# NFS-Root Mini-HowTo

par Andreas Kostyrka, `andreas@ag.or.at`

Version 8, 8 Août 1997

(Adaptation française par Eric Dumas `dumas@Linux.EU.Org`). Ce mini HowTo présente comment configurer une station **Linux** "sans" disque, qui monte sa racine via NFS. La dernière version de ce mini HowTo peut toujours être trouvée et récupérée sur le site `ftp://sunsite.unc.edu/pub/Linux/docs/HOWTO/mini/NFS-Root`, ou sur n'importe quel miroir près de chez vous.

## 1 Copyright



### 1.1 Contributeurs

- *Avery Pennarun* (`apenwarr@foxnet.net`) : comment amorcer la machine sans LILO ;
- Ofer Maor (`ofer@hadar.co.il`) : a écrit un mini-HowTo bien meilleur pour configurer les stations de travail ne possédant pas de disque ;
- Christian Leutloff (`leutoff@sundancer.tng.oche.de`) : informations sur netboot.

## 2 Présentation générale

En général, on peut rencontrer les problèmes suivants concernant une station de travail :

- elle doit récupérer sa propre adresse IP, et si besoin, également le reste de la configuration Éthernet ;
- elle doit connaître le serveur NFS et le chemin de montage de sa partition racine.

L'implémentation actuelle de *NFSROOT* dans le noyau **Linux** (à partir de la version 1.3.7x) autorise les “solutions” suivantes :

- l'adresse IP peut être trouvée par RARP, ou la configuration Éthernet complète peut être passée au noyau via des paramètres à LILO ou LOADLIN ;
- le chemin NFS pour monter la partition peut être passé au noyau via des paramètres. Si cela n'est pas fait, alors le noyau suppose que le serveur RARP est également le serveur NFS, et utilise le chemin par défaut (le chemin par défaut est codé en dur dans le noyau : `/tftpboot/`*adresse-IP de la machine*.).
- la configuration du client peut être trouvée par BOOTP.

Avant de commencer à configurer un environnement sans disque, vous devez décider si vous allez amorcer la machine en utilisant LILO ou LOADLIN. L'avantage de les utiliser est la souplesse. L'inconvénient est la rapidité. Amorcer un noyau **Linux** sans LILO est plus rapide.

# 3 Configurer le serveur

## 3.1 Compiler les noyaux

Inclure le support RARP dans le noyau du serveur est sûrement une très bonne idée. Vous devez absolument l'inclure si vous allez amorcer sans donner des paramètres au noyau. D'un autre côté, cela ne vous aidera pas vraiment si le client n'est pas sur le même sous réseau que le serveur.

Le noyau de la station de travail doit posséder les éléments suivant au minimum :

- système de fichiers NFS inclu (ce n'est pas la peine de compiler le système de fichiers ext2 : un module suffira) ;
- "Root on NFS" doit être activé ;
- le gestionnaire Éthernet pour la carte réseau de la station doit être inclue dans le noyau ;
- en fonction de vos besoin, il est possible que vous ayez à inclure les protocoles RARP ou BOOTBP pour Nfs-Root (voir les questions posées lors de la configuration du noyau après avoir activé NFS).

Si la station de travail sera amorcée sans aucun paramètre passé au noyau, vous devez également fixer le périphérique de la *racine* à `0:255`. Pour faire cela, il suffit de créer un fichier de périphérique avec :

`mknod /dev/nfsroot b 0 255`.

Après avoir crée un tel fichier de périphérique, vous pouvez fixer le périphérique racine pour l'image du noyau avec :

`rdev`*image-noyau* `/dev/nfsroot`.

## 3.2 Création du système de fichiers racine

### 3.2.1 Copier le système de fichiers

Attention : bien ces instructions peuvent très bien fonctionner chez vous, elles ne sont peut être pas très bien adaptées dans un environnement de production. Consultez le mini-HowTo NFS-Root-Client de Ofer Maor (`ofer@hadar.co.il`) pour une meilleur solution.

Après avoir décidé où placer la racine de l'arborescence, il suffit de la créer avec par exemple :

`mkdir -p` *répertoire*

et

`tar cClf / - | tar xpCf` *répertoire* `-`.

Si votre noyau s'amorce sans LILO, alors la racine doit être `/tftpboot/`*adresse-IP*. Si cela ne vous plait pas, il suffit de le changer dans le fichier Makefile dans les sources du noyau. Recherchez et modifiez la ligne `NFS_ROOT = -DNFS_ROOT`. Si vous modifiez cela, vous devrez alors recompiler le noyau.

#### 3.2.2 Changer la racine du système de fichiers

Maintenant, supprimez les fichiers inutiles et vérifiez les scripts situés dans `/etc/rc.d`. Certains points sont vitaux :

- il est important que le périphérique `eth0` soit configuré. La station de travaille est lancée avec une interface `eth0` au moins configurée partiellement. Donner comme adresse IP à la station l'adresse du serveur n'est pas vraiment une chose vraiment intelligente à faire (comme cela est arrivé une fois à l'auteur lors de ses essais...).
- un autre point important concerne le fichier `/etc/fstab` de la station de travail. Il doit être configuré pour des systèmes de fichiers nfs.
- ATTENTION : ne mélangez pas la racine du système de fichiers situé sur le serveur la racine du système de fichiers de la station de travail (j'ai déjà patché un fichier `rc.inet1` sur le serveur et je me demandais pourquoi la station de travail ne fonctionnait toujours pas.).

#### 3.2.3 Exporter le système de fichiers

Exporter le répertoire racine de la station de travail. Consultez la page de manuel `exports(5)`. Vous devriez également relancer les démons `nfsd` et `mountd` après ces modifications. Avec la RedHat, vous pouvez effectuer très simplement cette opération en lançant `/etc/rc.d/init.d/nfs stop` puis `/etc/rc.d/init.d/nfs start`.

#### 3.2.4 Configuration RARP

Configurer le serveur RARP quelque part sur le réseau. Si vous amorcez sans un paramètre *nfsroot*, le serveur RARP doit également être un serveur NFS. En principe, ce sera le cas. Pour cela, vous devrez utiliser un noyau possédant le support RARP.

Pour réaliser cette opération, lancez (et insérez-le quelque part dans un fichier `/etc/rc.d` du serveur !) :

`/sbin/rarp -s` *adresse-ip adresse-matériel*

où

- *adresse-ip* : est l'adresse IP de la station de travail ;
- *adresse-matériel* : est l'adresse Éthernet de la carte réseau de la station de travail.

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

Vous pouvez également utiliser un nom symbolique à la place de l'adresse IP, du moment où le serveur est capable de trouver l'adresse IP (fichier `/etc/hosts` ou résolution par le DNS).

#### 3.2.5 Configuration de BOOTP

Pour configurer BOOTP, vous devrez éditer le fichier `/etc/bootptab`. Consultez les pages de manuel *bootpd(8)* et *bootptab(5)*.

#### 3.2.6 Trouver les adresses matérielles

Je ne connais pas l'adresse de la carte ! Comment la trouver ?

- amorcez avec la disquette de boot, et regardez la ligne où votre carte réseau est identifiée. Elle contient normalement six octets en hexadécimal, qui devraient normalement correspondre à l'adresse de la carte.
- amorcez la station avec un système d'exploitation qui utilise TCP/IP. Ensuite, lancez un `ping` depuis le serveur sur la station. Regardez enfin le cache ARP en exécutant `/sbin/arp -a`.

# 4 Amorcer la station de travail

## 4.1 Utiliser une ROM bootable

Comme je ne l'ai pas utilisé par moi-même, je ne peut donc vous donner que les conseils suivants (merci à Christian Leutloff, `leutloff@sundancer.tng.oche.de`).

- utiliser des bootroms "normale" ;
- utiliser le paquetage `netboot` écrit par Gero Kuhlmann, et qui est disponibles pour Linux, avec des informations supplémentaires. netboot est récupérable sur les miroirs Linux, ou dans le paquetage de la Debian (`netboot-0.4`).
- lire attentivement la documentation fournie avec votre bootrom ;
- vous devrez probablement à activer le démon `tftpd` sur votre serveur, mais cela dépend un peu de la manière dont votre bootrom charge le noyau ;
- toute information concernant les vendeurs de bootrom pour Linux sont les bienvenues, étant donné que tout le monde n'a pas forcément accès à un grossiste en prom (tout spécialement en Europe, puisque c'est là où je suis)

## 4.2 Utiliser simplement une disquette

Si vous avez exporté la racine du système de fichier avec un nom correcte et que votre serveur NFS est également le serveur RARP (ce qui implique que les deux machines soient sur le même sous-réseau), alors il suffit tout simplement d'amorcer la machine en utilisant un noyau qui aura été copié sur la disquette (par exemple avec `cat`). Vous devez fixer le périphérique racine dans le noyau à 0:255. Cela suppose au préalable que le répertoire racine sur le serveur soit `/tftpboot/IP-Address` (cette valeur peut être modifée en recompilant le noyay).

## 4.3 Utiliser un *bootloader* et RARP

Donnez au noyau tous les paramètres nécessaire lorsque vous souhaitez amorcer la machine, et ajoutez :

`nfsroot=`*adresse-ip-serveur*`:`*/chemin d'accès*

où *adresse-ip-serveur* est l'adresse IP du serveur NFS et */chemin d'accès* est le chemin d'accès au répertoire racine.

Quelques astuces :

- lorsque vous utilisez LILO, pensez à utiliser la caractéristique "lock". Tapez une seule fois correctement tous les paramètres et ajoutez "lock". La prochaine fois que vous amorcerez la machine, LILO lancera un timeout directement et amorcera la machine sans plus attendre.
- lorsque vous générez une disquette d'amorçage spécifique à une station, vous pouvez également utiliser l'option `append=` dans le fichier `lilo.conf`.

### 4.4 Utiliser un *bootloader* sans RARP

En plus de `nfsroot`, il est nécessaire de donner en argument au noyau :

`nfsaddrs=`*wst-IP*:*srv-IP*: *gw-IP*:*netm-IP*:*nommachine*

Le noyau va alors configurer `eth0` avec les paramètres donnés :

- wst-IP : adresse IP de la machine ;
- srv-IP : adresse IP du serveur NFS ;
- gw-IP : adresse IP de la passerelle ;
- netm-IP : masque réseau ;
- nommachine : nom de la machine.

## 5 Problèmes connus

### 5.1 /sbin/init ne se lance pas

Un problème fréquent avec `/sbin/init` est que certaines distributions récentes sont fournies avec une version du programme `init` dynamiquement lié. Donc, vous devez fournir une configuration correcte concernant le répertoire `/lib` au client. Une solution assez simple consiste à remplacer /sbin/init (pour le client) par un programme statiquement lié "Hello World". De cette manière, vous pouvez déterminer si c'est bien la cause du problème, ou bien un problème plus grave.

### 5.2 Problèmes avec le répertoire /dev

Lors de l'amorçage de la machine, si vous obtenez tout un tas de messages d'erreurs concernant les ttys, vous devriez alors lancer un `MAKEDEV` sur le client dans le répertoire `/dev`. Certaines rumeurs font part que cela ne fonctionne pas avec certains serveurs qui utilisent des numéros de périphériques codés sur 64 bits. Contactez-moi si vous avez ce genre de problème. Une solution possible consiste à créer un petit disque mapé en méoire (ram disc) contenant le répertoire `/dev` et de réinstaller les i-noeuds des périphériques à chaque fois.

## 6 Pour plus de renseignements...

- il existe un client `BOOTP` : *ftp://sunsite.unc.edu/system/Network/admin/bootpc.v045.tgz*.

  Avec `initrd` (inclus dans Linux 2.0), cela devrait fonctionner assez simplement pour les stations sans disque. Ce démon est en fait un choix judicieux surtout si vous avez besoin d'une configuration particulière.

- Pour les amorçages basés sur bootpd, cela n'est pas vraiment nécessaire puisque Linux 2.0 contient également l'option d'utiliser BOOTP au lieu de RARP. Pour être plus précis, vous pouvez inclure les deux options dans le noyau lors de sa configuration, et la réponse la plus rapide sera celle choisie.
- dans le répertoire `Documentation` des sources du noyau, vous pourrez y trouver un fichier documentant les systèmes NFS-Root ;
- il existe un patch quelque part qui permet de swapper *via* NFS. On me l'a envoyé (durant une période très surchargée), mais j'ai comme qui dirait perdu le mail...

Vous le trouverez probablement sur le site `http://www.linuxhq.com/`< dans la partie “patches non-officiels”.

Ma clef PGP publique peut être consultée en effectuant un finger à l'adresse `andreas@ag.or.at`. Il s'agit de :

```
F1 F7 43 D5 07 C4 6C 87  BF 6B 33 A2 2C EE 5A F9
```