Source Control for People Who Don't Like Source Control

Jim Weirich Chief Scientist / EdgeCase jim@edgecase.com @jimweirich



Source Control for People

Jim Weirich Chief Scientist / EdgeCase jim@edgecase.com @jimweirich



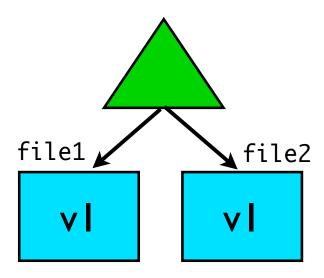


Your Mission

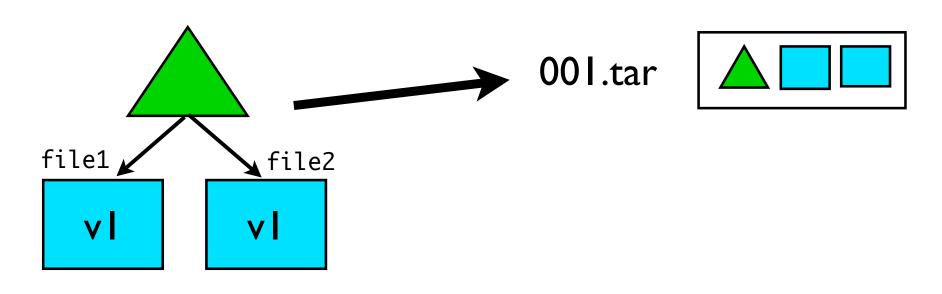
Design and Build A Source Code Control System Custom
Source
Control

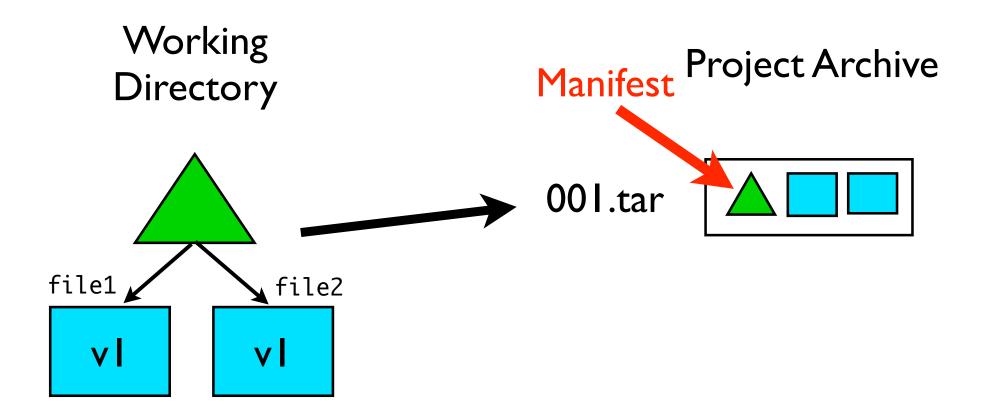
Custom
Source
Control

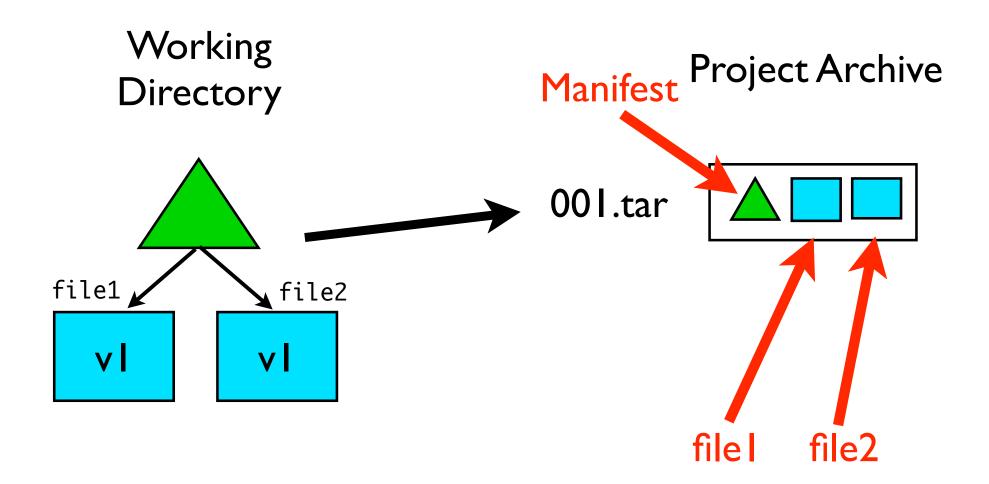
In the Old Days...



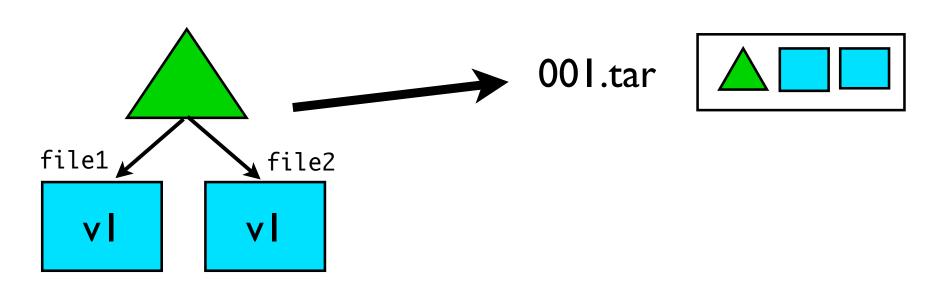
Project Archive



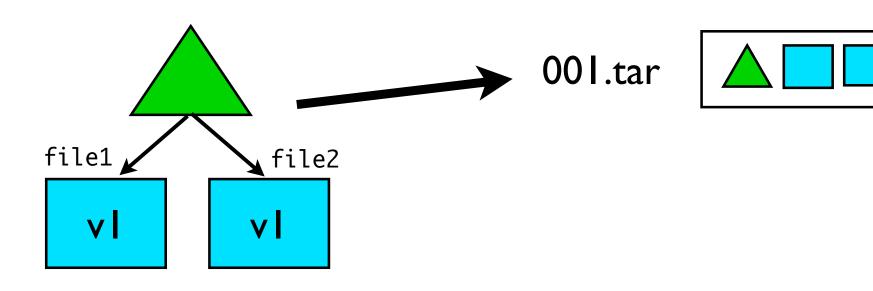




Project Archive



Project Archive



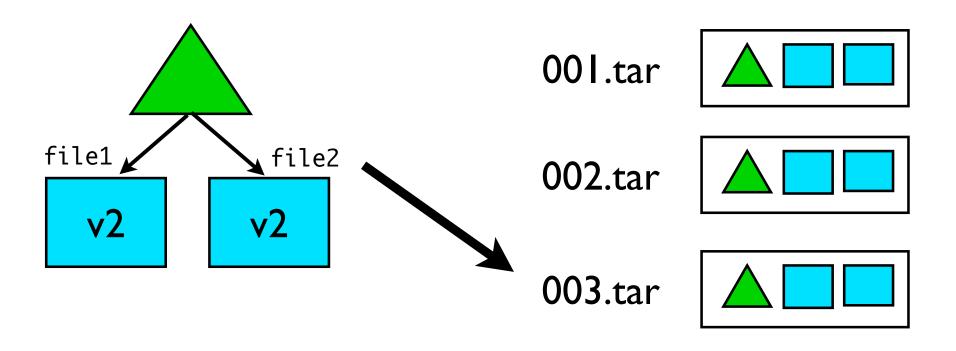
\$ csc snap

Project Archive



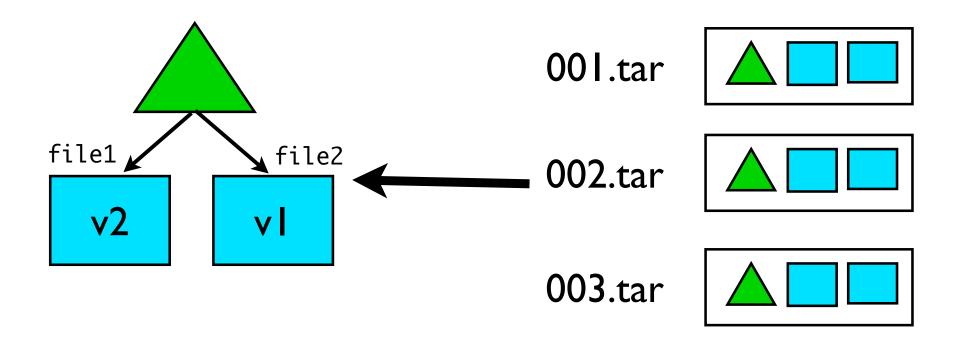
\$ csc snap

Project Archive



\$ csc snap

Project Archive



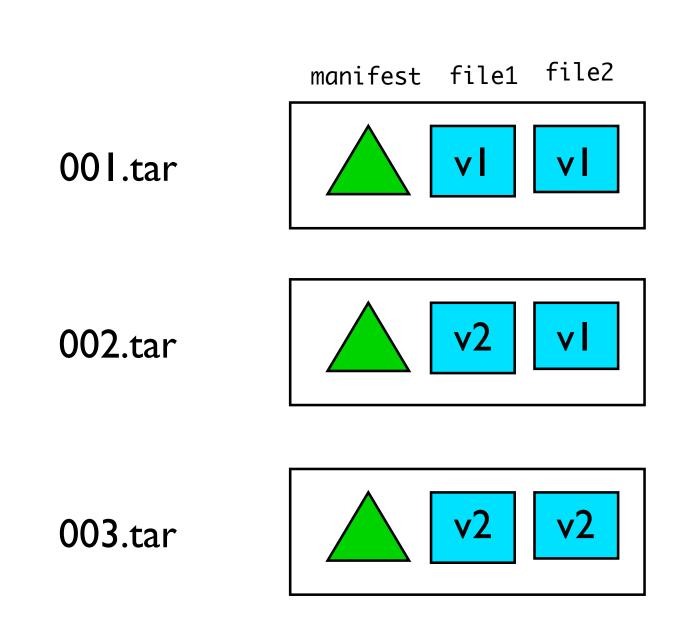
\$ csc checkout 002

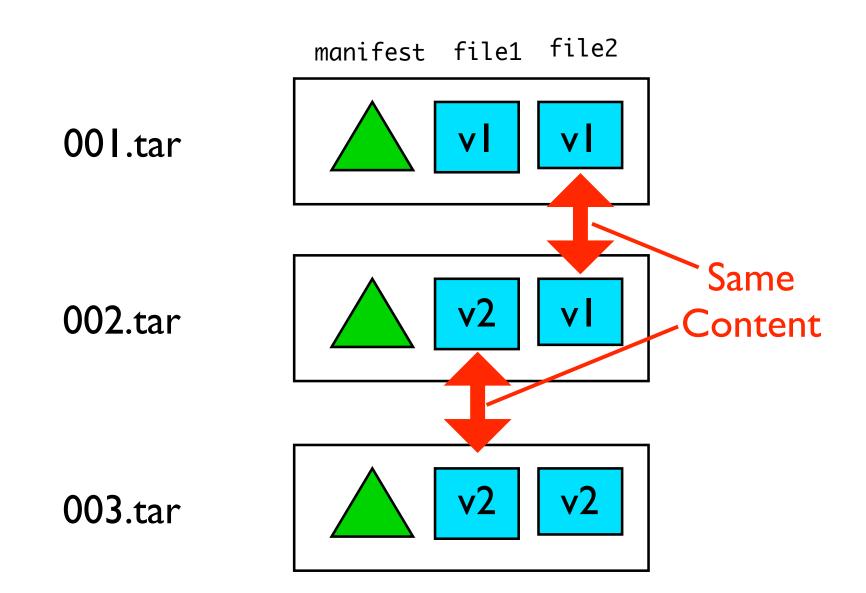
Two Basic Commands

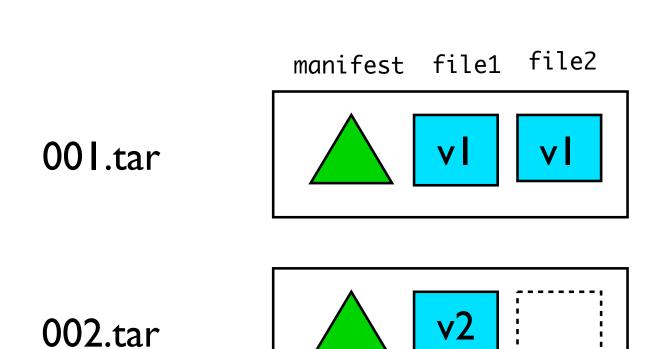
- Snap
- Checkout

Very Easy/ Very Simple

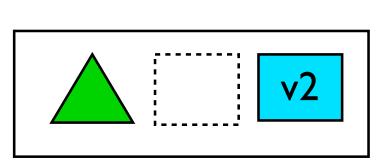
Rather Wasteful



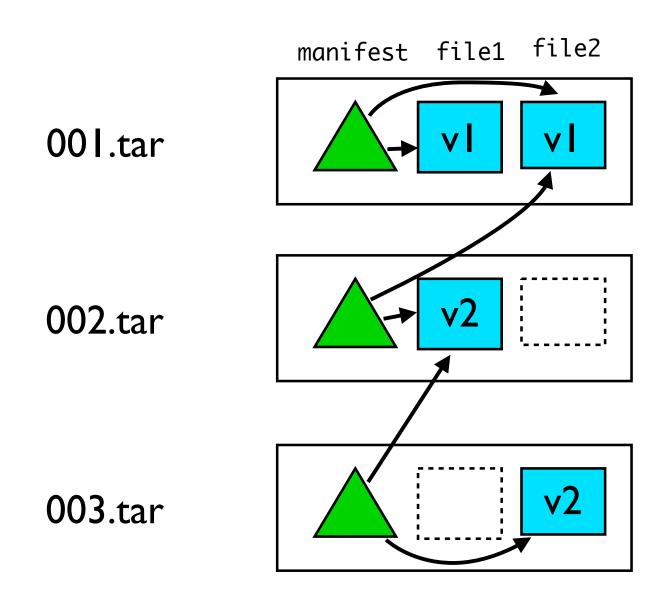


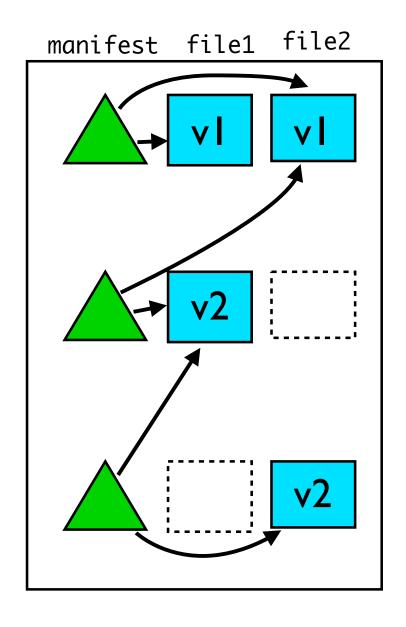


003.tar



20

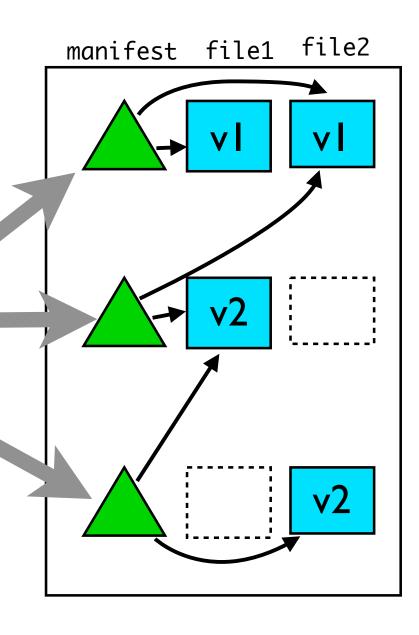




22

Note:

Each snapshot is totally defined by the manifest file for that snapshot.



File Name Contents

```
manifest △ / (001)
file1 ☐ file1 (V1)
file2 ☐ file2 (V1)
manifest △ /
manifest △ /
file1 ☐ file1 (V2)
file2 ☐ file2 (V2)
```

File Name Contents manifest \triangle / (001) file1 | | file1 (V1) file2 | file2 (V1) manifest \triangle / manifest \triangle / file1 file1 (V2) file2 (V2) file2 **CONFLICT!**

File Name Contents

File Name Contents

File Name Contents

Note:

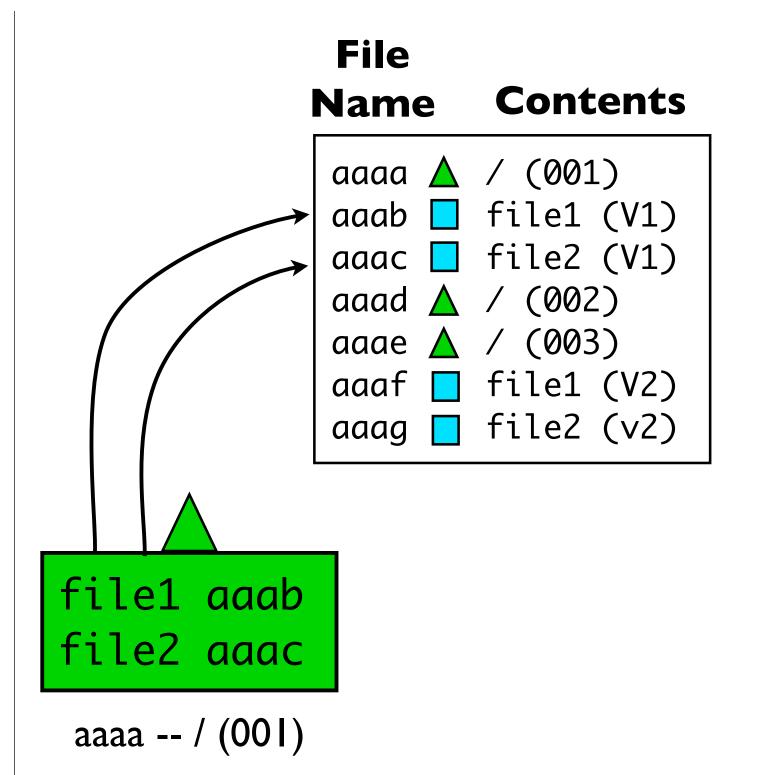
The Project Archive is just a directory with files in it.


```
$ 1s
aaaa aaab aaac aaad aaae aaaf aaag
```

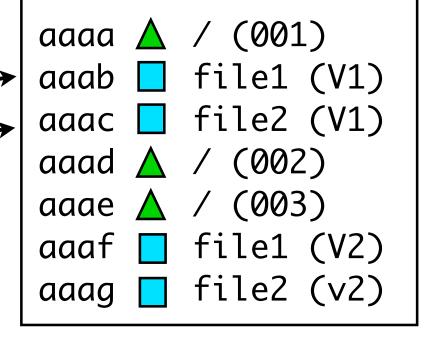
Project Archive

- Is just a directory with files in it
- The archive files have **arbitrary** names.
 - The file names can be anything as long as they are **unique**
- We track the **real file** names in the manifest files.

File Name Contents



File Name Contents



file1 aaab file2 aaac

aaaa -- / (00 I)

Note:

A manifest file is just a text file listing all the files it contains and their IDs.

File Name Contents

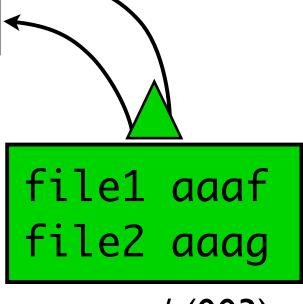
File Name Contents

```
aaaa 🛕 / (001)
aaab I file1 (V1)
aaac I file2 (V1)
aaad \triangle / (002)
aaae 🛕 / (003)
aaaf [ file1 (V2)
aaag | file2 (v2)
   file1 aaaf
   file2 aaac
    aaad -- / (002)
```

File Name Contents

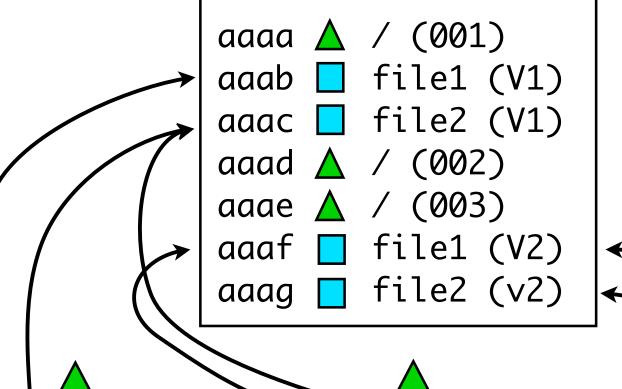
```
aaaa  \( \lambda \) (001)
aaab  \( \lambda \) file1 (V1)
aaac  \( \lambda \) file2 (V1)
aaad  \( \lambda \) (002)
aaae  \( \lambda \) (003)
aaaf  \( \lambda \) file1 (V2)
aaag  \( \lambda \) file2 (v2)
```

File Name Contents



aaae -- / (003)

File Name Contents



file1 aaab file2 aaac

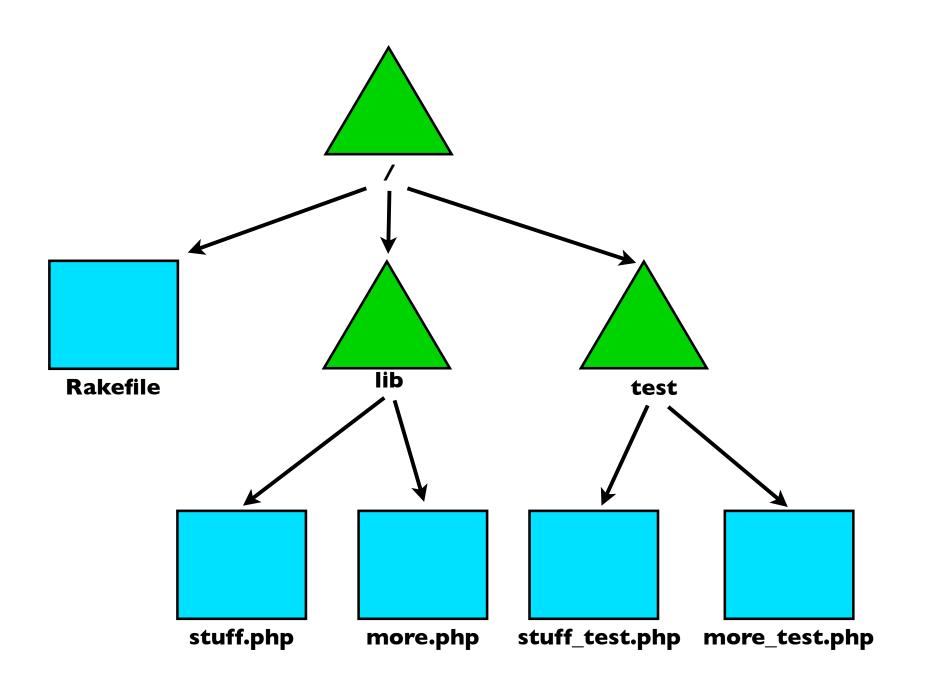
aaaa -- / (001)

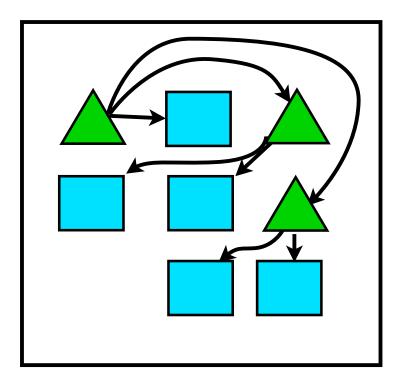
file1 aaaf file2 aaac

aaac -- / (002)

file1 aaaf file2 aaag

aaae -- / (003)

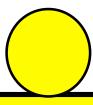




Metadata would be nice

What Metadata?

- The Root tree of the snapshot
- Name/EMail of the person committing the snapshot
- Date/Time of the snapshot
- Some comments describing the snapshot



tree aaae author Jim <<u>me@email.com</u>> 1246996320 -0400

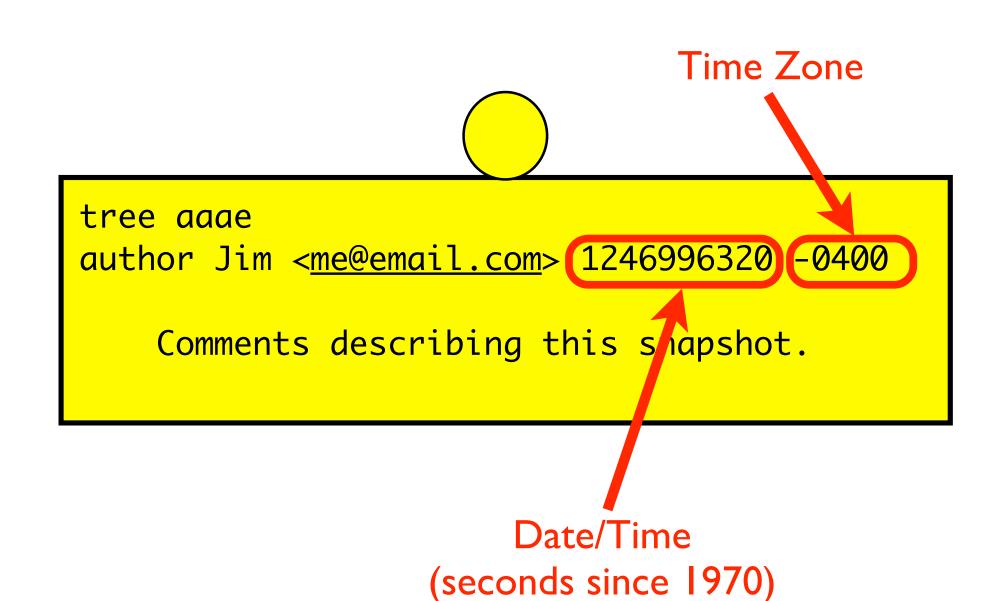
Comments describing this snapshot.

Root Manifest for Snapshot

tree aaae author Jim <<u>me@email.com</u>> 1246996320 -0400

Comments describing this snapshot.

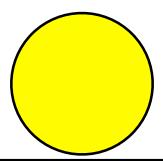
Author & EMail tree aaae author Jim < me@email.com > 1246996320 -0400 Comments describing this snapshot.





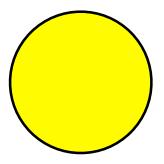
Comments describing this snapshot.

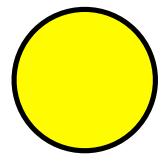


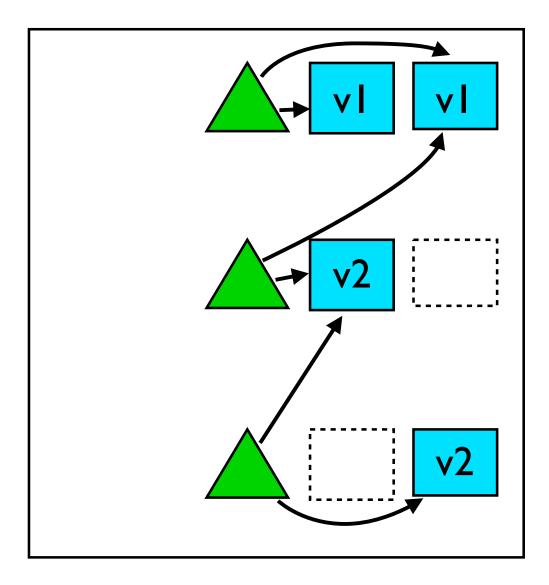


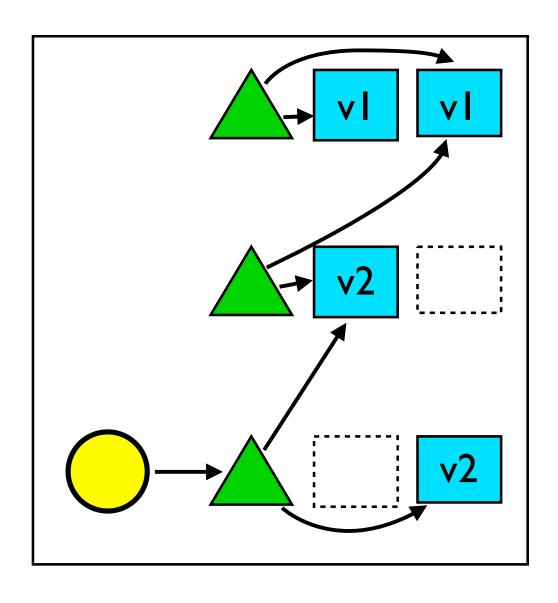
tree aaae author Jim <<u>me@email.com</u>> 1246996320 -0400

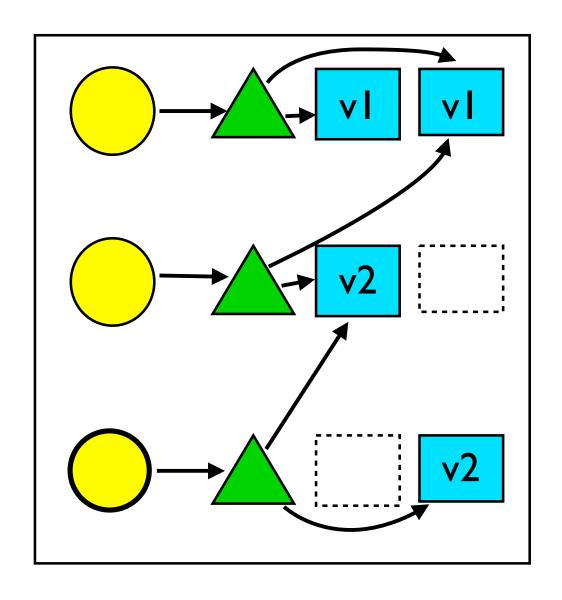
Comments describing this snapshot.

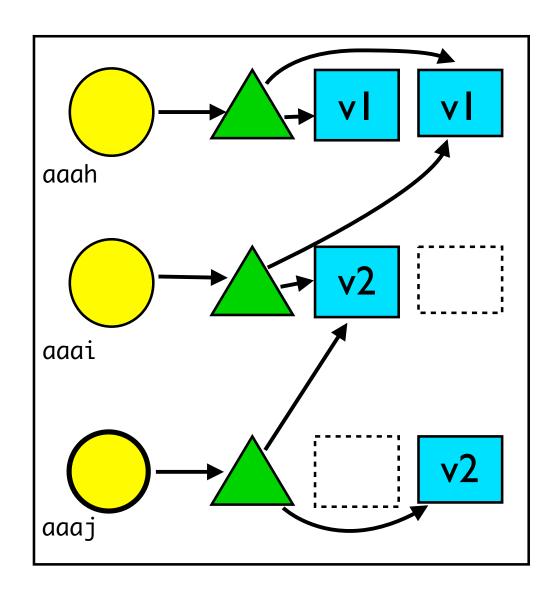












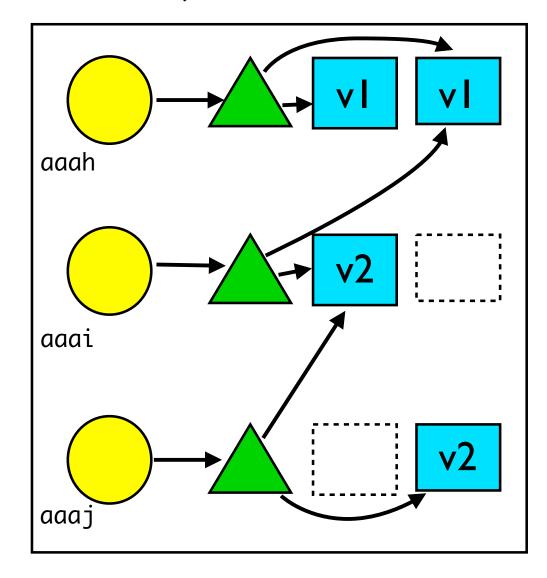
tree aaae author Jim <<u>me@email.com</u>> 1246996320 -0400

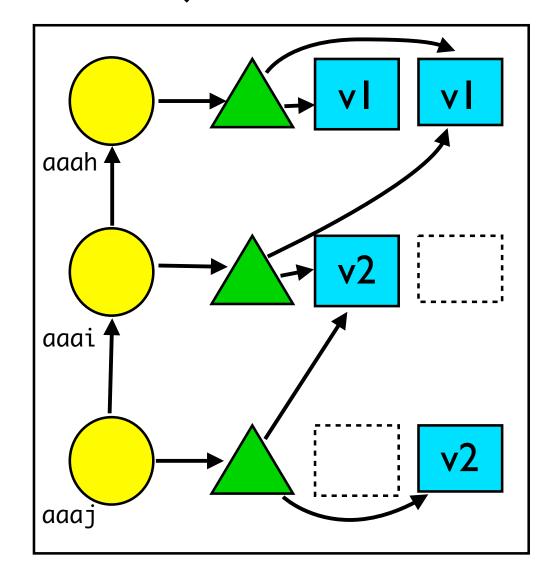
Comments describing this snapshot.

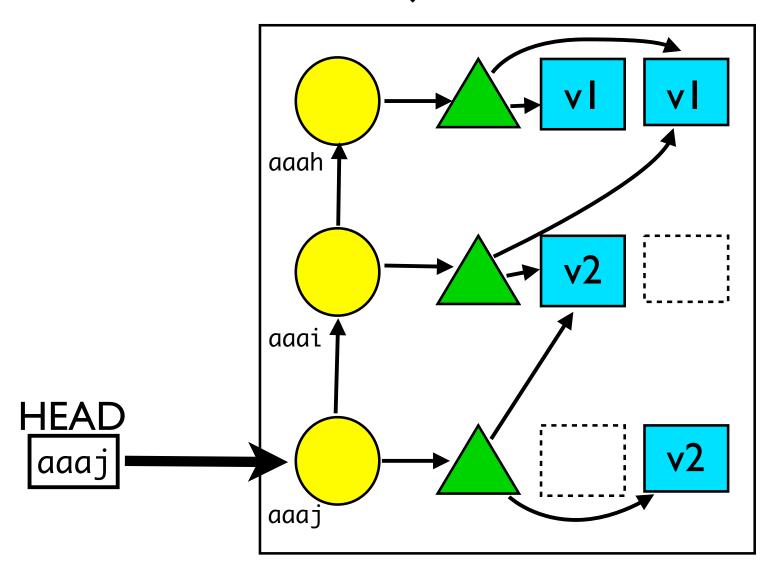
tree aaae parent aaai

author Jim <<u>me@email.com</u>> 1246996320 -0400

Comments describing this snapshot.

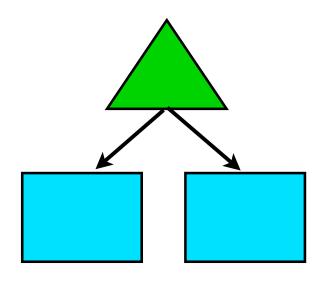




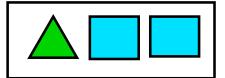


Summary So Far ...

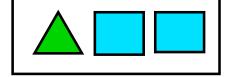
Project Archive



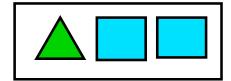
001.tar

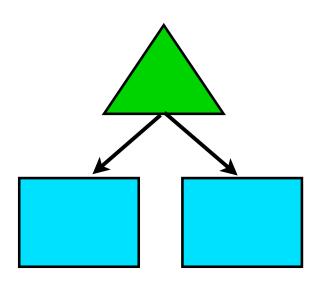


002.tar

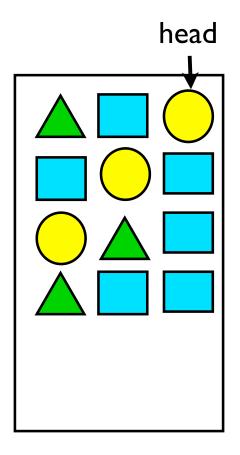


003.tar

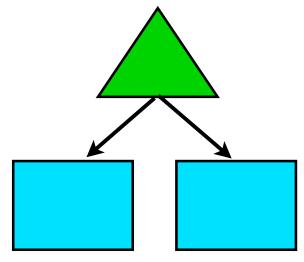




Project Archive

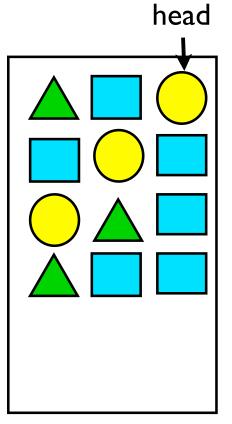


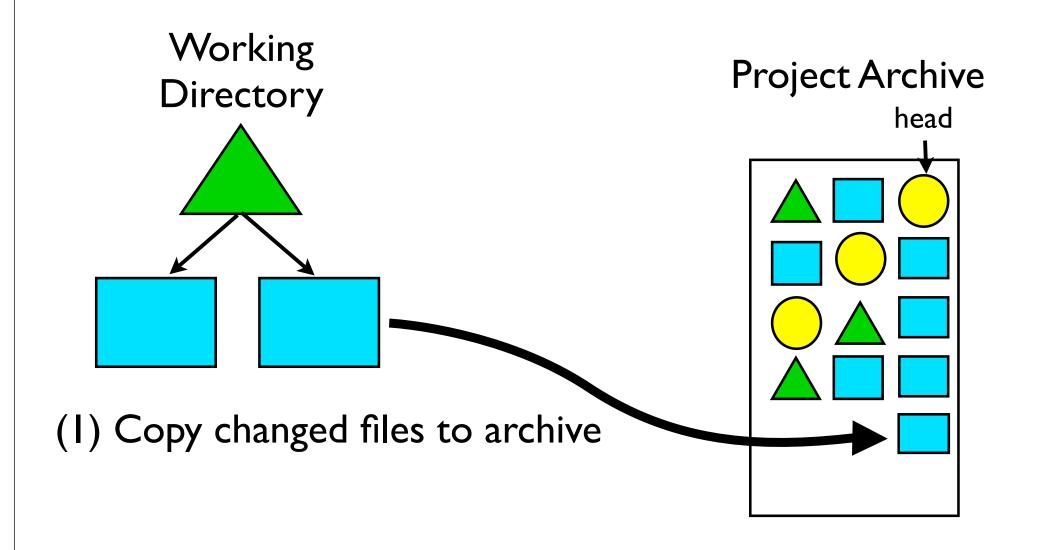
How do we create a snapshot now?



(I) Copy changed files to archive

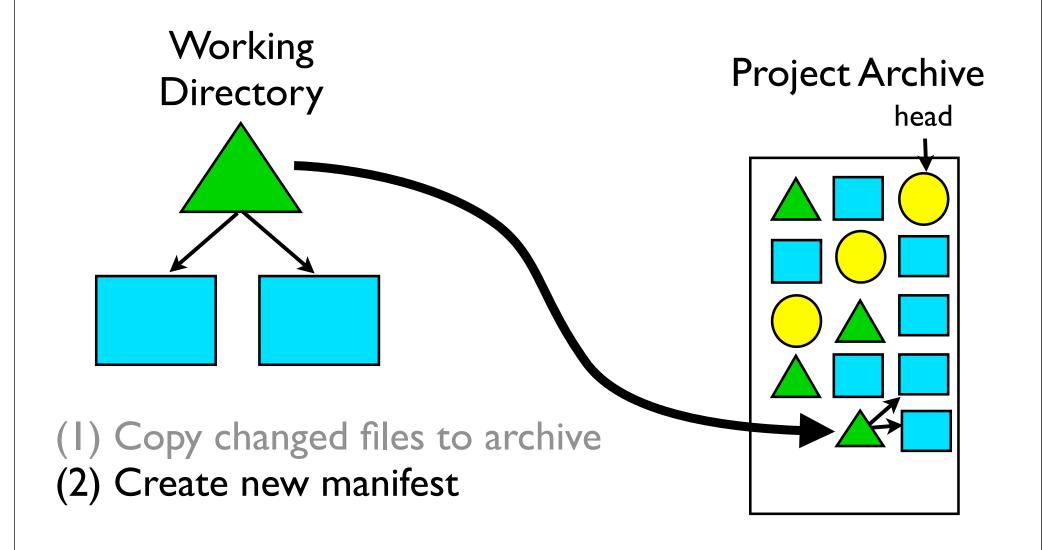
Project Archive





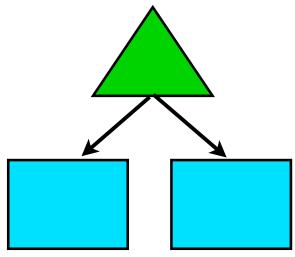
- (I) Copy changed files to archive
- (2) Create new manifest

Project Archive head



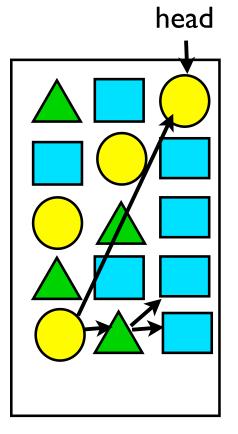
- (I) Copy changed files to archive
- (2) Create new manifest
- (3) Create new snapshot record

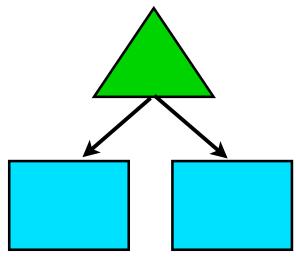
Project Archive head



- (I) Copy changed files to archive
- (2) Create new manifest
- (3) Create new snapshot record

Project Archive

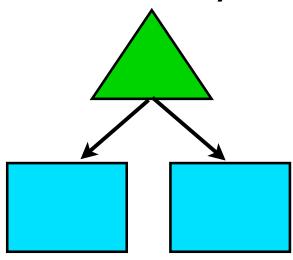




- (I) Copy changed files to archive
- (2) Create new manifest
- (3) Create new snapshot record
- (4) Readjust HEAD

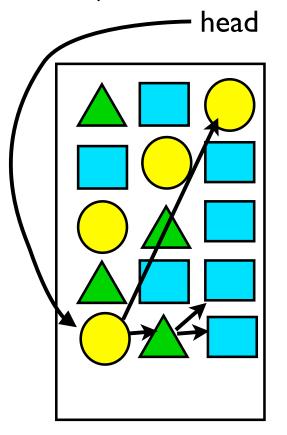
Project Archive

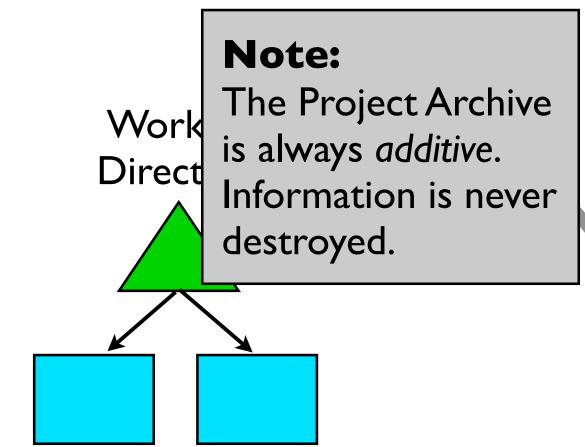
head



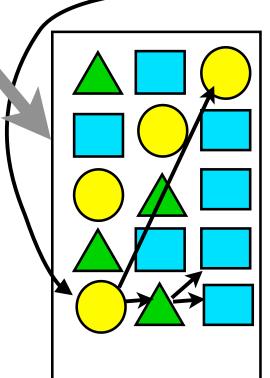
- (I) Copy changed files to archive
- (2) Create new manifest
- (3) Create new snapshot record
- (4) Readjust HEAD

Project Archive

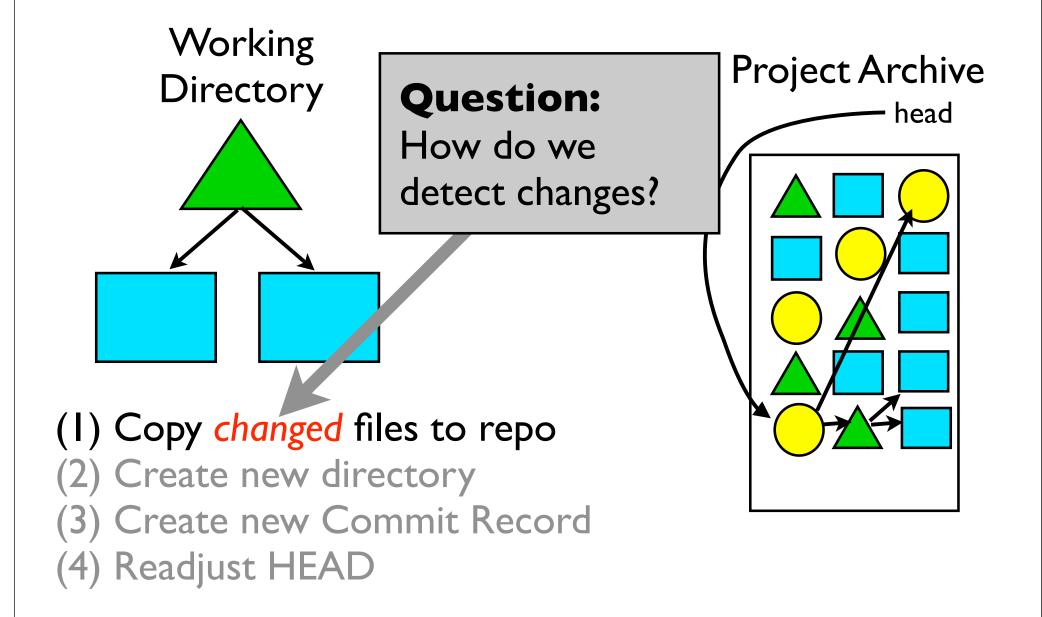




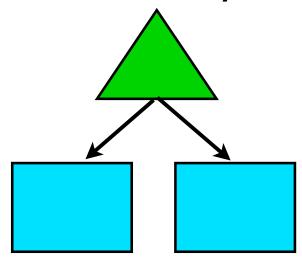
· head



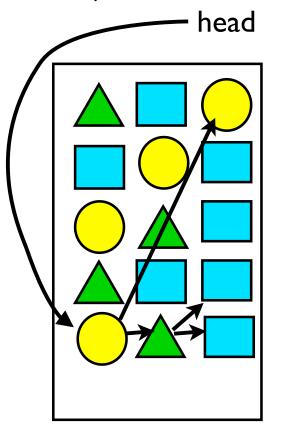
- (I) Copy changed files to archive
- (2) Create new manifest
- (3) Create new snapshot record
- (4) Readjust HEAD

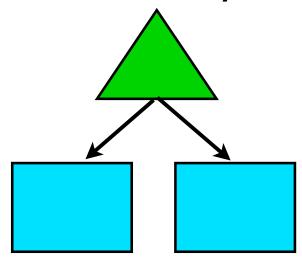


Friday, September 18, 2009 56

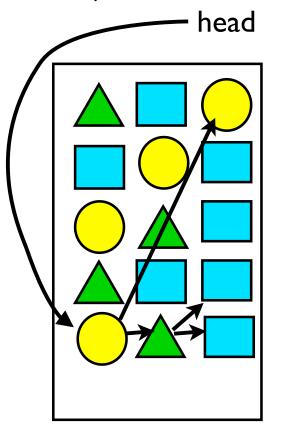


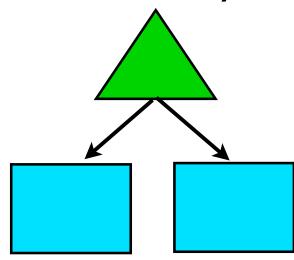
- (I) Copy changed files to repo
- (2) Create new directory
- (3) Create new Commit Record
- (4) Readjust HEAD



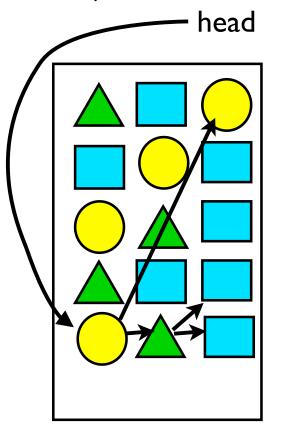


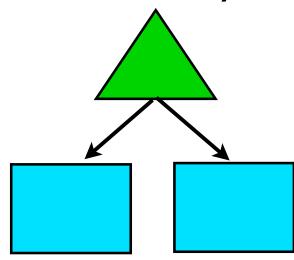
- (I) Copy changed files to repo
- (2) Create new directory
- (3) Create new Commit Record
- (4) Readjust HEAD



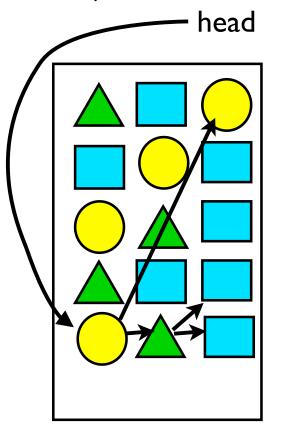


- (I) Copy changed files to repo
- (2) Create new directory
- (3) Create new Commit Record
- (4) Readjust HEAD





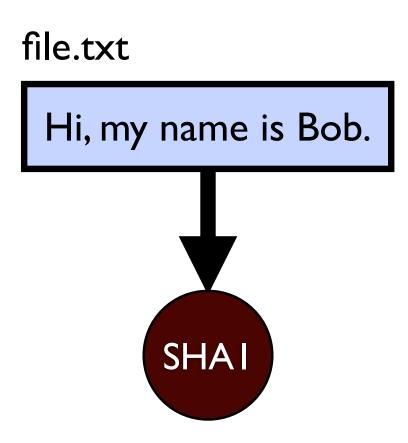
- (I) Copy changed files to repo
- (2) Create new directory
- (3) Create new Commit Record
- (4) Readjust HEAD

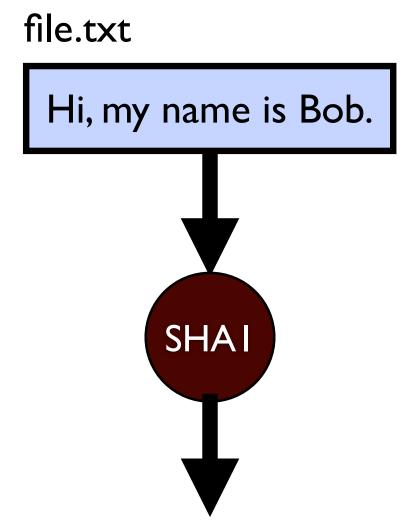


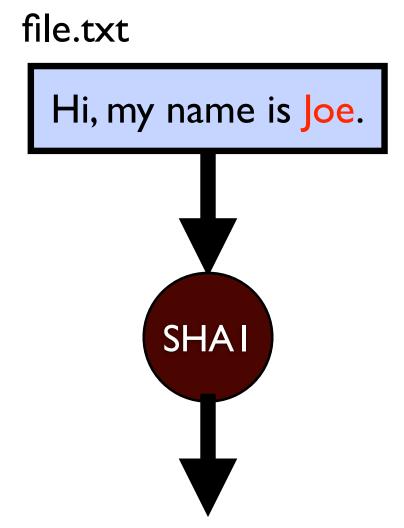
Detecting Duplication

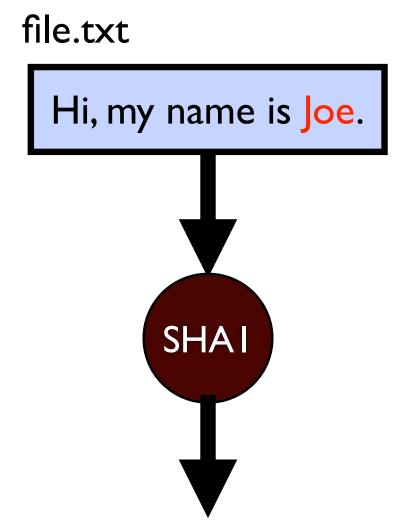
file.txt

Hi, my name is Bob.

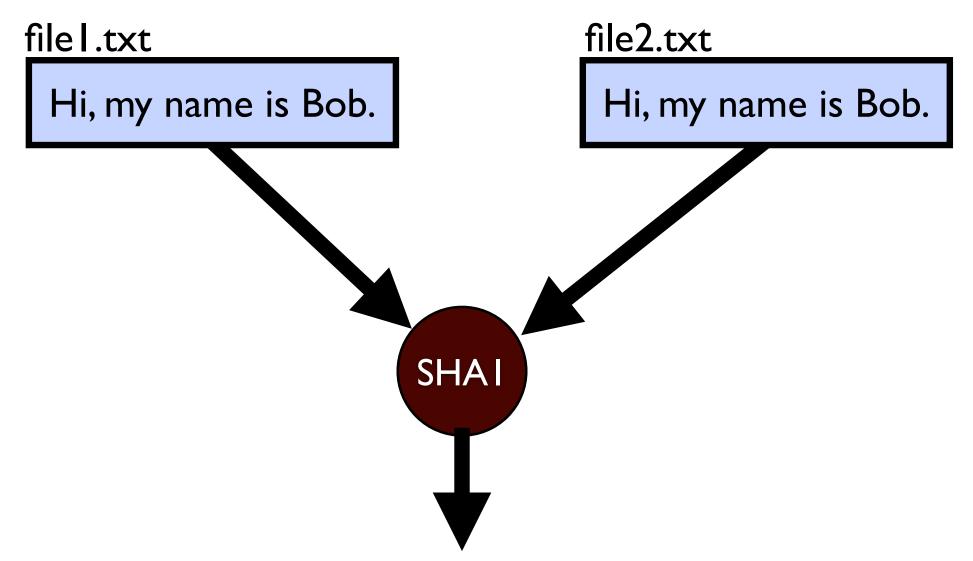


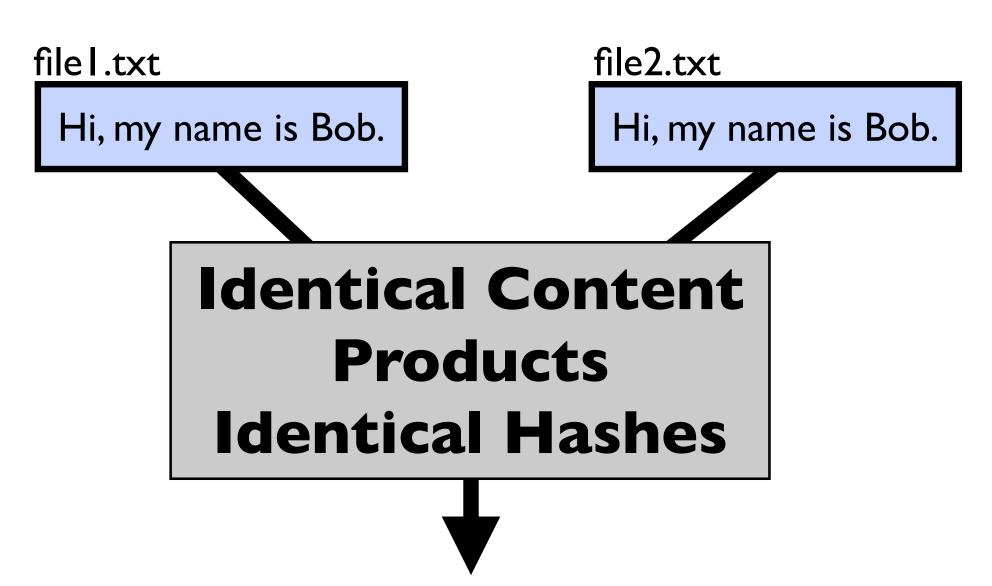






"53f2509e877de3cad85c547a0c2a22b43c5d81eb"





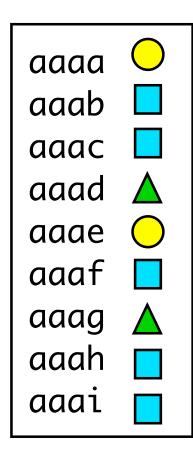
File Names in the Archive

File Names in the Archive

- Arbitrary
- Unique

Idea:

Use the SHA1 Hash for the File Name



Friday, September 18, 2009 66

HEAD

75e3dee3503e7ad5...

```
75e3dee3503e7ad5...

1586ad8975e3dee3...

f6da34481586ad89...

f99b89d6f6da3448...

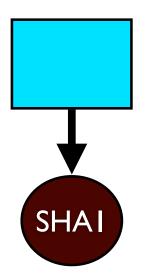
503d6749f99b89d6...

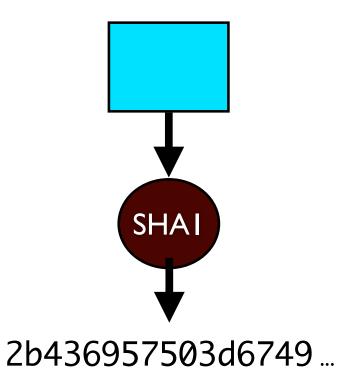
2b436957503d6749...

944ff522944ff522...

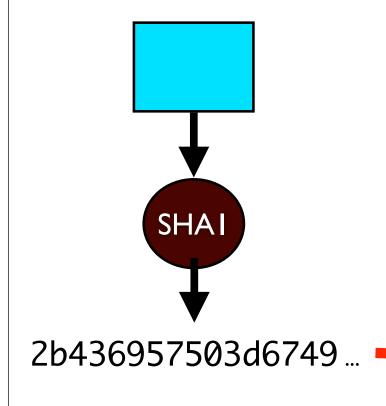
6c024e25944ff522...

503e7ad56c024e25...
```





Friday, September 18, 2009 68



Match!

Project Archive

75e3dee3503e7ad5...

1586ad8975e3dee3...

f6da34481586ad89...

f99b89d6f6da3448...

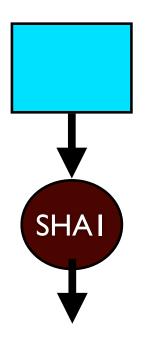
503d6749f99b89d6...

2b436957503d6749...

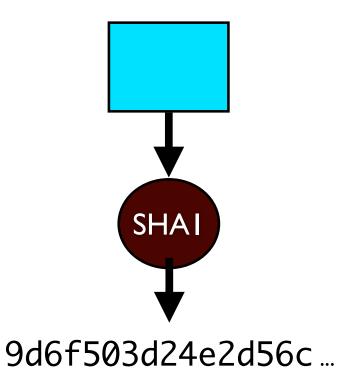
944ff522944ff522...

6c024e25944ff522...

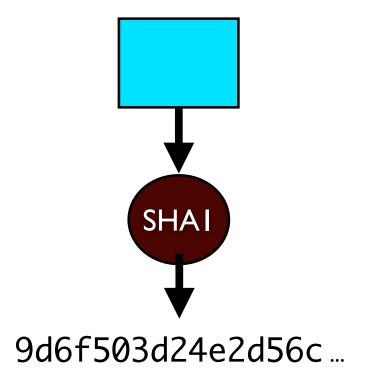
503e7ad56c024e25...



75e3dee3503e7ad5 1586ad8975e3dee3 f6da34481586ad89 f99b89d6f6da3448 503d6749f99b89d6 2b436957503d6749 944ff522944ff522 6c024e25944ff522
6c024e25944ff522 5 03e7ad56c024e25



75e3dee3503e7ad5 1586ad8975e3dee3 f6da34481586ad89 f99b89d6f6da3448 503d6749f99b89d6 2b436957503d6749 944ff522944ff522 6c024e25944ff522
6c024e25944ff522 503e7ad56c024e25



No Match!

Project Archive

```
75e3dee3503e7ad5...

1586ad8975e3dee3...

f6da34481586ad89...

f99b89d6f6da3448...

503d6749f99b89d6...

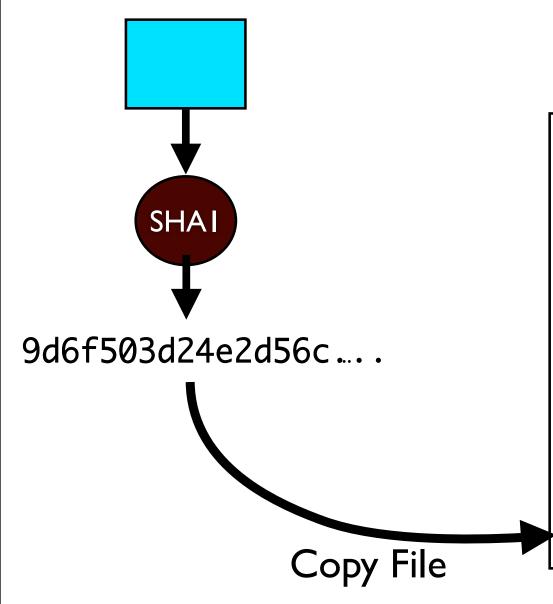
2b436957503d6749...

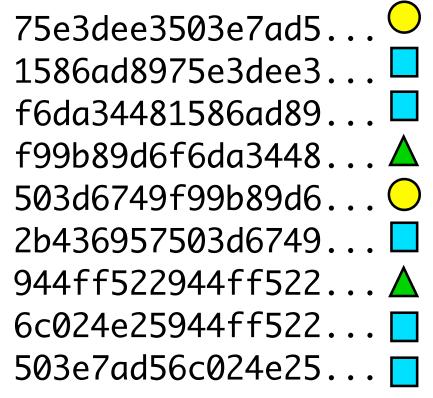
944ff522944ff522...

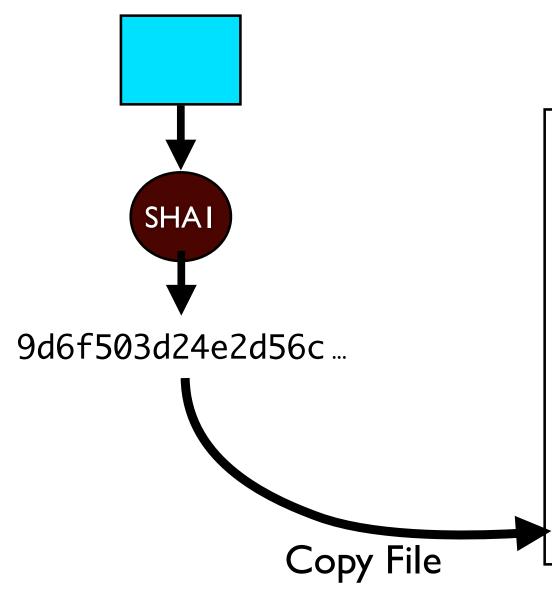
6c024e25944ff522...

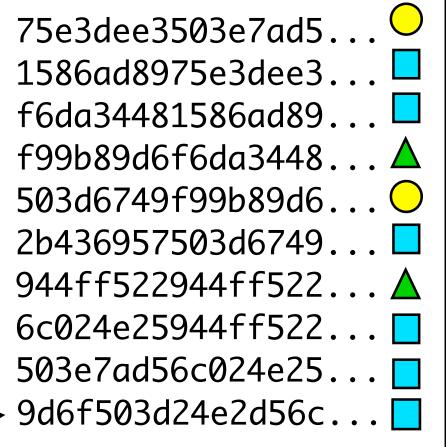
503e7ad56c024e25...
```

Friday, September 18, 2009 69

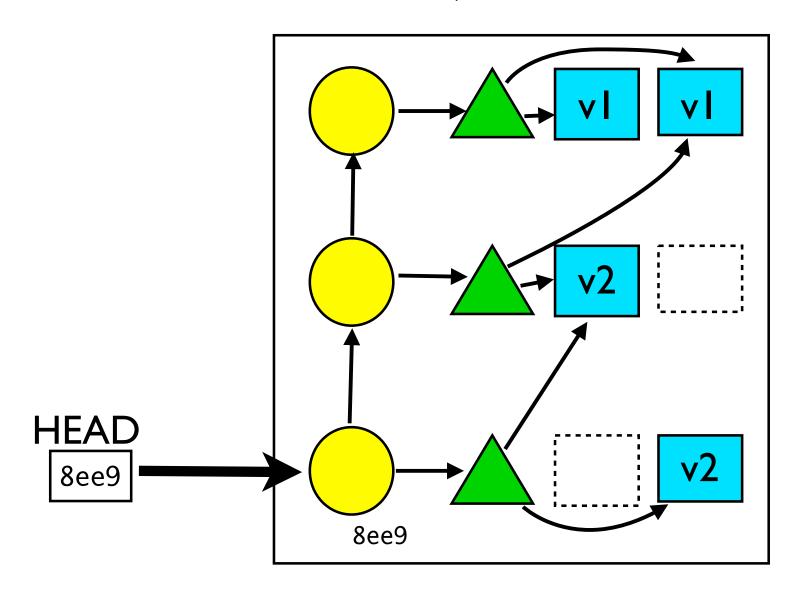


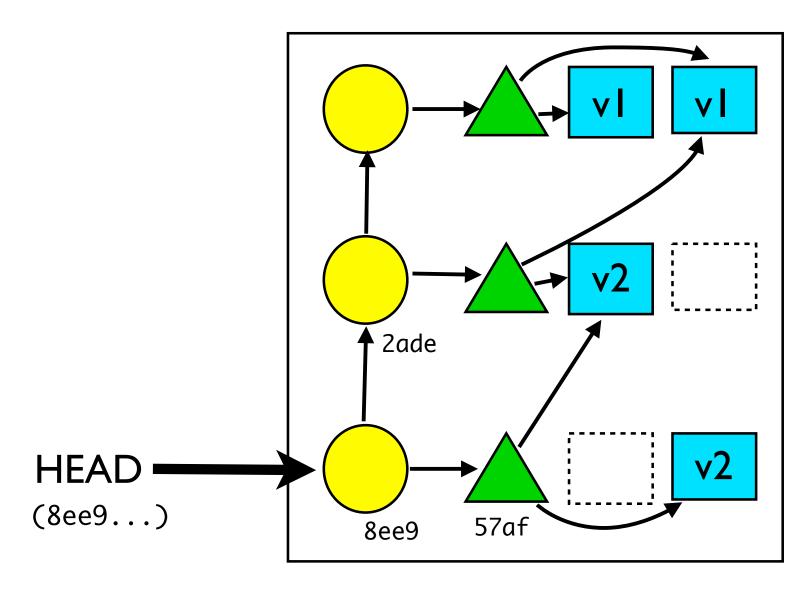


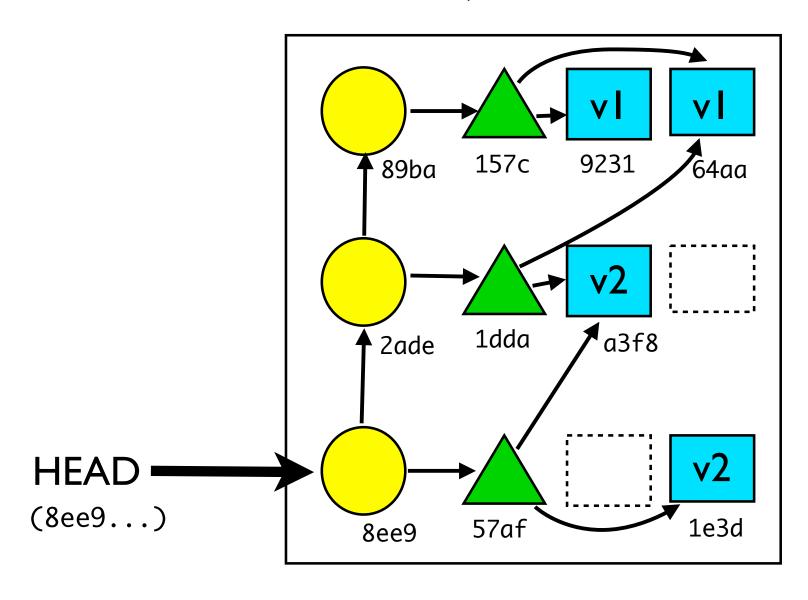


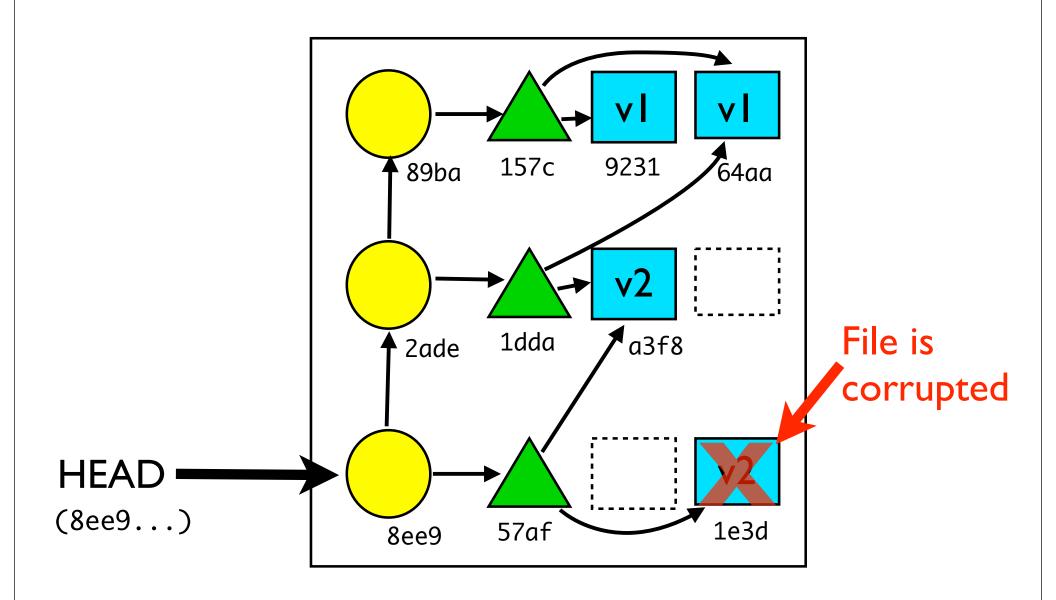


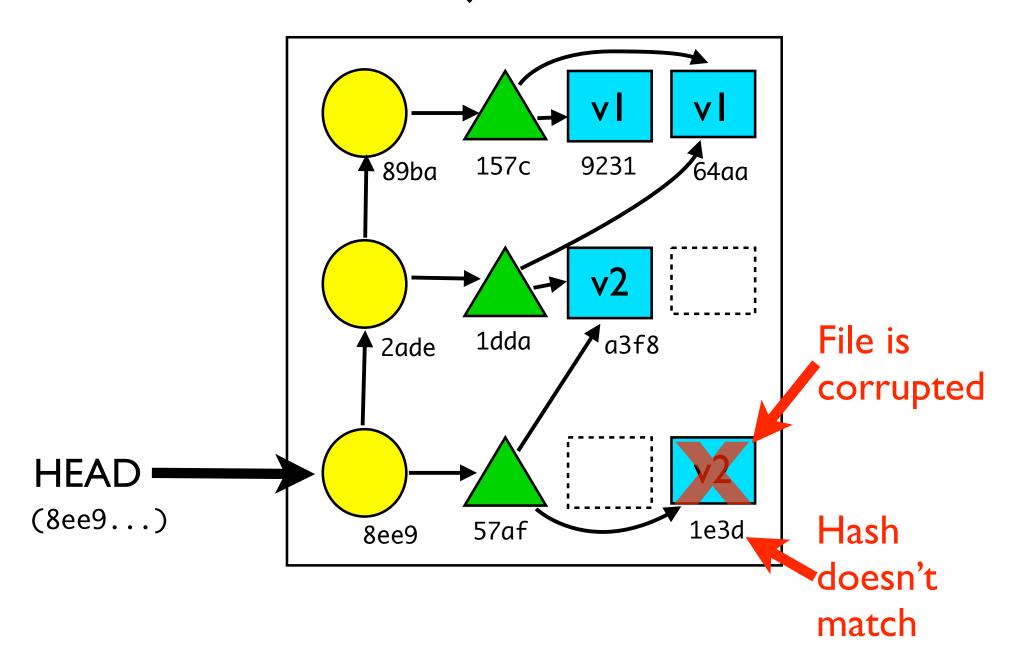
Project
Archive
Integrity

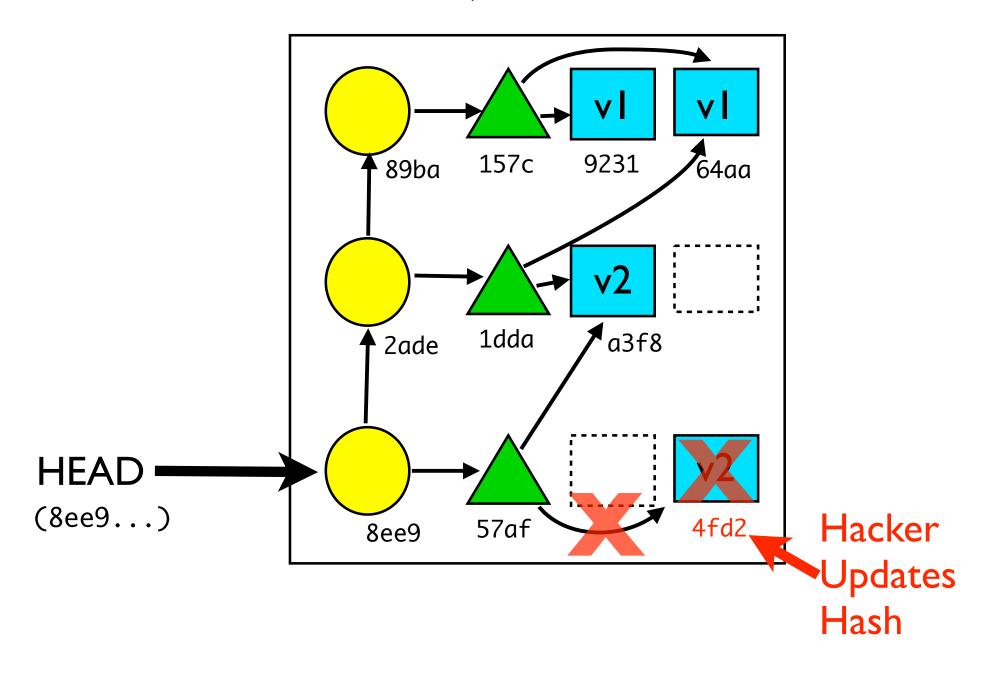


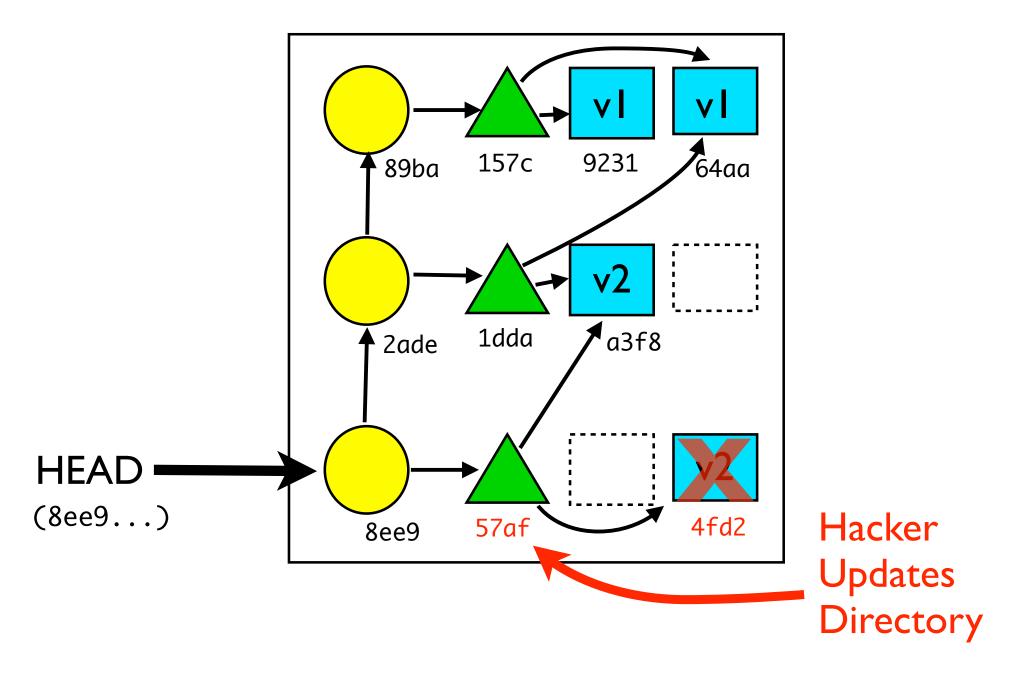


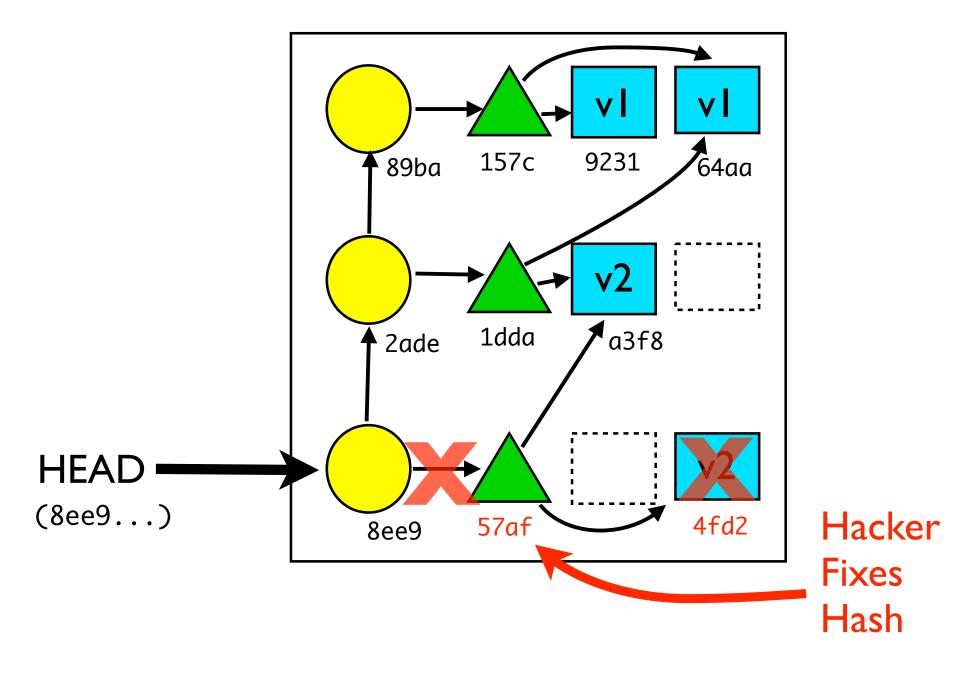


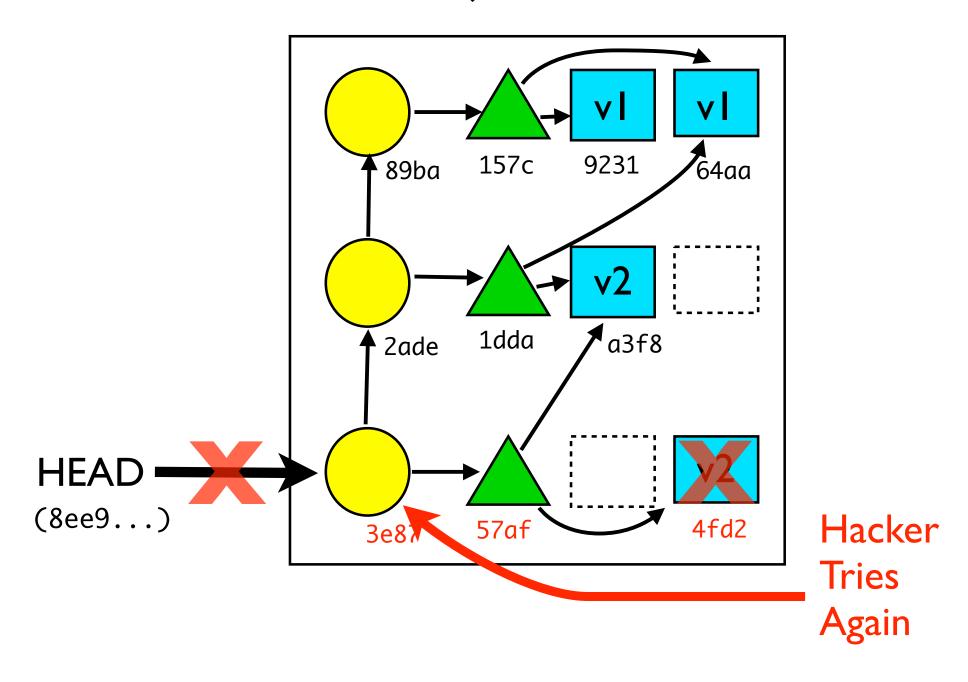






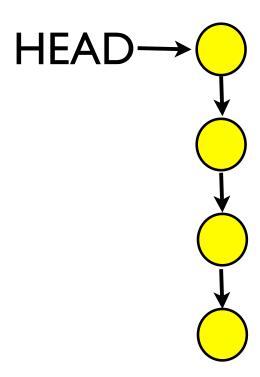


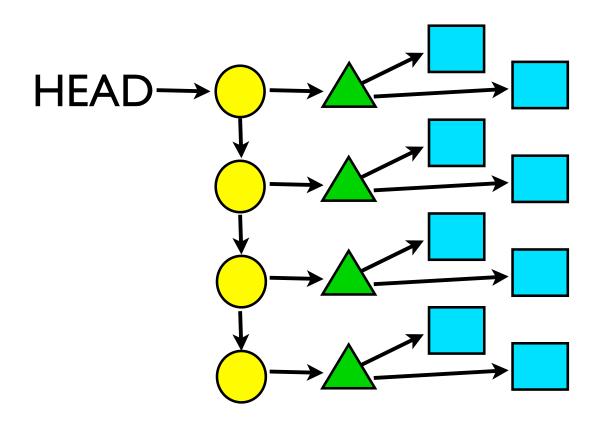


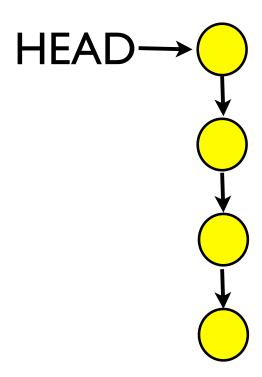


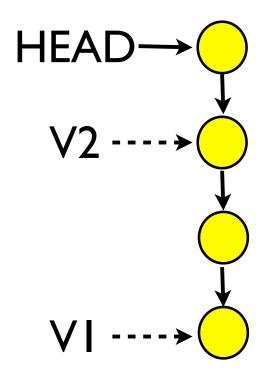
Tracking Snapshots

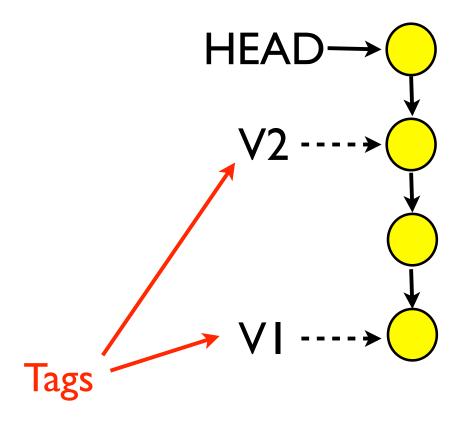
Need to keep track of versions by name.

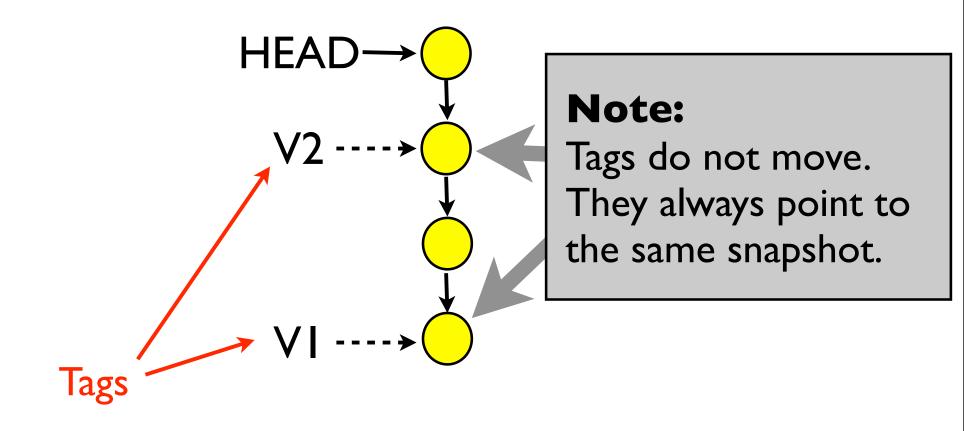


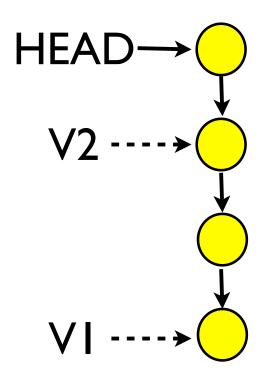




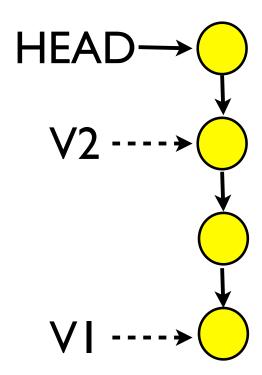


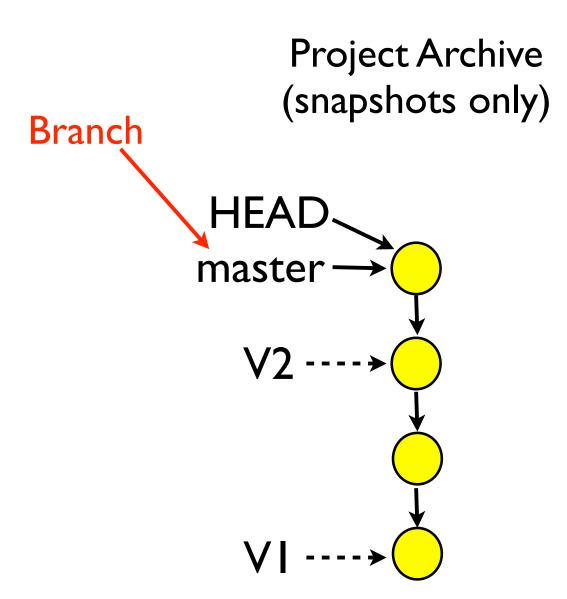


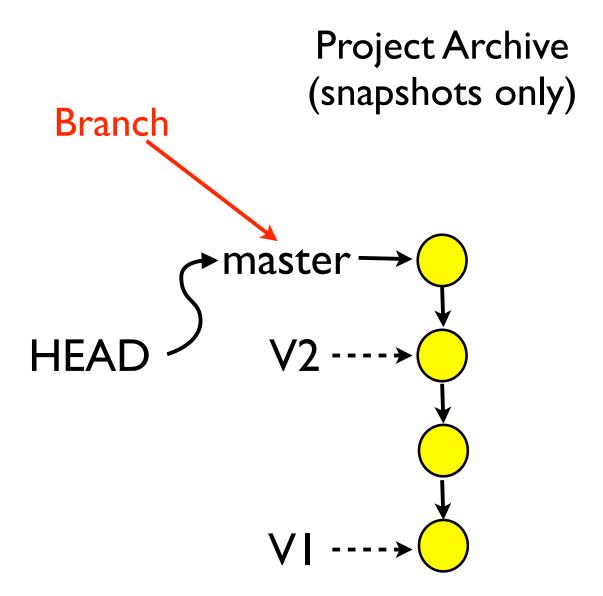


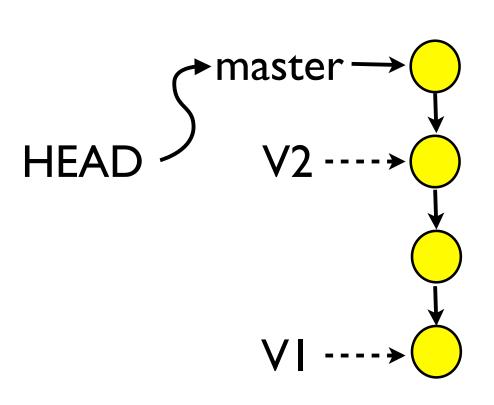


Need to handle changes to older versions while allowing development on the latest version.

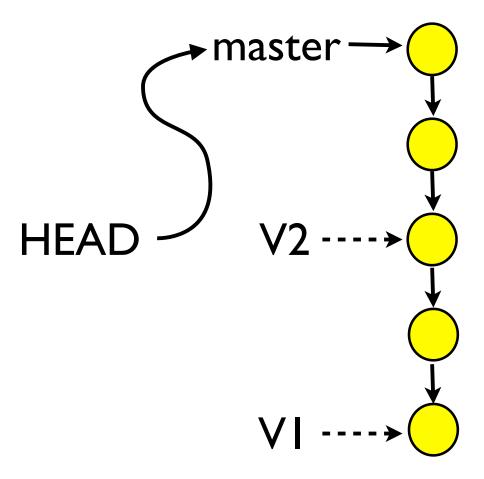




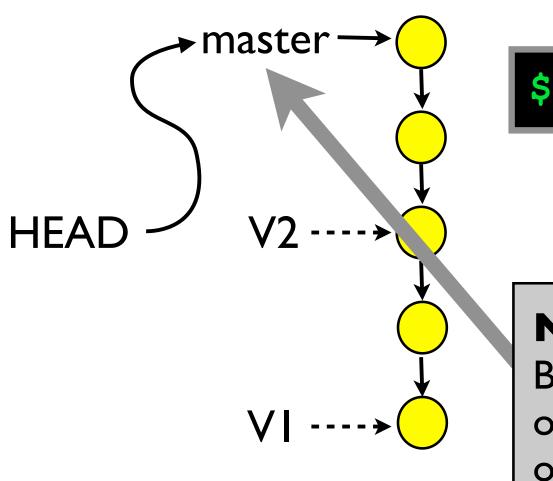








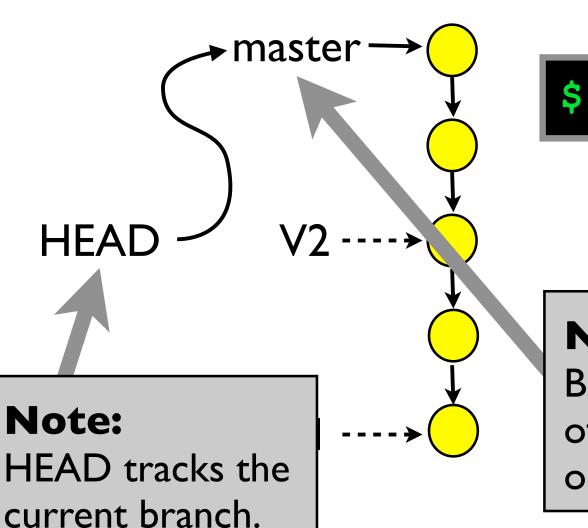
\$ csc snap



\$ csc snap

Note:

Branches keep track of the last snapshot on that branch.

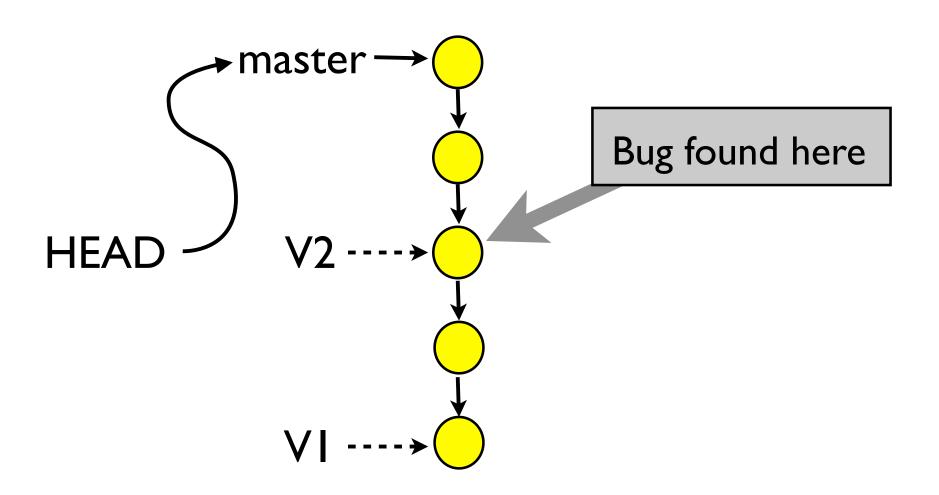


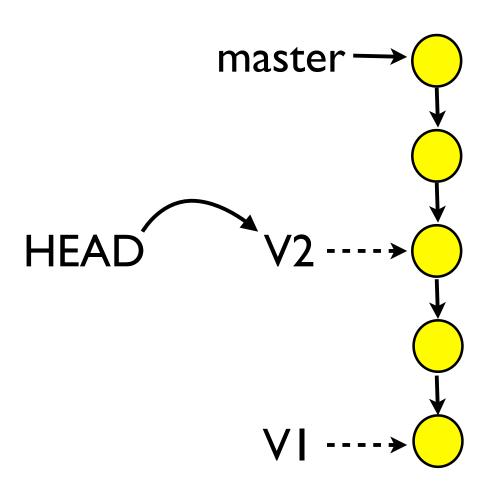
\$ csc snap

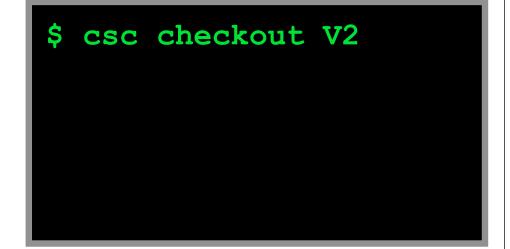
Note:

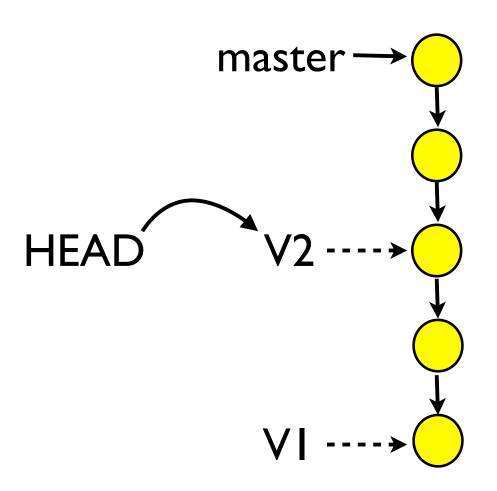
Branches keep track of the last snapshot on that branch.

Oops! We found a bug

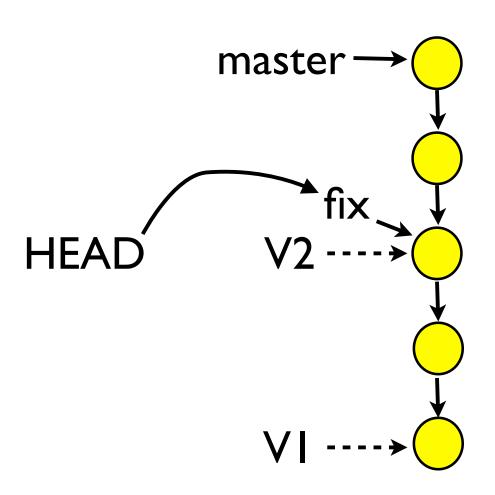




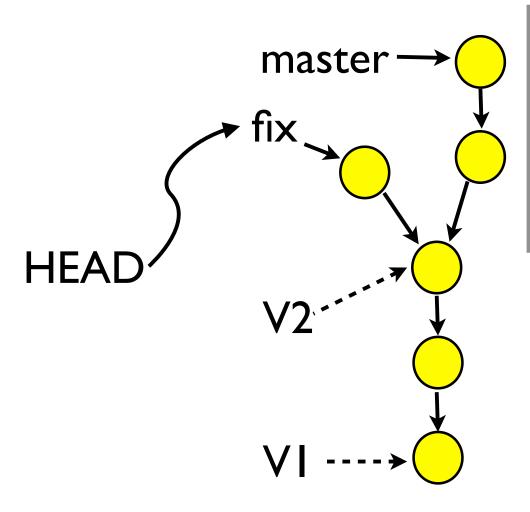




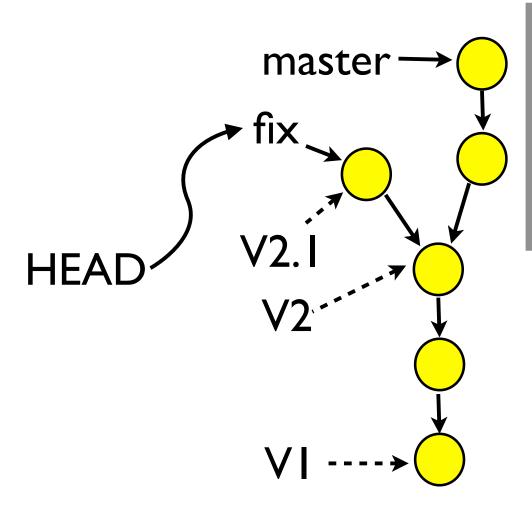
```
$ csc checkout V2
$ edit files
```



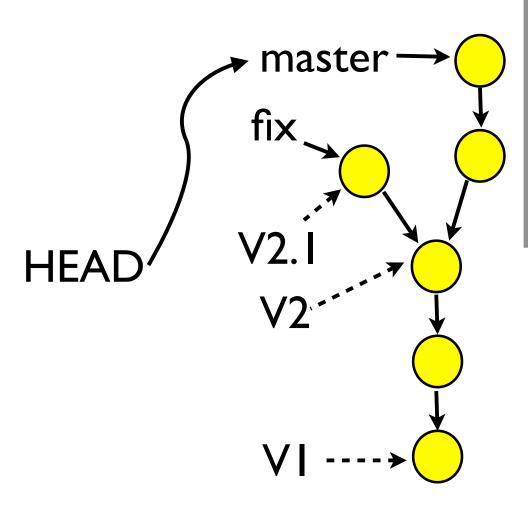
```
$ csc checkout V2
$ edit files
$ csc branch fix
```



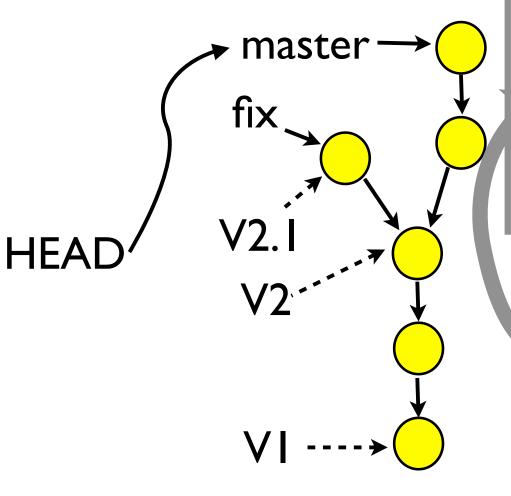
```
$ csc checkout V2
$ edit files
$ csc branch fix
$ csc snap
```



```
$ csc checkout V2
$ edit files
$ csc branch fix
$ csc snap
$ csc tag V2.1
```



```
$ csc checkout V2
$ edit files
$ csc branch fix
$ csc snap
$ csc tag V2.fix
$ csc checkout master
```

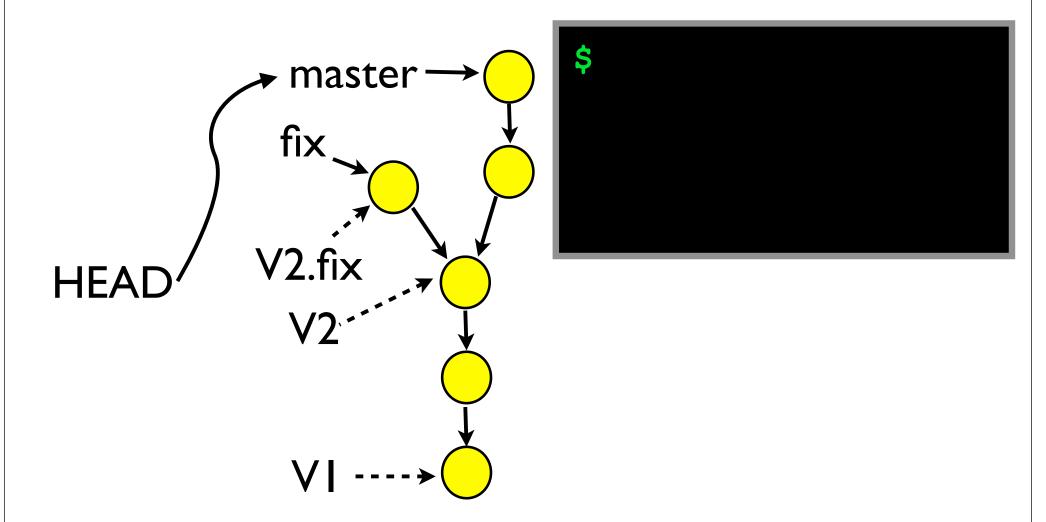


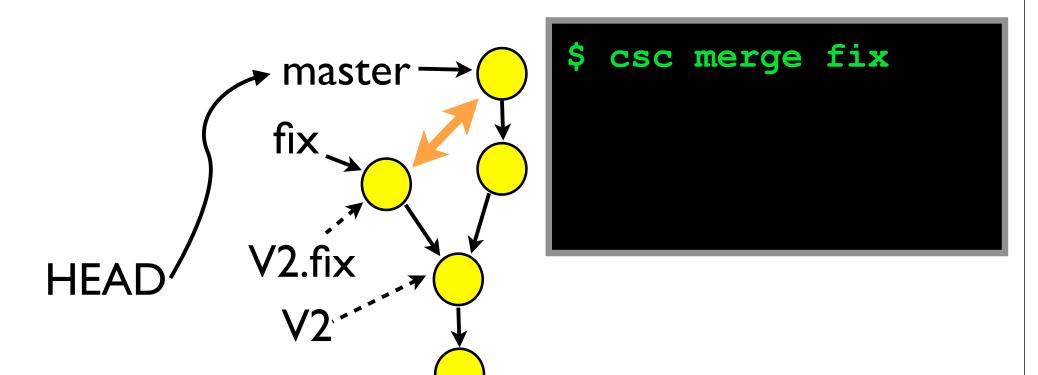
```
$ csc checkout V2
$ edit files
csc branch fix
csc snap
$ csc tag V2.fix
$ csc checkout master
```

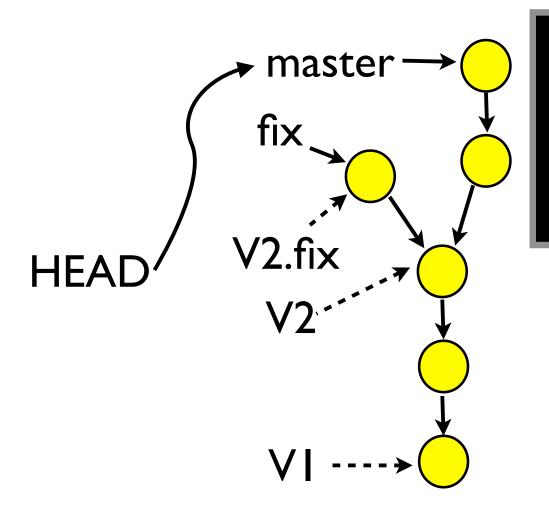
Note:

We didn't create a branch until after we edited files.

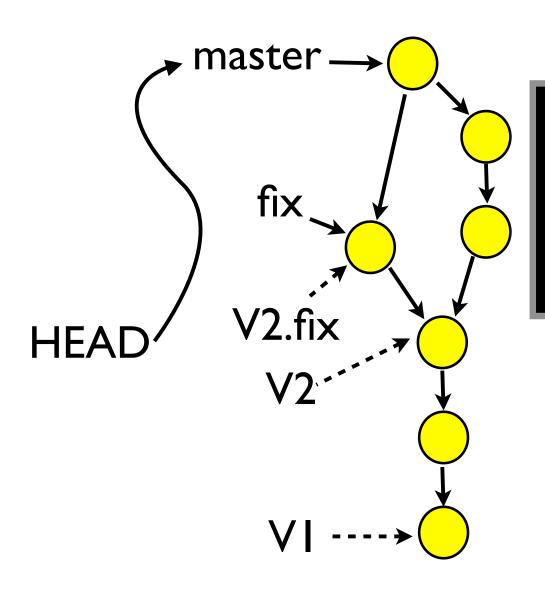
Merge Bug Fixes to Master



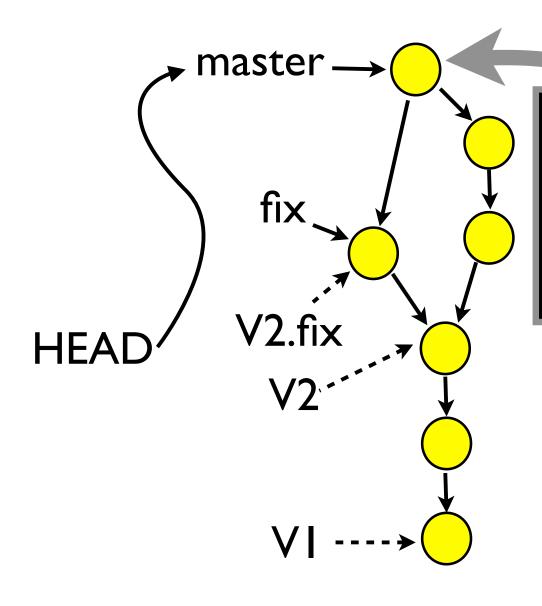




\$ csc merge fix
(resolve conflict)



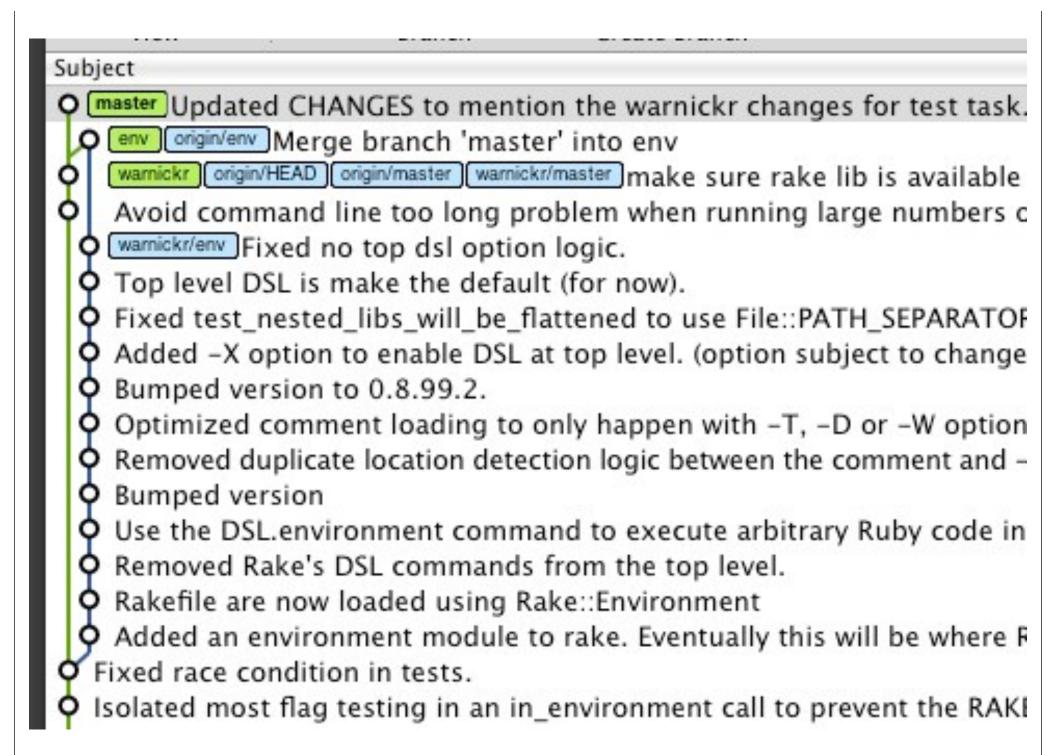
\$ csc merge fix
(resolve conflict)
\$ csc snap



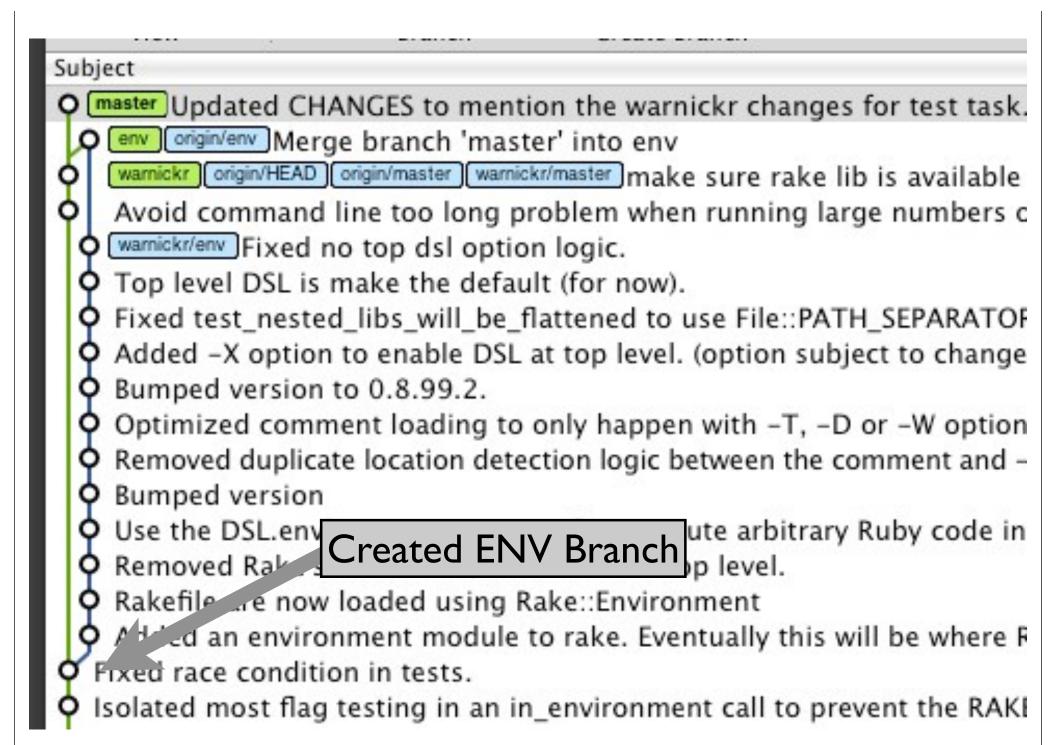
\$ csc merge fix (resolve conflict) \$ csc s.ap

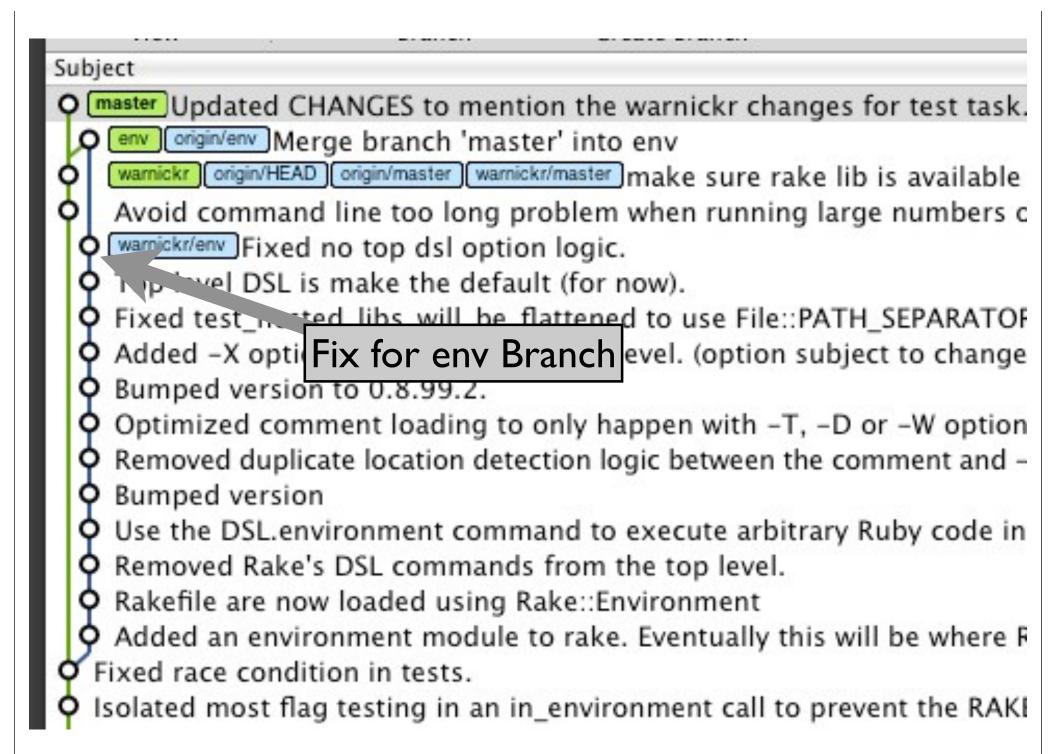
Note:

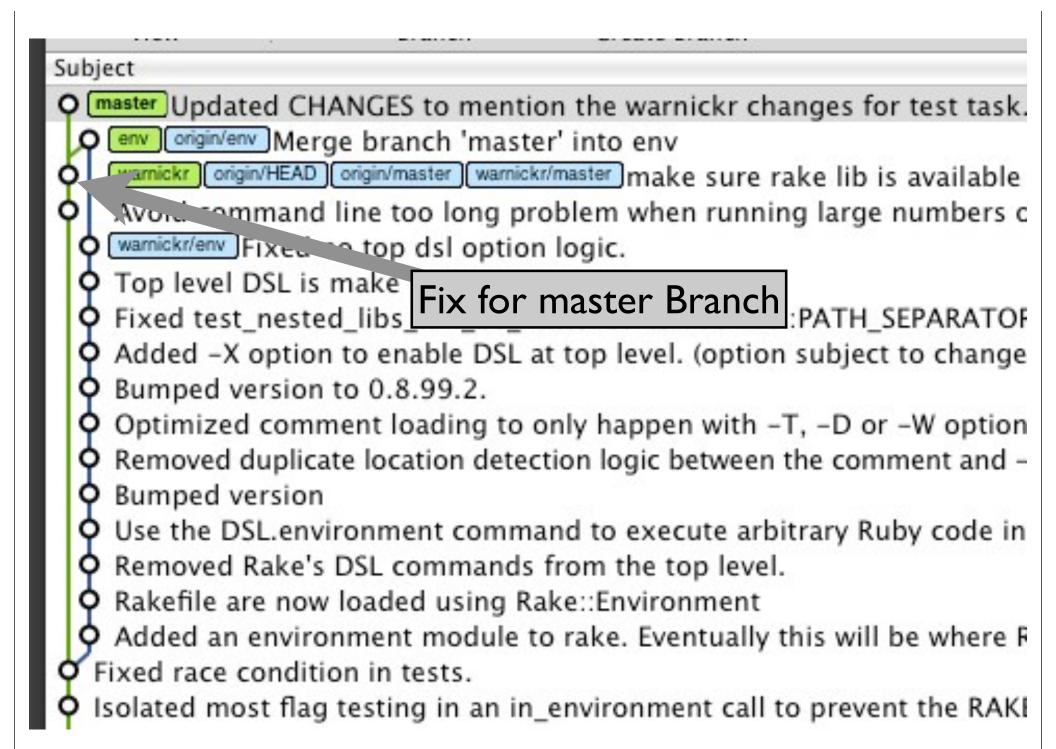
This snapshot has two parents. The changes from both parents have been merged together.

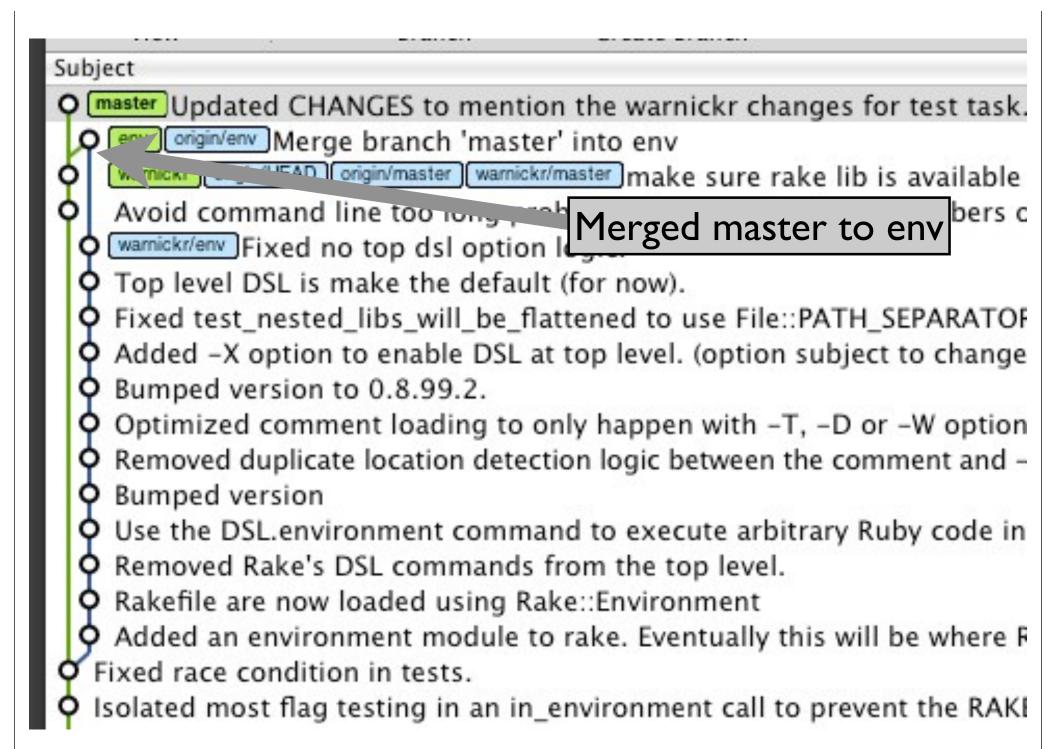


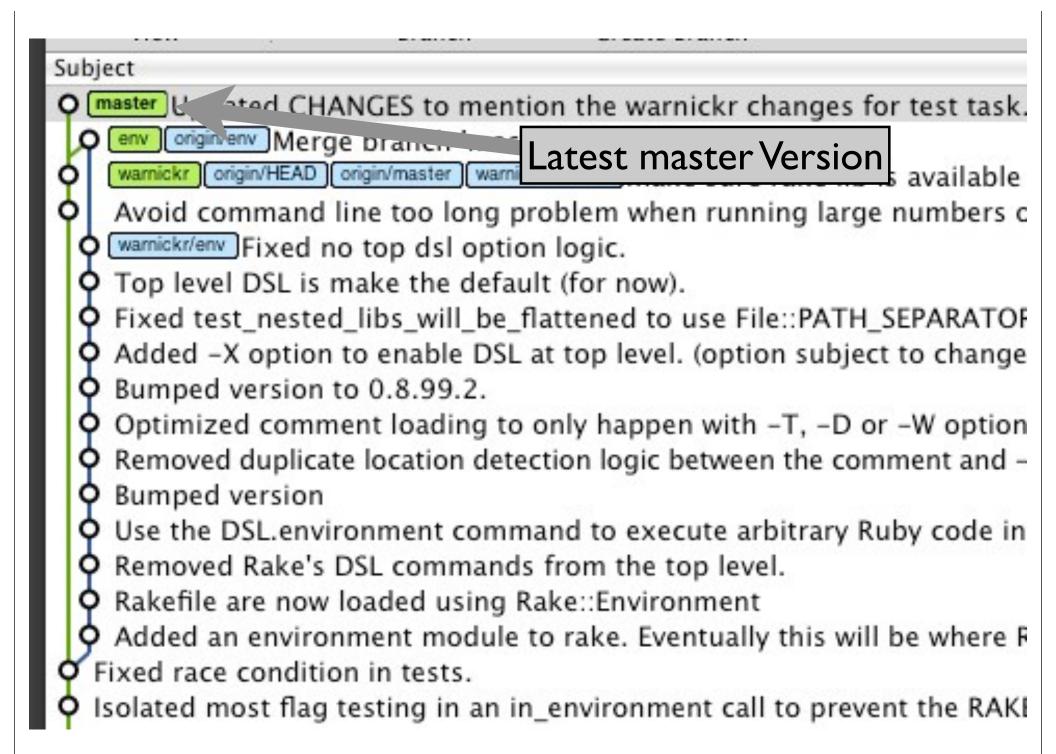
109

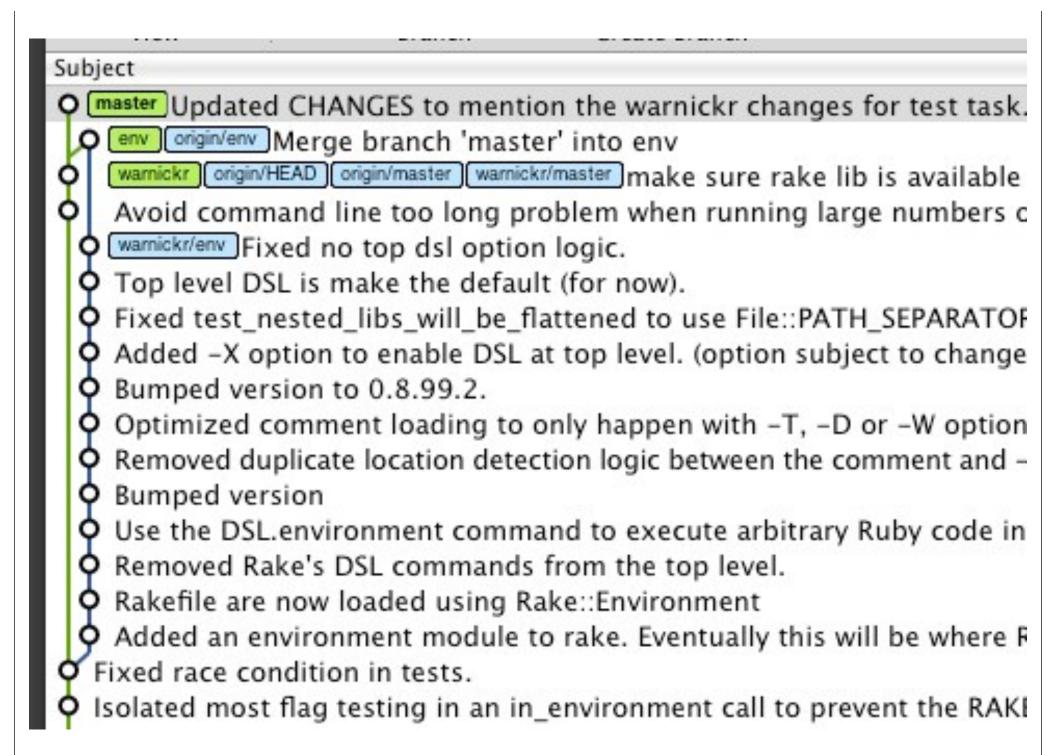




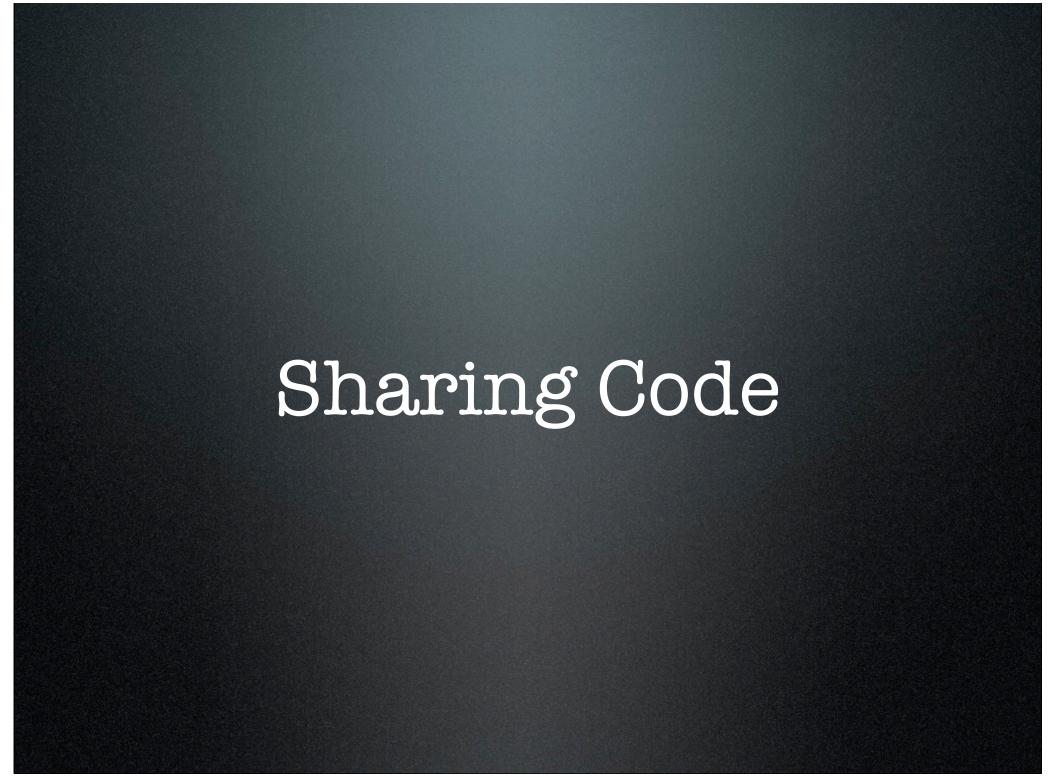


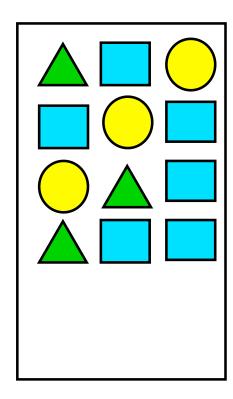






109

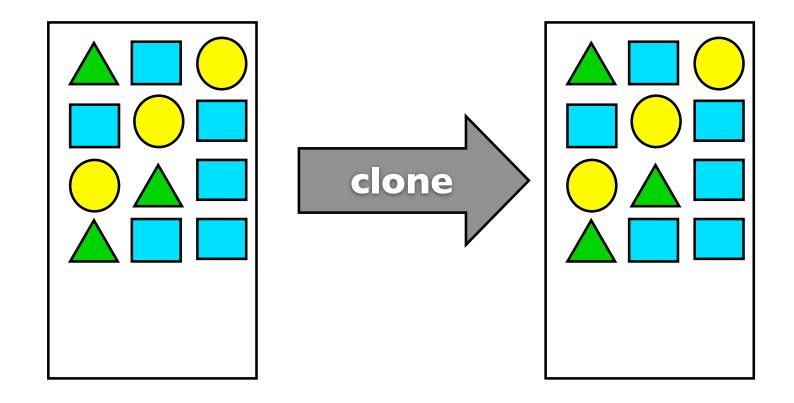




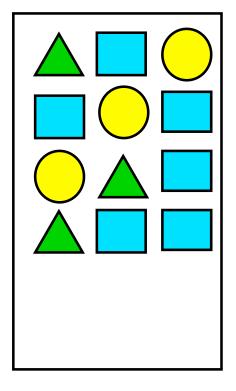
Friday, September 18, 2009

111

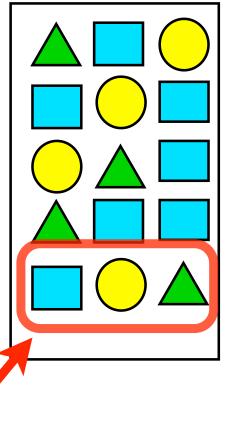
Your Archive



you\$ csc clone url_for_my_repo

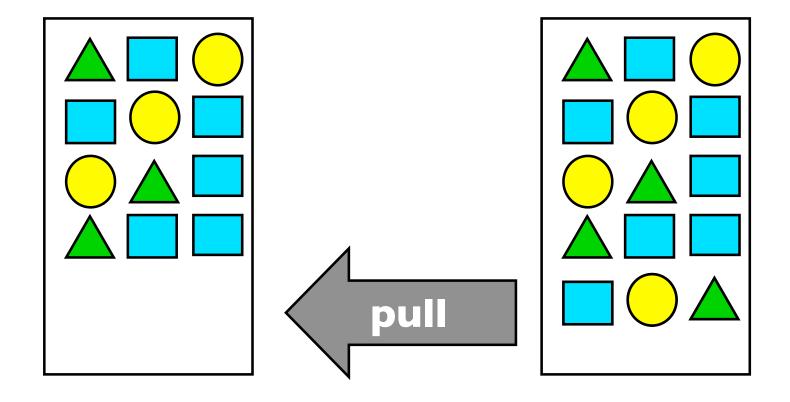


Your Archive



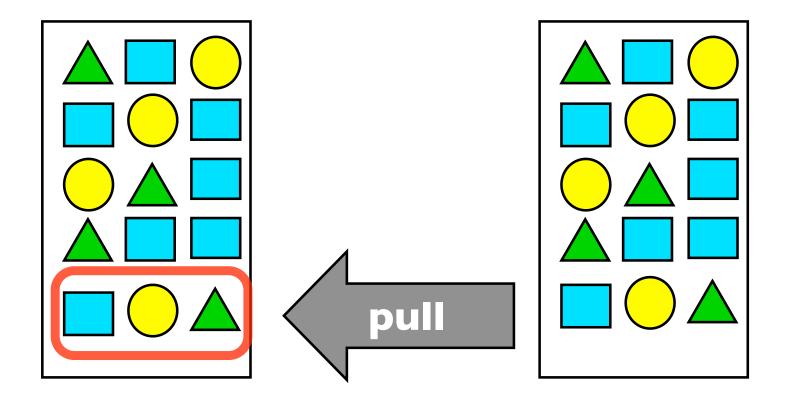
You make changes

Your Archive



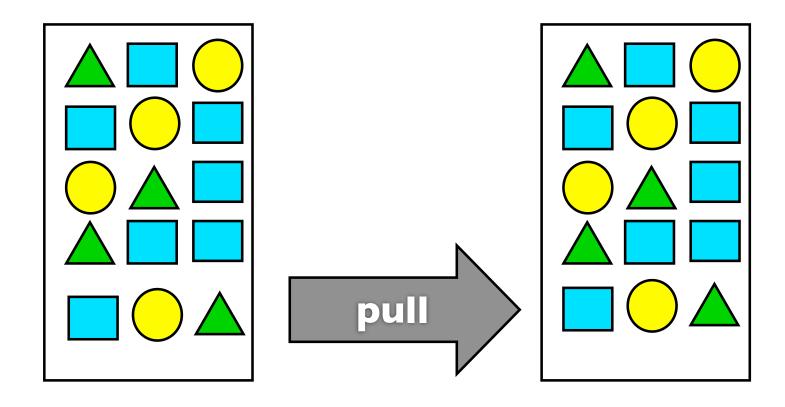
me\$ csc pull url_for_my_repo

Your Archive



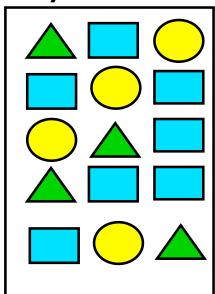
me\$ csc pull url_for_my_repo

Your Archive

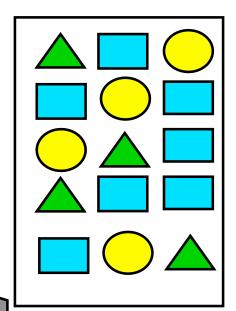


you\$ csc pull url_for_my_repo

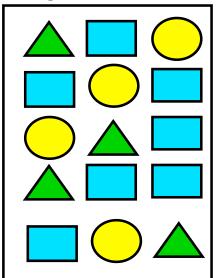
Common Remote Uses



Remote Archive

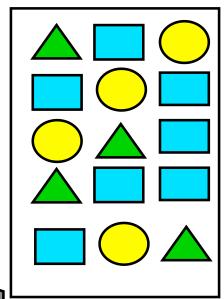


My Archive

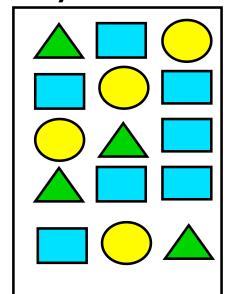




Shared Remote Archive



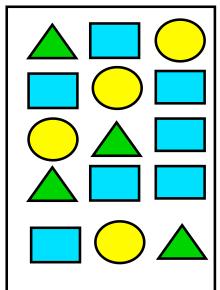
My Archive



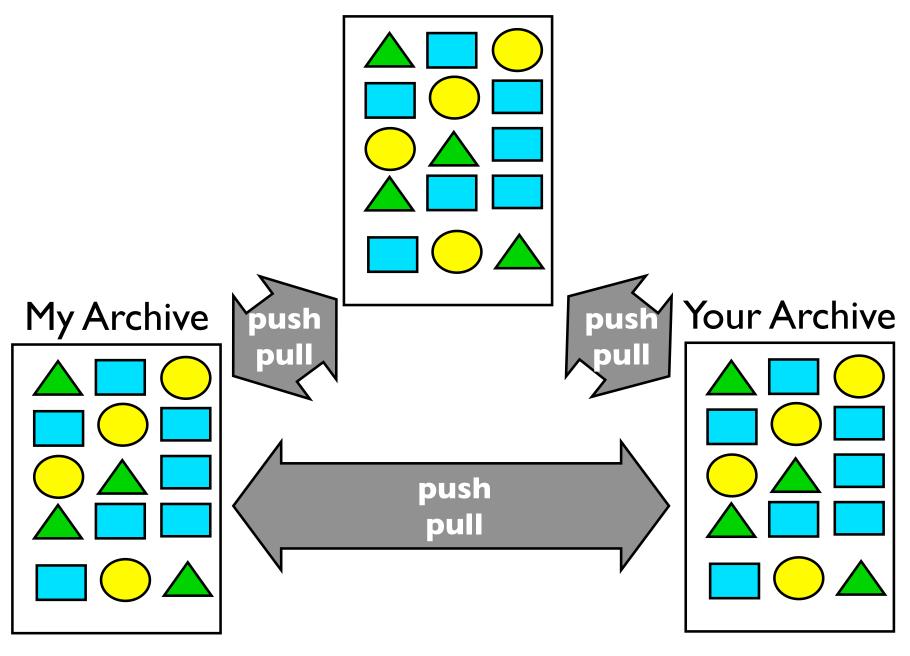
push pull

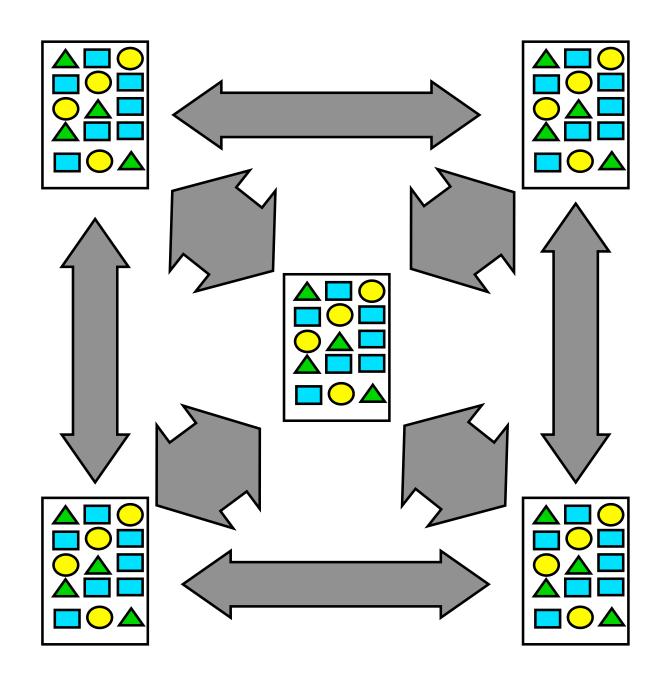
push pull

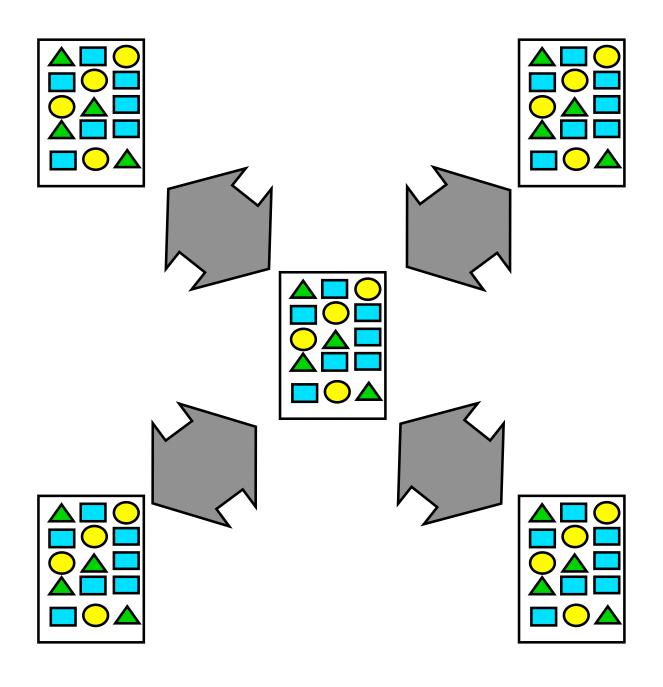
Your Archive

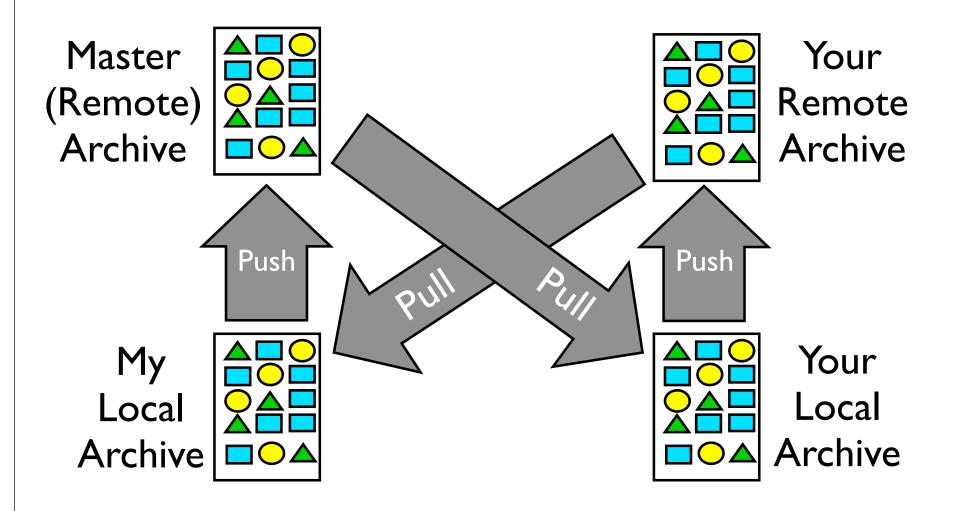


Shared Remote Archive

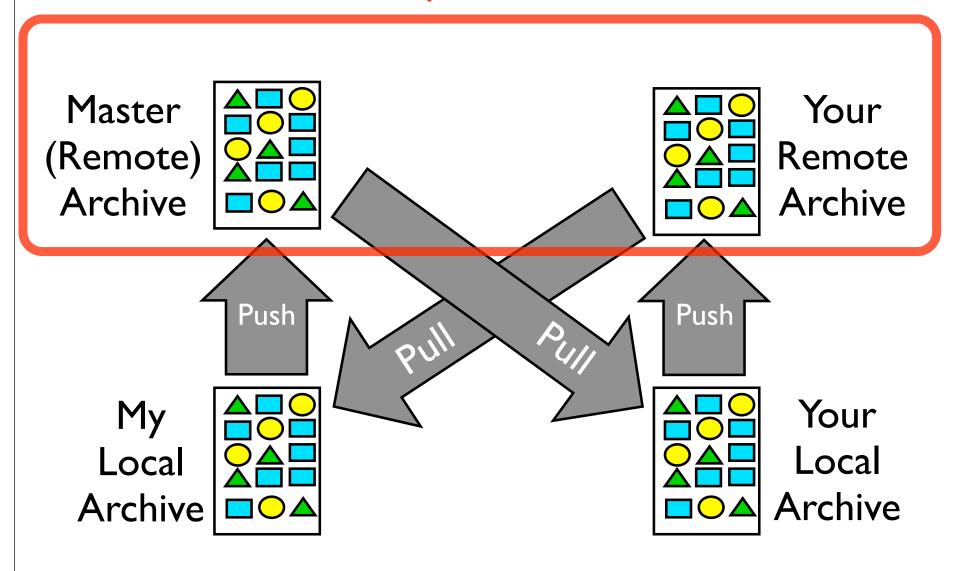


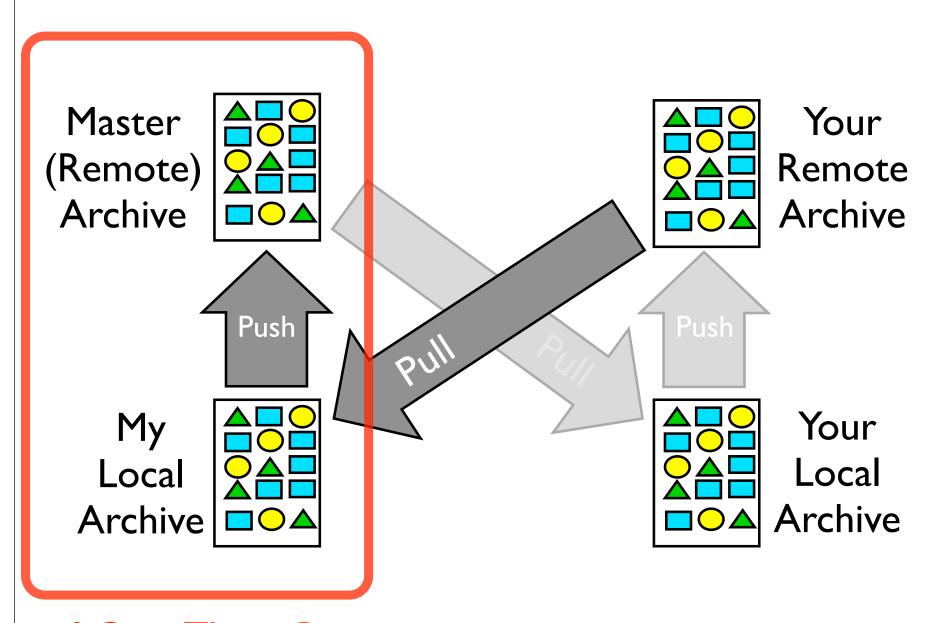




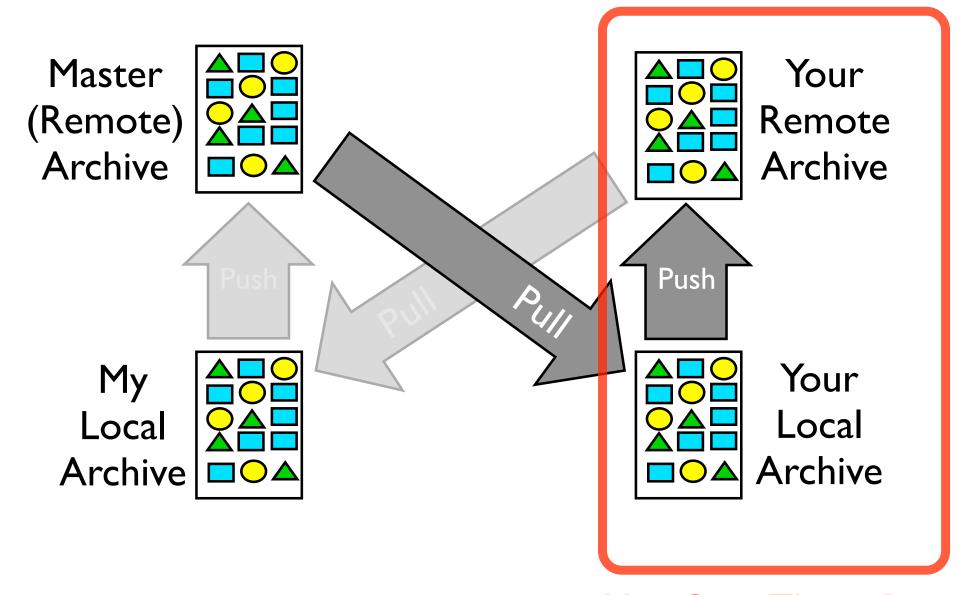


These Repos are Public

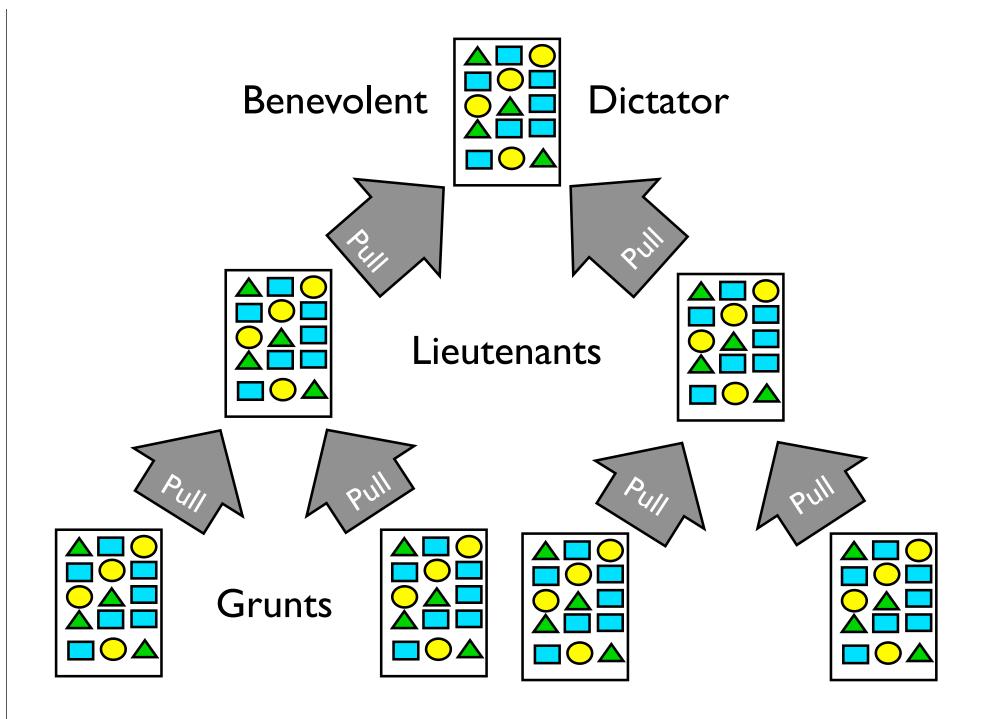


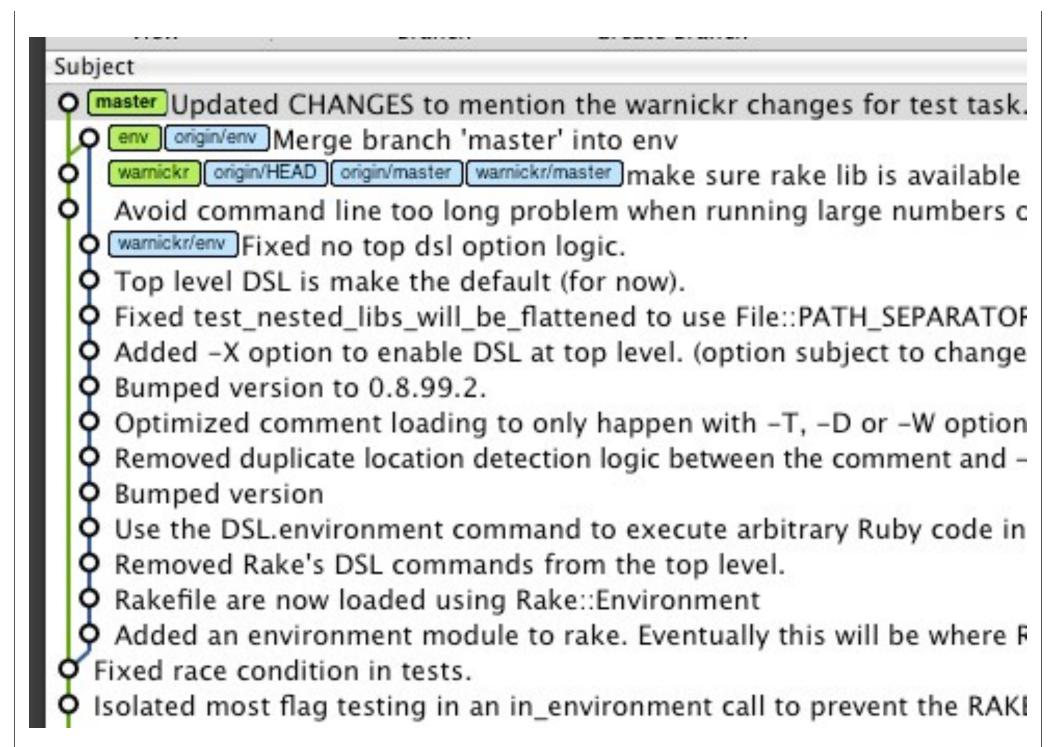


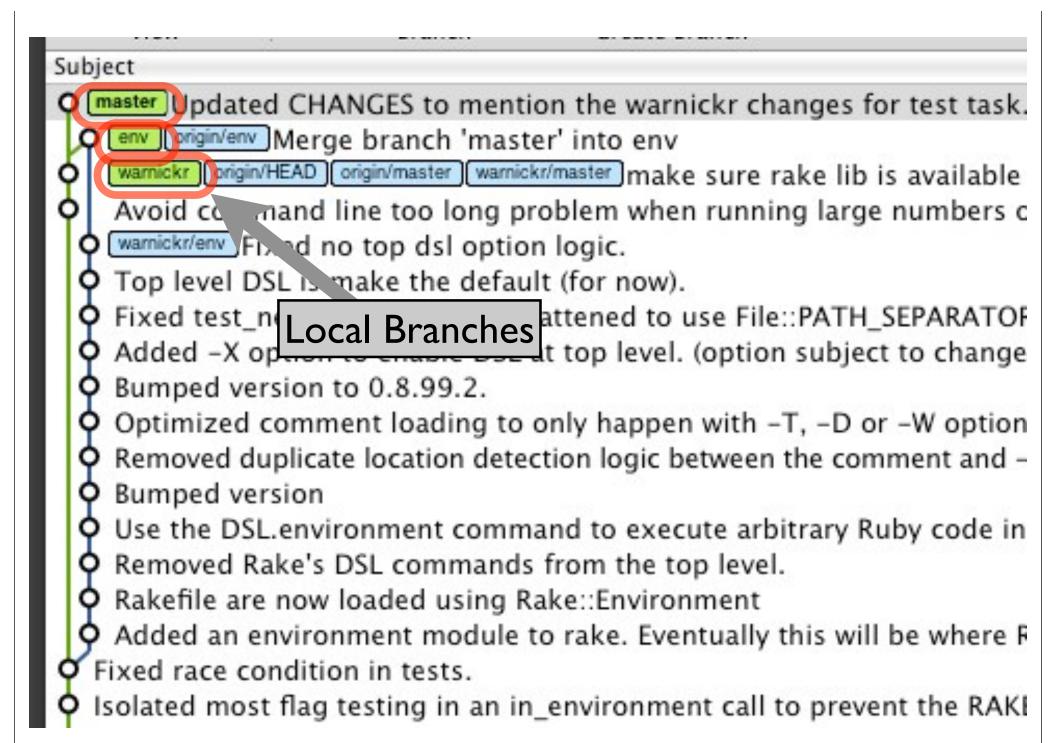
I Own These Repos

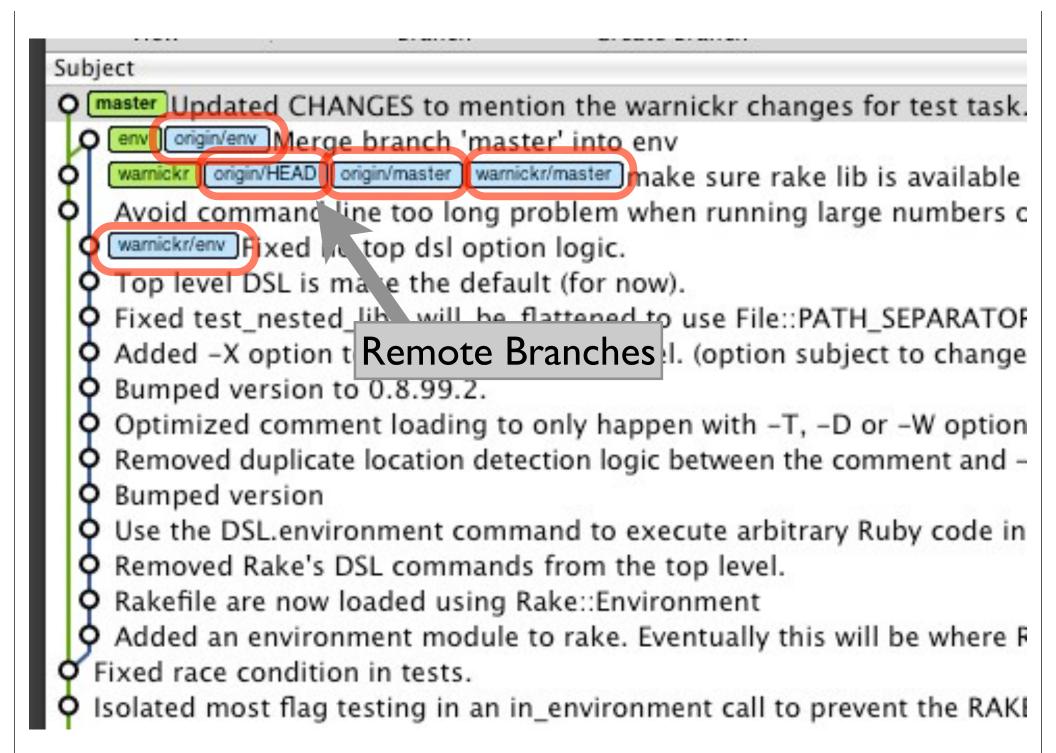


You Own These Repos











CSC Features

- Complete revision history
- Efficient Storage
 - very lightweight project archives
- Guaranteed integrity
- Flexible Sharing
- Cheap Branching
- Easy Tagging

CSC Features

- Complete revision history
- Efficient Storage
 - very lightweight project archives
- Guaranteed integrity
- Flexible Sharing
- Cheap Branching
- Easy Tagging







```
csc == git
snapshot == commit
project archive == repository
manifest == tree (or directory)
```

```
csc snap == git commit -a
csc checkout name == git checkout name
csc branch name == git checkout -b name
csc tag tagname == git tag tagname
```

Differences from Git

138

Csc

- We have only:
 - Working Directory
 - Repository (project archive)

Git

- Git Uses
 - Working Directory
 - Staging Area
 - Repository

Csc

- We have only:
 - Working Directory
 - Repository (project archive)
- File Based

- Git Uses
 - Working Directory
 - Staging Area
 - Repository
- Content Based

Csc

• One level object file storage

- Files stored in .git/objects
 - Uses a 2-level directory
 - SHA1 hash is the directory/file name

Csc

• One level object file storage

• Plain Files

- Files stored in .git/ objects
 - Uses a 2-level directory
 - SHA1 hash is the directory/file name
- Encoded/Compressed files

Csc

• Just the basic features

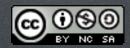
- init/log/status/diff
- reset/rebase
- cherry-pick
- bisect
- ... lots of others



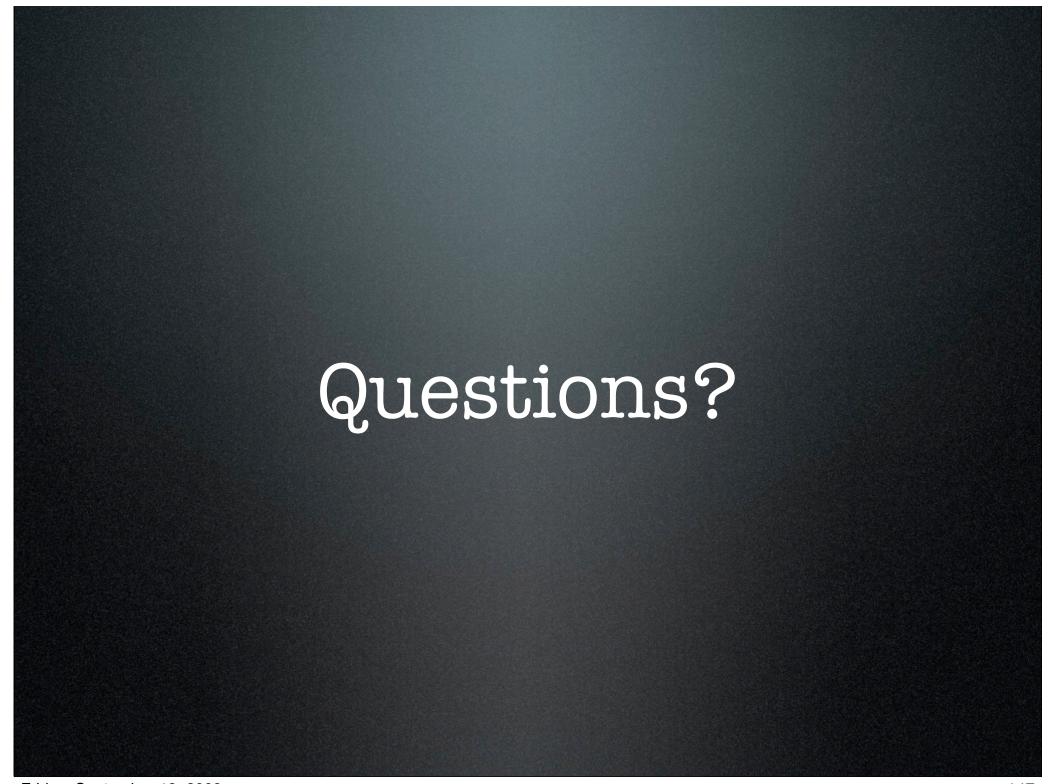
References

- Git from the Bottom Up (John Wiegley)
 - http://ftp.newartisans.com/pub/ git.from.bottom.up.pdf
- Git Internals -- PeepCode PDF (Scott Chacon)
 - http://peepcode.com/products/git-internals-pdf
- Any presentation by Scott Chacon
 - http://github.com/schacon/git-presentations

Source Control for People Who Don't Like Source Control
by Jim Weirich
is licensed under a
Creative Commons Attribution-Noncommercial-Share Alike 3.0
United States License.
Based on a work at github.com.

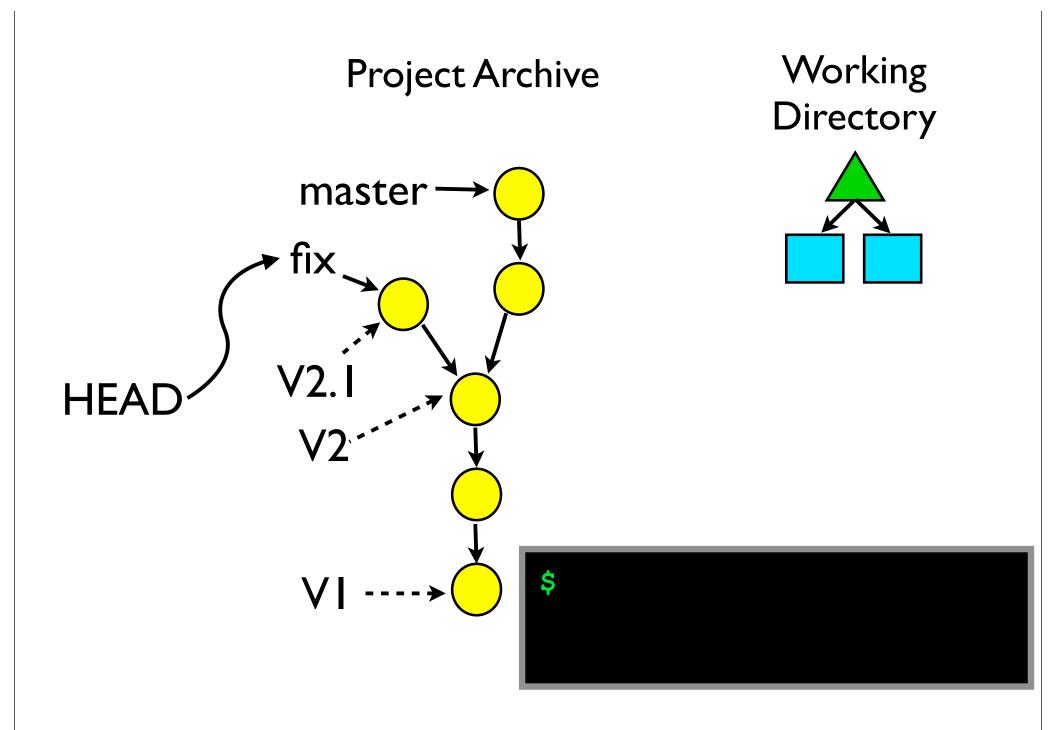


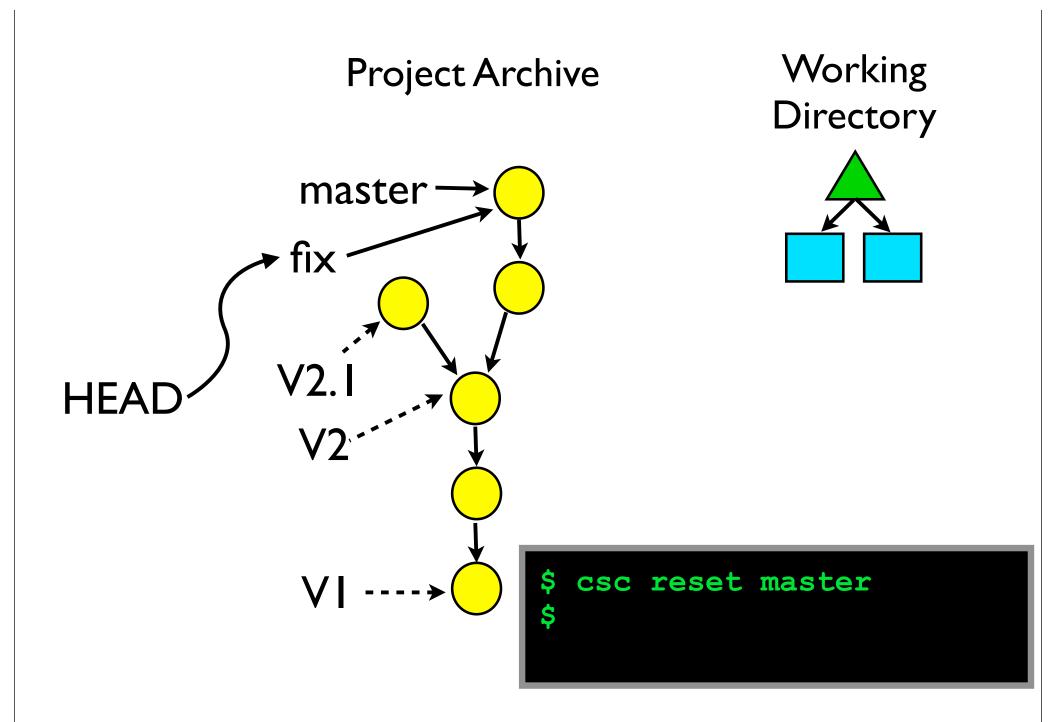
git://github.com/jimweirich/presentation_source_control.git

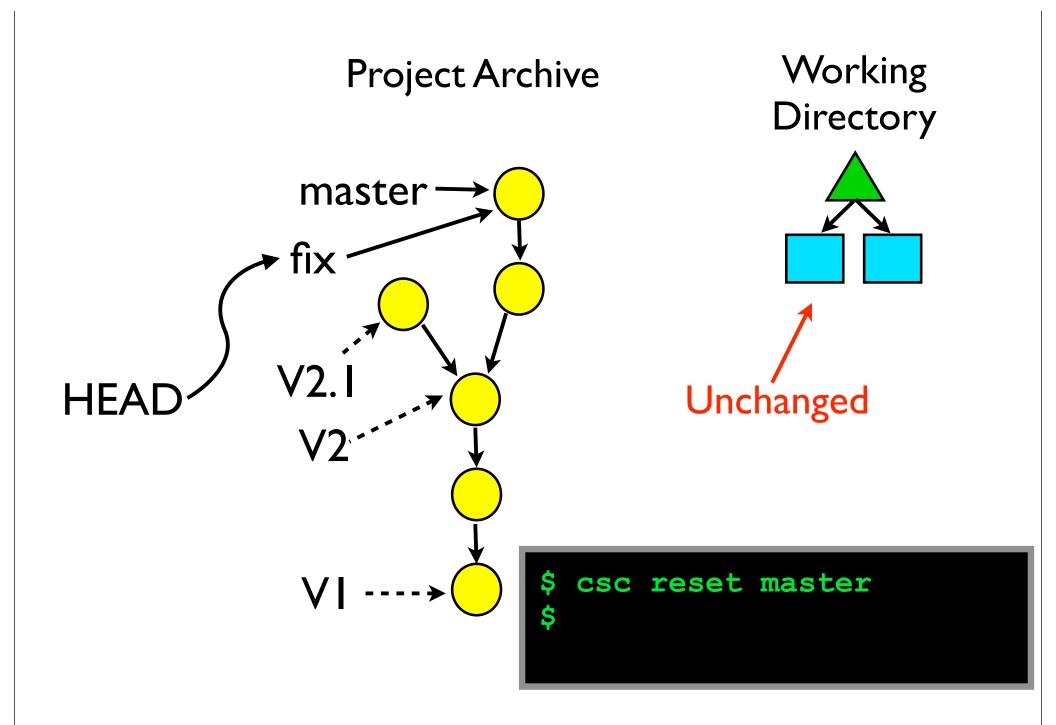


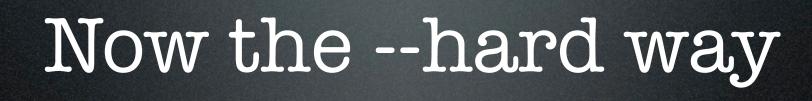
Appendix A -- Reset

148

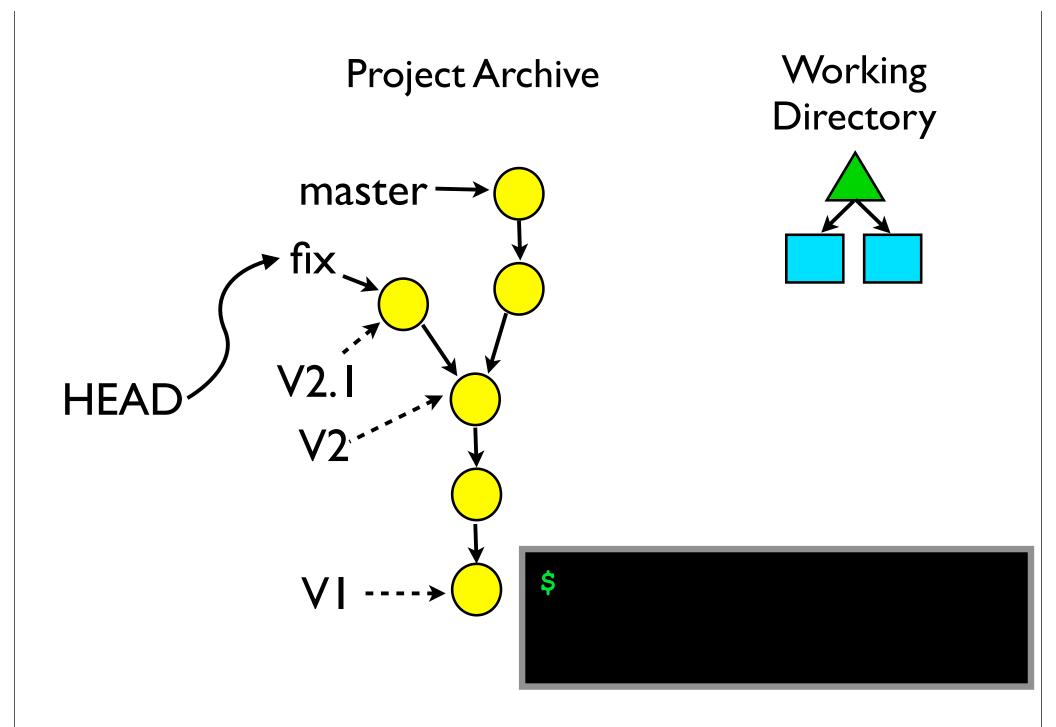


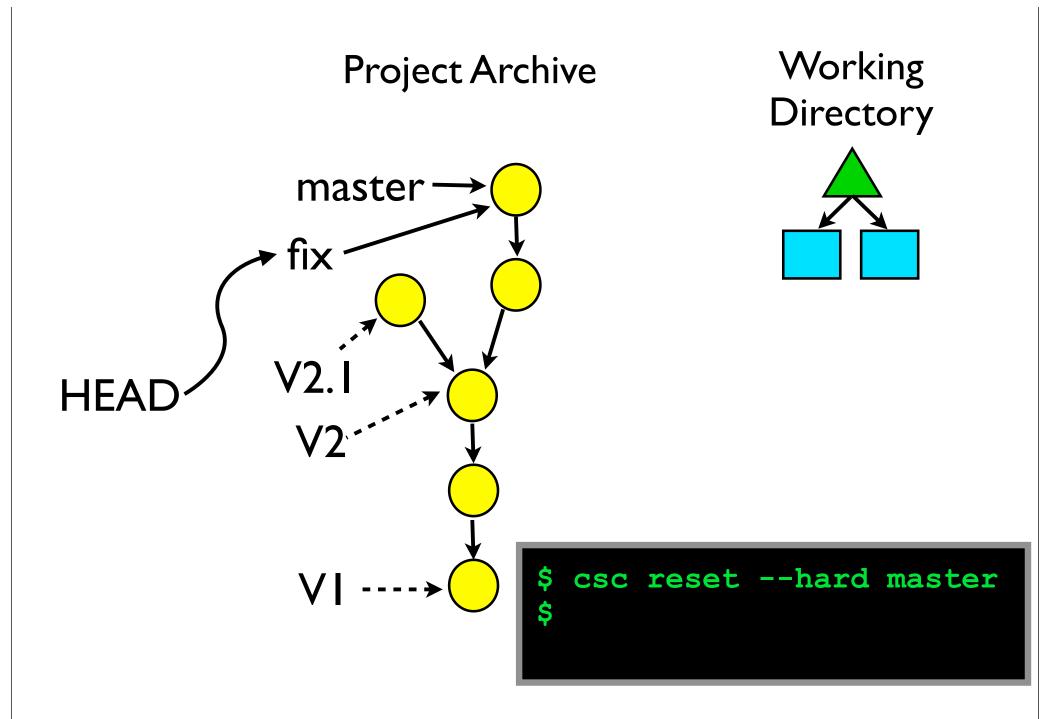


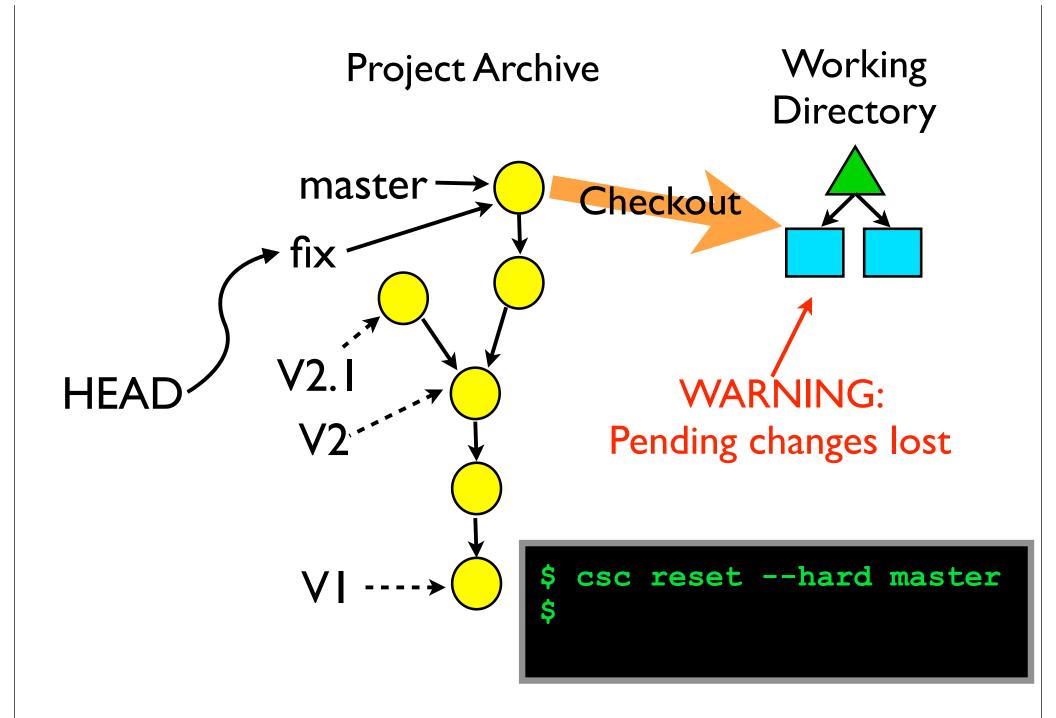




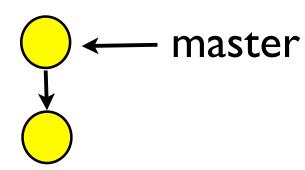
152

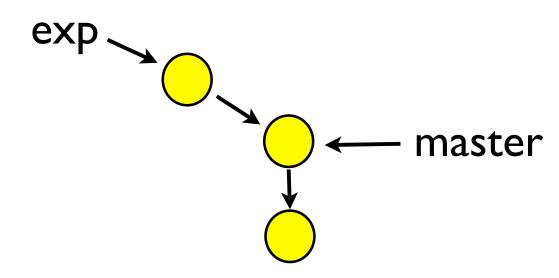


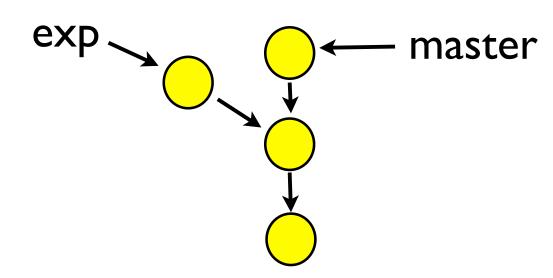


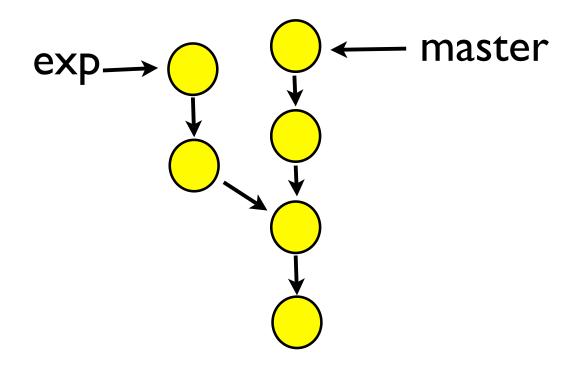


Appendix B -- Rebase



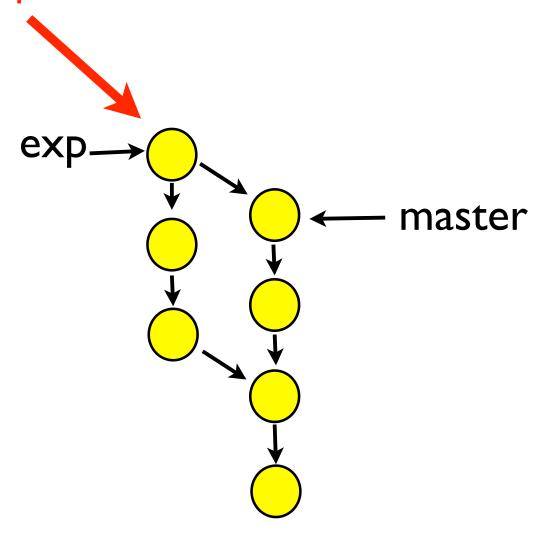




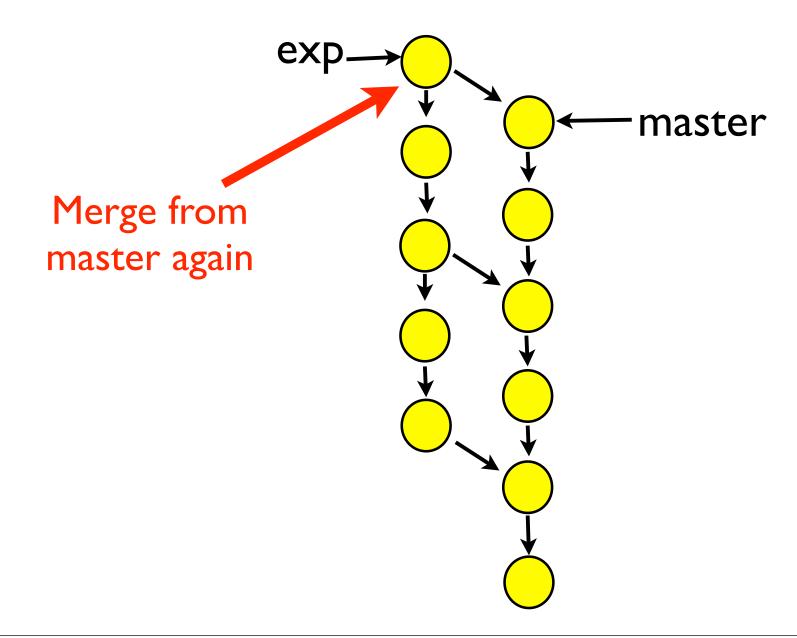


Rebase (merging option)

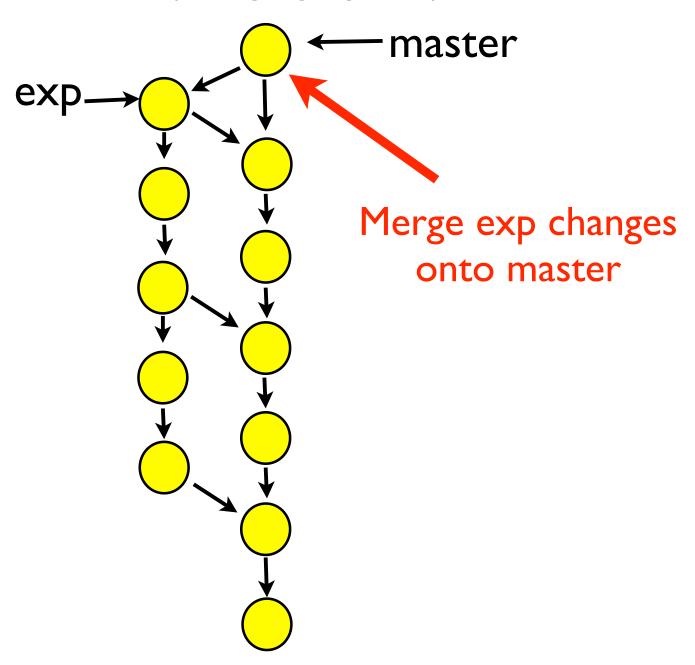
Merge changes on master to exp



Rebase (merging option)

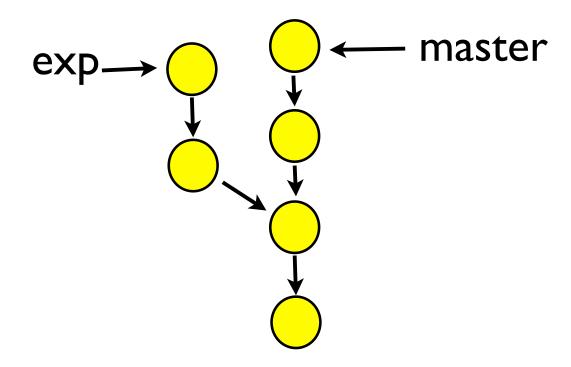


Rebase (merging option)

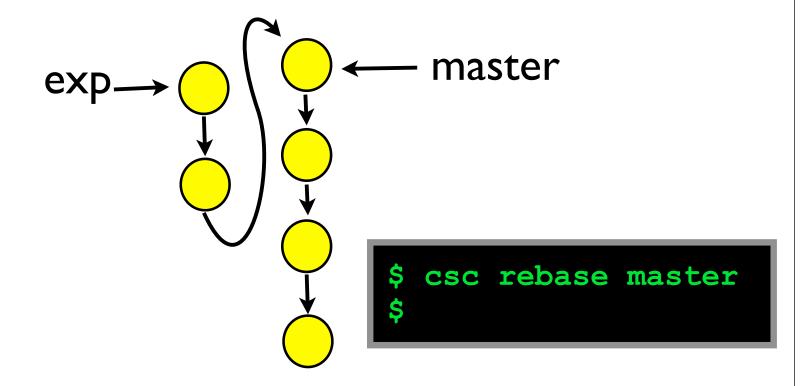


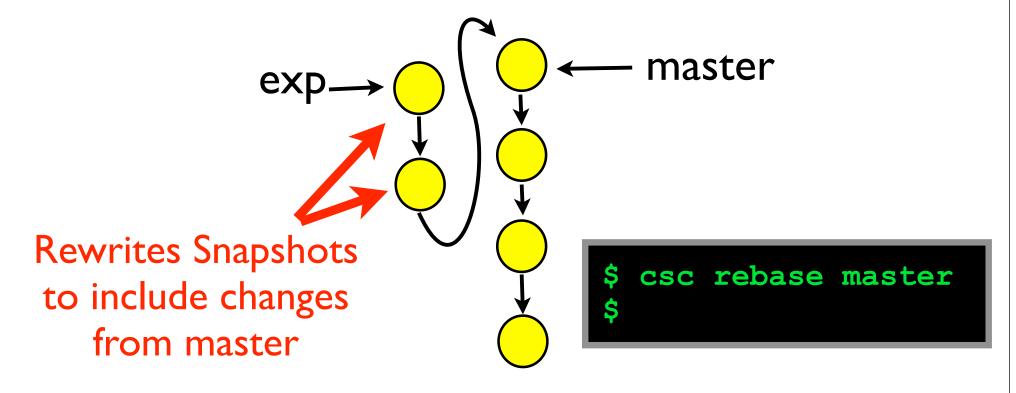
163

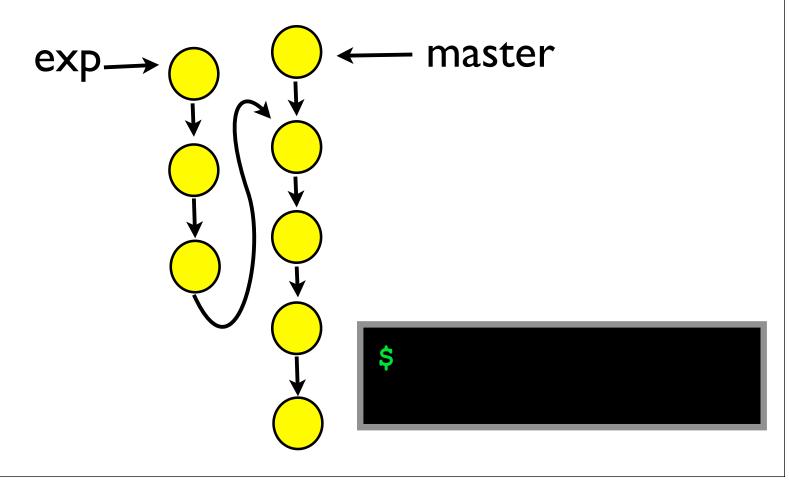


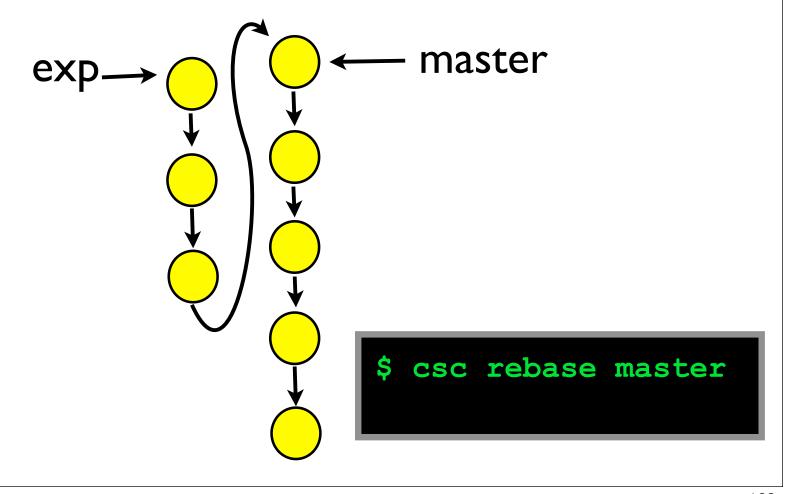


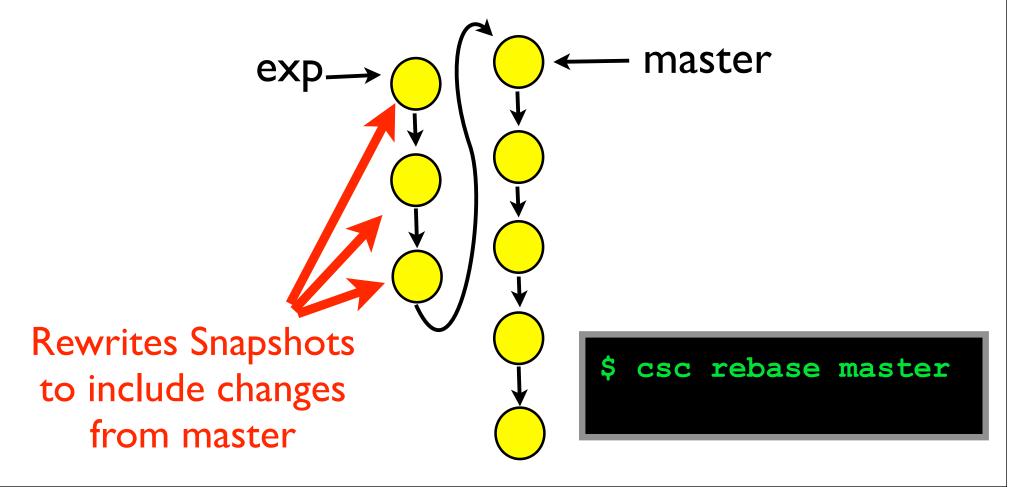
165

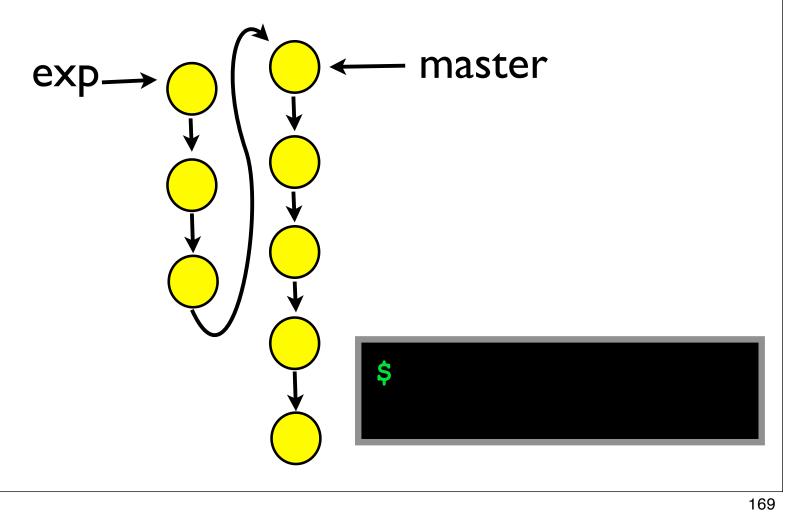


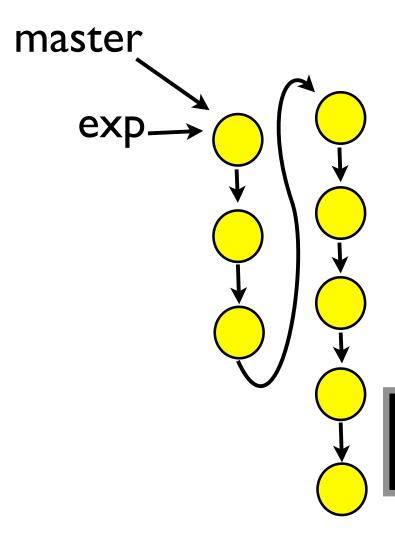












\$ csc checkout master
\$ csc merge exp

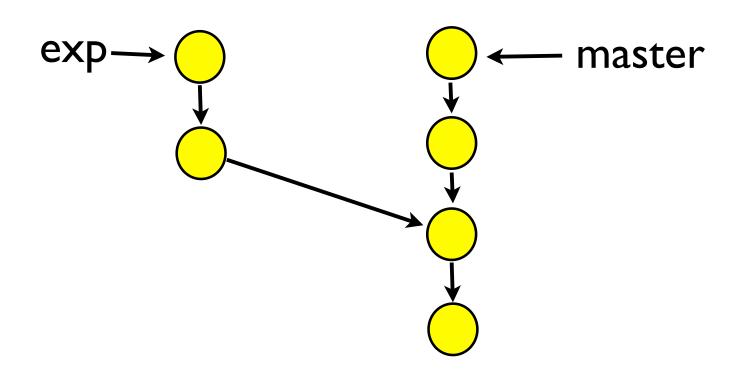
WARNING:

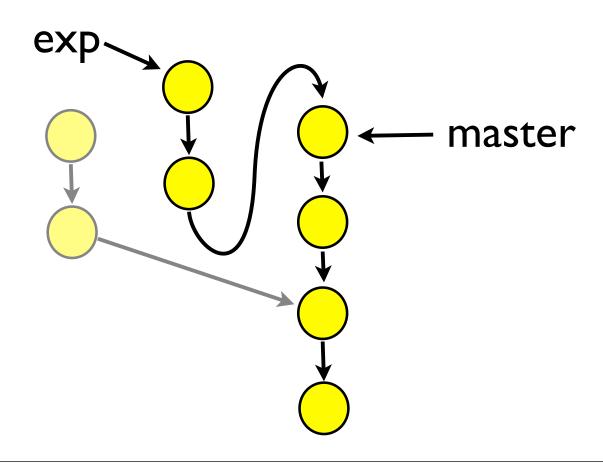
Conflicts may occur during rebase

WAIT!

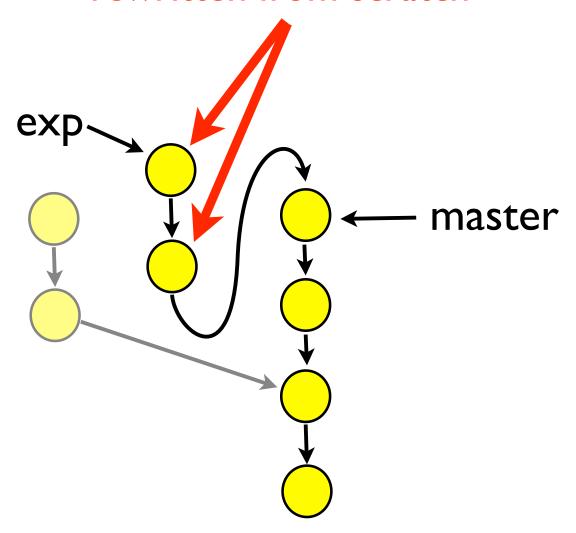
Rebase rewrites history?
Isn't that dangerous?

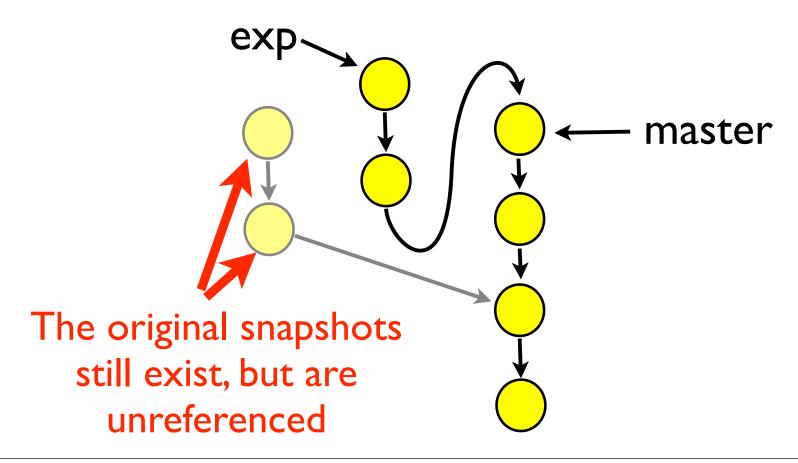


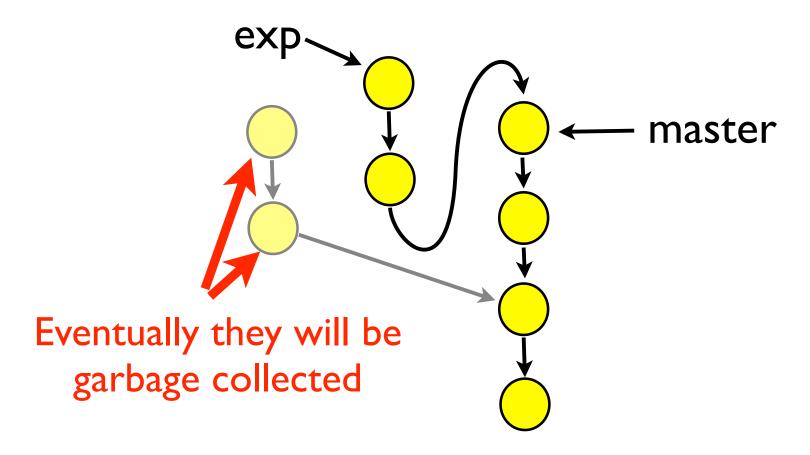




These snapshots are rewritten from scratch







WARNING:

Do **not** rebase shared branches!

178