FALL 2018

TEXAS
The University of Texas at Austin

APPLIED ENGINEERING DATA ANALYSIS, OPTIMIZATION AND VISUALIZATION

Getting started with TACC

JOSHUA RHODES, PHD
Research Fellow/Adjunct Professor, The University of Texas at Austin

TACC is the Texas Advanced Computing Center

- Started in 1960s with a CRAY CDC 6600
  - Your phone is more powerful than this thing now
- TACC became TACC in 2001
- Mostly housed at the PRC campus
  - Visualization lab on Main Campus

TACC is available for you as part of UT

- Computing is free
- Your advisor needs to get a project set up
- TACC can be brought on for research projects and as computing support for grants
- Deeper dives into using TACC, by TACC
  - https://learn.tacc.utexas.edu/

Logging into TACC

- From Terminal
  - ssh <username>@maverick.tacc.utexas.edu
  - Password
  - 2-factor authentication

The first thing you see is your account balance

```
--------------------------------------------------------------------------------
---------------------- Project balances for user joshdr ----------------------
| Name           Avail SUs     Expires | Name           Avail SUs     Expires |
| ME397M-Applied-Engin      75000  2018-06-30 | WebberEnergy        4497  2018-03-31 |
------------------------- Disk quotas for user joshdr -------------------------
| Disk         Usage (GB)     Limit    %Used   File Usage       Limit   %Used |
| /home             1.0      10.0    10.28         4722      750100    0.63 |
| /work           696.7    1024.0    68.03       123818     3000000    4.13 |
--------------------------------------------------------------------------------
```

$Home vs. $Work directory

- $Home: where you run programs from
  - 10GB storage
- $Work: where you store your data
  - 1TB storage
- You can access your data in work from home
  - i.e. function running in $home can read file in $work

You land on your home folder

- home/XXXXX/<tacc_username>
- The same commands we learned in Lecture 2:Terminal apply here
  - cd, mkdir, ls, rm, rmdir, etc.

When you first log on, you are on a login node

- YOU CANNOT RUN PROGRAMS ON THE LOGIN NODE – YOU WILL GET KICKED OFF
- You can make folders, move files, edit file, install programs, queue files to run

There is a way to get a compute node all to yourself

- You can create an interactive development session
  - idev –m 90
  - -m 90 = get computer for 90 min, default = 30 min
    - Will kick you off and not save anything at end of time!!

So now that I am in, what do I do?

- You have to load programs (modules) before you can use them
  - Like double-clicking on an icon
- To check what you can run
  - module avail

What modules are available?: module avail

```
login1.maverick(1008)$ module avail

------------------------------------------------------ /opt/apps/intel14/mvapich2_2_0/modulefiles ------------------------------------------------------
   Rstats/3.0.3          petsc/3.5-complexdebug    petsc/3.5-cxxcomplexdebug    petsc/3.5-uni             python/2.7.6
   paraview/4.3.1        petsc/3.5-cxx             petsc/3.5-cxxdebug           petsc/3.5-unidebug        visit/2.7
   petsc/3.5-complex     petsc/3.5-cxxcomplex      petsc/3.5-debug              petsc/3.5          (D)    visit/2.9    (D)

---------------------------------------------------------- /opt/apps/intel14/modulefiles -----------------------------------------------------------
   boost/1.51.0    gsl/1.16    hdf5/1.8.12 (m)    mvapich2/2.0b (L)    nco/4.5.4    netcdf/4.2.1.1    udunits/2.2.19

------------------------------------------------------------ /opt/apps/xsede/modulefiles -------------------------------------------------------------
   CTSSV4            TERAGRID-paths      cue-comm         globus-5.0        jdk32/1.7.0       tginfo/1.1.3     xdusage/1.0
   GLOBUS-5.0        apache-ant/1.6.5    cue-login-env    globus/5.2.5      pacman/3.29       tgproxy/0.9.1
   TERAGRID-BASIC    ctssv4              cue-math         gsissh/4.3        teragrid-basic    tgresid/2.3.4
   TERAGRID-DEV      cue-build           cue-tg           gx-map/0.5.3.3    teragrid-dev      uberftp/2.6

---------------------------------------------------------- /opt/apps/modulefiles -----------------------------------------------------------
   Rstudio/0.98.501        cuda/7.5              gcc/5.4.0                lmod/7.7.1            pylauncher/2.1            vapor/2.5.0
   Rstudio/1.0.153  (D)   cuda/8.0     (g,D)    git/2.7.0                luatools/1.1          qt/4.8.4                  vtk/6.1.0
   amira/6.0.1            cudnn/4.0             idev/1.0                 matlab/2013a          remora/1.7                xalt/1.8    (L)
   autotools/1.0          ddt/5.0.1             idl/7.0.6                matlab/2015a          sanitytool/1.3
   cmake/2.8.12.2         ffmpeg/2.1.4          idl/8.4           (D)    matlab/2017a    (D)   settarg/7.7.1
   cmake/3.7.1      (D)   gcc/4.7.1             intel/14.0.1.106  (L,D)  mcr/v81               swr/17.0
   cuda/6.5               gcc/4.9.1      (D)    intel/15.0.3             ncl_ncarg/6.3.0       tacc-singularity/2.3.1
   cuda/7.0               gcc/4.9.3             launcher/2.0             ospray/1.3.1          tacc_tips/0.5

------------------------------------------------------------ /opt/modulefiles ------------------------------------------------------------
```

Is that it?

- Check for ALL installed programs
  - module spider
- Check for installed programs that start with R
  - module spider R
  - >

We have multiple versions of R installed and available, v3.4.0 has dependencies (so it d/n show up earlier)

```
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
 Rstats:
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
    Versions:
       Rstats/3.0.3
       Rstats/3.4.0

-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
 For detailed information about a specific "Rstats" module (including how to load the modules) use the module's full name.
 For example:

    $ module spider Rstats/3.4.0
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------


-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
 RstatsPackages: RstatsPackages/3.4.0
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

   You will need to load all module(s) on any one of the lines below before the "RstatsPackages/3.4.0" module is available to load.

     intel/15.0.3  mvapich2/2.1

   Help:
     This is the R statistics (RstatsPackages) package built on October 04, 2017.

     It includes the following accessory packages including but not limited to:
     Rmpi, snow, snowfall
     pdbMPI, pbdSLAP, pbdBASE, pbdDMAT, pbdDEMO, pbdNCDF4, pmclust
     multicore
     doMC, doSNOW, doMPI, doParallel
     BH, bigmemory, biganalytics, bigtabulate, synchronicity
     Rdsm, SparseM, slam, cluster, randomForest, bit, ff, mchof
     BioConductor (base installation plus some common packages)
     ggplot2, rjags/r2jags, rgdal, rstan

     The RstatsPackages modulefile extends the PATH and LD_LIBRARY_PATH paths as appropriate.

     Version 3.4.0
```

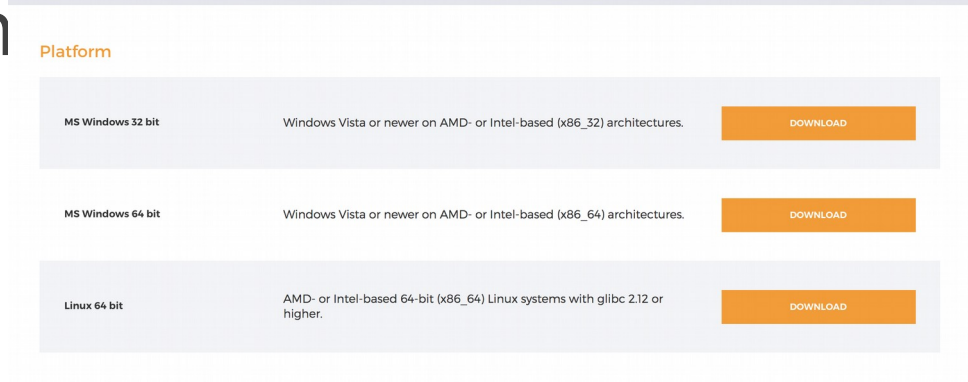
For the packages we used to build our R homework, we need R v3.4.0

- We must load the dependencies first

```
login1.maverick(1003)$ module load intel/15.0.3  mvapich2/2.1

The following have been reloaded with a version change:
  1) intel/14.0.1.106 => intel/15.0.3     2) mvapich2/2.0b => mvapich2/2.1

login1.maverick(1004)$ module avail

------------------------------------------------------ /opt/apps/intel15/mvapich2_2_1/modulefiles -------------------------------------------------------
   Rstats/3.4.0            amber/14.0     fftw3/3.3.4               paraview/5.4.1    python/2.7.9    vtk/7.0.0 (D)
   RstatsPackages/3.4.0    fftw2/2.1.5    parallel-netcdf/4.3.3.1   phdf5/1.8.16      visit/2.12

---------------------------------------------------------- /opt/apps/intel15/modulefiles ----------------------------------------------------------
   cxx11/4.9.1    gsl/1.16    hdf5/1.8.16    impi/5.0.3    mvapich2/2.1 (L)    netcdf/4.3.3.1    python/2.7.9    python/2.7.13 (D)

------------------------------------------------------------ /opt/apps/xsede/modulefiles ------------------------------------------------------------
   CTSSV4            TERAGRID-paths      cue-comm         globus-5.0        jdk32/1.7.0       tginfo/1.1.3     xdusage/1.0
   GLOBUS-5.0        apache-ant/1.6.5    cue-login-env    globus/5.2.5      pacman/3.29       tgproxy/0.9.1
   TERAGRID-BASIC    ctssv4              cue-math         gsissh/4.3        teragrid-basic    tgresid/2.3.4
   TERAGRID-DEV      cue-build           cue-tg           gx-map/0.5.3.3    teragrid-dev      uberftp/2.6
```

Installing your own software

- Here we will install GAMS
  - https://www.gams.com/24.8/docs/userguides/userguide/_u_g__u_n_i_x__i_n



You can create shortcuts for yourself

- alias gams=~/opt/gams/gams25.0_linux_x64_64_sfx/gams
- alias gamslib=~/opt/gams/gams25.0_linux_x64_64_sfx/gamslib

To use TACC for more than 90 minutes, you will need to submit a job to the queue

- There is a sequence to get in the queue
  - Shell script
  - R script

Aside: you can use VIM to make some quick edits

You can use a utility like CyberDuck to make life easier