***叢集電腦使用問題綜合說明***

目錄

[*1.* *使用者注意事項(必讀):* 1](#_Toc426299943)

[*2.Hyper Cluster 作業系統* 4](#_Toc426299944)

[ *跟PC的連結* 4](#_Toc426299945)

[ *一般使用者常見的問題:* 4](#_Toc426299946)

[ ***Linux 簡易操作指令*** 8](#_Toc426299947)

[ *一般使用者:* 8](#_Toc426299948)

[ ***vi使用說明*** 10](#_Toc426299949)

[*3.應用軟體與程式* 11](#_Toc426299950)

[ ***執行R*** 11](#_Toc426299951)

[ *R (S –Plus) Programs* 12](#_Toc426299952)

[ *一般使用者常見的問題:* 12](#_Toc426299953)

***叢集電腦使用問題綜合說明***

# *使用者注意事項(必讀):*

* ***第一重要規則***：執行JOB時，一定要在node上，不可以在Server上直接執行，否則可能造成Server當機。*（rsh node04*: 進入node04。）
* 若要執行大型R程式，請依照下列標準流程操作：

1. 確定程式在單一核心下，需耗費記憶體大小(先測試1/100

的資料量)

1. 選定合適、足夠記憶體的node。
2. 以*qsub*送R程式(詳見第12頁)。"Do you want to execute the program R (Y/N)?"，請務必輸入Y。
3. 其詢問 "Please specify the memory usage (GB) ?"，輸入欲限制的記憶體大小，**沒有進行此步驟，將可能造成CLUSTER 記憶體耗盡而當機**(若發生節點當機，請立即通知管理員)。
4. 若發現39 node無法滿足記憶體需求，請通報管理員，由管理員請示老師是否開通49帳戶。

註：若要執行平行運算，請先知會管理員，說明占用node與所需時間，由管理員公告。未知會者，管理員得直接刪除程式。

* 39的系統硬體：

39包含1個server 和3個node，詳細記憶體規格如下，請依照使用需求選取適當node。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Memory Allocation on Cluster 39** | | | | | |
| **Server** |  |  | **Node04** | **Node05** | **Node06** |
| **4GB** |  |  | **16GB** | **64GB** | **64GB** |

# *使用者規範(必讀):*

* **Cluster使用原則**

1. **Cluster 是公共空間，所以請大家培養公德心**
2. 一般而言，Cluster是要給有大量記憶體需求的人使用，如果記憶體用量不大，不建議使用，一方面不需要(算的不會比較快)，並將壓縮到其他人使用權利。
3. 目前的開放空間是 IP 39，這個伺服器上面是一架 server 跟3個 node，一個 node 有8顆核心，總計有24個core，若單純只是想開平行運算加速，但沒有實質記憶體的需求，將會壓縮到其他人使用權利。

4. 因為 Cluster 是公共空間，所以不要將所有資源都占滿。

* **Cluster使用方式**

1. 登入

(1) 登入各個 node 時，請由 server 登入，若要切換到其他 node 時，

請用 exit先退回server 再切換至 node。

(2) 登入一個 node， 請先以 top 查看該 node 於記憶體與cpu使用情形

a. 於 cpu 請看第三列 Cpu(s) 之 %us 使用百分比，或再按 1 查看 cpu 詳細被使用情形

b. 於記憶體請看第4列 Mem，free的部分，若記憶體空間不足以進行 job 請切換 node

2. CPU

(1) 大量 core 的使用主要來自平行運算，若 job 記憶體需求不大(單一工作為2 GB 以下)，建議至 node04執行。

(2) 若是跨 node 平行運算，請先確認所有 node 均沒有使用者，並且使用crm確認各 node 均有資源負載該 job。

**3. RAM**

**(1) 了解自己 job 使用的記憶體大小，選擇適當的 node。**

(2) 盡量使用排程，限制記憶體進行工作

(3) 於39，每位使用者單一工作的記憶體限制為，每個 node 記憶體之75%，超出限制將直接中斷程式。

* **Cluster使用需留意的公德心**

1. **若並非以排程進行工作，請在工作完畢後關閉程式，以釋放記憶體。**
2. **如果使用 crtl + Z 將工作丟至背景，請記得以fg切回關閉工作，或以 ps 找出相關 PID，再以 kill -9 PID 刪除工作，以釋放記憶體**

* 執行其他非R程式，亦可以指令＂*pvmem=*＂指定使用記憶體之上限，以velvet為例如下（在node06上使用一個processor，記憶體限制10GB）：

*#PBS -l nodes=node06:ppn=1 -l pvmem=10G*

*cd /data4/yfwu/velvet\_1.1.05\_k91*

*time ./velveth /data6/xiaogeng/assembly\_time\_test/r20/ 61 -fasta -shortPaired /data5/seq/110527/fa/s\_1\_seq.fa /data5/seq/110527/fa/s\_2\_seq.fa /data5/seq/110527/fa/s\_3\_seq.fa -shortPaired2 /data5/seq/110527/fa/s\_5\_seq.fa /data5/seq/110527/fa/s\_6\_seq.fa /data5/seq/110527/fa/s\_7\_seq.fa /data5/seq/110527/fa/s\_8\_seq.fa*

*sleep 60*

*time ./velvetg /data6/xiaogeng/assembly\_time\_test/r20/ -cov\_cutoff 1 -ins\_length 350 -exp\_cov 300 -ins\_length2 450 -exp\_cov 300*

*sleep 30*

*cd /data6/xiaogeng/assembly\_time\_test/r20/*

*rm -f Graph2 LastGraph Roadmaps Sequences PreGraph*

* 目前檔案放置目錄規劃

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 目錄 | 讀取權限 | 存入權限 | 總大小 | 備註 |
| /home | ALL | ALL | 2 TB |  |
| /data1  /data1/特定目錄 | 特定目錄: 特定群組 | 特定目錄: 特定群組 | 0.5TB |  |
| /data2  /data2/特定目錄 | 非特定目錄:ALL  特定目錄: 特定群組 | Administrator  Administrator | 2TB | 目前genedata存放處。 |
| /data3/public  /data3/特定目錄 | public:ALL  特定目錄:特定群組 | public:ALL  特定目錄:特定群組 | 2TB | 目前Gene group成員可在此自由存取資料。  Data-mining group 可在此存取上課資料。 |
|  |  |  |  |  |

* 如研究需求，請到**/data2/public and data2/**特定目錄
* 群組規劃，為有效掌控Cluster硬碟容量配置，並使不同身份別能滿足個人容量需求，因此規劃群組類別，並分配可使用容量，當個人可使用的容量快滿時，系統會發出警告訊息，使用者需注意。

當出現以下訊息時，表示使用量快滿了：

md9: warning, user block quota exceeded.

當出現以下訊息時，表示超過最大容量，且無法再寫入：

md9: write failed, user block limit reached.

每個人(Account)至少屬於一個群組，下表為個人在該群組下的可使用最大容量限制：

|  |  |  |  |
| --- | --- | --- | --- |
| 群組名稱 | 個人最大容量(GB) | 群組名稱 | 個人最大容量(GB) |
| SSTeacher | 100 | SLmaster | 15 |
| STeacher | 30 | SMmaster | 15 |
| SResearch | 100 | DMTeacher | 15 |
| Teacher | 10 | Master | 10 |
| Administrator | 5 | GeneSG | 10 |
| Postdoc | 15 | Doctor | 15 |

* 為有效節省硬體容量空間，除了老師的Account之外，其它身份的Account都會有使用期限。如需要程式執行極大量的Input and output 可以使用/tmp，避免server上的 raid loading過重，以致於程式無法正常執行

使用方式: 將欲執行的程式複製到/tmp底下,並將結果回傳至/data2或是/home底下

注意: 使用/tmp將失去RAID及時備份的保護，若是硬碟出問題，結果將遺失

# *2.Hyper Cluster 作業系統*

## *跟PC的連結*

### *一般使用者常見的問題:*

#### *Q: 如何登入使用Hyper Cluster?*

ANS:

請正確安裝下列檔案:

1. Xming-6-9-0-31-setup.exe (Xming主程式)

2. Xming-fonts-7-5-0-11-setup.exe (Xming字型檔)

3. Xming-portable-PuTTY-7-3-0-6-setup.exe (Putty 主程式)

4. FTP軟體 (CuteFTP, filezilla, ssh, ..., etc)

注意: 安裝Xming字型檔時，請自行選擇安裝所有字型。

安裝檔案完畢後，執行以下步驟:

Step 1: 開啟Xming。

Step 2: 開啟putty 後:

• 在Category 中點選Connection 中SHH 選項下的X11選項, 在其視窗內勾選Enable X11 forwarding, 然後在X display location 中鍵入:0.0。

• 在Category中點選Session選項, 在其視窗中Host Name( IP address)鍵入140.116.52.39。

• 點選open 後即進入cluster。

Step 3: 進入後, 鍵入帳號及密碼。

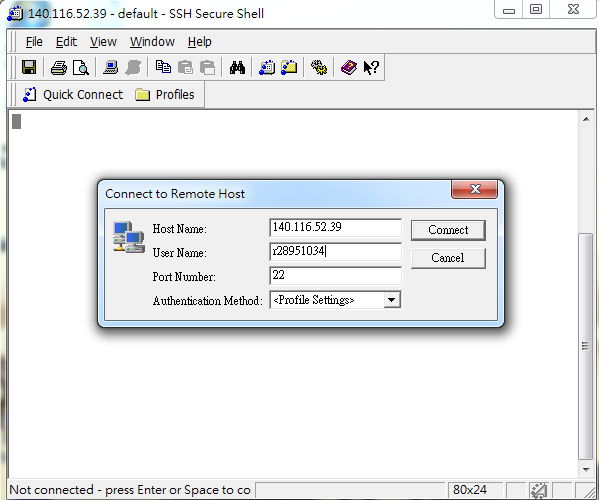
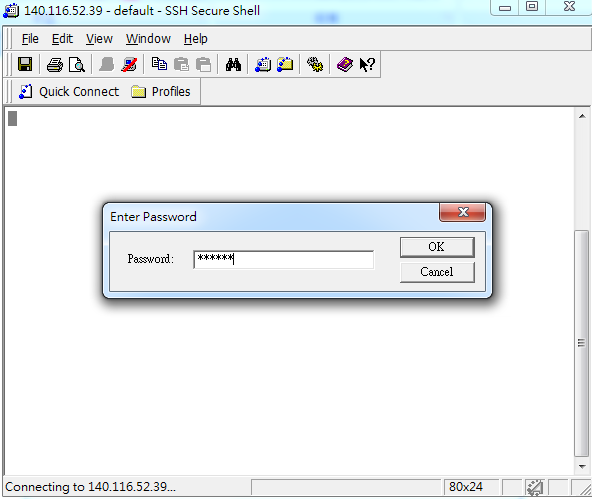
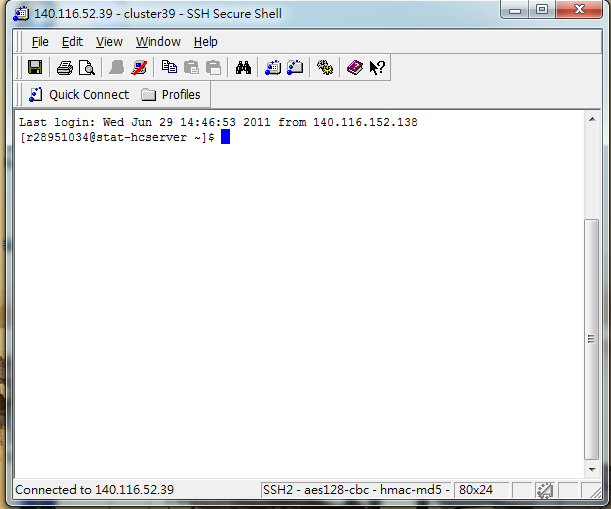
#### *Q: 如何登入使用Hyper Cluster?*

1.從成大統計系家珍網頁下載「Cluster軟體」

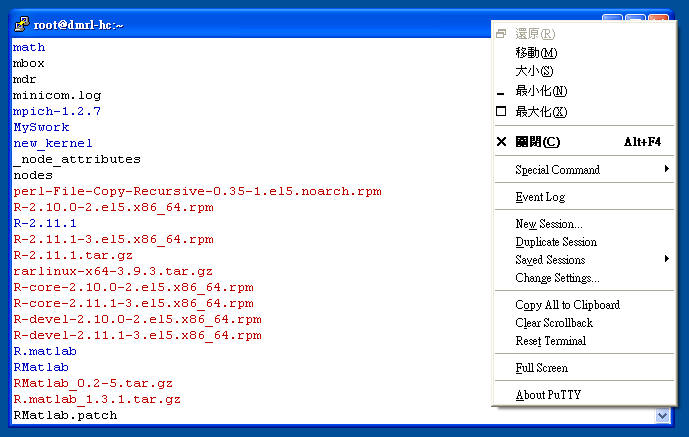
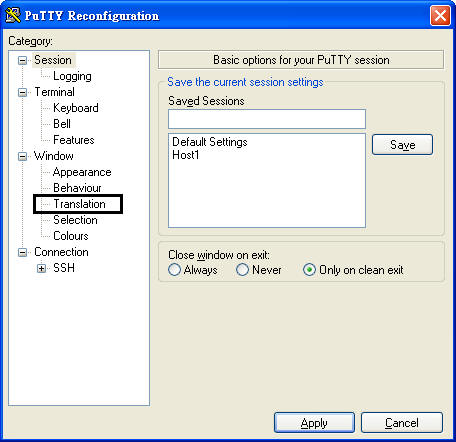
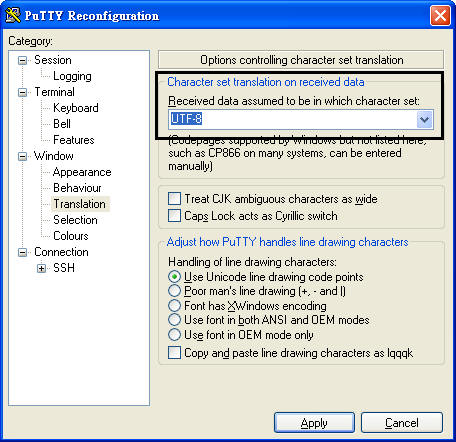
2.執行SSHSecureShellClient-3.2.9.exe

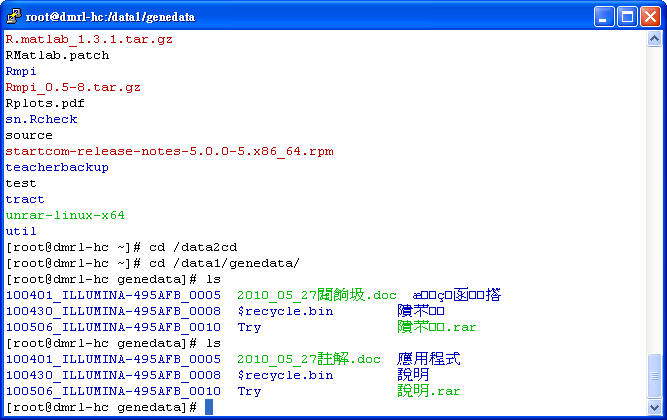
3.產生「SSH Secure Shell Client」及「SSH Secure File Transfer Client」。前者可以登入cluster，後者可以在pc及cluster間進行檔案傳輸

4.執行「SSH Secure Shell Client」，點選File → Quick Connect



#### *Q:在Linux系統上，如何讓檔案顯示中文名稱:*

由於 MS 系統預設編碼模式為 Big5 對於中文顯示上與國際慣用的 UTF-8 不同，因此連線主機後，如下圖所示，再 Putty 邊框上按下滑鼠右鍵：  
  
蹦出的選單請選擇 " Change Settings "，會出現下圖：  
  
接著，請於 Window 分頁中點選 Translation項目，開啟編碼頁面：  
  
  
將 Character 改為 UTF-8 即可，底下為修改前後的差異



* ***Linux 簡易操作指令***
* *一般使用者:*

|  |  |
| --- | --- |
| **系統瀏覽** | **說明** |
| [scchen@stat-hcserver~]$ | 登入後所在的位置 /home/scchen |
| [scchen@stat-hcserver ~]$**w** | 全部使用者的情況 |
| [scchen@stat-hcserver ~]$**rsh** *node04*  Last login: Sat May 3 12:33:02 from stat-hcserver.cluster | rsh 是在node04的機器上執行某些指令，而把結果傳回 local host |
| [scchen@stat-hcserver ~]$ qstat | 觀看有幾個job在nodes上執行工作 |
| [scchen@stat-hcserver ~]$ ssh –X node04 | 在noe04 中使用X window(進入node前輸入此指令，會直接進入node) |
| [scchen@stat-hcserver ~]$ pbsnodes -a | 檢查各nodes 的狀態 |
| [scchen@stat-hcserver ~]$ top | 查看目前所在node的狀態與記憶體使用量 |
| [scchen@stat-hcserver ~]$ more cc.txt | 查看檔案cc.txt的內容 |

|  |  |
| --- | --- |
| **檔案管理指令** | **說明** |
| [scchen@node04 ~]$**pwd** | 查尋所在的位置 |
| [scchen@node04 ~]$**mkdir** *newfile* | 建立名為newfile的資料夾 |
| [scchen@node04 ~]$**ls** | 列出目錄下的檔案及子目錄 |
| [scchen@node04 ~]$**cd** *newfile* | 改變工作目錄至newfile |
| [scchen@node04 newfile]$**cd** **..** | 回到上一層工作目錄 |
| [scchen@node04 ~]$**vi** *newdata.txt* | 利用vi指令可開啟或建立檔案 |
| [scchen@node04 ~]$**rm** *newdata.txt* | 移除檔案 |
| [scchen@node04 ~]$**rmdir** *newfile* | 移除newfile資料夾 |
| [scchen@node04 /]$**cd /** | 改變工作目錄至根目錄 |
| [scchen@node04 /]$**cd** /home/scchen | 改變工作目錄至/home/scchen |
| [scchen@node04 /]$**cp** *file1 file2* | 將file1複製至file2 |
| [scchen@node04 /]$**find / -name** *test* | 在根目錄下名字包含test的檔案及子目錄名稱都會被列出來 |
| [scchen@node04 /]$ tar -zxvf 檔案名稱  其他解壓縮方式可參考下列網站:  <http://www.dotblogs.com.tw/phoenixwu/archive/2008/05/22/4103.aspx> | 在該目錄下，執行解壓縮方式 |
| [scchen@node04 /]$ du –m –s /home/使用者名稱 | 察看目錄內所有檔案使用掉的空間的情況，-m 或-s擇一使用。 |
| [yfwu@node02 s\_4\_1\_sequence.txt]$ df | 查詢目前各硬碟使用狀況 |
| [yfwu@node02 /]$ ls -al | 查詢目前資料夾權限 |
| [yfwu@node05 data4]$ du -m | 查詢/data4/底下的所有檔案大小 |

|  |  |
| --- | --- |
| **執行程式** | **說明** |
| [scchen@node04 ~]$ **$F90 $F90FLAGS** *filename***.f – o** *filename* **$LINK\_FNL\_STATIC** | Compiler fortran using IMSL |
| [scchen@node04 ~]$**f77** *filename***.f –o** *filename* **-lm** | Compiler fortran 77 |
| [scchen@node04 ~]$**./***filename* | 執行程式 |
| [scchen@node04 ~]$**./***filename* **&** | 直接將程式送至背景執行 |
| [scchen@node04 ~]$**./***filename*  [1]+ Stopped ./shift  [scchen@node04 ~]$**bg** | 按**Ctrl+z**將正在執行的程式送進背景執行，再按**bg** 送至背景執行 |
| [scchen@node04 ~]$**ps**  PID TTY TIME CMD  1195 pts/1 00:00:34 shift | 或**ps aux**查尋processes |
| [scchen@node04 ~]$**kill -9** *1195* | 終止Process ID (PID)的執行 |
| [scchen@node04 ~]$**exit** | 離開 |
| [yfwu@node03 ~]$ **time** 程式指令 | 紀錄程式執行時間 |

\*\*\*\*如何觀察程式執行時所使用memory量

Case1.

以crm查看memory方式

Case2.

top -d 1去觀看每秒使用情況,要跳出時按q即可  
或是watch -n 1 free去查看每秒記憶體使用情況要跳出時要  
按ctrl+c

\*\*\*執行程式時, 若data放置的位置與程式執行位置在不同硬碟(EX:data1/&data2/)容易造成執行失敗或是出現錯誤

* ***vi使用說明***

若需使用vi來編輯文字檔案(例如.txt 或 .sh)

Step 1: 鍵入 vi cc.txt 即可建立並開啟檔案cc.txt

Step 2: 按 Insert 即可開始編輯。

Step 3: 編輯完畢後，按 Esc + :w 再按Enter即可存檔。

Esc + :q 再按Enter即可離開。

Esc + :wq 再按Enter即可存檔並離開。

備註: 此vi編輯器的取代性很高，若無特別需要，一般直接在windows環境下直

接利用記事本編輯 .txt 或 .sh檔，再利用FTP軟體上傳至cluster，會比較

方便與快速。

備註: 若需要更多vi指令，可參考。

# *3.應用軟體與程式*

在Hyper Cluster 下所使用的程式如下:

程式

[**R**](#R)

[**FORTAIN**](#fortran)

[**MATLAB**](#matlab)

[**其他**](#其他)

* ***執行R***

在目錄下輸入 *R* 會開啟預設之R-3.0.2 ;

而使用方式與一般在Windows使用R的操作方式相同,可以在putty下直接

鍵入欲操作指命。

* 欲安裝R package，請洽系統管理者。

## *R (S –Plus) Programs*

### *一般使用者常見的問題:*

* **Batch**

*R*

使用Cluster 39時：

#### *Q: 如何將已編輯好的程式(\_\_.R)檔送到CLUSTER WORK呢?*

ANS:

1. 一次送一個JOB

Step1: 將以編輯好的程式(\_\_.R)檔上傳到Cluster，

假設檔名為 test.r，上傳路徑為 /home/hyfu/Rcode/

Step2: 編寫\_\_\_.sh檔，並將其檔案上傳到Cluster，

假設檔名為 test.sh，上傳路徑為 /home/hyfu/Rcode/sh/

#PBS -l nodes=node04:ppn=1

cd /home/hyfu/Rcode/

R CMD BATCH test.r test.out

#若要使用其他版本R，可將"R CMD" 改為 "R-3.1.0CMD"

Step3: 執行 test.sh

[hyfu@stat-hcserver ~]$ qsub /home/hyfu/Rcode/sh/test.sh

P.S~查看是否有執行成功

[hyfu@ stat-hcserver ~]$ qstat

Job id Name User Time Use S Queue

---------------- ---------------- ---------------- -------- - -----

6263. stat-hcserver test.sh hyfu 00:00:12 R workq

P.S~想刪除正在執行的JOB

[hyfu@ stat-hcserver ~]$ qdel 6263

P.S~想查看 test.out

[hyfu@ stat-hcserver ~]$ more test.out

R version 2.8.1 (2008-12-22)

Copyright (C) 2008 The R Foundation for Statistical Computing

ISBN 3-900051-07-0

……

1. 一次送多個JOB

Step1: 將以編輯好的程式(\_\_.R)檔上傳到Cluster，

假設檔名為 test\_job22.r、test\_job23.r、…、test\_job26.r，

上傳路徑為 /home/hyfu/Rcode/

Step2: 編寫\_\_\_.txt檔，並將其檔案上傳到Cluster。檔內的四個參數為

1.指定運算節點的機器名稱

2.程式欲執行的記憶體數量

3.程式的input檔案

4.程式的output檔案.

假設檔名為 test.txt，上傳路徑為 /home/hyfu/Rcode/txt/

node04 5G test\_job25.r test\_job25.out

node05 5G test\_job26.r test\_job26.out

Step3: 編寫\_\_\_.sh檔，並將其檔案上傳到Cluster，

假設檔名為 test.sh，上傳路徑為 /home/hyfu/Rcode/sh/

#!/bin/bash

while read P1 P2 P3

do

sleep 1s

JOB= `qsub -N pbs.sh -l nodes=${P1}:ppn=1 << EOJ

cd /home/hyfu/Rcode/

R CMD BATCH ${P2} ${P3}

EOJ

`

echo "$JOB"

done < /home/hyfu/Rcode/txt/test.txt

exit

Step4: 執行 test.sh

[hyfu@ stat-hcserver ~]$ chmod 755 /home/hyfu/Rcode/sh/test.sh

[hyfu@ stat-hcserver ~]$ /home/hyfu/Rcode/sh/test.sh

6267. stat-hcserver

6268. stat-hcserver

#### *Q: 如何在CLUSTER中使用Rmpi平行計算程式?*

ANS:

目前Rmpi測試之後是可以使用,測試的步驟如下：

(1)請用一般使用者的身份執行下列步驟

先編寫一個檔案,此範例是使用bhost為檔名,檔名部份使用者可以自己取名

語法是：

機器名稱 cpu=數量

機器名稱 cpu=數量

.... cpu=數量

stat-hcserver cpu=數量 (stat-hcserver一定要寫入)

範例如下:

[infowrap@stat-hcserver ~]$ cat bhost

node04.cluster cpu=8

node03.cluster cpu=8

stat-hcserver cpu=8

(2)在執行R之前先使用lam,用法如下：

lamboot -v filename (filename就是步驟一之檔名)

[infowrap@stat-hcserver ~]$ lamboot -v bhost

LAM 7.1.2/MPI 2 C++/ROMIO - Indiana University

n-1<17020> ssi:boot:base:linear: booting n0 (node04.cluster) n-1<17020> ssi:boot:base:linear: booting n1 (node03.cluster) n-1<17020> ssi:boot:base:linear: booting n2 (stat-hcserver) n-1<17020> ssi:boot:base:linear: finished

(3)執行R

若要執行平行程式,請先下library(Rmpi),整個流程如下：

> library(Rmpi)

> mpi.spawn.Rslaves(nslaves=8)

8 slaves are spawned successfully. 0 failed.

master (rank 0, comm 1) of size 9 is running on: stat-hcserver

slave1 (rank 1, comm 1) of size 9 is running on: node04

slave2 (rank 2, comm 1) of size 9 is running on: node04

slave3 (rank 3, comm 1) of size 9 is running on: node04

slave4 (rank 4, comm 1) of size 9 is running on: node04

slave5 (rank 5, comm 1) of size 9 is running on: node04

slave6 (rank 6, comm 1) of size 9 is running on: node04

slave7 (rank 7, comm 1) of size 9 is running on: node04

slave8 (rank 8, comm 1) of size 9 is running on: node04

> mpi.remote.exec(mpi.get.processor.name())

$slave1

[1] "node04"

$slave2

[1] "node04"

$slave3

[1] "node04"

$slave4

[1] "node04"

$slave5

[1] "node04"

$slave6

[1] "node04"

$slave7

[1] "node04"

$slave8

[1] "node04"

> n<-3

> mpi.remote.exec(double,n)

X1 X2 X3 X4 X5 X6 X7 X8

1 0 0 0 0 0 0 0 0

2 0 0 0 0 0 0 0 0

3 0 0 0 0 0 0 0 0

> mpi.quit()

MPI\_Recv: process in local group is dead (rank 1, comm 4)

MPI Recv: process in local group is dead (rank 4, SSI:coll:smp:local comm for CID 7

MPI Recv: process in local group is dead (rank 1, SSI:coll:smp:local comm for CID 7

MPI\_Recv: process in local group is dead (rank 2, SSI:coll:smp:local comm for CID 7

(4)執行完後請下lamhalt指令

[infowrap@stat-hcserver ~]$ lamhalt