

# Installation de serveur Web

Thomas LEDUC

Thomas.Leduc@masi.ibp.fr

<http://quartz.dgs.jussieu.fr:8080/>

Mars 1997

## Contents

|          |                                                             |          |
|----------|-------------------------------------------------------------|----------|
| <b>1</b> | <b>Les serveurs Web</b>                                     | <b>2</b> |
| 1.1      | Introduction - raison d'être . . . . .                      | 2        |
| 1.2      | Petite mise au point . . . . .                              | 2        |
| 1.3      | Choix du serveur - critères - comparatif . . . . .          | 2        |
| 1.4      | Présentation du produit . . . . .                           | 3        |
| 1.5      | Variables principales - explications . . . . .              | 3        |
| 1.6      | Redirection automatique de pages . . . . .                  | 3        |
| 1.7      | Les alias . . . . .                                         | 4        |
| 1.8      | Les messages d'erreur . . . . .                             | 4        |
| <b>2</b> | <b>Les procédures CGI</b>                                   | <b>4</b> |
| <b>3</b> | <b>Et en pratique...</b>                                    | <b>5</b> |
| 3.1      | Les fichiers de configuration du NCSA HTTPd 1.5.1 . . . . . | 6        |
| 3.1.1    | Le fichier httpd.conf . . . . .                             | 6        |
| 3.1.2    | Le fichier srm.conf . . . . .                               | 6        |
| 3.1.3    | Le fichier access.conf . . . . .                            | 6        |

# 1 Les serveurs Web

## 1.1 Introduction - raison d'être

Un serveur HTTP est un logiciel qui, lorsqu'il est lancé sur une machine reliée à Internet, permet à tout utilisateur d'Internet de se connecter sur cette machine avec un client HTTP tel que Netscape Navigator, NCSA Mosaic, Microsoft Internet Explorer... et d'obtenir des informations, en particulier des documents HTML.

On parle de serveur HTTP parce que le rôle d'un tel logiciel est de répondre à des requêtes et d'envoyer des pages d'information en utilisant le protocole HTTP (HyperText Transfer Protocol). Néanmoins, ces serveurs font souvent beaucoup plus que cela... à tel point qu'on pourrait se demander si l'appellation "serveur HTTP" n'est pas un peu réductrice. On parlera également de "Serveur WWW" ou de "HTTPd", le 'd' signifiant ici daemon, terme utilisé dans la terminologie UNIX pour désigner un programme qui attend en permanence des événements (ici des requêtes HTTP) pour entreprendre des actions (ici envoyer des documents par HTTP).

Les bénéfices que vous tirerez de la mise en oeuvre de votre propre serveur HTTP sont multiples : économie des frais d'hébergement et de maintenance, contrôle direct du site (pas d'intermédiaire), des statistiques de connexion et des scripts CGI...

Ainsi, les possibilités offertes à l'administrateur d'un serveur WWW sont bien plus larges que celles de la personne qui se contente d'écrire quelques pages HTML et de les placer sur un serveur géré par quelqu'un d'autre. Aussi, si vous en avez le temps et les moyens, vous devriez probablement commencer à envisager la possibilité d'administrer votre propre serveur...

## 1.2 Petite mise au point

Attention, dans ce chapitre, nous "délaissions" le point de vue "client du service Web de l'Internet" pour celui (plus créatif) de "serveur Web". Les documents que vous allez mettre en ligne par le biais de votre serveur pourront alors théoriquement être accédés depuis n'importe quel client Web, indépendamment de sa plate-forme cliente. Par contre, vous ne pourrez raisonnablement pas faire abstraction de votre type de plate-forme si vous adoptez un point de vue serveur, ne serait-ce que parce que tous les serveurs ne fonctionnent pas forcément sur tous les types de plate-forme.

## 1.3 Choix du serveur - critères - comparatif

A ce jour, vous avez le choix entre plus de 50 serveurs HTTP dont un certain nombre sont gratuits. Plusieurs critères matériels, logiciels et financiers sont à prendre en compte lors du choix :

- disponibilité du serveur pour votre type de plate-forme, votre type de système d'exploitation et vos numéros de versions,
- niveau et qualité du support technique
- prix
- enregistrement des "Logs"
- richesse du protocole
  - support des inclusions faites par le serveur (SSI)
  - accès aux variables d'états du serveur depuis les CGI
- sécurité

- restrictions ou interdictions d'accès par noms de domaines ou numéros IP de machines
- restrictions d'accès à certaines parties de l'arborescence de fichiers
- encryptage et authentification par clef publique
- support de SSL et S-HTTP
- autres fonctionnalités
  - outils de configuration graphiques
  - outils de maintenance graphiques
  - maintenance à distance pendant le fonctionnement du serveur
  - outils de mesures de performances temps-réel
  - implémentation d'autres protocoles (ftp, gopher...)
  - inclusion d'un moteur de recherche
  - gestion des répertoires utilisateurs
  - utilisation en serveur proxy avec ou sans cache
  - couplage avec une base de données...

Comparatif des serveurs du marché disponibles à l'adresse suivante : <http://webcompare.iworld.com/>

## 1.4 Présentation du produit

Répertoires : htdocs (répertoire principal), cgi-bin (programmes CGI), logs (fichiers d'enregistrements d'erreurs et d'accès), icons (répertoires d'images), conf (fichiers de configuration), src (fichiers des sources, contient parfois l'exécutable httpd lui-même), support

## 1.5 Variables principales - explications

|                |                                                                                                                                                    |
|----------------|----------------------------------------------------------------------------------------------------------------------------------------------------|
| ServerName     | nom FQDN ou adresse IP de la machine serveur                                                                                                       |
| Port           | numéro du port (80 par défaut) sur lequel le serveur est à l'écoute de requêtes de clients potentiels                                              |
| ServerRoot     | emplacement du répertoire principal du serveur                                                                                                     |
| DocumentRoot   | emplacement du répertoire principal de document du serveur                                                                                         |
| ServerAdmin    | adresse électronique du Webmaster                                                                                                                  |
| User,Group     | nom d'utilisateur sous lequel tournera le démon httpd (ce nom d'utilisateur et de groupe permet aussi d'associer des droits d'accès pour le démon) |
| ErrorLog       | emplacement des fichiers d'erreurs de Log                                                                                                          |
| TransferLog    | emplacement des fichiers d'accès                                                                                                                   |
| UserDir        | répertoire par défaut pour chaque compte utilisateur                                                                                               |
| DirectoryIndex | fichier d'index par défaut                                                                                                                         |
| Alias          | raccourcis mnémoniques de noms de répertoires du serveur                                                                                           |
| ScriptAlias    | raccourcis mnémoniques de noms de répertoires de scripts du serveur                                                                                |

## 1.6 Redirection automatique de pages

A ne pas confondre avec les alias...

Dans le fichier srm.conf, vous pouvez placer des directives de redirection de fichier (après déplacement de votre page sur un autre site Web, par exemple) vers des sites distants ou non.

Par exemple, sur le serveur Web de la machine hector, vous pouvez placer dans le fichier srm.conf la ligne suivante :

Redirect /quartz/ http://quartz.dgs.jussieu.fr:8080/TMP/

## 1.7 Les alias

A ne pas confondre avec les redirections, ce sont des chaînes de caractères mnémoniques qui permettent d'abréger des chemins un peu longs à retenir. On peut en souvent en déclarer une "infinité" dans le fichier srm.conf. Exemple :

Alias /icons/ /home/stock/leduc/httpd/icons/

permet de déclarer un alias vers le répertoire des images, et :

ScriptAlias /cgi-bin/ /home/stock/leduc/httpd/cgi-bin/

permet de déclarer un alias vers le répertoire des scripts CGI.

## 1.8 Les messages d'erreur

Les messages d'erreur sont normalisés sur les serveurs http dans la mesure o ils sont numérotés. Il est important de traduire en français les messages d'erreur car cela peut être fait facilement sur tous les serveurs. Le message d'erreur est envoyé aux navigateurs sous forme d'un code HTML, que vous pouvez modifier à loisir.

Regardons ces messages d'erreur :

| Numéro de message | Signification                                                                                                     |
|-------------------|-------------------------------------------------------------------------------------------------------------------|
| 301               | Le document a été déplacé de façon permanente                                                                     |
| 302               | Le document a été déplacé de façon temporaire                                                                     |
| 304               | Le document n'a pas été modifié, il est donc possible d'utiliser la copie du cache Netscape ou d'un proxy serveur |
| 400               | L'adresse du document contient une erreur de syntaxe                                                              |
| 401               | Vous n'êtes pas autorisé à accéder au document                                                                    |
| 402               | L'accès au document est soumis à un paiement                                                                      |
| 403               | Vous êtes interdit d'accès sur le serveur                                                                         |
| 404               | L'URL demandée est valide mais n'est pas sur le serveur                                                           |
| 405               | La méthode de requête du formulaire n'est pas autorisée                                                           |
| 406               | La requête n'est pas acceptée par le serveur                                                                      |
| 407               | Une autorisation proxy est nécessaire                                                                             |
| 408               | Le temps d'attente pour accéder à la page demandée a expiré                                                       |
| 500               | Une erreur interne du serveur est survenue                                                                        |
| 501               | Une requête faite au serveur n'est pas supportée par celui-ci                                                     |
| 502               | Mauvaise passerelle d'accès                                                                                       |
| 503               | Service non disponible                                                                                            |
| 504               | Temps d'accès à la passerelle d'accès expiré                                                                      |

## 2 Les procédures CGI

Ecrire des procédures CGI demande de connaître la programmation, cette activité est plutôt réservée à un public averti.

Le développement peut être en C, en Pascal, en Basic ou en tout autre langage informatique.

Il existe cependant un langage très riche et très adapté à la manipulation des fichiers, c'est le langage PERL.

Le langage PERL regroupe la puissance des langages C, Shell Script, awk, sed et grep. Ceci ne dira rien aux néophytes mais pour les connaisseurs du système UNIX, cela évoque les langages les plus puissants et les plus astucieux

Un cours sera dédié à l'écriture de CGI.

### 3 Et en pratique...

La manipulation d'installation d'un démon HTTP se fait en plusieurs étapes :

- choix du démon et de sa version en fonction de la plate-forme notamment : Apache, Ncsa...
- choix du serveur ftp anonyme à partir duquel les sources (ou les exécutables) vont être rapatriés à l'aide d'un service Archie :  
`http://web.urec.fr/archie.html`  
`http://archie.univ-rennes1.fr/`
- Rappatriement du fichier ".tar.gz", ".tgz", ".tar.Z"... (à partir de ftp.ibp.fr ou directement de `http://www.apache.org/`, `http://hoohoo.ncsa.uiuc.edu/` ou `http://www.ora.com/...`)  
Détarrage et décompression :

```
gzip -dc *.tar.gz | tar xvf -
```

- A partir de là, deux options possibles selon que vous avez rapatrié des sources ou des binaires. Si ce sont des binaires (ou des exécutables), passez directement au point suivant, sinon, passez sous ./src et entrez successivement les commandes Configure et make...en "priant" pour que la compilation s'effectue sans problème !
- Il s'agit alors de configurer votre démon pour votre propre plate-forme. Pour ce faire, rendez-vous dans le sous-répertoire conf et recopiez les fichiers access.conf-dist, httpd.conf-dist et srm.conf-dist au même endroit en les renommant : access.conf, httpd.conf et srm.conf respectivement.
- éditez alors httpd.conf en vue de le modifier. Il va alors vous falloir spécifier un nouveau port d'écoute pour votre démon httpd (si vous n'êtes pas root ce port est alors de numéro supérieur à 1023, il doit forcément valoir moins de 65535), un nom d'utilisateur et de groupe utilisateur, un nom de répertoire d'installation du démon (répertoire à partir duquel il sera possible d'accéder directement aux sous-répertoires contenant les fichiers de configuration, de log, d'erreurs...), le nom de la machine serveur...
- éditez ensuite srm.conf en vue de le modifier. Dans ce fichier, vous allez préciser le répertoire principal à partir duquel vous délivrerez l'information mais vous y préciserez aussi tous les alias (ou raccourcis) permettant de pointer sous d'autres répertoires annexes (icons, cgi-bins...).
- éditez enfin access.conf en vue de le modifier et établissez-y les autorisations d'accès pour chacun de vos répertoires délivrant de l'information.
- Pour lancer le serveur, il ne vous reste plus alors qu'à entrer la commande :

```
src/httpd -f /home/cestac/leduc/httpd/conf/httpd.conf
```

### 3.1 Les fichiers de configuration du NCSA HTTPd 1.5.1

#### 3.1.1 Le fichier httpd.conf

|             |                                                                                                                                                                                      |
|-------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ServerType  | standalone ou inetd                                                                                                                                                                  |
| Port        | 80 si vous êtes super-utilisateur de votre machine serveur et un entier supérieur à 1024 sinon.                                                                                      |
| User        | nom de l'utilisateur sous lequel ce démon est lancé (seul le super-utilisateur de la machine serveur peut faire tourner le serveur HTTP sous un autre nom que le sien).              |
| Group       | groupe de l'utilisateur sous lequel ce démon est lancé (seul le super-utilisateur de la machine serveur peut faire tourner le serveur HTTP sous un autre nom de groupe que le sien). |
| ServerName  | permet de préciser le nom (FQDN) de la machine serveur.                                                                                                                              |
| ServerAdmin | adresse électronique du Webmaster                                                                                                                                                    |
| ServerRoot  | emplacement du répertoire principal du serveur                                                                                                                                       |
| ErrorLog    | nom du fichier des traces d'accès ayant générés une erreur.                                                                                                                          |
| TransferLog | nom du fichier des traces d'accès au serveur.                                                                                                                                        |
| PidFile     | nom du fichier qui contient le numéro de processus Unix qui permet de contrôler le démon.                                                                                            |

#### 3.1.2 Le fichier srm.conf

|                |                                                                                                                                                                                                                         |
|----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| DocumentRoot   | emplacement du répertoire principal de document du serveur                                                                                                                                                              |
| UserDir        | répertoire par défaut pour chaque compte utilisateur                                                                                                                                                                    |
| Redirect       | directives de redirection de fichier (après déplacement de votre page sur un autre site Web, par exemple) vers des sites distants ou non.                                                                               |
| Alias          | A ne pas confondre avec les redirections, ce sont des chaînes de caractères mnémoniques qui permettent d'abrégier des chemins un peu longs à retenir. On peut en souvent en déclarer une "infinité".                    |
| ScriptAlias    | permet de déclarer un alias vers des répertoires de scripts CGI.                                                                                                                                                        |
| DirectoryIndex | nom du fichier qui sera lu par défaut à l'entrée dans un nouveau répertoire.                                                                                                                                            |
| ReadmeName     | spécifie le nom (générique) du fichier qui sera lu par défaut à chaque chargement d'un répertoire dépourvu de fichier répondant au nom spécifié dans la variable DirectoryIndex.                                        |
| HeaderName     | spécifie le nom (générique) du fichier qui sera lu par défaut à chaque chargement d'un répertoire dépourvu de fichier répondant au nom spécifié dans la variable DirectoryIndex. Son contenu sera placé en bas de page. |
| AccessFileName | nom du fichier qui sera lu dans chaque répertoire pour le contrôle des accès.                                                                                                                                           |
| ErrorDocument  | nom du fichier CGI qui sera lancé à chaque erreur spécifiée en premier argument.                                                                                                                                        |

#### 3.1.3 Le fichier access.conf