Matlab @Taiwania 3

Matlab 執行方式

以平行運算 *benchmark*為例

參考指引

目錄

零、說明 **2**

一、**pre-requirement** **2**

二、**Matlab** 運行方式 **3**

1. 準備 matlab 程式 (平行範例) 3
2. 設定 sbatch 需要的參數 4
3. 呼叫 slurm 執行 5

三、確認結果 **5**

四、實驗建議 **6**

五、參考 **- Slurm** 常用指令 **6**

零、說明

以下為如何在國網中心 Taiwania 3 透過 Slurm 送 matlab 平行程式 進行計算。

一、pre-requirement

- 以下操作需登入 **T3 登入節點**

```
+-------------------------------------------------------------------------+
|                                                                         |
|  Welcome to NCHC 2020 HPC System                                        |
|                                                                         |
|                                                                         |
|                                                                         |
+-------------------------------------------------------------------------+
|                                                                         |
| 1. All jobs, batch or interactive must be submitted to the scheduler.   |
|                                                                         |
| 2. DO NOT run computational jobs on login nodes.                        |
|    Login nodes are meant for light editing or compilation and for submitting|
|    jobs. Any job that runs for more than an hour* or is using significant|
|    memory and CPU within an hour should be submitted to Slurm either as |
|    interactive or batch job.                                            |
|                                                                         |
| 3. DO NOT run SLURM query commands CONTINUOUSLY and FREQUENTLY on login |
|    nodes. Any execution of interval less than 30 seconds and lasting more|
|    than 10 minutes is regarded as a system attack and will be terminated|
|    immediately without notification.                                    |
|                                                                         |
|                                                                         |
|   * Long-running jobs will be removed.                                  |
|  ** We reserve the right to limit access for the users who abuse the system.|
|                                                                         |
+=========================================================================+
[waue0920@lgn303 ~]$
```

- 執行wallet或get_su_balance 取得錢包餘額與計畫編號
  - ex : 下面範例中的 GOVxxx016
  - 計畫編號查詢方式請到 iservice 查詢，或執行 wallet

```
[waue0920@lgn303 ~]$ wallet
{"PROJECT_ID":"GC      7","PROJECT_NAME":"NCH                                    }
{"PROJECT_ID":"GC      6","PROJECT_NAME":"資                                     5348.0096"}
{"PROJECT_ID":"GC      6","PROJECT_NAME":"NCH                                    89.5205"}
{"PROJECT_ID":"GC      5","PROJECT_NAME":"民                                     "}
{"PROJECT_ID":"GC      4","PROJECT_NAME":"NCH                                    "}
{"PROJECT_ID":"GC      3","PROJECT_NAME":"NCH                                    74.9353"}
```

- 登入節點上僅供整理檔案與派送工作，不可直接執行長時間佔用CPU之工作，以免影響其他用戶使用之權益

二、Matlab 運行方式

```
[waue0920@lgn303 matlab_workspace]$ ls -al
total 34
drwxr-xr-x  2 waue0920 TRI107129  4096 Apr 21 16:37 .
drwx------ 44 waue0920 TRI107129 16384 Apr 21 16:37 ..
-rw-r--r--  1 waue0920 TRI107129   459 Apr 21 15:03 matlab-slurm.sb
-rw-r--r--  1 waue0920 TRI107129  1757 Apr 21 11:40 MPS_test.m
```

1. 準備 matlab 程式 (平行範例)

以下為matlab 的平行範例程式，可更換或修改

$ cat MPS_test.m

```
%% Test Script with Prime Factors Caculation
% Configuration
numWorkers = [1 2 4 6 16 32 64];    % Worker number list
resultPath = './';  % Report storage path
outputName = 'MATLAB_MPS_Report.csv';   % Report file name
logName = 'MPS_test_log.log';   % Log name

%% Start Logging
fid = fopen(logName,'a');
timestamp = datetime('now','Format', ...
    '==== yyyy年MM月dd日 HH:mm:ss====');   % Get current time
fprintf(fid,'%s\n',timestamp);

%% Test Algorithm
% Solve prime factors for each prime
primeNumbers = primes(uint64(2^18));    % Problem scale
compositeNumbers =
primeNumbers.*primeNumbers(randperm(numel(primeNumbers)));

% Pre-allocation
factors = zeros(numel(primeNumbers),2); % 2 factors
elapsed_time = zeros(size(numWorkers)); % Time spend

% Parallel for-loop
try
    fprintf(fid,'Start Testing...\n');
    for i = 1:numel(numWorkers)
        tic;
        parfor (idx = 1:numel(compositeNumbers), numWorkers(i))
            factors(idx,:) = factor(compositeNumbers(idx));
        end
```

```
        elapsed_time(i) = toc;
    end
    fprintf(fid,'Test Done!\n');
catch
    fprintf(fid,'Start  Failed!\n');
    fclose(fid);
end

%% Generate Report
speedup = elapsed_time(1)./elapsed_time;    % Calculate speed up compare to
singal worker
resultTable = array2table([numWorkers',elapsed_time',speedup'], ...
    'VariableNames',{'numWorkers','Time(sec)','Speedup'});

% Write out report
try
    fprintf(fid,'Generate Report...\n');
    writetable(resultTable,fullfile(resultPath,outputName))
    fclose(fid);
catch Msg
    fprintf(fid,'Generate Report Failed!!\n');
    fprintf(fid,'Error: \[%s\] \n',Msg.identifier);
    fprintf(fid,'Error Message: \[%s\] \n', Msg.message);
    fclose(fid);
```

2. 設定 sbatch 需要的參數

$ cat matlab-slurm.sb

```
#!/bin/bash -l
#SBATCH -A GOVxxx016   #計畫編號
#SBATCH -J the_job_name   #job_name 可自訂
#SBATCH -p ctest   # queue/partition name
#SBATCH -n 64   #共要執行的process數目
## SBATCH --ntasks-per-node=56 #每個節點執行的process數目，最多為56
#SBATCH -N 4        # Maximum number of nodes to be allocated
#SBATCH --switches=1 #maximum count of switches
#SBATCH -o %j.out   # Path to the standard output
#SBATCH -e %j.err  # Path to the standard error
ml matlab
matlab -nodisplay -nosplash < MPS_test.m
```

! 詳細可參考 iService核心運算說明, partition name 可參考 queue 與計算資源使用說明

3. 呼叫 slurm 執行

```
$ sbatch matlab-slurm.sb
```

確認是否已在排班中

```
$ squeue -u <username>
```

```
[waue0920@lgn303 matlab_workspace]$ squeue -u waue0920
             JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
           1430334     ctest waue_not waue0920 PD       0:00      1 (Resources)
           1430348     ctest waue_not waue0920 PD       0:00      1 (Resources)
           1430353     ctest waue_not waue0920 PD       0:00      1 (None)
[waue0920@lgn303 matlab_workspace]$
```

三、確認結果

當slurm 完成工作後，可以看到有幾個檔案出現

```
[waue0920@lgn303 matlab_workspace]$ ls -al
total 36
drwxr-xr-x  2 waue0920 TRI107129  4096 Apr 21 16:36 .
drwx------ 44 waue0920 TRI107129 16384 Apr 21 16:36 ..
-rw-r--r--  1 waue0920 TRI107129     0 Apr 21 14:41 1430334.err
-rw-r--r--  1 waue0920 TRI107129   506 Apr 21 14:43 1430334.out
-rw-r--r--  1 waue0920 TRI107129   220 Apr 21 15:26 MATLAB_MPS_Report.csv
-rw-r--r--  1 waue0920 TRI107129   459 Apr 21 15:03 matlab-slurm.sb
-rw-r--r--  1 waue0920 TRI107129   783 Apr 21 15:26 MPS_test_log.log
-rw-r--r--  1 waue0920 TRI107129  1757 Apr 21 11:40 MPS_test.m
```

以slurm 的執行結果而言，執行完會有 PID 的兩個檔案出現

- <jobid>.err
- <jobid>.out

以上面的範例程式而言，執行完會出現

- MATLAB_MPS_Report.csv
- MPS_test_log.log

範例程式平行後運行結果

```
[waue0920@lgn303 matlab_workspace]$ cat MATLAB_MPS_Report_n64.csv
numWorkers,Time(sec),Speedup
1,32.157246,1
2,6.767954,4.75139842853542
4,3.499362,9.18945967864999
6,2.377907,13.5233404838793
16,1.290861,24.911470716057
32,1.217706,26.4080541608566
64,1.197563,26.8522374188247
```

四、實驗建議

不同sbatch參數會得到不同結果，以下列舉一些參數的不同所獲得不同的結果

#SBATCH -n 1	#SBATCH -n 64 #SBATCH -N 4
numWorkers,Time(sec),Speedup 1,36.372204,1 2,13.225855,2.75008337835248 4,13.130152,2.77012817521077 6,13.115558,2.77321056412545 16,13.169758,2.76179744532891 32,13.117167,2.77287039190703 64,13.093352,2.77791386040794	numWorkers,Time(sec),Speedup 1,32.157246,1 2,6.767954,4.75139842853542 4,3.499362,9.18945967864999 6,2.377907,13.5233404838793 16,1.290861,24.911470716057 32,1.217706,26.4080541608566 64,1.197563,26.8522374188247

如何妥善的設置參數使效率提高，可參考 [iService核心運算說明]

五、參考 - Slurm 常用指令

- sinfo

$ sinfo

```
[waue0920@ln01-twnia2 matlab_workspace]$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
gtest        up      30:00      2   idle gn[1227-1228]
gp1d         up 1-00:00:00      5 drain* gn[1008,1024,1117,1122,1129]
gp1d         up 1-00:00:00     53  drain gn[1001-1007,1009-1023,1025-1030,1103-1116,1118-1121,1123-1128,1130]
gp1d         up 1-00:00:00     15    mix gn[1201-1210,1213-1215,1220-1221]
gp1d         up 1-00:00:00     11   idle gn[1211-1212,1216-1219,1222-1226]
gp2d*        up 2-00:00:00      4 drain* gn[1008,1117,1122,1129]
gp2d*        up 2-00:00:00     46  drain gn[1001-1007,1009-1022,1103-1116,1118-1121,1123-1128,1130]
gp2d*        up 2-00:00:00     15    mix gn[1201-1210,1213-1215,1220-1221]
gp2d*        up 2-00:00:00     11   idle gn[1211-1212,1216-1219,1222-1226]
gp4d         up 4-00:00:00      4 drain* gn[1008,1117,1122,1129]
gp4d         up 4-00:00:00     46  drain gn[1001-1007,1009-1022,1103-1116,1118-1121,1123-1128,1130]
gp4d         up 4-00:00:00     15    mix gn[1201-1210,1213-1215,1220-1221]
gp4d         up 4-00:00:00     11   idle gn[1211-1212,1216-1219,1222-1226]
express      up 4-00:00:00      5 drain* gn[1008,1024,1117,1122,1129]
express      up 4-00:00:00     53  drain gn[1001-1007,1009-1023,1025-1030,1103-1116,1118-1121,1123-1128,1130]
express      up 4-00:00:00     15    mix gn[1201-1210,1213-1215,1220-1221]
express      up 4-00:00:00     11   idle gn[1211-1212,1216-1219,1222-1226]
t3_gp4d      up 4-00:00:00      6   idle gpn[3005-3010]
```

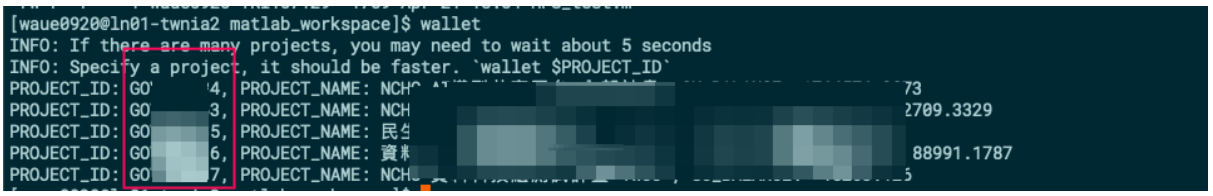
- squeue

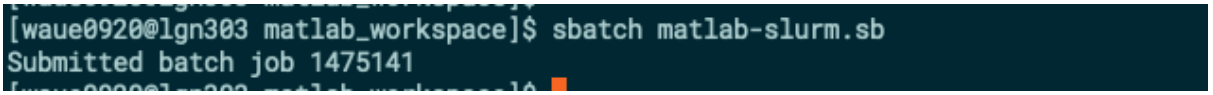
squeue -u `whoami`

```
[waue0920@ln01-twnia2 matlab_workspace]$ squeue -u `whoami`
             JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
            369049      gp1d waue_not waue0920  R       0:02      4 gn[1215-1218]
```

- wallet

$ wallet



- sbatch

$ sbatch <your_slurm_batch_file>