

校園豐富教育環境的網路開機方案

馮德歲

樹德科技大學資工系研究生
s95739103@mail.student.stu.edu.tw

洪朝貴

樹德科技大學資工系副教授
ckhung@mail.stu.edu.tw

摘要

本文主要在探討關於跨網域的大型 Linux 無磁碟機系統的運作與實作方法，其中包含系統運作原理與各種服務的架設。之所以採用這樣的架構主要是考量於今日多數的電腦教室並沒有真正的在輔助教育上扮演正確的角色，相較於微軟 Windows 上需要逐項另購的英數理化美音工等教育軟體。諸如 edubuntu 等自由的作業系統，內建即有豐富的教育軟體，極適校園使用。此無磁碟機系統已實際運作於樹德科技大學資訊工程系各實驗室與計算機中心電腦教室內的數百台電腦上。在不久的未來將能夠橫跨整個校園，並為校園內的教學環境提供更好的應用軟體與多元化的選擇服務。

關鍵詞：Linux，PXE，無磁碟機，網路開機

Abstract

We investigate the difficulties and technological solutions in building a large-scale deployment of Linux diskless computers located across several subnets. It is motivated by the poor use of computer resources in an educational setting. Typical campus computer labs running Windows have few educational software packages installed, in stark contrast to the wealth of such packages in labs running the Linux-based edubuntu operating system. This paper is based on the experiences obtained from one such actual deployment of hundreds of computers in the computer science department and computing center of Shu-Te University. It is expected to be extended to cover the entire campus, providing a much richer educational environment for teachers and students.

Keywords: Linux, PXE, Diskless, Network booting

1. 前言

一般(小學至大學)校園電腦環境雖有微軟授權，應用軟體仍僅限於辦公軟體；而英數理化美音工…等，的教育軟體則必須另外逐項採

購。由於軟體採購經費多半已投入作業系統及辦公軟體授權，一般學校在這方面的投資非常薄弱，造成校園電腦比較像是辦公場所的秘書用電腦；諸如幾何互動、光學力學和星象天文，等等校園真正需要的軟體反而甚少見到。

除了教育軟體嚴重缺乏以外，其實管理與個人化的問題更是層出不窮。學生常常需要到電腦教室取得固定的坐位來保留個人的設定。諸如: Internet Message(MSN、Yahoo Message、AIM、ICQ…etc)、佈景主題或 Bookmark 等常見的個人化資料。抑或者是完成到一半的作品、上次撰寫的程式和筆記。當然，一隻 Flash Disk 可以是一個解決的方法，可是卻不是一個好的方法。舉例而言，這包含了較昂貴的價格、容易遺失與易損壞等缺點。

電腦教室的硬體管理方面所造成的問題亦同，每一台電腦都有自己的硬碟，更新軟體或安裝系統都會照成為管理單位的一大困擾。學期中，系統不斷的因為各種問題而照成軟體的損壞，使得空下來的待修主機不斷的增加。

不論是個人化的系統，抑或是軟體的維護的問題，我們可以找到一個共同點，也就是那些運作都是在磁碟機上進行的。有鑑於此，一些國內知名的補習班會在上課時發硬碟讓學員使用。但這卻很難套用於校園內的資訊教學上，可想而知的是那將會帶來更多的問題。

拜現代網路設備突飛猛進的發展速度所賜，從市面上常見的 Gigabits 網路卡到近幾年將會成為主流的 10GE 網路設備。乙太網路的傳輸效能將直逼於目前 IDE 硬碟設備的實際存取速度。這樣的效能造就了上述諸多問題的一個解決方案。也就是說，事實上我們可以建立「一個完全不需要磁碟機的系統」，也就是本文後敘所稱的「無磁碟機系統」(Diskless System)。

Linux 系統的桌面應用環境就這幾年的發展而言，已讓新手不再有距離感。微軟視窗系統上亦有不少宣稱先進的功能來自於某一些早已使用在自由軟體上的專案。諸如，在 IE7 上才出現的 Tabbed browsing 即是 Mozilla 早在 2002 年時就已經在使用的功能。透過無磁碟機

系統的便利性使得使用者在 Linux 與其它作業系統之間的應用轉換之間更加順利。無磁碟機系統相較於一般的電腦設備而言，更具有簡易管理與移動性的優勢。從校園的電腦教學環境乃至於具有高效能網路的跨國企業，無磁碟機都是一個極佳的系統應用方案。

就校園教育環境而言，藉由此 Linux 無磁碟系統的設定與安裝，讓學生使用免費授權的自由軟體將使各科教學的內容更加豐富，而不會僅因為受限於高價的軟體使得學生的學習受到影響。

建構在 PXE 下的 Linux 無磁碟機系統其實早在數年前就廣為應用在叢集運算系統裡。大部份的個人電腦叢集系統需要管裡的節點，小從數十台乃至於大到成千上萬台。當建立這樣的環境時，採用無磁碟機系統將會是他們的選項之一。但無磁碟機系統的大眾化依然存在一些非專業網管難以進入的門檻。這一點由國家高速網路與計算中心所研發的 DRBL(企鵝龍自由軟體教學平台)解決了安裝套件與設定的問題。簡易的 DRBL 足夠應付較小型的教室環境，若將其規模設定在一所大學校園則需要做更深入的瞭解及一些客製化的設定。因為一般的無磁碟機系統的運作通常會盡量採用較單純的網路環境，而在較大的環境或透過 Internet 直接運作無磁碟機系統則是較少見的情況。後者為本方案所運作的環境。

(本文有另一個Web的版本，對於實作部份有提供較完整的詳細資料，該版本的網址為<http://ossug.org/~kevin/DRBL/>。)

2. 運作流程

2.1 基本架構:

早期的無磁碟機系統通常透過在網路卡內嵌各式各樣的 Boot Rom 運作，而最近這幾年則是透過 PXE 做為 Boot Rom。PXE 是 WfM 標準(Wired for Management)的一部分。WfM 是由英特爾 (Intel) 所提議的一套開放產業標準，是一套透過網路自動化管理 client-PC 的系統。WfM 在 1999 年後被 IPMI(智慧型平台管理介面，Intelligent Platform Management Interface)標準所取代。

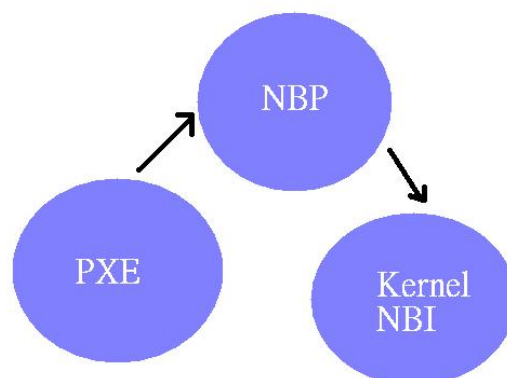
一般市面上所看到 2000 年後的個人電腦，不論主機板支援那一種標準(WfM 或 IPMI)，在 BIOS 裡面應該都包含了可啟用 PXE 系統。

2.2 PXE 遠端開機的流程:

1. DHCP(Dynamic Host Configuration Protocol) Client 發出 DHCPDISCOVER 訊號到 Server 的 Port 67。
2. Server 發出 DHCP OFFER 到 Client 的 Port 68。
3. Client 發出 DHCP REQUEST 向 Server 取得 NBP(Network Bootable Program)的檔名與 Server 位址。
4. Server 將資料透過 DHCPACK 送到 Client 的 Port 68。
5. Client 透過 TFTP 連線至 Server. 並發送 download 的訊號。
6. Client 抓取 NBP Image File 完成後進行初始化並執行。
7. NBP 開機。顯示多重開機選單後，再下載 Linux Kernel NBI(Network Bootable Image)。
8. Linux 開機。

PXE 在由 TFTP 取得 Image 這個動作時，Client 的記憶體是被限制在 640k 以下的，而 Boot Agent 需要 128K。所以當你的 Boot Image 大於 512K 時，在載入上就會發生問題。而 NBP 這個載入器才是用來進行真正的工作，也就是載入 Linux kernel Image。以 DRBL 的系統而言，它運行的流程如圖一所示。

PXELINUX 會透過 PXE 所取得的資訊找到 tftp 的 Server 後取得 Kernel NBI。透過 PXELINUX 載入 Linux Kernel 之後，此時所有的工作就會轉移到 Linux 上。



圖一、各系統開機順序

2.3 Linux 遠端開機的運作:

Linux 會在開機時透過 NFS(Network File System)掛上 root 檔案系統，並由 node 裡的 etc 目錄取得 fstab 檔案後掛載 dev 跟 proc 目錄及其它遠端目錄。

事實上，NFS 或 NIS(Network Information Service)都是屬於 RPC(Remote Procedure Calls)遠端程序呼叫的服務。NFS 在第二版跟第三版時都必須透過 RPC Portmap 代為指定 Port Number，NFSv4 則不需要。所以此時若無特殊設定，請不要使用任何 Deny port 的防火牆設定。

最後啟動 Xwindow，並執行 xdm 登入畫面，讓使用者可以自行透過 NIS 進行認證。

2.4 跨網域的運作:

欲將無磁碟機系統做為伺服器統一服務各網域間的連線，有兩種較簡單的做法可以做到:

1. 使用 DHCP Relay Agent，將服務集中於同一台主機。(DHCP Relay Agent 的服務通常由 Router 提供)
2. 在各 DHCP Server 內指定 Diskless 伺服器位址。

不論是採用哪一種做法，最重要的是使得 Client 取得 DHCP Option 當中所定義的 next-server，也就是主機的位址。(亦即 2.2 PXE 流程第 3.與 4.項)若系統的使用量非常大的話，也可以由此分別指出多台的 next-server 來分散負載。next-server 的主要用途是在於指出 TFTP 下載 Boot Image 的主機位址。若 DHCP 主機與 TFTP 主機在同一台的話，則內定 next-server 為 DHCP 主機的位址(同做法一)。

3. 實務設定

為了讓複雜的過程更加平易近人，我們將系統架設在 DRBL 之上。在實作的過程中仍需要進行大量的修改設定檔，好處是系統可以運用 DRBL 本身所產生的 Image 檔與各式各樣的 Script 用來維護系統的運作。DRBL 的運作與安裝請參考其官方網站，網址如下:

<http://drbl.nchc.org.tw/>。

本文關於設定的部份將會針對較大型的跨網域運作時所需要修改的設定進行介紹。在此設定的過程中所提到的部份，其實也能應用到支援其它無磁碟機環境的系統上，如 GEEXBOX 等多媒體影音撥放系統上。表一為 DRBL 與本系統之差異性。

表 1 DRBL 修改一覽表

	需修改	不需修改
NBI Image		●
DHCP	●	
TFTP		●
NFS	●	
NIS	●	
Script	●	
PXELinux	●	

3.1 DHCP

DHCP 服務採用的是 ISC(Internet Systems Consortium)所提供的 DHCP3 伺服器程式，並且在其設定檔內啟用了多項設定:

1. not authoritative 功能。not authoritative 的作用主要是使 DHCP 不發出 DHCPNAK 訊號，如果需求的架構上有超過一台 DHCP Server 或 DHCP Relay 的 Subnet 跟主機會處在同一個廣播區域內，可能需要考慮此設定。
2. Filename 用來讓 PXE 透過 tftp 載入 PXELinux 的 Image 檔案名稱。
3. 由於 SIS900 的網路卡所載入的 PXELinux Image 需要另外製作一個非標準的版本，所以我們需要利用 DHCP3 的 substring 字串函數截取 PXE 送出來的 vendor-class-identifier，此變數會記錄網路卡的 MAC Address。若是 SIS900 網路卡，他登記的 MAC Address 前三碼有可能為 00:07:95 或 00:0C:6E。
4. next-server 用來設定 TFTP 主機的位址。

3.2 TFTP

在本系統中採用的 tftp 為 tftpd-hpa，DRBL 系統使用的 Option 為-s 做為 chroot 之用。若沒有加上此功能，則需要在 DHCP 的 filename 裡設定完整的目錄與檔名。通常 TFTP 會出現問題主要是發生在設定以外的地方，如 UDP Port

69 在中間被防火牆等設備擋住的話，TFTP 將不能正常的運作。

3.3 NFS

需要修改 etc 目錄內的 exports 設定檔。原始的 DRBL 系統會針對所有的 Client 需要掛載的目錄建立單一的設定，其中每一個 Client 會用到 6 行以上的設定。由於本系統要建立在跨網域的情形之下，所以勢必需要修剪為較簡易的設定。其中由於 /usr、/opt 與 /tftpboot/node_root 屬於限制存取的目錄，可以考慮一次開放。而 /home、/var/spool/mail 與 /tftpboot/nodes 目錄則需要做較精確的切割。舉例而言，一次開放 192.168.0.0 下的 Netmask 255.255.0.0 網段，皆可使用 /opt 目錄的設定方法如下：

```
/opt 192.168.0.0/16(ro, sync, async,  
no_root_squash, subtree_check)
```

如此一來即可大量減少設定檔的內容，若要新增其它網域的主機也只需要簡單的設定並新增即可。

DRBL 的 node 在進行 mount nfs 時，有針對 NFSv3 在 fstab 裡進行最佳化的參數設定，Option 為 "rsize=8192, wsize=8192, tcp" 但這並不是 NFSv4 的最佳化參數。這個參數被寫在 /opt/drbl/sbin/目錄下的 gen_client_files.sh 檔案約 133 行的地方，適情況可能需要適當的調整它。若採用 NFSv4，其最佳化參數為 "rsize=32768, wsize=32768, proto=tcp"。

3.4 NIS

NIS 的資料是儲存在一種稱為 DBM 的資料庫內。也就是說，它不是即時的與系統同步。需要透過 makedbm 來建立 NIS，有部分的系統會在 adduser 或 deluser... 等等的指令設定當中自動執行此動作。關於設定的部份需要注意的是 ypserv.securenets，這個設定檔主要是用來做為 NIS 的連線來源限制，設定檔的內容只分成兩個部份。分別是 Subnet mask IP 與 Network IP。在跨網域設定時需要注意網域內的所有主機都必需被加入。

NIS 在 Client 運作時有兩種設定方式，分別是透過 IP 位址或 DNS(Domain Name System) 來定義 NIS Server 的位址，DRBL 系統選用的是後者的設定。然而此種方式運作時需要在每個 node 系統內的 /etc/hosts 或 DNS 裡建立主機

的資料。若要改成讓 node 透過 IP 來取得 NIS 服務，可以先到 /etc/yp.conf 將內容改為：

```
ypserver <主機的 IP 位址>
```

接著可以透過 DRBL 提供的程式將這個檔案覆蓋到所有 node 的 /etc 目錄下：

```
/opt/drbl/sbin/drbl-host-cp /etc/yp.conf \  
/etc
```

完成後所有的 node 即會改用 IP 與 Server 建立 NIS 的連線。

3.5 PXELinux

PXELinux 的設定檔位於 pxelinux.cfg 目錄內，DRBL 預設是 Linux 為預設的開機系統。但為了一般的校園內應用的需要，可將 label local 的相關設定內加入 MENU DEFAULT，即可改用本地端磁碟為預設開機。若使用者未選擇，經過 timeout 時間(系統單位 1 為 0.1 秒，預設值為 70)後即會自動載入 DEFAULT 所在的 label。開機畫面與背景圖案皆可修改。背景圖的規格必需為 640 x 480，8-bit/color RGBA，並且目前只支援 PNG 與 JPEG 格式，其參數為 MENU BACKGROUND。

3.6 重建與新增 node 目錄

透過 DRBL 提供的工具 gen_client_files.sh 可重新建立 /tftpboot/nodes/ 目錄內的所有 node。其中需要注意更改的 Option 分別為 host、nfsserver 與 nisserver 的位址。舉例而言若要新增或修改 192.168.100.1 這個 node 的資料，可以利用下面這個指令完成：

```
./gen_client_files.sh - host 192.168.100.1 \  
--nfsserver <主機的 IP 位址> --label drbl101 \  
--nisserver <主機的 IP 位址>
```

在重建或新增大量的 node 時，通常是自行撰寫 Script 大量批次執行即可。

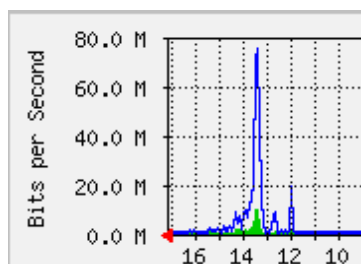
3.7 實務運作分析

本系統目前已實作於樹德科技大學資訊工程系與計算機中心數間電腦教室內，並隨時供應數百台的個人電腦與 Notebook 可在開機時任意選用遠端或 Local 端的磁碟機系統。

此無磁碟機系統的伺服器機型為 IBM blade HS21，採用 Intel Xeon 3.0Ghz 的 CPU、3GB 的記憶體和 72.4GB 容量的硬碟。就目前運作的情形而言，系統瓶頸已經出現在硬碟設

備的容量大小上了。無磁碟機系統在這種架構下所需要使用的容量其實非常的高，畢竟所有的使用者個人的資料都必需儲存在遠端的伺服器內。

圖二為 61 台 Client 同時開機三小時並進行授課所產生的流量表，顯示本系統所使用到的網路頻寬多數集中在開機的瞬間。若 Client 的記憶體足夠，開機後的運作將不至於感受到網路傳輸的延遲。



圖二、電腦教室在使用本系統時網路負載的 MRTG 截圖

4. 結論與未來發展方向

無磁碟機系統不同於早期的大型電腦搭配終端機形式的架構，所有的 Client 都是獨立運作的一台電腦。唯有磁碟機設備是完全使用網路檔案系統。這樣的架構除非是在特殊真正大量的讀取跟寫入可能會照成效能的低落外，一般在使用軟體時幾乎不會感覺任何不便。

雖然只是少了磁碟機，但所帶來的綜效卻遠大於此。學生不用攜帶磁碟片或其它 Storage 出門；不需要在上課前找一台正常的電腦；任何一台電腦登入後都會成為自己專屬的電腦；老師不需要在拿著光碟一一安裝整間教室；不需要擔心病毒的危害；有著自由的教學軟體授權，使學生能真正將英數理化美音工等課程融入於電腦教學的環境當中。

無磁碟機系統的運用到底能夠有多少優點很難將它全部列舉出來。然而在國內，無磁碟機系統僅運用在少數的單位，通常是大學的研究室、高中或國中小的電腦教室。這樣的情況所面臨到的問題就是國內技術文件的缺乏與相關技術人員的數量。

不可否認的，就系統本身與未來的發展方向而言。無磁碟機系統仍存在著許多問題與改善的空間。舉例而言，目前已知的就有系統的安全性問題與針對各種網路系統環境配合的技術發展。這些需要克服的問題將需要更多人的參與，期盼此系統的運作方式能夠帶來更多的參與者。

參考文獻

- [1] Droms, R., "Dynamic Host Configuration Protocol", RFC 2131, 1997.
- [2] Intel, Corp., SYSTEMSOFT, Corp., "Preboot Execution Environment (PXE) Specification", Version 2.1, Intel Corp, 1999.
- [3] IETF, "Network File System Version 4 Protocol", RFC 3530, 2003.
- [4] NCHC 自由軟體實驗室, "DRBL 指令表", <http://drbl.nchc.org.tw/one4all/desktop/drbl-cmd-list.php>, NCHC, 2006.
- [5] Peter Anvin, H., "Can I send information to PXELINUX via the dhcp response?", <http://syslinux.zytor.com/pxe.php#special>, Official Website of PXELINUX.
- [6] Sollins, R. K., "THE TFTP PROTOCOL", REVISION 2, RFC 1350, 1992.
- [7] Sun Microsystems, Inc., "RPC: Remote Procedure Call: Protocol Specification Version 2", RFC 1057, 1988.