Tutorial: Run ABCluster on Computer Clusters with a Queueing System

Author: Jun Zhang
zhangjunqcc@gmail.com

Contents

- Prologue
- Serial isomer+Gaussian using Torque
- Parallel isomer+Gaussian using Torque
- Epilogue

Prologue

Prologue

- ABCluster is an efficient software for the global optimization of chemical clusters.
- This tutorial will introduce how to run isomer and lego on a computer cluster with a queueing system. The reader is assumed to be familiar with ABCluster. Otherwise, please refer to the ABCluster manual and ppt tutorials on:

http://www.zhjun-sci.com/software-abcluster-download.php

Run ABCluster on a Computer Cluster

- The isomer and lego components of ABCluster can call any computational chemistry software directly. But if you want isomer and lego to call the software in a queueing system on a computer cluster, you need a few more steps.
- Linux computer cluster and Torque queueing system.

Serial isomer+Gaussian using Torque

Global Minimum of B_7^-

- Example: Search the global minimum of B_7^- at B3LYP/6-31g(d) level of theory.

b7.inp

```
b7              # Result file name
B 7             # Cluster file name
cube 3 3 2      # Structure types
30              # Maximal number of calculations
>>>>
xyz2gaussian optfile $inp$ > $xxx$.gjf
g09 < $xxx$.gjf > $xxx$.log 2>/dev/null
gaussian2xyz $xxx$.log > $out$
rm $xxx$.gjf $xxx$.log
>>>>
```

optfile

```
%nproc=16
%mem=20GB
#N B3LYP/6-31g(d) scf(xqc,novaracc)
opt(MaxCycles=100)

opt

-1 1
>>>>

>>>>
```

- Without a queueing system:

nohup isomer b7.inp > b7.out &

Queueing System

- To run ABClsuter on a computer cluster, assume it uses Torque. The script of Gaussian job is given below. To submit the job, run: qsub submit.pbs

submit.pbs

```
#!/bin/bash

#PBS -l nodes=1:ppn=32
#PBS -l walltime=48:00:00
#PBS -q normal
#PBS -e $PBS_JOBID.err
#PBS -o $PBS_JOBID.out
#PBS -N se8

export GAUSS_SCRDIR=/scratch/sciteam/zhang9
fn=filename
g09 < ${fn}.gjf > ${fn}.log
```

We will use sed to replace this line to the true filename

ABCluster+Torque

- To enable ABCluster with Torque, you can use the following script: jobhold.sh

Submit the job and obtain its ID. Please modify this according to your system.

Check if the job is finished every 10 seconds

```
#!/bin/bash

# Submit and get job ID.
me=`whoami`
jobid=`qsub submit.pbs | tail -n 1`

# Detect if the job is finished.
sleep_time=10 # in second
jobstate="Q"
while [ -n "${jobstate}" ]
do
   sleep ${sleep_time}
   jobstate=`qstat -u ${me} | grep ${jobid}`
done
```

ABCluster+Torque

- Now modify b7.inp:

b7.inp

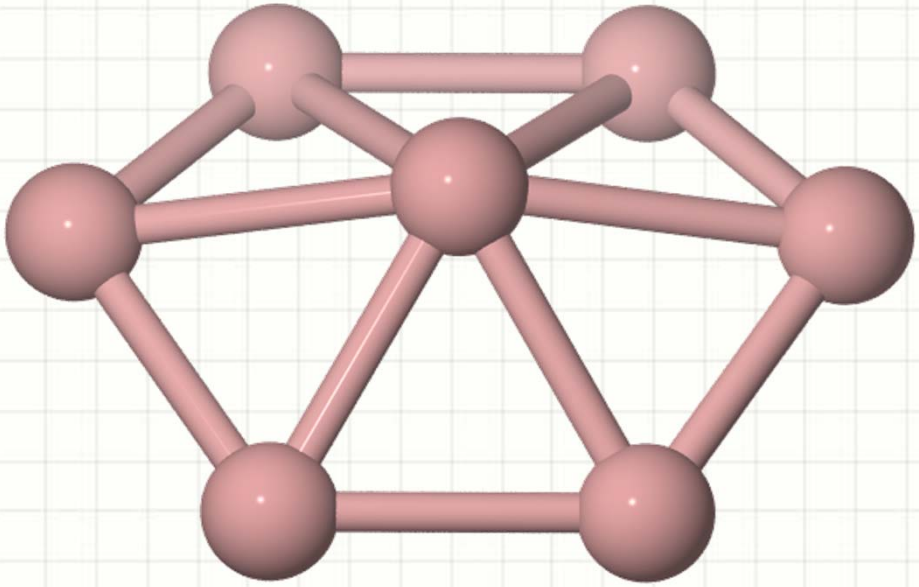
Replace the file name

```
b7              # Result file name
B 7             # Cluster file name
cube 3 3 2      # Structure types
30              # Maximal number of
calculations
>>>>
xyz2gaussian optfile $inp$ > $xxx$.gjf
sed -i '11c fn=$xxx$' submit.pbs
./jobhold.sh
gaussian2xyz $xxx$.log > $out$
rm $xxx$.gjf $xxx$.log
>>>>
```

Submit

- Run this command on the control node:

  nohup isomer b7.inp > b7.out &
- Now isomer will use Torque to submit jobs.
- This is the final global minimum:



Parallel isomer+Gaussian using Torque

Parallelization of ABCluster

- From version 2.0, ABCluster can run jobs in parallel with SSH.
- Nodes must be able to connect each other using SSH without password.

Parallelization of ABCluster

- It is not always a good idea to run ABCluster with too many nodes. For example, to do 50 calculations, one can use 10 nodes to run 10 calculations at a time, or use 50 nodes to run all the calculations at one time. However, the latter is inefficient, because in the first way, the following calculations can use information obtained from previous calculations to generate better initial structure guess. In the latter way, all searches will use random intial guess.

Global minimum of B_7^-

- To use 6 nodes to search the global minimum of B_7^- at B3LYP/6-31g(d) level of theory, we will write the hostname of the nodes in a file called nodes.

b7.inp

```
b7              # Result file name
B 7             # Cluster file name
cube 3 3 2      # Structure types
30              # Maximal number of calculations
>>>>
xyz2gaussian `pwd`/optfile $inp$ > $xxx$.gjf
g09 < $xxx$.gjf > $xxx$.log 2>/dev/null
gaussian2xyz $xxx$.log > $out$
rm $xxx$.gjf $xxx$.log
>>>>
nodes
```

Absolute path!

nodes

```
6
node01
node01
node02
node02
node03
node03
```

- Without Torque: nohup isomer b7.inp > b7.out &

ABCluster+Torque

- With Torque, you do not explicitly SSH the computational nodes. You can change the 6 names to the hostname of your login node, or simply "localhost."

b7.inp

```
b7              # Result file name
B 7             # Cluster file name
cube 3 3 2      # Structure types
30              # Maximal number of calculations
>>>>
xyz2gaussian `pwd`/optfile $inp$ > $xxx$.gjf
sed -i '11c fn=$xxx$' `pwd`/submit.pbs
`pwd`/jobhold.sh
gaussian2xyz $xxx$.log > $out$
rm $xxx$.gjf $xxx$.log
>>>>
nodes
```

nodes

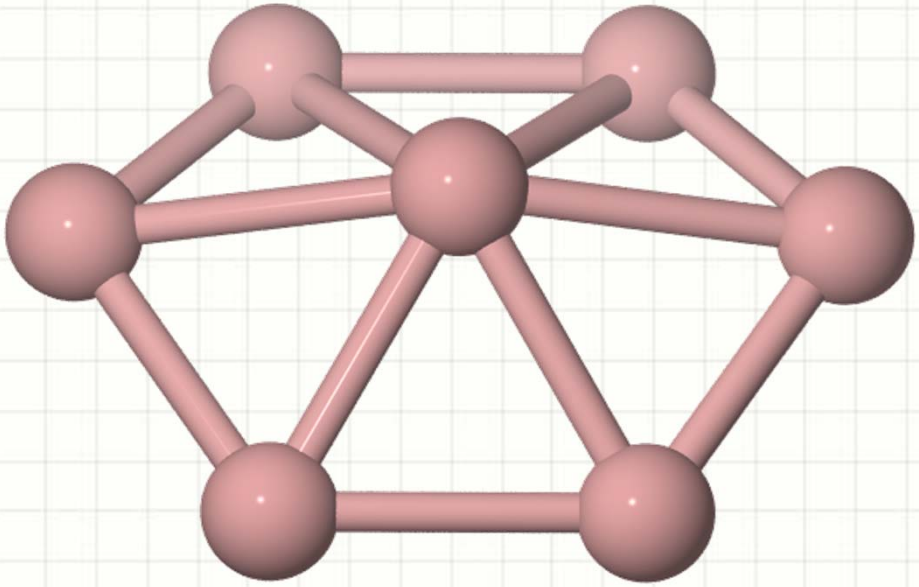
```
6
localhost
localhost
localhost
localhost
localhost
localhost
```

Submit

- Run this command on the control node:

  nohup isomer b7.inp > b7.out &

- Now isomer will use Torque to submit 6 jobs at one time!



Epilogue

Now I believe readers can use isomer and lego to call Gaussian or other software with queueing systems.

For the lattest version of ABCluster:

www.zhjun-sci.com

For any suggestions:

zhangjunqcc@gmail.com


THANK YOU!