Setting up your R environment

Thomas Hauser
Director of Research Computing
thomas.hauser@colorado.edu

https://github.com/ResearchComputing/Parallelization_Workshop/tree/master/Day4-Parallel_R

Starting R on RMACC Summit

- Convention on my slides
  - `$` is bash prompt
  - `>` is R prompt
- Login to the RC login nodes and then a compile node on RMACC Summit

```
$ ssh yourIdentikey@tutorial-login.rc.colorado.edu
$ ssh scompile
```

- Load the appropriate module

```
$ module load R
$ module list
```

- Which modules are loaded in your environment?

R Startup Files

- `.Renviron`
  - Setting up environment variables
- `.Rprofile`
  - Lines of R code every time R starts up
- First `.Renviron` and then `.Rprofile`

Location of R Startup Files

- Three folders
  - R_HOME – where R is installed
    ```
    > R.home()
    ```
  - HOME – User home directory
    ```
    > Sys.getenv("HOME")
    ```
  - R's current working directory
    ```
    > getwd()
    ```

Short Quiz – R environment

- What is RHOME on RMACC Summit?

R package installation

- Functionality can be added through additional packages
- Packages should be installed in your project directory

```
$ ls /projects/$USER
```

- R package path (2 ways to do it)
  1. Setting in .Renviron
     - Recommendation: /project/$USER/R

     ```
     R_LIBS=/projects/YOURIDENTIKEY/R
     ```
  2. Better in .Rprofile

     ```
     sinfo <- Sys.info()
     user <- sinfo["user"]
     userpath <- paste("/projects/", user, "/R", sep="")
     .libPaths(c(userpath, .libPaths()))
     ```
- R puts user package in first element of `.libPath`

Setup your .Rprofile

- What's your .libPaths()?
- Create the directory for your R packages

```
$ mkdir /projects/$USER/R
```

- Add the following lines to your .Rprofile in $HOME

```
sinfo <- Sys.info()
user <- sinfo["user"]
userpath <- paste("/projects/", user, "/R", sep="")
.libPaths(c(userpath, .libPaths())
```

- What's your .libPaths() now?

Installing R packages

- Quick hands-on Exercise
  ```
  > sessionInfo()
  ```
- To install packages use
  ```
  > install.packages("ggplot2")
  ```
- Use the package
  ```
  > library(ggplot2)
  > sessionInfo()
  ```
- Test the library
  ```
  > pdf("diamonds.pdf")
  > ggplot(diamonds, aes(x=carat, y=price)) + geom_point()
  > dev.off()
  ```

Install Rmpi

```
#!/bin/bash
module purge
module load R
module load openmpi/1.10.2


wget https://cran.r-project.org/src/contrib/Rmpi_0.6-6.tar.gz


R CMD INSTALL Rmpi_0.6-6.tar.gz --configure-args=" \
--with-Rmpi-include=$CURC_OPENMPI_INC \
--with-Rmpi-libpath=$CURC_OPENMPI_LIB \
--with-Rmpi-type=OPENMPI" \
--library=/projects/$USER/R --no-test-load
```

- Run the following in the workshop directory

```
$ cd $HOME/Parallelization_Workshop/Day4-Parallel_R
$ bash ./installRmpi.sh
```

Install pbdMPI

```
#!/bin/bash
module purge
module load R
module load openmpi/1.10.2


wget https://cran.r-project.org/src/contrib/pbdMPI_0.3-3.tar.gz


R CMD INSTALL pbdMPI_0.3-3.tar.gz --configure-args=" \
  --with-mpi-type=OPENMPI \
  --with-mpi-include=$CURC_OPENMPI_INC \
  --with-mpi-libpath=$CURC_OPENMPI_LIB" \
  --library=/projects/$USER/R --no-test-load
```

- Run the following in the workshop directory

```
$ cd $HOME/Parallelization_Workshop/Day4-Parallel_R
$ bash ./installpbdMPI.sh
```

References

- Efficient R programming
  - Colin Gillespie
  - Robin Lovelace
  - *2017-04-10*
  - https://csgillespie.github.io/efficientR/

Questions?

- Instructor: Thomas Hauser
- Email rc-help@colorado.edu
- Twitter: CUBoulderRC

- Link to survey on topic R-setup:
  - http://tinyurl.com/curc-survey16

- R setup
- Slides: https://github.com/ResearchComputing/Parallelization_Workshop/tree/master/Day1

License