
Chapitre 2

Les systèmes de fichiers

1. Vision globale de l'organisation des disques

Pour l'utilisateur, le système de fichiers Unix apparaît comme une **arborescence de répertoires** dont la racine se nomme `/`.

Cette arborescence globale est très fréquemment constituée de plusieurs sous-arborescences qui sont rassemblées, lors du démarrage du système, par une opération dite de **montage**.

■ Remarque

*Une sous-arborescence est baptisée **filesystem**. Nous conserverons par la suite ce terme qui est bien ancré dans le vocabulaire des unixiens et plus précis qu'une traduction du style **système de fichiers**.*

Selon les versions, les administrateurs système disposent de deux types d'organisation des disques. Il s'agit, d'une part, des **partitions** et, d'autre part, des **volumes logiques**. Un filesystem correspond concrètement à une structure physique qui occupe l'espace disque correspondant, soit à une partition, soit à un volume logique. Plusieurs types de filesystems sont disponibles selon les versions.

L'organisation physique des disques doit être complètement transparente pour les utilisateurs et les programmes. Sa maîtrise constitue une compétence essentielle de l'administrateur système.

Dans tous les cas, Unix requiert une région disque complémentaire baptisée historiquement **zone de swap** mais nommée plus souvent aujourd'hui **espace de pagination**. Cet espace permet l'implémentation de la mémoire virtuelle. Il est géré directement par le noyau et ne comporte pas, bien entendu, de structure de filesystem.

Signalons enfin que certaines sous-arborescences peuvent parfois se trouver physiquement sur une autre machine. L'opération de montage se fait alors à travers le réseau grâce notamment à l'application **NFS** (*Network FileSystem*), standard de fait TCP/IP pour le partage de fichiers.

■ Remarque

Dans le cadre de cet ouvrage destiné à l'utilisateur du système, nous allons simplement présenter rapidement les concepts associés à l'organisation des disques.

1.1 Organisation classique en partitions

Dès l'installation du système lui-même ou lors d'un ajout ultérieur, un disque est divisé en **partitions** baptisées aussi **sections** ou encore **slices**.

Ces partitions sont des régions **contiguës d'un seul disque** physique. Leur **taille n'est pas modifiable** après la création.

Une structure de filesystem sera générée ensuite sur l'intégralité de la partition.

La taille d'une partition est déterminée définitivement lors de sa création. Un découpage judicieux du disque demande donc, de la part de l'installateur, planification et expérience. Il convient, en effet, d'effectuer des choix cohérents au niveau des tailles des différents filesystems concernés.

En complément de celles destinées à accueillir des filesystems, il convient de prévoir une partition destinée à la gestion de la mémoire virtuelle (*zone de swap* ou *espace de pagination*).

1.1.1 Mode bloc et mode raw bloc

Les entrées/sorties disques peuvent s'effectuer de deux façons : le mode **bloc** et le mode **raw bloc**.

Le mode bloc correspond à l'utilisation normale de la partition par un service de tampons (*buffers*) gérés par le noyau. Cela correspond à des mécanismes d'écritures différées et de lectures anticipées qui conviennent à la quasi-totalité des programmes.

Dans le cadre de certains logiciels ou de certaines commandes, il est néanmoins possible d'effectuer des entrées/sorties directes sans utiliser le système de tampons du noyau.

Pour ce faire, les partitions possèdent deux noms différents en tant que fichier spécial.

■ Remarque

Nous faisons ici une référence en avant à la notion de fichier spécial, détaillée dans la suite de ce chapitre. Un fichier spécial précise le nom d'un périphérique. Le fait de désigner ce fichier spécial dans une commande appropriée active implicitement le programme de gestion (driver) de ce périphérique.

Le nom de la partition en mode bloc correspond donc aux commandes usuelles et à l'utilisation en tant que filesystem.

Le nom de la partition en mode *raw bloc* est disponible pour une utilisation en tant qu'espace disque dédié à un logiciel. Dans ce cas de figure, cette partition ne contient pas de filesystem. Sa gestion est faite sous la responsabilité de ce logiciel qui la considère comme un grand fichier séquentiel. Le cas concret le plus répandu est celui d'un S.G.B.D. qui gèrerait directement l'espace disque correspondant à une base de données. On parle alors communément, dans le jargon des administrateurs, d'une **raw partition**.

Il faut noter enfin que, même si la partition contient un filesystem, certaines commandes administratives, de par leur fonctionnement interne, utilisent aussi l'entrée en mode *raw bloc*.

Remarque

Les versions **Linux** ne conservent pas cette distinction historique entre les deux types de fichiers spéciaux pour une même partition. L'entrée en mode bloc est la seule disponible. La commande administrative **raw** permet cependant d'effectuer, si nécessaire, une association avec une entrée raw bloc.

1.1.2 Visualisation des partitions (Solaris, Linux)

Les versions Unix concernées possèdent une commande spécifique pour visualiser ou modifier l'organisation d'un disque en partitions.

Sous Solaris, la commande interactive **format** (réservée à l'administrateur) est un utilitaire complet de maintenance des disques et comprend évidemment une rubrique dédiée aux partitions.

Sous Linux, la commande **fdisk** est un utilitaire de gestion des partitions et comprend évidemment une possibilité d'affichage.

Exemple

```
$ fdisk -l /dev/sda
Disk /dev/sda: 160.0 GB, 160041885696 bytes
255 heads, 63 sectors/track, 19457 cylinders, total 312581808 sectors
Units = sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk identifier: 0x0007d654

   Device Boot      Start         End      Blocks   Id  System
/dev/sda1    *          2048        499711       248832   83   Linux
/dev/sda2                501758       31258095      156039169    5   Extended
/dev/sda5                501760       31258095      156039168   8e   Linux LVM
$
```

Les noms des partitions Linux sont, la plupart du temps, de la forme **/dev/sdxy** où *x* est une lettre comprise entre *a* et *h* qui désigne le disque et *y* un numéro qui désigne la partition.

Les matériels de type PC sont limités à quatre partitions physiques. Il est possible de créer trois partitions primaires et une partition étendue au sein de laquelle on pourra créer autant de partitions logiques supplémentaires que nécessaire. Les quatre premiers chiffres sont associés aux partitions physiques.

Les partitions numérotées à partir de 5 correspondent aux partitions logiques créées au sein de la partition physique étendue.

1.2 Organisation en volumes logiques

Par rapport aux partitions classiques, les volumes logiques permettent de s'affranchir de certaines limitations.

Dans un premier temps, les disques doivent être regroupés dans des **groupes de volumes**. À l'intérieur d'un groupe de volumes, l'administrateur décide d'une organisation en **volumes logiques**. Il s'agit de régions **non forcément contiguës d'un ou plusieurs disques**. Elles **pourront être agrandies** après la création.

Une structure de filesystem se crée, cette fois, sur un volume logique et le comportement Unix habituel est ensuite complètement reproduit.

La plupart des versions permettent aujourd'hui une organisation en volumes logiques et les administrateurs disposent alors de beaucoup plus de souplesse.

La gestion des disques par volumes logiques est native dans les versions *AIX*. *HP-UX* propose également cette organisation depuis longtemps. *Linux* et *Solaris* disposent aussi d'une implémentation. La mise en œuvre se fait par des commandes administratives spécifiques qui mettent cependant en jeu une terminologie commune ou très voisine.

1.2.1 Terminologie du LVM (Logical Volume Manager)

Donnons quelques définitions résumées en conservant volontairement les termes anglo-saxons :

VG (Volume Group)

Le groupe de volumes est une entité obligatoire. Il s'agit d'un ensemble de disques sur lesquels seront créés des volumes logiques. L'intérêt majeur de la notion de groupe de volumes réside sans doute dans la fonctionnalité d'*export-import*. Cette opération permet en effet de déplacer un ensemble de disques d'une machine vers une autre en conservant tous les filesystems présents sur ces disques.

PV (*Physical Volume*)

Un volume physique correspond à un disque complet. Celui-ci doit être intégré à un groupe de volumes avant de pouvoir être utilisé.

LV (*Logical Volume*)

Un volume logique est un espace disque destiné, la plupart du temps, à contenir un filesystem. Il s'agit bien de l'équivalent d'une partition classique, plus souple en termes de gestion.

PP (*Physical Partition*) ou PE (*Physical Extent*)

Ces deux termes sont synonymes (PP : terminologie AIX, PE : terminologie HP-UX). Il s'agit d'unités d'allocation d'espace disque (8 ou 16 Mo par exemple) utilisées pour constituer des volumes logiques. La taille de ces unités se choisit lors de la création du groupe de volumes.

LP (*Logical Partition*) ou LE (*Logical Extent*)

Ces deux termes sont synonymes (LP : terminologie AIX, LE : terminologie HP-UX). Une partition logique (LP ou LE) correspond par défaut à une partition physique (PP ou PE). Il est cependant possible d'indiquer qu'une partition logique sera associée à deux, voire trois, partitions physiques pour mettre ainsi implicitement en œuvre une fonctionnalité de copies redondantes de l'information (**mirroring**) ou d'écritures en parallèle (**striping**).

■ Remarque

Il est important de noter, dans ces définitions résumées, que le terme de partition prend ici un sens complètement différent de celui des organisations classiques. L'équivalent fonctionnel (amélioré) de la partition classique est bien le volume logique. La confusion pourrait apparaître dans certaines implémentations telles que Linux où une partition classique peut être utilisée comme un PV (disque physique) afin d'être intégrée à un groupe de volumes.