Expanding Serial Analysis with Slurm Arrays

Christopher Coffey

Northern Arizona University

[nau.edu/hpc](nau.edu/hpc)

Serial job limitations

- Serial jobs are limited to one core as they are not multi-threaded
- How to take advantage of a huge system then without threading, or MPI?

Expanding Serial job analysis

Solution:

- If the job's and their data are independent then we can

- Submit a bunch of similar serial jobs each with their own job script

- Create a for loop which aids in the above:

  for i in `seq 1 1000`; do sbatch jobscript_$i.sh;done

  - Demo!

Expanding Serial job analysis

- That's all I have folks, any questions?
  - … wait no theres a better way!

Expanding Serial job analysis

- The previous example works, but we have to maintain a jobscript file for each job. What a pain!

- Slurm arrays can help!

- Create many jobs from one job script! Even 100's of thousands!

- Can be used in many different scenarios:
  - One analysis program, many data inputs
  - One analysis program, one data input, many constraints
  - One input, many programs
  - Many commands

Slurm Arrays!

Slurm Arrays

1. Job script is created

analysis
--array=8

2. Job script is submitted

analysis
--array=8

slurm
workload manager

3. Job is launched with eight instances running in parallel

| analysis 123456_1 | analysis 123456_2 | analysis 123456_3 | analysis 123456_4 |
| analysis 123456_5 | analysis 123456_6 | analysis 123456_7 | analysis 123456_8 |

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

Example Slurm Arrays

Demos!

Summary

- Job arrays are an easy way to achieve some parallelism from single core jobs

- According to the man page, Slurm supports job arrays of size 4M
  - slurm.conf config value: MaxArraySize

- I've tested job arrays of 40K, some sites support 100K

Questions?

Thank you!

Demo files:
- goo.gl/XNiiev

7bb15a56486c84a87f16039e4ec352c1 (md5sum) - slurm_arrays.tar

Slides:
- goo.gl/H5wTJx

Christopher Coffey
Northern Arizona University
chris.coffey@nau.edu

NORTHERN ARIZONA UNIVERSITY

Whoa, lets slow it down a bit!

- If after creating the huge job arrays you've decided that you, or your admin would prefer you to run only "n" at a time, instead of all at once (if you were permitted).

- In your array declaration use "%n" to limit how many can be run at a time

  Example: --array=1-24%4   # only allow 4 jobs at a time to run from your array