บทที่ 12 การทำ Linux Cluster

เครื่องมือและอุปกรณ์

ก่อนที่จะเริ่มทำระบบ Linux Cluster นั้นควรจะมีการเตรียมอุปกรณ์ต่าง ๆ ให้พร้อมเสียก่อน เพื่อความสะดวกในการปฏิบัติงาน สำหรับอุปกรณ์ต่าง ๆ ที่ใช้มีดังนี้

1. เครื่องหลักหรือเครื่อง Server จำนวน 1 เครื่อง
   - CPU 486 DX2 100 MHz
   - RAM 16 M
   - Hard Disk 1.7 G
   - Lan Card
   - CD-ROM
   - ฟลอปปีดิสก์ไดรฟ์
2. เครื่องลูกข่ายหรือเครื่อง Client จำนวน 1 เครื่อง หรือ มากกว่าก็ได้
   - CPU 486 DX2 66 MHz
   - RAM 12 M
   - Hard Disk 1.2 G
   - Lan Card เครื่องละ 1 ตัว
   - ฟลอปปีดิสก์ไดรฟ์
3. Swiching hub (10/100 BaseT) เป็นอุปกรณ์เชื่อมระหว่างเครื่อง Server และเครื่องลูกข่าย
4. สาย UTP เป็นสายเชื่อมระหว่างเครื่อง Server และ เครื่อง Client ในระบบ โดยต่อเข้ากับ Hub และ Lan Card ของเครื่อง Server และเครื่อง Client
5. ซอฟต์แวร์ที่ใช้ - ใช้ระบบปฏิบัติการ Linux RedHat (Linux TLE 3)
   - MC (Midnight Commander)
   - NFS (Network File System)
   - Crond Deamon

- MPI (Message Passing Interface) เป็นโปรแกรมที่ใช้พัฒนาการเขียนโปรแกรมแบบขนาน
- SCMS (SMILE Cluster Management System) เป็นโปรแกรมที่ใช้บริหารระบบ Cluster

หมายเหตุ

- เครื่องคอมพิวเตอร์ที่ใช้อาจจะมากหรือน้อยกว่านี้ก็ได้ แต่อย่างน้อยที่สุดคือ 2 เครื่องโดยแบ่งเป็นเครื่อง Server ส่วนอีกเครื่องหนึ่งที่เหลือจะเป็นเครื่อง Client
- เครื่องคอมพิวเตอร์อาจจะเปลี่ยนแปลงไปตามความเหมาะสม โดยเฉพาะเครื่อง Server ควรใช้เครื่อง ที่มี Processor ที่มีความเร็วสูงและมี RAM มาก
- Lan Card ที่ใช้ต้องมี ROM (ROM ธรรมดาไม่ใช่ ROM boot)
- CD-ROM ของเครื่องลูกข่ายไม่จำเป็นต้องมีก็ได้อาจจะติดตั้งที่เครื่อง Server เพียงเครื่องเดียว แล้ว ใช้งานร่วมกันผ่านระบบเน็ตเวิร์กได้ แต่ฟลอปปีดิสก์ไดรฟ์ควรจะมีทุกเครื่อง เพื่อที่จะสามารถบูทจากแผ่นไปแก้ไขระบบได้ง่ายในกรณีเกิดปัญหาขึ้น

ขั้นตอนการติดตั้งระบบ Linux Cluster

เมื่อเตรียมเครื่องมือและอุปกรณ์ต่าง ๆ สำหรับติดตั้งระบบ Linux Cluster เรียบร้อยแล้ว ขั้นตอนต่อไปก็จะ เป็นการติดตั้งระบบใช้งาน โดยจะกล่าวถึงการติดตั้งระบบทั้งหมด 2 โหนด โดยมีโหนด xserver (เครื่อง Server) เป็นเครื่องตัวหลักทำหน้าที่เป็นไฟล์เซิร์ฟเวอร์และไฟล์วอลล์ของระบบส่วนที่เหลือเป็นโหนดลูกข่าย คือ xclient1 ทำหน้าที่เป็นส่วนช่วยสนับสนุนการประมวลผลข้อมูล สำหรับขั้นตอนการติดตั้งระบบ Linux Cluster นั้น แบ่งได้ดังนี้

1. การติดตั้งเครื่อง Server (xserver)
2. การติดตั้งเครื่อง Client (xclient1)
3. การติดตั้งโปรแกรมแบบขนาน MPI
4. การติดตั้งโปรแกรมบริหารระบบ SCMS

การติดตั้งเครื่อง Server (xserver)

1. ลงระบบปฏิบัติการ Linux ที่เครื่อง xserver เลือกประเภท Server Type
2. ในขั้นตอนการเซต network ให้ทำการเลือกการ์ดแลนด้วย เช่น NE2000 compatible แต่ก่อนอื่น ควรเซตหาค่า I/O และ IRQ ในโหมด DOS ก่อนเพื่อให้ทราบค่า I/O และ IRQ
3. เมื่อใส่ค่า I/O และ IRQ เรียบร้อยแล้ว กำหนดค่าต่าง ๆ ดังนี้

```
hostname : xserver
domainname : bj.ac.th
IP address : 10.0.0.99
netmask : 255.255.255.0
root 's password : linuxserver
```

เมื่อใส่ password แล้วรอจนกว่าลงระบบ Linux เสร็จ

4. เข้าสู่ระบบ Linux แล้วลงโปรแกรมที่ใช้งานในเครื่อง xserver โดยลงจากแผ่น CD Linux TLE 3 โดยใช้คำสั่งทำดังนี้

```
# mount /dev/cdrom /mnt
# cd /mnt/RedHat/RPMS
# rpm -i cro*
# rpm -i inet*
# rpm -i mc*
```

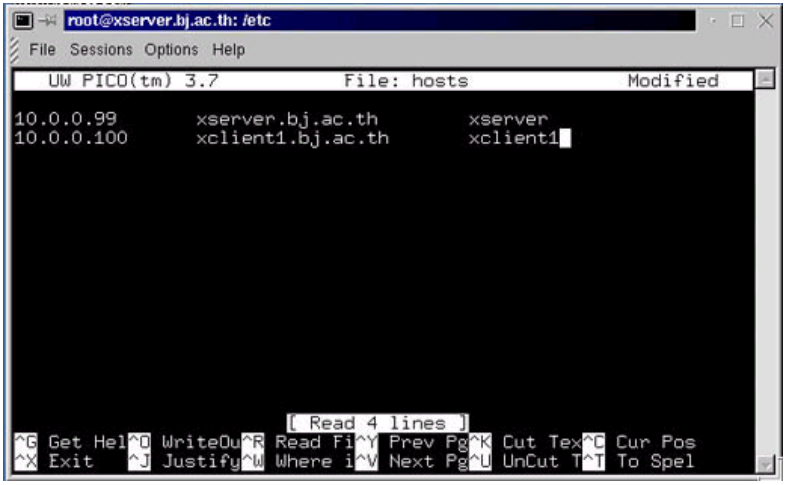
5. เพิ่มชื่อเครื่อง xserver และเครื่องลูกข่าย โดยแก้ไขในไฟล์ /etc/hosts ดังนี้

```
# pico /etc/hosts
```

พิมพ์ชื่อเครื่อง xserver และ เครื่องลูกข่ายลงไปให้ครบ ดังนี้

```
10.0.0.99 xserver xserver.bj.ac.th
10.0.0.100 xclient1 xclient1.bj.ac.th
```

เสร็จแล้วเซฟ ดังรูป



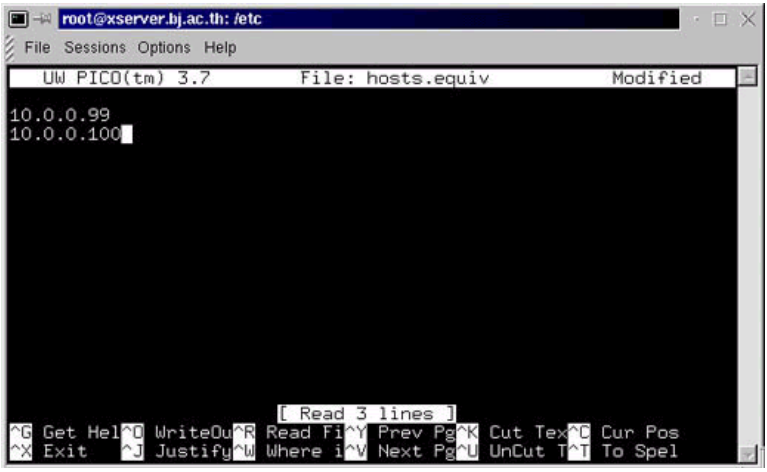
จากนั้นสร้างไฟล์ /etc/hosts.equiv และ /root/.rhosts ขึ้นมาใหม่แล้วทำการเพิ่มเฉพาะชื่อ IP address ของเครื่องทุกเครื่องลงไป โดย

```
# pico /etc/hosts.equiv
```

แล้วเพิ่มชื่อ IP address ของทุกเครื่องลงไป

```
10.0.0.99
10.0.0.100
```

ดังรูป



และ

```
# cp /etc/hosts.equiv /root/.rhosts
```

แล้วเพิ่มชื่อ IP address ของทุกเครื่องลงไปเช่นเดียวกับไฟล์ /etc/hosts.equiv

6. เซต NFS (Network File System) and Crond Daemon

6.1 ก่อนทำการเซตควร ทำการเปิดใช้บริการ (System service) โดยสั่ง

```
# setup
```

6.2 เลือกหัวข้อ System service แล้วทำเครื่องหมาย * ที่ crond , nfs และ inet

ข้อแนะนำ *ทำเครื่องหมาย * ในบริการทั้งหมดก็ได้*

6.3 สร้าง directory ชื่อ x ที่ /usr เพื่อใช้ในการแชร์ไฟล์ให้กับทุกเครื่องในระบบ โดยสั่ง

```
# mkdir /usr/x
```

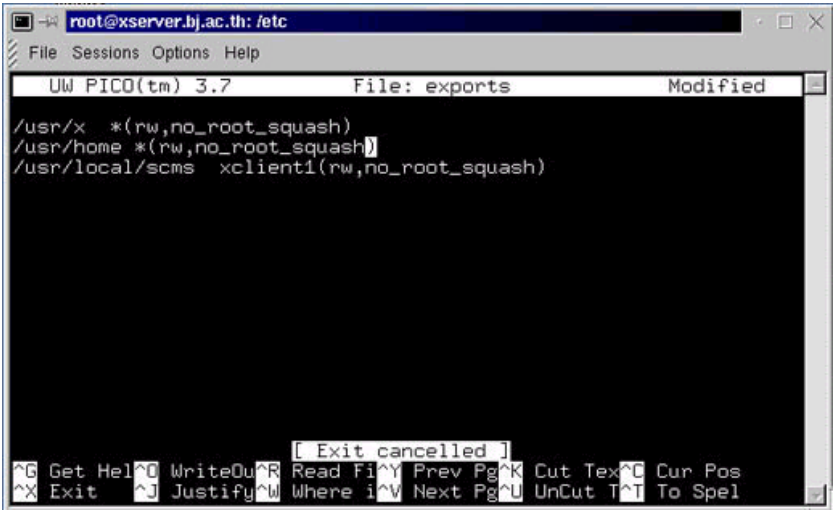
6.4 ทำการแก้ไขข้อมูลในไฟล์ /etc/exports เพื่อให้ทุกเครื่องใช้ห้อง x ,ห้อง home และห้อง scms เป็น sharedirectory โดยสั่ง

```
# pico /etc/exports
```

แล้วเพิ่ม

```
/usr/x *(rw,no_root_squash)
/home *(rw,no_root_squash)
/usr/local/scms *(rw,no_root_squash)
```

ลงไป (บรรทัดนี้สำหรับผู้ที่ติดตั้งโปรแกรม scms) เสร็จแล้วเซฟ



6.5 ทดสอบการใช้งานของ NFS โดยไปที่ /etc/rc.d/init.d แล้วใช้คำสั่ง

```
# ./nfs stop
# ./nfs start
```

หากขึ้น [OK] ทั้งหมดก็ใช้ได้

สำหรับการติดตั้งเครื่อง xserver ก็มีรายละเอียดเพียงเท่านี้ ขั้นตอนต่อไปจะเป็นการติดตั้งเครื่อง Client เพื่อใช้เป็นส่วนช่วยเพิ่มความเร็วในการประมวลผลข้อมูล

การติดตั้งเครื่องลูกข่าย (xclient)

1. ลงระบบปฏิบัติการ Linux ที่เครื่อง xclient เลือกประเภท Server Type
2. เซตการ์ดแลนตามชนิดของไดร์เวอร์ พร้อมตั้งค่า I/O, IRQ และกำหนดค่าต่าง ๆ ดังนี้

```
hostname : xclient1
domainname : bj.ac.th
IP address : 10.0.0.100
netmask : 255.255.255.0
root's password : linuxclient
```

3. ทำการลงโปรแกรมเพิ่ม คือ mc จากแผ่น CD Linux TLE 3
4. แก้ไขไฟล์ /etc/hosts แล้วเพิ่มชื่อเครื่อง xserver และ xclient ไปจนครบ ดังนี้

```
# pico /etc/hosts
```

เพิ่มชื่อเครื่องในระบบให้ครบ

```
10.0.0.99 xserver xserver.bj.ac.th
10.0.0.100 xclient1 xclient1.bj.ac.th
```

จากนั้นสร้างไฟล์ .rhosts ที่ห้อง root แล้วเพิ่มเฉพาะชื่อของทุกเครื่องลงไป

5. เซต NFS (Network File System)

5.1 ทำการสร้างห้องที่ใช้ share file และ directory

```
# mkdir /usr/x
```

5.2 แก้ไขไฟล์ /etc/fstab

```
# pico /etc/fstab
```

แล้วเพิ่ม

```
xserver:/usr/x /usr/x nfs defaults 0 0
xserver:/home /home nfs defaults 0 0
```

เสร็จแล้วเซฟแล้วใช้คำสั่ง

```
# mount -a
```

ซึ่งเป็นการ mount file system ใหม่ทั้งหมด ทำดังนี้ทุกโหนดหากไม่มี error ใด ๆ ก็ใช้ได้

ข้อสังเกต

ที่ห้อง /usr/x ของครื่อง xclient ทุกเครื่อง สามารถ share file และ directory กับ /usr/x ที่เครื่อง xserver แต่ข้อมูลจริง ๆ ทั้งหมดจะถูกเก็บอยู่ใน Hard Disk ที่เครื่อง xserver

โปรแกรม MPI เป็นโปรแกรมสำหรับใช้ในการพัฒนาการเขียนโปรแกรมแบบขนาน โดยสามารถทำการ ดาว์นโหลดโปรแกรมได้ที่เว็บไซต์ " http:// smile.cpe.ku.ac.th/beowulf " และเพื่อเป็นการประหยัดเนื้อ ที่ฮาร์ดดิสก์ควรลงไว้เฉพาะที่เครื่อง xserver เท่านั้น เพราะเครื่องลูกข่ายในระบบสามารถ share file และ directory ได้ที่ /usr/x ในเครื่อง xserver เมื่อทำการดาว์นโหลดโปรแกรมมาพร้อมแล้วทำดังนี้

1. สร้างห้อง gz ไว้เก็บไฟล์ mpi.gz โดยใช้คำสั่ง

```
# mkdir /gz
```

2. ทำการระเบิดไฟล์ mpi.gz โดยมาที่ห้อง gz แล้วใช้คำสั่ง

```
# tar zxvf mpi.gz
```

3. จะได้ห้อง mpich-1.2.0 ขึ้นมา เข้าไปห้องนี้ แล้วทำการ config ดังนี้

```
# mkdir /usr/x/mpi
 # ./configure -prefix=/usr/x/mpi -arch=LINUX -disable-f77-disable-
    f90modules
```

ซึ่งเป็นการ config โปรแกรม mpich ลงในห้อง /usr/x/mpi โดยใช้สถาปัตยกรรมแบบ LINUX และไม่ใช้ภาษา fortran

4. ทำการ make สร้างโปรแกรม ซึ่งรายละเอียดการทำงานของคำสั่งอ่านได้จากไฟล์ make.log

```
# make >& make.log
# make install
```

เสร็จแล้ว เข้าห้อง /usr/x/mpi จะพบว่ามีไฟล์อยู่เป็นจำนวนมาก

5. แก้ไขข้อมูลในไฟล์ /usr/x/mpi/share/machine.LINUX โดย

```
# pico /usr/x/mpi/share/machine.LINUX
```

ทำการลบข้อความทั้งหมดในไฟล์ออก แล้วเพิ่มชื่อเครื่องที่ใช้ร่วมในการประมวลผลในระบบ cluster ลงไป เสร็จแล้วเซฟมาที่ console mode

การทดสอบและใช้งาน MPI เบื้องต้นเนื่องจาก MPI จะทำงานผ่าน remote shell (rsh) ดังนั้นเราต้องจะทดสอบก่อนว่า MPI ที่เราติดตั้งไป นั้นใช้งานได้หรือไม่ โดยที่เครื่อง xclient ทุกเครื่อง ลอง logout ออกแล้ว login เข้าระบบใหม่ แล้วทำดังนี้

```
# rsh xserver hostname
```

ถ้าใช้งานได้ที่เครื่อง xclient จะขึ้น xserver.bj.ac.th ส่วนที่เครื่อง xserver ลองทำเช่นเดียวกันคือ

```
# rsh xclient1 hostname
```

ถ้าใช้งานได้ก็จะขึ้น xclient1.bj.ac.th เช่น

ทดลองการคอมไพล์โปรแกรมแบบขนาน

เมื่อเรามีโปรแกรมที่เขียนแบบขนานโปรแกรมหนึ่ง สามารถทำการคอมไพล์และเรียกใช้งานได้ ดังนี้

1. เข้าห้อง /usr/x/mpi/examples แล้วสั่งคอมไพล์โดย (สมมติว่าเราต้องการคอมไพล์โปรแกรม cpi.c)

```
# ../bin/mpicc -o cpi cpi.c
```

2. จะได้ไฟล์ cpi ที่สามารถรันได้ แล้วลองรัน โดย

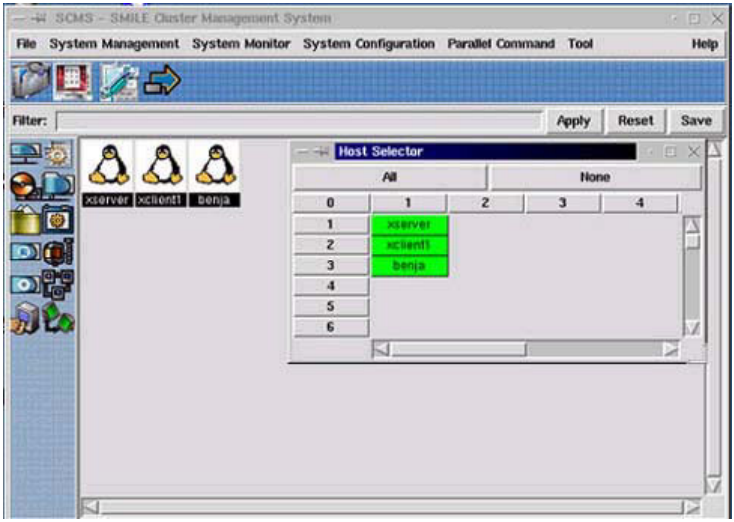
```
# ./mpirun -np 2 cpi
```

ซึ่งเป็นการรันโปรแกรม cpi โดยเรียกเครื่องในระบบมาช่วยกันประมวลผล 2 เครื่อง และเราสามารถเรียก จำนวนเครื่องมาช่วยในการประมวลผลได้โดยเปลี่ยนตัวเลขจากเลข 2 เป็นจำนวนที่ต้องการ สำหรับรายละเอียดและขั้นตอนการติดตั้งระบบ Linux Cluster ก็มีเพียงเท่านี้

หากท่านสนใจที่จะศึกษา รายละเอียดเพิ่มเติม สามารถศึกษาได้ที่ เวบไซต์ http://smile.cpe.ku.ac.th/beowulf และ เวบไซต์ http:// www-unix.msc.anl.gov/mpi/mpich

การติดตั้งและใช้งาน SCMS

โปรแกรม SCMS เป็นโปรแกรมสำหรับใช้ในการบริหารระบบสำหรับ Admin



โดยสามารถทำการดาว์นโหลดโปรแกรมได้ที่เว็บไซต์ " http://smile.cpe.ku.ac.th/software "

เมื่อทำการดาว์นโหลดโปรแกรมมาพร้อมแล้วทำดังนี้

1.ให้ copy ไฟล์ scms-1.2.2.tar.gz ไปไว้ที่ /usr/local โดยใช้คำสั่ง

```
# cp scms-1.2.2.tar.gz /usr/local
```

2. ทำการระเบิดไฟล์ scms-1.2.2.tar.gz โดยใช้คำสั่ง

```
# tar zxvf scms-1.2.2.tar.gz
```

แล้วทำการ config ดังนี้

```
# make x86 SRC=$PWD
# make install DEST=/usr/local/scms
```

3. จากนั้นเมื่อออกจากห้อง scms-1.2.2 จะเห็นห้อง scms

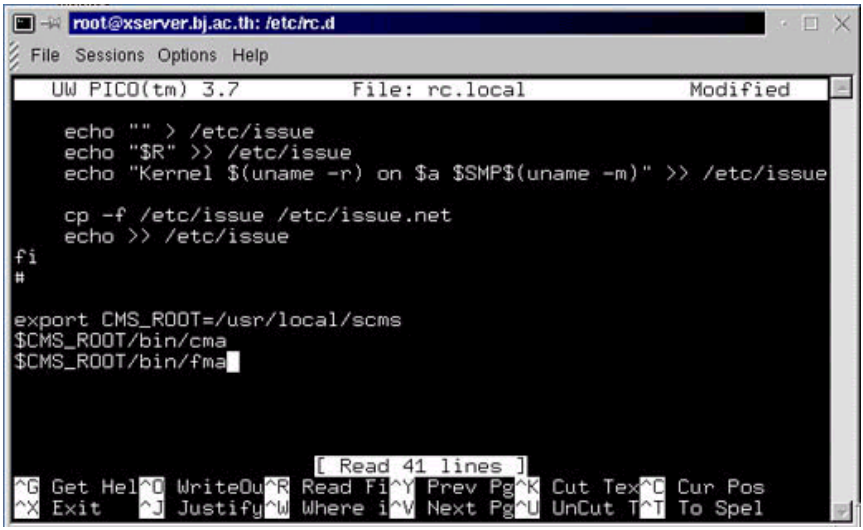
4. เข้าไปแก้ไขไฟล์ rc.local ที่ /etc/rc.d/ โดยใช้คำสั่ง

```
# pico /etc/rc.d/rc.local
```

แล้วเพิ่ม

```
export CMS_ROOT=/usr/local/scms
    $CMS_ROOT/bin/cma
    $CMS_ROOT/bin/fma
```

ดังรูป



สำหรับเครื่องลูกข่ายนั้นให้ไปแก้ไฟล์ที่ /etc/fstab โดยเพิ่มบรรทัดต่อไปนี้ลงไป

```
xserver:/usr/local/scms /usr/local/scms nfs default 0 0
```

และแก้ไขไฟล์ที่ /etc/rc.d/rc.local แล้วเพิ่ม

```
export CMS_ROOT=/usr/local/scms
$CMS_ROOT/bin/cma
```

เข้า X-Window แล้วเรียก #scms ผ่านทาง terminal โปรแกรมก็จะถูกเรียกขึ้นมา

การทดสอบและใช้งาน MPI เบื้องต้น

เนื่องจาก MPI จะทำงานผ่าน remote shell (rsh) เราจะทดสอบก่อนว่าใช้งานได้หรือเปล่า โดยที่เครื่อง xclient ทุกเครื่อง ลอง logout แล้ว login เข้าระบบใหม่ แล้วทำดังนี้

```
#rsh xserver hostname
```

- ถ้าได้จะขึ้น xserver.bj.ac.th ส่วนที่เครื่อง xserver สั่ง

```
# rsh xclient1 hostname
```

- ถ้าได้จะขึ้น xclient1.bj.ac.th เหมือนกัน

หมายเหตุ เราสามารถใช้งานในฐานะ root เท่านั้น

ทดลองการคอมไพล์โปรแกรมแบบขนาน

1. เข้าห้อง /usr/x/mpi/examples เราจะคอมไพล์ cpi.c โดยสั่ง

```
# ../bin/mpicc -o cpi cpi.c
```

2. จะได้ไฟล์ cpi ที่สามารถรันได้ แล้วลองรันโดย

```
# ./mpirun -np 2 cpi
```

- เป็นการรันโปรแกรม cpi โดยเรียกเครื่องในระบบมาช่วยกัน ประมวลผล 2 เครื่อง

การพัฒนา SBC (Small Benja Cluster) ในขั้นต่อไป

1. เปลี่ยนจาก Hub ธรรมดา (ความเร็ว 10 mbps ที่ใช้อยู่เป็นswitching Hub หรือ Fast ethernet switching ความเร็ว 100 mbps)

2. พัฒนาไปสู่เครือข่ายความเร็วสูงอย่างแท้จริง เช่น ใช้ Fiber Optic

3. พัฒนาโปรแกรมแบบขนาน โดยใช้ MPI ให้รัน โปรแกรมได้ในโหมด X window จะเป็นการใช้ศักยภาพของ cluster มากยิ่งขึ้น

4. เครื่องลูกข่ายไม่ต้องใช้ Hard Disk โดยนำ โปรแกรม diskless มาใช้ โดยให้เครื่องลูกข่ายบูตเครื่องจาก แผ่นหรือ Rom boot สำหรับ Linux จากการ์ดแลน

5. เครื่องที่นำมาใช้ในระบบ cluster นี้ควรเป็น ไมโครคอมพิวเตอร์สมรรถนะสูง เช่น CPU Pentium II 500 MHz , RAM 128 M เป็นจำนวนมาก เช่น 8,16,32,64,128 หรือ 256 เครื่อง เป็นต้น

ข้อจำกัดที่พบ

1. เนื่องจากได้ลองใช้โปรแกรม diskless แล้ว แต่ไม่สำเร็จเพราะโปรแกรมมีข้อจำกัด ที่ลูกข่ายต้องมี Rom- boot สำหรับบูต Linux ในการ์ดแลน
2. โปรแกรม diskless ต้องใช้กับเครื่องที่เป็น Pentium ขึ้นไป
3. สามารถศึกษาเพิ่มเติมที่ http://smile.cpe.ku.ac.th และที่ http://www.beowulf.org

ใบงานที่ 12

ชื่อ…… ชั้น……

เลขที่……คะแนน…….

1. ครูอธิบายหลักการทำงานของเครื่องคอมพิวเตอร์แบบขนาน
2. จากนั้นอธิบายความแตกต่างระหว่างคอมพิวเตอร์ที่ใช้กันอยู่กับคอมพิวเตอร์ที่จะต่อแบบขนาน
3. ครูอธิบายความหมายของคำว่า Linux Cluster และโปรแกรมที่ใช้ทำ
4. ครูแจกใบความรู้ที่ 12
5. ครูให้นักเรียนแบ่งกลุ่ม โดยแต่ละกลุ่มจะมีเครื่องคอมพิวเตอร์ 3 เครื่อง ดังนี้
    1. เครื่องตัวแม่ 1 เครื่อง
    2. เครื่องตัวลูก 2 เครื่อง
6. จากนั้นให้นักเรียนติดตั้งระบบ Network
7. ทำการติดตั้ง Linux Cluster
8. ทดสอบการใช้งาน และให้คะแนน