SGE info

Walter K Vogel <vogelw@onid.oregonstate.edu> Thu, Dec 22, 2016 at 7:03 PM
To: Richard Rodrigues <dr.richrodrigues@gmail.com>

Rich,

I left something off the array job description. It generates and accesses an environmental variable: ${SGE_TASK_ID}

I am attaching a recent script that uses this variable is several locations to process different input files. Best explained by example.
I am am attaching an entire run script as it was recently run.

There are some things in this script that you should just ignore. The module system is not officially supported on our infrastructure and of course the binary is in my $HOME directory.

You see how I used the ${SGE_TASK_ID} variable in three different commands in succession.

On Dec 22, 2016, at 6:22 PM, Richard Rodrigues <dr.richrodrigues@gmail.com> wrote:

Cool, this is definitely helpful! :)

Thanks.

-Rich

On Thu, Dec 22, 2016 at 5:03 PM, Walter K Vogel <vogelw@onid.oregonstate.edu> wrote:

> Rich,
>
> Start an interactive session on a queue; example one slot only:
>
> qrsh -q transkingdom
>
> for 10 slots
>
> qrsh -q transkingdom -pe thread 10
>
> To submit a batch script use the 'qsub' i.e.
>
> qsub ./script
>
> note that the script needs to be executable itself (this is what I do)
> or run by an executable, for example:
>
> bash bashscript
> or
> perl perlscript.pl
>
> EITHER on the command-line OR in the "script" OR a mixture of both location,
> several options must be set.
> For example consider the following:

```
#!/bin/bash

# ------------------------------------------------------------------
----------- #
# SGE commands
            #
# ------------------------------------------------------------------
----------- #
#$ -cwd
#$ -V
#$ -S /bin/bash
#$ -N runname
#$ -q transkingdom
#$ -l mem_free=10G
#$ -pe thread 32

Command
Command2
```

The above script will run Command followed by Command2 on transkingdom in parallel using 32 slots. It will also only start if there is at least 10G RAM available. The commands could be anything, a unix command, an other executable script, etc.

Note that SGE options have a '#$' at the start of the line. For an array job use the '-t' option. For example

```
#!/bin/bash

# ------------------------------------------------------------------
----------- #
# SGE commands
            #
# ------------------------------------------------------------------
----------- #
#$ -cwd
#$ -V
#$ -S /bin/bash
#$ -N runname
#$ -q transkingdom
#$ -l mem_free=10G
#$ -pe thread 32
#$ -t 1-50

Command
Command2
```

The following will do the same but launch only on the all.q (the only queue with a .q in its name) AND require the reef host.

```
#!/bin/bash

# ------------------------------------------------------------------
----------- #
```

```
# SGE commands
            #
# ------------------------------------------------------------------
---------- #
#$ -cwd
#$ -V
#$ -S /bin/bash
#$ -N runname
#$ -q all.q
#$ -l h=reef
#$ -pe thread 32
#$ -t 1-50

Command
Command2
```

see that the ‘-l’ option sets requirements. In the above instance the host ‘reef’ is required. Note here I left off the free memory requirement its optional after all. This will just run regardless of the memory situation (which may not be what you want). You can have multiple ‘-l’ options set as in

```
#$ -l h=reef
#$ -l mem_free=50G
```

Hope the above helps, feel free to ask questions if this isn’t clear.

Walter

--
Rich

md-run_bio
2K