# File concept:

## Khái niệm:

* Là đơn vị lưu trữ logic trên bộ nhớ ngoài. Đối vs người dùng thì file là đơn vị phân bố nhỏ nhất của disk, chỉ có thể ghi vào ổ đĩa thông qua file.
* File là 1 dãy bits, bytes, dòng, bản ghi mang ý nghĩa được định nghĩa bởi người tạo ra.
* Gồm 2 kiểu chính là:
  + Data files: text (ASCII) và binary.
  + Program: source and object files.

Table

Description automatically generated

## Thuộc tính của file:

Text, letter

Description automatically generated

A picture containing text

Description automatically generated

* Thuộc tính của file được lưu trong cấu trúc directory hay các directory entry.

Diagram

Description automatically generated

## Thao tác với file:

* Tạo file (Create)
* Ghi file (Write)
* Đọc file (Read)
* Thay đổi vị trí trong file (Seek)
* Xóa file (Delete)
* Xóa dữ liệu của file (Truncate)

### Note:

Thao tác duyệt tìm directory entry tốn thời gian.

* Để giải quyết thì có bảng mở file: Chứa thông tin của file đang được mở.

Khi không sử dụng file nữa thì OS xóa entry của file tương ứng trong bảng mở file.

# Cấu trúc thư mục file:

Hệ thống file tự động gói và dỡ gói các dữ liệu các dòng bytes thành các block (Có kích thước tương đương vs sector của đĩa, thường là 512 bytes).

1 file thì có thể được lưu trữ trong nhiều block

Do đọc theo từng block -> internal fragmentation (Các bytes cuối của block cuối cùng của file ko để trống).

## Phân vùng:

* Đĩa được chia thành nhiều phân vùng: partitions, minidisks, volumes
* Mỗi phân vùng được xử lý như vùng lưu trữ phân biệt, và có thể chứa 1 OS riêng.

## Các thao tác với thư mục:

* Tìm kiếm file
* Tạo file
* Xóa file
* Liệt kê thư mục
* Đổi tên file
* Duyệt hệ thống file

## Các loại cấu trúc thư mục:

* Thư mục 1 mức: Cấu trúc đơn giản, các file nằm trong cùng 1 thư mục; Nhưng khi số người dùng và số file lớn, khả năng trùng tên file cao.

Chart

Description automatically generated

* Thư mục 2 mức: Mỗi người dùng sử dụng thư mục riêng, và khi làm việc chỉ duyệt thư mục riêng. Giải quyết được vấn đề trùng tên và chỉ hiệu quả khi người dùng độc lập. Khó khăn khi dùng chung file.

Timeline

Description automatically generated

* Thư mục cấu trúc cây: Là cấu trúc thư mục phổ biến nhất. Cấu trúc có thư mục gốc, mọi file trong hệ thống này có đường dẫn độc nhất. Xóa thư mục con dẫn tới xóa hết cây con của nó.
  + Đường dẫn (Path name) có 2 loại: đường dẫn tuyệt đối và đường dẫn tương đối.

Diagram

Description automatically generated

* Thư mục dùng chung: User này có thể link đến file của user khác. Khi duyệt thư mục, 1 file có thể được duyệt nhiều lần.

Diagram

Description automatically generated

# Cài đặt tổ chức file system:

## Phương pháp duyệt file:

* Danh sách tuyến tính với con trỏ tới các khối dữ liệu.
* Bảng băm cùng vs danh sách tuyến tính

## Phương pháp phân phối vùng lưu trữ:

* Phân phối liên tục: file được lưu trữ trong các khối nhớ liên tiếp nhau
  + Tốc độ truy cập nhanh
  + Khó khăn khi muốn tăng kích thước file; Cách lựa chọn vùng nhớ đều dẫn tới phân mảnh ngoài (Các vùng nhớ liên tiếp còn lại ko đủ cho yêu cầu lưu trữ)
* Phân phối liên kết: File được phân phối các ô nhớ không liên tục, cuối mỗi khối là con trỏ để trỏ tới khối tiếp theo.
  + Tốc độ truy nhập chậm, phải định vị từ đầu
  + Liên kết khối có thể bị hỏng dẫn tới mất dữ liệu, hoặc liên kết sai.
* Phân phối chỉ mục: Mỗi file có một khối chỉ mục chính (index block) chứa danh sách các khối dữ liệu của file.
  + Không gây hiện tượng phân mảnh ngoài
  + Cho phép truy nhập trực tiếp
  + Luôn cần ít nhất 2 block cho 1 file dù file rất nhỏ, block dữ liệu và block chỉ mục.

# Tìm hiểu về Ext file system: