Comprendre la gestion de versions avec GIT

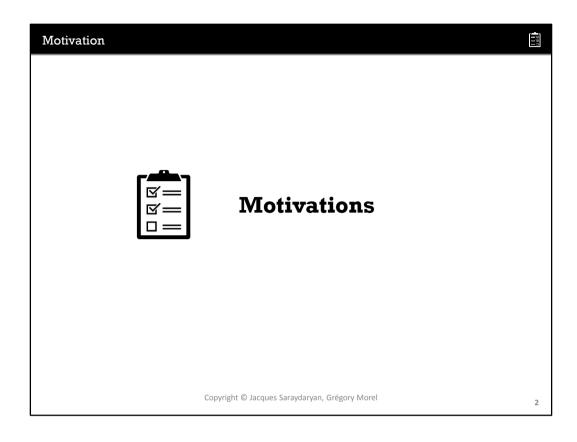
J. Saraydaryan, G. Morel

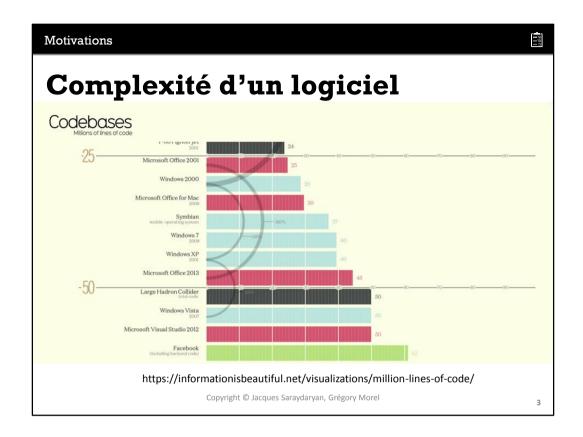






Copyright © Jacques Saraydaryan, Grégory Morel





- Les logiciels sont de plus en plus complexes, contiennent de plus en plus de lignes de codes, et sont donc de plus en plus difficiles à maintenir et faire évoluer
- On n'est plus sur un modèle où un développeur travaille seul de 0 dans son garage
 : les projets sont désormais collaboratifs, beaucoup de code est réutilisé
 (« forké ») à partir de code de projets existants
- => Il est important de pouvoir échanger, modifier, collaborer efficacement sur du code informatique

Motivations

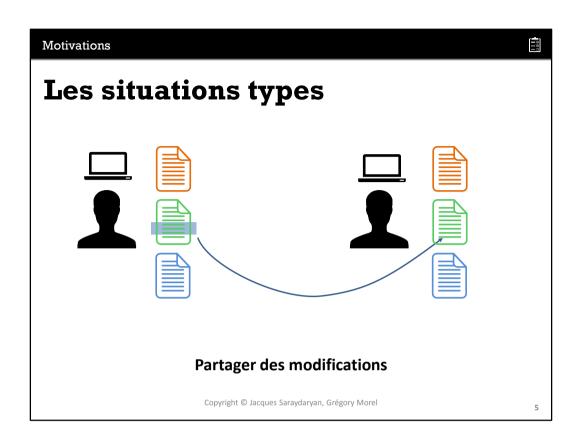


La collaboration dans le dev.

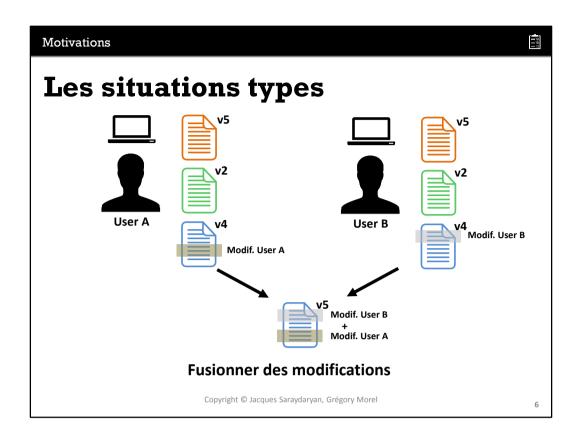
- ☐ Suivre l'évolution d'un code source
 - Revenir à une version antérieure
 - Comprendre les différentes modifications effectuées
 - Sauvegarder le travail effectué
 - Documentation du code
- ☐ Collaboration, travailler à plusieurs
 - Partager les modifications effectuées
 - Savoir qui a effectué des modifications
 - Fusionner les modifications de plusieurs développeurs.



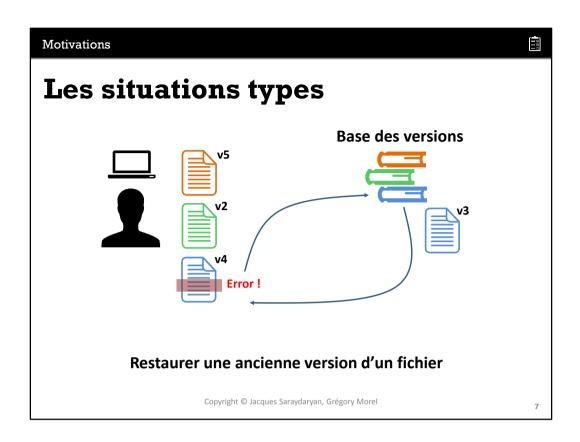
Copyright © Jacques Saraydaryan, Grégory Morel



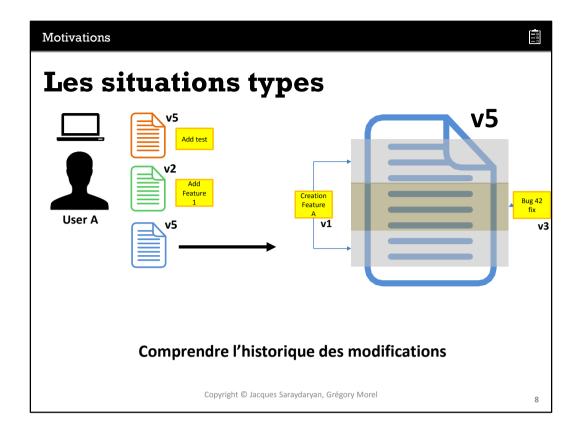
Comment partager des modifications avec nos collaborateurs ? Par mail ?!? En utilisant un répertoire partagé ?!?



Mais quid si deux personnes travaillent sur le même fichier simultanément ?



Parfois, on aimerait aussi pouvoir revenir à l'état précédent, ou une *version précédente* d'un fichier...



Il devient alors important, quand on gère plusieurs versions de nos fichiers, d'identifier rapidement ce qui change d'une version à l'autre.









Les logiciels de gestions de version

Copyright © Jacques Saraydaryan, Grégory Morel

)





Définition

■ Définition

Outil permettant de garantir le suivi de versions d'un ensemble de fichiers et fournissant des outils permettant de naviguer / commenter / fusionner les différentes versions de fichiers.

- Architecture
 - Locale :

le suivi des versions des fichiers est **uniquement** assuré sur la machine **locale**

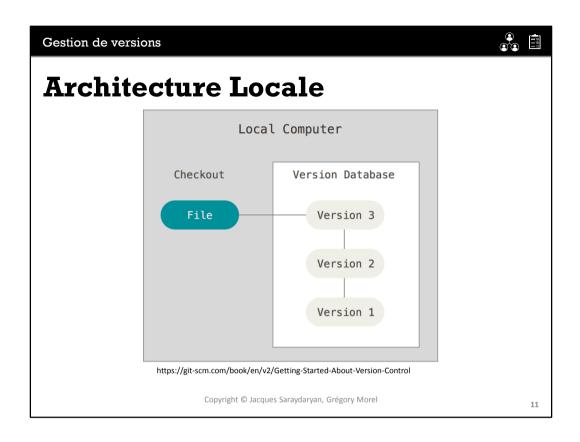
Centralisée:

le suivi et le maintien de version est garanti par un serveur central

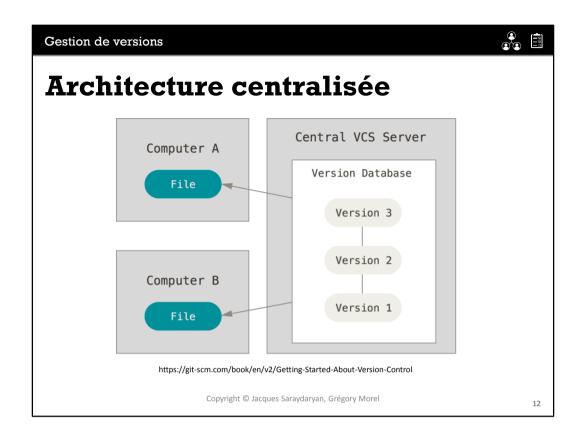
■ Décentralisée :

chaque développeur possède l'ensemble des versions des documents et informe le réseau de collaborateurs de toutes modifications

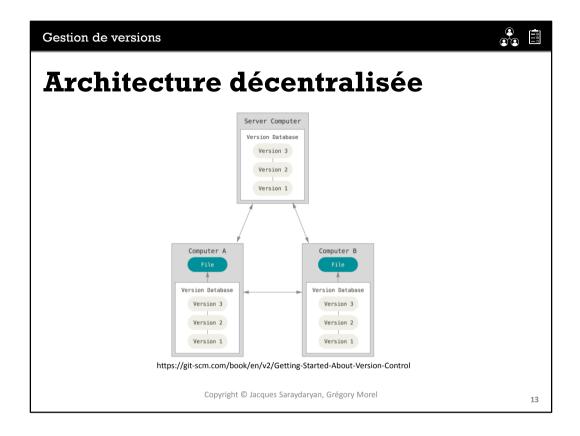
Copyright © Jacques Saraydaryan, Grégory Morel



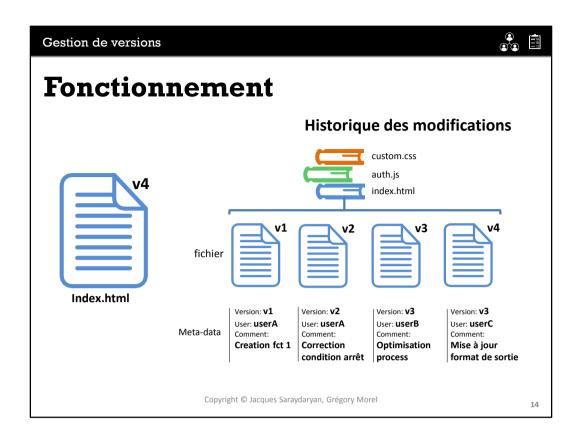
Dans une architecture locale, on dispose sur notre machine d'une base de données de toutes les versions des fichiers. Ca peut être utile quand on travaille seul, mais l'intérêt est limité si on souhaite collaborer avec d'autres développeurs...



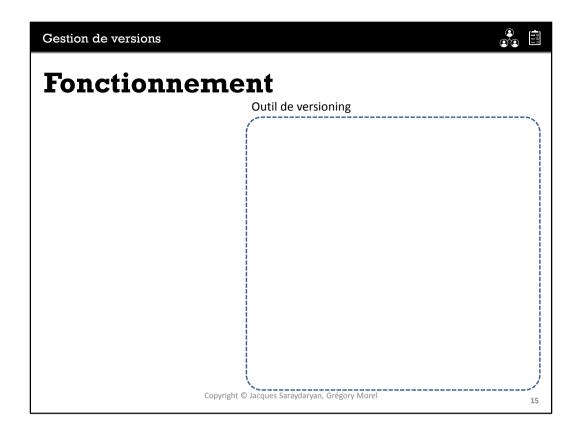
L'architecture centralisée permet de pallier ce problème. Il s'agit de l'architecture qu'on rencontre dans les anciens systèmes de gestion de versions, type SVN.

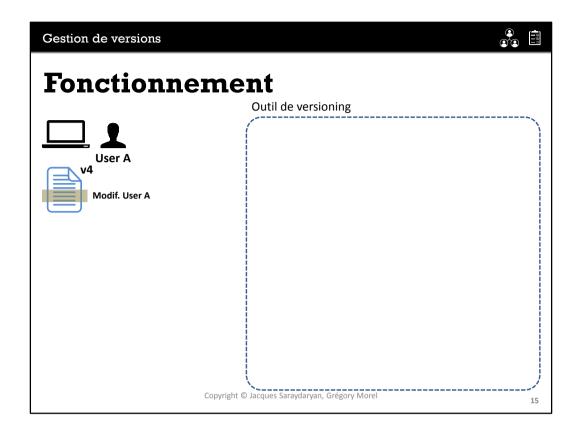


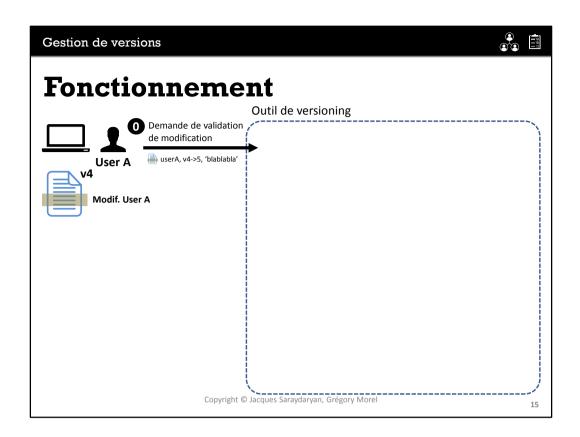
- L'inconvénient des systèmes centralisés, c'est qu'il faut un compte sur le serveur central. Avec les architectures décentralisées, il suffit de cloner (ou forker) un projet pour commencer à travailler dessus (avec tout l'historique du projet en local) et ensuite de proposer sa contribution (pull request) au dépôt principal (mainteneur principal du projet).
- Par ailleurs, la décentralisation apporte beaucoup plus de flexibilité : comme tous les contributeurs disposent de l'intégralité du projet en local, chacun devient en quelques sortes un serveur et la collaboration devient beaucoup plus simple.
- La décentralisation de Git a ainsi beaucoup apporté au développement des logiciels libres.

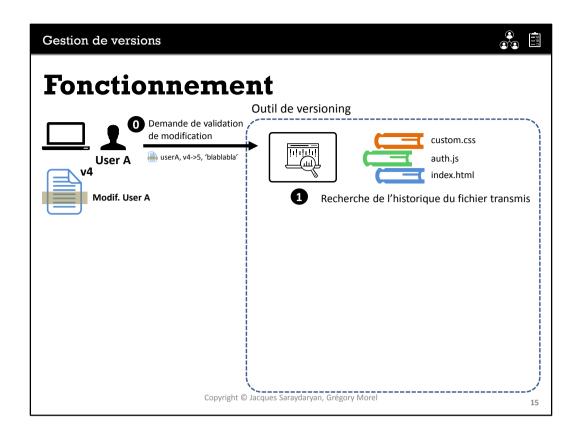


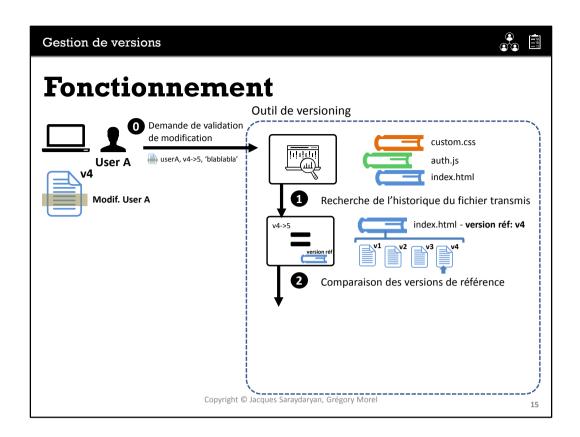
Pour chaque fichier versionné, un historique complet des versions est disponible. Et chaque version contient des *méta-données* : numéro de version, qui a créé cette version et des commentaires sur les changements apportés

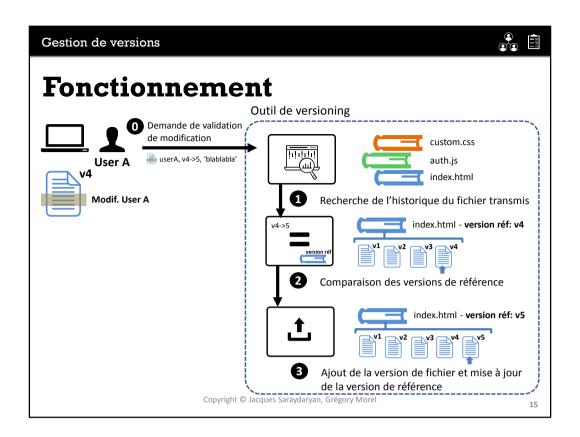
















Les systèmes de gestion de version

☐ Beaucoup d'outils disponibles ! (gratuits et payants)

Outil	Туре	Description	Projets qui l'utilisent
<u>cvs</u>	Centralisé	C'est un des plus anciens logiciels de gestion de versions. Bien qu'il fonctionne et soit encore utilisé pour certains projets, il est préférable d'utiliser SVN (souvent présenté comme son successeur) qui corrige un certain nombre de ses défauts, comme son incapacité à suivre les fichiers renommés par exemple.	OpenBSD
SVN (Subversion)	Centralisé	Probablement l'outil le plus utilisé à l'heure actuelle. Il est assez simple d'utilisation, bien qu'il nécessite comme tous les outils du même type un certain temps d'adaptation. Il a l'avantage d'être bien intégré à Windows avec le programme <u>Tortoise SVN</u> , là où beaucoup d'autres logicles s'utilisent surtout en ligne de commande dans la console. Il y a un <u>tutoriel SVN</u> sur OpenClassrooms.	Apache, Redmine, Struts
Mercurial	Distribué	Plus récent, il est complet et puissant. Il est apparu quelques jours après le début du développement de Git et est d'ailleurs comparable à ce dernier sur bien des aspects. Vous trouverez un <u>tutoriel sur Mercurial</u> sur OpenClassrooms.	Mozilla, Python, OpenOffice.org
<u>Bazaar</u>	Distribué	Un autre outil, complet et récent, comme Mercurial. Il est sponsorisé par Canonical, l'entreprise qui édite Ubuntu. Il se focalise sur la facilité d'utilisation et la flexibilité.	Ubuntu, MySQL, Inkscape
<u>Git</u>	Distribué	Très puissant et récent, il a été créé par Linus Torvalds, qui est entre autres l'homme à l'origine de Linux. Il se distingue par sa rapidité et sa gestion des branches qui permettent de développer en parallèle de nouvelles fonctionnalités.	Kernel de Linux, Debian, VLC, Android, Gnome, Qt

Source: https://openclassrooms.com/fr/courses/1233741-gerez-vos-codes-source-avec-git

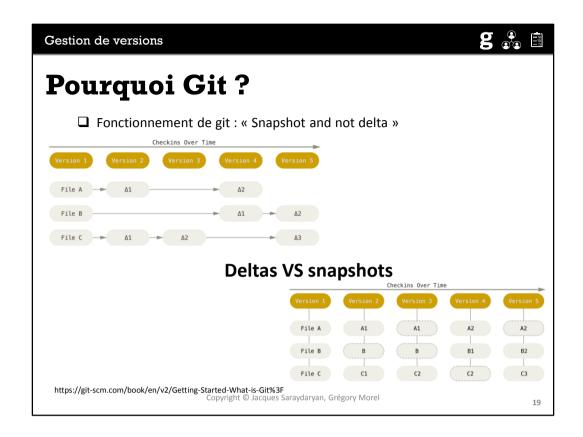
Copyright © Jacques Saraydaryan, Grégory Morel



Gestion de versions Git ☐ Créé en 2005 par Linus Torvalds (le créateur de Linux) Objectifs Gratuit Système rapide Conception simple Usage possible de développement non linéaire (branches de développement en parallèle) Pleinement distribué Support de gros projet possible (e.g. kernel linux) Propriétés « Snapshot and not delta » Les opérations de git sont principalement locales L'intégrité des composants est garantie (checksum) Principalement que des fonctions d'ajout (très difficile de modifier des éléments validés avec git) 3 états principaux pour des fichiers : modified - staged - commited Copyright © Jacques Saraydaryan, Grégory Morel 18

Quelques références :

https://git-scm.com/book/en/v2/Getting-Started-What-is-Git%3F https://medium.com/@shalithasuranga/how-does-git-work-internally-7c36dcb1f2cf https://git-scm.com/video/what-is-git



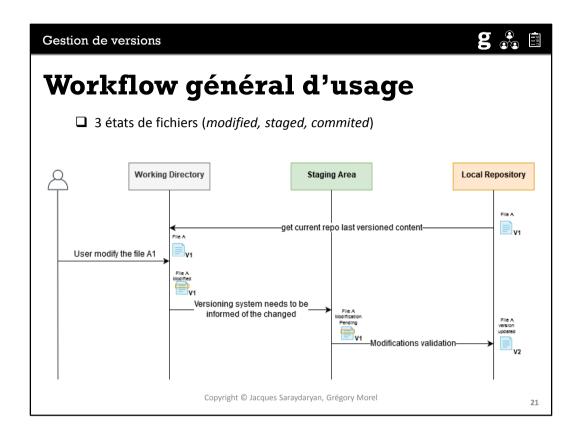
Les anciens systèmes de gestion de version (SVN...) se basaient sur des « Deltas » : d'une version d'un fichier à l'autre, seules les *différences* étaient enregistrées, ce qui permet d'économiser de l'espace disque tout en permettant de reconstituer chaque version d'un fichier.

Git adopte une approche différente : à chaque nouvelle version d'un projet, l'intégralité des fichiers modifiés est enregistrée (les fichiers non modifiés pointent simplement sur le contenu de la version précédente)

Ce point de vue est justifié par le fait que le coût de l'espace disque a considérablement chuté depuis les premiers systèmes de gestion de version ; par ailleurs, ceci simplifie le développement par « branches » qu'on verra plus loin.

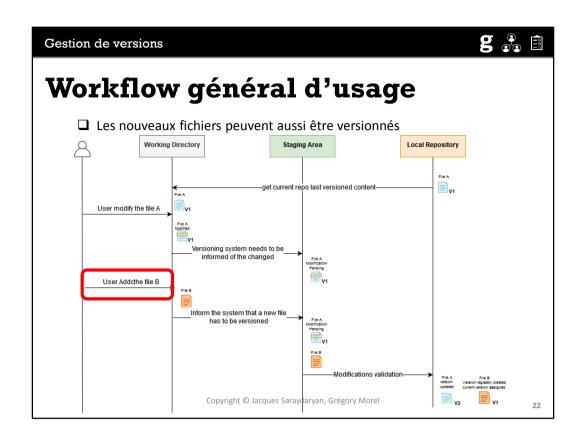
Sources:

https://medium.com/@shalithasuranga/how-does-git-work-internally-7c36dcb1f2cf https://git-scm.com/video/what-is-git

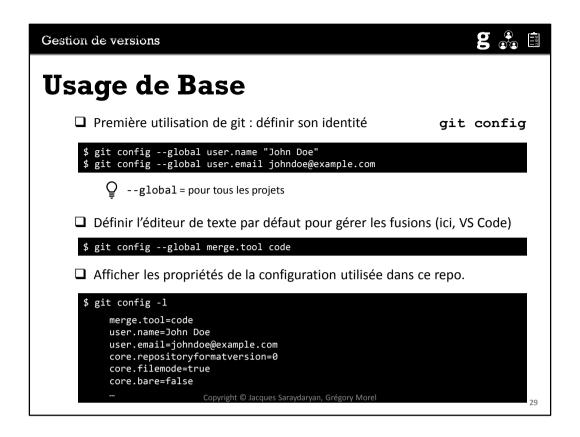


- 1. On commence par récupérer le code (checkout) depuis un dépôt (repository)
- 2. L'utilisateur modifie le fichier A
- 3. Le fichier modifié se trouve alors en état « staged » ou « attente de validation »
- 4. L'utilisateur valide ces modifications (commit), ce qui aboutit à la création d'une nouvelle version dans le système, qui devient la référence pour ce fichier

https://medium.com/@shalithasuranga/how-does-git-work-internally-7c36dcb1f2cf https://git-scm.com/video/what-is-git https://www.guora.com/What-is-Git-and-why-should-I-use-it



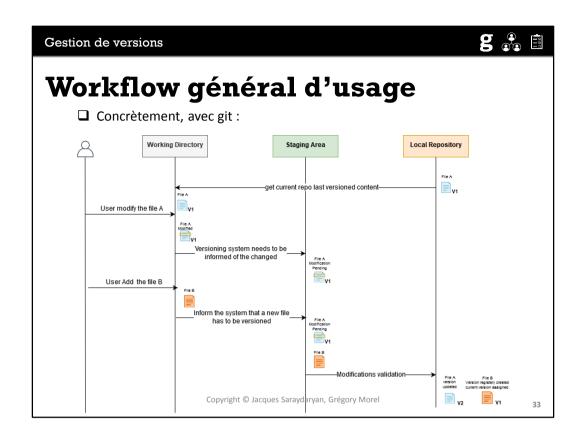
Il est bien sûr aussi possible de **créer** un nouveau fichier à versionner (ici, le fichier B)



https://git-scm.com/book/en/v1/Getting-Started-First-Time-Git-Setup



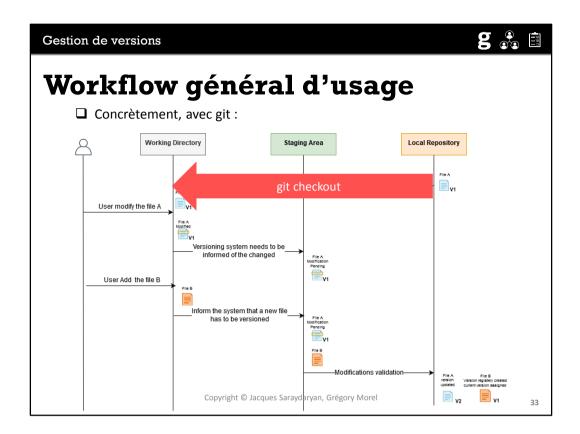


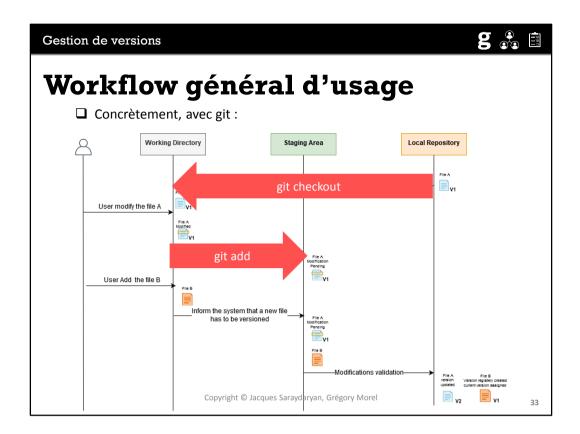


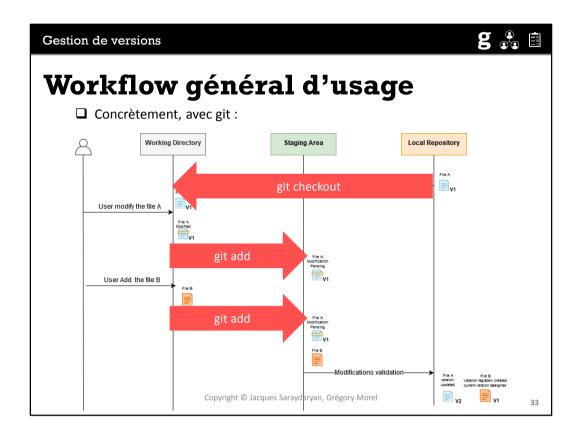
Remarque : il n'est pas indispensable de faire un git add sur un fichier déjà versionné

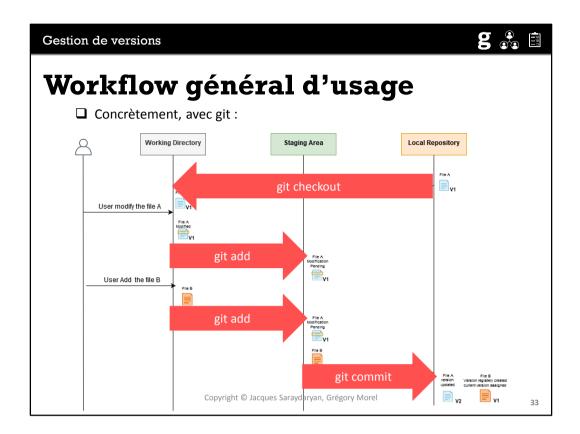
https://medium.com/@shalithasuranga/how-does-git-work-internally-7c36dcb1f2cf https://git-scm.com/video/what-is-git

https://www.quora.com/What-is-Git-and-why-should-I-use-it









Gestion de versions



Usage de Base

☐ Ajout d'un fichier

git add

 Lorsqu'on souhaite versionner un nouveau fichier ou un fichier existant, il est nécessaire d'en informer git

\$ git add myNewFile1

Plusieurs fichiers peuvent être ajoutés en même temps

\$ git add myNewFile1 myNewFile2 myNewFile3

Tous les fichiers non versionnés peuvent être ajoutés d'un coup

\$ git add --all

Une fois les fichiers ajoutés ils doivent être validés (commités)

\$ git commit myNewFile1 myNewFile2 myNewFile3 -m " validation of new files"

Copyright © Jacques Saraydaryan, Grégory Morel

Gestion de versions



Usage de Base

☐ Validation d'un fichier

git commit

 La modification d'un fichier existant nécessite une validation pour être versionnée

\$ git commit file1 -m "my first file modification"

La validation de plusieurs fichiers est également possibles

\$ git commit file1 file2 file3 -m " multi files commit operation"

Ainsi que la validation de l'ensemble des fichiers modifiés

\$ git commit -a -m "commit operation of all files"

Copyright © Jacques Saraydaryan, Grégory Morel



Usage de Base

☐ Comparaison version courante/version validée

git diff

 Permet de visualiser les différences entre le fichier courant non validé et la version du fichier validée

\$ git diff file1

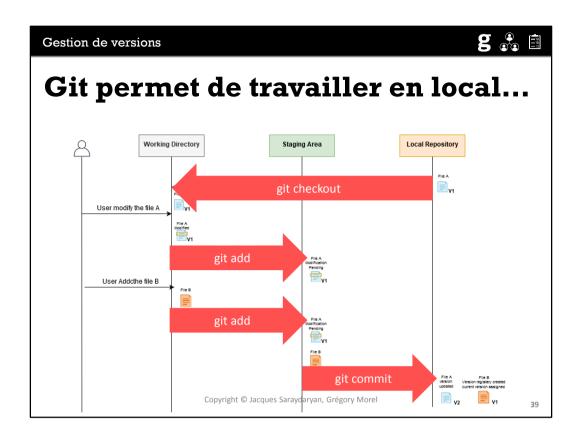
Gestion de versions

■ Exemple :

```
$ git diff Storage.java
index 781c9d1..4ff95cc 100644
--- a/src/com/ci/myShop/controller/Storage.java
+++ b/src/com/ci/myShop/controller/Storage.java
@@ -2,19 +2,21 @@ package com.ci.myShop.controller;
import com.ci.myShop.model.Item;

import java.util.HashMap;
-import java.util.Map;

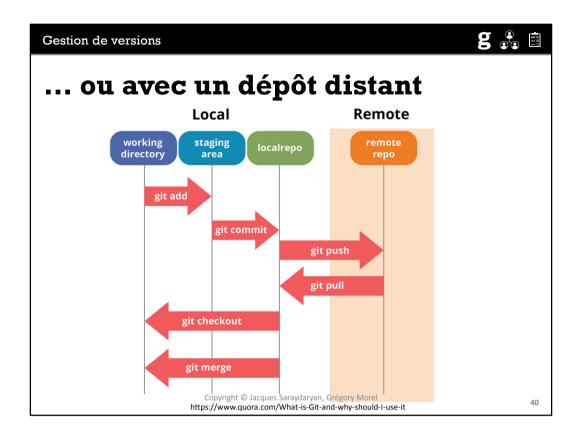
+/**
+Class managing ItemList
+*/
public class Storage {
+//Map of item
    private Map<Integer, Item> itemList;
```



Jusqu'à présent, le workflow était celui-ci, avec un développement essentiellement en *local*.

https://medium.com/@shalithasuranga/how-does-git-work-internally-7c36dcb1f2cf https://git-scm.com/video/what-is-git

https://www.quora.com/What-is-Git-and-why-should-I-use-it



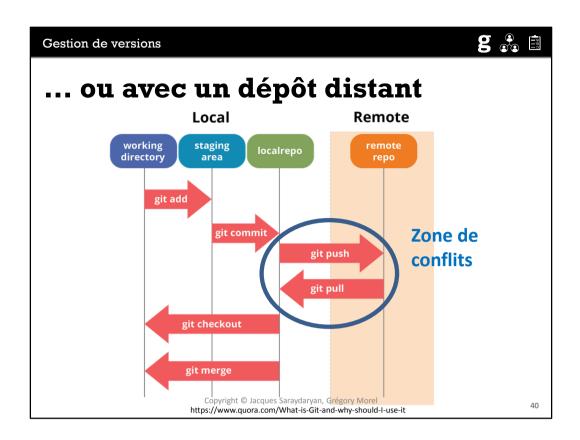
Mais l'intérêt des outils de gestion de version est surtout de pouvoir travailler à plusieurs, sur un même dépôt **distant**.

L'envoi des fichiers vers le dépôt distant se fait par la commande **git push**, et la récupération des dernières versions présentes sur le dépôt distant par la commande **git pull**.

Evidemment, c'est lors de ces opérations **push** et **pull** que des **conflits** pourront apparaître, si deux utilisateurs ont modifié un fichier et essaient de le pousser vers le dépôt distant, ou si je récupère un fichier qui a été modifié par quelqu'un d'autre depuis que j'ai moi-même travaillé sur ce même fichier en local.

https://medium.com/@shalithasuranga/how-does-git-work-internally-7c36dcb1f2cf https://git-scm.com/video/what-is-git

https://www.quora.com/What-is-Git-and-why-should-I-use-it



g 🚱 🗎

Usage de Base

Gestion de versions

☐ Récupération des données distantes

git pull

```
$ git pull
    remote: Enumerating objects: 4, done.
    remote: Counting objects: 100% (4/4), done.
    remote: Compressing objects: 100% (2/2), done.
    remote: Total 3 (delta 0), reused 0 (delta 0)
    Unpacking objects: 100% (3/3), done.
    From https://gitlab.com/jsaraydaryan/test-ci
        cc2e658.afc40e7 master -> origin/master
    Updating cc2e658.afc40e7
    Fast-forward
    fileA | 1 +
        1 file changed, 1 insertion(+)
        create mode 100644 fileA
```

☐ Envoi des données modifiées localement sur le serveur distant git push

```
$ git push
Enumerating objects: 4, done.
Counting objects: 100% (4/4), done.
Compressing objects: 100% (2/2), done.
Writing objects: 100% (3/3), 295 bytes | 98.00 KiB/s, done.
Total 3 (delta 0), reused 0 (delta 0)
To https://gitlab.com/jsaraydaryan/test-ci_git
afc40e7..1a58ccf master -> master
```



On peut **modifier** le **dernier** commit (par exemple si on veut modifier le message du commit, si on a oublié un fichier, etc.)



Il est également possible d'annuler toutes les modifications apportées en local sur un fichier, et de revenir à la version du dépôt local.

(En gros, on repasse de l'état staged à l'état commited)

Gestion de versions g	
Usage de Base	
 Annulation des modifications (locales / commit privé) git reset HE Annulation du dernier commit (soft) 	E A D
\$ git reset HEAD	i
Copyright © Jacques Saraydaryan, Grégory Morel	45

Il est même possible de revenir à un état avant un commit donné, avec git reset <commit>.

Précision : les commits sont annulés **sur le dépôt** ; mais dans le répertoire de travail, on conserve les modifications apportées sur les fichiers. Si on veut vraiment revenir à l'état initial, il faut forcer **git reset** avec l'option --hard.

Rem. : git reset employé sans n° de commit réinitialise la zone de transit (*staging area*).

Gestion de versions



Usage de Base

☐ Annulation des modifications (locales / commit privé) git reset HEAD

Annulation du dernier commit (soft)

\$ git reset HEAD

Annulation d'autres commit

Head : dernier commit

Head : avant dernier commit

Head^^ : avant avant dernier commit Head^2 : avant avant dernier commit (autre notation)

D6d98923868578a7f38dea79833b56d0326fcba1 : numéro commit (précis)

D6d9892 : numéro commit (notation courte)

Copyright © Jacques Saraydaryan, Grégory Morel



Usage de Base

☐ Annulation des modifications (locales / commit privé) git reset HEAD

Annulation du dernier commit (soft)

\$ git reset HEAD

Annulation d'autres commit

Head : dernier commit

Head^: avant dernier commit

Head^^ : avant avant dernier commit

Head~2: avant avant dernier commit (autre notation)

D6d98923868578a7f38dea79833b56d0326fcba1: numéro commit (précis)

D6d9892 : numéro commit (notation courte)

Seuls les commits sont annulés, vos fichiers restent les mêmes

Annulation du dernier commit (hard)

\$ git reset --hard HEAD /!\ Annule les commits et perd tous les changements

Copyright © Jacques Saraydaryan, Grégory Morel



Usage de Base

☐ Annulation des modifications (distantes / commit public) git revert

Annulation d'un commit publié sur le dépôt distant

\$ git revert 5478cc1

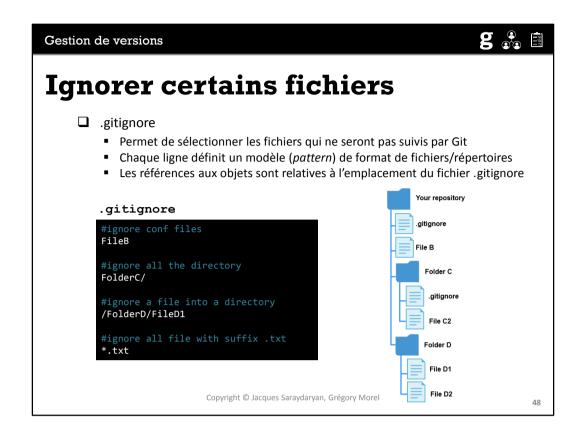
Attention! Remettre votre dossier de travail dans l'état du commit 5478cc1

En réalité, crée une **nouvelle version** qui efface les modifications effectuées depuis la version spécifiée

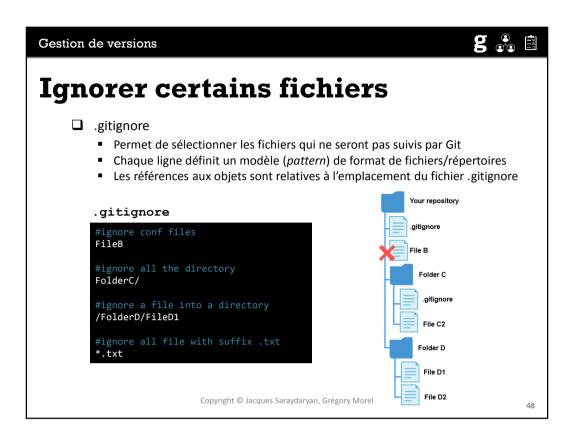


Les pushs sont définitifs, les modifications publiques sont enregistrées et ne peuvent pas être supprimées

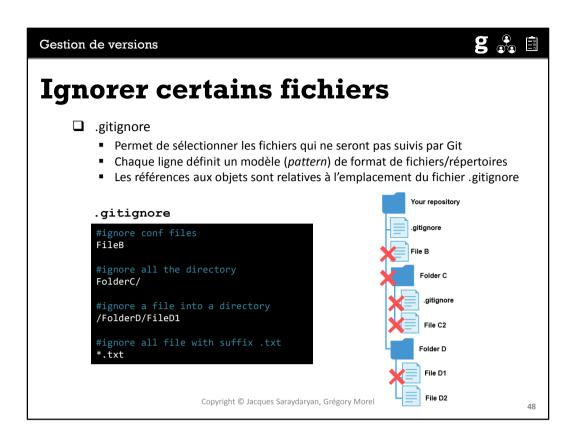
Copyright © Jacques Saraydaryan, Grégory Morel

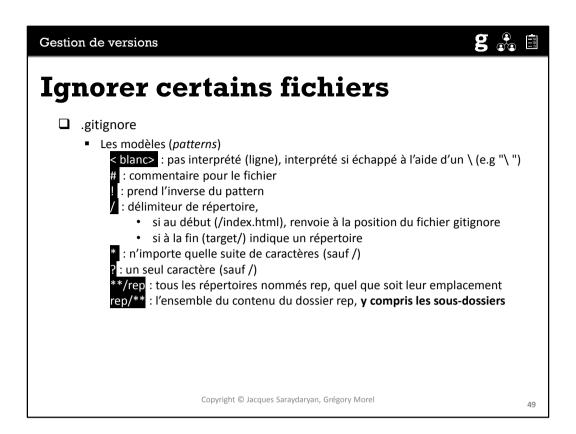


https://git-scm.com/docs/gitignore









https://git-scm.com/docs/gitignore





Résolution des conflits

☐ Conflit

Un conflit survient lorsque 2 utilisateurs ont modifié la même zone d'un fichier et que le système de versioning n'est pas capable de fusionner ces éléments. Il n'est alors pas possible d'appliquer les modifications sur le serveur.

Copyright © Jacques Saraydaryan, Grégory Morel

Gestion de versions



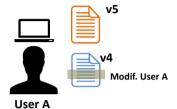




Résolution des conflits

☐ Conflit

Un conflit survient lorsque 2 utilisateurs ont modifié la même zone d'un fichier et que le système de versioning n'est pas capable de fusionner ces éléments. Il n'est alors pas possible d'appliquer les modifications sur le serveur.



Copyright © Jacques Saraydaryan, Grégory Morel



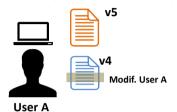




Résolution des conflits

☐ Conflit

Un conflit survient lorsque 2 utilisateurs ont modifié la même zone d'un fichier et que le système de versioning n'est pas capable de fusionner ces éléments. Il n'est alors pas possible d'appliquer les modifications sur le serveur.





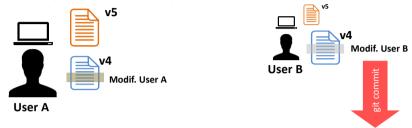
Copyright © Jacques Saraydaryan, Grégory Morel

Gestion de versions

Résolution des conflits

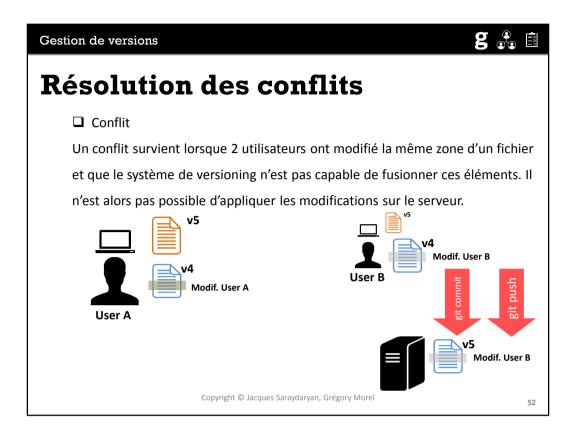
☐ Conflit

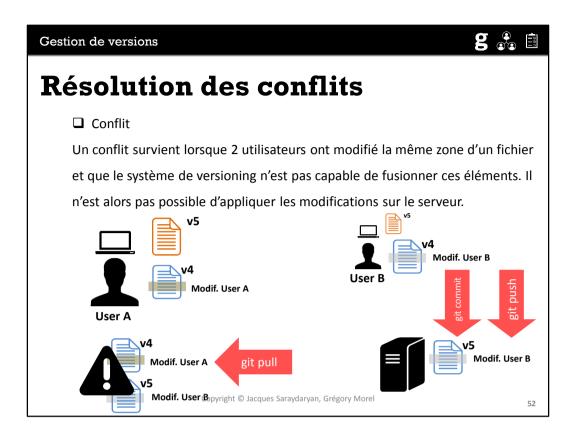
Un conflit survient lorsque 2 utilisateurs ont modifié la même zone d'un fichier et que le système de versioning n'est pas capable de fusionner ces éléments. Il n'est alors pas possible d'appliquer les modifications sur le serveur.



Copyright © Jacques Saraydaryan, Grégory Morel

74





Gestion de versions





Résolution des conflits

☐ Résolution du conflit



Exemple



User A

Hello the is the fileToMerge I am A i modified the file here This is the text of end of the file

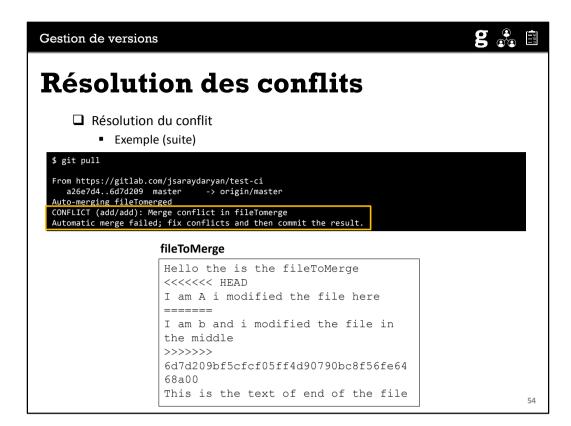
Hello the is the fileToMerge I am b and i modified the file in the middle

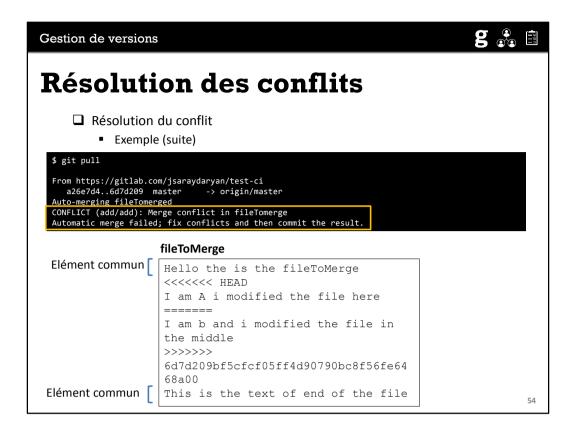
This is the text of end of the file

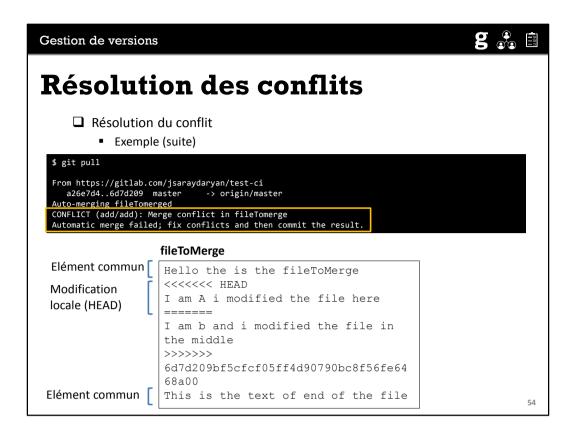
\$ git commit fileToMerge "B modified file" \$ git push

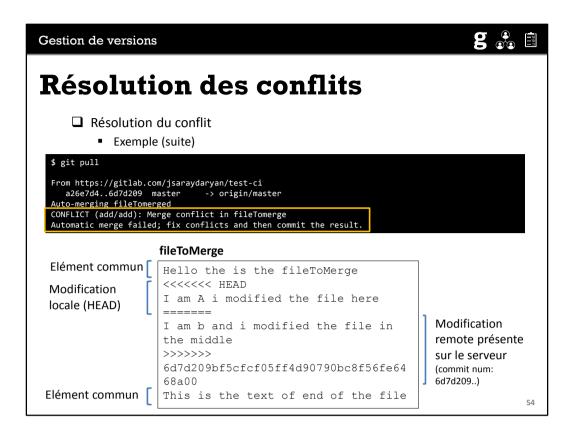
\$ git commit fileToMerge "A modified file"
\$ git push

hint: Updates were rejected because the remote contains work that you do hint: not have locally. This is usually caused by another repository pushing hint: to the same ref. You may want to first integrate the remote changes hint: to the same integrate the remote changes hint: (e.g., 'git pull ...') before pushing again. Copyngnt⊕Jacques Saraydaryan, Grégory Morel













Résolution des conflits

- ☐ Résolution du conflit
 - Exemple (suite)

fileToMerge

Hello the is the fileToMerge

I am A i modified the file here $\overline{}$

I am b and i modified the file in the middle

6d7d209bf5cfcf05ff4d90790bc8f56fe6468a00

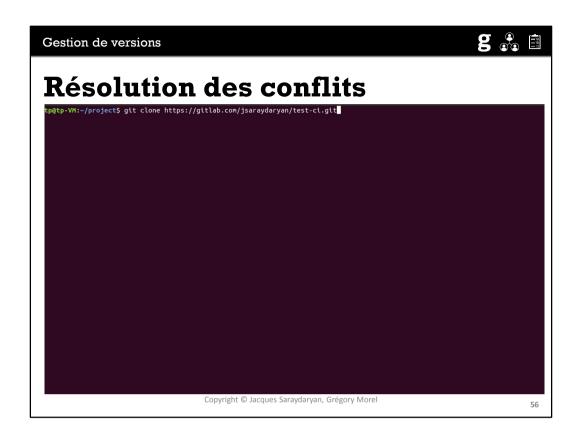
This is the text of end of the file

fileToMerge

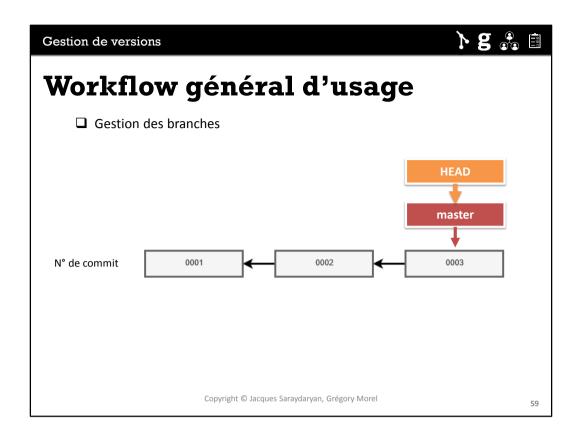
Hello the is the fileToMerge
I am A i modified the file here
This is the text of end of the file

\$ git add fileToMerge \$ git commit -m "merge the current file, with my modification" [master 28bb9d3] merge the current file, with my modification \$ git push

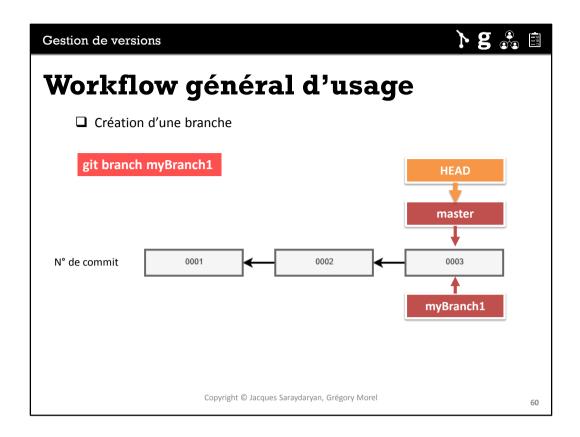
Copyright © Jacques Saraydaryan, Grégory Morel





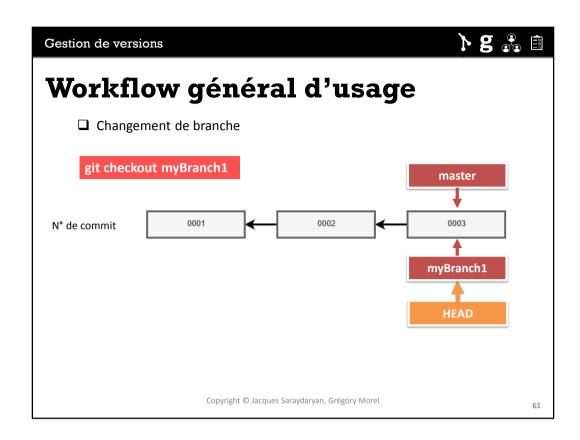


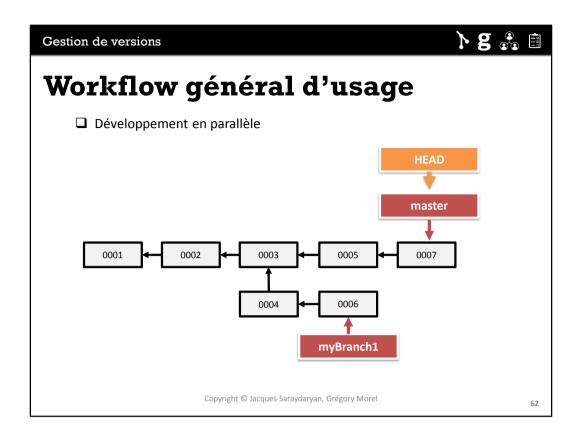
L'enchaînement de différents commits constitue une **branche**. Il existe toujours au moins une branche, créée lors de l'initialisation du dépôt, appelée **master**. Comme on le verra par la suite, il peut exister plusieurs branches ; la branche sur laquelle on est en train de travailler est identifié par un « pointeur » appelé **HEAD**



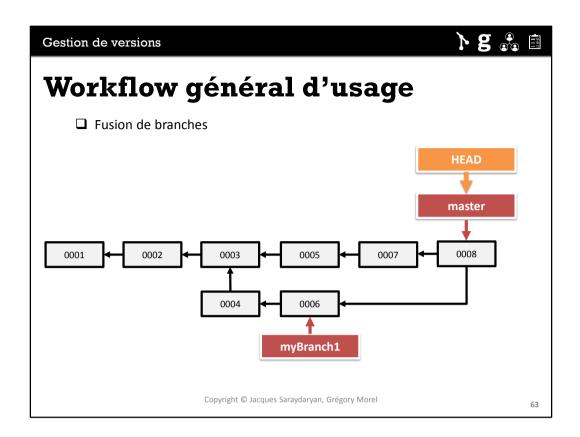
On peut créer une nouvelle branche pour démarrer un développement **parallèle** dans une application (par exemple, pour développer une fonctionnalité expérimentale).

Pour cela, on commence par créer une branche avec la commande **git branch**, qui part de la dernière version de la branche initiale. Pour l'instant les deux branches sont donc identiques.





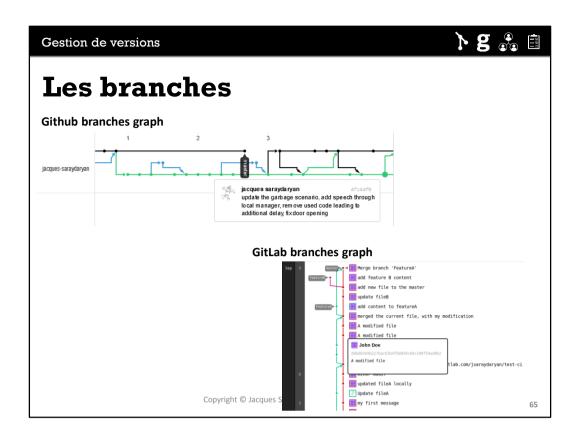
Les développements sur les deux branches peuvent alors avoir lieu en parallèle.



Les deux branches, ayant subi des développement en parallèle, **diffèrent**. On peut choisir de reporter les modifications de l'une dans l'autre, par exemple les modification faites dans *myBranch1* sur *master*.



https://git-scm.com/book/en/v1/Git-Branching-Basic-Branching-and-Merging







Questions?

Copyright © Jacques Saraydaryan, Grégory Morel

О