

5 - HPC quick start

Robert Settlage

2020-11-05

Here are some minimal steps to get on ARC clusters and validate your account is setup properly.

VPN

If off VT's Blacksburg (and soon Roanoke) campus, you must have the VPN going. Without getting too geeky, this basically makes it appear that you are on campus by routing your internet traffic through VT servers. In case you are wondering, most of the articles you are wanting to read that are looking for a subscription will magically show up with your VPN on.

https://vt4help.service-now.com/sp?id=kb_article&sys_id=d5496fca0f8b4200d3254b9ce1050ee5

ARC user name and password

I have set this up for you. Please user your pid and password like you were logging into Canvas.

Access the clusters

We are going to use Open OnDemand (OOD). Via your browser (Chrome or Firefox), go to: <https://ood.arc.vt.edu/pun/sys/dashboard>.

Validate Rstudio on ARC works for you

From the OOD menus, choose "Interactive Apps" and Rstudio. Fill out the form to match the below. We are using `partition=normal_q` and `Account=stat5526-fall2020`.

Click launch, wait for the Rstudio button to show up blue. Click on it to launch Rstudio. Create a new Rmarkdown file, choose the PDF option, and just knit the default document to PDF. If the PDF pops up, things are working.

Validate you can start a job from the command line

Here we are going to create a script we will submit to the scheduler manually.

In Rstudio, choose file, new, text file. Copy and paste the code at the end of this document in that file and save it as `SLURM_basic_job.sh`.

Now, back in OOD, choose the menu item "Clusters", then "TinkerCliffs Shell Access". In that window, type:

```
SBATCH SLURM_basic_job.sh
```

Hopefully it will say something like job ### submitted where the ### number is the assigned job id. If it does, you can watch its progress by typing:

```
squeue and looking for you and that job id.
```

Rstudio

This app will launch a Rstudio Server in a container on one node.

R version and package set

R: 4.0.0 keras

This defines the R version and basic packages installed. You can install more in your session.

Account

stat5526-fall2020|

- The allocation you would like to use for SLURM.

Reservation

Partition

normal_q

- To request a GPU enabled queue, preface it with v100_. Example:
v100_normal_q

Number of hours (min-1, max-48)

4

Number of nodes (min-1, max-2)

1

Number of cores per node (min-1, max-128)

128

☐ I would like to receive an email when the session starts

Launch

Figure 1: OOD setup

When the job no longer shows up in the squeue list, you can look for the output in the current directory by typing:

`ls -l` and looking for a file named `slurm-###.out`. You can view that output by opening it via the Rstudio interface OR by simply typing `cat slurm-###.out`.

```
#!/bin/bash

#####
##
## sbatch SLURM_basic_job.sh
##
#####

##### job customization
#SBATCH --job-name=R_test
#SBATCH -N 1
#SBATCH -c 8                #this sets SLURM_CPUS_ON_NODE
#SBATCH -t 00:15:00
#SBATCH -p normal_q
#SBATCH -A stat5526-fall2020
###SBATCH --mail-user=rsettlag@vt.edu #uncomment and change if you want emails
#SBATCH --mail-type=FAIL

#####
date
cluster=${SLURM_CLUSTER_NAME}
echo running on $cluster cluster in node `hostname`
echo in job $SLURM_JOB_ID
echo with tmpfs $TMPFS
echo with tmpdir $TMPDIR

## starting the script, need day
start=$(date +%s)
today=$(date '+%D %T')
echo $today

#####

module load containers/singularity
singularity exec --bind $TMPFS:/tmp,/sys/fs/cgroup/cpuset/slurm/uid_$(id -u):/job_sets \
    /projects/arcsingularity/ood-rstudio-basic_4.0.0.sif Rscript -e "5+3"
```