

# Introduction au système de fichiers

Par asma soltani

le 26/10/2017

[asma.soltani@gmail.com](mailto:asma.soltani@gmail.com)

Ce document est sous licence [Creative Commons: Licence d'attribution en partage identique et non commercial 4.0](https://creativecommons.org/licenses/by/4.0/).

## Objectifs généraux:

Ce chapitre, permet à l'apprenant de s'initier aux concepts suivants :

- Exemples des systèmes de gestion de fichiers
- La représentation physique d'un fichier et d'un répertoire
- L'allocation utilisée par un système de fichiers

## Sommaire:

Introduction  
Fichiers  
Répertoires  
Allocation  
Exemples de systèmes de fichiers

## Introduction

Le **système de fichiers** est la portion du système d'exploitation qui s'occupe de la gestion des mémoires de masse. Il prend en charge le stockage des fichiers sur le disque dur, le rangement de ceux-ci dans des répertoires, l'ouverture ou la fermeture de fichiers/répertoires, et bien d'autres choses encore. La gestion des partitions est aussi assez liée au système d'exploitation.

## Fichiers:

Les fichiers sont de différente nature selon ce qu'ils contiennent. La convention historique est que le suffixe de leur nom indique cette nature **mais ce mécanisme n'est toutefois pas 100% robuste**. Exemples :

- Mon\_programme.java
- Mon\_programme.jar
- Mon\_document.doc
- Ma\_video.avi
- Ma\_photo.png

Les programmes informatiques qui transforment les données sont aussi matérialisés ou tout simplement contenus dans des fichiers. Il y a donc deux grandes catégories de fichiers : les fichiers de données « statiques » et les fichiers exécutables ou « de commande » (programmes), qui eux sont « dynamiques ». Les suffixes des programmes sont .bat et .cmd pour les fichiers de commande ou encore .com, .exe, .obj, .lib, .dll, .class, .jar... dans le système Windows.

Pour Windows, on peut obtenir cette information comme suit : `echo %PATHEXT%`.

Dans le système UNIX, ce sont .obj, .lib, .class, .jar... mais un suffixe n'est pas le critère déterminant pour être un « programme ». Sous UNIX, le plus sûr moyen de différencier fichiers programme et fichiers normaux est d'utiliser l'option -F dans la commande ls.

## Répertoires

Les répertoires sont juste des fichiers particuliers. Ils contiennent des répertoires et des fichiers contrairement à des fichiers « terminaux » qui ne contiennent que des données et des programmes. Il y a de nombreux moyens de distinguer un répertoire d'un fichier terminal :

- UNIX : `ls -aF | grep /` (ou aussi `ls -aF | grep '/'`, *i.e.*, on recherche le caractère '/' dans le flux de données produit par `ls -aF`)
- Windows : `dir /ad`

Les répertoires sont créés, détruits et manipulés (navigation) préférablement de manière graphique (*Bureau*). Toutefois les manipulations via la souris ne sont jamais que la mise en oeuvre des commandes suivantes :

- Création, destruction : `mkdir, rmdir`
- Visualisation du contenu d'un répertoire
  - UNIX : `ls -aF`
  - Windows : `dir /a /q`
- Visualisation profonde (ou récursive)
  - UNIX : `ls -R`
  - Windows : `dir /s` (voir aussi `tree`)
- Navigation : `cd`



Le premier système de fichier ayant été utilisé sur un système d'exploitation Microsoft est le système FAT, qui utilise une table d'allocation de fichiers (en anglais FAT, File Allocation Table). La table d'allocation de fichiers est en fait un index qui liste le contenu du disque, afin d'enregistrer l'emplacement des fichiers sur celui-ci (c'est une sorte de sommaire). Etant donné que les blocs qui constituent un fichier ne sont pas toujours stockés de manière contiguë sur le disque (c'est ce que l'on appelle la fragmentation), la table d'allocation permet de conserver la structure du fichier en créant des liens vers les blocs constitutifs du fichier. Le système FAT est un système 16 bits permettant de décrire un fichier par un nom d'une longueur de 8 caractères et une extension qui en comporte 3.

Le système de fichiers FAT est donc un système 16 bits, cela signifie qu'il ne peut pas adresser les clusters (un cluster ou bloc est la plus petite unité de stockage d'un système de fichiers) sur plus de 16 bits. Le nombre maximum de clusters repérables avec le système FAT est ainsi de  $2^{16}$ , soit 65536 clusters. Or, étant donné qu'un cluster est constitué d'un nombre fixé (4,8,16,32, ...) de secteurs de 512 octets contigus, la taille maximale d'une partition FAT se trouve en multipliant le nombre de clusters par la taille d'un cluster. Avec des clusters d'une taille 32Ko, la taille maximale d'une partition FAT est donc de 2Go ( $32 \times 2^{16} = 2097152 \text{Ko} = 2048 \text{Mo} = 2 \text{Go}$ ).

D'autre part, un fichier ne peut occuper qu'un nombre entier de clusters, c'est-à-dire que si un fichier occupe plusieurs clusters, le dernier sera occupé en partie, et la place inoccupée restante est autant de place perdue. Par conséquent plus la taille d'un cluster est réduite, moins il y a de gaspillage de place. On estime qu'un fichier gaspille en moyenne la moitié d'un cluster, cela signifie que sur une partition de 2Go 16Ko seront perdus par fichier...

- clusters de 512 octets : partitions de 2 à 32 Mo ( $512 \times 2^{16} = 32 \text{Mo}$ ) ;
- clusters de 1 Ko : partitions de 4 à 64 Mo ( $1024 \times 2^{16} = 64 \text{Mo}$ ) ;
- clusters de 2 Ko : partitions de 8 à 128 Mo ;
- clusters de 4 Ko : partitions de 16 à 256 Mo ;
- clusters de 8 Ko : partitions de 32 à 512 Mo ;
- clusters de 16 Ko : partitions de 64 à 1 024 Mo, c'est-à-dire 1 Go ;
- clusters de 32 Ko : partitions de 128 à 2 048 Mo, c'est-à-dire 2 G

### La table d'allocation des fichiers

La Table d'Allocation de Fichiers est une liste de valeurs numériques permettant de décrire l'allocation des clusters d'une partition, c'est-à-dire l'état de chaque cluster de la partition dont elle fait partie. La table d'allocation est en fait un tableau dont chaque cellule correspond à un cluster. Chaque cellule contient un chiffre qui permet de savoir si le cluster qu'elle représente est utilisé par un fichier, et, le cas échéant, indique l'emplacement du prochain cluster que le fichier occupe. On obtient donc une chaîne FAT, c'est-à-dire une liste chaînée de références pointant vers les différents clusters successifs, jusqu'au cluster de fin de fichier. Chaque entrée de la FAT a une longueur de 16 ou 32 bits (selon qu'il s'agit d'une FAT16 ou d'une FAT32). Les deux premières entrées permettent de stocker des informations sur la table elle-même, tandis que les entrées suivantes permettent de référencer les clusters. Certaines entrées peuvent contenir des valeurs indiquant un état du cluster spécifique. Ainsi la valeur 0000 indique que le cluster n'est pas utilisé, FFF7 permet de marquer le cluster comme défectueux pour éviter de l'utiliser, et les valeurs comprises entre FFF8 et FFFF spécifient que le cluster contient la fin d'un fichier. Chaque partition contient en réalité deux copies de

la table, stockées de manière contiguë sur le disque, afin de pouvoir la récupérer si jamais la première copie est corrompue.

### **L'évolution FAT32 :**

Utilisant des adresses sur 28 bits (et non pas 32!), il permet de constituer des unités d'allocation de taille réduite sur des disques de taille importante

La capacité des partitions peut s'élever jusqu'à 2 To :

- clusters de 4 ko : partitions de 1 To ( $2^{28} \times 4\text{ko} = 1\text{ To}$ );
- clusters de 8 ko : partitions de 2 To ;

### **L'évolution exFAT (FAT64) :**

Le système de fichiers exFAT (Extended File Allocation Table) est l'évolution du système FAT32 pour les mémoires flash et les supports de stockage externes (disques durs et assimilés) de grandes capacités. Cela permet notamment d'avoir des fichiers dont la taille dépasse 4Go (la limite de FAT32) et des partitions jusqu'à 64 Zio (1 Zio est équivalent à environ 1 milliard de Tio ou 1,1 milliards de To, 1 073 741 824 plus exactement) contre 2 To pour FAT32.

### **La suite : NTFS :**

NTFS (New Technology File System) est un système de fichiers développé par Microsoft Corporation pour sa famille de systèmes d'exploitation Windows NT, à partir de Windows NT 3.1, Windows 2000 et tous leurs successeurs à ce jour (XP, Server 2003, 7, 8, etc.). NTFS est le successeur du système de fichiers FAT comme système de fichiers de prédilection des systèmes d'exploitation Microsoft Windows. Il est sur 64 bits. Le nom des fichiers peut aller jusqu'à 255 caractères. Il rajoute une notion de sécurité

Taille des clusters = 512b, 1kb, 2kb, 4kb, 8kb, ...64kb

Taille max : 16 EiB

L'exbibyte est une unité d'information de stockage informatique.

Il est une contracture de EXa Binary BYte. Son abréviation est EiB. 1 exbibyte =  $2^{60}$  bytes

La limite théorique de la taille d'une partition est de 16 exaoctets (17 milliards de To), mais la limite physique d'un disque est de 2To

