

Bioconductor's GIT transition

Implementation details

Nitesh Turaga
Bioconductor Core
Roswell Park Cancer Institute

Bioc2017, July 26th 2017
Dana Farber Cancer Institute

The Challenge

SVN

REDUCE COMPLEXITY

GIT

1300+ Software packages

300+ Experiment Data packages

15 years worth of commit history

30+ Workflow packages

Manifest Files

Maintainer passwords/access
rights

Build System

The Challenge

SVN

REDUCE COMPLEXITY

GIT

1300+ Software packages

300+ Experiment Data packages

14 years worth of commit history

30+ Workflow packages

Manifest Files

Maintainer passwords/access
rights

Build System

1300+ software packages

300+ experiment data

30+ workflow

Simplify manifests

Gitolite + git pre-receive hooks

Build system change to use git

The Solution

git.bioconductor.org

Access to server

1. SSH (read + write) permissions
2. HTTPS (read only) permissions

Maintainers:
Please submit your SSH keys

IMPORTANT!!!!

www.bit.ly/submit-bioc-ssh-keys

Commit history

- Preserve entire commit history,
- e.g.:
 - `git log master` ## for complete log
 - `git log RELEASE_X_Y` ## for release log

svn log

```
~/D/tmp >>> svn log https://hedgehog.fhcrc.org/bioconductor/trunk/madman/Rpacks/affy | tail -n 30
```

```
-----  
r469 | dfcimm2 | 2002-02-16 05:36:46 -0500 (Sat, 16 Feb 2002) | 2 lines
```

```
...still trying to update INDEX...
```

```
-----  
r456 | dfcimm1 | 2002-02-14 12:32:19 -0500 (Thu, 14 Feb 2002) | 2 lines
```

```
*** empty log message ***
```

```
-----  
r437 | dfcimm1 | 2002-02-09 19:02:31 -0500 (Sat, 09 Feb 2002) | 2 lines
```

```
*** empty log message ***
```

```
-----  
r422 | dfcimm2 | 2002-02-08 07:44:51 -0500 (Fri, 08 Feb 2002) | 2 lines
```

```
I think the function image only works from R-1.2.0
```

```
-----  
r405 | dfcimm1 | 2002-02-06 17:03:12 -0500 (Wed, 06 Feb 2002) | 2 lines
```

```
Initial revision
```

2002 first commit to `affy`

git log

```
~/D/t/affy >>> git log | tail -n 45
```

master

```
[commit 99c617a389eaca6d924481ad252078252267a4fd
Author: unknown <unknown>
Date: Sat Feb 16 10:36:46 2002 +0000
```

```
...still trying to update INDEX...
```

```
git-svn-id: file:///home/git/hedgehog.fhcrc.org/bioconductor/trunk/madman/Rpacks/affy@469 bc3139a8-67e5-0310-9ffc-ced21a209358
```

```
commit ccb9c56aed432fbe895acb569c19c20f2b1fb146
Author: unknown <unknown>
Date: Thu Feb 14 17:32:19 2002 +0000
```

```
*** empty log message ***
```

```
git-svn-id: file:///home/git/hedgehog.fhcrc.org/bioconductor/trunk/madman/Rpacks/affy@456 bc3139a8-67e5-0310-9ffc-ced21a209358
```

```
commit f444ca3079d19bee1a350f6b3f4add51a0d253e1
Author: unknown <unknown>
Date: Sun Feb 10 00:02:31 2002 +0000
```

```
*** empty log message ***
```

```
git-svn-id: file:///home/git/hedgehog.fhcrc.org/bioconductor/trunk/madman/Rpacks/affy@437 bc3139a8-67e5-0310-9ffc-ced21a209358
```

```
commit 8a94f15eef0870f3220c654a1bd3320bfd0e4cb1
Author: unknown <unknown>
Date: Fri Feb 8 12:44:51 2002 +0000
```

```
I think the function image only works from R-1.2.0
```

```
git-svn-id: file:///home/git/hedgehog.fhcrc.org/bioconductor/trunk/madman/Rpacks/affy@422 bc3139a8-67e5-0310-9ffc-ced21a209358
```

```
commit ec9af7aac8ea189658910f7a62f5e4d7c0048b63
Author: unknown <unknown>
Date: Wed Feb 6 22:03:12 2002 +0000
```

```
Initial revision
```

```
git-svn-id: file:///home/git/hedgehog.fhcrc.org/bioconductor/trunk/madman/Rpacks/affy@405 bc3139a8-67e5-0310-9ffc-ced21a209358
```

```
[~/D/t/affy >>> git branch -a
* master
remotes/origin/HEAD -> origin/master
remotes/origin/RELEASE_1_0
remotes/origin/RELEASE_1_0_branch
remotes/origin/RELEASE_1_4
remotes/origin/RELEASE_1_4_branch
remotes/origin/RELEASE_1_5
remotes/origin/RELEASE_1_6
remotes/origin/RELEASE_1_7
remotes/origin/RELEASE_1_8
remotes/origin/RELEASE_1_9
remotes/origin/RELEASE_2_0
remotes/origin/RELEASE_2_1
remotes/origin/RELEASE_2_10
remotes/origin/RELEASE_2_11
remotes/origin/RELEASE_2_12
remotes/origin/RELEASE_2_13
remotes/origin/RELEASE_2_14
remotes/origin/RELEASE_2_2
remotes/origin/RELEASE_2_3
remotes/origin/RELEASE_2_4
remotes/origin/RELEASE_2_5
remotes/origin/RELEASE_2_6
remotes/origin/RELEASE_2_7
remotes/origin/RELEASE_2_8
remotes/origin/RELEASE_2_9
remotes/origin/RELEASE_3_0
remotes/origin/RELEASE_3_1
remotes/origin/RELEASE_3_2
remotes/origin/RELEASE_3_3
remotes/origin/RELEASE_3_4
remotes/origin/RELEASE_3_5
remotes/origin/master
```

git branch -a

Implementation

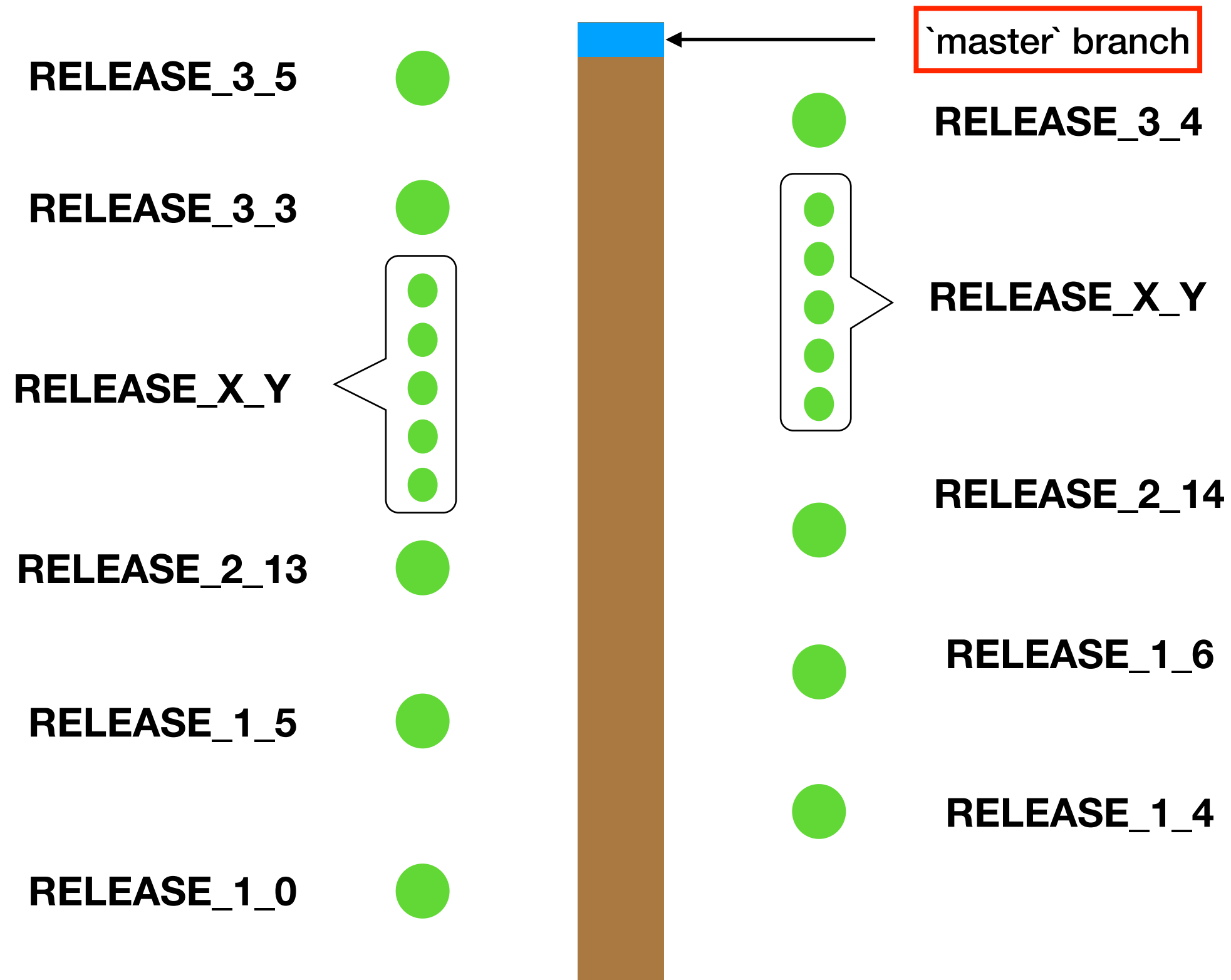
Implementation:

git svn clone



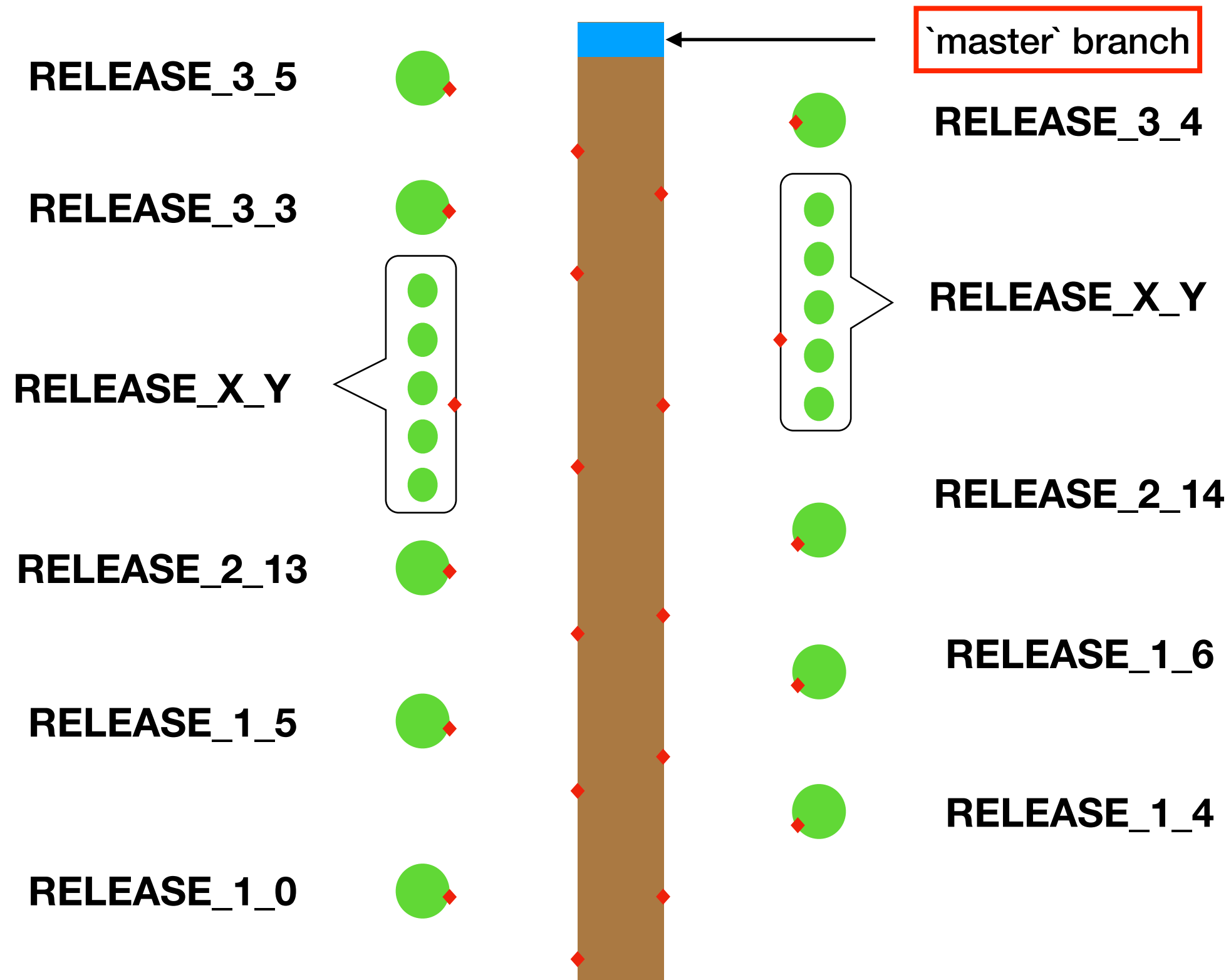
Implementation: Add release branches

Add orphan branch points



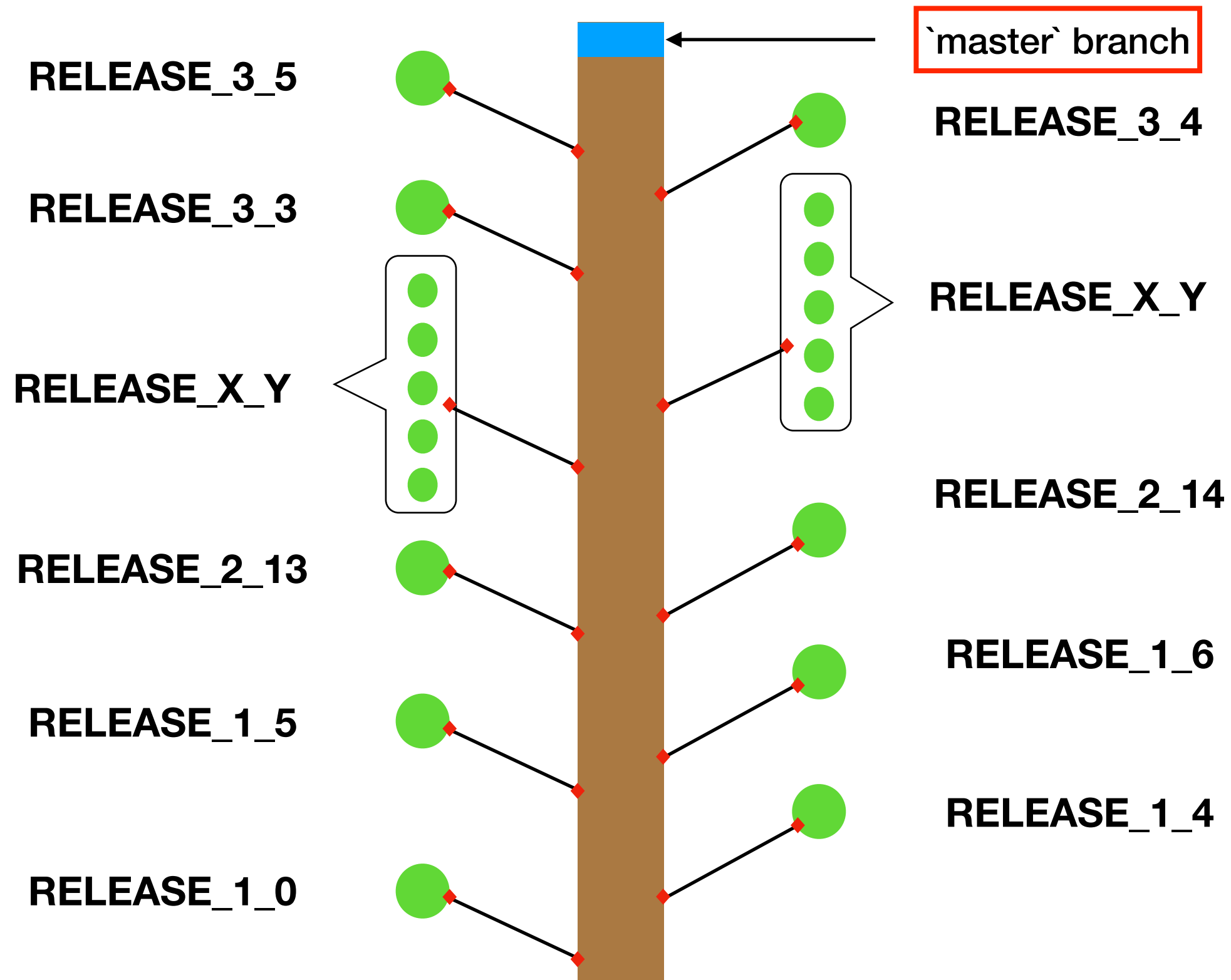
Implementation: Add commit history

find branch points



Implementation: Add commit history

`graft` and `filter-branch`



More Implementation

- Data experiment packages
- Workflow packages
- Manifest files
- Git server setup

Maintainers:
submit your SSH keys

IMPORTANT!!!!

www.bit.ly/submit-bioc-ssh-keys

The amazing Bioconductor core team

Martin Morgan

Hervé Pagès

Valarie Obenchain

Lori Shepherd

Marcel Ramos

Daniel Van Twisk

Jim Hester

+

Community: Alpha testers, Beta testers

Thank you.

Questions?

 **GitHub : nturaga**