ELIXIR Webinar

Bioconda and BioContainers

Tool deployment and sustainability in science

Björn Grüning
Bioconda/BioContainer community

www.elixir-europe.org

Tool deployment & sustainability in science

Question ⟶ Development of code ⟶ Deployment

Tool deployment & sustainability in science

Question ⟶ Development of code ⟶ Packaging ⟶ Deployment

Package managers are charged with the task of **finding**, **installing**, **maintaining** or **uninstalling** software packages upon the **user**'s command.

Tool deployment & sustainability in science

Question ⟶ Development of code ⟶ Packaging ⟶ Deployment

CP⬡N
pip Bioconductor
R CRAN

yum
apt-get
docker
Biocontainers
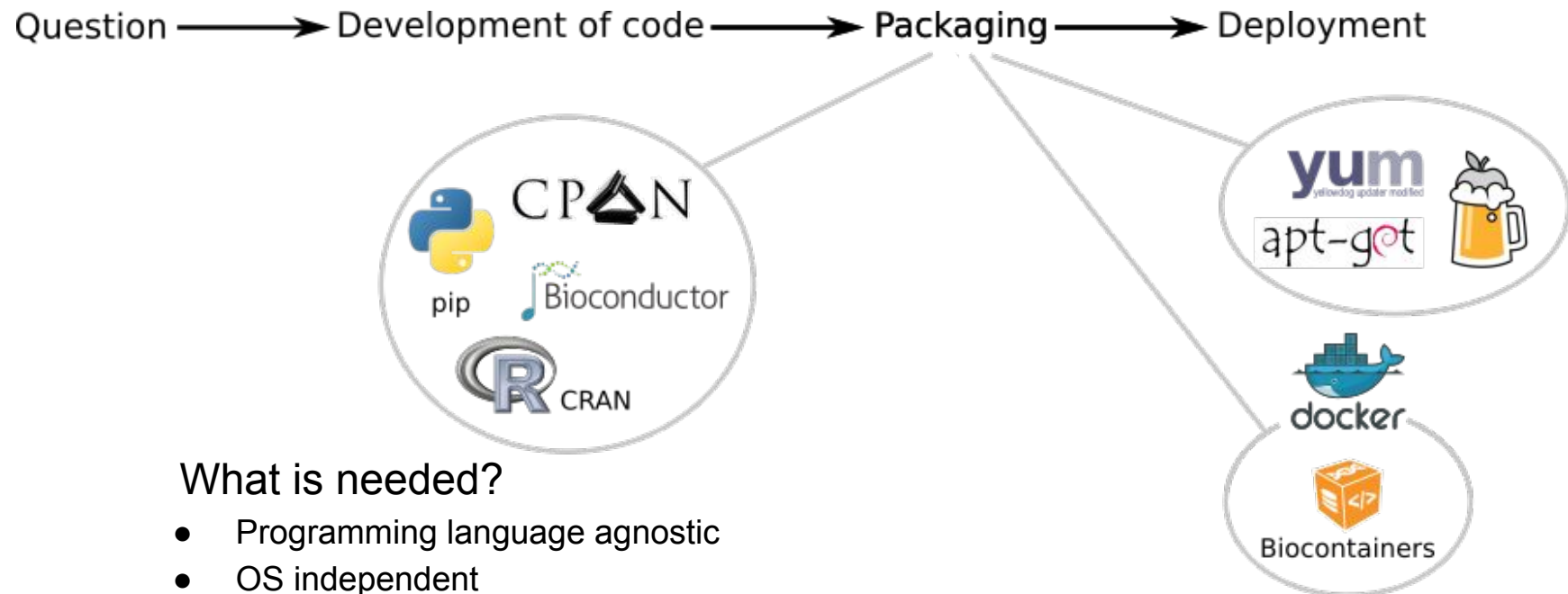
What is needed?
- Programming language agnostic
- OS independent
- No root privileges needed
- Management of multiple version
- HPC and Cloud compatible
- easy to maintain

elixir

CONDA

- Open Source package manager
- Independent of any programming language and OS
- Fast, robust and easy package installation

```
$ conda install deeptools
```

- Installation and management of multiple versions

elixir

CONDA

```yaml
package:
  name: deeptools
  version: '2.4.3'

source:
  fn: deepTools-2.4.3.tar.gz
  url: https://pypi.python.org/packages/40/4b/2f443a9d0d4201d3f152c3a9f291075f4b2cd5322d29c796/deepTools-2.4.3.tar.gz
  md5: 1751e860c5436d1973d7621ebdf00c6a

requirements:
  build:
    - python
    - setuptools
    - numpy >=1.9.0
    - scipy >=0.17.0
    - matplotlib >=1.4.0
    - pysam >=0.8.2
    - py2bit >=0.2.0
    - gcc
  run:
    - python
    - pybigwig >=0.2.3
    - numpy >=1.9.0
    - scipy >=0.17.0
    - matplotlib >=1.4.0
    - pysam >=0.8.2
    - py2bit >=0.2.0

test:
  imports:
    - deeptools

  commands:
    - bamCompare --version

about:
  home: https://github.com/fidelram/deepTools
  license: GPL3
  summary: A set of user-friendly tools for normalization and visualzation of deep-sequencing data
```

meta.yaml

```bash
#!/bin/bash

$PYTHON setup.py install
```

build.sh

eliXir

CONDA

- Build packages

```
$ conda build packages/deeptools/
```

- building a scientific community

BIOCONDA

- using a unified build environment

```
$ ./simulate-travis.py --docker --package deeptools
```

- joining other Conda communities **CONDA-FORGE**

https://bioconda.github.io | https://conda.io/docs

elixir

BIOCONDA®

Bioconda is a channel for the conda package manager specializing in bioinformatics software. Bioconda consists of:

- a repository of recipes hosted on GitHub
- a build system that turns these recipes into conda packages
- a repository of >1500 bioinformatics packages ready to use with `conda install`
- Over 130 contributors that add, modify, update and maintain the recipes

Browse packages in the bioconda channel: Available packages

The conda package manager has recently made installing software a vastly more streamlined process. Conda is a combination of other package managers you may have encountered, such as pip, CPAN, CRAN, Bioconductor, apt-get, and homebrew. Conda is both language- and OS-agnostic, and can be used to install C/C++, Fortran, Go, R, Python, Java etc programs on Linux, Mac OSX, and Windows.

Conda allows separation of packages into separate repositories, or *channels*. The main *defaults* channel has a large number of common packages. Users can add additional channels from which to install software packages not available in the defaults channel. Bioconda is one such channel specializing in bioinformatics software.

Using bioconda

bioconda supports only 64-bit Linux and Mac OSX.

1. Install conda

Bioconda requires the conda package manager to be installed. If you have an Anaconda Python installation, you already have it. Otherwise, the best way to install it is with the Miniconda package. The Python 3 version is recommended.

See also:

- What's the difference between Anaconda, conda, and Miniconda?
- The conda FAQs explain how it's easy to use with existing Python installations.

https://bioconda.github.io
https://conda.io/docs

elixir

BIOCONDA

BIOCONDA
Navigation

Contributing

Bioconda is completely dependent on contributors to add, update, and maintain recipes. Every little bit helps! Below are instructions for one-time setup as well as a general procedure to follow for each recipe you'd like to add.
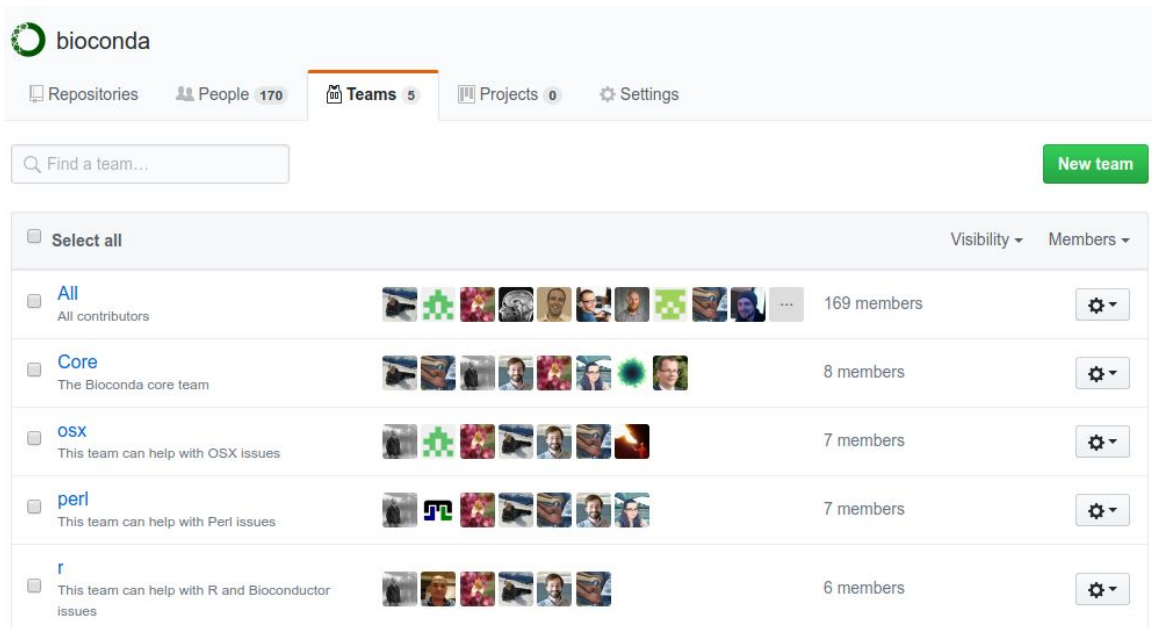
- One-time setup
  - Git and GitHub (one-time setup)
  - Install conda and Docker (one-time setup)
  - Request to be added to the bioconda team
- Contributing a recipe
  - Update repo and requirements
  - Write a recipe
  - Test locally
  - Push changes, wait for tests to pass, submit pull request
  - Use your new recipe
- Troubleshooting failed recipes
  - Reading bioconda-utils logs on Travis-CI
  - HTTP 500 errors
  - HTTP 404 errors
  - ZLIB errors
  - /usr/bin/perl or /usr/bin/python not found
  - Troubleshooting failed mulled-build tests
- Build system
- Guidelines for **bioconda** recipes
  - bioconda recipe checklist
  - Python
  - R (CRAN)
  - R (Bioconductor)
  - Java
  - Perl
  - C/C++
  - General command-line tools
  - Metapackages
  - Other examples of interest
  - Name collisions
  - Tests
  - Link and unlink scripts (pre- and post- install hooks)
  - Versions

https://bioconda.github.io
https://conda.io/docs

BIOCONDA®

- 2245 package (2017.03)
- contribute at *https://github.com/bioconda/bioconda-recipes*
- community driven

Community

208 Contributors   3700 merged PR
in 1.5 years

But the new cool kid is called **Containers**

But the new cool kid is called **Containers**

Question ⟶ Development of code ⟶ Packaging ⟶ Deployment

CONDA
BIOCONDA
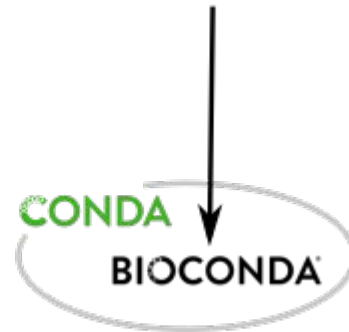
docker

Mulled
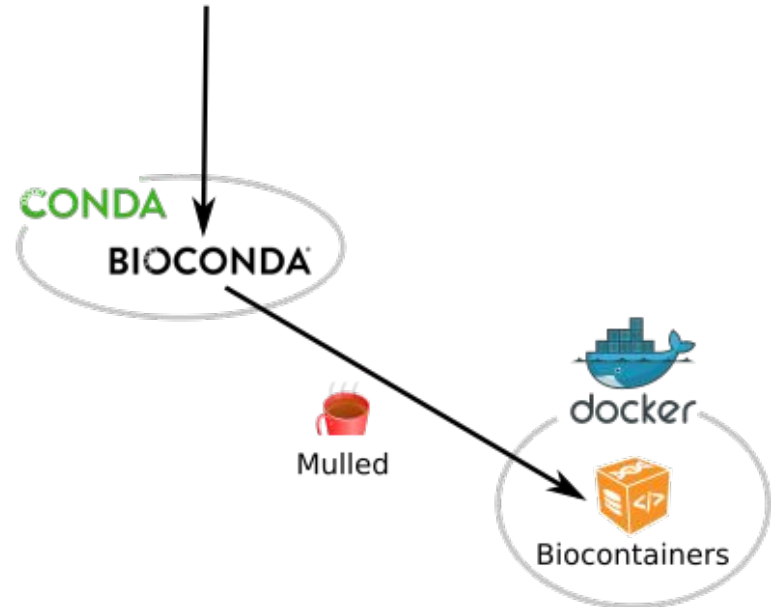
Biocontainers

<u>Mulled</u>
- A layer donning approach to build containers
- without Dockerfile
- completely automatised
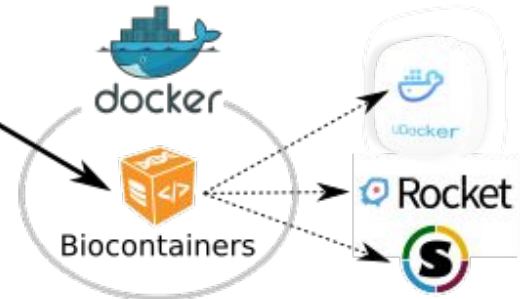- powered by Travis CI

elixir

Biocontainers

http://biocontainers.pro

- defining standards and best practises for containers
- developing training material
- developing documentation

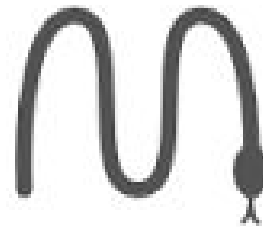# But the new cool kid is called **Containers**

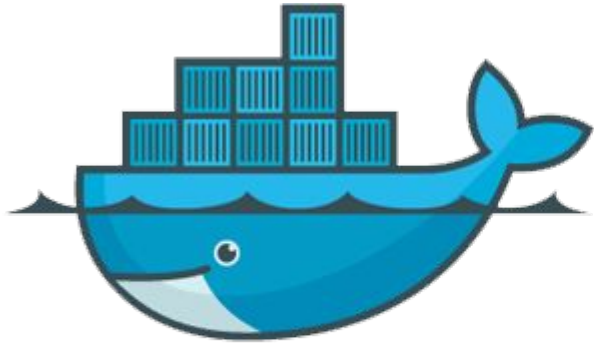Question ⟶ Development of code ⟶ Packaging ⟶ Deployment

CONDA
BIOCONDA

mulled

docker

Biocontainers

uDocker
Rocket
S

Get all the other nice technologies for free
- rkt
- singularity
- modules

eli𝑥ir

Who is using it?

Thanks!

Yasset Perez-Riverol
Felipe Leprevost ...

Rolf Backofen and the RBC

Johannes Köster
Ryan Dale ...