

---

# Using Docker and Singularity

Lionel Spinelli - May 16th, 2017






## ■ A first problem : R

- R exists in several version from 2.x to 3.4
- Each version offers lot of packages
- Packages appear at a specific version and can be deprecated after a specific version
- Each scientific project requires a specific list of packages but accumulating packages on a single version can produce conflict
- Scientific projects often start with a chosen version of R that have to be upgraded during the project, with the risk to impact other project/user
- Some people work on several projects that uses different R version due to availability of packages or starting time
- On shared computers it is complex to have several R version/packages group installed

# Taking advantage of Docker hub



## Repositories (278)

		Stars	Pulls	Details
	<a href="#">rocker/rstudio</a> public   automated build	125 STARS	100K+ PULLS	<a href="#">&gt; DETAILS</a>
	<a href="#">opencpu/rstudio</a> public   automated build	22 STARS	50K+ PULLS	<a href="#">&gt; DETAILS</a>
	<a href="#">harmish/rstudio</a> public   automated build	0 STARS	589 PULLS	<a href="#">&gt; DETAILS</a>
	<a href="#">mobilizingcs/rstudio</a> public   automated build	0 STARS	10K+ PULLS	<a href="#">&gt; DETAILS</a>
	<a href="#">mccahill/rstudio</a> public   automated build	3 STARS	180 PULLS	<a href="#">&gt; DETAILS</a>

# Taking advantage of Docker hub

```
1 FROM ubuntu:16.04
2
3 MAINTAINER Spinelli Lionel (lionel.spinelli@univ-amu.fr)
4
5 # #####
6 # INSTALLING R FROM SOURCES
7 # #####
8
9 ARG R_VERSION
10 ARG BUILD_DATE
11 ENV BUILD_DATE ${BUILD_DATE:-}
12 ENV R_VERSION ${R_VERSION:-3.3.2}
13 ENV LC_ALL en_US.UTF-8
14 ENV LANG en_US.UTF-8
15 ENV TERM xterm
16
17 ## dependencies
18 RUN apt-get update \
19     && apt-get -y install python2.7-dev python-pip build-essential pkg-config libffi-dev libmagickwand-dev python-cffi
20
```

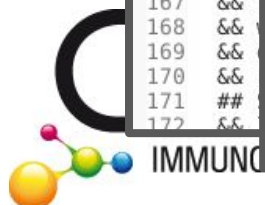
# Taking advantage of Docker hub

```
1 FROM ubuntu:16.04
2
3 MAINT 132 # #####
4 133# INSTALLING Latex for sweave compilation
5 # ### 134 # #####
6 # INS 135
7 # ### 136 RUN apt-get update \
8 137 && apt-get -y install texlive-full
9 ARG R 138
10 ARG E 139 # #####
11 ENV E 140 # INSTALLING RSTUDIO SERVER
12 ENV R 141 # #####
13 ENV L 142
14 ENV L 143 ARG RSTUDIO_VERSION
15 ENV T 144 ARG PANDOC_TEMPLATES_VERSION
16 145 ENV PANDOC_TEMPLATES_VERSION ${PANDOC_TEMPLATES_VERSION:-1.18}
17 ## de 146
18 RUN a 147 ## Add RStudio binaries to PATH
19 && 148 ENV PATH /usr/lib/rstudio-server/bin:$PATH
20 149
150 ## Download and install RStudio server & dependencies
151 ## Attempts to get detect latest version, otherwise falls back to version given in $VER
152 ## Symlink pandoc, pandoc-citeproc so they are available system-wide
153 RUN apt-get update \
154 && apt-get install -y --no-install-recommends \
155 file \
156 git \
157 libapparmor1 \
158 libcurl4-openssl-dev \
159 libedit2 \
160 libssl-dev \
161 lsb-release \
162 psmisc \
163 python-setuptools \
164 sudo \
165 wget \
166 && RSTUDIO_LATEST=$(wget --no-check-certificate -q0- https://s3.amazonaws.com/rstudio-server/current.ver) \
167 && [ -z "$RSTUDIO_VERSION" ] && RSTUDIO_VERSION=$RSTUDIO_LATEST || true \
168 && wget -q http://download2.rstudio.org/rstudio-server-${RSTUDIO_VERSION}-amd64.deb \
169 && dpkg -i rstudio-server-${RSTUDIO_VERSION}-amd64.deb \
170 && rm rstudio-server-*-amd64.deb \
171 ## Symlink pandoc & standard pandoc templates for use system-wide
172 && ln -s /usr/lib/rstudio-server/bin/pandoc/pandoc /usr/local/bin \
```

hon-cffi

# Taking advantage of Docker hub

```
1 FROM ubuntu:16.04
2
3 MAINT 132 # #####
4 133 # INSTALLING Latex for sweave compilation
5 # ### 134 # #####
6 # INS 135
7 # ### 136 RUN apt-get update \
8 137 && apt-get -y install texlive-full
9 ARG R 138
10 ARG E 139 # #####
11 ENV E 140 # INSTALLING RSTUDIO SERVER
12 ENV R 141 # #####
13 ENV L 142
14 ENV L 143 ARG R
15 ENV T 144 ARG P
16 145 ENV P
17 ## de 146
18 RUN a 147 ## Ad
19 && 148 ENV P
20 149 |
    150 ## Do
    151 ## At
    152 ## Sy
    153 RUN a
    154 &&
    155 f
    156 g
    157 l
    158 l
    159 l
    160 l
    161 l
    162 p
    163 p
    164 s
    165 w
    166 &&
    167 &&
    168 &&
    169 &&
    170 &&
    171 ##
    172 &&
    215 # #####
    216 # INSTALLING R PACKAGES FOR PROJECT
    217 # #####
    218
    219 RUN apt-get update \
    220 && export DEBIAN_FRONTEND=noninteractive \
    221 ## && apt-get -y install xorg openbox
    222 && apt-get -y install libx11-dev freeglut3 freeglut3-dev libpng-dev
    223 COPY R_package_install_V1.0.R /
    224 RUN su rstudio \
    225 && Rscript /R_package_install_V1.0.R
    226
    227 # #####
    228 # INSTALLING PDF VIEWER
    229 # #####
    230
    231 RUN apt-get update \
    232 && apt-get -y install evince
    233
    234 CMD ["/init"]
```

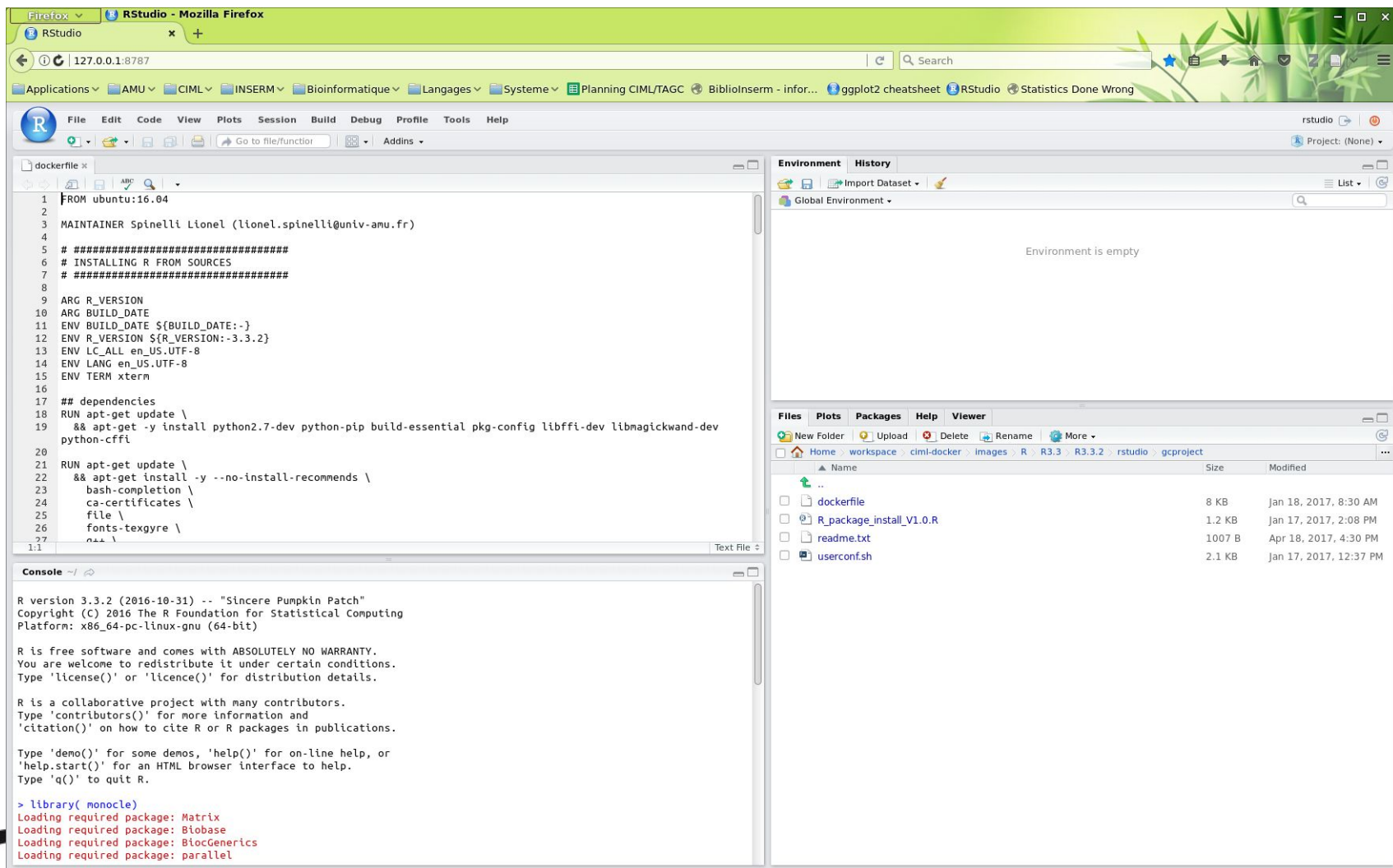


# Building our own Docker R + Rstudio image

```
1
2 This image contains:
3
4 - R 3.3.2
5 - Rstudio server (installation requires the userconf.sh file)
6 - Packages for the GCProject from Pierre Milpied. The packages are installed using the R_Package_install_VXX.R scripts
7
8
9 # #####
10    COMPILE THE IMAGE
11 # #####
12
13 docker build -t ciml/rstudio_gcproject ~/workspace/ciml-docker/images/R/R3.3/R3.3.2/rstudio/gcproject/
14
15 # #####
16    RUN THE IMAGE
17 # #####
18 |
19 docker run -d -p 8787:8787 -v /home/$USER/workspace:/home/rstudio/workspace --network=host -e USERID=$UID ciml/rstudio_gcproject
20
21 # #####
22    CONNECT TO RSTUDIO
23 # #####
24
25 In an Internet browser, type as url : http://127.0.0.1:8787 and use the login/password: rstudio/rstudio
26
27 # #####
28    NOTES
29 # #####
30
31 - To use knitr PDF compilation instead of Sweave, you have to go into Rstudio menu Tools->Global Options->Sweave->Weave Rnw files with.. and select "knitr".
32
```



# Using R + Rstudio Docker image



The screenshot displays the RStudio interface within a Docker container. The top panel shows the RStudio menu and toolbar. The main editor area contains a Dockerfile with the following content:

```
1 FROM ubuntu:16.04
2
3 MAINTAINER Spinelli Lionel (lionel.spinelli@univ-amu.fr)
4
5 #####
6 # INSTALLING R FROM SOURCES
7 # #####
8
9 ARG R_VERSION
10 ARG BUILD_DATE
11 ENV BUILD_DATE ${BUILD_DATE:-}
12 ENV R_VERSION ${R_VERSION:-3.3.2}
13 ENV LC_ALL en_US.UTF-8
14 ENV LANG en_US.UTF-8
15 ENV TERM xterm
16
17 ## dependencies
18 RUN apt-get update \
19     && apt-get -y install python2.7-dev python-pip build-essential pkg-config libffi-dev libmagickwand-dev
20     python-cffi
21
22 RUN apt-get update \
23     && apt-get install -y --no-install-recommends \
24     bash-completion \
25     ca-certificates \
26     file \
27     fonts-texgyre \
28     nss \
29     tzdata \
30     xz-utils
```

The Environment pane on the right shows "Global Environment" and "Environment is empty". The Files pane at the bottom right shows the directory structure:

Name	Size	Modified
..		
dockerfile	8 KB	Jan 18, 2017, 8:30 AM
R_package_install_V1.0.R	1.2 KB	Jan 17, 2017, 2:08 PM
readme.txt	1007 B	Apr 18, 2017, 4:30 PM
userconf.sh	2.1 KB	Jan 17, 2017, 12:37 PM

The Console pane at the bottom shows the R version and installed packages:

```
R version 3.3.2 (2016-10-31) -- "Sincere Pumpkin Patch"
Copyright (C) 2016 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> library(monocle)
Loading required package: Matrix
Loading required package: Biobase
Loading required package: BiocGenerics
Loading required package: parallel
```



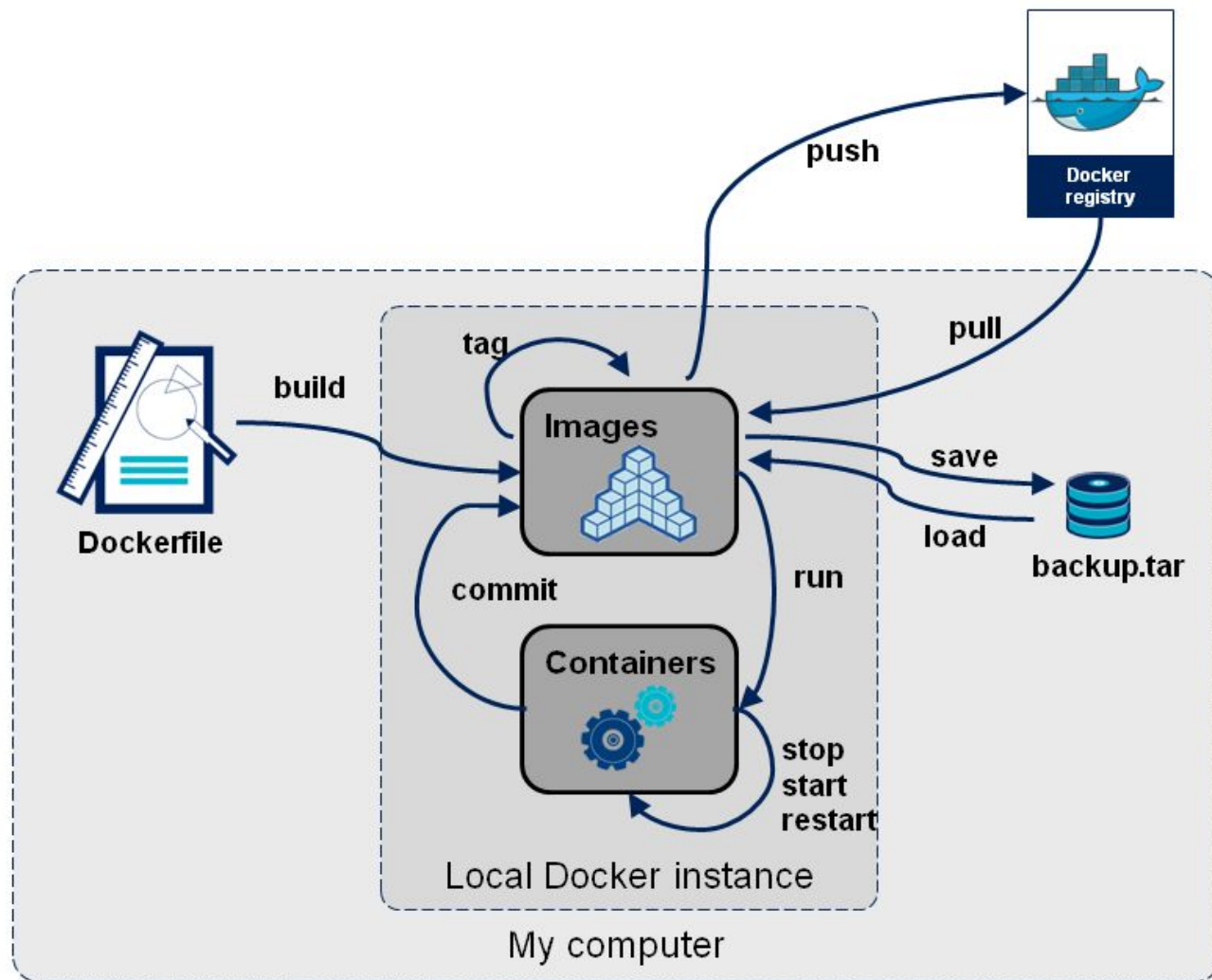
# Sharing to community : sharing sources

The screenshot shows the Eclipse IDE interface with the following components:

- Project Explorer:** Shows a project structure for 'ciml-docker' containing sub-projects for 'doc', 'images', 'R', and 'R3.3.2'. The 'R3.3.2' sub-project contains 'rstudio' and 'gcproject' sub-projects. The 'gcproject' sub-project contains 'dockerfile', 'R\_package\_install\_V1.0.R', 'readme.txt', and 'userconf.sh'.
- Editor:** Displays the 'readme.txt' file with the following content:

```
1 This image contains:
2
3
4 - R 3.3.2
5 - Rstudio server (installation requires the userconf.sh file)
6 - Packages for the GCProject from Pierre Milpied. The packages are installed using the R_Package_install_VXX.R scripts
7
8
9 # #####
10 COMPILE THE IMAGE
11 # #####
12
13 docker build -t ciml/rstudio_gcproject ~/workspace/ciml-docker/images/R/R3.3/R3.3.2/rstudio/gcproject/
14
15 # #####
16 RUN THE IMAGE
17 # #####
18
19 docker run -d -p 8787:8787 -v /home/$USER/workspace:/home/rstudio/workspace --network=host -e USERID=$UID ciml/rstudio_gcproject
20
21 # #####
22 CONNECT TO RSTUDIO
23 # #####
24
25 In an Internet browser, type as url : http://127.0.0.1:8787 and use the login/password: rstudio/rstudio
26
27 # #####
28 NOTES
29 # #####
30
31 - To use knitr PDF compilation instead of Sweave, you have to go into Rstudio menu Tools->Global Options->Sweave->Weave Rnw files v
32
33
34
```
- Console:** Shows 'No consoles to display at this time.'
- Problems:** Shows 'No problems to display at this time.'
- Target Platform State:** Shows 'No target platform state to display at this time.'

# Sharing to community : sharing images



## ■ A second problem : sudo rights

- Docker requires sudo rights to execute
- Inside the container, the user is sudo
- Not a problem on local machine
- Problem on shared machine
- Problem on clusters (local or mésocentre)

## A solution : singularity



- Same idea of Docker : images and containers
- Same functionalities than Docker
- Somehow less confined
- No need to be sudo to run containers
- User inside the container is the same as user launching it
- Singularity image is a simple file : even easier to share than docker
  
- **Directly building Singularity image seems more difficult**
- But very easy way to convert docker container to singularity image

# Importing Singularity R + Rstudio image

```
1 |
2 # =====
3 #   Convert Docker container to Singularity
4 # =====
5
6 # The name of the Docker image
7 DOCKER_IMAGE=image_name
8 # The size in MiB
9 SIZE=3600
10
11 # Create the Singularity file
12 sudo singularity create --size $SIZE $DOCKER_IMAGE.img
13
14 # Start the Docker container
15 sudo docker run --name $DOCKER_IMAGE -d $DOCKER_IMAGE sleep 1800
16
17 # Convert the container to Singularity
18 sudo docker export $DOCKER_IMAGE | sudo singularity import $DOCKER_IMAGE.img
19
20
21 # =====
22 #   Running a singularity image
23 # =====
24
25 # Launch a single command inside the image
26 singularity exec <image>.img <command to execute>
27
28 # Launch an interactive shell
29 singularity shell <image>.img
30
31 # Launch an predetermined command into the image
32 singularity run <image>.img
```

## Usage of singularity at AMU Mésocentre

```
lspinelli@euclide:~$
lspinelli@euclide:~$ ssh lspinelli@login.ccamu.u-3mrs.fr
lspinelli@login.ccamu.u-3mrs.fr's password:
Last login: Thu May  4 11:03:11 2017 from ciml-005.univ-mrs.fr

              ( )
            /  \
           /    \
          /      \
         /        \
        /          \
       /            \
      /              \
     /                \
    /                  \
   /                    \
  /                      \
 /                        \
/                          \
 \                          /
  \                        /
   \                      /
    \                    /
     \                  /
      \                /
       \              /
        \            /
         \          /
          \        /
           \      /
            \    /
             \  /
              \ /

INFORMATIONS : - Site Internet : https://mesocentre.univ-amu.fr
               - Tous les services fonctionnent normalement

*-----*
| On project a005: 784.4/5000 (15.7%) hours have been consumed
| You are using  0/4882 MB ( 0%) on /home
| You are using 0.26/9.00 TB ( 2%) on /scratch
*-----*

[lspinelli@login ~]$ cd /scratch/lspinelli/
[lspinelli@login lspinelli]$ ls -al
total 256
drwxr-x---  6 lspinelli a005          512 May  3 11:20 .
drwxr-xr-x 241 root      root        131072 May  3 13:53 ..
drwxrwxr-x  7 lspinelli lspinelli  131072 May  4 18:34 EVlab
drwxrwxr-x  4 lspinelli a005          512 May  4 10:59 Reference
drwxrwxr-x  2 lspinelli a005          512 May  3 11:20 Singularity
drwxrwxr-x  3 lspinelli a005          512 May  3 11:35 Software
[lspinelli@login lspinelli]$ ls Singularity/
cellranger1.3.1.img
[lspinelli@login lspinelli]$ █
```

## Usage of singularity at AMU Mésocentre

```
GNU nano 2.0.9 File: evlab_s1_cellranger_singularity.oar
#!/bin/bash
#OAR -n evlab_s1_cellranger_singularity_4cores
#OAR -l nodes=1/core=4,walltime=30:00:00
#OAR -p smp AND nodetype like 'SMP512Gb'
#OAR -O evlab_s1_cellranger_singularity_4bcores.%jobid%.out
#OAR -E evlab_s1_cellranger_singularity_4bcores.%jobid%.out
#OAR -p singularity='YES'

singularity exec /scratch/lspinelli/Singularity/cellranger1.3.1.img /scratch/lspinelli/Software/launch_cellranger.sh /scratch/lspinelli/EVlab/s1 /$
```



# Usage of singularity at AMU Mésocentre

```
[lspinelli@login EVlab]$ oarsub -S ./evlab_s1_cellranger_singularity.oar
[ADMISSION RULE] Modify resource description with type constraints
[MEDIUM QUEUE] This job is routed into the medium queue
[JOB PROJECT] No project selected using a005
OAR_JOB_ID=692841
[lspinelli@login EVlab]$ oarstat -fj 692841
Job_Id: 692841
  job_array_id = 692841
  job_array_index = 1
  name = evlab_s1_cellranger_singularity_4cores
  project = a005
  owner = lspinelli
  state = Waiting
  wanted_resources = -l "{type = 'default'}/host=1/core=4,walltime=30:0:0"
  types =
  dependencies =
  assigned_resources =
  assigned_hostnames =
  queue = medium
  command = ./evlab_s1_cellranger_singularity.oar
  launchingDirectory = /scratch/lspinelli/EVlab
  stdout_file = evlab_s1_cellranger_singularity_4bcores.692841.out
  stderr_file = evlab_s1_cellranger_singularity_4bcores.692841.out
  jobType = PASSIVE
  properties = ((singularity='YES') AND cluster='YES') AND shortnode = 'NO') AND drain='NO'
  reservation = None
  walltime = 30:0:0
  submissionTime = 2017-05-15 09:20:19
  cpuset name = lspinelli 692841
  initial_request = oarsub -S ./evlab_s1_cellranger_singularity.oar; #OAR -n evlab_s1_cellranger_singularity_4cores; #OAR -l nodes=1/core=4,wallt
ime=30:00:00; #OAR -p smp AND nodetype like 'SMP512Gb'; #OAR -O evlab_s1_cellranger_singularity_4bcores.%jobid%.out; #OAR -E evlab_s1_cellranger_si
ngularity_4bcores.%jobid%.out; #OAR -p singularity='YES'
  message = R=4,W=30:0:0,J=B,N=evlab_s1_cellranger_singularity_4cores,Q=medium,P=a005 (Karma=0.003,quota_ok)
  scheduledStart = 2017-05-15 09:42:06
  resubmit_job_id = 0
  events =
```

# Conclusion

GO docker!

simple to build, simple to use, quite simple to share

GO singularity!

compatible with Docker, simple to use, extra-simple to share,  
usable on AMU mésocentre

!!! Let's share between institutes !!!