**NFS and Data Oddities**

1. Overview

Last week, I implemented a common NFS container and collected system metrics from pre-alignment jobs. NFS was chosen to overcome permission issues on Chameleon containers which caused Slurm to be unable to monitor containers. However, it ends with a new problem, namely Slurm detects that the container size is 1KB when in reality it is hundreds of GB. In order not to linger on this issue, I proceeded to the task of collecting system metrics from pre-alignment jobs. However, in the data that has been obtained there are oddities for which I have not been able to find the cause.

1. Replacing chameleon container with NFS

Previously I used the chameleon container to store input files and intermediate files from jobs. However, it turns out that slurm accounting cannot monitor the container because there are permission restrictions in it that we cannot modify. As an alternative solution to this problem, we can use NFS. Using NFS, we mount a directory of compute nodes in the directory of the master node. The advantage of this method is that we can freely set the permissions of the directory that we use as a common container. Therefore, I decided to use NFS.

Once NFS is implemented, Slurm can monitor this common container. However, a new problem arose. Slurm only detects that the container contains only 1KB even though the container has hundreds of GB. This must have happened because of a configuration error in the Slurm that I applied. However, since I'm not sure I can solve it quickly, I decided to skip this problem and move on to the next task, which is collecting system metrics from pre-alignment jobs.

1. NFS automation script

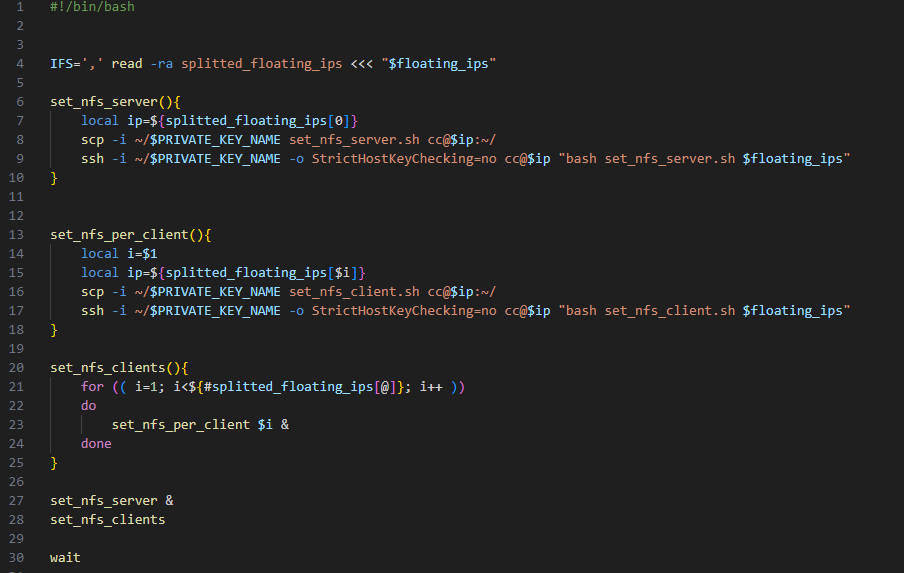


Figure 1. Set nfs for all nodes.



Figure 2. Set NFS for server node.

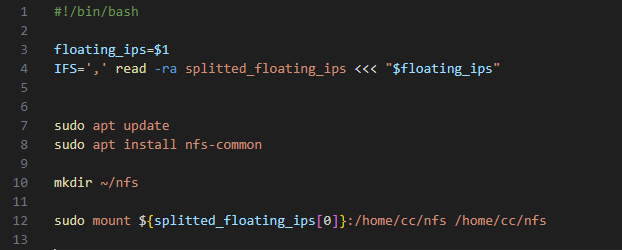


Figure 3. Set NFS for client nodes

1. Data oddities

The pre-alignment script is run by varying the amount of memory and CPUs allocated based on the combination of sets M (in GB) and C, where M={ 1, 2, 4, 8, 16, 32, 64} and C={1, 2 , 4, 8, 16, 32}. From this, we derive the systems metrics for 48 jobs. However, there are oddities in the data obtained. In the AveDiskRead and AveDiskWrite columns there are many values 0 even though the tasks were executed successfully. In addition, there are many values that are repeated in the AveDiskWrite column even though the values in this column are continuous values. Based on initial observations, in this column there are only 13 unique values for all the data collected.

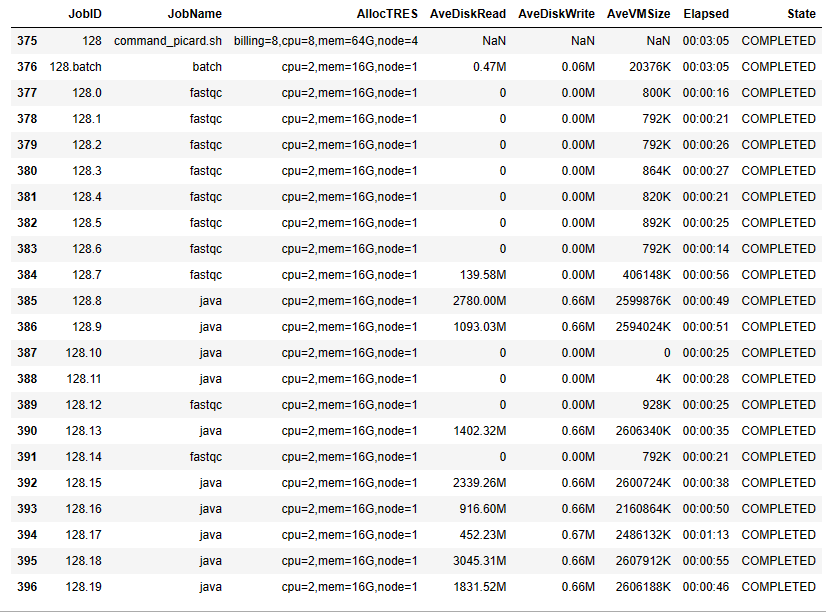


Figure 4. Sample of data oddities.