NFS-Root Mini-Howto

作者: Andreas Kostyrka, *andreas@ag.or.at*
譯者: Jing-Jong SHYUE, *shyue@sonoma.com.tw*
V8, 8 August 1997

這份 Mini HOWTO 將試著解釋如何設定一個 "無磁碟" 的 Linux 工作站. 且這個工作站的根目錄檔案系統是以 NFS 的方式掛載. 這份文件的最新版可以在 ftp://sunsite.unc.edu/pub/Linux/docs/HOWTO/mini/NFS-Root 或任一個靠近你的 sunsite 映射站台找到.

目錄

1 版權聲明(此段不譯)

(c) 1996 Andreas Kostyrka (e9207884@student.tuwien.ac.at or andreas@ag.or.at)

Unless otherwise stated, Linux HOWTO documents are copyrighted by their respective authors. Linux HOWTO documents may be reproduced and distributed in whole or in part, in any medium physical or electronic, as long as this copyright notice is retained on all copies. Commercial redistribution is allowed and encouraged; however, the author would like to be notified of any such distributions.

All translations, derivative works, or aggregate works incorporating any Linux HOWTO documents must be covered under this copyright notice. That is, you may not produce a derivative work from a HOWTO and impose additional restrictions on its distribution. Exceptions to these rules may be granted under certain conditions; please contact the Linux HOWTO coordinator at the address given below.

In short, we wish to promote dissemination of this information through as many channels as possible. However, we do wish to retain copyright on the HOWTO documents, and would like to be notified of any plans to redistribute the HOWTOs.

If you have questions, please contact Andreas Kostyrka <mailto:andreas@ag.or.at>, the author of this mini-HOWTO, or Tim Bynum, the Linux HOWTO coordinator, at <mailto:linux-howto@sunsite.unc.edu> via email.

1.1 相關支援

- Avery Pennarun <apenwarr@foxnet.net> (如何啓動 Linux 而不使用 **LILO**)
- Ofer Maor <ofer@hadar.co.il> (提供一個較好的 "建立無磁碟工作站" HOWTO.)
- Christian Leutloff <leutloff@sundancer.tng.oche.de> (提供網路啓動的資訊.)

2 一般概論

一般來說, 對這樣的工作站會有下列的問題:

- 系統必須知道他專用的 IP 位置, 並且視需要而重設 Ethernet 的組態.
- 系統必須知道 *NFS* 伺服器的位置, 與根目錄的掛載路徑.

目前核心 (1.3.7x) 對於 *NFSROOT* 的實作提供了下面的解決方案:

- IP 位置可以利用 *RARP* 取得; 完整的 Ethernet 組態則可以用 **LILO** 或 **LOADLIN** 的參數傳遞給核心.
- 要被掛載的 *NFS* 路徑可以利用核心的參數來傳遞. 如果沒有在此指定, 核心會假設 *RAPR* 伺服器同時也是 *NFS* 伺服器, 並且用核心編譯時的預設路徑. (目前的預設路徑是 `/tftpboot/`<機器的 *IP* 位置>.
- 客户端的組態也可以利用 *BOOTP* 取得.

在開始設定無磁碟環境之前, 你必須決定是否要使用 **LILO** 或 **LOADLIN** 來啓動系統. 用這種方式的好處是較有彈性, 缺點則是速度較慢. 不使用 **LILO** 來開機的速度會較快. 你必須自行決定.

3 伺服器的設定

3.1 編譯核心

把 *RARP* 支援編譯到核心中是個好方法. 如果在你不想在啓動時傳遞額外的參數給核心, 你必須將他編譯到核心中. 但如果客户端不在同一個子網路區段中, 這個方式便沒有作用.

工作站的核心則必須至少有下列的支援:

- *NFS* 檔案系統支援. (不需要把 *ext2* 檔案系統的支援編譯到核心中, 以模組方式編譯就夠了.)
- 啓動 "Root on NFS" 的功能.
- 網路卡的 Ethernet 驅動程式必須被編譯在核心中.
- 依據你的需求而將 NFS 根目錄系統下面的 *RARP* 或 *BOOTP* 編譯到核心中. (也就是當 make config 執行時, 在 NFS 選項之後的設定.)

如果在你不想在工作站開機時傳遞參數給核心, 你必須把根目錄設備設定成 0:255. 做法是用 `mknod /dev/nfsroot b 0 255` 建立一個裝置控制檔. 有了這個檔案後, 你便可以把核心內的根目錄路徑以 `rdev <kernel-image> /dev/nfsroot` 指定.

3.2 建立根目錄檔案系統

3.2.1 複製檔案系統

警告: 這些指令可能可以正常作用, 但並不表示可以在其他環境下作用. 對於建立根目錄檔案系統的較佳方式請參考 *Ofer Maor* `<ofer@hadar.co.il>` 所寫的 *NFS-Root-Client mini howto.*

決定了根目錄的放置路徑後, 可以用下面的指令建立 `mkdir -p <路徑>`, 接著執行
`tar cClf / - | tar xpCf <路徑> -`.

如果你不用 LILO 啓動核心, 根目錄路徑必須置於 /tftpboot/<*IP* 位置>. 如果你不喜歡這個位置, 你可以修改核心原始碼中最頂層的 Makefile, 找到類似下面這行並加以修改

```
NFS_ROOT = -DNFS_ROOT="\"/tftpboot/%s\""
```

如果你改了這個, 你必須重新編譯核心.

3.2.2 根目錄系統所需的改變

首先清除不需要的檔案, 並且檢查 /etc/rc.d 的指令稿. 下面是一些重點:

- 首先是 eth0 的設定. 工作站啓動時最少要啓動 eth0, 讓工作站使用伺服器的 IP 位置可不是一個好主意. (作者早期便犯過這個錯誤)
- 另一個重點是 /etc/fstab 的設定. 他應該設成 nfs 檔案系統.
- 警告: 不要把伺服器的根目錄系統與客户端工作站的根目錄系統弄混了. (我曾經改了伺服器上的 rc.inet1, 並懷疑爲何工作站仍然不能運作)

3.2.3 將檔案系統分享出去

要將根目錄系統分享給工作站的指令請參照 `exports(5)`. 在改變設定檔後, 你必須重新起動 nfsd 與 mountd. 在 RedHat 下, 你只要執行 `/etc/rc.d/init.d/nfs stop ; /etc/rc.d/init.d/nfs start` 就可以了.

3.2.4 設定 RARP

在網路上必須有一個 *RARP* 伺服器. 如果啓動時同時沒有設定 nfsroot 參數, 則 *RARP* 伺服器必須設在 *NFS* 伺服器上. 通常都將 *RARP* 伺服器設在 *NFS* 伺服器上. 要完成這個設定, 你的核心必須要有 *RARP* 支援.

並且必須執行下面的指令 (通常放在 `/etc/rc.d` 裡面):

`/sbin/rarp -s` <*ip* 位置> <硬體位置>

其中

ip 位置
: 是工作站的 IP 位置, 而

硬體位置
: 則是工作站網路卡的乙太網路位置.

例如: `/sbin/rarp -s 131.131.90.200 00:00:c0:47:10:12`

你也可以使用符號名稱來取代 IP 位置, 但伺服器必須能找到該名稱所對應的 IP 位置. (設定在 /etc/hosts 中, 或是經由 *DNS* 查詢)

3.2.5 數定 BOOTP

要設定 *BOOTP*, 你必須編輯 `/etc/bootptab`. 請參閱 *bootpd(8)* 以及 *bootptab(5)* 説明頁.

3.2.6 找尋硬體位置

如果我不知道網路卡的硬體位置, 我該如何查詢?

- 利用你所製作的核心磁片開機, 並且找尋網路卡被確認的部分. 通常都會有 6 組 16 進位的位元組 , 這應該就是網路卡的硬體位置.
- 用其他支援 TCP/IP 的作業系統啓動工作站. 然後從伺服器 ping 這個工作站, 接著利用下面的指令查詢伺服器上的 ARP 快取.
  `/sbin/arp -a`

4 啓動工作站

4.1 使用開機唯讀記憶體

由於我沒有用過這個惡劣的方法，我只能提供你下面的方向 (Christian Leutloff <leutloff@sundancer.tng.oche.de> 提供):

- 你不能使用 "一般的" 開機唯讀記憶體.
- Gero Kuhlmann 提供一個 netboot 套件, 製作給 Linux 用的開機唯讀記憶體與相關資訊. 你可以在你當地的 Linux 映射站台找到 netboot 套件; 或是從 Debian 套件中取得 (netboot-0.4).
- 仔細閱讀你的開機唯讀記憶體所附的手冊.
- 你可能必須啓動伺服器上的 tftpd, 但這依據你的開機唯讀記憶體載入核心的方式而定.
- 找尋你所使用的開機唯讀記憶體公司提供的 *Linux* 修補, 但大多數的人沒有可程式化唯讀記憶體燒錄機 *:(*
  *(*特別是我住的歐洲*)*歡迎找我, 我這邊有燒錄器.

4.2 使用原始的核心磁片

如果你的根目錄檔案系統已經用預設的名稱正確的分享出去, 而且你的 *NFS* 伺服器同時也是 *RARP* 伺服器 (這些機器也都在同一個子網路區段中), 你可以直接把核心複製到磁片上, 並以那片磁片開機. (你必須在核心中把根目錄設備設成 0:255) 這個方法假設你的根目錄檔案系統存放於伺服器的 /tftpboot/*IP* 位置 路徑之下. (這個路徑可以經由修改編譯核心時的設定而改變.)

4.3 使用 bootloader 與 *RARP*

在開機時給核心所有必要的參數, 並且增加 nfsroot=<伺服器 *IP* 位置>:<掛載路徑> 這組參數. 其中 伺服器 *IP* 位置 是 NFS 伺服器的 IP 位置, 而 掛載路徑 是根目錄檔案系統存放的路徑.

技巧:

- 當使用 **LILO** 時, 考慮使用 "鎖定" 功能: 只要在第一次輸入完整的參數, 並加上 "lock" 這個參數 , 下一次開機時讓 LILO 逾時.
- 要製作工作站的專用開機磁片時, 你可以在 lilo.cong 中使用 append= 功能.

4.4 使用 bootloader 但不使用 *RARP*

另一種使用 nfsroot 的方法是給核心下面的參數: nfsaddrs=<*wst-IP*>:<*srv-IP*>:<*gw-IP*>:<*netmask*>:<*hostname*> 核心將會依據所給的參數來設定 eth0.

wst-IP

工作站的 IP 位置

srv-IP

NFS 伺服器的 IP 位置

gw-IP

道器的 IP 位置

netmask

子網路遮罩

hostname

工作站的名稱

5 已知的問題

5.1 /sbin/init 無法啓動.

一個常見的問題是 /sbin/init 在某些套件中是動態連結的, 所以你必須提供一個正確的 /lib 給客户端工作站. 一個簡單的方法是把 /sbin/init 換成靜態連結的 "Hello World" 程式, 如此便可以確認問題是否來自動態連結.

5.2 /dev 的問題.

如果你開機時收到一堆關於 tty 的錯誤訊息, 你必須在客户端工作站的 /dev 路徑下執行 "MAKEDEV" 程式. 有些謠言說這個方法不能用在 64 位元的伺服器上. 如果你遇到了, 請告訴我你的伺服器作業系統版本. 一個可能的解決方式是在系統啓動初期建立一個小型的 /dev 記憶體磁碟, 並且重新設定裝置的節點位置.

6 其他主題

- 在 `ftp://sunsite.unc.edu/system/Network/admin/bootpc.v045.tgz` 有 BOOTP 的客户端程式. 這個程式包含了在 Linux 2.0 中的 initrd 程式, 他可以很好的應用在無磁碟工作站上. 對於更特定的設定, initrd 是個很好的選擇.
- 在 Linux 2.0 中包含了 BOOTP 的選項而不是 RARP, 所以可能不需要另外的 bootpd.
- 在核心原始檔的說明文件中, 有一個關於 NFS 根目錄系統的文件.
- 曾經有人寄給我一個修補程式, 讓你可以經由 NFS 來設定置換區, 可是我弄丢了 :(你可能可以在 http://www.linuxhq.com/ 的非官方修補區中找到這個修補檔.
- andreas@ag.or.at 的 PGP 公開金鑰指紋是: F1 F7 43 D5 07 C4 6C 87 BF 6B 33 A2 2C EE 5A F9.