Source Control for People Who Don't Like Source Control

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



Your Mission

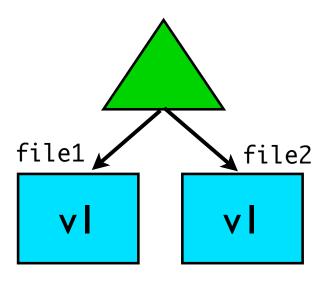
Your Mission

Design and Build A Source Code Control System

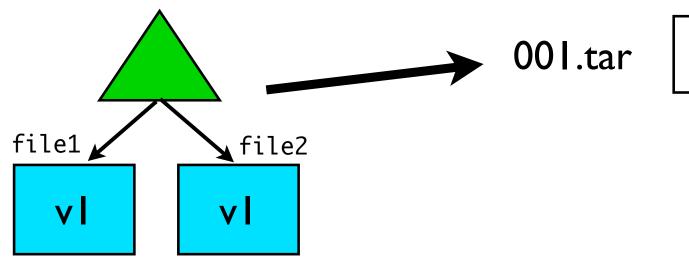
Custom Source Control

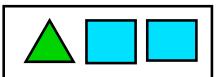
Custom
Source
Control

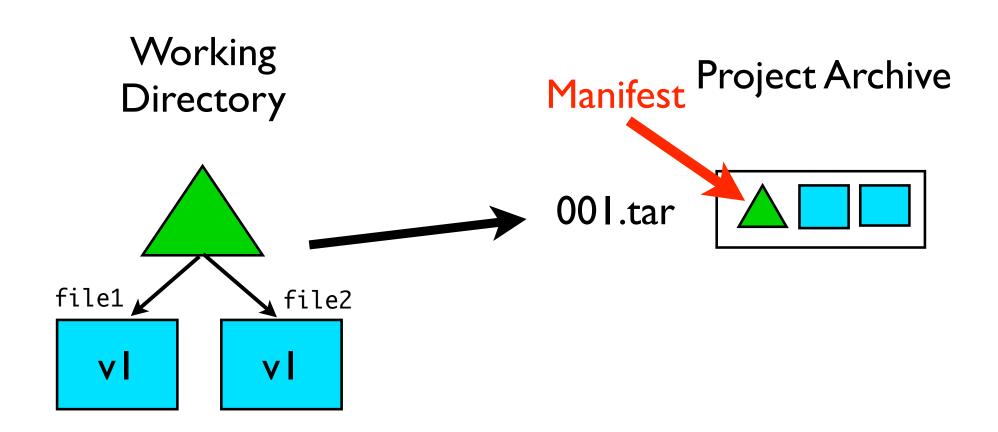
In the Old Days ...

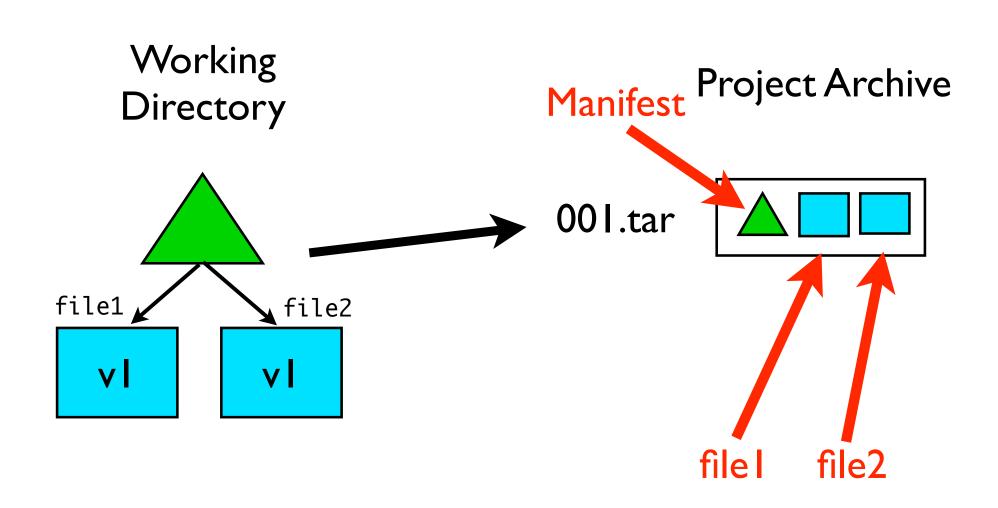


Project Archive

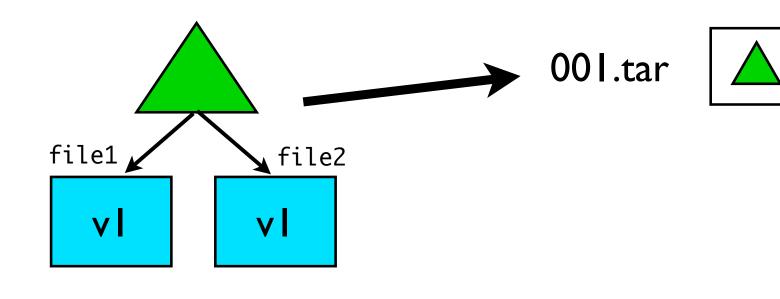




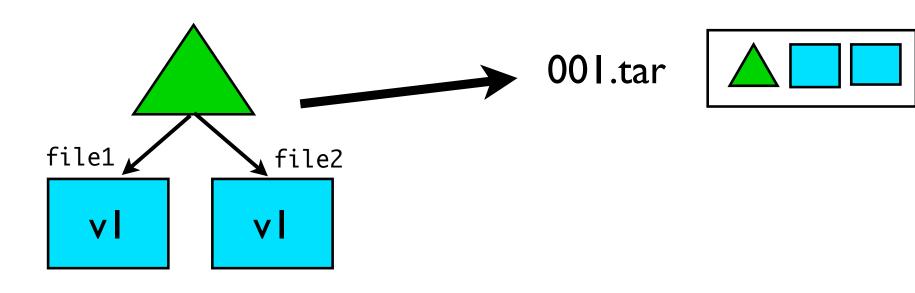




Project Archive



Project Archive



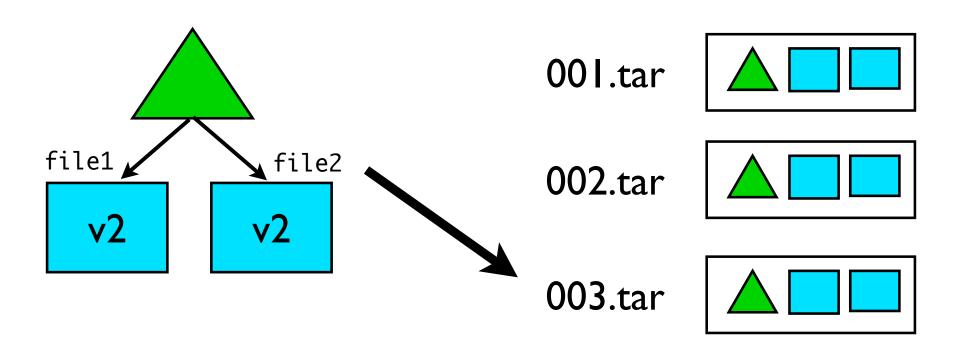
\$ csc snap

Project Archive



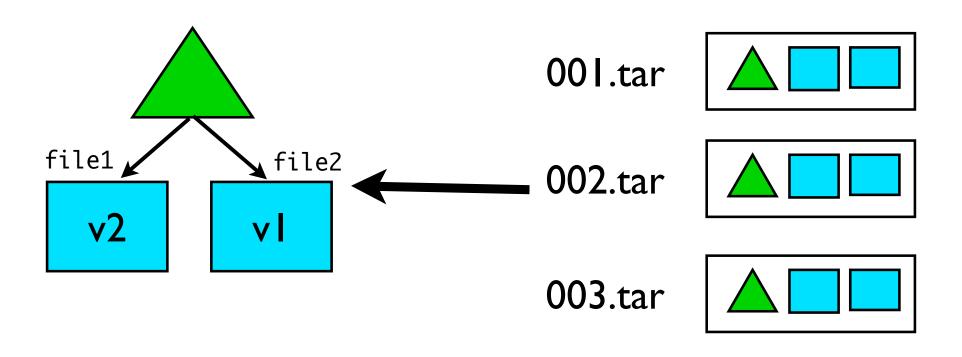
\$ csc snap

Project Archive



\$ csc snap

Project Archive



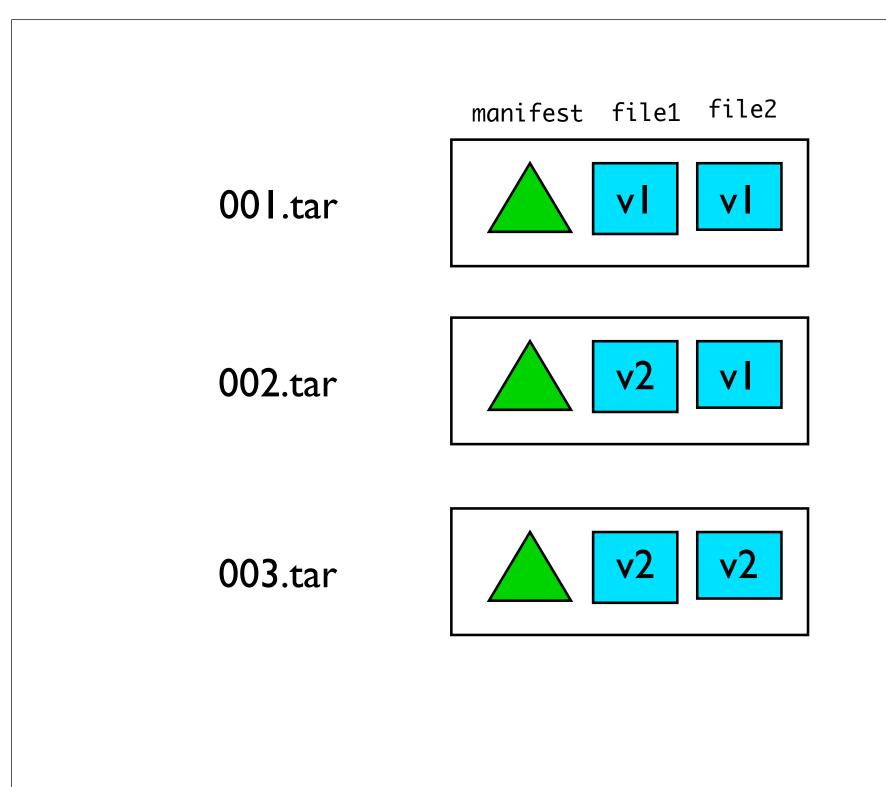
\$ csc checkout 002

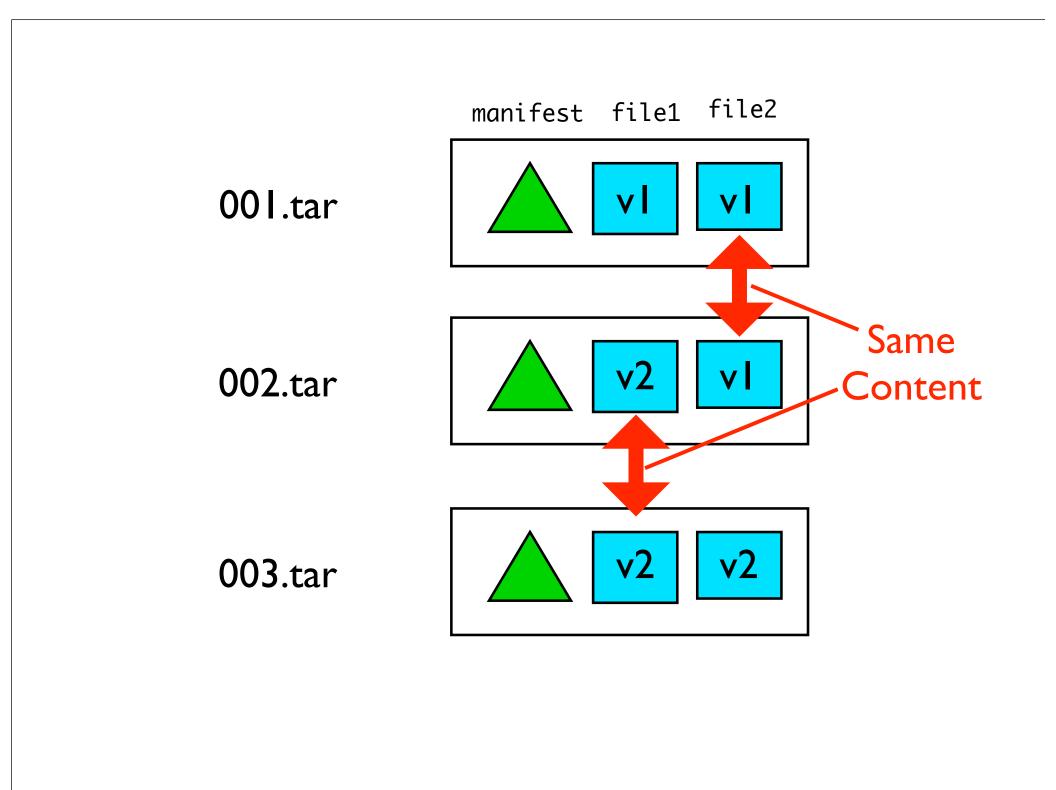
Two Basic Commands

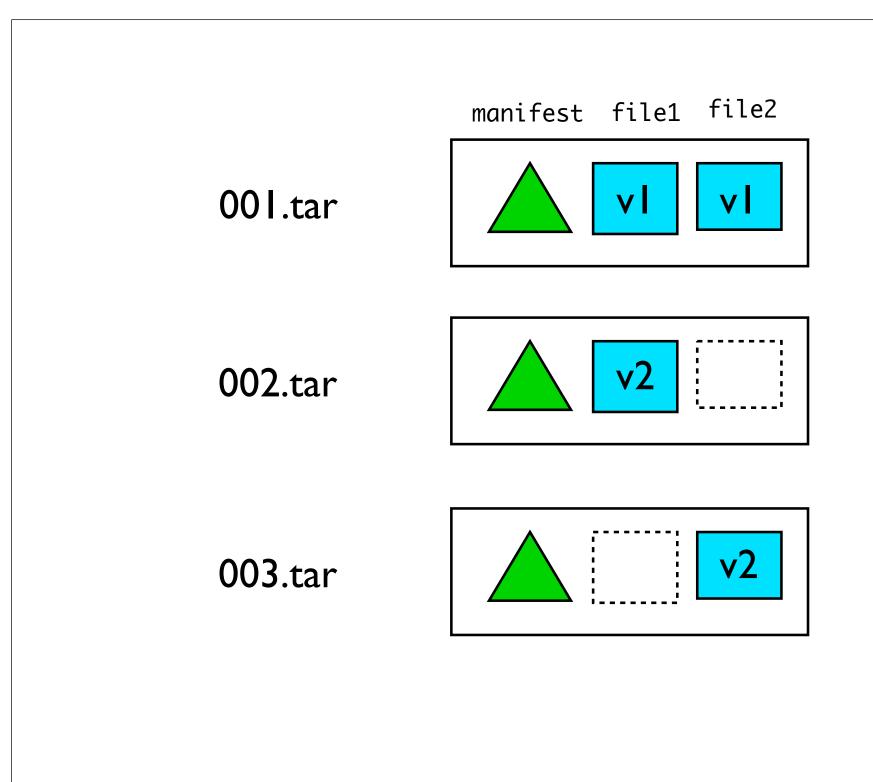
- Snap
- Checkout

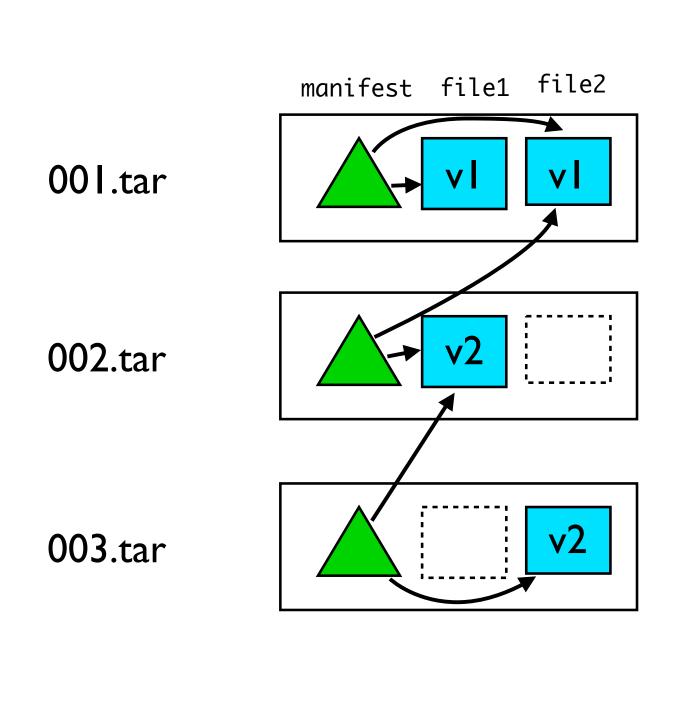
Very Easy/ Very Simple

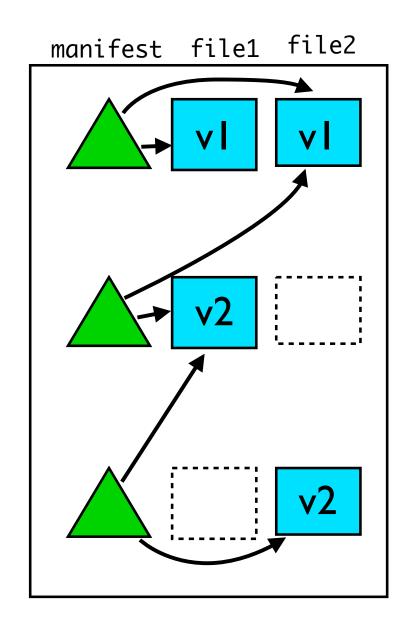
Rather Wasteful





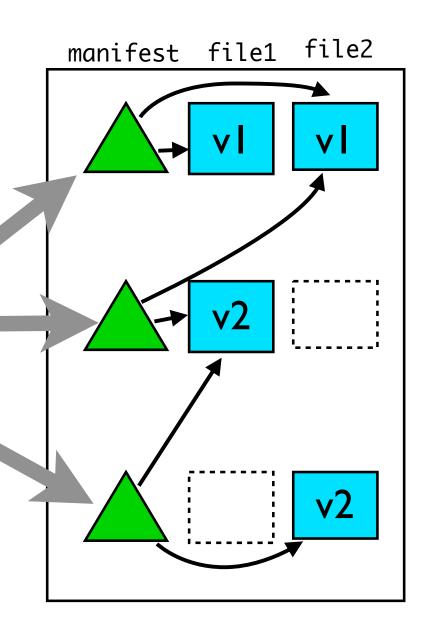






Note:

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



File Name Contents manfest \triangle / (001) file1 | | file1 (V1) file2 | Tile2 (V1) manifest \triangle / manifest \triangle / file1 (V2) file1 file2 (V2) file2 **CONFLICT!**

Note:

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

File Contents Name **△** / (001) aaaa file1 (V1) aaab file2 (V1) aaac aaad aaae 🛕 / aaaf | file1 (V2) file2 (V2) aaag

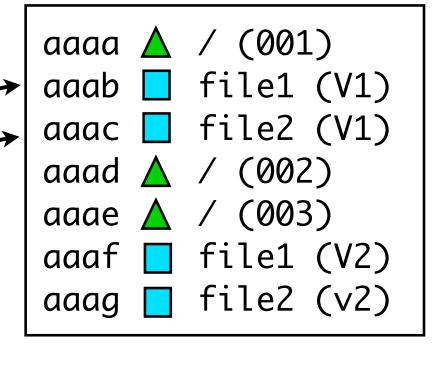
```
$ ls
aaaa aaab aaac aaad aaae aaaf aaag
```

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

aaaa -- / (00 I)



file1 aaab file2 aaac

aaaa -- / (00 I)

Note:

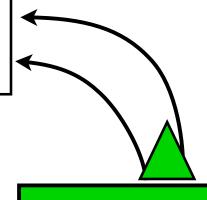
A manifest file is just a text file listing all the files it contains and their IDs.

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

aaad -- / (002)

File Name Contents

File Name Contents



file1 aaaf file2 aaag

aaae -- / (003)

File Name Contents

file1 aaab file2 aaac

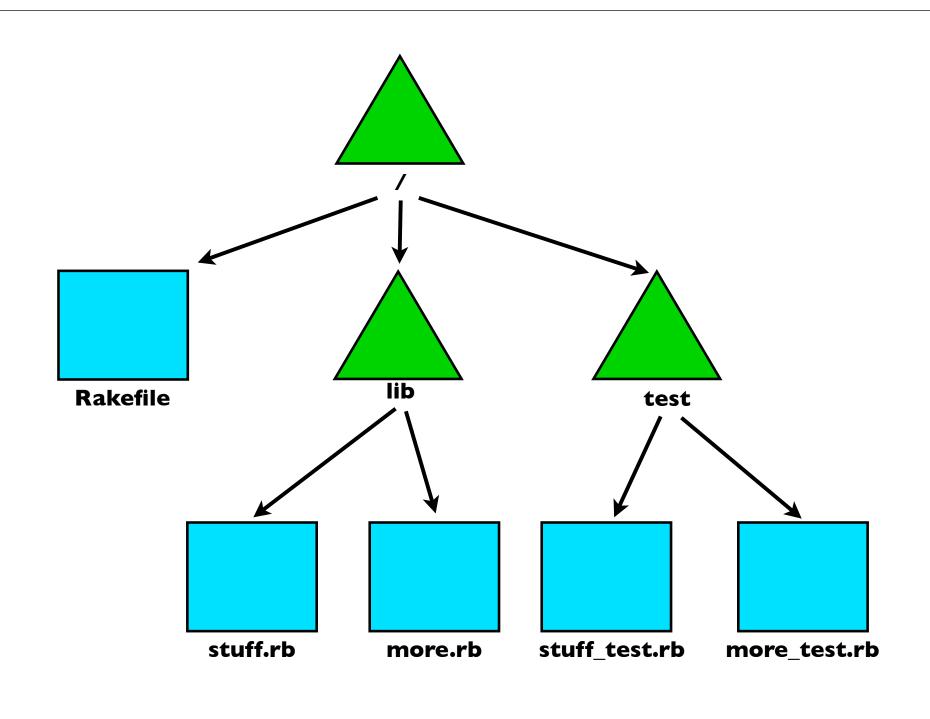
aaaa -- / (00 I)

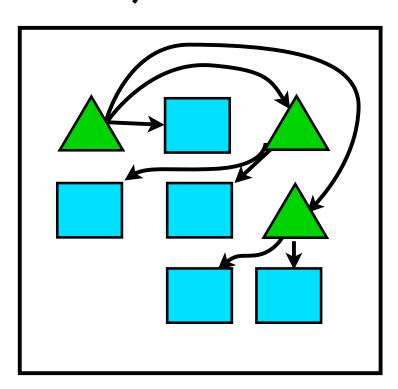
file1 aaaf file2 aaac

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

Root Manifest for Snapshot

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

Author & EMail

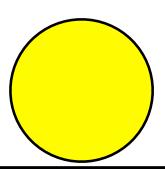
tree daae

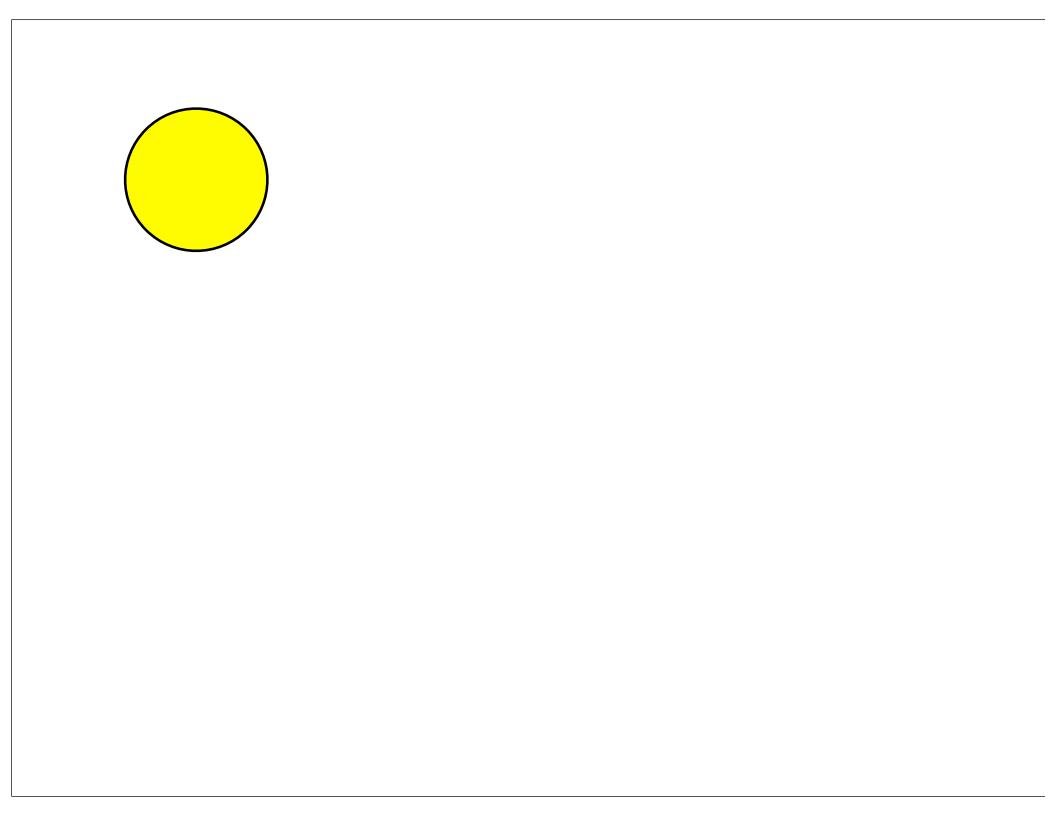
author Jim <<u>me@email.com</u>> 1246996320 -0400

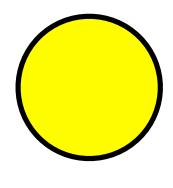


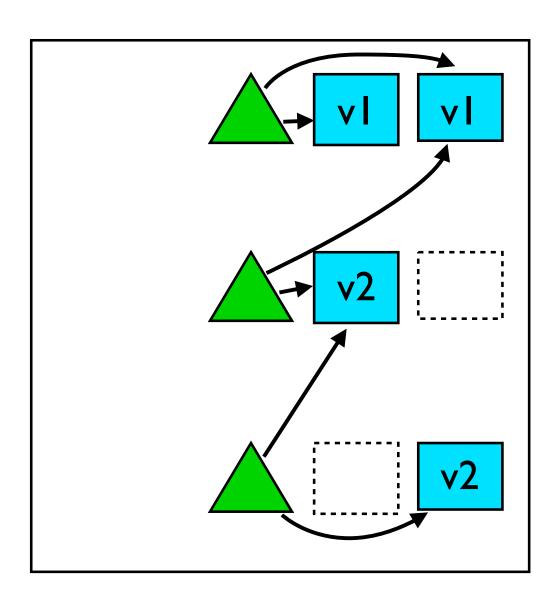
Date/Time (seconds since 1970)

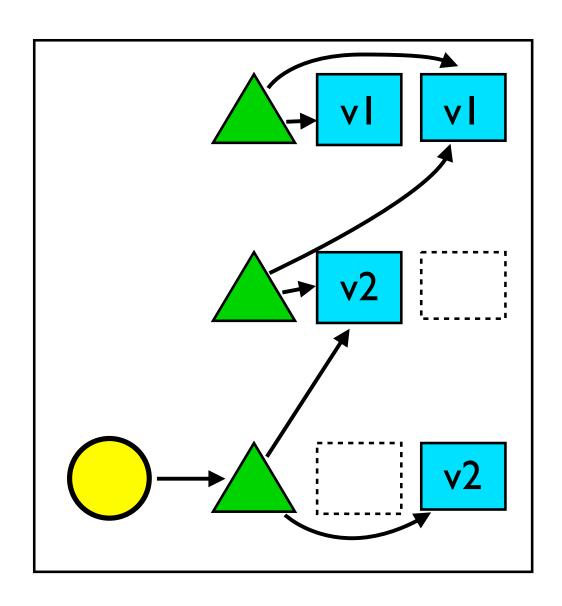


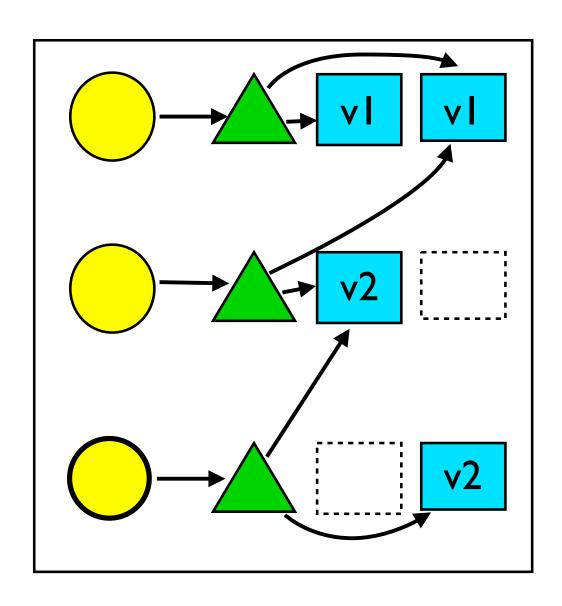


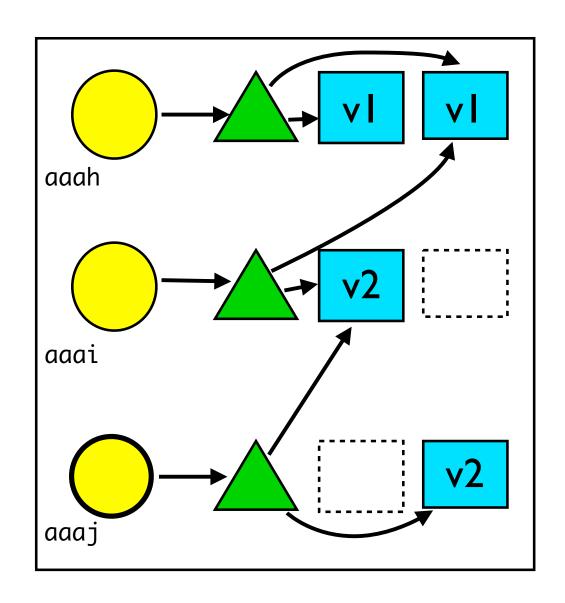




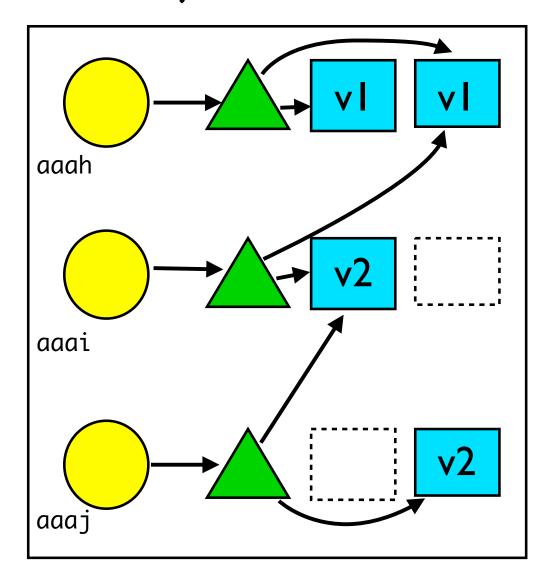


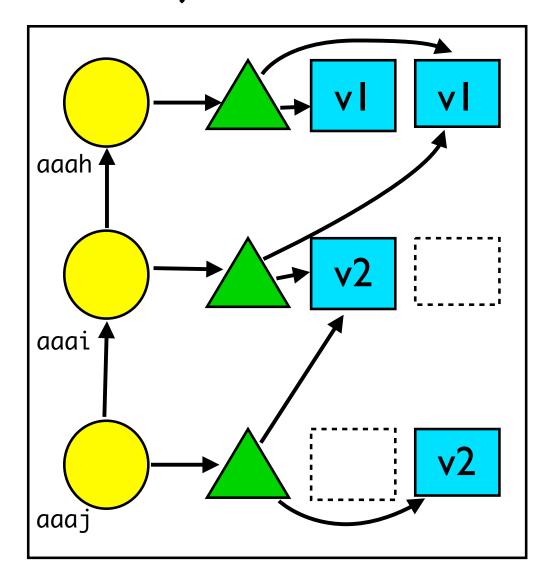


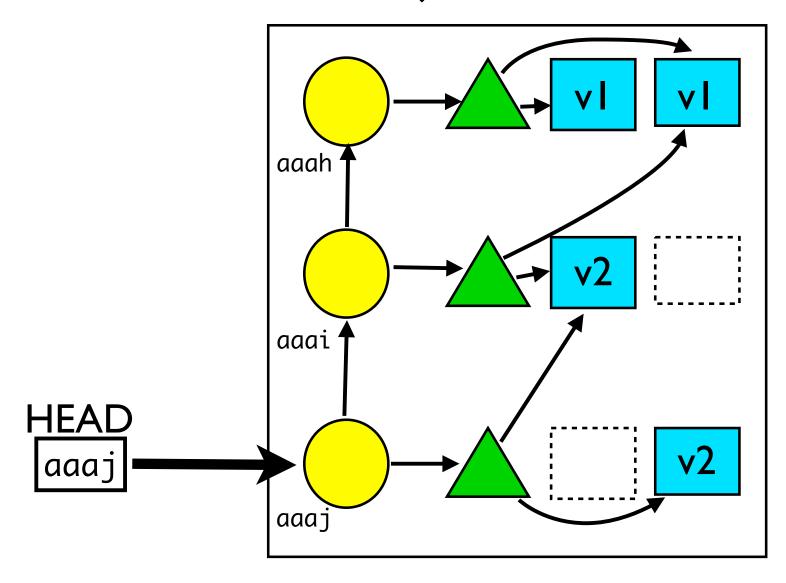




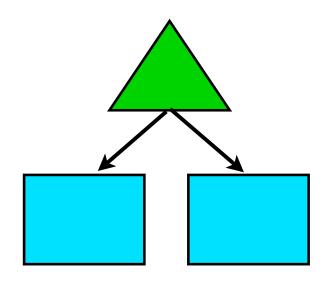
tree aaae
parent aaai
author Jim <me@email.com> 1246996320 -0400







Summary So Far ...

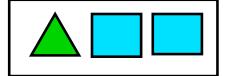


Project Archive

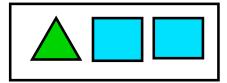
001.tar

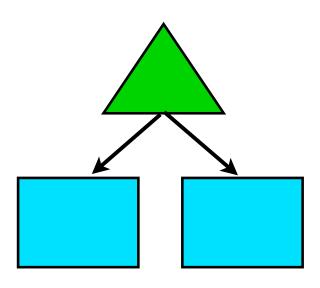


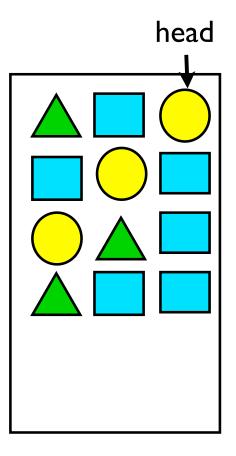
002.tar



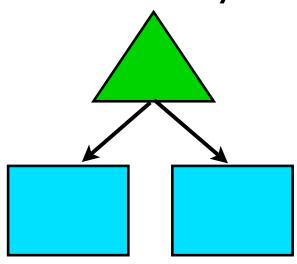
003.tar



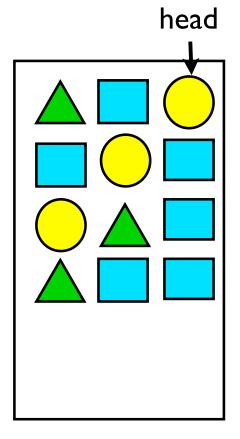


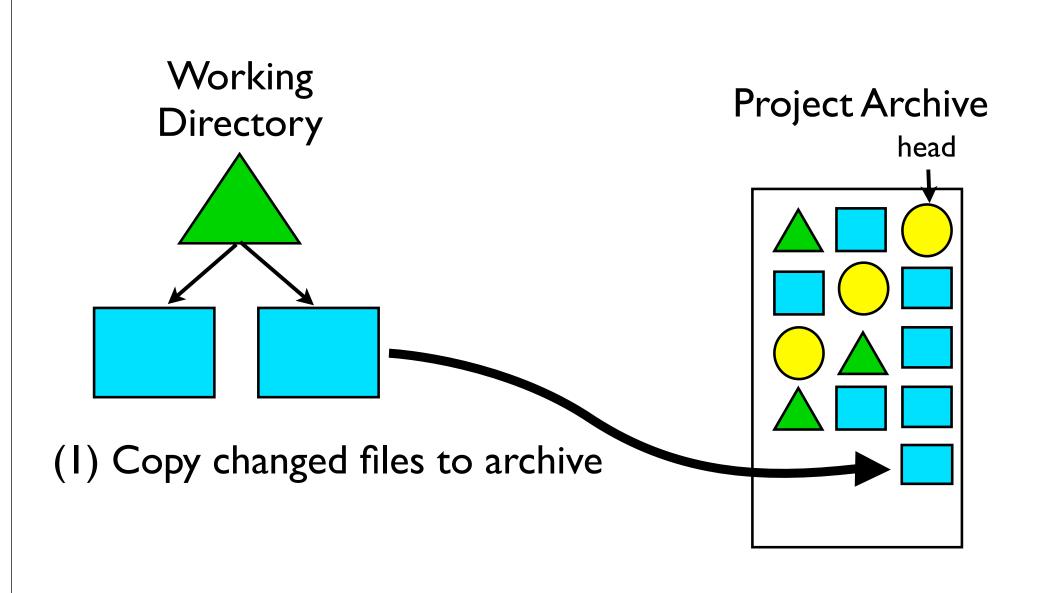


How do we create a snapshot now?



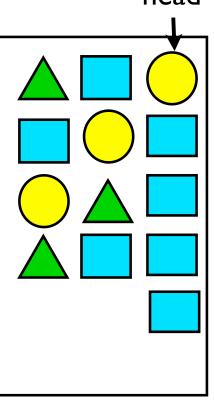
(I) Copy changed files to archive

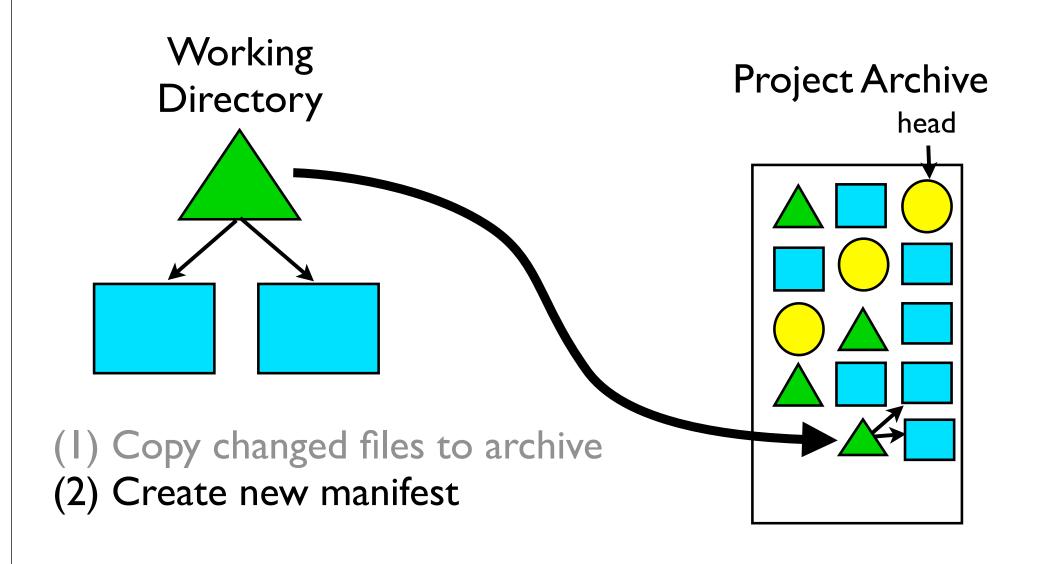




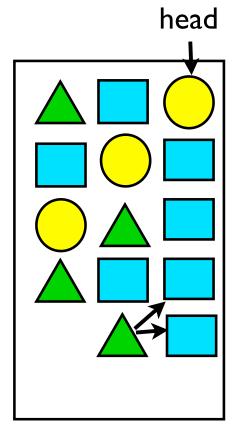
- (I) Copy changed files to archive
- (2) Create new manifest

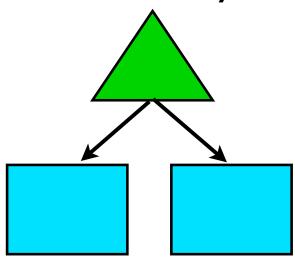
Project Archive head



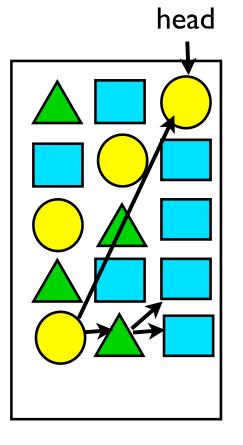


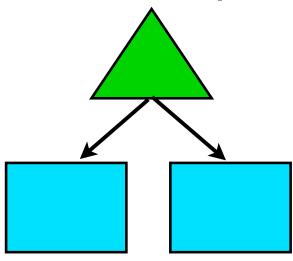
- (I) Copy changed files to archive
- (2) Create new manifest
- (3) Create new snapshot record



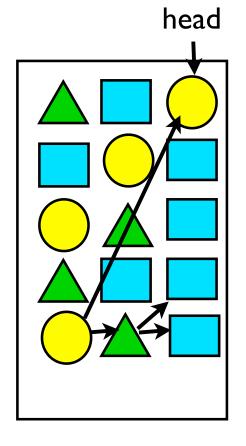


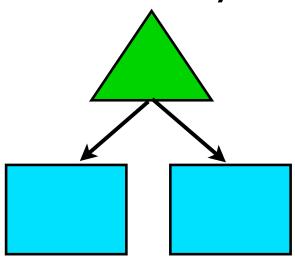
- (I) Copy changed files to archive
- (2) Create new manifest
- (3) Create new snapshot record



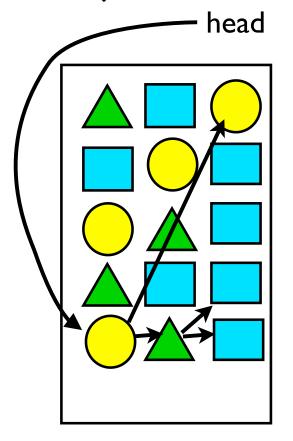


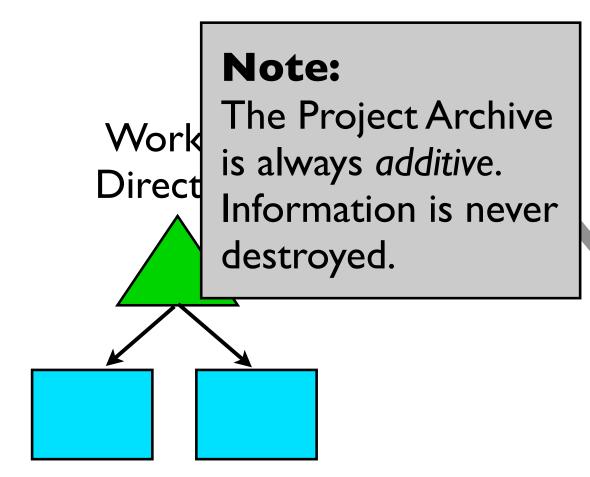
- (I) Copy changed files to archive
- (2) Create new manifest
- (3) Create new snapshot record
- (4) Readjust HEAD





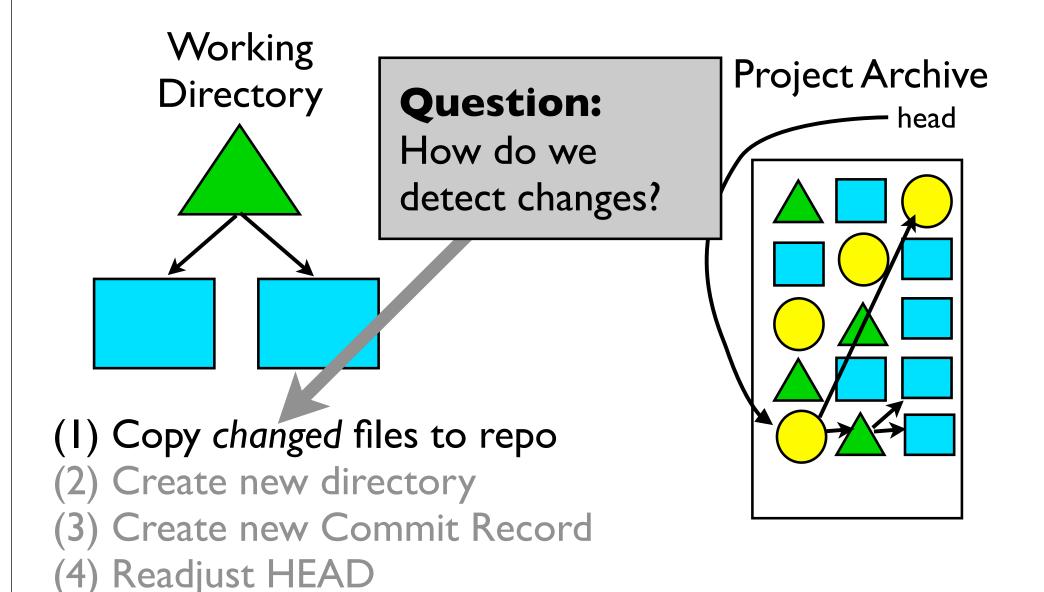
- (I) Copy changed files to archive
- (2) Create new manifest
- (3) Create new snapshot record
- (4) Readjust HEAD

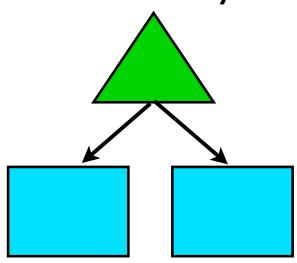




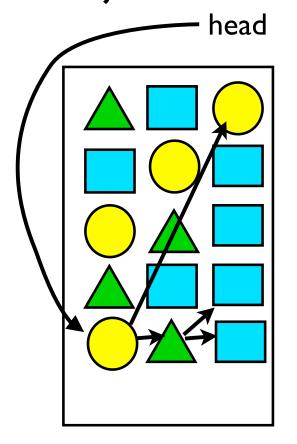
head

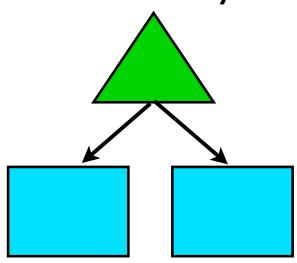
- (I) Copy changed files to archive
- (2) Create new manifest
- (3) Create new snapshot record
- (4) Readjust HEAD



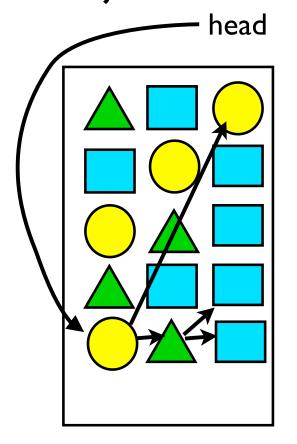


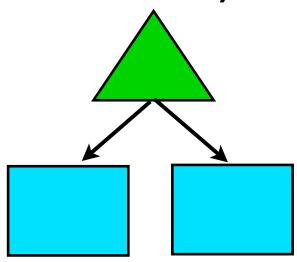
- (I) Copy changed files to repo
- (2) Create new directory
- (3) Create new Commit Record
- (4) Readjust HEAD



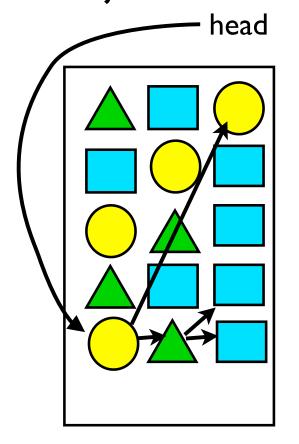


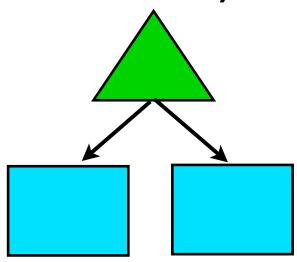
- (I) Copy changed files to repo
- (2) Create new directory
- (3) Create new Commit Record
- (4) Readjust HEAD



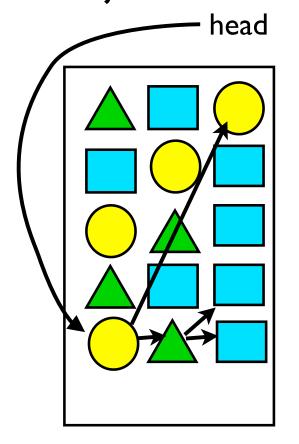


- (I) Copy changed files to repo
- (2) Create new directory
- (3) Create new Commit Record
- (4) Readjust HEAD





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

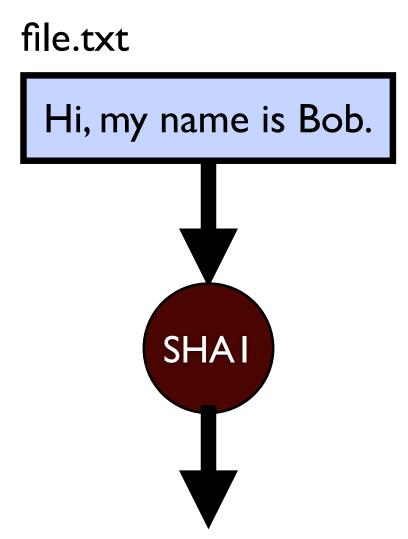


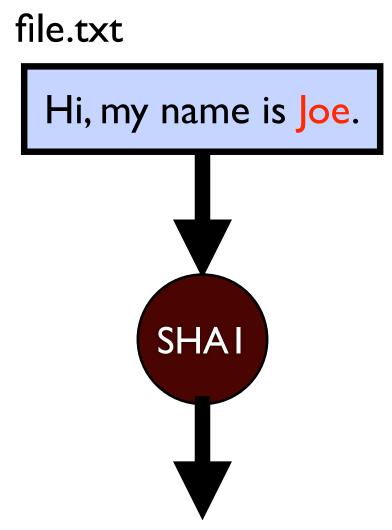
Detecting Duplication

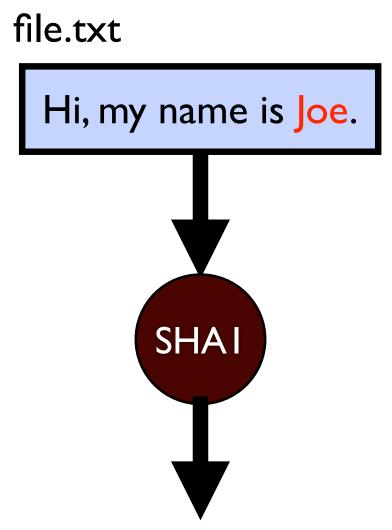
file.txt

Hi, my name is Bob.

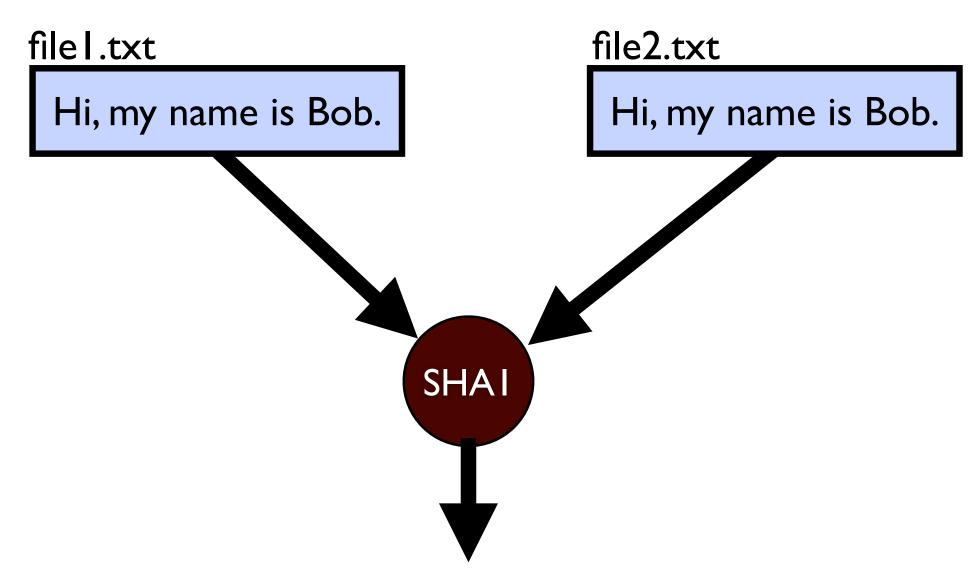
file.txt Hi, my name is Bob. SHAI







"53f2509e877de3cad85c547a0c2a22b43c5d81eb"





Hi, my name is Bob.

file2.txt

Hi, my name is Bob.

Identical Content Products Identical Hashes

"55dc96ede241511e35f1973957bbd51f37f3762b"

File Names in the Archive

File Names in the Archive

- Arbitrary
- Unique

Idea:

Use the SHA1 Hash for the File Name

aaaa 🔾 aaab aaac 🔲 aaad 🛕 aaae 🔵 aaaf 🔲 aaag 🛕 aaah 🔲 aaai

```
75e3dee3503e7ad5...

1586ad8975e3dee3...

f6da34481586ad89...

f99b89d6f6da3448...

503d6749f99b89d6...

2b436957503d6749...

944ff522944ff522...

6c024e25944ff522...

503e7ad56c024e25...
```

HEAD

75e3dee3503e7ad5...

```
75e3dee3503e7ad5...

1586ad8975e3dee3...

f6da34481586ad89...

f99b89d6f6da3448...

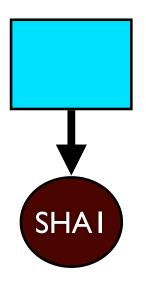
503d6749f99b89d6...

2b436957503d6749...

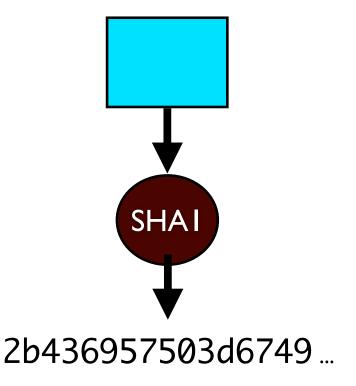
944ff522944ff522...

6c024e25944ff522...

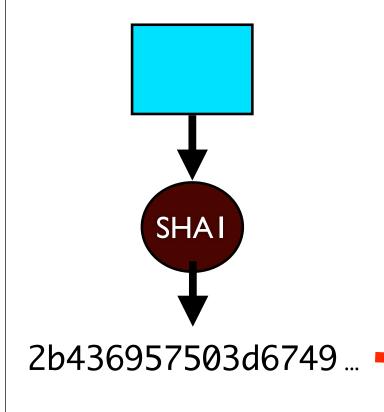
503e7ad56c024e25...
```



```
75e3dee3503e7ad5...
1586ad8975e3dee3...
f6da34481586ad89...
f99b89d6f6da3448... 🛕
503d6749f99b89d6...
2b436957503d6749...
944ff522944ff522... A
6c024e25944ff522...
503e7ad56c024e25...
```



```
75e3dee3503e7ad5...
1586ad8975e3dee3...
f6da34481586ad89...
f99b89d6f6da3448... 🛕
503d6749f99b89d6...
2b436957503d6749...
944ff522944ff522... A
6c024e25944ff522...
503e7ad56c024e25...
```



Match!

Project Archive

75e3dee3503e7ad5...

1586ad8975e3dee3...

f6da34481586ad89...

f99b89d6f6da3448...

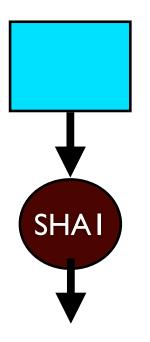
503d6749f99b89d6...

2b436957503d6749...

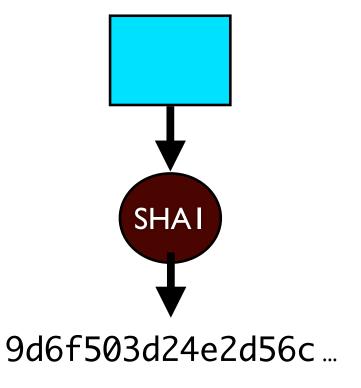
944ff522944ff522...

6c024e25944ff522...

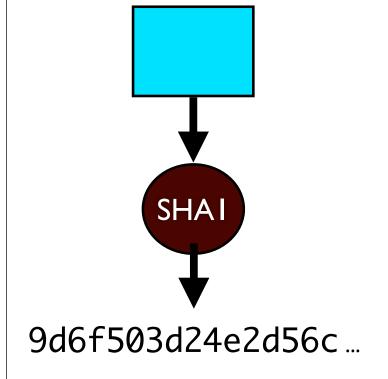
503e7ad56c024e25...



```
75e3dee3503e7ad5...
1586ad8975e3dee3...
f6da34481586ad89...
f99b89d6f6da3448... 🛕
503d6749f99b89d6...
2b436957503d6749...
944ff522944ff522... A
6c024e25944ff522...
503e7ad56c024e25...
```



```
75e3dee3503e7ad5...
1586ad8975e3dee3...
f6da34481586ad89...
f99b89d6f6da3448... 🛕
503d6749f99b89d6...
2b436957503d6749...
944ff522944ff522... A
6c024e25944ff522...
503e7ad56c024e25...
```



No Match!

```
75e3dee3503e7ad5...

1586ad8975e3dee3...

f6da34481586ad89...

f99b89d6f6da3448...

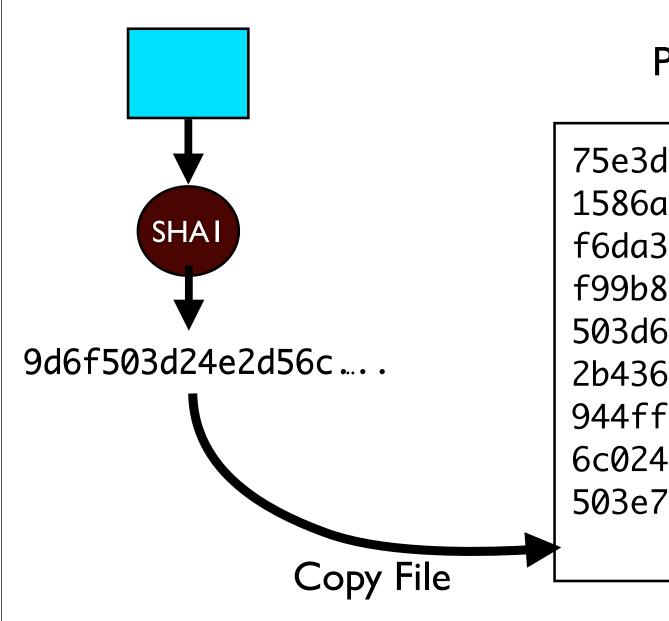
503d6749f99b89d6...

2b436957503d6749...

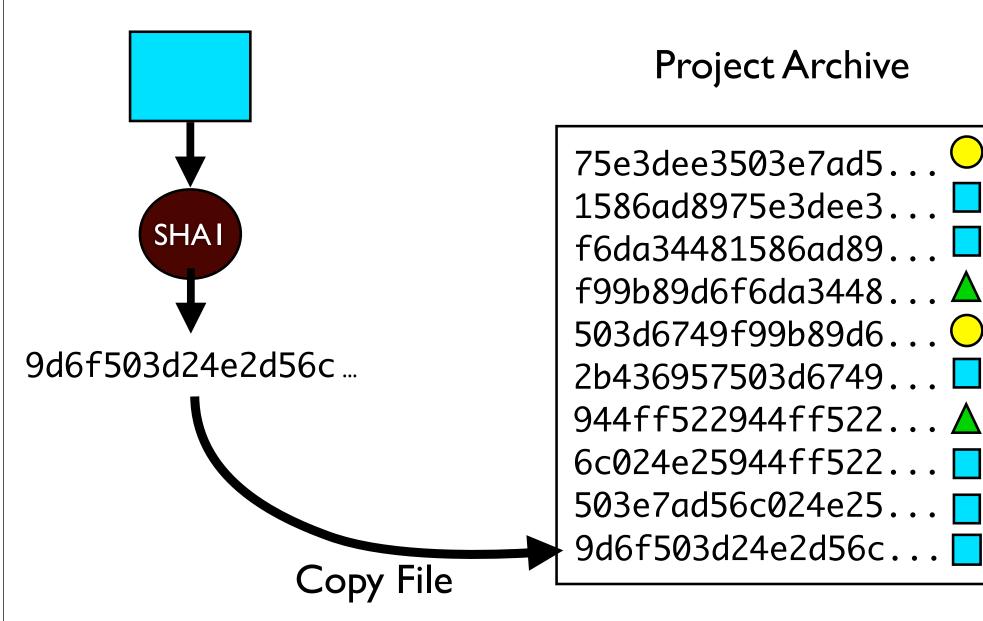
944ff522944ff522...

6c024e25944ff522...

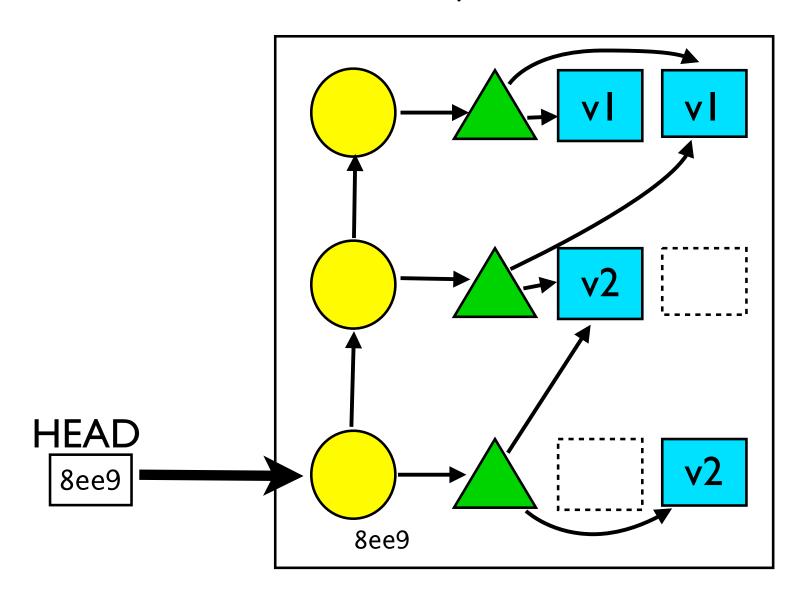
503e7ad56c024e25...
```

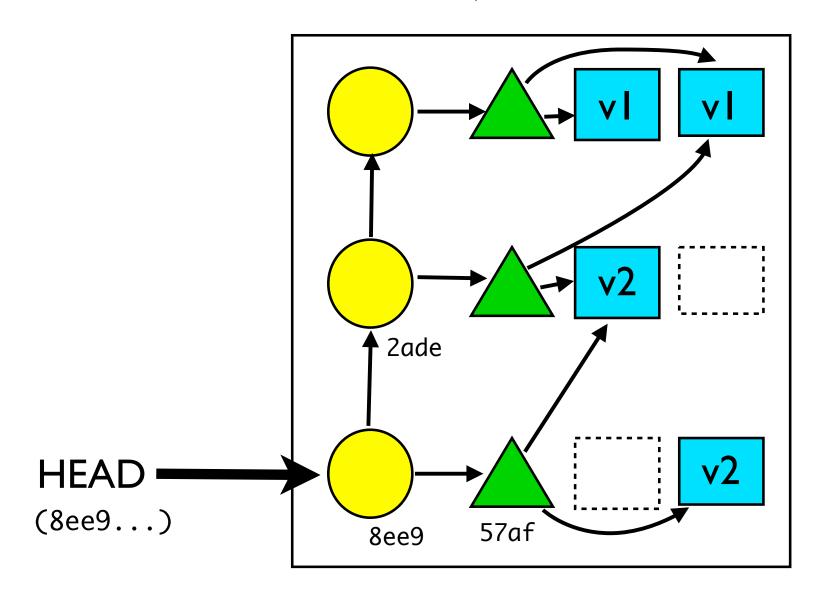


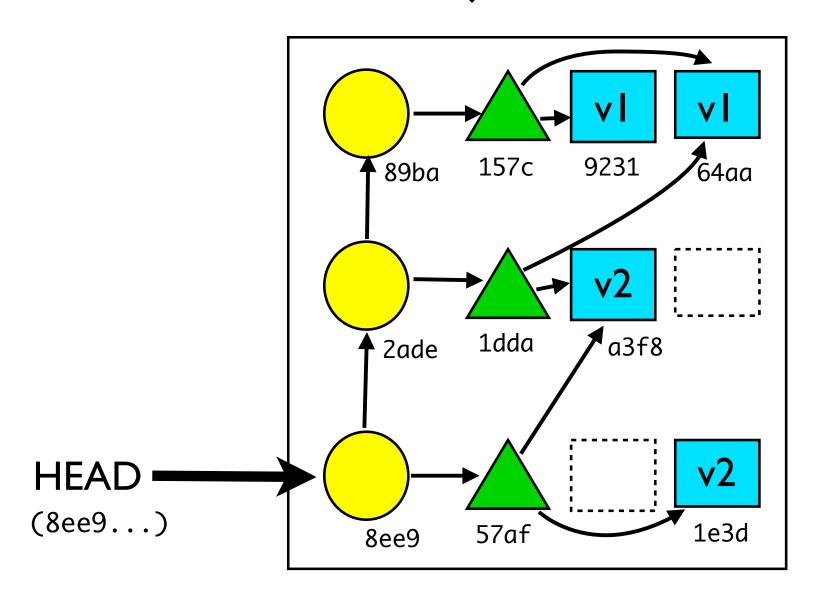
```
75e3dee3503e7ad5...
1586ad8975e3dee3... L
f6da34481586ad89...
f99b89d6f6da3448... 🛆
503d6749f99b89d6...
2b436957503d6749...
944ff522944ff522... A
6c024e25944ff522...
503e7ad56c024e25... F
```

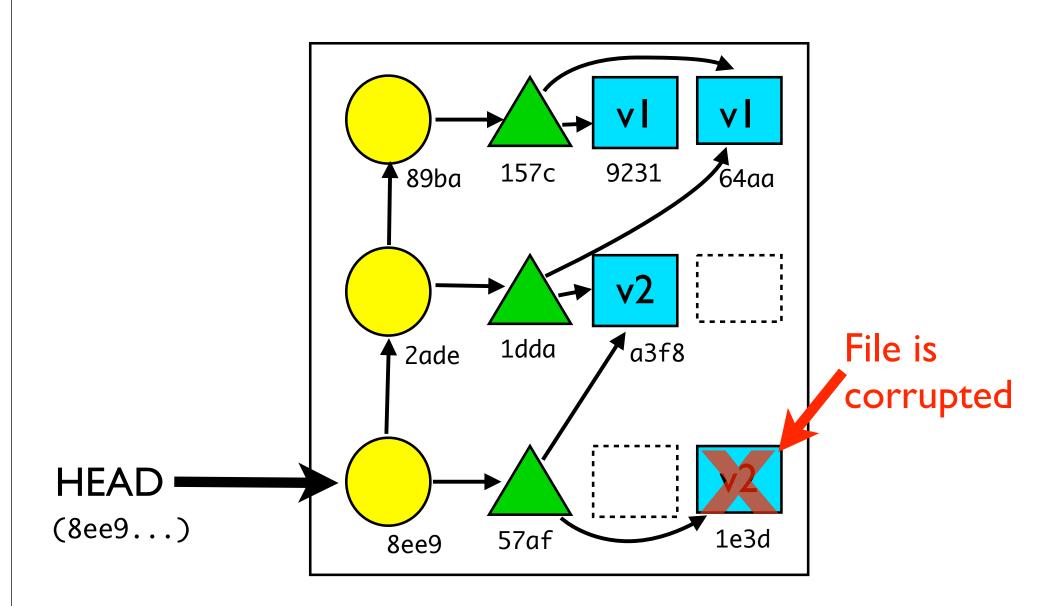


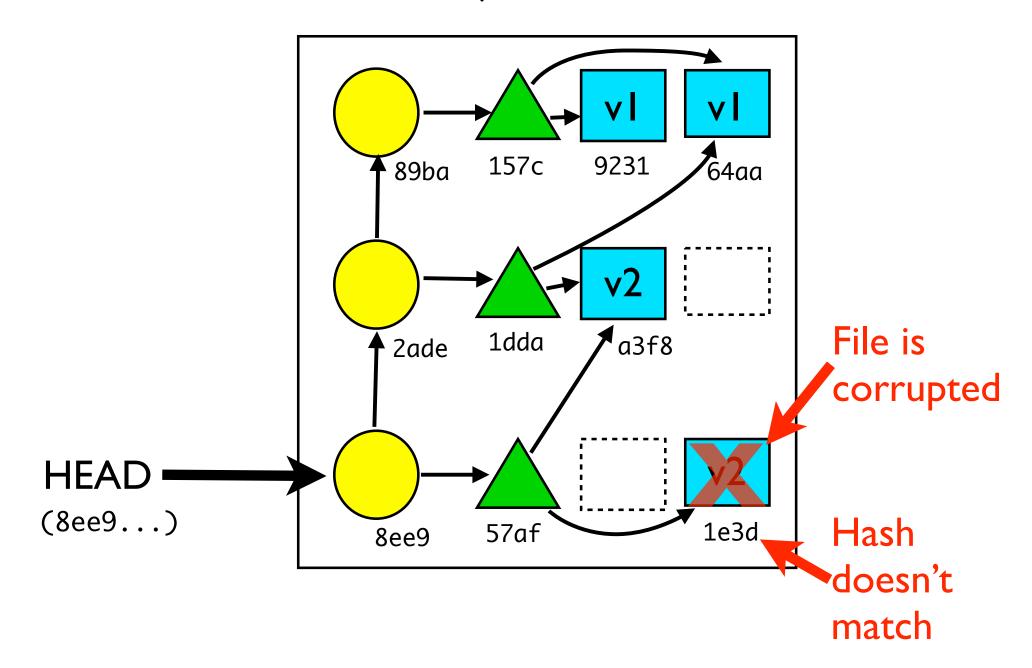
Project
Archive
Integrity

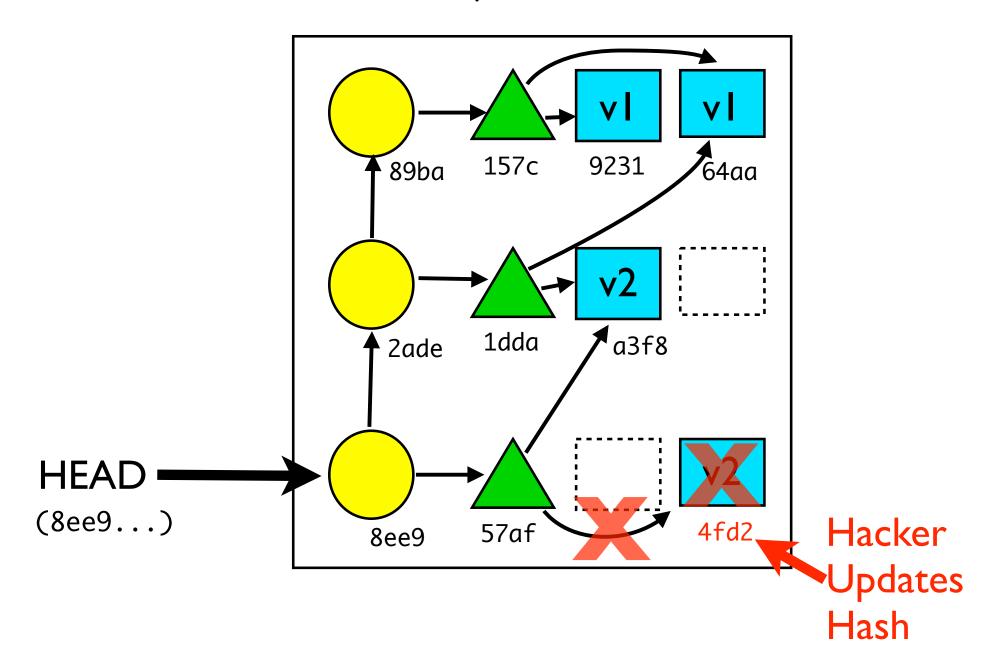


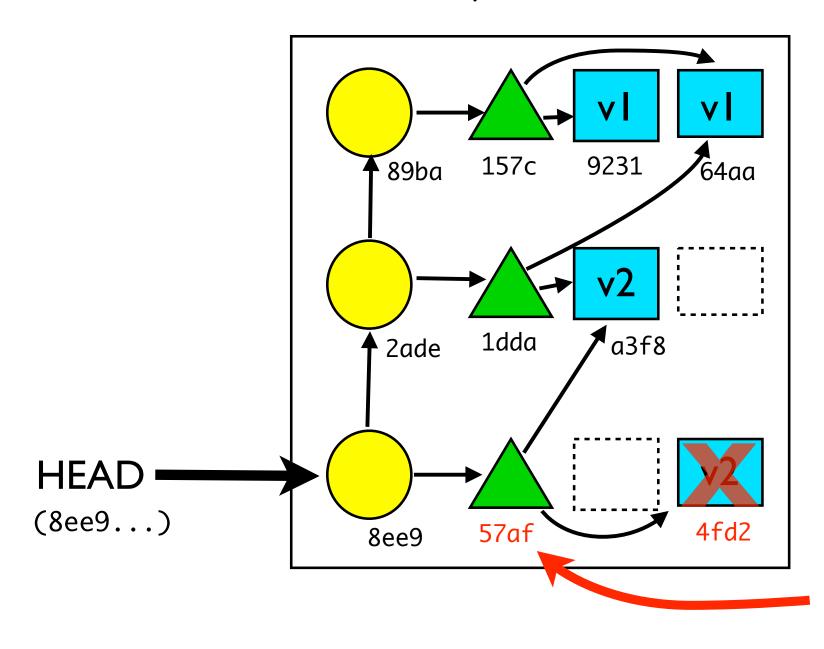




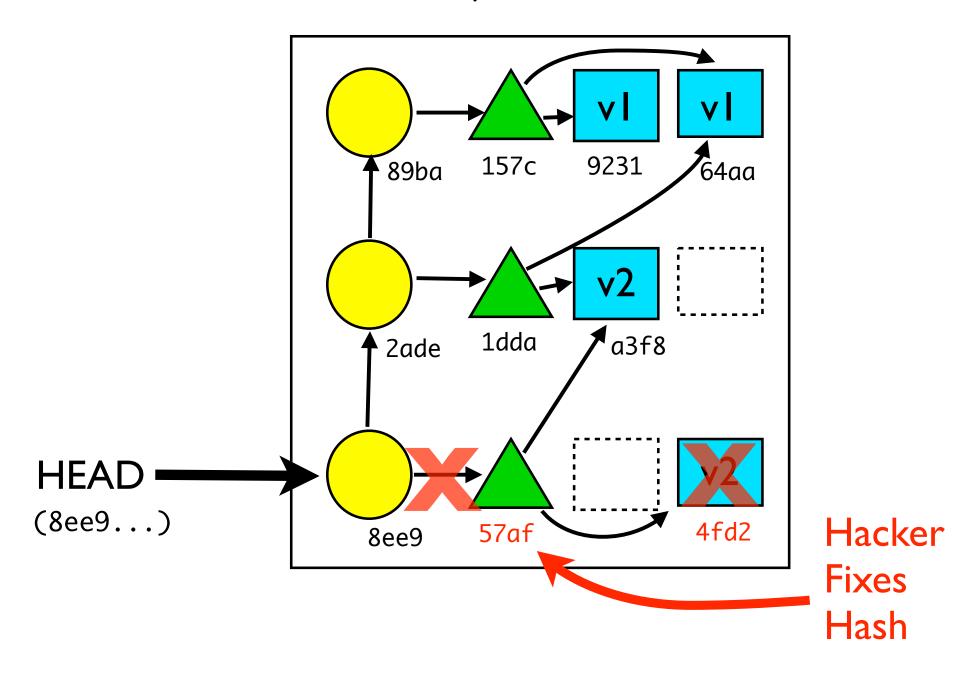


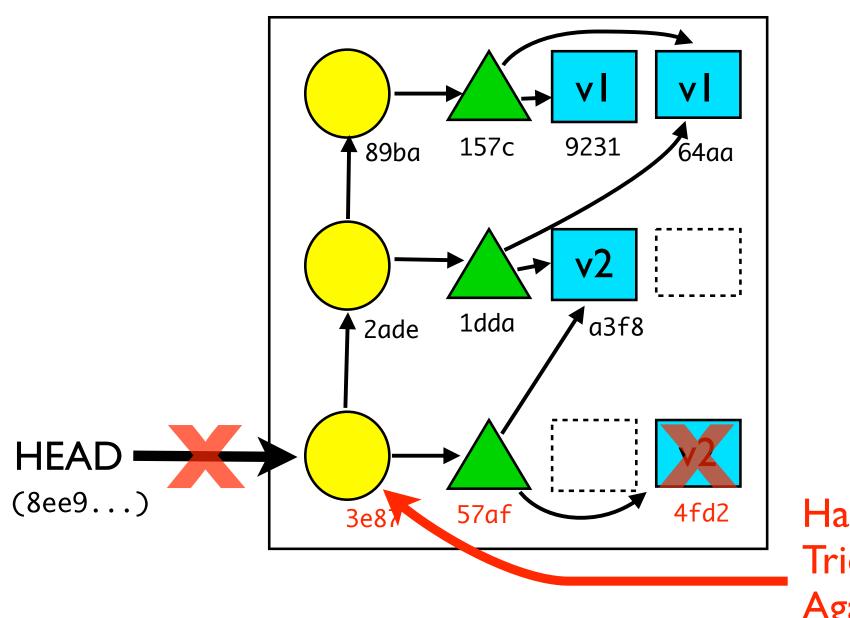






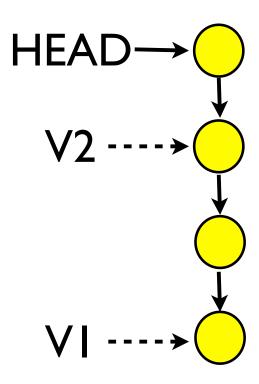
Hacker
Updates
Directory

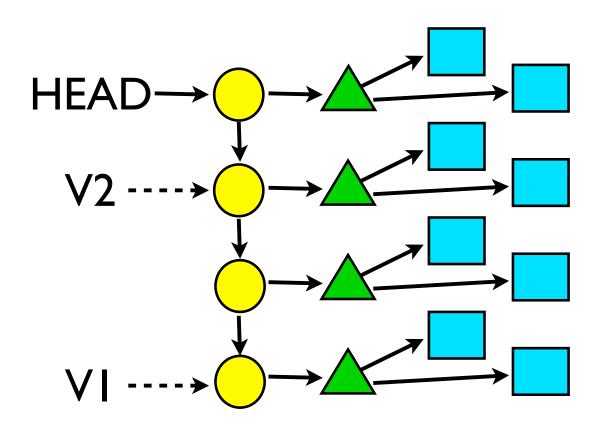


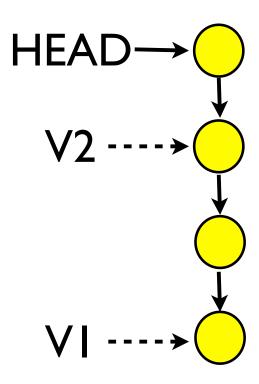


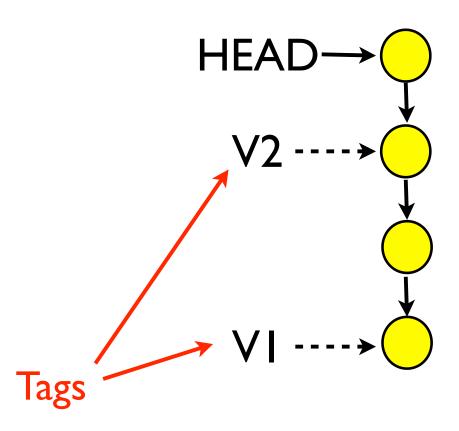
Hacker Tries Again

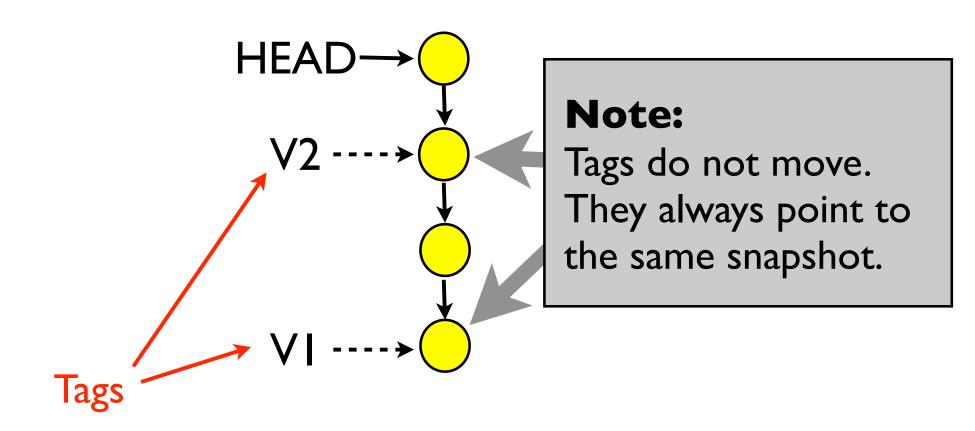
Tracking Snapshots

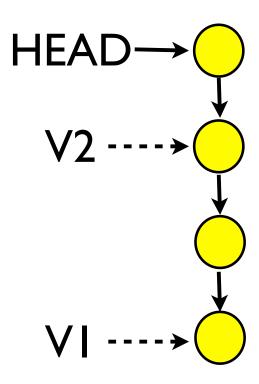


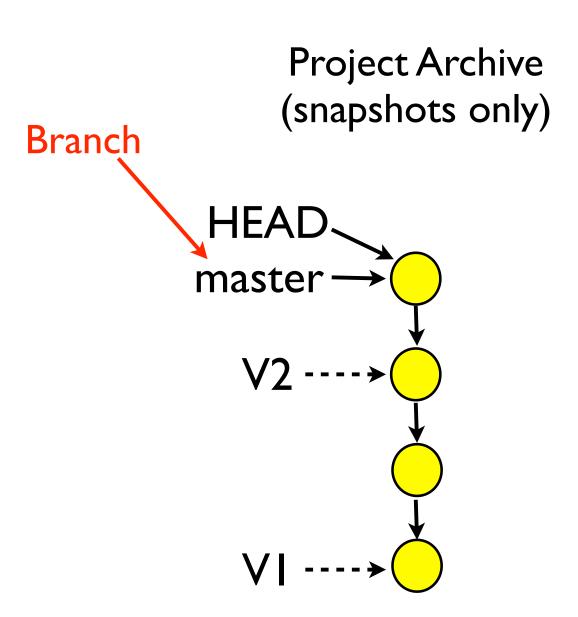


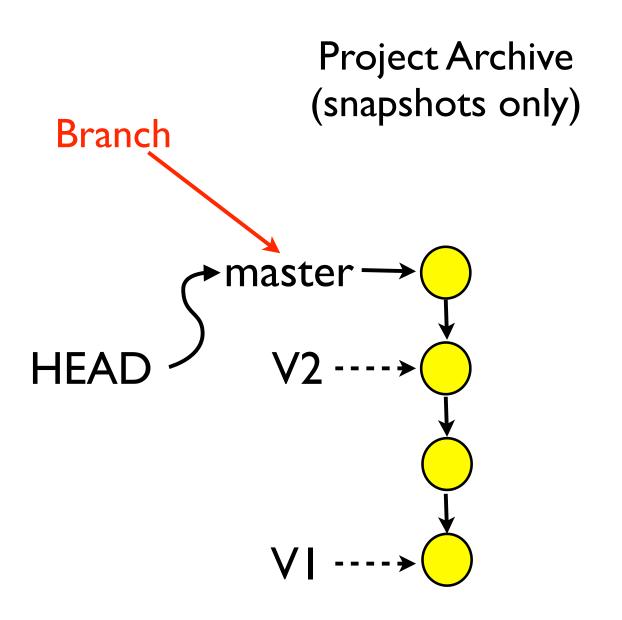


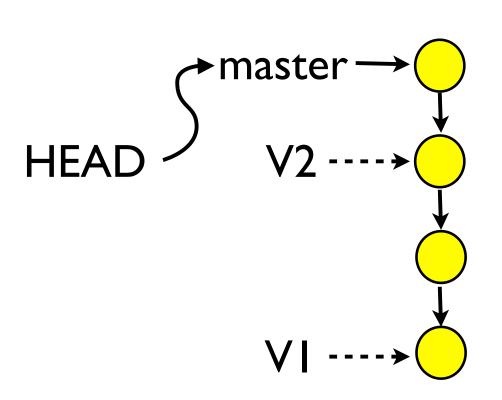




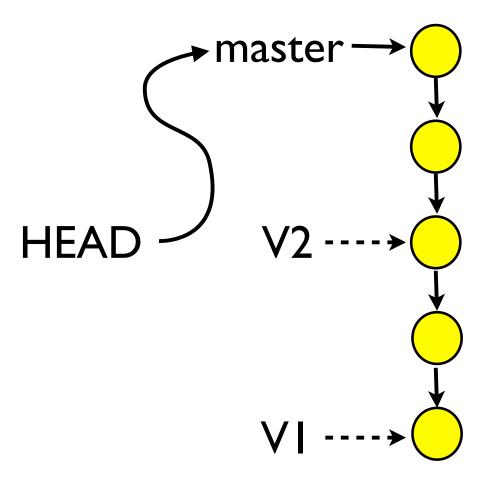




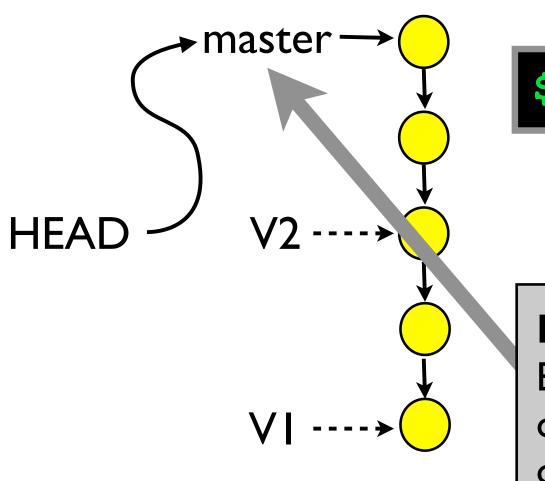








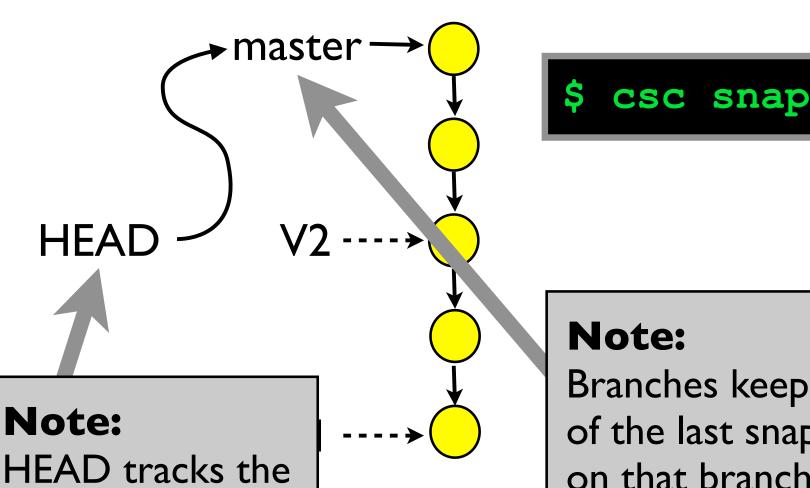
\$ csc snap



\$ csc snap

Note:

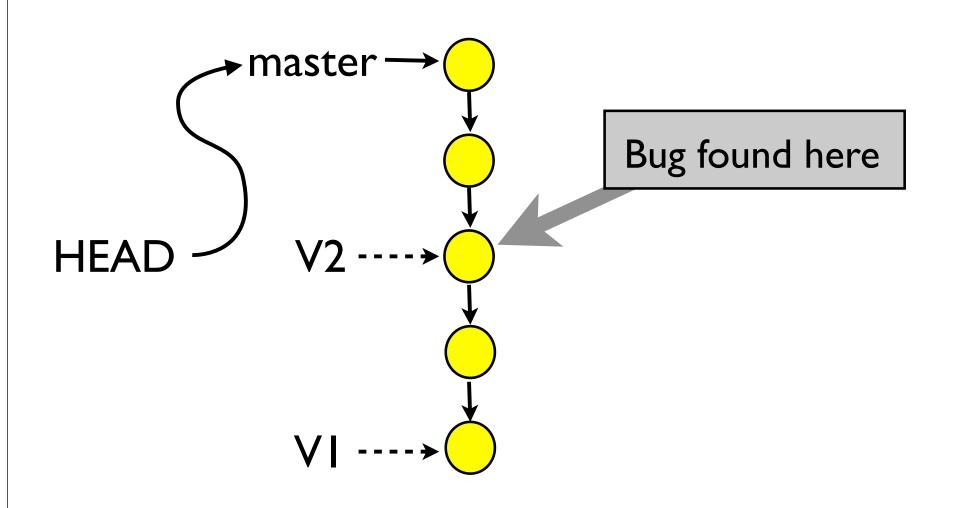
Branches keep track of the last snapshot on that branch.

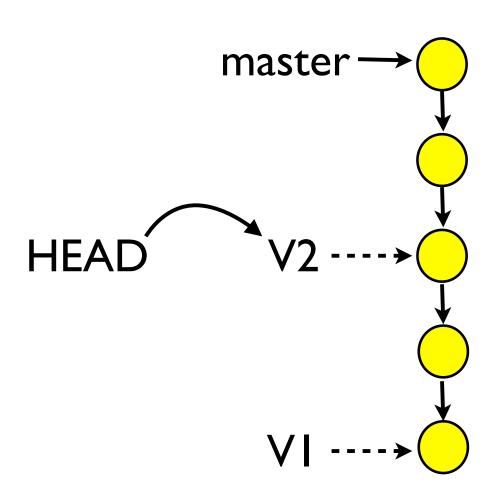


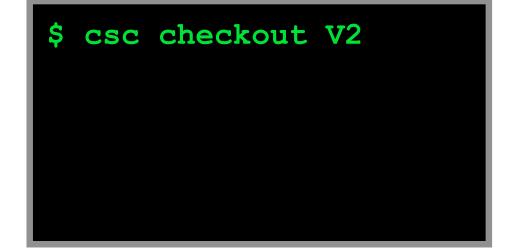
current branch.

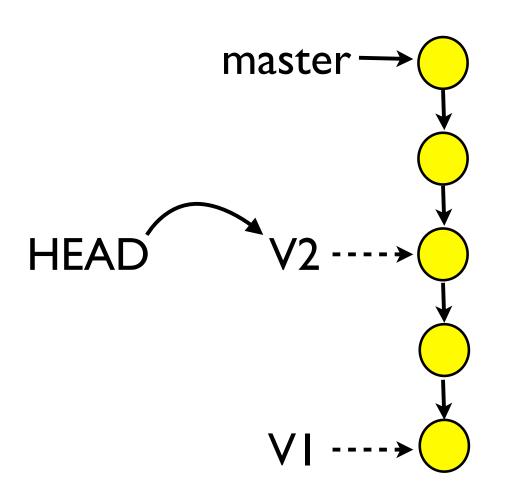
Branches keep track of the last snapshot on that branch.

Oops! We found a bug

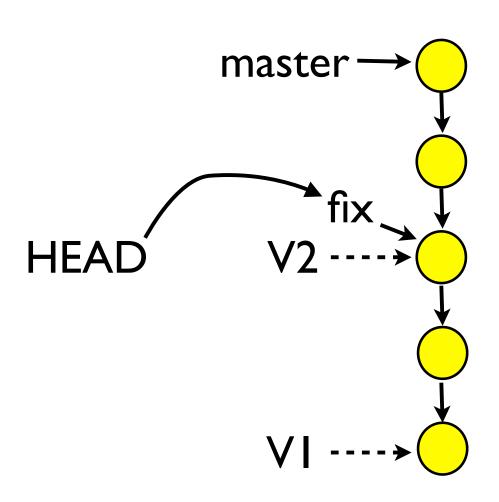




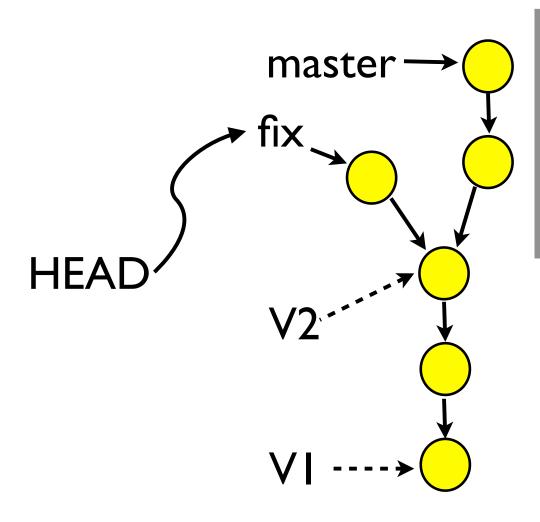




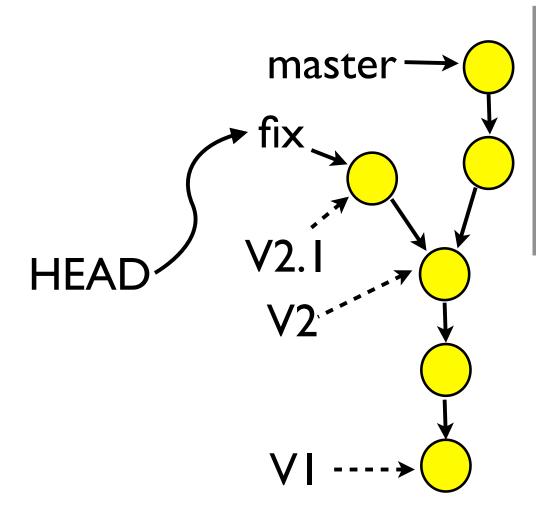
```
$ csc checkout V2
$ edit files
```



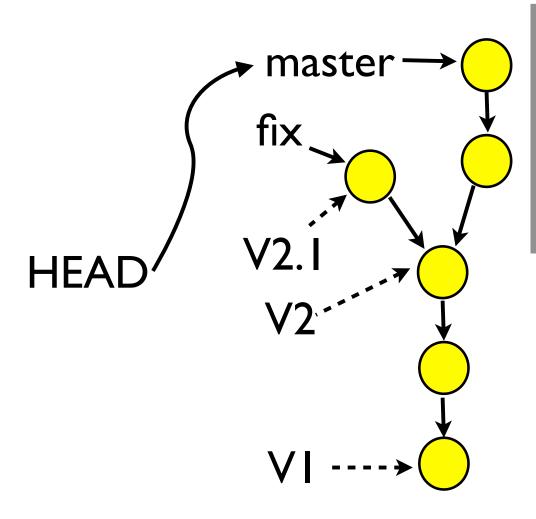
```
$ csc checkout V2
$ edit files
$ csc branch fix
```



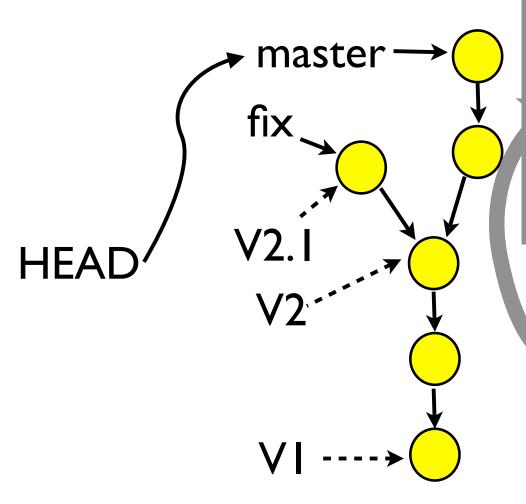
```
$ csc checkout V2
$ edit files
$ csc branch fix
$ csc snap
```



```
$ 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
```

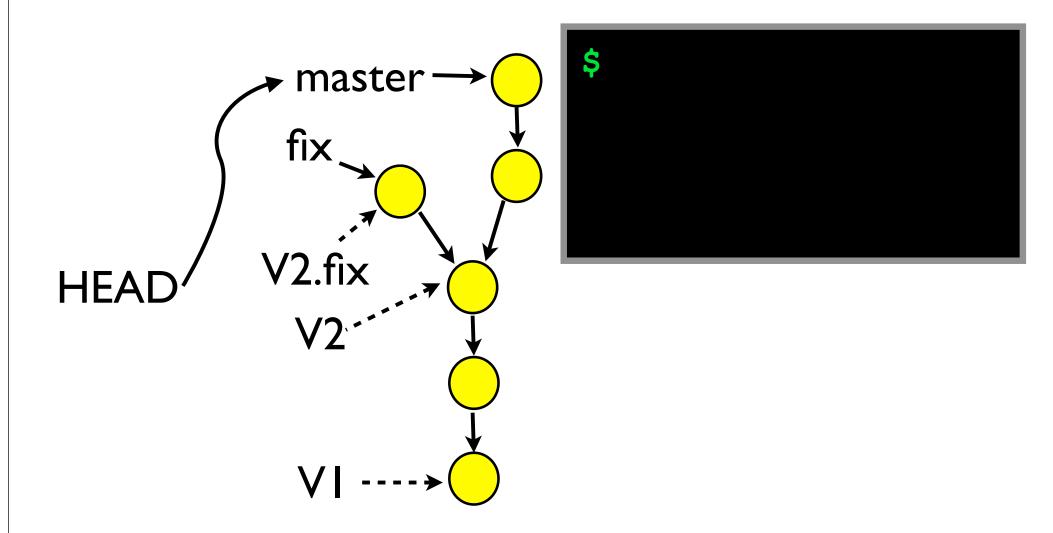


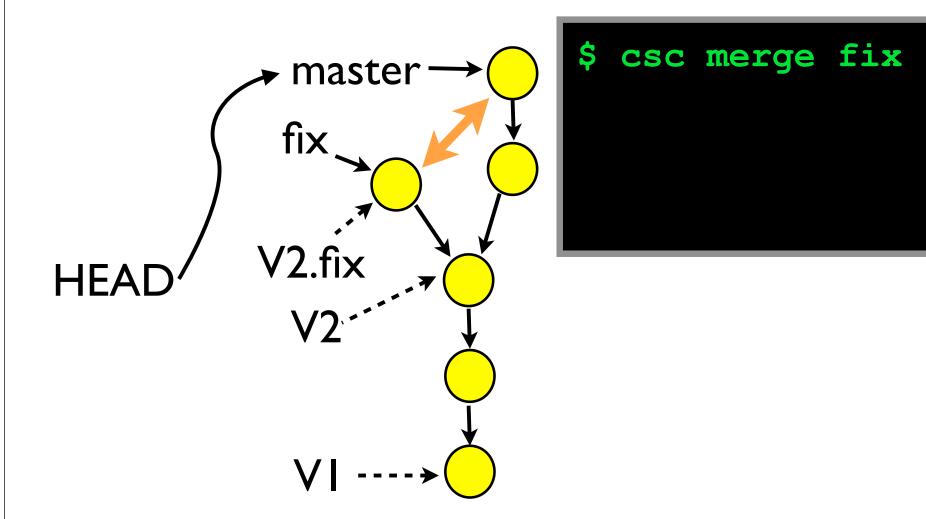
```
$ csc checkout V2
$ edit files
csc branch fix
csc snap
$ csc snap
$ csc tag V2.fix
$ csc checkout master
```

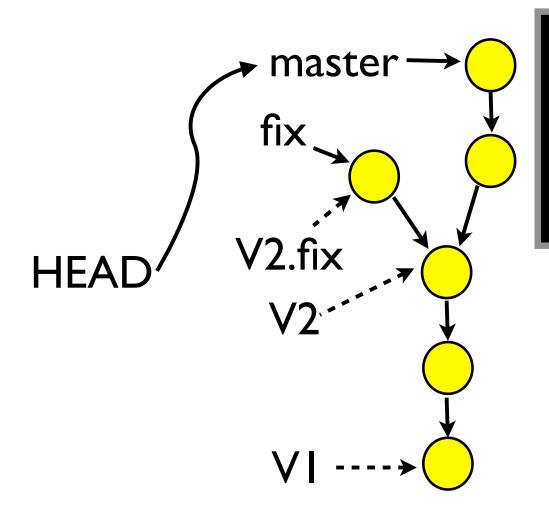
Note:

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

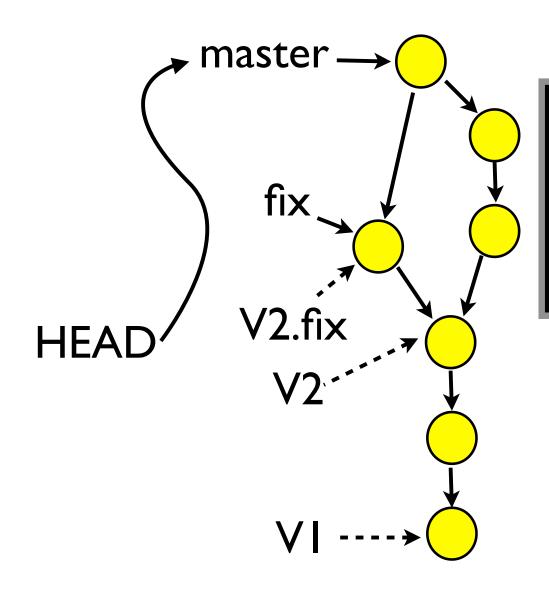
Merge Bug Fixes to Master



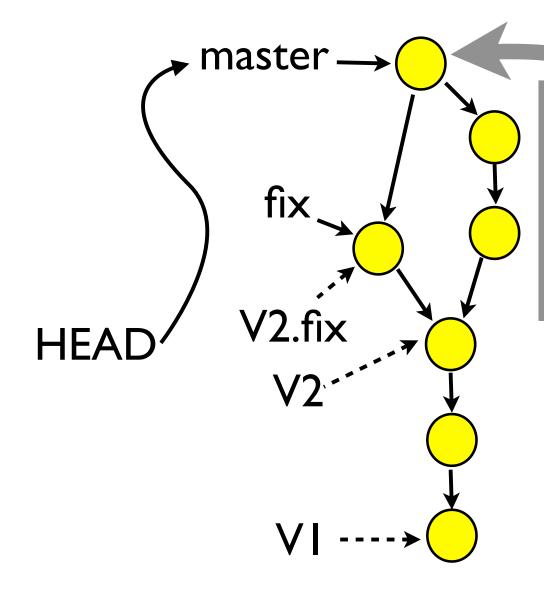




\$ csc merge fix
(resolve conflict)



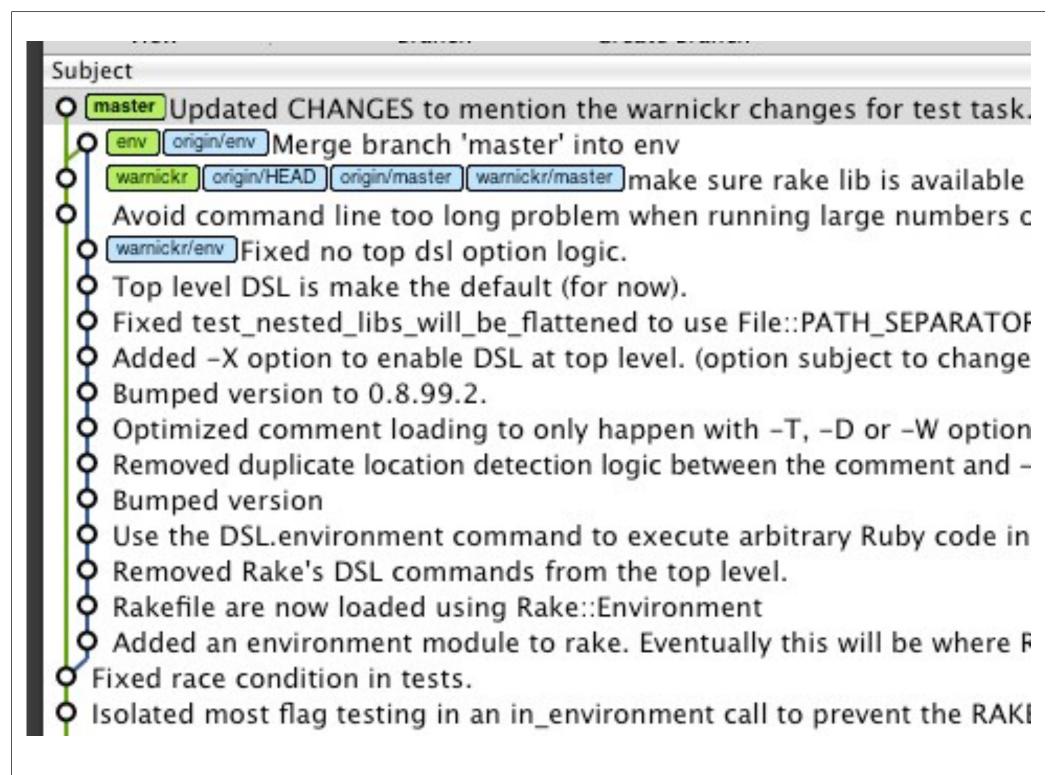
\$ csc merge fix
(resolve conflict)
\$ csc snap

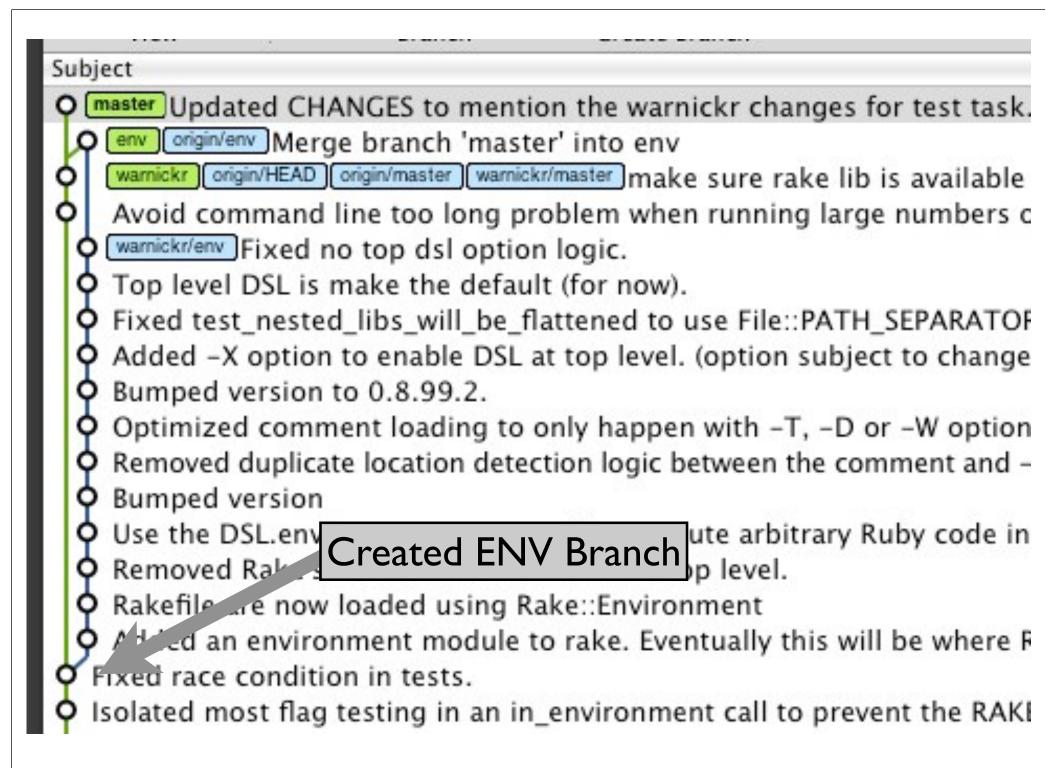


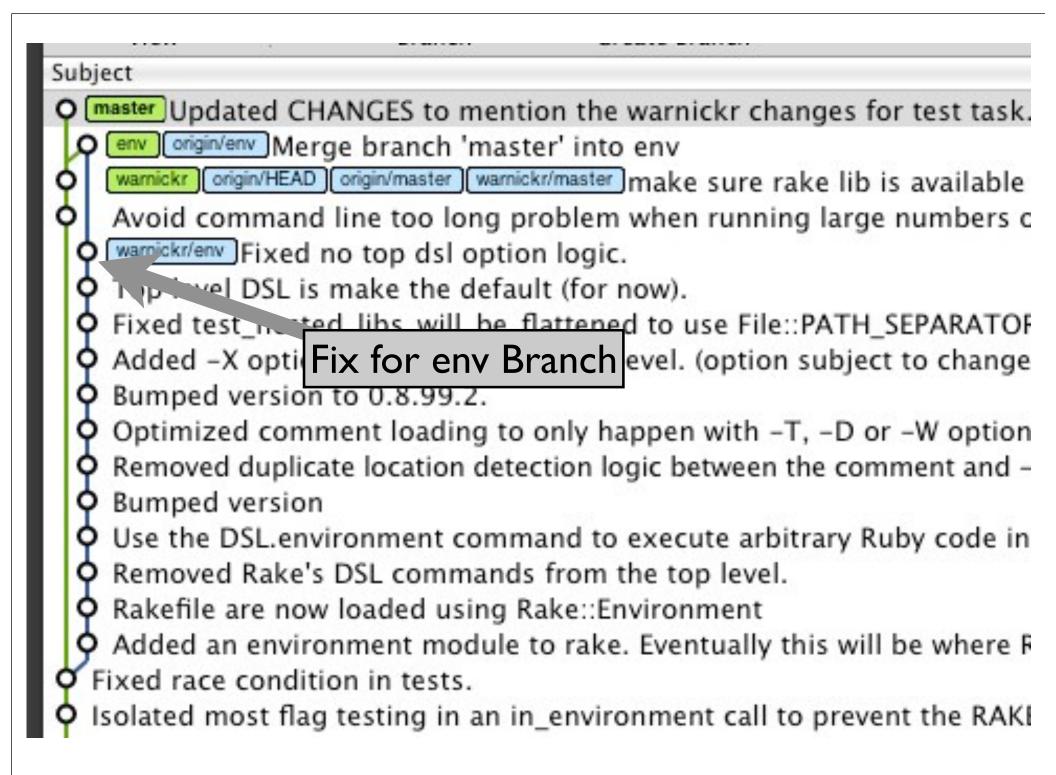
\$ csc merge fix
(resolve conflict)
\$ csc s.ap

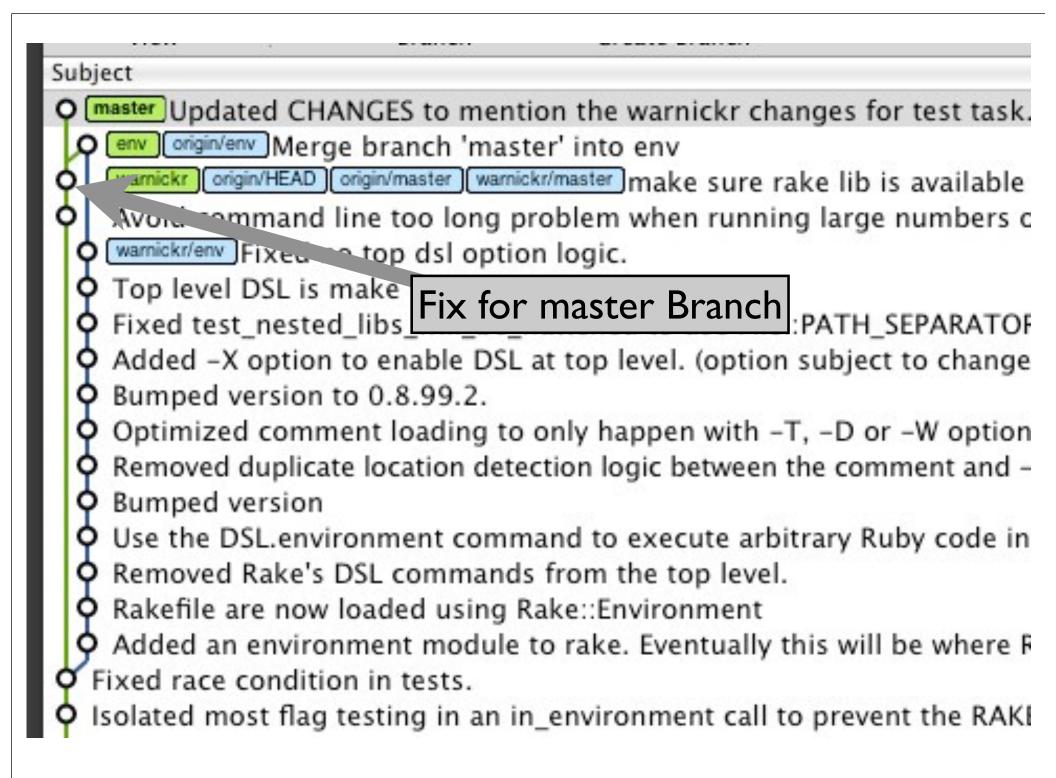
Note:

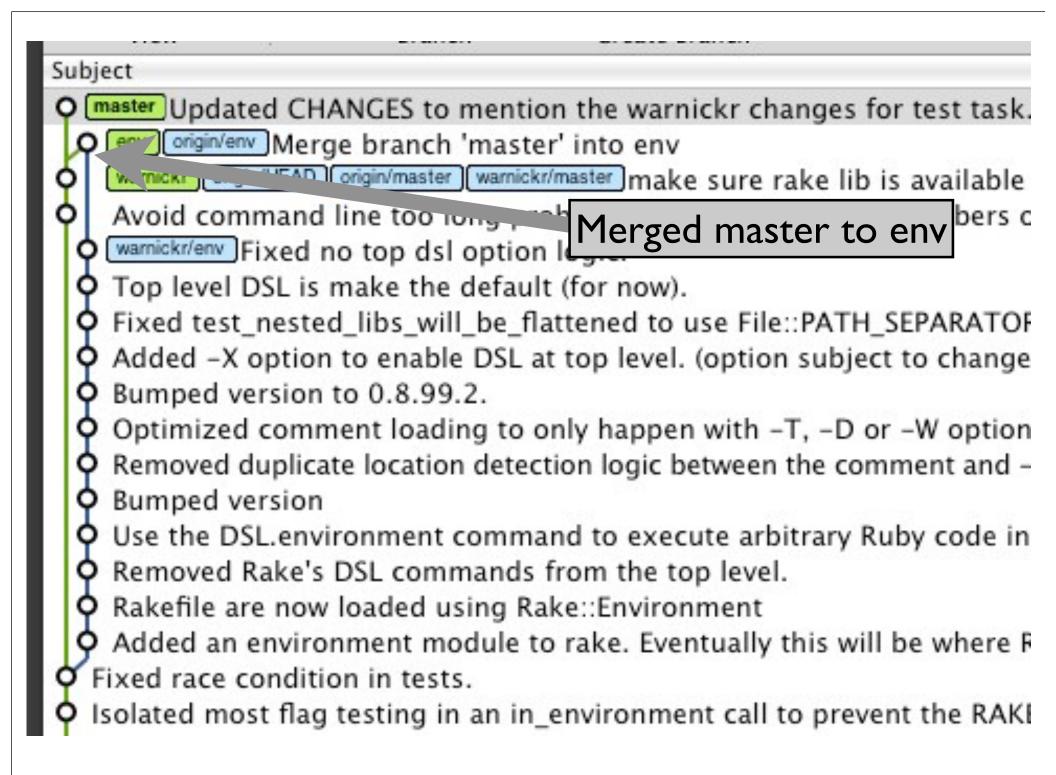
This snapshot has two parents. The changes from both parents have been merged together.

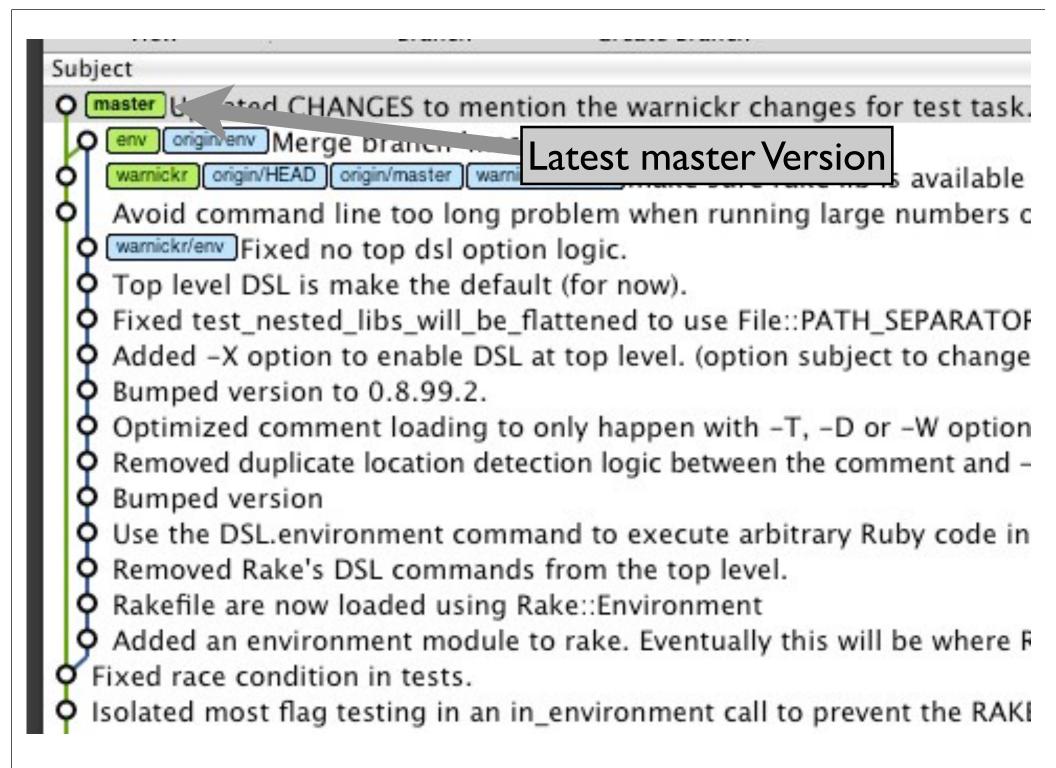


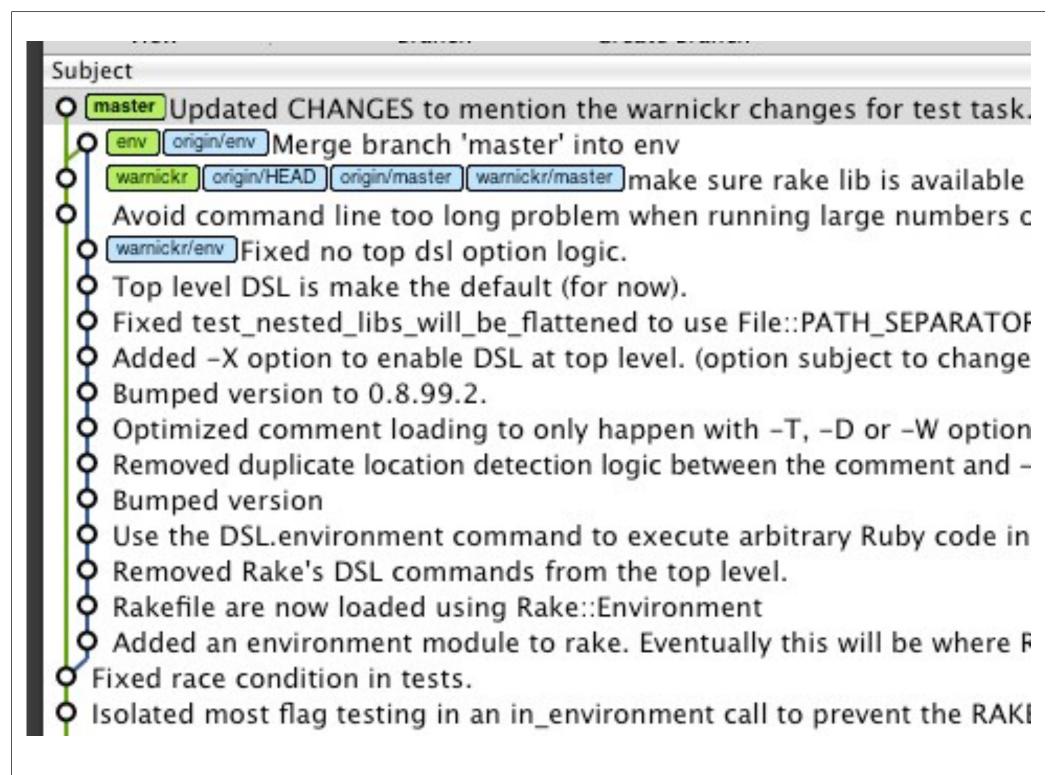




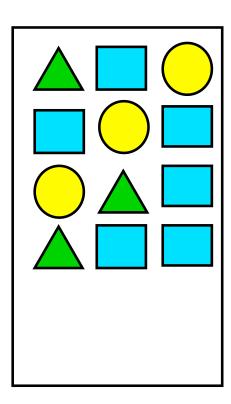




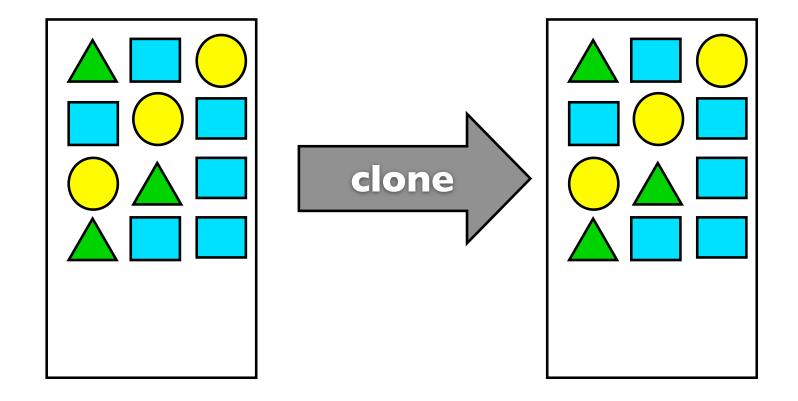




Sharing Code

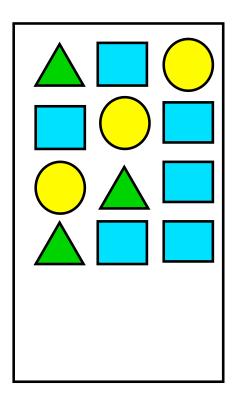


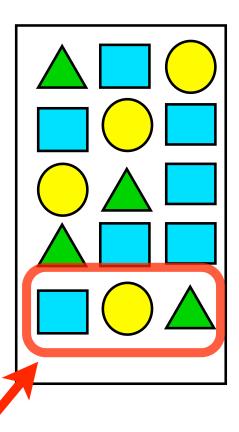
Your Archive



you\$ csc clone url_for_my_repo

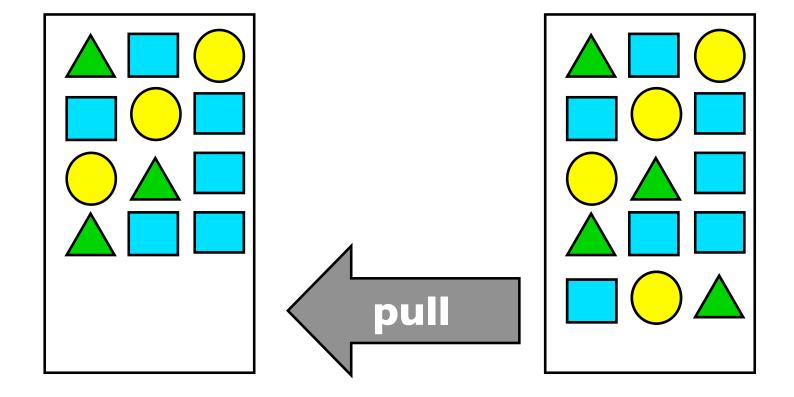
Your Archive





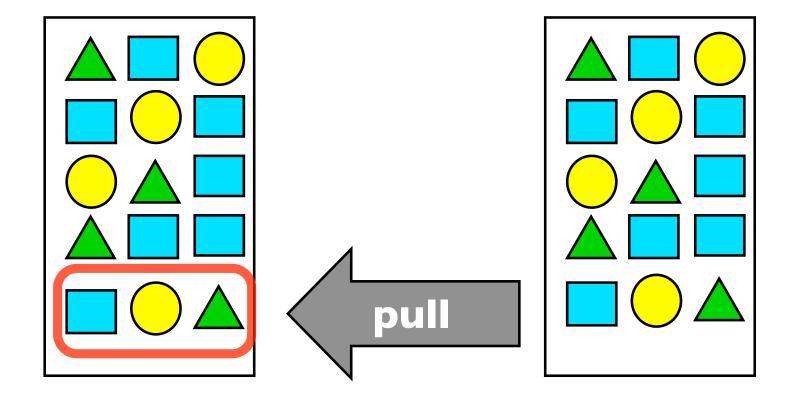
You make changes

Your Archive



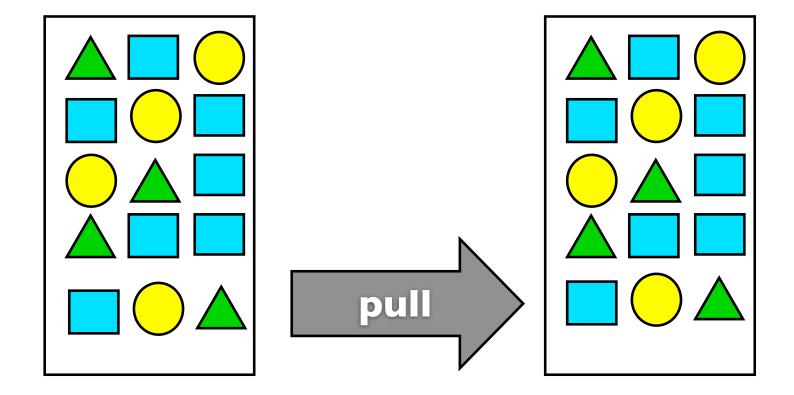
me\$ csc pull url_for_my_repo

Your Archive



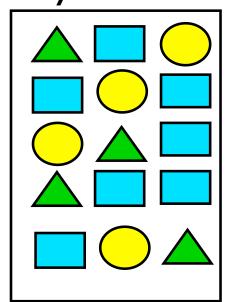
me\$ csc pull url_for_my_repo

Your Archive

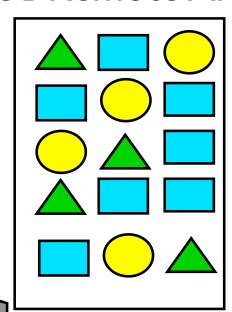


you\$ csc pull url_for_my_repo

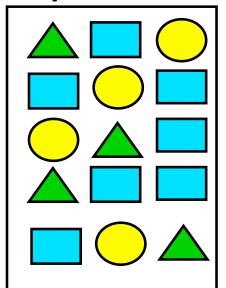
Common Remote Uses



Shared Remote Archive

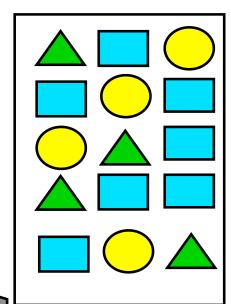


My Archive

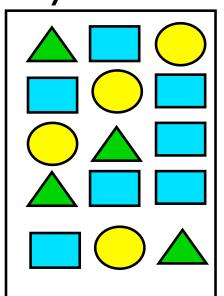


push pull

Shared Remote Archive

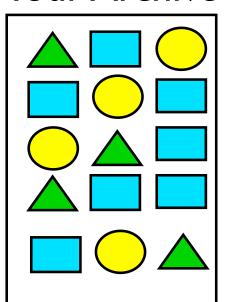


My Archive

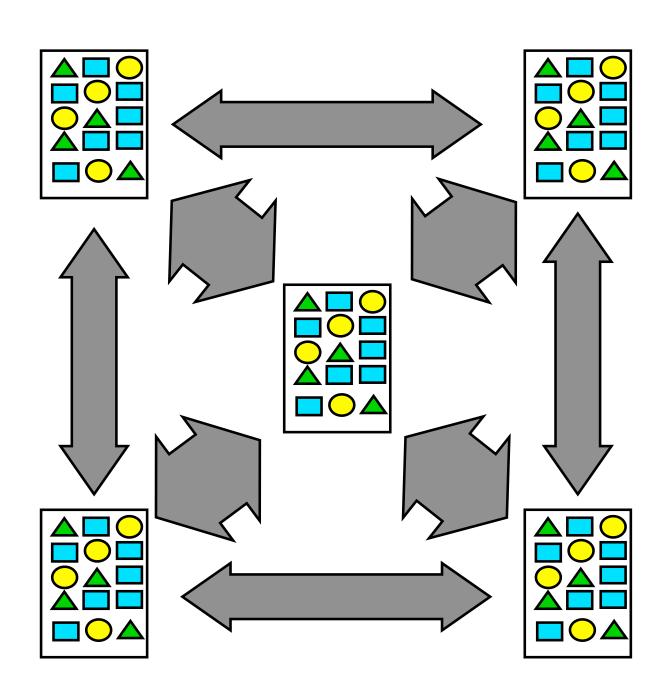


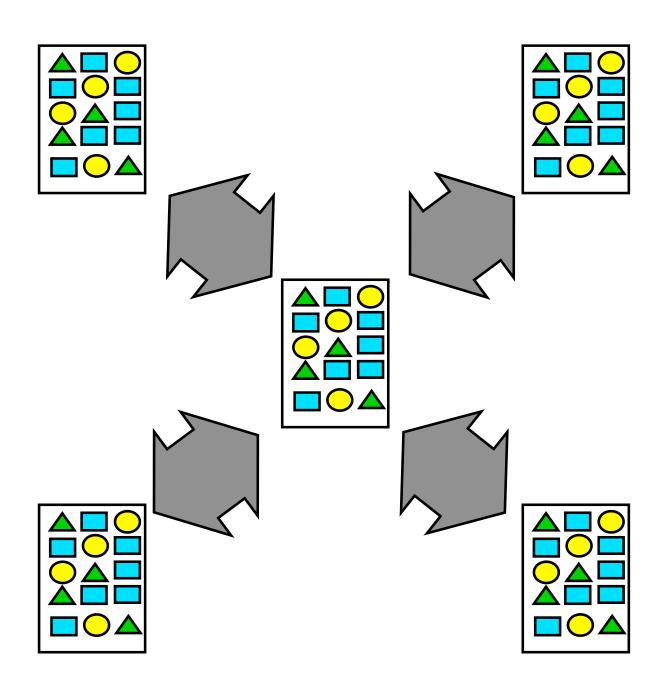
push pull push pull

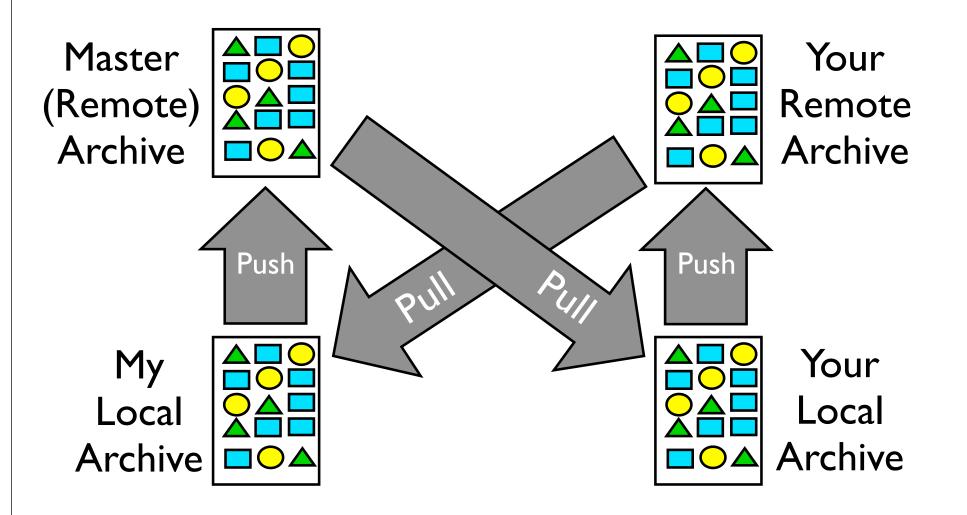
Your Archive



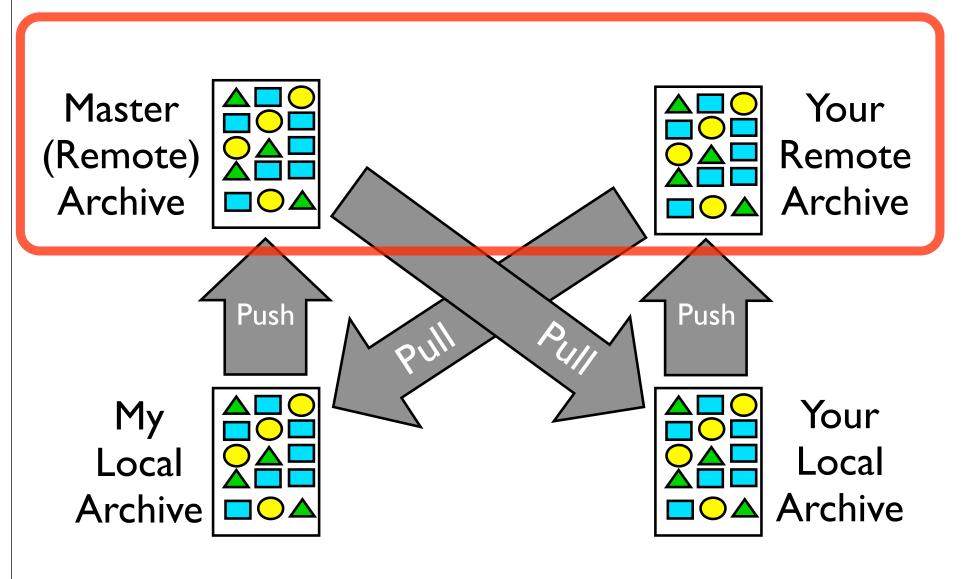
Shared Remote Archive Your Archive My Archive push push pull

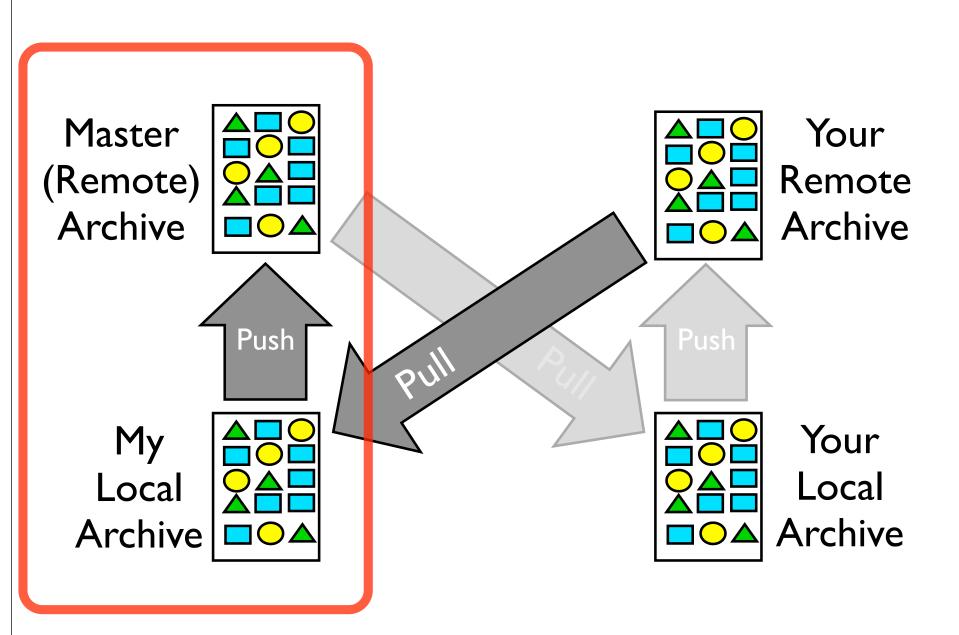




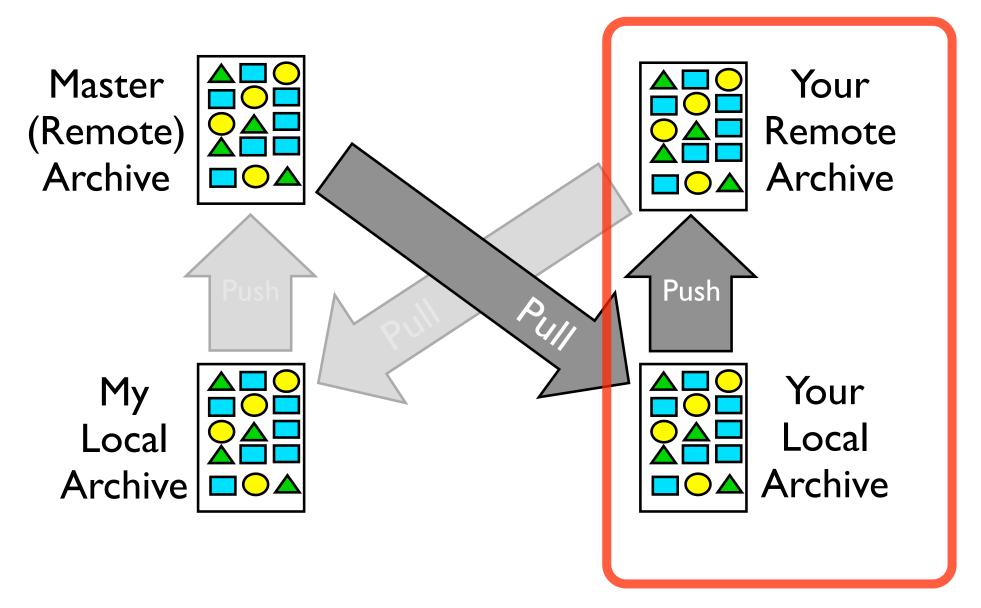


These Repos are Public

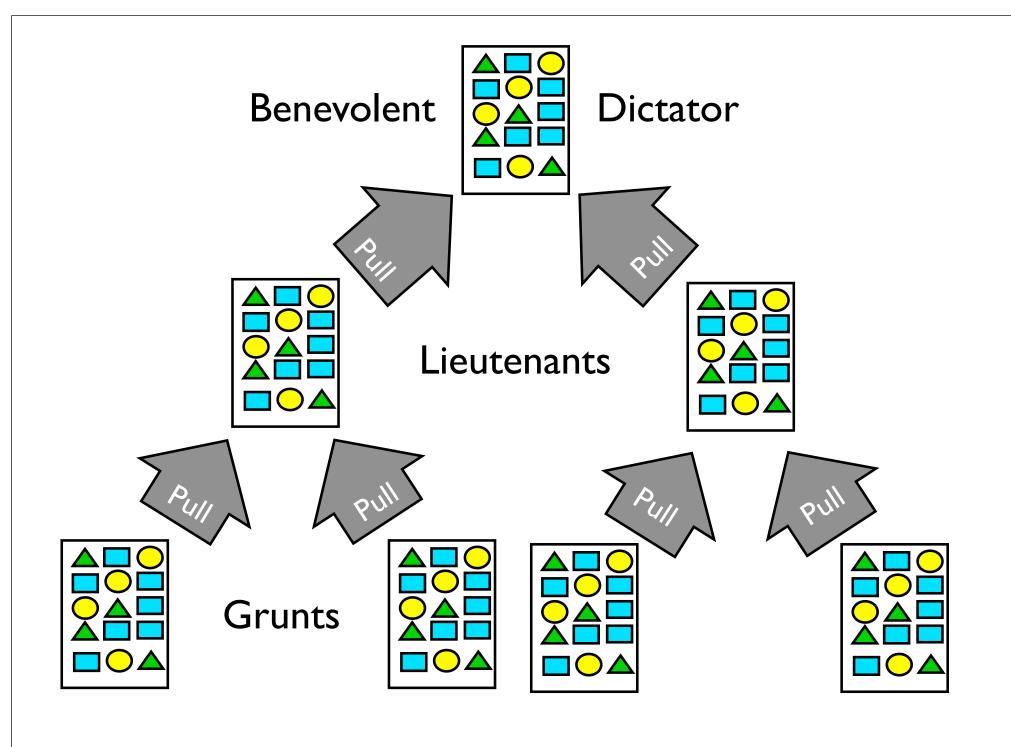


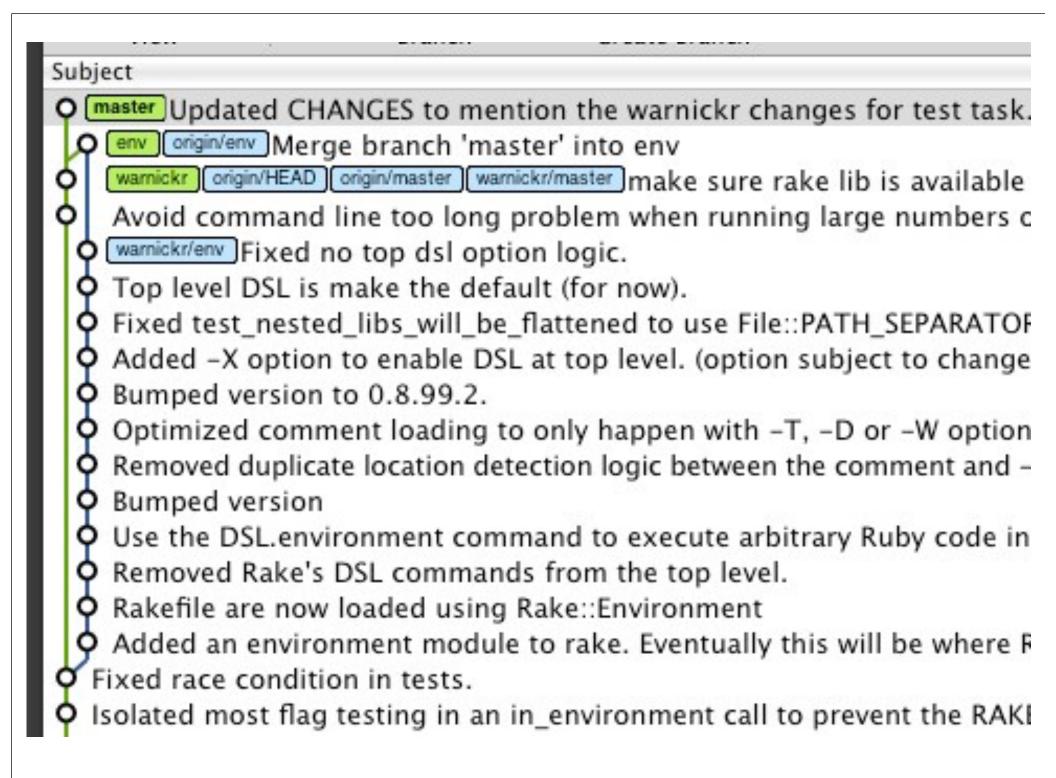


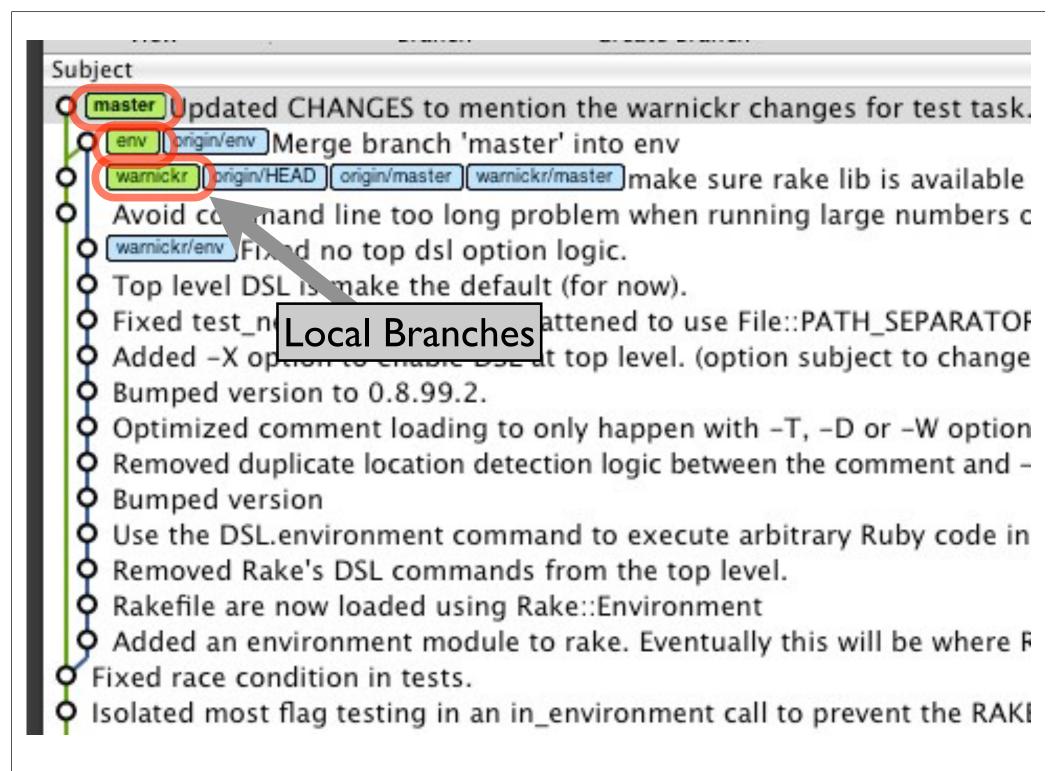
I Own These Repos

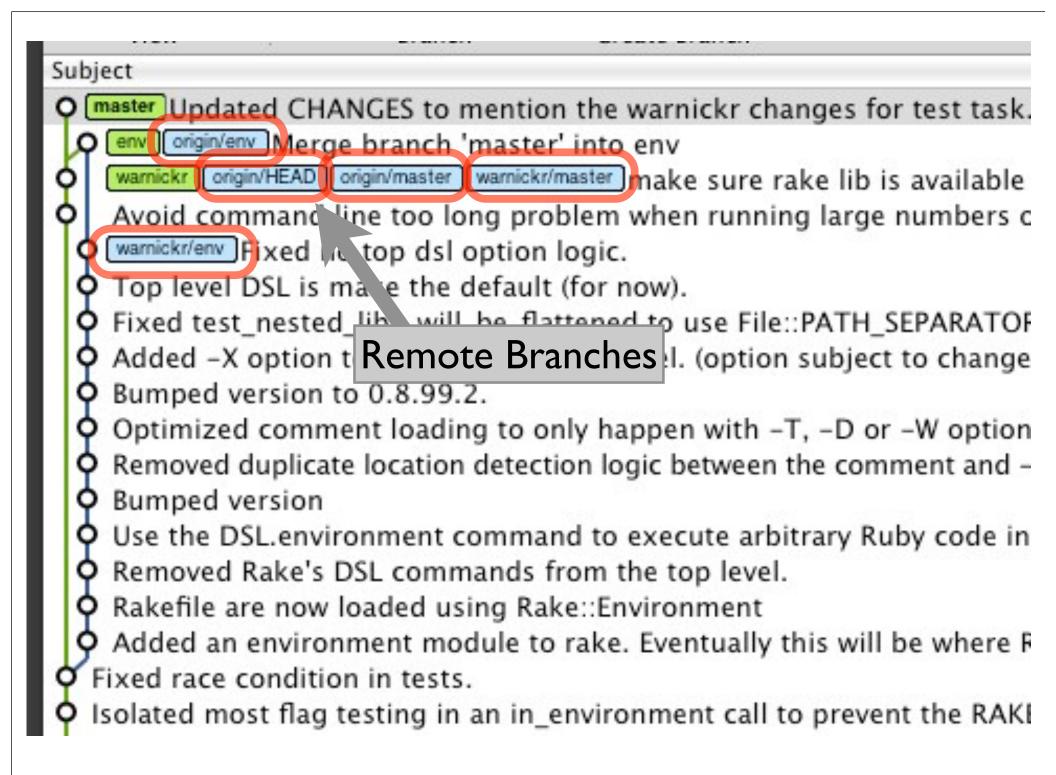


You Own These Repos









Wrap Up

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

Sounds Like ...

CSC

GIT

```
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

Us

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

- Git Uses
 - Working Directory
 - Staging Area
 - Repository

Us

One level object file storage

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

Us

Push/Pull entire repos

- Push/Pull individual branches
- Fetch without merging to local branches
- Metadata to map branches to remote branches

Us

• Just the basic features

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

Demo

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

Questions?