

Git & github

Gestion collaborative des sources

Laurent Jouanneau Miage d'Evry octobre 2012

Licence d'utilistation : http://creativecommons.org/licenses/by-sa/3.0/









Principes des CVCS/DVCS



Comment sauvegarder?

- Dans un projet, les fichiers évoluent, sont créés, sont effacés
- Comment faire pour garder des versions intermédiaires de ces fichiers, notamment à chaque version du projet ?
 - 1) On fait une sauvegarde à la main dans des répertoires différents, dans des disques différents
 - → fastidieux, sujet à erreurs, collaboration difficile
 - 2) On utilise un logiciel dédié à cette tâche : un « VCS »



Définition

- VCS = Version Control System
- Un gestionnaire de version de sources stocke les changements successifs effectués sur des fichiers
- Les changements sont enregistrés dans une base de donnée spéciale appelée « dépôt ».
- C'est l'utilisateur qui lance l'enregistrement quand il le désire : il « commit ».
- Un enregistrement dans la base = un état à un instant t du projet = « changeset » ou « commit »



Pourquoi

- « Versionner » les changements sert à
 - Pouvoir tester des solutions et revenir en arrière
 - Retrouver les modifications qui ont provoqué des régressions, des bugs
 - Trouver les changements effectués pour une fonctionnalité similaire à celle que l'on veut développer → comprendre le fonctionnement de l'application
 - Permettre au système de gestion de pouvoir fusionner le travail de plusieurs développeurs → c'est un outils de collaboration indispensable pour le développement



CVCS et DVCS

- Il y a deux types de gestionnaire de source :
 - Les systèmes centralisés
 - Les systèmes décentralisés



Central VCS

- Central Version Control System
- Un serveur stocke toutes les modifications
- Le client dialogue avec le serveur pour toute opération
- Le client ne possède qu'une version des fichiers
- schema



Central VCS

- Avantages
 - Tout le monde a accès aux dernières modifications
 - Réglage fin des droits d'accès
- Inconvénients
 - Nécessité d'avoir accès au serveur pour travailler, pour commiter et faire la plupart des opérations (log, diff...)
 - L'historique est stocké à un seul endroit
 - Serveur inaccessible, base de donnée corrompue, détruite : gros impact sur le développement
- Exemple de logiciels : CVS, Subversion



Distributed VCS

- Distributed Version Control System
- Chaque utilisateur a sa propre copie du dépôt (un clone), avec tout l'historique et toutes les données
- Chacun peut récupérer les modifs d'un autre clone
- Il peut y avoir un dépôt « central », qui sert de référence
- Ex: Mercurial, Git, Bazaar, ...
- schema



Distributed VCS

Avantages

- Clonage → peu de chance de perdre les sources
- Pas besoin d'être connecté pour travailler
- Plus récents, donc plus évolués, plus puissants (merge ...)
- Expérimentation plus aisée et historisée

Inconvénients

- Un peu plus complexe à apprendre
- Moins évident à administrer un dépôt central (droits..)



Introduction sur Git



Pourquoi Git

- DVCS
- Moderne, puissant, performant
- Devient le VCS le plus utilisé
- Popularité « boostée » par les fonctionnalités offertes par le site Github, facilitant les collaborations



Historique

- Créé par Linus Torvalds
- En remplacement du VCS Bitkeeper (non libre) pour gérer le projet Linux, les autres VCS libres n'ayant pas les fonctionnalités voulues
- Prototype développé en 3 jours et sortie le 6 avril 2005
- Version 1.0 en décembre 2005
- Pour comparaison : CVS en 1990, Visual SourceSafe en 1994, Subversion et Bitkeeper en 2000, Mercurial en 2005, Bazaar en 2007.



Téléchargement et installation

- Git est avant tout un outil en ligne de commande
 - http://git-scm.com
- Téléchargement : http://git-scm.com/download/
 - Pour windows cela installe http://msysgit.github.com/
- Pour Linux, préférez passer par le système de dépôt de la distribution (apt-get install...)
- Suivre les instructions d'installation
- Ouvrez un terminal, c'est prêt.
 - git help



Fonctionnement interne de Git



Stockage

- Ne stocke pas les différences entre chaque version
- Stockes chaque version des fichiers dans une sorte de système de fichiers
- Intégrité du contenu → checksum
- Ajout d'information, pas de suppression

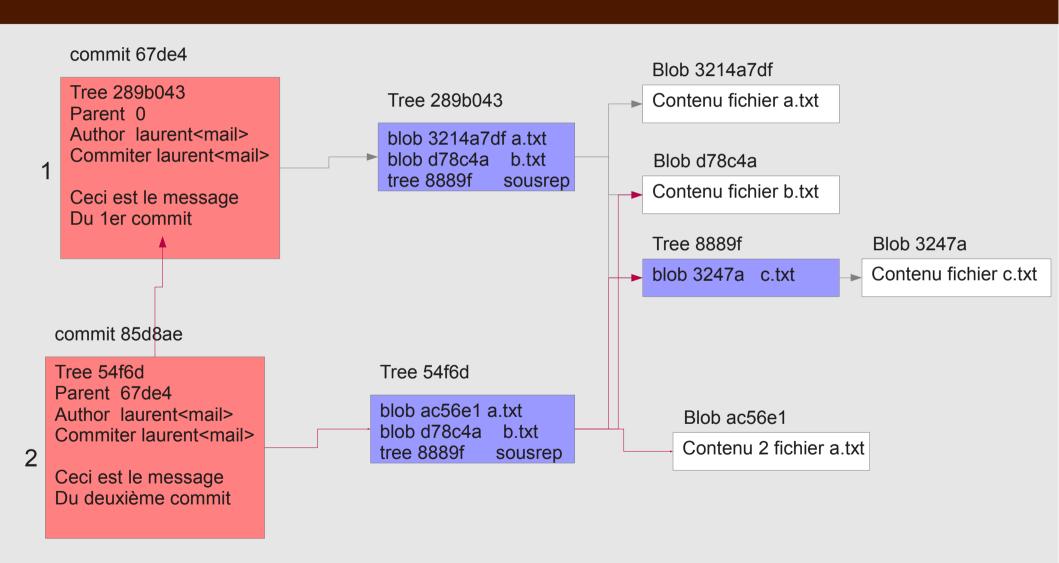


Stockage

- La base de donnée Git contient quatre type d'objets
- Ils sont tous identifiés par un nombre hexa qui est le checksum sha1 de leur contenu
- **Blob**: contient le contenu d'un fichier. Il ne stockes ni de nom, ni autres attributs git show 5529b198
- Tree: liste les fichiers d'un répertoire (nom + id du blob correspondant + permissions + type) git 1s-tree 6894ac45
- Commit: pointe vers le tree correspondant et le(s) commit(s) parent(s). Contient également message + auteur + date
- Tag : point vers un objet (commit), et identifié par un nom



stockage





Utilisation basique de Git



Configuration de Git

- Avant toute chose, il est préférable de configurer git pour indiquer :
 - Son nom
 - Son email
- git config --global user.name "Pierre Durant"
 git config --global user.email "Pierre@durant.name"
- Ces informations seront utilisées dans les commits
- Sockées dans ~/.gitconfig (--global) ou dans le fichier « config » du dépôt

```
git config --help

git config --global core.editor vim

git config --global color.ui true
```



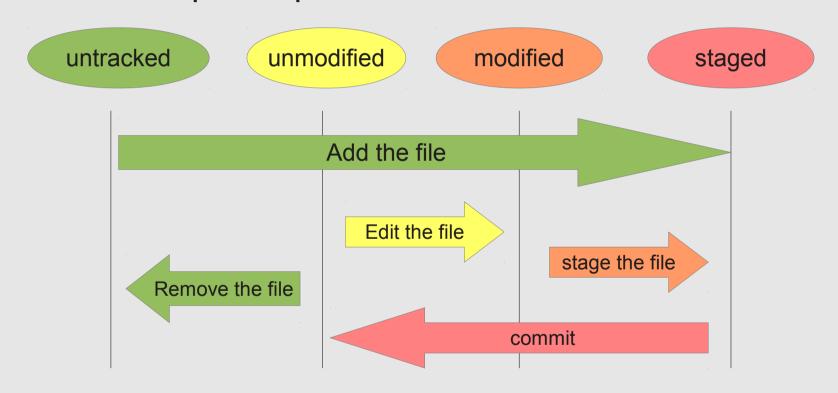
Création d'un dépôt

- On veut commencer à historiser les changements des fichiers d'un répertoire (d'un projet)
- On se place dans un répertoire puis : git init
- → répertoire .git = la « base de donnée » du dépôt



Principes de fonctionnement

Un fichier passe par 4 états



Pour savoir l'état des fichiers : git status



Principe de fonctionnement

- git add: pour ajouter un nouveau fichier dans le dépôt, ou pour indiquer de prendre en compte les modifications d'un fichier dans le prochain commit
- git rm: pour supprimer un fichier
- git mv : pour renommer un fichier
- git commit: pour valider les ajouts de fichiers ou de modifications de fichier, et valider les suppressions de fichier. Cette commande s'accompagne toujours d'un message expliquant le changement.



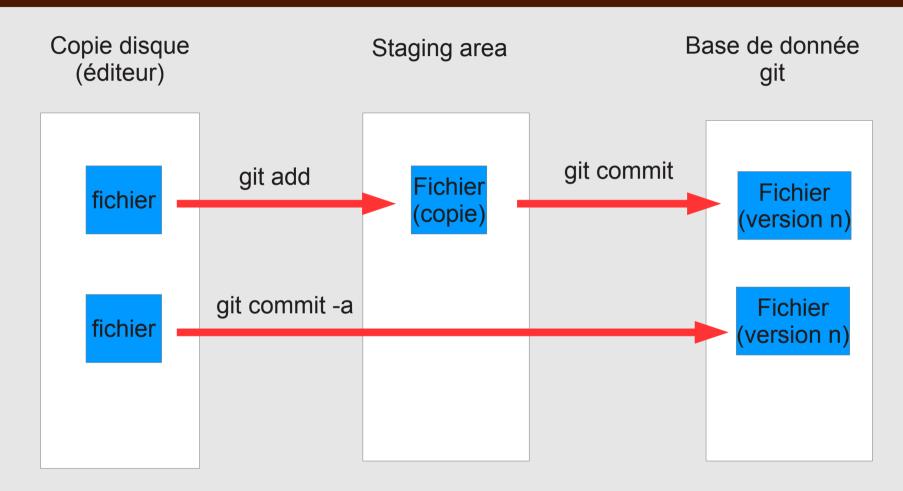
Staging area

- « zone » qui contient les fichiers en « staged »
- Elle contient donc la version des fichiers tels qu'ils étaient lors du git add.
- Ce sont ces versions qui seront sauvegardées lors du commit
- → permet de sélectionner les modifications pour le prochain commit
- Si un fichier « staged » est modifié avant le commit : il a alors les deux états, « modified » et « staged » correspondant à deux versions différentes du fichier. Seule la version « staged » est prise en compte dans le prochain commit.
- Pour commiter des changements sans passer par la staging area :

git commit -a



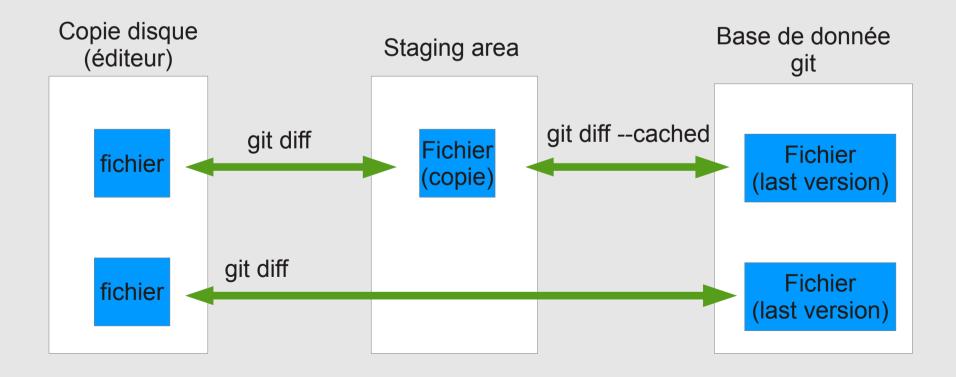
Staging area





Voir les modifications

- git diff: pour voir les modifications des fichiers ayant l'état « modified »
- git diff --cached : pour voir les modifications des fichiers ayant l'état « staged »





Annuler des modifications

- D'un fichier « modified »
- D'un fichier « staged »
- Modifier le dernier commit
- Annuler le dernier commit

```
git reset --hard HEAD^
```

attention, peut être dangereux

```
git checkout -- fichier.txt

git reset HEAD fichier.txt

git commit --amend
```

Commit

- Un commit est toujours accompagné par un texte, un message
 - Une première ligne faisant office de résumé
 - Si on veut ajouter des détails (pourquoi ces modifications, conséquences à prendre en compte par la suite etc...), plusieurs lignes de texte séparé du résumé par un saut de ligne.
- Paramètre « -m » en ligne de commande ou dans l'éditeur qui s'ouvre automatiquement

git commit -m "correction bug #1234 : crash lorsque v=0"



Quand commiter?

- Un commit = une étape d'un développement
 - Une petite fonctionnalité
 - Une étape de refactorisation
 - Correction d'un bug
- Pas de commit pour chaque modification de fichiers
- Pas de commit non plus contenant une tonne de modifications. Il y a forcément des étapes qui ont été faites.
- Faire en sorte que les modifications enregistrées dans un commit fassent sens : il faut pouvoir avoir un historique clair et utilisable = pouvoir identifier qui quoi quand pourquoi.



Ignorer des fichiers

- Fichier .gitignore
- Contient la liste des fichiers à ne pas prendre en compte dans le dépôt
- Exemple :

```
TAGS
.DS_Store
_dist/
*.komodoproject
.komodotools
```



Historique d'un dépôt

- git log
- gitk



Tags

- Étiquette pour retrouver facilement un commit particulier
- Le plus souvent pour étiqueter le commit qui finalise la sortie d'une version du projet.
- Pour étiqueter le dernier commit : git tag VERSION_1.0
- Lister les tags : git tag
- Mettre à jour le code source correspondant à un tag :

```
git checkout VERSION_1.0
```

• Supprimer un tag : git tag -d VERSION_1.0



Utilisation des branches avec Git



Qu'est-ce qu'une branche?

- Conceptuellement :
 - c'est une suite de commit qui a été réalisée en parallèle à une autre suite de commit
 - L'historique a donc divergé à un instant T
 - Il peut y avoir plusieurs branches en même temps. Elles ont chacune un nom
 - Branche principale : « master »
 - Une branche peut rejoindre une autre branche : « fusion de branche », ou « merge »



Pourquoi des branches?

- Création d'une branche pour le développement d'une fonctionnalité, la correction d'un bug, ou expérimenter : la branche principale reste ainsi dans un état « stable »
- Branches pour maintenir des versions stables : une branche par version, dans lesquelles on ne corrige que les bugs.
- Branches pour des versions particulières d'un projet
- Avec Git, la creation/fusion/suppression des branches est très peu couteux (disque, perf), on peut donc en utiliser tant qu'on veut



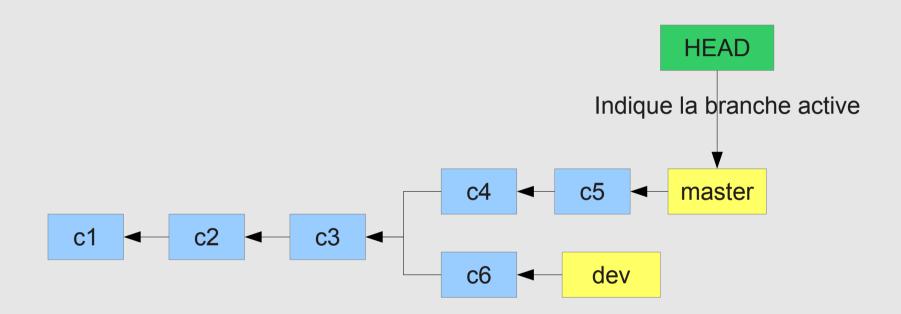
Qu'est-ce qu'une branche?

Techniquement :

- C'est un pointeur vers un commit, en l'occurrence, le dernier commit d'une suite de commit.
- Ce pointeur est modifié à l'ajout d'un nouveau commit, si ce dernier est le successeur du commit pointé.
- Pointeur spécial : HEAD, qui pointe vers le pointeur de la branche active



Qu'est-ce qu'une branche?





Créer / activer une branche

• Pour créer une branche :

```
git branch foo 5defc
```

- Le point de départ de la branche est le commit courant
- Pour activer une branche :

```
git checkout foo
```

- HEAD est mis à jour

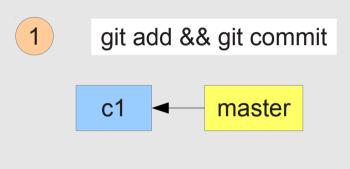
git checkout 84a96

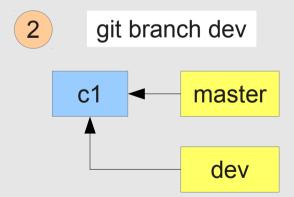
- Les sources du répertoire sont mis à jour
- Les prochains commits seront rattachés à la branche courante
- Création + activation en même temps : git checkout -b foo 65df
- Lister les branches

git branch

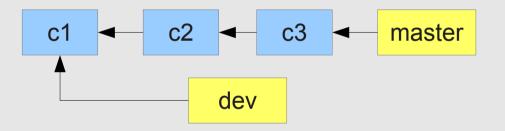


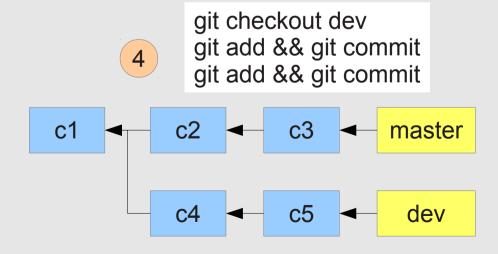
Utilisation des branches





git add && git commit git add && git commit





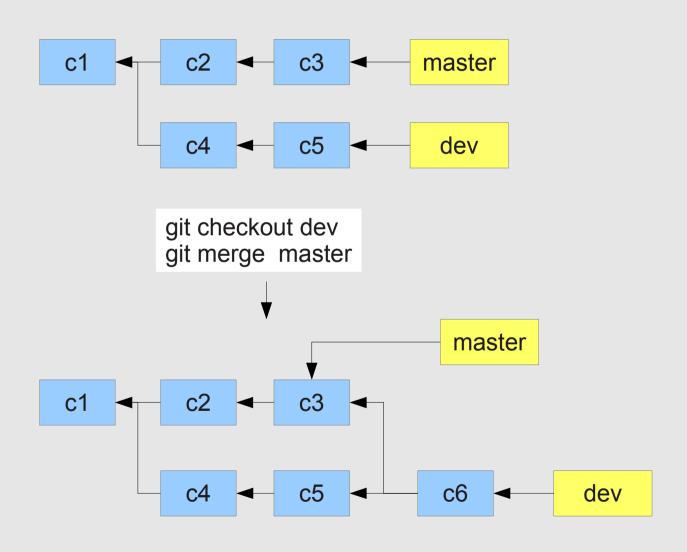


Fusion de branches

- But : intégrer les changements d'une branche (« test » par exemple) dans une autre (« master » par ex)
- Actions :
 - Se mettre dans la branche dans laquelle on veut intégrer les commits de l'autre branche
 git checkout master
 - Fusionner avec la commande merge
- La commande merge créé un nouveau commit dont les changements sont le résultat de la fusion



Fusion de branche





Conflits de fusion

- La fusion peut échouer à cause de conflits
- Des marqueurs ont été ajouté dans les fichiers concernés, permettant de repérer les parties de code en conflits

- Il faut alors corriger ces fichiers
- « git add »
- Puis commiter

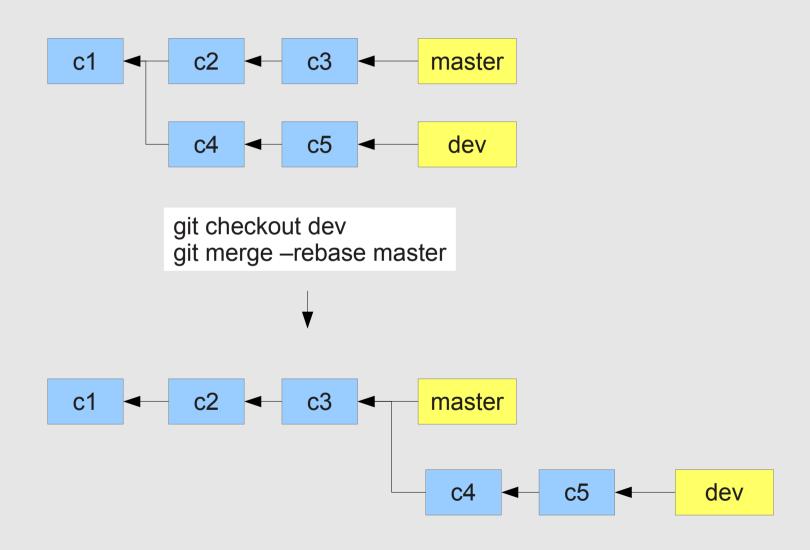


Rebasing

- Autre manière de fusionner les changements d'une autre branche
- Les modifications d'une branche à fusionner sont recommitées à la suite des commits de la branche courante
- → historique plus clair
- Important : les commits sont recréés ! Aussi, si la branche à fusionner a déjà été publiée, le rebase n'est pas possible, risque de disfonctionnement.
- Ne pas faire un rebase dans une branche dont tout les commits ont déjà été publiés (push)



rebasing





Développement collaboratif avec Git



Clonage d'un dépôt

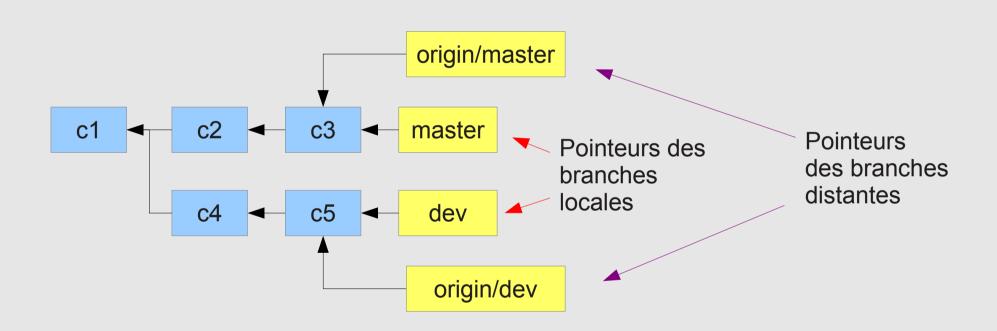
- Quand on travaille à plusieurs, il faut s'échanger les modifications, les commits
- En général : création du dépôt local en clonant un dépôt distant

git clone git@github.com:laurentj/comete2011.git

- Il y a alors les branches locales, et les branches distantes
- En local, les branches distantes se nomment « aliasdepot/nombranche »
- Le dépôt d'où l'on a cloné s'appelle « origin ».
- Le nom peut être changé



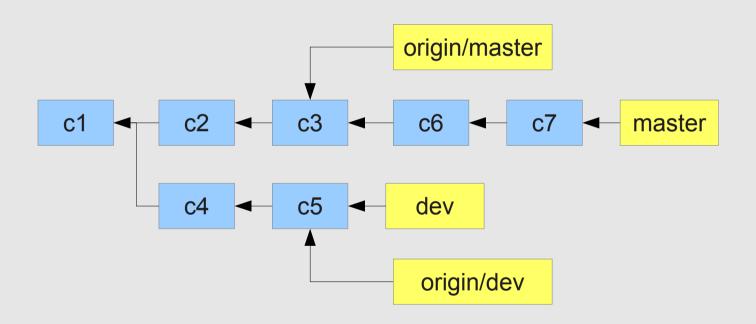
Clonage d'un dépôt





Clonage d'un dépôt

Aprés des commits (c6 & c7) dans master





Mises à jour : pull & push

Pour pousser les modifications vers le dépôt distant :

```
git checkout mabranche
git push origin branchedistante
```

• Pour récupérer des modifications :

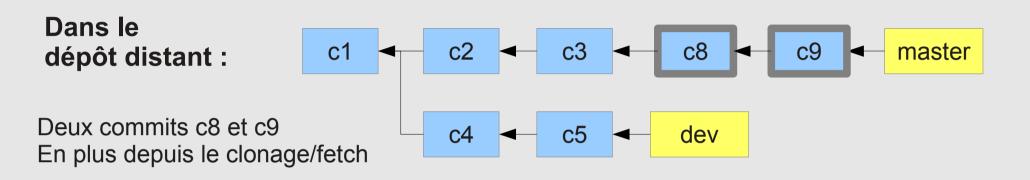
```
git checkout mabranche git pull origin branchedistante
```

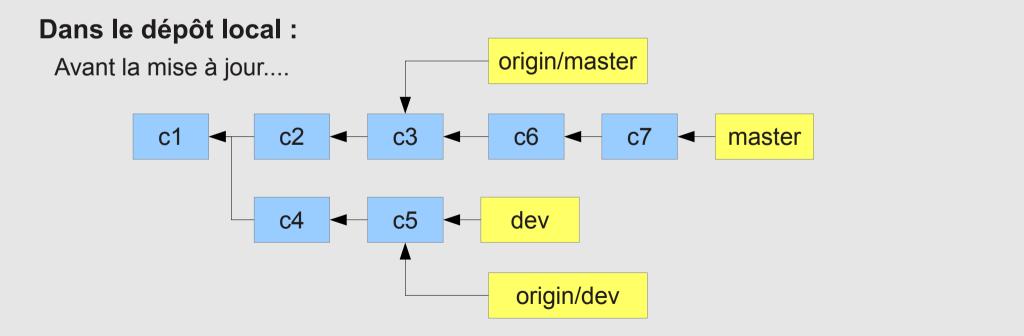
- Pull = fetch + merge
- Avec rebase (les commits locaux sont déplacés après le dernier commit de la branche distante)

```
git pull --rebase origin master
```



Fetch + merge #1

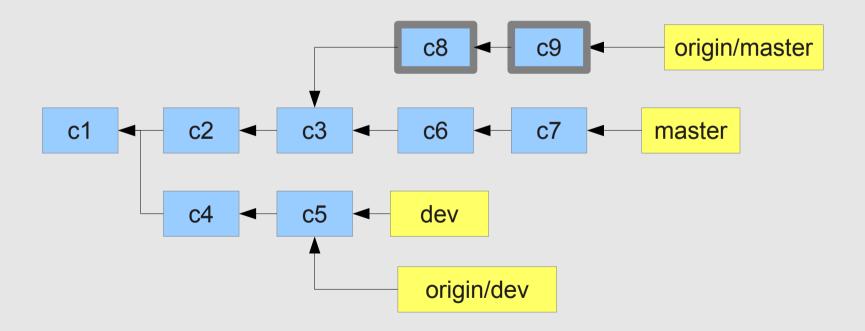






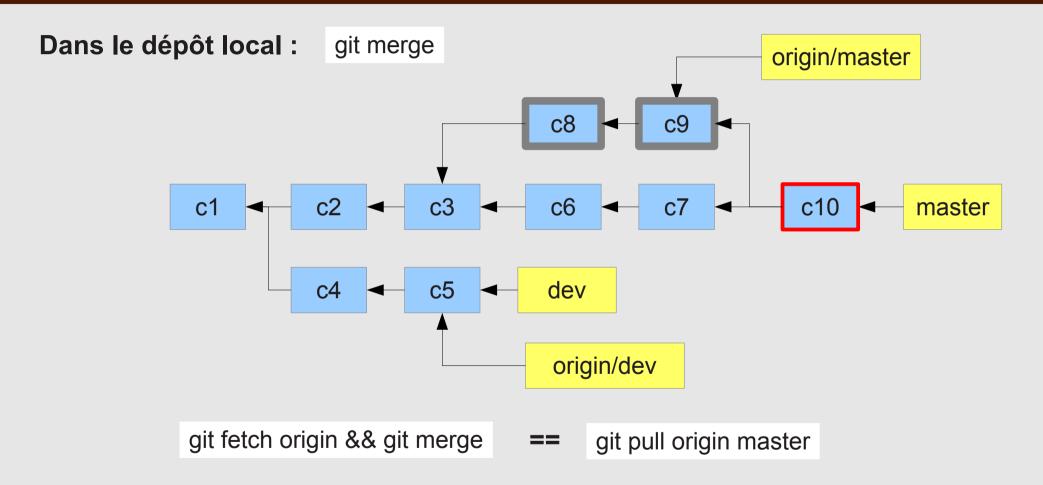
Fetch + merge #2

Dans le dépôt local : git fetch origin master





Fetch + merge #3 == pull

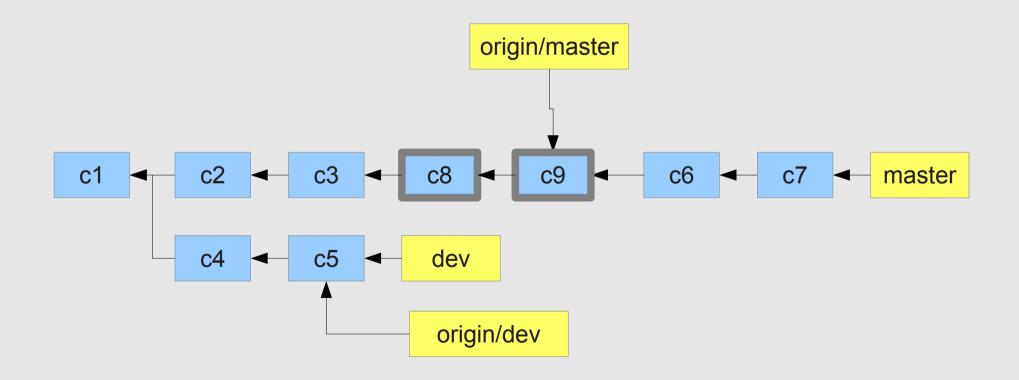


Après un git push, le dépôt distant contiendra les mêmes commit qu'en local



Fetch + rebase

Dans le dépôt local : git rebase (À la place de merge)



git fetch origin && git rebase

==

git pull –rebase origin master



Plusieurs dépôts distants

- À la place de « origin », on peut indiquer des chemins ou des URLS d'autres dépôts distants : on peut s'échanger des modifications entre plusieurs dépôts
- On peut créer des alias de ces URLS
- Gestion de ces alias et des informations des dépôts distant : « git remote »

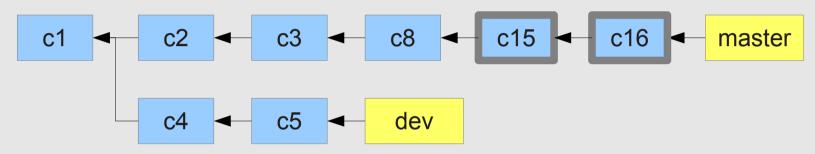
```
git remote # liste des dépôts
git remote add unAlias URLdepot

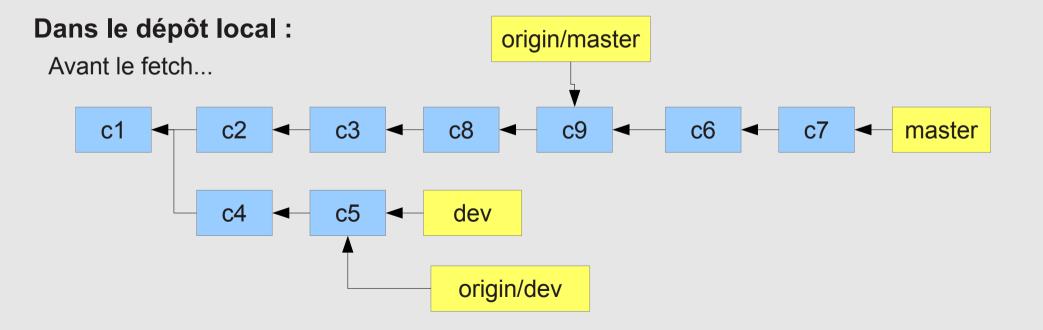
→ git pull unAlias master
```



2 dépôt distants

Dans le 2ieme dépôt distant (celui de paul) :



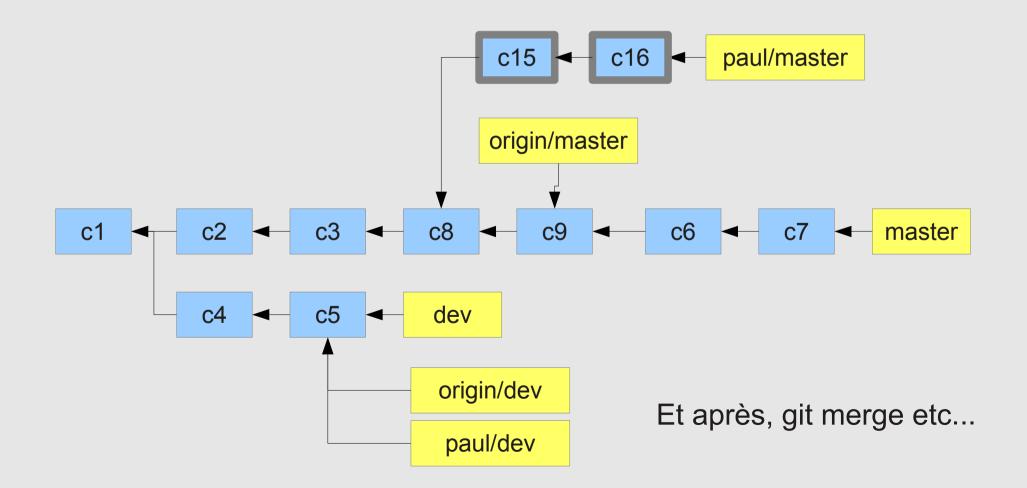




2 dépôt distants #2

Dans le dépôt local :

git remote add paul git@paul.com:projet.git git fetch paul





Workflows



Workflow de branches

- Suivant les besoins d'un projet, sa criticité, ses développeurs, l'utilisation des branches va être différentes
- Plusieurs manière de travailler sont possibles



Workflow de branches simple

- La branche master contenant le projet « stable »
- Une branche « develop » pour y effectuer le développement des fonctionnalités en cours et les stabiliser. Quand « develop » est stable → fusion dans « master »
- Une branche « topic » issue de « develop » pour les nouveaux développement. Fusion avec « develop » quand cela devient utilisable
- Pour les petits projets

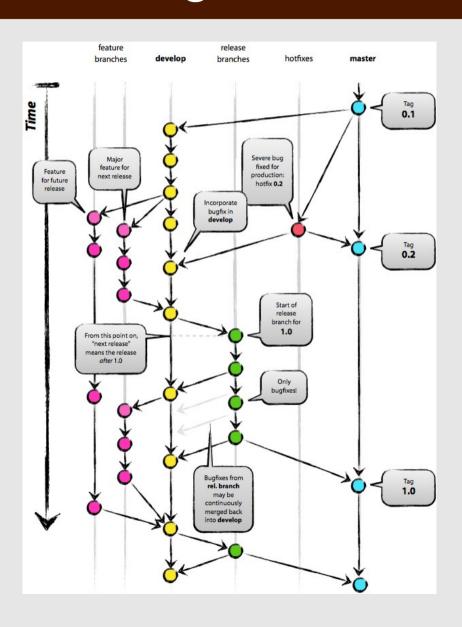


Workflow de branche #2

- Une branche par sujet de développement
- Une branche « master » dans laquelle on fusionne les dernières nouveautés
- Des branches pour chaque version stables publiées, où l'on y fait que des corrections de bugs
- D'autres workflows existent et plus complexe. Ex :
 « gitflow »
 - http://nvie.com/posts/a-successful-git-branching-model/
 - https://github.com/nvie/gitflow



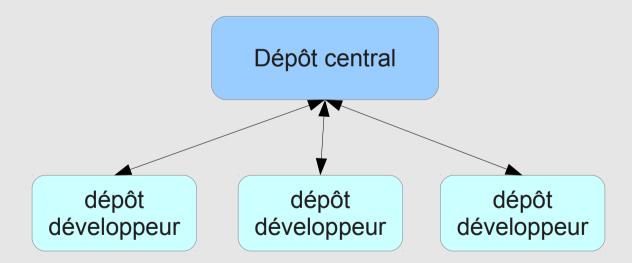
gitflow





Workflows collaboratifs

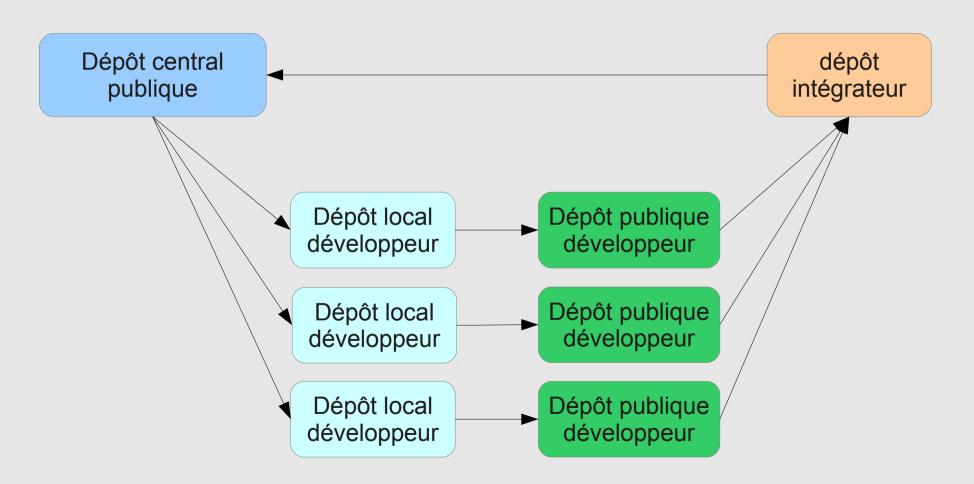
- But : définir la manière de collaborer
- En voici des exemples
- Workflow centralisé :





Workflows collaboratifs

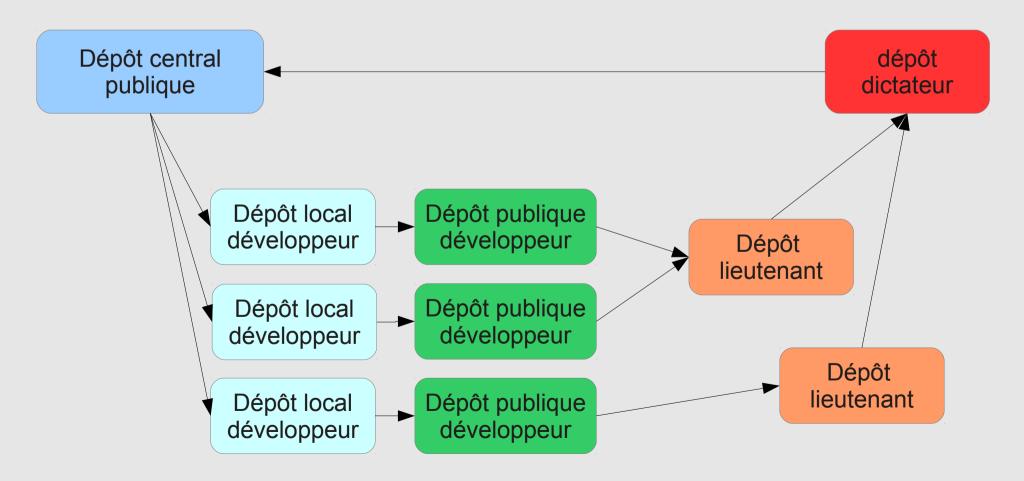
Workflow d'intégration





Workflows collaboratifs

 Workflow dictatorial : délégation d'intégration à des « lieutenants »





Github



Github.com

- Site favorisant la collaboration du développement d'un projet (privé ou publique)
- Interface utilisateur facilitant
 - Le clonage des dépôts (« fork »)
 - La fusion des contributions
 - La revue de code
- Utilisé par de nombreux projets libres
- Équivalent « libre » : gitorious.org
- Équivalent pour Mercurial : bitbucket.org

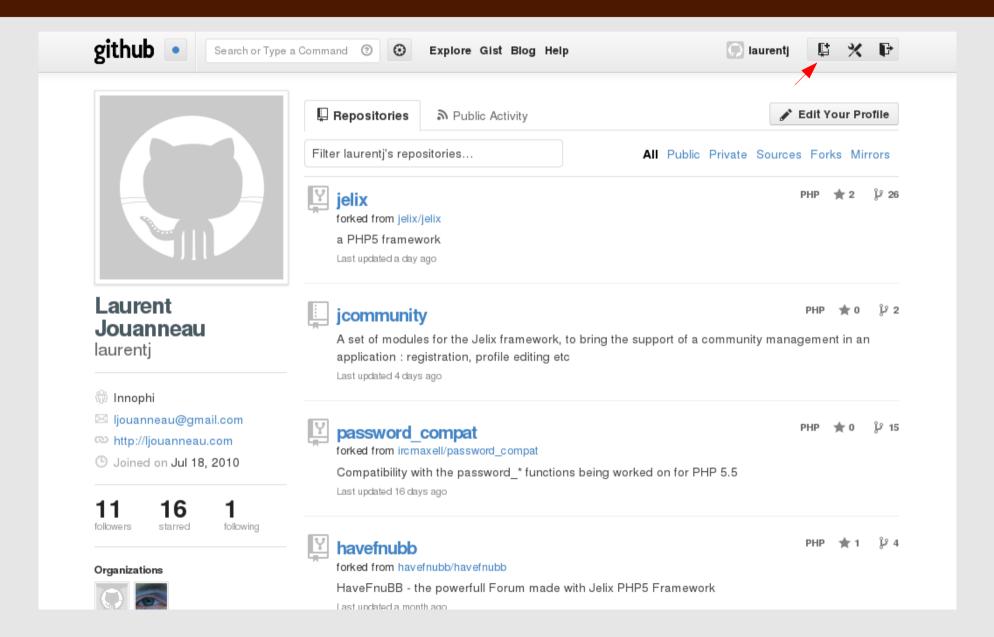


Compte

- Création d'un compte
 - Indiquez une clé ssh : facilite l'authentification

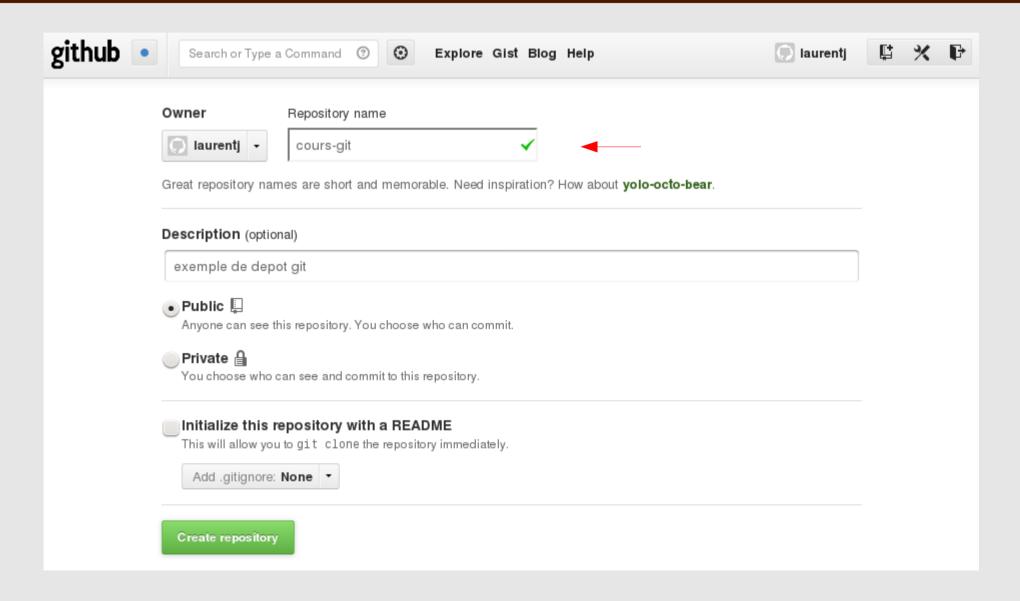


Compte & Création dépôt



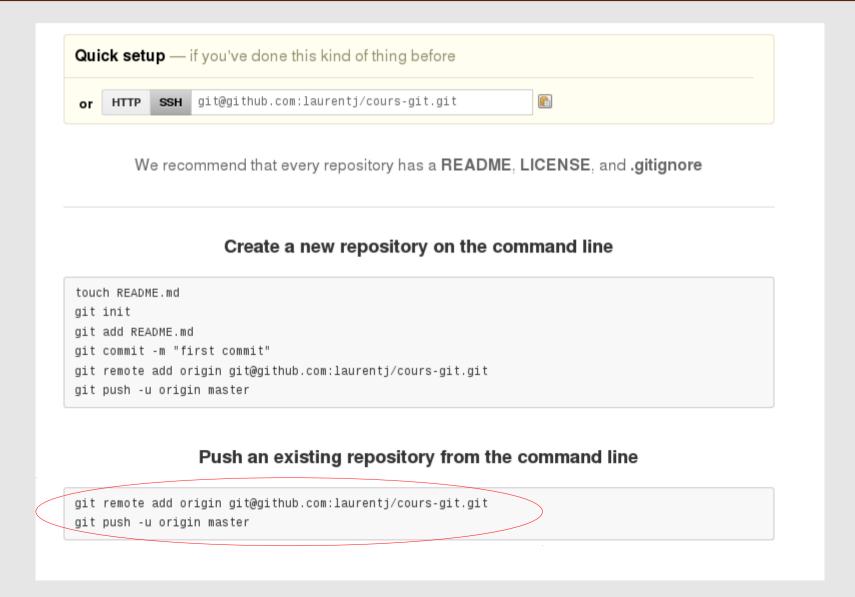


Création dépôt





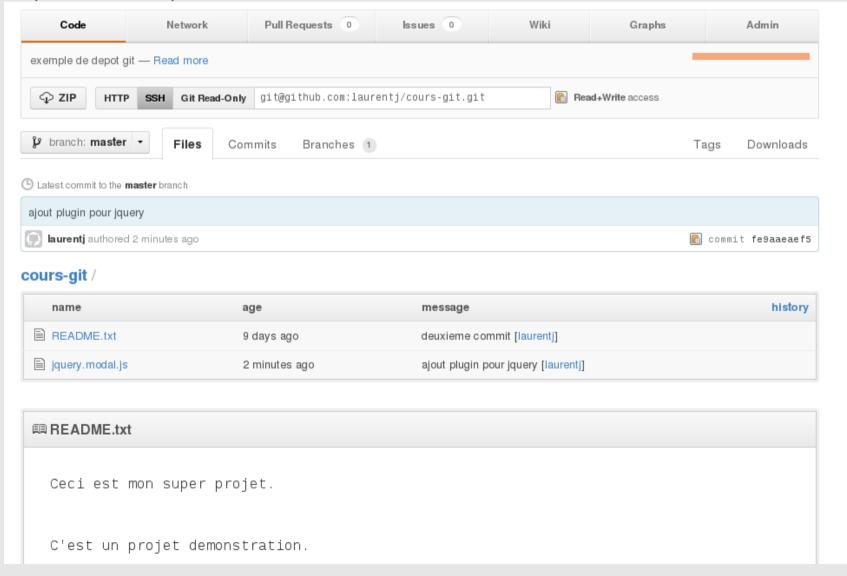
Création dépôt





Navigateur de code

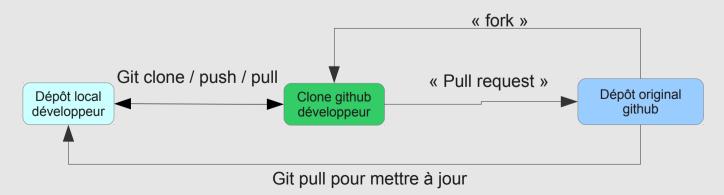
Aprés un push vers le dépôt nouvellement créé...





Contribuer

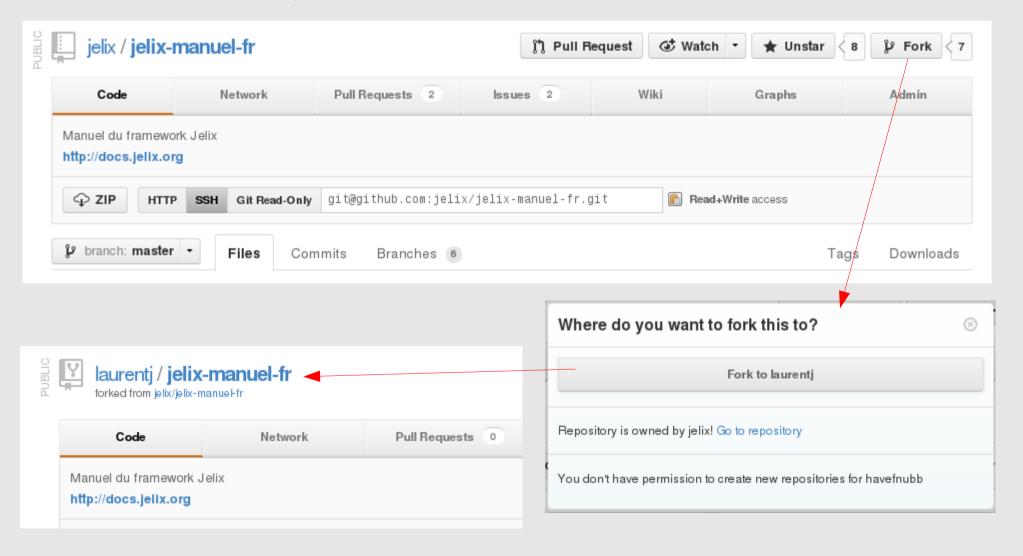
- La première chose : « forker » le dépôt cible
 - → on a un clone dans son compte
- On clone ce clone sur sa machine en local
- On commit
- On « push » vers son clone github
- On fait un « pull request » vers le dépôt original sur github





Fork

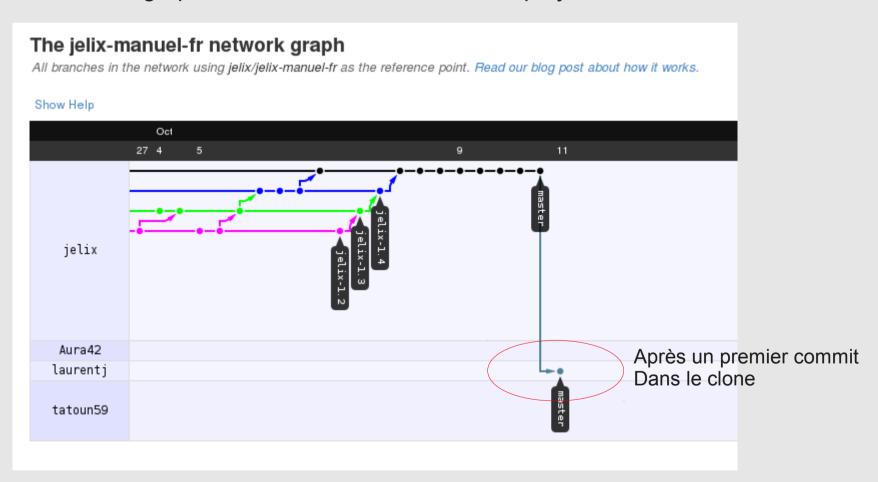
Pour contribuer à un projet : cloner le dépôt sur son compte → « fork »





Visualiser les clones

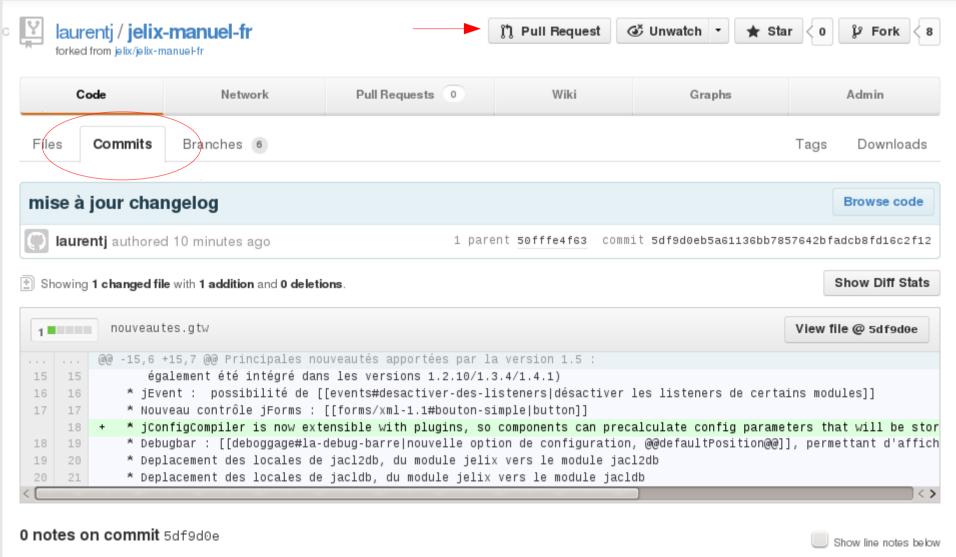
Onglet « network » : graphe des commits / branches, du projet et de ses clones





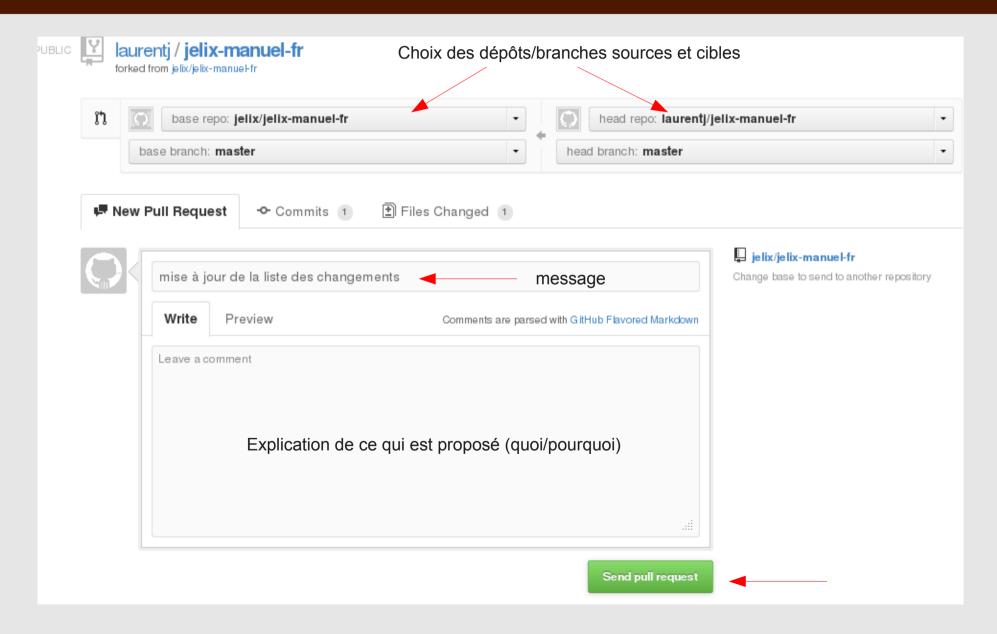
Pull request

On veut proposer une modification



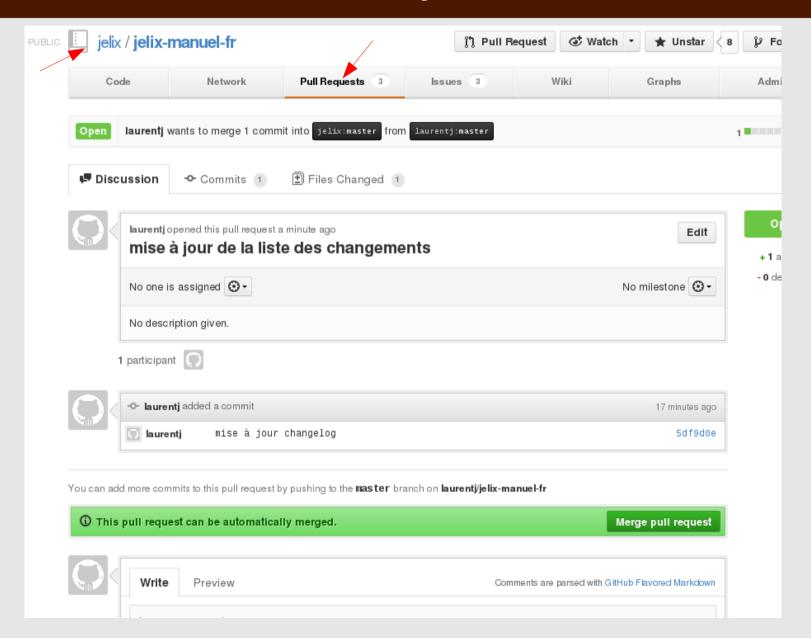


Pull request



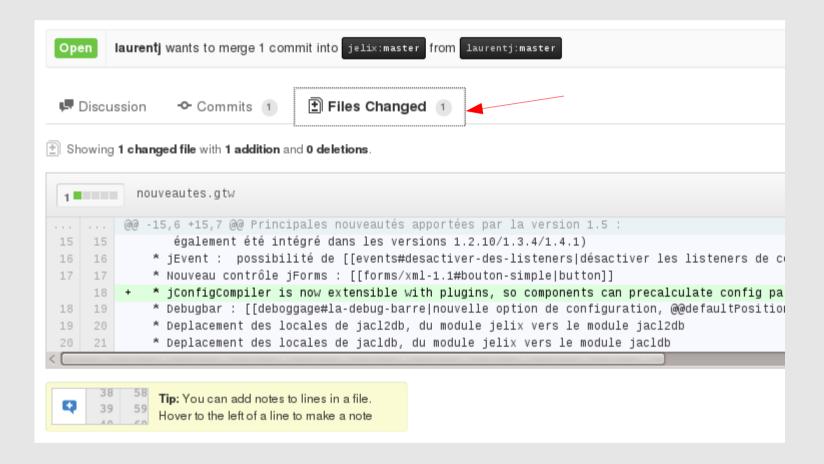


Pull Request





PR: voir modifications



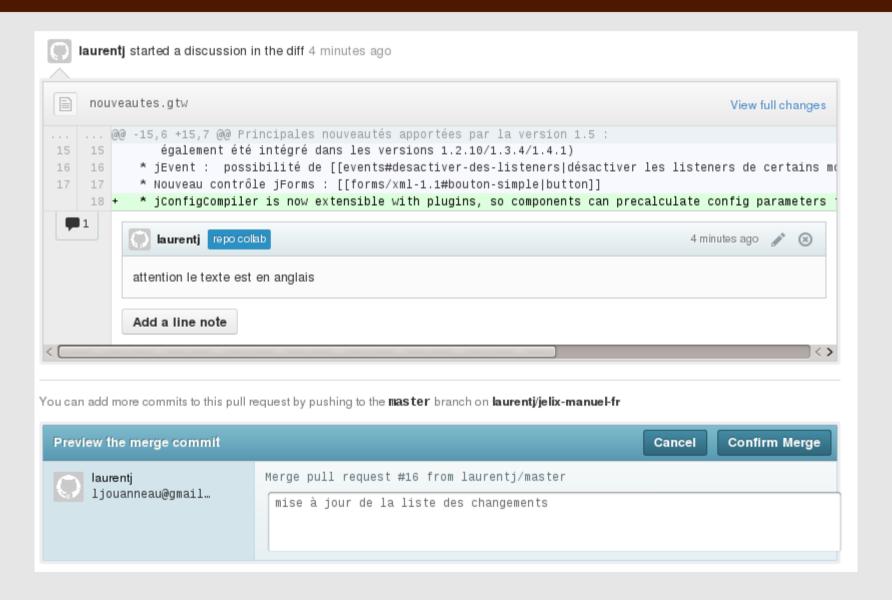


PR: Commenter des modifications



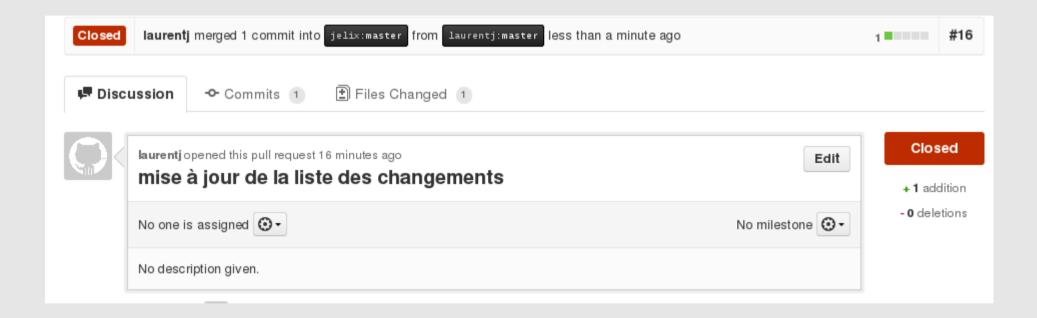


Pull Request: fusion





Pull Request : résultat





Pull request : graph

