R on HPC

NUS-HTPC

Module module avail R

GNU C++	Intel Optimized
R-3.0.2	R-3.2.3_intel
R-3.1.2	R-3.4.3_intel
R-3.2.5	
R-3.5.0	
R-3.6.0	

Install R packages

Load R module on the login node

$ module load R -3.5.0

Launch R

$ R

Install packages

$ install.packages("dplyr")

- Choose 0-Cloud [https] mirror for CRAN

Custom r environment

Use miniconda to create your own R conda environment

1) Load miniconda environment in login node

```
$ module load miniconda
```

1) Using conda create a conda environment

```
$ conda create   -n myenv   -c conda -forge r  -tidyverse
```

1) Activate conda env and use R

```
$ source activate myenv
$ R
```

R serial job script

```
#!/bin/bash

#PBS -P Project_Name_of_Job
#PBS -j oe
#PBS -N Job_Name

#PBS -q serial
#PBS -l select=1:ncpus=1:mem=10GB

source /etc/profile.d/rec_modules.sh
module load R  -3.2.5

Rscript mycode.R
```

```
#!/bin/bash

#PBS -P Project_Name_of_Job
#PBS -j oe
#PBS -N Job_Name

#PBS -q serial
#PBS -l select=1:ncpus=1:mem=10GB

source /etc/profile.d/rec_modules.sh
module load miniconda
Source activate myenv

Rscript mycode.R
```

Run R in parallel

Mostly can run multiprocessing in single node

parallel

foreach

doSNOW

Multi node parallelization require MPI

Rmpi

pdbMPI

https://bookdown.org/rdpeng/rprogdatascience/parallel -computation.html

R- Parallel job script

```
#!/bin/bash

#PBS -P Project_Name_of_Job
#PBS -j oe
#PBS -N Job_Name

#PBS -q parallel24
#PBS -l select=1:ncpus=24:mem=50GB

source /etc/profile.d/rec_modules.sh
module load R  -3.2.5

Rscript myparallelcode.R
```

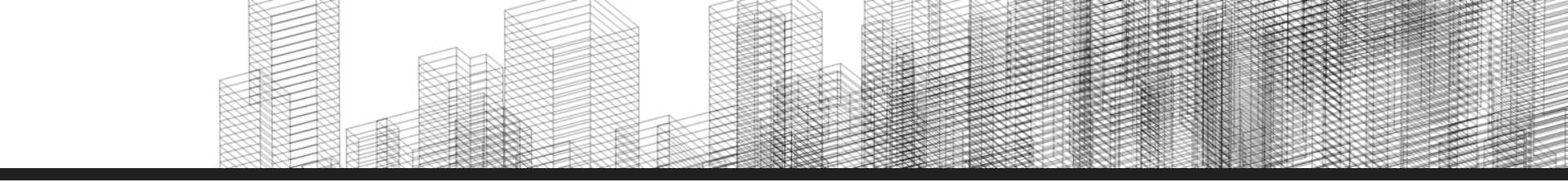
```
#!/bin/bash

#PBS -P Project_Name_of_Job
#PBS -j oe
#PBS -N Job_Name

#PBS -q parallel24
#PBS -l select=1:ncpus=24:mem=50GB

source /etc/profile.d/rec_modules.sh
module load miniconda
Source activate myenv

Rscript myparallelcode.R
```



Additional help

nTouch

https://ntouch.nus.edu.sg/

Search for HPC Enquiries