del tempo sarà impossibile capire cosa si era fatto senza controllare il codice.

L’opzione -s (“signed”) aggiunge la firma dell’autore nella descrizione:

```
Signed-off-by: user1 <user1@gmail.com>
```

```
1 ordine più logico nella sezione di branch, checkout
2
3 gli esempi di codice nella sezione dei comandi branch e checkout segue
4 l'ordine dei comandi che si fanno normalmente per creare e eliminare
5 branch
6
7 # Immetti il messaggio di commit per le modifiche. Le righe che iniziano
8 # con '#' saranno ignorate e un messaggio vuoto interromperà il commit.
9 #
10 # Data: Thu Feb 20 22:05:02 2020 +0100
11 #
12 # rebase interattivo in corso su a0ecc94
13 # Ultimo comando eseguito (1 comando eseguito):
14 # reword 69ec3f1 ordine più logico nella sezione di branch, checkout
15 # Nessun comando rimanente.
16 # Attualmente stai modificando un commit durante il rebase del branch 'master' su 'a0e
   cc94'.
17 #
18 # Modifiche di cui verrà eseguito il commit:
19 #      modificato: git.pdf
20 #      modificato: git.tex
21 #
~
~
```

Figure 1: scrittura del testo di un commit in vim

4.3 Add, rm, reset, status

I file coinvolti dal commit devono essere prima selezionati con add. In questo modo si entra nella **staging area**. è uno stato intermedio che sta prima di un commit per tracciare le modifiche momentanee in caso di piu' prove.

Se ad esempio si modificano functions.cpp, functions.hpp e main.cpp:

```
$ git add functions.cpp functions.hpp
$ git commit -m "added get function"
```

In questo caso le modifiche di main.cpp non verranno aggiunte al commit added get function.

Qualsiasi modifica effettuata su functions.cpp o functions.hpp dopo l'utilizzo di add verrebbe anch'essa esclusa dal commit.

Git add si comporta in modo indifferente sia per file appena creati che per le modifiche di file già esistenti.

L'opzione -a ("all") passata al comando di commit include automaticamente tutte le modifiche attualmente pendenti.

Alcune opzioni utili per add:

- -A aggiunge qualsiasi modifica all'area di staging
- . come -A ma non aggiunge la rimozione dei file
- -u non aggiunge i nuovi file

Il comando git rm fa il contrario di add, mentre reset pulisce completamente la staging area.

Per vedere quali modifiche sono nella staging area e quali invece non sono ancora state aggiunte con add:

```
$ git status
Sul branch master
Il tuo branch è aggiornato rispetto a 'origin/master'.
```

Modifiche di cui verrà eseguito il commit:

(usa "git restore --staged <file>..." per rimuovere gli elementi dall'area di staging)

```
modificato: git.pdf
modificato: git.tex
```

Modifiche non nell'area di staging per il commit:

```
(usa "git add <file>..." per aggiornare gli elementi di cui sarà eseguito il commit)
(usa "git restore <file>..." per scartare le modifiche nella directory di lavoro)
modificato:      README.md
```

Questo comando mostra anche informazioni relative al branch su cui si è Posizionati e se si è aggiornati rispetto al remote (vedere 5.3).

4.4 Push

Permette di caricare un numero illimitato di commit su un branch di un repository remoto. Le modifiche locali vengono unite a quelle remote.

Git non permette di effettuare un push se la **storia**, intesa come sequenza di commit, del branch remoto non è uguale al branch locale, escludendo i commit appena aggiunti fatti in locale. Se ci si trova in questa situazione occorre effettuare un riallineamento, (pull/rebase).

```
$ git push
Username for 'https://github.com': Stivvo
Password for 'https://Stivvo@github.com':
To https://github.com/Stivvo/msTest.git
 ! [rejected]      master -> master (fetch first)
error: push di alcuni riferimenti su 'https://github.com/Stivvo/msTest.git' non riuscito
Gli aggiornamenti sono stati rifiutati perché il remoto contiene delle
modifiche che non hai localmente. Ciò solitamente è causato da un push
da un altro repository allo stesso riferimento. Potresti voler integrare
le modifiche remote (ad es. con 'git pull ...') prima di eseguire
nuovamente il push.
Vedi la 'Nota sui fast forward' in 'git push --help' per ulteriori
dettagli
```

4.5 Merge, Pull, fetch

Merge permette di fondere le modifiche di due branch, locali o remoti (o anche commit diversi dello stesso branch) come ad esempio quando devo unire nel branch principale le modifiche fatte su un branch per correggere dei bug. Se si effettua un pull come suggerito sopra, avverrà infatti un merge. Se git rileva dei conflitti segnala all'utente di risolverli manualmente.

Fetch permette di aggiornare lo stato dei branch in remoto per controllare se ci sono branch nuovi o magari nuovi commit sul branch al quale si sta lavorando per evitare di rimanere disallineati con il repository di riferimento.

Pull è in sostanza un fetch seguito da un merge, ed è quello che capita di utilizzare più spesso. In generale pull fonde ciò che si trova al momento sul branch del repository remoto con quello locale.

Pull e fetch chiamati senza argomenti vanno a prelevare la versione remota del branch su cui si è localmente.

```
<<<<<< HEAD
class FirstClass {
=====
class SecondClass {
>>>>>> 4ceb8e7c4fe78b59c00be99418f54280df19078c
```

Questo è il caso in cui mentre il repository locale rimaneva indietro di alcuni commit la classe è stata rinominata in FirstClass da un primo sviluppatore. Nel frattempo un secondo ha fatto un push di un commit in cui l'ha chiamata SecondClass. Quando il primo sviluppatore si trova a dover fare il pull delle modifiche del secondo, per risolvere i conflitti di merge deve scegliere tra la versione locale (HEAD) e quella dell'altro, identificata dal codice hash (vedere 4.7) del commit che ha rinominato la classe in SecondClass. Il prossimo push sarà quello del commit di merge, automaticamente creato da git.

4.6 Checkout, branch, merge

Git branch senza opzioni viene utilizzato per creare un nuovo branch locale. La creazione del branch develop:

```
$ git branch develop
```

Per posizionarsi su develop:

```
$ git checkout develop
M git.pdf
M git.tex
Si è passati al branch 'develop'
```

Se si prova a eseguire un push dal branch appena creato occorre aggiungerlo alla lista dei branch remoti:

```
fatal: Il branch corrente develop non ha alcun branch upstream.
Per eseguire il push del branch corrente ed impostare il remoto come upstream, usa
```

```
git push --set-upstream origin develop
```

L'opzione `-all` mostra tutti i branch locali e remoti. Il branch seguito dall'asterisco è quello su cui si è posizionati correntemente.

```
$ git branch --all
* develop
  master
  temp
remotes/origin/HEAD -> origin/master
remotes/origin/develop
remotes/origin/master
```

In questo caso, `temp` è solamente locale.

L'opzione `-d` invece elimina un branch. Non è possibile eliminare il branch corrente:

```
$ git branch -d develop
error: Impossibile eliminare il branch 'develop' di cui è stato eseguito
il checkout in '/home/stefano/prog/GitNoob2ProIta'
```

Per eliminare lo stesso branch anche dal repository remoto:

```
$ git push -d origin develop
```

L'opzione `-b` di `checkout` crea un branch se quello passato come parametro non esiste, utilizzando quindi prima un `git branch` e poi un `git checkout`.

Per fondere due branch si utilizza ovviamente `merge`. In questo caso si vuole portare le modifiche di `develop` su `master`:

```
$ git checkout master
$ git merge develop
Updating f42c576..3a0874c
Fast-forward
 git.tex | 2 ++
 1 file changed, 2 insertions(+)
```

4.7 Log, Diff

Molto di quanto appena spiegato sarebbe inutile se non si potesse vedere la storia dei commit.

`Git log` mostra l'intera storia dei commit visualizzata nel pager (vedere 6) di default. Le lunghe serie di caratteri sono i codici **hash**, univoci per ogni commit. L'output del comando mostra anche a quale commit puntano la `HEAD` e i repository remoti (se abbiamo dei commit di cui non abbiamo fatto ancora il push è probabile che quest'ultima sia più indietro rispetto ad `HEAD`).

Questo comando può generare molto output. Sarà più semplice trovare un commit utilizzando l'opzione `pretty=oneline`, che assegna una sola riga ad ogni commit. Dopodiché sarà utile passare l'hash del commit trovato come argomento di `log`, per vedere informazioni più dettagliate. L'output di `log` escluderà semplicemente tutti i commit precedenti a quello.

È anche possibile ricercare il testo del commit interessato passandolo come argomento di `log` subito dopo `--grep=`:

```

commit 4b64be5218bed736d357d61471b87c4f5363d954 (HEAD -> master, origin/master, origin/HEAD)
Author: Stivvo <entattis15@itisvinci.com>
Date:   Sun Feb 23 17:37:09 2020 +0100

    git log

commit 559f87cd529b50efe231f45aa38afb7bfc0cdfbb
Author: Stivvo <entattis15@itisvinci.com>
Date:   Sat Feb 22 17:04:00 2020 +0100

    tagging

commit e640eb5a51bee98b0372ddf28c704d330019ae79
Author: Stivvo <entattis15@itisvinci.com>
Date:   Sat Feb 22 16:38:52 2020 +0100

    organizzazione link

    - link divisi in subsection
    - inizio tag

commit 43615496b72a11c0c1339c34093599292fb1acbc
Author: Stivvo <entattis15@itisvinci.com>
Date:   Sat Feb 22 13:48:35 2020 +0100

    tornare ad un commit precedente

    - migliore spiegazione di come uscire da detached head mantenendo le
      modifiche (reset)
    - revert come strada alternativa

```

Figure 2: scrittura del testo di un commit in vim

```

$ git log --grep='git log'
commit 4b64be5218bed736d357d61471b87c4f5363d954 (HEAD -> master, origin/master, origin/HEAD)
Author: Stivvo <entattis15@itisvinci.com>
Date:   Sun Feb 23 17:37:09 2020 +0100

```

```
git log
```

Si può ottenere la lista dei commit in cui è stata aggiunta o rimossa una determinata stringa all'interno dei file del repository passandola come parametro dopo l'opzione -S ("string").

Se invece si è interessati a vedere tutti i commit effettuati da una stessa persona:

```
$ git log --author="Stivvo"
```

5 la cartella .git

La cartella .git si trova nella root del repository e contiene tutti i file utilizzati da git, tra cui informazioni sui branch, sui commit. Nei sistemi operativi unix una cartella preceduta da un punto è nascosta e quindi occorre utilizzare il parametro -a di ls per poterla vedere.

```

$ ls .git/
branches/ COMMIT_EDITMSG config description FETCH_HEAD HEAD hooks/ index
info/ logs/ objects/ ORIG_HEAD packed-refs refs/

```

è molto importante la cartella refs. Contiene:

```

$ ls .git/refs
heads/ remotes/ tags/

```

5.1 heads

In git una head è un riferimento ad un branch o ad un commit di un determinato branch, locale o remoto. Una lista delle head disponibili:

```
$ ls .git/refs/heads
master/ develop/
```

HEAD è un file che punta all'ultimo commit del branch in cui si è attualmente posizionati nel repository locale.

```
$ cat .git/HEAD
ref: refs/heads/master
```

Nel caso in cui si voglia ritornare ad un commit precedente si entra nello stato di *deattached head*, ovvero facendo il checkout su uno specifico commit (identificato con il suo codice hash).

```
$ git checkout 8b10ce361a08e03179d46bab5d691148805bf8d8
Nota: esegui il checkout di '8b10ce361a08e03179d46bab5d691148805bf8d8'.
```

Sei nello stato 'HEAD scollegato'. Puoi dare un'occhiata, apportare modifiche sperimentali ed eseguirne il commit, e puoi scartare qualunque commit eseguito in questo stato senza che ciò abbia alcuna influenza sugli altri branch tornando su un branch.

Se vuoi creare un nuovo branch per mantenere i commit creati, puoi farlo (ora o in seguito) usando l'opzione `-c` con il comando `switch`. Ad esempio:

```
git switch -c <nome nuovo branch>
```

Oppure puoi annullare quest'operazione con:

```
git switch -
```

HEAD si trova ora a b0451d9 immagine scrittura commit

Se si vuole mantenere i commit fatti in questo stato è buona cosa spostarsi su un nuovo branch come suggerito.

Se si sceglie di rimanere sullo stesso, non si può effettuare direttamente il push dei commit effettuati in questo stato, perché non si è di fatto posizionati su nessun branch:

```
$ git push
fatal: Attualmente non sei su un branch.
Per eseguire ora il push della cronologia che ha condotto
allo stato corrente (HEAD scollegato), usa
```

```
git push origin HEAD:<nome del branch remoto>
```

Il comando suggerito da git serve per caricare le modifiche effettuate in detached head direttamente sul branch remoto, come spiegato nella sezione 5.3. è molto probabile che non funzioni, perché andrebbe ad eliminare delle modifiche remote successive al commit in cui si è entrati in detached head. Quindi è necessario aggiungere l'opzione `-f` ("force") a push se si vuole eliminarle.

Questo non risolve lo stato di detached head. Occorre infatti ritornare sul proprio branch (in questo caso master), che però contiene ancora i commit che vogliamo eliminare: un push li riporterebbe sul repository remoto. Quindi dopo aver fatto `git checkout master`, tornando sul branch da cui ci si era distaccati, si può utilizzare `reset`, che è come un pull forzato che invece di fare un merge sovrascrive il branch remoto su quello locale:

```
$ git reset --hard origin/master
```

Oppure si può sempre clonare nuovamente il progetto, ma è sempre la soluzione peggiore.

C'è un modo migliore per ritornare a un commit precedente, senza modificare i commit già pushati:

```
$ git revert 0552dd1c6e3c11c8c5246836e9994e6fcd431a0f..HEAD
$ git commit -m "torno al commit precedente"
```

In questo modo il commit `torno al commit precedente` conterrà delle modifiche che riportano allo stato del commit di cui si è specificato il codice come argomento di `revert`. `..HEAD` indica che si ripristinano le modifiche effettuate da quel commit fino a `HEAD` (questo intervallo può dunque comprendere diversi commit), ovvero lo stato corrente del branch.

5.2 Tags

Possono essere assegnati ad un commit in cui si è raggiunto un traguardo nel progetto (ad esempio 1.3.5).

Per visualizzare i tag:

```
$ git tag
v1.0
v2.0
```

Per creare un tag basta dare al comando un argomento, che sarà il nome del tag:

```
$ git tag v2.1
```

I tag possono essere utilizzati in questo modo per descrivere piccoli progressi nello sviluppo. Per segnare il punto di una release è bene utilizzare l'opzione `-a` (“annotated”):

```
$ git tag -a v3.0
```

In questo modo, verrà aperto l'editor di default per poter inserire informazioni su ciò che le novità portate da quella release. Come per i commit, l'opzione `-m` permette di scrivere un breve titolo senza aprire l'editor.

Per visualizzare queste informazioni:

```
$ git show v3.0
commit ca82a6dff817ec66f44342007202690a93763949
Author: User <user@email.com>
Date:   Mon Mar 17 21:52:11 2020 -0700
```

```
    new release 3.0!
```

Si può aggiungere un tag ad un qualsiasi commit precedente specificando il suo codice:

```
$ git tag -a v1.2 9fceb02
```

I tag possono anche essere eliminati con l'opzione `-d` e si può fare il checkout su uno specifico tag come si fa con i commit.

5.3 Remote

Un remote è il percorso di un repository remoto, di solito è quello del repository sul server. Il primo remote che viene utilizzato in un repository viene chiamato di default **origin** (non è obbligatorio), ma ne possono essere aggiunti altri.

... comando

Otteniamo una lista dei remote disponibili. Di default le operazioni come pull sottointendono che si voglia utilizzare il remote origin, ma l'operazione può essere eseguita su qualsiasi altro remote (ad esempio `git pull nuovoRemote`).

In un remote può essere aggiunto (`git remote add`), rinominato (`git remote rename`) o rimosso (`git remote remove`).

6 Installazione e configurazione

```
less -r config core.pager ""
```


7 Github

8 Approfondimenti

8.1 I submodule

8.2 Revisionare i commit

rebase, ammend

8.3 Gitignore

9 fonti, link utili

9.1 pagine ufficiali di git

- pro git
- tutti i comandi
- vantaggi di git

9.2 correlato

- benefici dei version controllo
- norme sulla scrittura dei commit
- norme sull'assegnazione dei nomi alle versioni
- guida a git log

9.3 stack overflow

- che cos'è head
- differenza tra pull e fetch
- annullare add
- parametri di add
- sovrascrivere con pull
- checkout di un branch remoto
- eliminare un branch
- tornare a commit precedenti
- cercare il messaggio dei commit
- cercare il testo modificato dai commit
- grep del testo modificato dai commit
- git log sempre colorato
- commit di una persona
- rinominare un branch
- cos'è head
- detached head

- risolvere una detached head
- differenza tra head, master e origin
- tipi di head
- push origin head
- uscire da detached head