HPC Coffee hour

Slides: https://rcdata.nau.edu/hpcpub/workshops/hpc_coffee_hr.pdf

Schedule: https://in.nau.edu/arc/hours

Support email: ask-arc@nau.edu

NAU NORTHERN ARIZONA UNIVERSITY

Topics

- Who are we?
- Updates
- Tiers of storage
- Pending jobs
- Data portal (Chris)
- Slurm arrays (Joseph)
- Efficient use of resources (Joseph)
- Cluster Metrics (Joseph)
- Spack (Joseph)
- Globus (Jason)
- Job script archiver (Jason)
- Interactive vs batch jobs (Jason)
- Your topics!

NAU NORTHERN ARIZONA UNIVERSITY

Who are we?

- https://in.nau.edu/arc/our-team/

- Chris
- Keith
- Joseph
- Jason
- Mike
- Alex (student)

NORTHERN ARIZONA UNIVERSITY

General Updates

- Just completed a quarterly maintenance before spring semester on Jan 9th
- Rearranged our website and improved our documentation
  - https://nau.edu/arc
- We have two login nodes:
  - wind.hpc.nau.edu (monsoon.hpc.nau.edu)
    - Everyone but classroom users allowed
  - rain.hpc.nau.edu

NORTHERN ARIZONA UNIVERSITY

Software stack Updates

- Jupyterlab app added to ondemand
- Vscode app added to ondemand
- Mambaforge our recommended python module (vs anaconda)
- Updated singularity module, now called “Apptainer”
  - Use this to create your own containers in solving complex software dependencies

NAU NORTHERN ARIZONA UNIVERSITY

Tiers of storage

- /home – fast but small (20GB)
  - Keep jobscripts and small executables here
  - Backed up to tape
  - Daily snapshots located here: /home/.snapshot/<DATE>/<USERID>
- /scratch – fast and large (15TB)
  - Default location!
  - Very fast, can handle parallel writes
  - No backups
- /projects - slower and large (5TB+)
  - Storing data long-term
  - Slower storage, not for doing high input / output from many jobs
  - Should work ok for reference in jobs, but not for manipulating (writing) data
  - No backups
  - Snapshots available upon request!
  - Data can be accessible via our data portal!

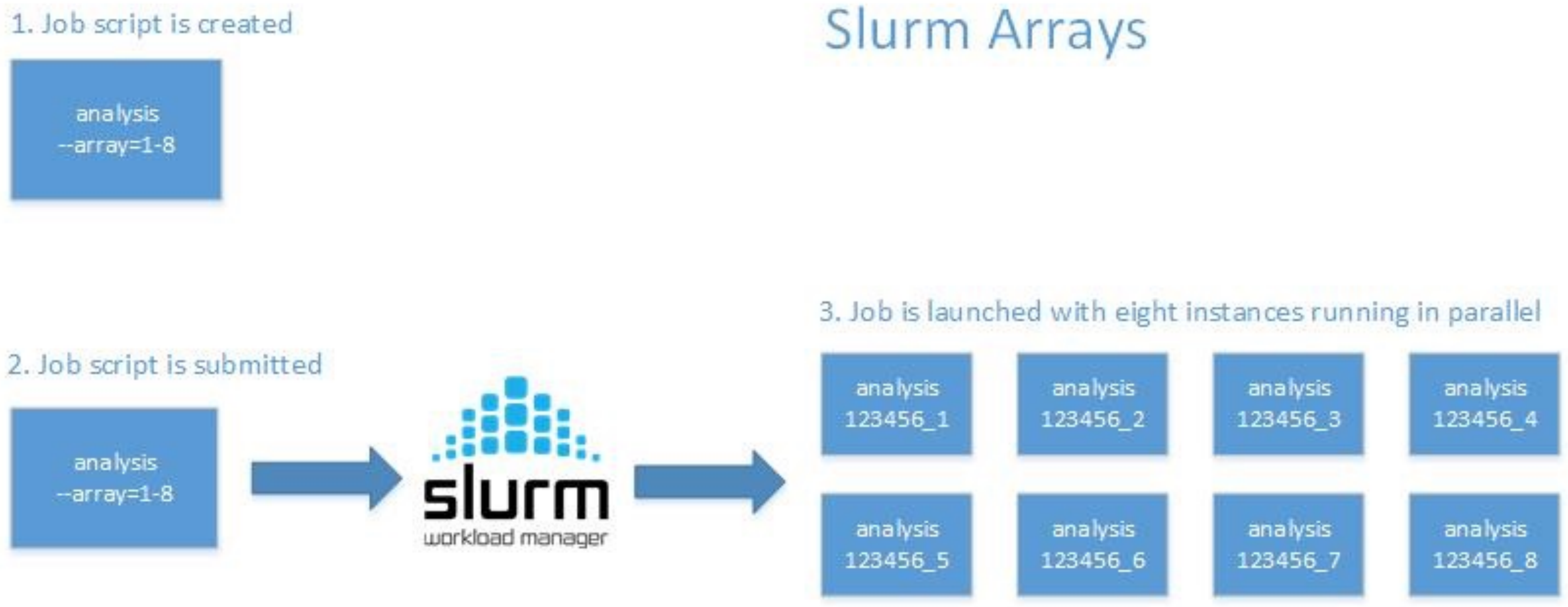
Data portal

- demo

Slurm Arrays (Joseph)

- Awesome for researchers with many inputs and one analysis program
- Achieve parallelism from single core jobs

Slurm Arrays



Useful environment variables

SLURM_ARRAY_JOB_ID:	the job array's ID (parent)
SLURM_ARRAY_TASK_ID:	the id of the job array member n (child)
%A	
%a	

NORTHERN ARIZONA UNIVERSITY

Slurm Arrays Exercise

- From your scratch directory: “/scratch/nauid”
- tar xvf /common/contrib/examples/bigdata_example.tar
- cd bigdata
- edit the file “job_array.sh” so that it works with your nau id replacing all NAUID with yours
- Submit the script “sbatch job_array.sh”
- Run “squeue”, notice there are 5 jobs running, how did that happen!

Efficient use of resources

- https://metrics.hpc.nau.edu/doppler
- jobstats
- user_efficiency
- group_efficiency
- jobstats -S 2/1/24
- By creating efficient jobs:
  - Jobs start faster for you and everyone!
  - Your group gets more resources!
    - As your jobs use less of your groups allocated minutes
  - Not always simple though as we know!

NAU NORTHERN ARIZONA UNIVERSITY

Cluster Utilization

- How do you know what the current cluster utilization is?
  - Metrics : https://metrics.hpc.nau.edu
  - sinfo
    - Shows you the state of the cluster and the nodes
  - squeue
  - gpu_status

Spack

- What it is
  - A package manager to resolve complex software dependencies
- How to use it
  - spack info
  - spack list
  - spack find
  - spack install
  - spack load
- Demo

Globus (Jason)

- https://in.nau.edu/arc/globus/
- Brief 5 min demo

Interactive vs Batch jobs

- demo

Jobs script archive

- We backup a copy of your job scripts here:
  - /common/jobscript_archive/<userid>/year/month
- Handy dandy script to retrieve an archived script:
  - showscript <jobid>

Your turn for topics!

What shall we talk about now?

More tidbits … if no questions

- Apptainer
- TRES minutes

NORTHERN ARIZONA UNIVERSITY

Apptainer (Joseph)

- Singularity has been renamed
- Howto create a container from docker hub
- Howto create a container from a docker file
- Sandbox vs image
- Launching command from container
- Demo

TRES run minutes

- What the heck is that!?
- A number which limits the total number of remaining resource minutes which your ***running*** jobs can occupy.
- Enables flexible resource limiting
- Staggers jobs
- Increases cluster utilization
- Leads to more accurate resource requests

- Sumofjobs(resource * timelimit remaining)

NORTHERN ARIZONA UNIVERSITY

TRES run minutes

- TRES
  - Trackable Resource
- We limit groups to a certain number of cpu, memory, and gpu minutes
  - CPU: 4.4 million minutes
  - Memory: 13.7PiB minutes
  - GPU: 29,160 minutes
- This has nothing to do with your priority, rather, the amount of resources your group has access to in real time!

NORTHERN ARIZONA UNIVERSITY

Examples

- 14400 = 10 jobs, 1 cpu, 1 day in length
- 144000 = 10 jobs, 10 cpu, 1 day in length
- 720000 = 10 jobs, 10 cpu, 5 days in length
- 720000 = 1000 jobs, 1 cpu, ½ day in length
- 1105920 = 1 job, 1024 cpus, 18 hrs in length

Questions?

- Check your groups resource min usage:
  - sshare -l

NORTHERN ARIZONA UNIVERSITY

TRES run minutes (demo)

- Say, groupA’s total cpu minute limit is: 5000
- Example, groupA submits three jobs
  - Job1:
    - 1 core
    - 1 day timelimit (1440 minutes)
    - 1 GB memory
  - Job2:
    - 2 core
    - 1 days (1440 minutes)
    - 16 GB memory
    - 2880 minutes total !
  - Job 3:
    - 1 core
    - 1 day (1440 minutes)
    - 1GB memory

NORTHERN ARIZONA UNIVERSITY

TRES run minutes

- Assuming there are available monsoon resources
- How many jobs start?
- How many cpu minutes are in use?
- When is job 3 ELIGIBLE to start?

TRES run minutes

- Assuming there are available monsoon resources
- How many jobs start?
  - 2
- How many cpu minutes are in use?
  - 1440+2880 = 4320
- When is job 3 ELIGIBLE to start?
  - After ~6 hours (6*60 = 360), and 2 jobs (360*2) = 720 minutes

  - We have only 5000-4320 = 680 minutes available initially
  - After ~ 1/4 day goes by (360 minutes) * 2 (two jobs) = 720 minutes
  - 680 + 720 = 1400
  - After another 40 minutes we'll have 1440 at which point job starts

NORTHERN ARIZONA UNIVERSITY