Resources at CRG: Software in the cluster

Luis Expósito

luis.exposito@crg.eu

Agenda

- Introduction
- Software provided by the OS
- Using modules (Easybuild)
- R packages

Introduction

- SIT: Gabriel González, Rodny Hernández, Luis Expósito
- Administration services and support to CRG scientific community.
- Cluster and storage documentation:
  - www.linux.crg.es
- Creating a ticket or service call:
  - Mail to: sit_support@request.crg.es
  - Intranet (Help/support? - Open a Service Call Ticket /Open a SIT Ticket (HPC and Storage))

Introduction

- Each group has its own software area where you can install your software: /software/$groupname/el$release
  For example: /software/as/el7.2

- It's a network file system directory and it's available from all the cluster nodes

- Use that directory and not your home /users/… , so it can be accessed by other users from your or other groups
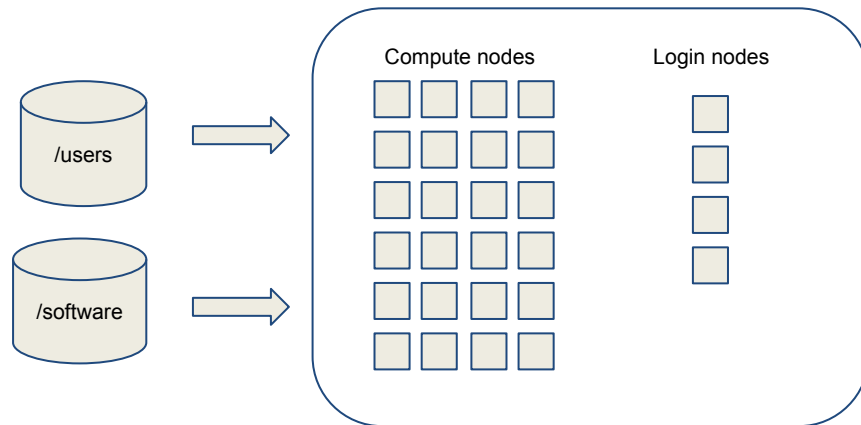
Agenda

- Introduction
- **Software provided by the OS**
- Using modules (Easybuild)
- R packages

Software provided by the OS

- The cluster runs Scientific Linux 7.2 + some extra repositories:
  - SL Base
  - SL Security
  - EPEL
  - REMI

- With this OS and this extra repositories we provide thousands of packages.

- We do periodically upgrades (once a month)

- Possible problem? Out-of-date (but stable) software

Software provided by the OS

- Looking for specific software? Look for in our repositories:

```
[lexposito@ant-login8 ~]$ yum list gcc
Loaded plugins: langpacks, priorities, protectbase
0 packages excluded due to repository protections
Installed Packages
gcc.x86_64                        4.8.5-4.el7                        @anaconda/7.2
```

- If the listed package fits your needs open a ticket and we'll install it in the system for you.

- No need to modify anything from your env.

Software provided by the OS

When you need a software that is not available in our repositories, or you need a newer version what can you do?

Software provided by the OS

When you need a software that is not available in our repositories, or you need a newer version what can you do?

- Install it yourself from source code

- Use some kind of installation management tool like:
    - Pythonz
    - Pyenv
    - Perlbrew
    - Anaconda
    - Easybuild

Agenda

- Introduction
- Software provided by the OS
- **Using modules (Easybuild)**
- R packages

EasyBuild - Introduction

- Compiling an application could be difficult so...

- EasyBuild is a software build and installation framework that allows you to install software in a structured, repeatable and robust way

- A software for installing other software:
    - In a specific path
    - Takes care of installing dependencies
    - Creates environment files with all the needed variables for your program that you easily load with modules

- It can install more than 500 packages. Newer versions of scientific software :
  https://github.com/hpcugent/easybuild/wiki/List-of-supported-software-packages

EasyBuild - Usage

First, load the easybuild:

module use /software/as/el7.2/EasyBuild/CRG/modules/all

module avail
module list
Module spider <module>
module load <module>
module unload <module>

EasyBuild - Usage

- You can print the list of modules:

```
[lexposito@ant-login8 ~]$ module avail
----------------------- /software/as/el7.2/EasyBuild/CRG/modules/all -----------------------
   Autoconf/2.69-foss-2016a
   Autoconf/2.69                                    (D)
   Automake/1.15-foss-2016a
   Automake/1.15                                    (D)
   Autotools/20150215-foss-2016a
   Autotools/20150215                               (D)
   BLAST/2.2.26-Linux_x86_64
…………
…………
```

EasyBuild - Usage

- You can search for an application:

```
[lexposito@ant-login8 ~]$ module spider BLAST

--------------------------------------------------------------------------------------------------
 BLAST: BLAST/2.2.26-Linux_x86_64
--------------------------------------------------------------------------------------------------
  Description:
   Basic Local Alignment Search Tool, or BLAST, is an algorithm for comparing primary biological
   sequence information, such as the amino-acid sequences of different proteins or the nucleotides
   of DNA sequences. - Homepage: http://blast.ncbi.nlm.nih.gov/

  This module can be loaded directly: module load BLAST/2.2.26-Linux_x86_64

  Help:
   Basic Local Alignment Search Tool, or BLAST, is an algorithm
   for comparing primary biological sequence information, such as the amino-acid
   sequences of different proteins or the nucleotides of DNA sequences.
      - Homepage: http://blast.ncbi.nlm.nih.gov/
```

easybuild

EasyBuild - Usage

```
[lexposito@ant-login8 ~]$ module load BLAST
```

```
[lexposito@ant-login8 ~]$ module avail
------------------------ /software/as/el7.2/EasyBuild/CRG/modules/all ----------------------
   Autoconf/2.69-foss-2016a
   Autoconf/2.69                                    (D)
   Automake/1.15-foss-2016a
   Automake/1.15                                    (D)
   Autotools/20150215-foss-2016a
   Autotools/20150215                               (D)
   BLAST/2.2.26-Linux_x86_64                        (L)
…………
```

```
[lexposito@ant-login8 ~]$ module list
Currently Loaded Modules:
  1) BLAST/2.2.26-Linux_x86_64
```
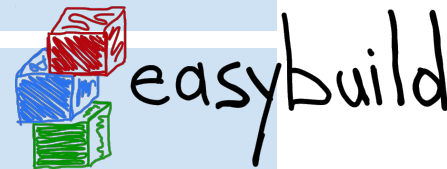
EasyBuild - Usage

- Check if the application is working:

[lexposito@ant-login8 ~]$ **blastall**
blastall 2.2.26   arguments:
  -p Program Name [String]
….

- With the command **which** we can check that we are using the blastall loaded by the easybuild:

[lexposito@ant-login8 ~]$ **which blastall**
/software/as/el7.2/EasyBuild/CRG/software/BLAST/2.2.26-Linux_x86_64/bin/blastall

EasyBuild - Usage

- You can unload the module:

[lexposito@ant-login8 ~]$ **module unload BLAST**
[lexposito@ant-login8 ~]$ **module list**
No modules loaded

Agenda

- Introduction
- Software provided by the OS
- Using modules (Easybuild)
- **R packages**

R packages

- The R version is the 3.3.2 (the one that comes with SL 7.2)
- We are providing a common repository for R packages (modules).
  - The repository is in /software/R/packages/
  - We maintain R repository on a best effort basis. (We are not R experts)
  - We install R packages following the repository option.

  - We only install packages using source code if ALL dependencies are satisfied and do not conflict with already installed packages from "official" repositories.

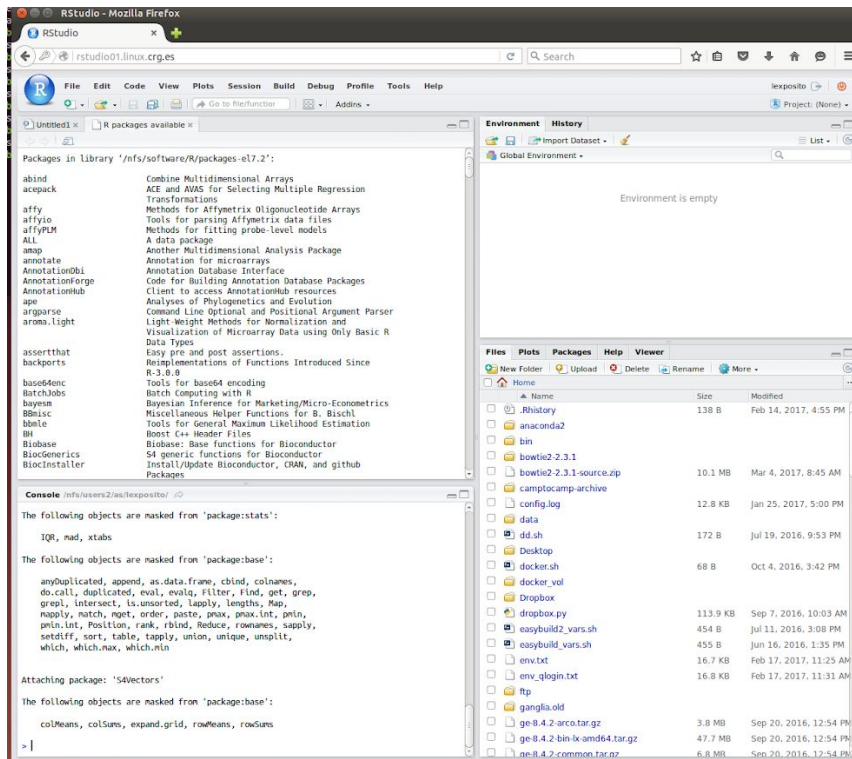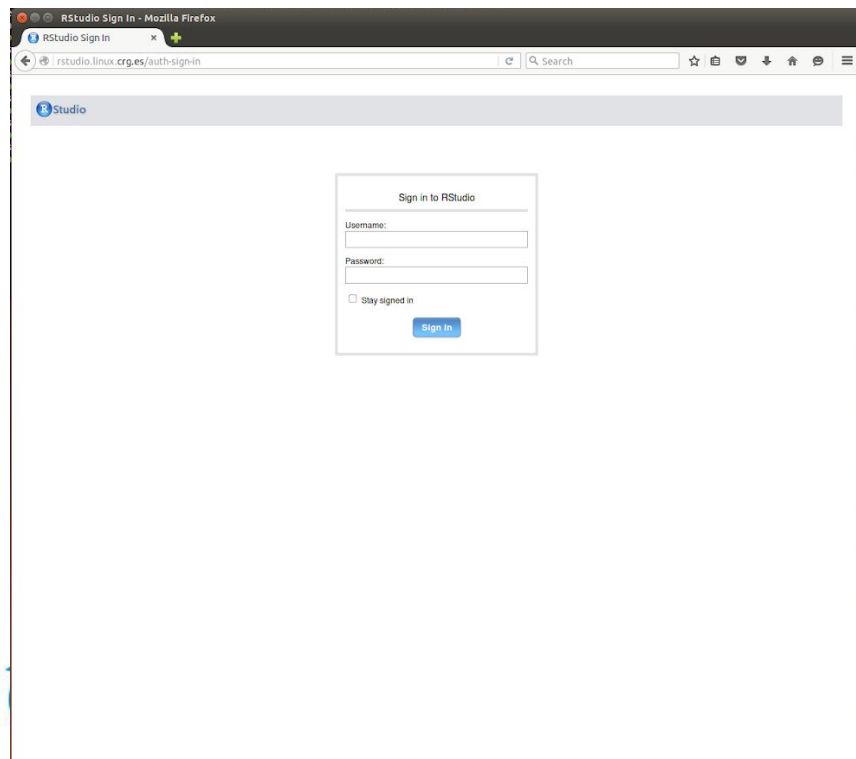- You can create your own R repository in your software area.

R packages

- If you want us to install a module open a RT ticket.

- R is updated when the OS updates it.

- Easy to use in your submit scripts:
  export R_LIBS=/software/R/packages/

- Dedicated server of rstudio: http://rstudio.linux.crg.es

R - Rstudio