

PC-Cluster 架設圖解教學
作業系統-linux-openSuSE 11.2

編寫者：吳泰賢(TAI HSIEN WU)

2010/11/15 更新

目錄

1. 環境介紹	P. 3
2. 架構圖	P. 3
3. 各種 Server 介紹	P. 3
4. 設定 Master	P. 4
5. 設定 Nodes	P. 10
6. 設定 ssh 免密碼登入	P. 12
7. 額外設定	P. 12
8. 附錄	P. 12

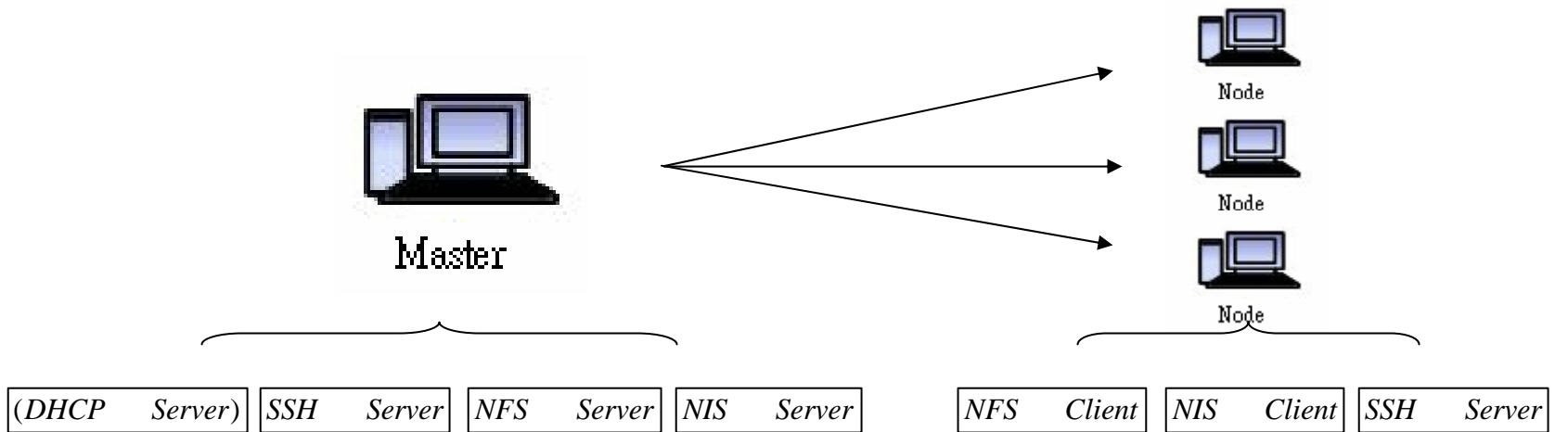
1. 環境介紹：

(1)OS：openSuSE 11.2 64bit GNOME Desktop-English

(2)Computer 數量：Master 一台(主機名稱：master)、Node 一台(主機名稱：node2)

(3)Master 有兩張網路卡，對外為固定 IP，對內為虛擬 IP：192.168.1.1

2. 架構圖：



3. 各種 Server 介紹：

(1) DHCP Server → 可製造虛擬 ip 分配給 Nodes

(2) NFS Server → 可開放本機的磁碟給 Nodes 存取

(3) NIS Server → 可將本機的帳號分享至其他 Nodes

(4) SSH Server → 開放 SSH 遠端連線功能，當中 Master 的功能為對外連線，而 Nodes 的功能則為和 Master 溝通的橋樑

* NIS Server 的功能是將帳號分享給其他的 Slave，若你的 Nodes 很少，亦或是帳號只有一個，功能確實不大，不用安裝也可架起 Cluster，但若你的(Nodes)很多，帳號又不只一個的時候，NIS Server 確實扮演著重要的角色。

* DHCP Server 並非必要，可以手動設定 IP 位置。

* 本篇文章**粗體**即為終端機下輸入的指令。

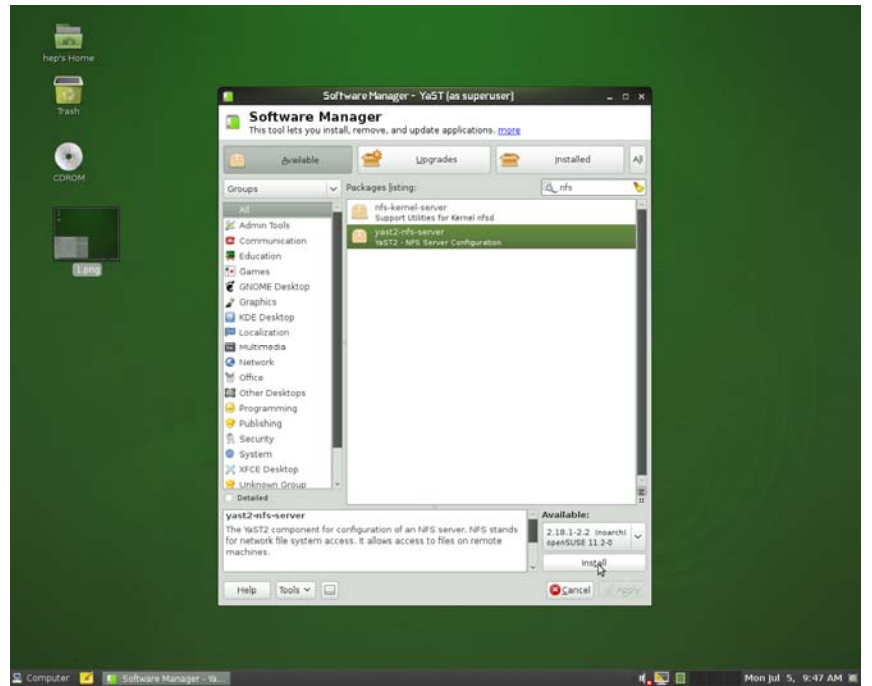
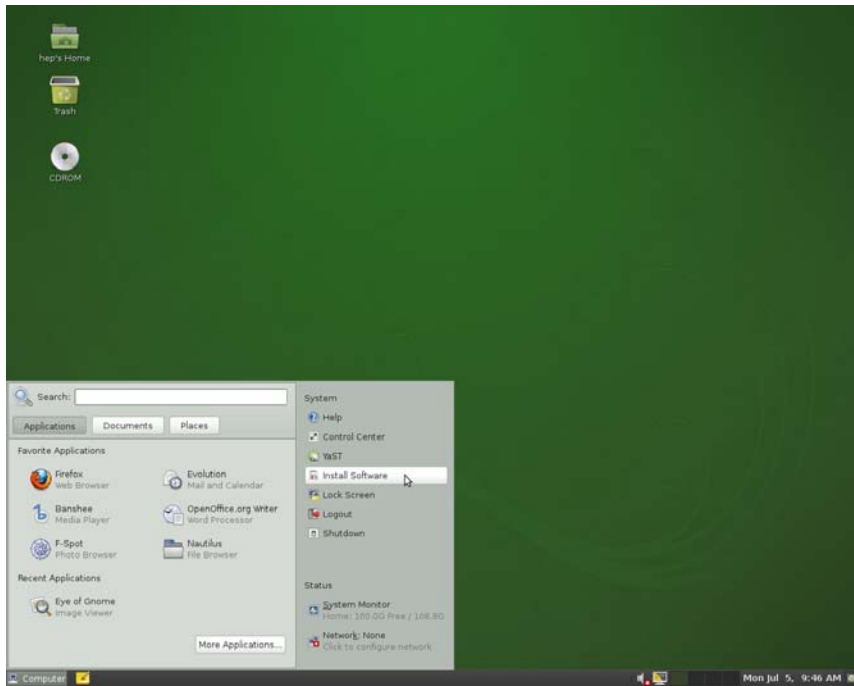
* 筆者在大專期間，實驗室需要，自行上網查詢相關資料(如：鳥哥的私房菜、Marty's Linux Cluster 架設日誌等...)而在實驗室架設出基本的 cluster，經過一段時間的磨練和成長，解決了一些零零散散的問題，累積心得進而寫出這邊 cluster 架設教學文章，不過學海無涯，深深感到自己對 linux 系統的了解只是滄海一粟，若在下述中有任何的錯誤或改進，歡迎來信共同討論和進步，以下是小弟的聯絡方式。

高雄師範大學物理學系 吳泰賢(TAI HSIEN WU) anigiss@hotmail.com

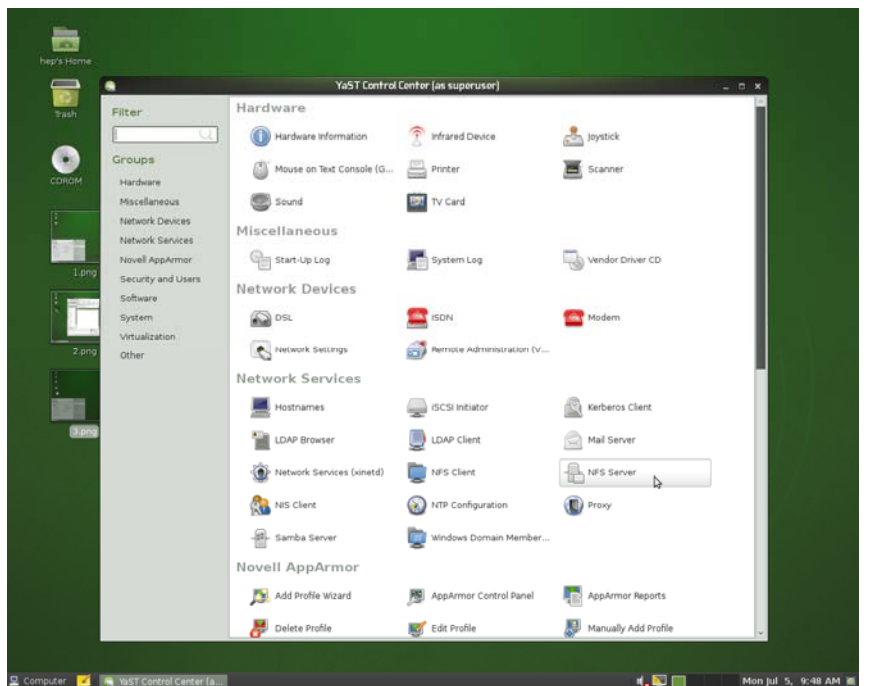
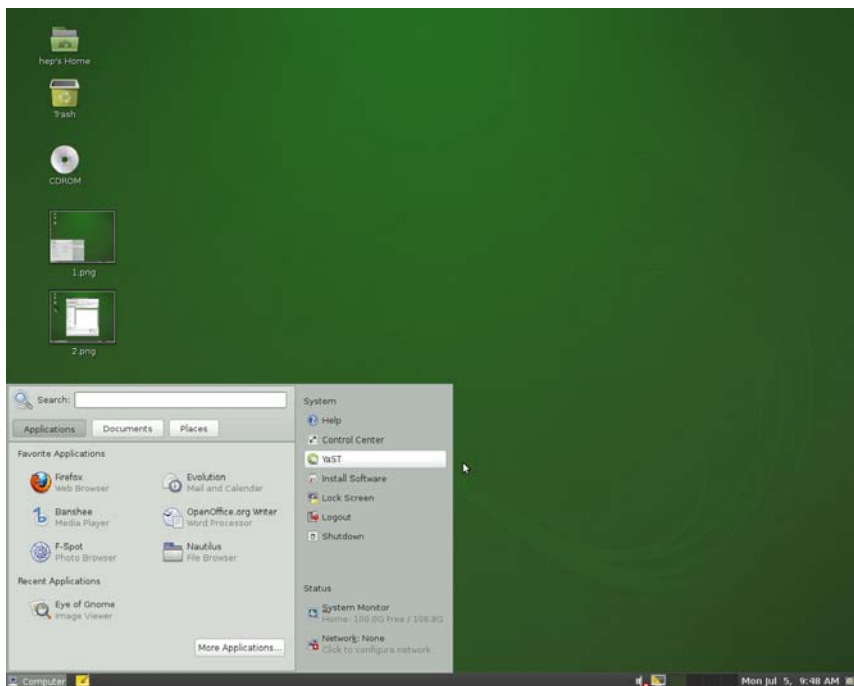
4. 設定 Master :

(1) 設定 NFS Server

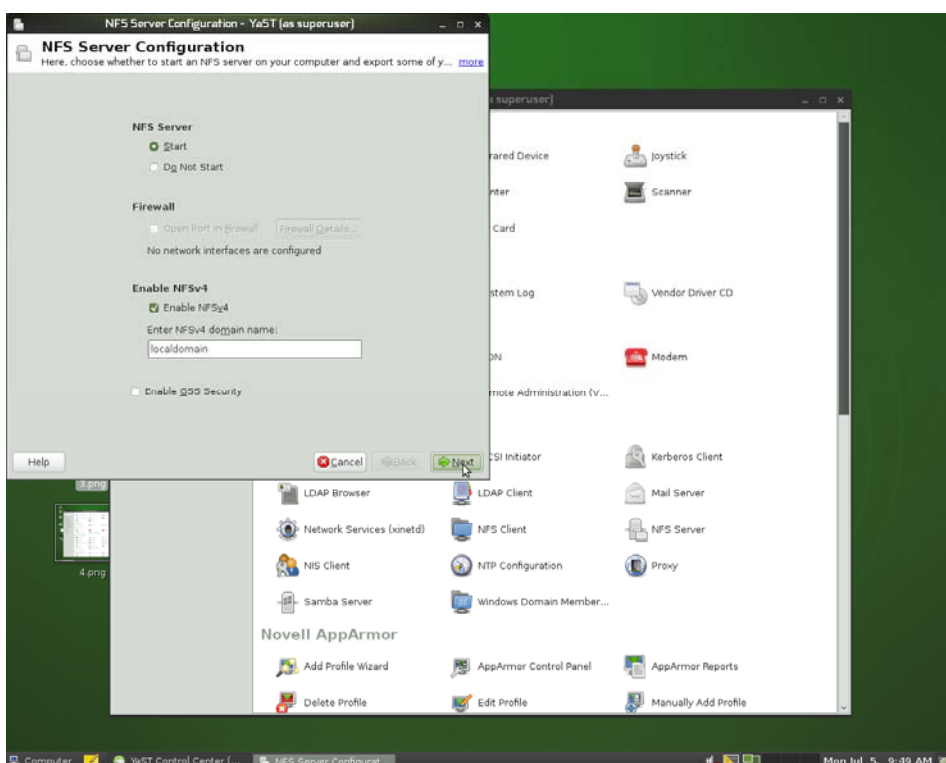
<1>利用「Install Software」功能，搜尋「nfs」，安裝「yast2-nfs-server」。



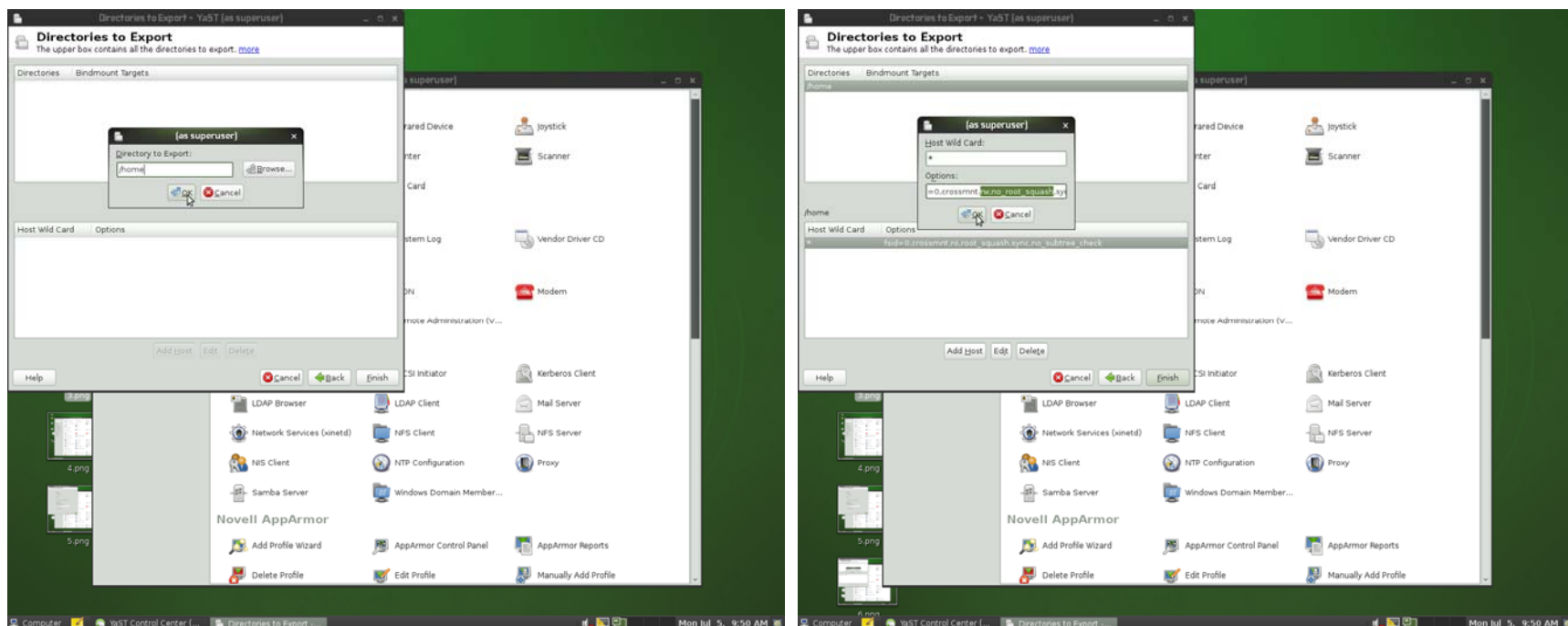
<2>利用「YaST」功能，進入 NFS Server 設定。



<3>依照下圖畫面設定，並選擇「Next」，其中 domain name 可自行設定，此處採預設狀態。

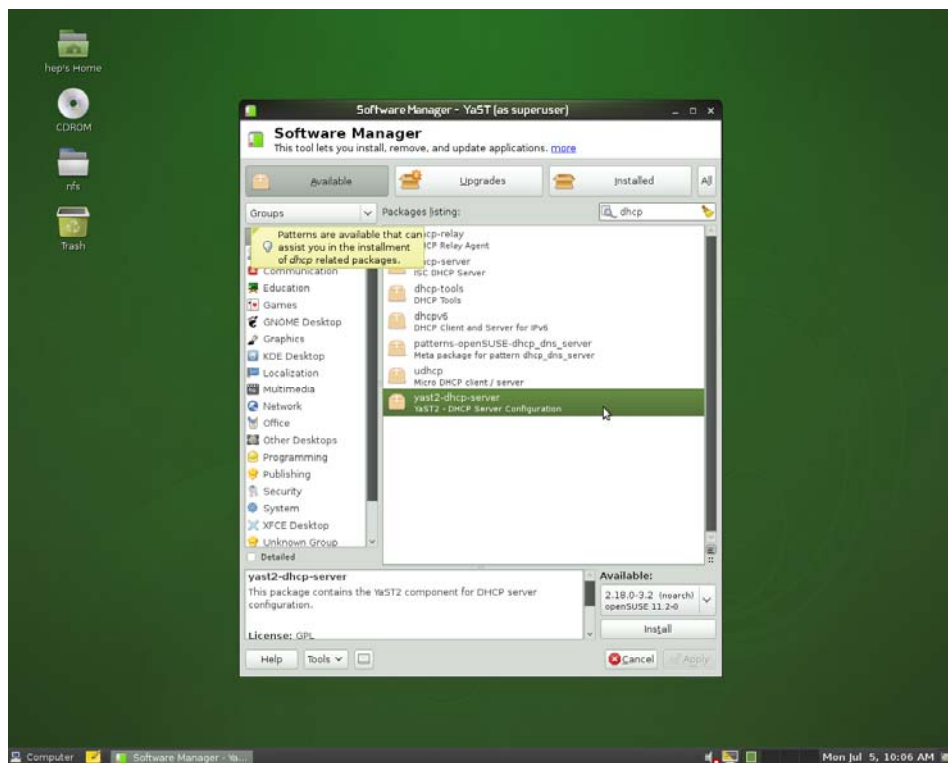


<4>選擇欲分享出去的目錄，以下以/home 為範例，輸入完分享目錄，會出現設定參數選項，「Host Wild Card」預設為「*」，表示分享給任何一台電腦，而「Options」在 cluster 架設中，我們習慣將預設的「ro,root_squash」改為「rw,no_root_squash」，見下圖。

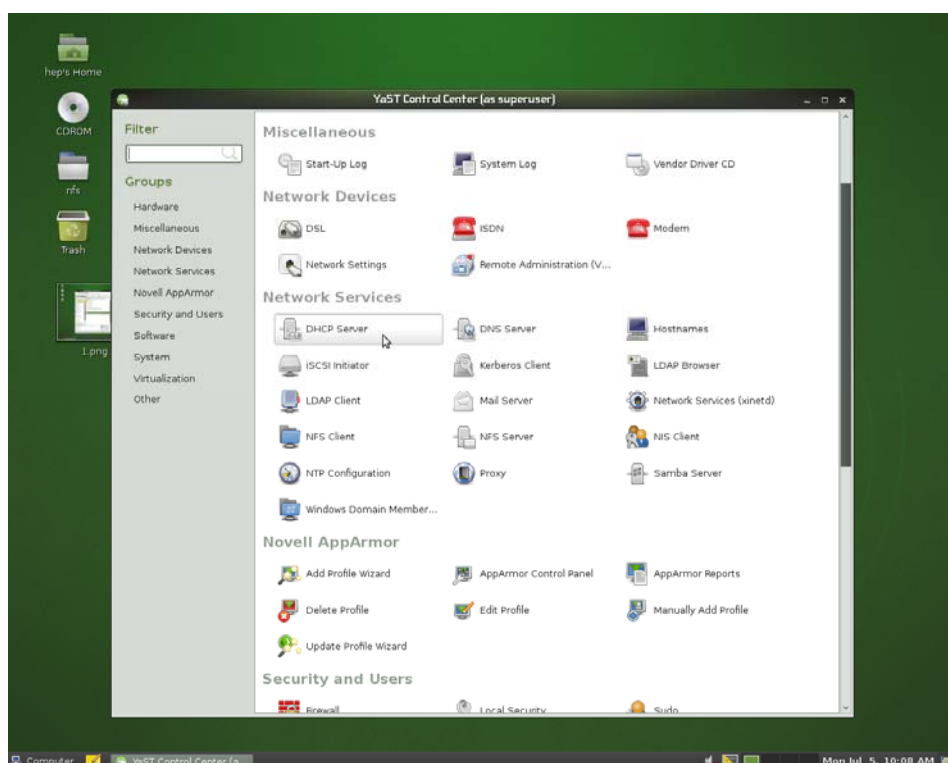


(2)設定DHCP Server

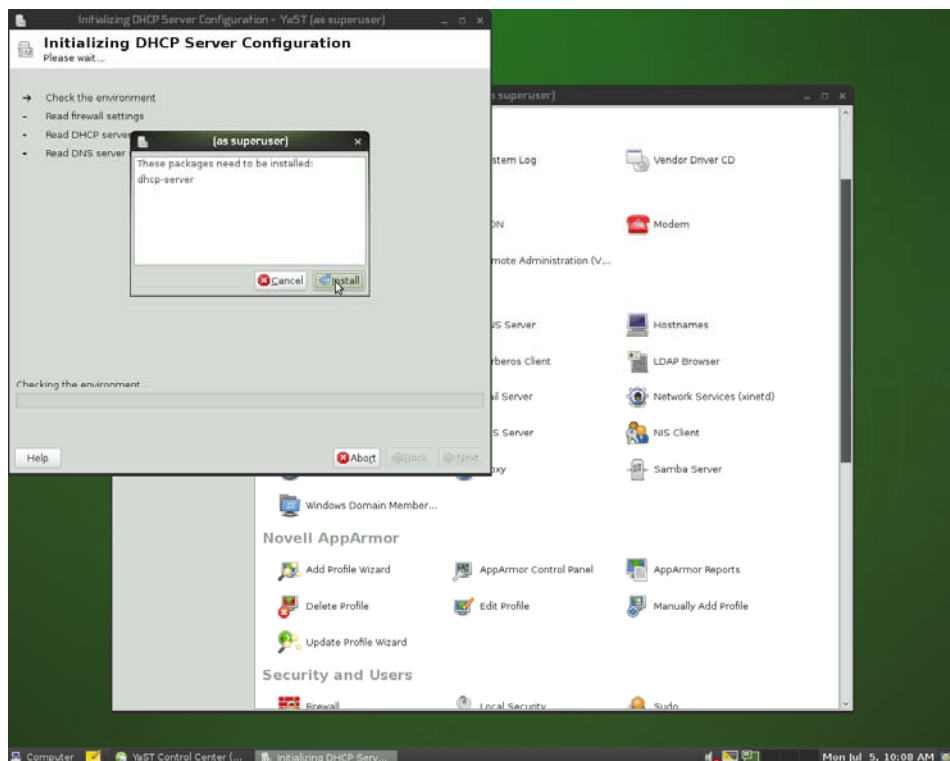
<1>利用「Install Software」搜尋「dhcp」，安裝「yast2-dhcp-server」。



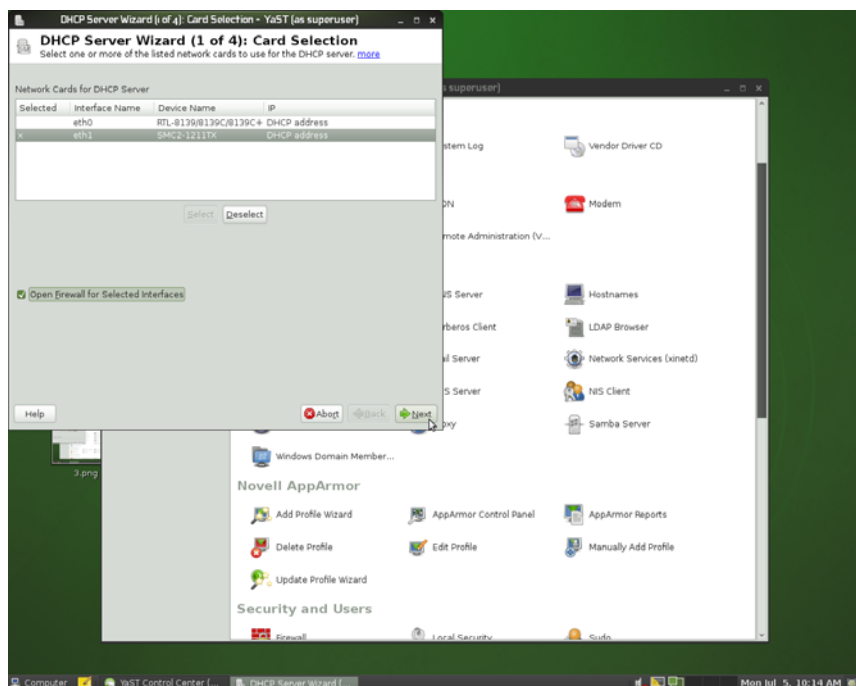
<2>利用「YaST」設定 DHCP Server，見下圖。



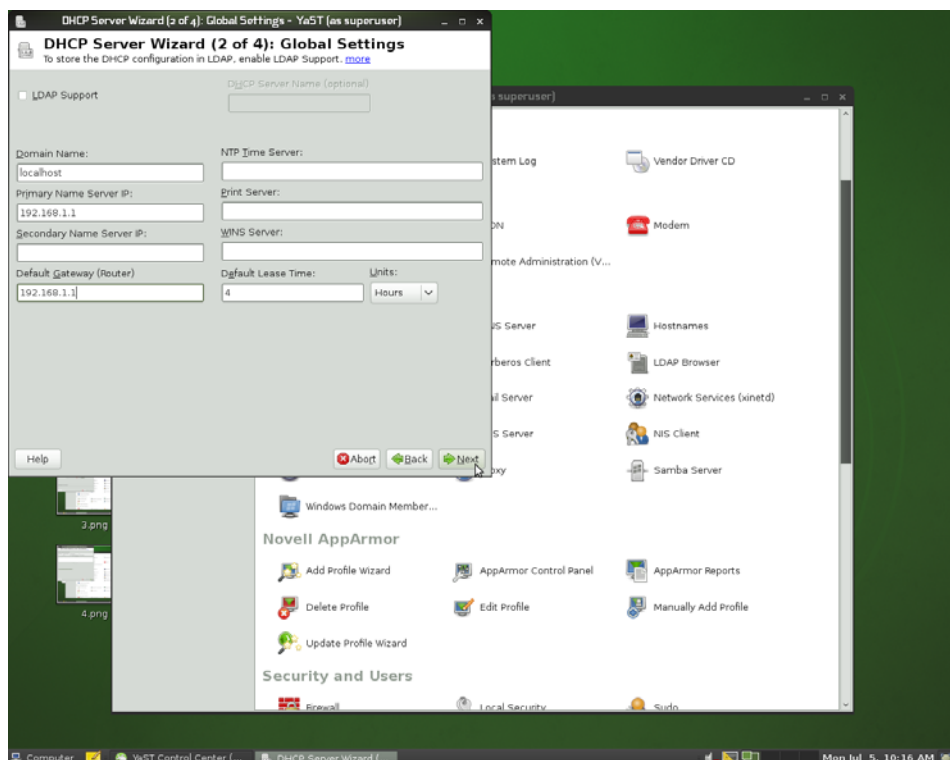
<3>在設定時，若遇到套件不足，openSUSE 會相當人性化的告知，並詢問您是否需要安裝，選擇「Install」即可。



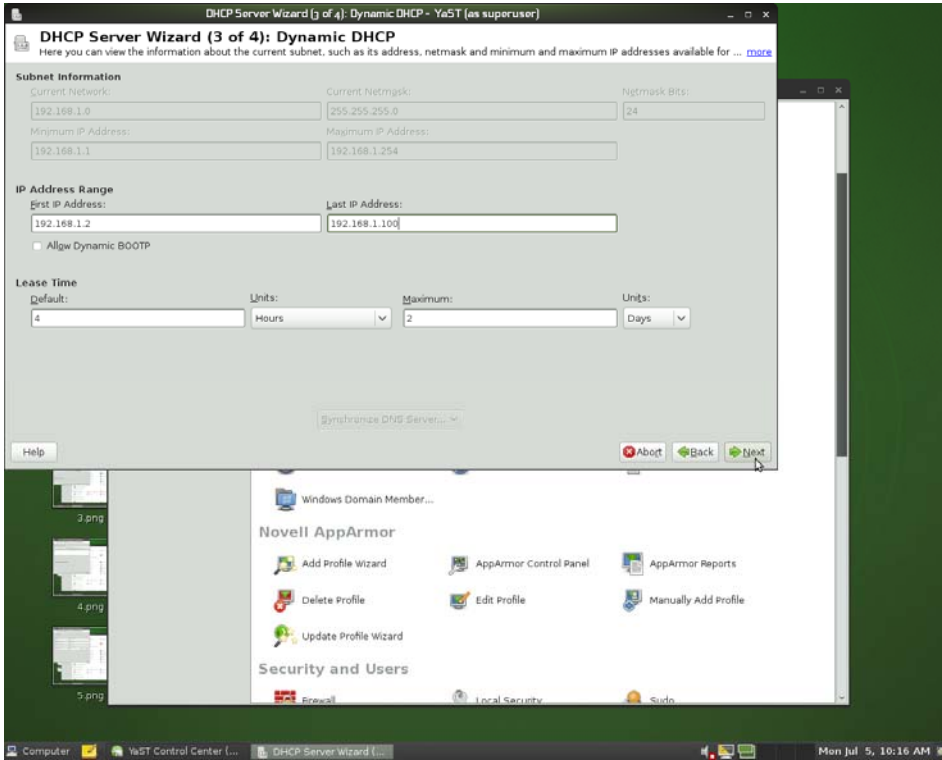
<4>開始進行設定，Master 有兩張網路卡，選擇對內那張設定為 DHCP Server，並開啓此 Server 在防火牆的通路(由於我們在安裝時，已經對內那張網路卡的防火牆關閉，所以其實此舉動無異議)。



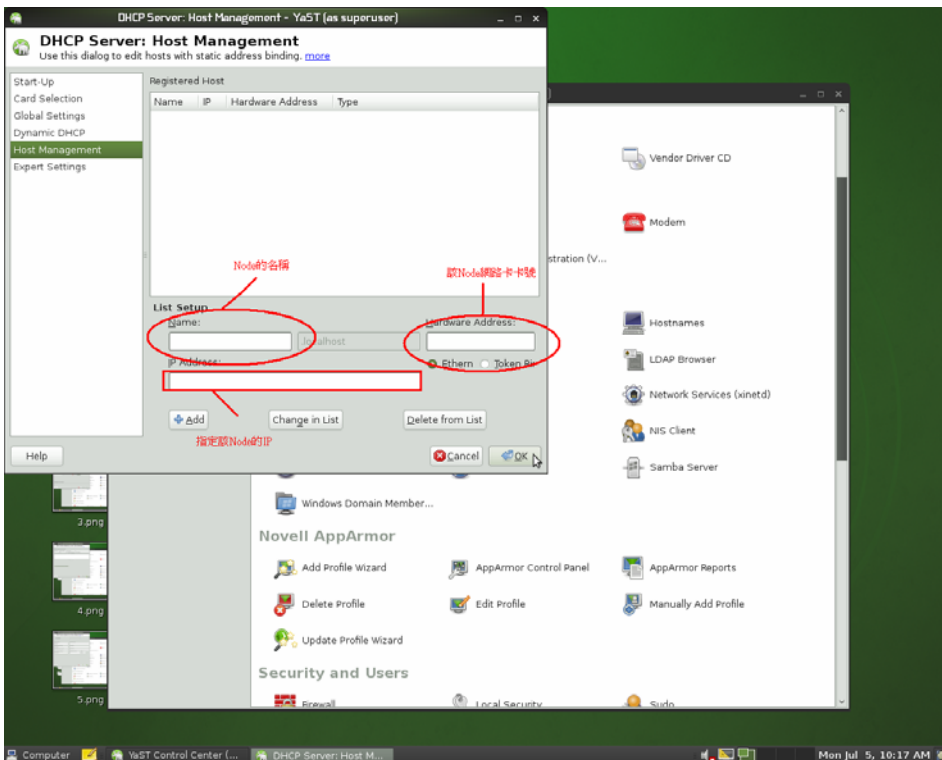
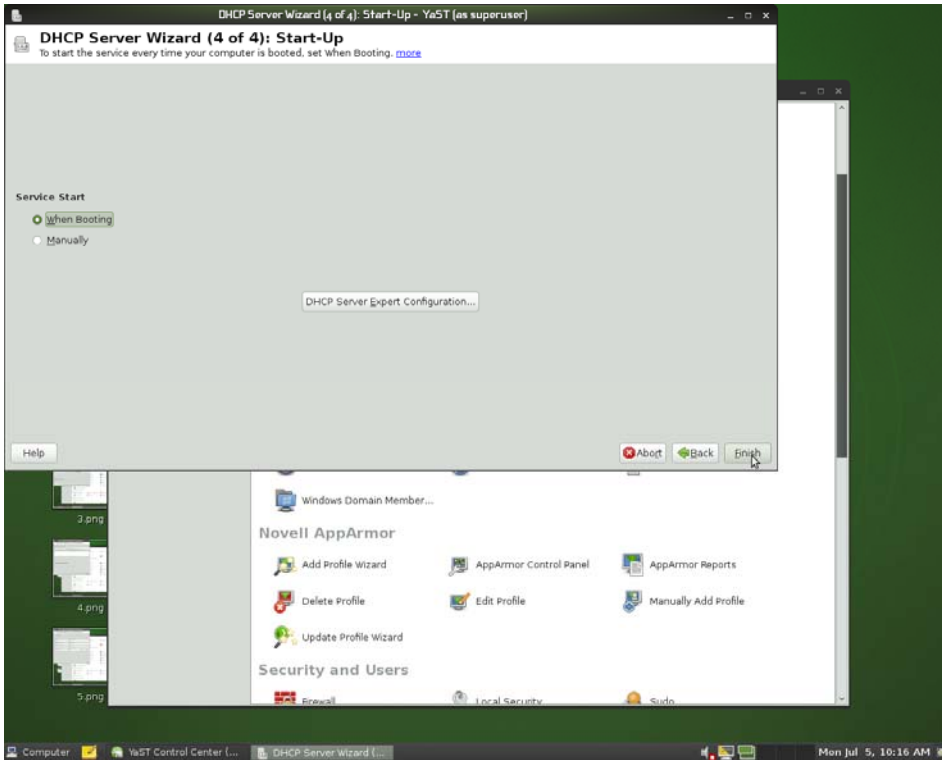
<5>設定好 DHCP Server 的內容，「Domain Name」可自行設定，「Primary Name Server IP」設定為對內網路卡的 IP(我們之前以將其設定為 192.168.1.1)，「Gateway」設定和「primary Name Server IP」相同，詳見下圖。



<6>設定您欲創造的虛擬 IP 範圍，此處我們設定為 192.168.1.2~192.168.1.100，見下圖。

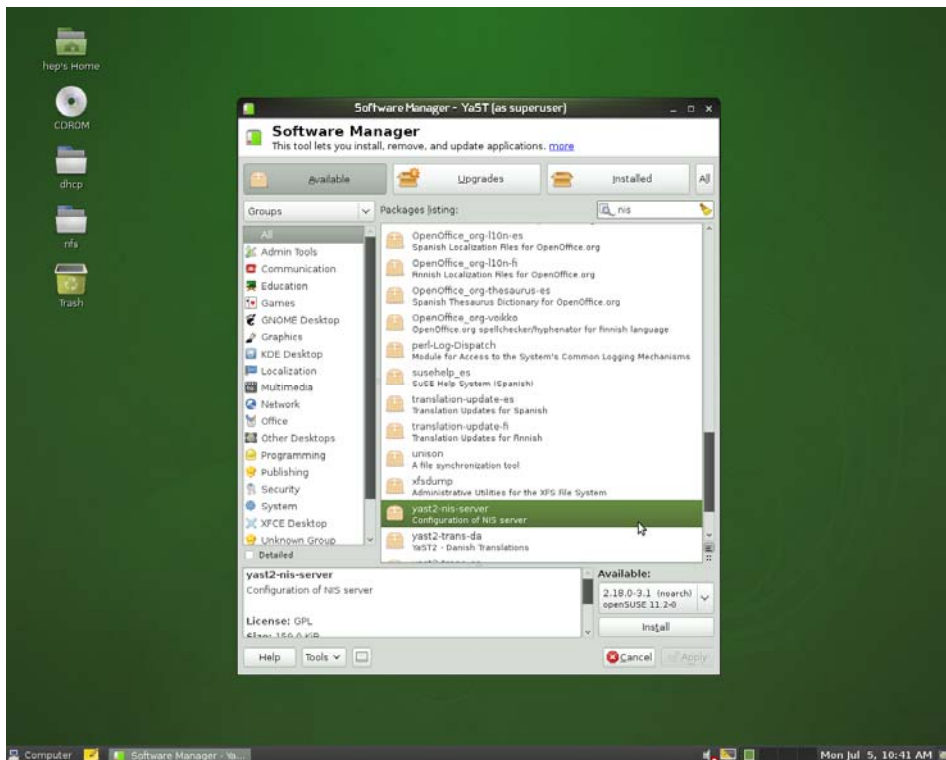


<7>選擇開啓即啓動 DHCP Server，並可利用「DHCP Server Expert Configuration」更準確的設定那台 Node 的 IP 指定爲多少，依據本人經驗，嚴格指定每台 Node 分配到的虛擬 IP，在整個 cluster 會較穩定，因此建議大家執行此步驟。

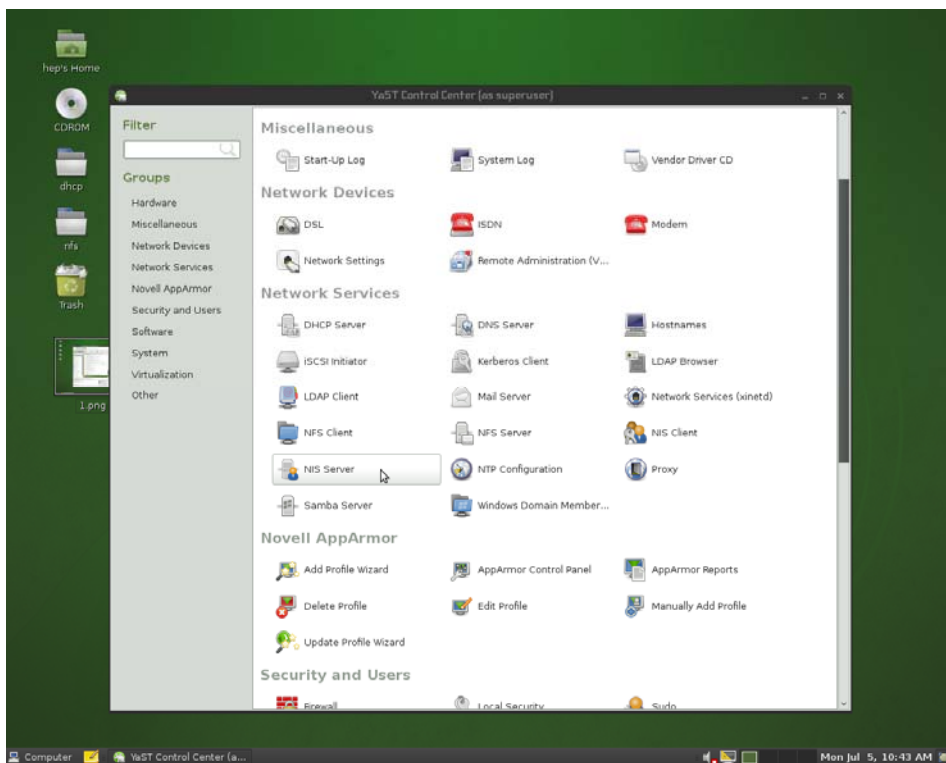


(3)設定NIS Server

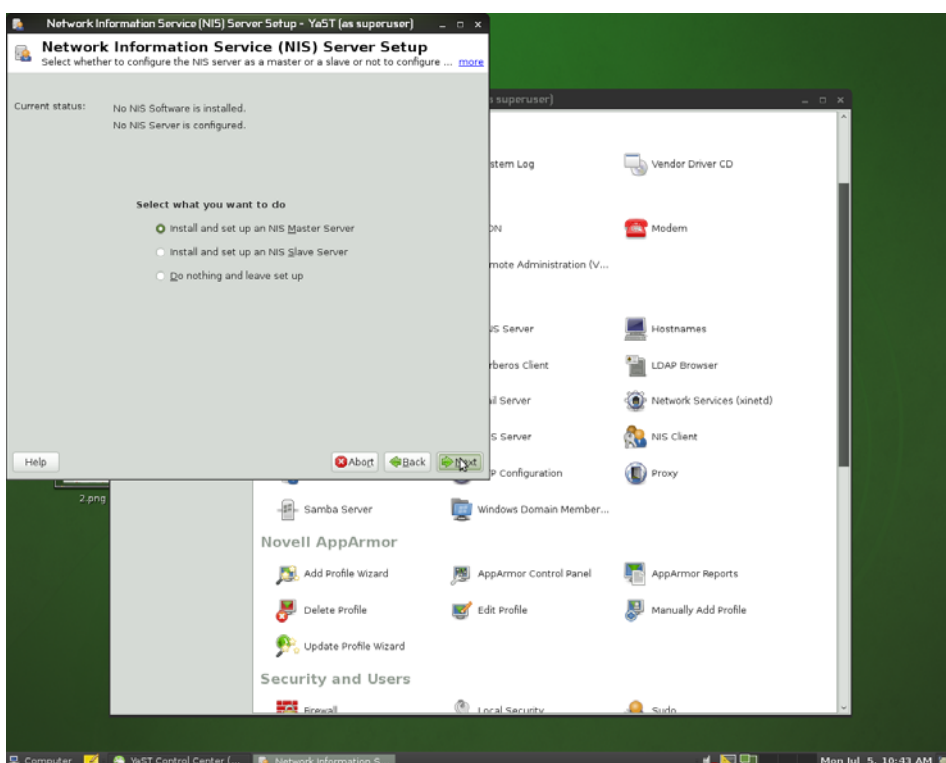
<1>利用「Install Software」搜尋 nis，並安裝「yast2-nis-server」。



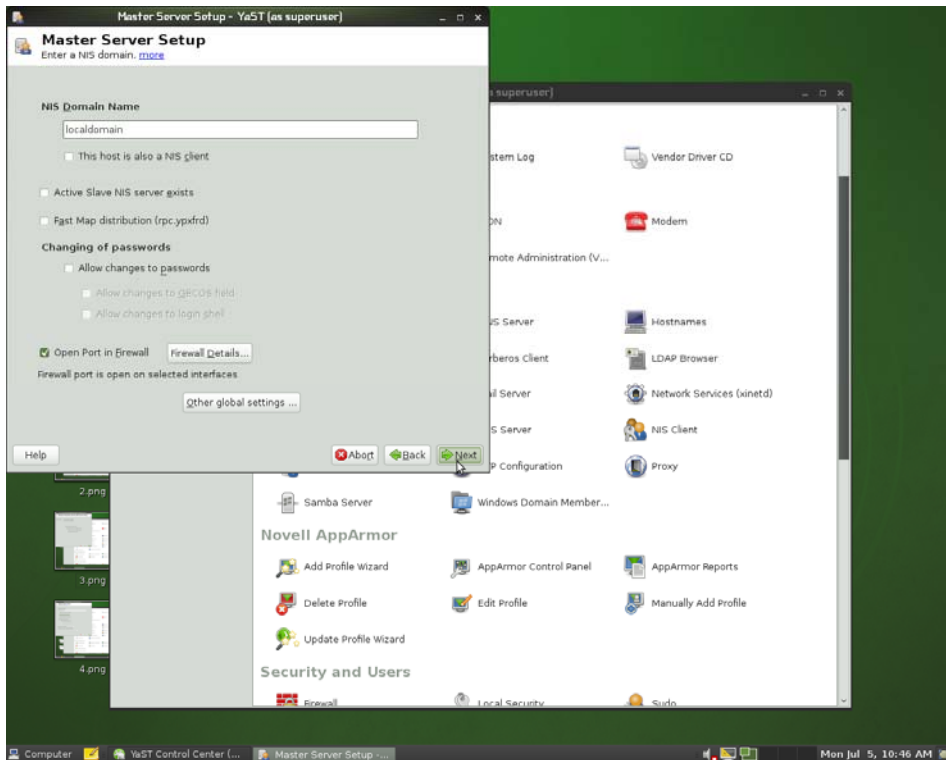
<2>在「YaST2」中，設定「NIS Server」。



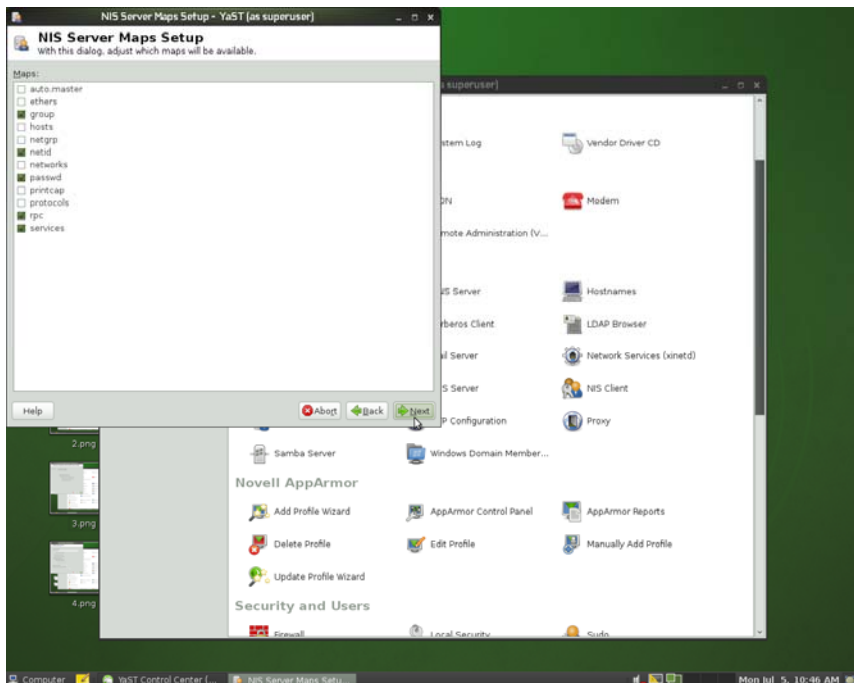
<3>選擇「Install and set up an NIS Master Server」，並將自動開始安裝不足的套件。



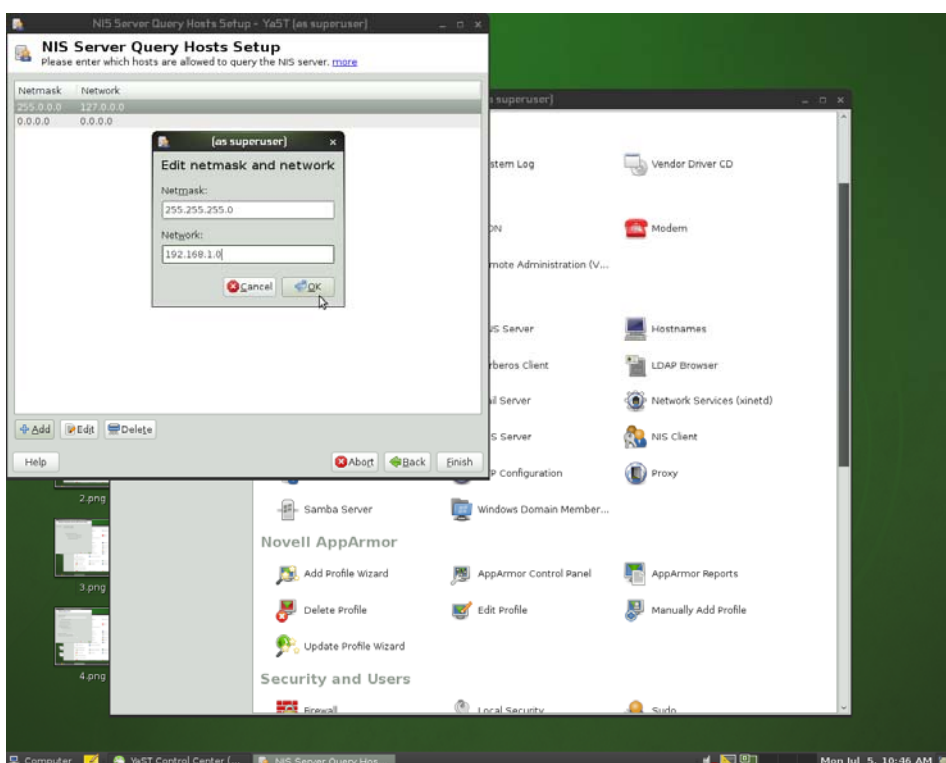
<4> 「NIS Domain Name」可自行設定，詳細如下圖設定。



<5>Maps 使用預設即可，若遇到預設為全部沒有打勾的情況，可以選擇「Abort」，再重頭設定一次，而「Install and set up an NIS Master Server」會變成「Reconfigure NIS Master Server」，繼續上述步驟後，Maps 即會出現如下圖的預設值。



<6>先將預設的兩組資料刪除後，再選擇「Add」，並編輯「Netmask」為 255.255.255.0，「Network」為 192.168.1.0(由於我們的主機對內 IP 為 192.168.1.1，因此我們對內整個網域即為 192.168.1.0)。(PS.圖中兩筆預設記得刪除，否則搭配 MPICH2 時會有問題)



(4) Master 額外設定

<1>由於 SSH Server 在安裝 openSuSE 即設定開啓(openSuSE11.2 預設為關閉, 安裝時請自行打開), 因此此處跳過, 有關 SSH Server 更詳細教學, 可參考「鳥哥的 Linux 私房菜」。

<2>先以 root 登入, 執行 `vi /etc/hosts`, 設定如下圖(由於此處我們 Master 對內虛擬 IP 為 192.168.1.1, 而 node2 對內 IP 為 192.168.1.2)。

```
192.168.1.1    master
192.168.1.2    node2
```

<3>大致上 Master 的設定就結束了!!

5. 設定 Nodes :

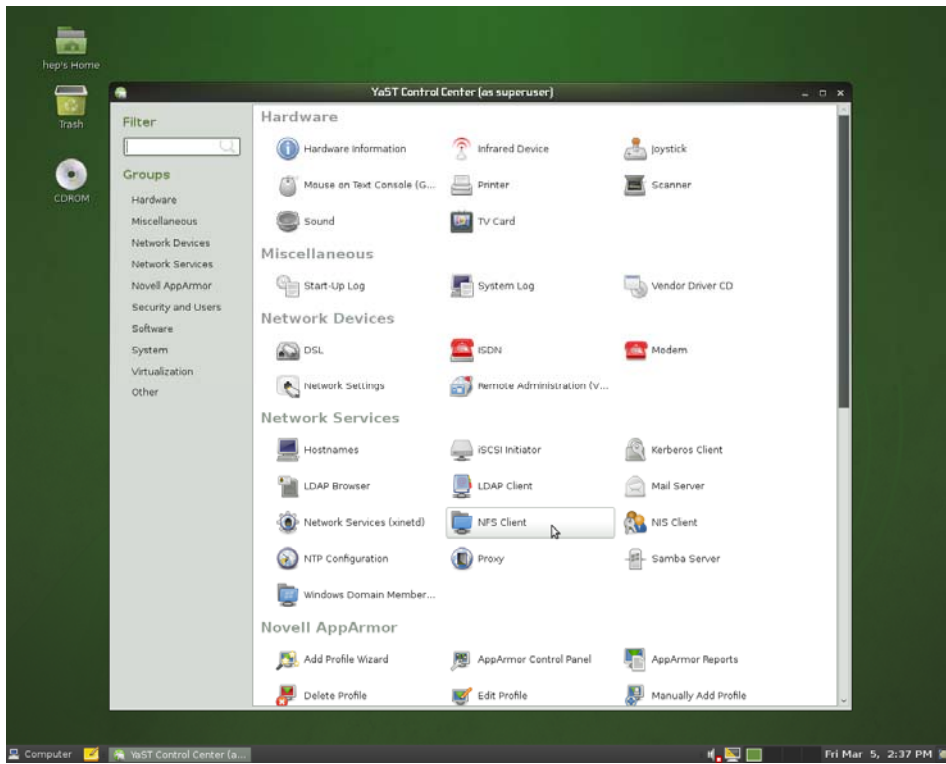
(1) 設定 hosts

以 root 登入, 執行 `vi /etc/hosts`, 設定如下圖。

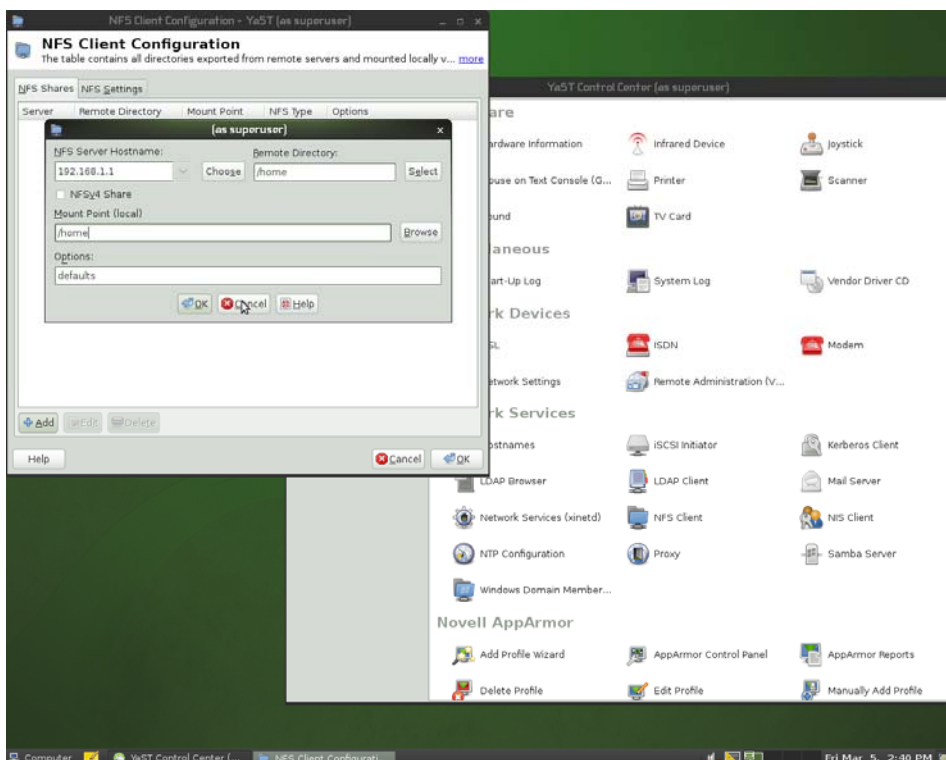
```
192.168.1.1    master
192.168.1.2    node2
```

(2) 設定 NFS Client

<1>利用「YaST2」選擇「NFS Client」。

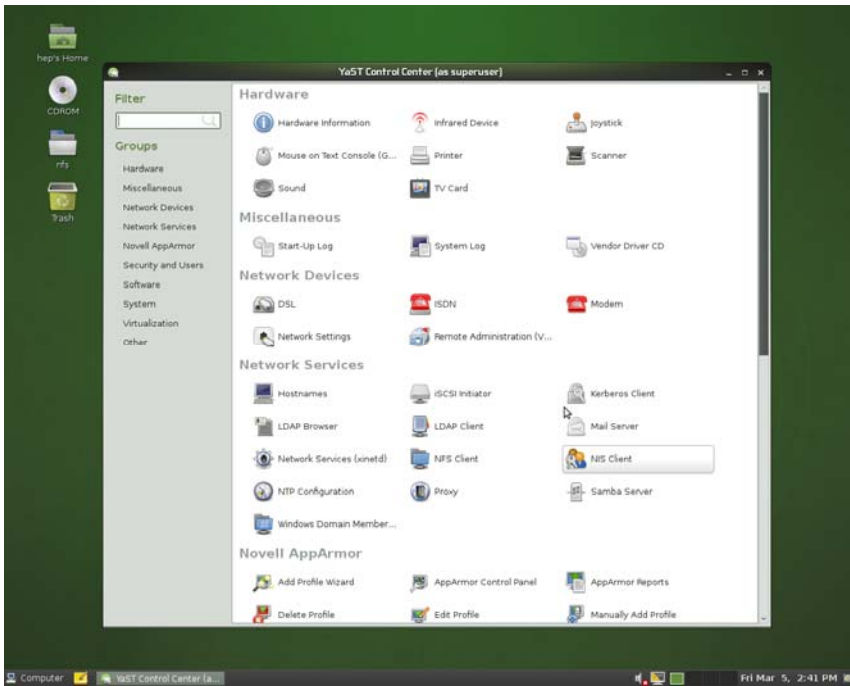


<2>利用「Add」設定, 將 Master 中的硬碟空間載入到本機(此處即為 node2)上, 其中「NFS Server Hostname」設定為 Master 的 IP, 「Remote Directory」為 Master 分享的資料夾名稱, 「Mount Point (local)」為掛載到本機的位置(可跟本機資料夾相同, 會蓋過去)。

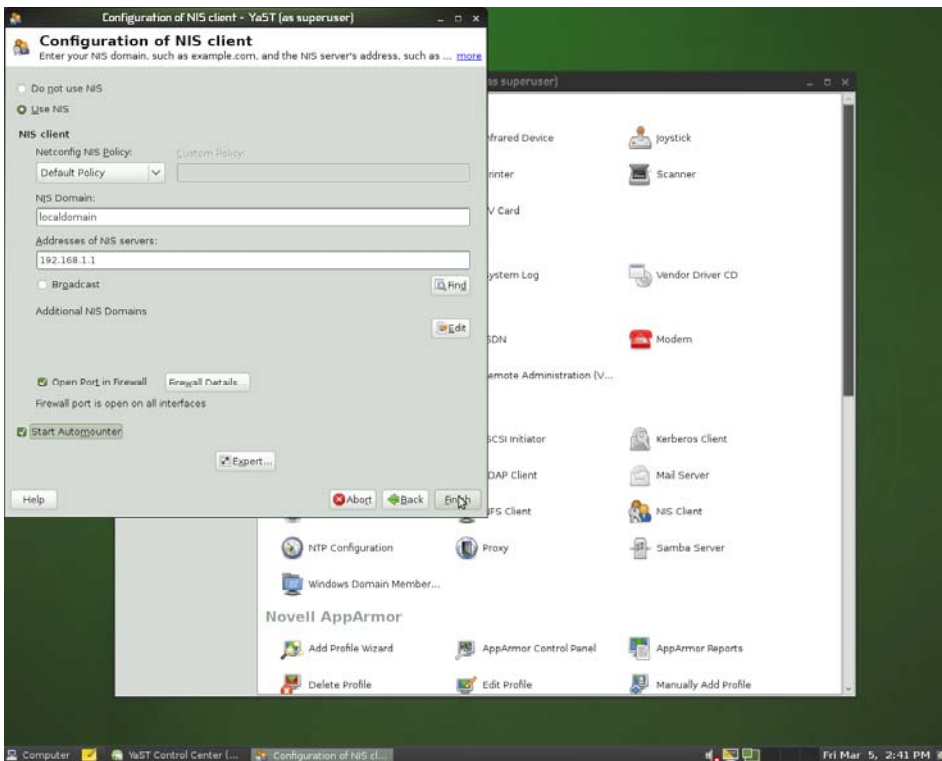


(3)設定NIS Client

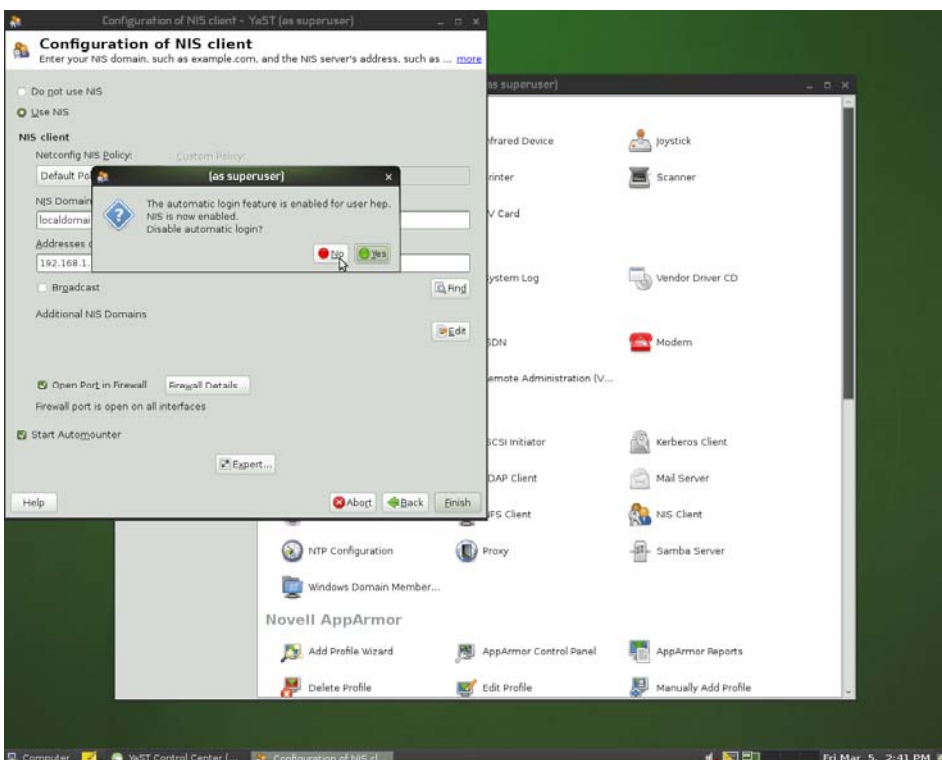
<1>利用「YaST2」設定「NIS Client」。



<2>如下圖設定，記得「NIS Domain」要和當初設定 Master 中 NIS Server 對應，「Addresses of NIS servers」輸入 Master 對內虛擬 IP(此處即為 192.168.1.1)，「Start Automounter」記得打勾。



<3>此時會和開機時自動預設帳號有衝突，可自行選擇是否要取消自動登入預設帳號。



(3)測試Nodes設定是否成功

<1>測試 NFS Server，在終端機下執行 `df`，若出現如下圖訊息，即表示成功。

```
hep@linux-qpma:~  
File Edit View Terminal Help  
Directory: /home/hep/Desktop  
Fri Mar 5 14:44:00 CST 2010  
hep@node2:~/Desktop> cd  
hep@node2:~> df  
Filesystem          1K-blocks      Used Available Use% Mounted on  
/dev/sda2           74848828    3378240 67668408   5% /  
udev                767008        824    766184   1% /dev  
gvfs-fuse-daemon   114034816   3421376 104820800   4% /home/hep/.gvfs  
/dev/sr0            4392280     4392280      0 100% /media/CDROM  
/dev/sdb1          312568640   65556020 247012620  21% /media/Transcend  
192.168.1.1:/home  114034816   3421376 104820800   4% /home  
hep@node2:~>
```

<2>使用 Master，建立一個測試帳號(ex. Sam)，輸入 `ssh node2`，此時尚須輸入密碼(後面教學即將設定成無需密碼)，輸入密碼登入至 node2 的終端機，輸入 `su Sam`，並輸入密碼，若可成功切換使用者帳號，表示 NIS Server 以正常運作。

6. 設定 ssh 免密碼登入：

由於 ssh 除了主機對外的遠端連線外，cluster 內部溝通，亦適用 ssh(有些教學使用 rsh，效果相差不遠)，因此對於內部網域的 ssh 連線，必須設定成不需密碼登入，方可使 MPICH、MPICH2 等平行化資料庫運作。

- (1) 在 Master 上，輸入 `ssh-keygen -t rsa`，接下來的問題皆按 Enter 即可，回到指令列後，輸入下列指令。
- (2) `cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys`(記得我們當初有把 master 的 home 分享並蓋過 node2 的 home，否則要在更複雜)。
- (3) 設定完成，可以在測試利用 ssh 連至 node2，看是否需要密碼，若已經不需使用密碼，表示成功!!
- (4) 每個使用者，皆須執行此設定，該帳號才可以免密碼登入，也就是說有幾個使用者，就要設定幾次。

7. 額外設定：

其實我們的 cluster 已經設定完成，cluster 只是將很多電腦串在一起，若程式的 code 沒有寫入平行的相關語法，其實 cluster 就完全是一台基本的電腦，並不會因為串了很多台 nodes，就跑的特別快，因此我們架設最主要的目的，就是建立一個平行處理的系統平台，再搭配 MPICH2 等資料庫，以達到平行的效果，此步驟的設定，是筆者自己在架設過程中，碰到問題查資料所獲得的經驗，也就是必須把 Master 和 nodes 中的 `/etc/hosts` 設定的很精確，如下圖

Master：

```
127.0.0.1      localhost  
  
140.          master master  
192.168.1.1   master master  
192.168.1.2   node2  
192.168.1.3   node3
```

Nodes：

```
127.0.0.1      localhost  
192.168.1.1   master  
192.168.1.2   node2 node2  
192.168.1.3   node3
```

如此一來，再使用 MPICH2 時才不會出現類似下列的錯誤

```
mpdboot_xxxhost (handle_mpd_output 386): failed to handshake with mpd on slavel; recvd output={}
```

8. 附錄

我們推薦一種分散式的 shell，在 Suse 中，可使用「安裝軟體」來進行安裝，其名稱叫作「pdsh」，在做一些集體的指令，比如說全部 node 都要重開機，即非常方便。

pdsh 使用方法：`pdsh -R ssh -w <你的 nodes 主機名稱>` 指令；下面為 Example

```
pdsh -R ssh -w node2,node3,node4 reboot
```

則你的 node2~node4 都會執行 reboot 這個命令。