UNIVERSITY OF SCIENCE
FACULTY OF INFORMATION TECHNOLOGY

ASSIGNMENT 03.04

SUBJECT: OPERATING SYSTEM

Class: 21CLC03

Students: Trần Nguyên Huân - 21127050

Nguyễn Hoàng Phúc - 21127671

Instructors: Thái Hùng Văn, Đặng Trần Minh Hậu

HO CHI MINH CITY – 2023

Contents

I. Tasks assignment table, completion level

Full name	Student Id	Tasks	Completion level
Trần Nguyên Huân	21127050	- Do research and write report part IV	100%
Nguyễn Hoàng Phúc	21127671	- Do research and write report part II and III	100%

II. Definition of File

- Files are logical units of informatinon created by processes.

- A file is a logical entity that gathers related information together to help retrieve related information to make it easy to access this information without regard to how they are stored on the storage device.

- The most important part of a file is the content it contains. In addition to the content, files often have attributes that make file management easier and more efficient.

III. File Attributes (2)

- Every file has a name and data, and operating systems also include additional information, known as attributes or metadata, which varies across systems.

- Some common attributes include those related to file protection, such as access permissions and passwords, as well as flags that control specific properties like whether the file is hidden or has been backed up recently.

- Record-length, key-position, and key-length fields are only present in files that can be looked up using a key.

- Additionally, various time fields keep track of when the file was created, accessed, and modified, which is useful for purposes like determining if a source file needs to be recompiled.

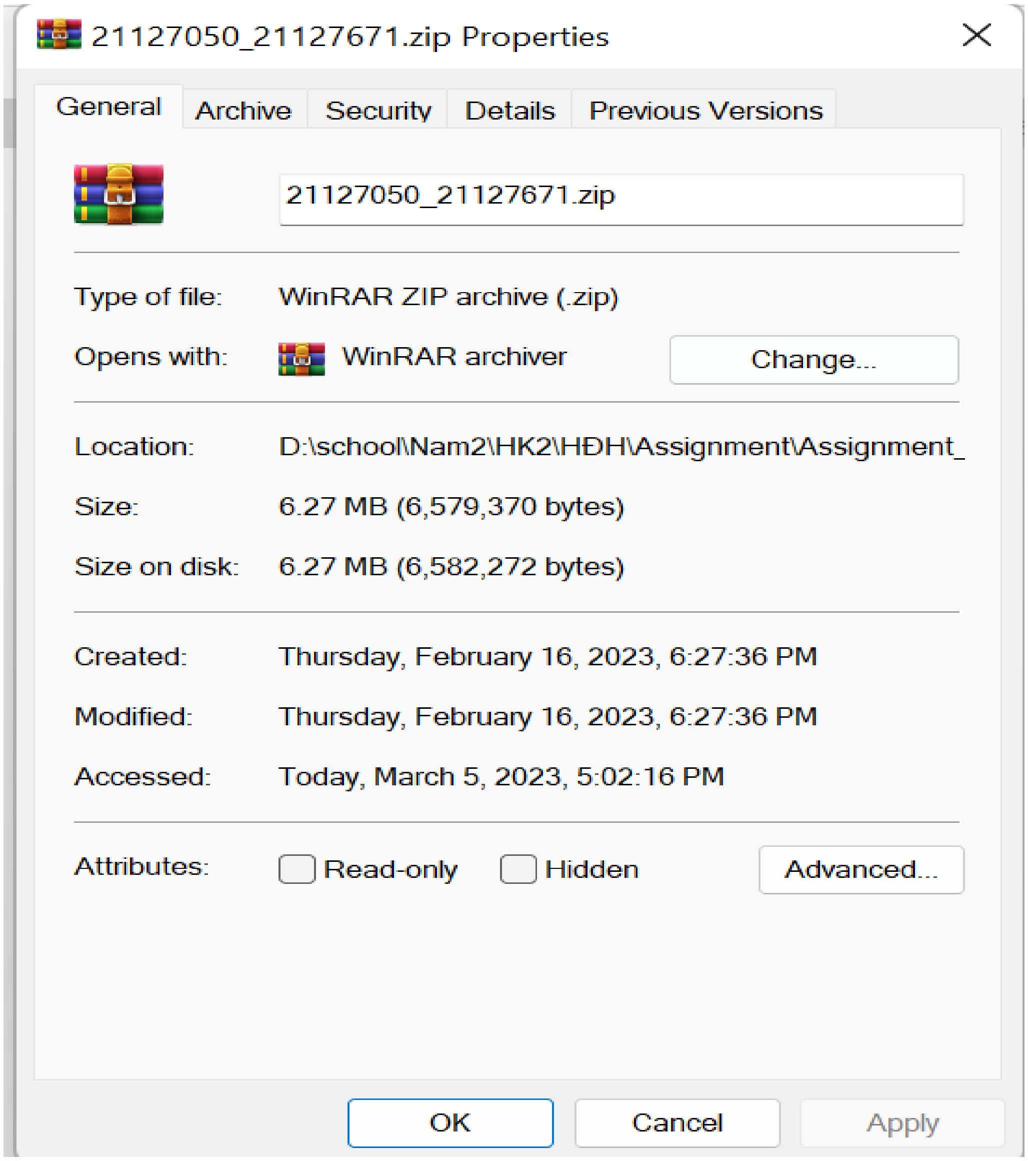
- Finally, the current size of the file is recorded, which is particularly important in old mainframe operating systems where the maximum size needed to be specified when the file was created.

- Here are some file attributes (1):

Attribute	Meaning
File name	Most operating systems today support filenames consisting of two parts: the name and the extension separated by a period. The length of each section depends on the storage organization of each operating system. Usually (as on Windows), the extension helps the operating system to recognize what file to run the corresponding program. However, for some operating systems (such as UNIX), the extension is only conventional so that it is easy for users to identify what file it is and can have multiple extensions, for example taptinen.tar. gz. Table II-1 shows some file extensions.
Protection	Who (user, group users, …) can access the file and in what way (read, write, execute…)
Password	Password needed to access the file
Creator	ID of the person who created the file
Owner	ID of current owner
Read-only flag	0 for read/write; 1 for read only
Hidden flag	0 for normal; 1 for do not display in listings
System flag	0 for normal files, 1 for system file
Archive flag	0 for has been backed up, 1 for needs to be backed up
ASCII/binary flag	0 for ASCII file; 1 for binary file
Random access flag	0 for sequential access only; 1 for random access
Temporary flag	0 for normal; 1 for delete file on process exit
Lock flags	0 for unlocked; nonzero for locked
Record length	Number of bytes in a record
Key position	Offset of the key within each record

Key length	Number of bytes in the key field
Creation time	Date and time the file was created
Time of last access	Date and time the file was last accessed
Time of last change	Date and time the file was last changed
Current size	Number of bytes in the file
Maximum size	Number of bytes the file may grow to

- Figure III-1 shows some file attributes on Windows.



21127050_21127671.zip

Type of file: WinRAR ZIP archive (.zip)

Opens with: WinRAR archiver

Location: D:\school\Nam2\HK2\HĐH\Assignment\Assignment_

Size: 6.27 MB (6,579,370 bytes)

Size on disk: 6.27 MB (6,582,272 bytes)

Created: Thursday, February 16, 2023, 6:27:36 PM

Modified: Thursday, February 16, 2023, 6:27:36 PM

Accessed: Today, March 5, 2023, 5:02:16 PM

Attributes: □ Read-only □ Hidden

Figure III-1. *Some file attributes on Windows*

IV. Some features of the file

IV.1. File Naming

- A file is a way to store information on a disk and retrieve it later, shielding the user from the details of how it's stored. When creating a file, a process gives it a name, which can be used by other processes to access it. File naming rules vary between operating systems, but typically allow one to eight letters, digits, and special characters. Some systems distinguish between upper and lowercase letters, while others don't. Windows 95 and 98 used the MS-DOS file system, which allowed for one to eight-character file names and an optional one to three-character extension. Windows NT, 2000, XP, Vista, 7, and 8 support both the FAT file systems and NTFS, which has different properties and supports Unicode file names. ExFAT is a modern Microsoft file system that is optimized for flash drives and large file systems, and can be read and written by OS X. Many operating systems support two-part file names, with the extension indicating something about the file, such as its format or content. Some of the more file extensions and their meanings are shown in Fig. IV-1.

Extention	**Meaning**
.bak	Backup file
.c; .cc; .java; .pas; .asm; .a	The file contains the source code of a programming language such as C, C++, Java...
.gif	Compuserve Graphical Interchange Format image
.hlp	Help file
.html	World Wide Web HyperText Markup Language document
.o; .obj	Object file (compiler output, not yet linked)
.pdf	Portable Document Format file
.ps	PostScript file
.tex; .txt; .rtf; .doc; .wp; .txt	The file contains text characters that can be read and understood by the user
.bmg; .jpg; .png	The file contains image
.mpeg; .mov; .rm; .mp3; .avi	The file contains image, sound
.arc; .zip; .rar; .tar	Compressed archive
.bat; .sh	File containing instructions to be executed sequentially.
.exe; .com; .bin; or no extension	File containing machine code, generated after linking object files, is the only file

	type that is executable by the operating system
.lib; .a; .so; .dll	The file contains a library of procedures that support programming

Figure IV-1. *Some typical file extensions. (1)*

IV.2. File Structure

Files can be structured in different ways, and three common types are shown in Fig. IV-2.

- The first type, as depicted in Fig. IV-2(a). Figure IV-2(a) depicts a simple file structure. The file's storage data is just an unstructured byte sequence. The semantics and organization of bytes of a file depend on the user program or the operating system. For example, the executable file (.exe) will have a way of organizing this byte sequence so that the Windows operating system can understand it; text files (.doc) have a different way of organizing this byte sequence so that the MSWord program can understand it. Most operating systems today (Windows, Linux) use this approach.
- The second type, as shown in Fig. IV-2(b). Figure II-2(b) depicts the file structure used on computers of the past decades (mainframe computers used punch cards). In this model, a file is divided into consecutive fixed-size records so that parallel processing of records can be made. However, this model is no longer used on current computer systems.
- The third type, as shown in Fig. IV-2(c), is a tree of records with a key field in a fixed position, sorted on the key field to allow rapid searching for a specific key. This model is used on some large mainframe computers for commercial data processing and is quite different from the unstructured byte streams used in UNIX and Windows.

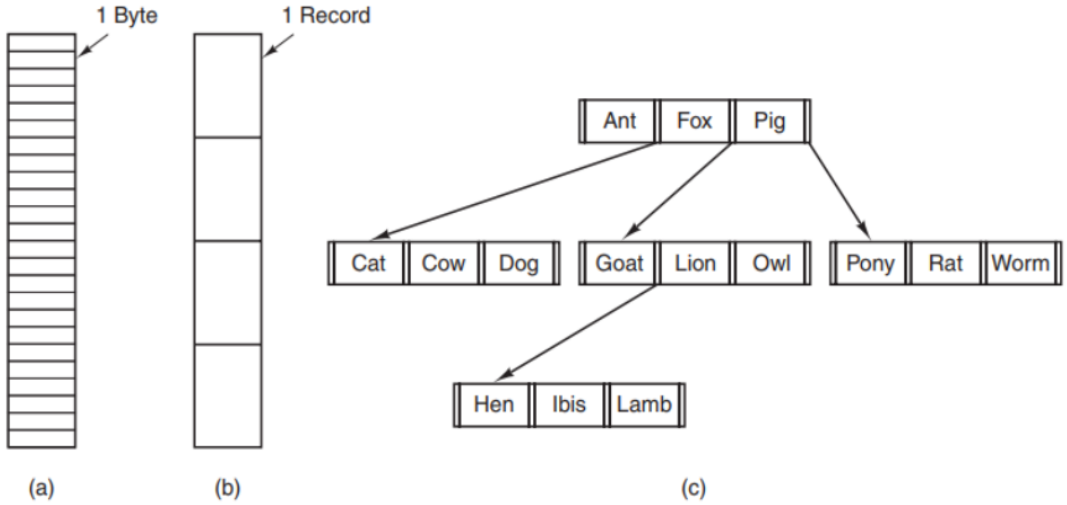


Figure IV-2. *Three kinds of files. (a) Byte sequence. (b) Record sequence. (c) Tree. (1)*

IV.3. File Types (3)

- Many operating systems, including UNIX and Windows, have different types of files, such as regular files and directories, character special files, and block special files. Regular files contain user information and can be either ASCII files or binary files.

- ASCII files consist of lines of text, while binary files do not and have an internal structure known to programs that use them.

- Executable binary files have a specific format that includes a header, text, data, relocation bits, and symbol table, as shown in Fig. IV-3(a). Archive binary files contain a collection of library procedures with headers full of binary numbers, as depicted in Fig. IV-3(b).

- Every operating system must recognize at least one file type, which is its own executable file. However, having strongly typed files can cause problems for users, such as being unable to compile an output file with the wrong extension. While this kind of user-friendliness may help novices, it can be frustrating for experienced users who need to circumvent the operating system's restrictions.

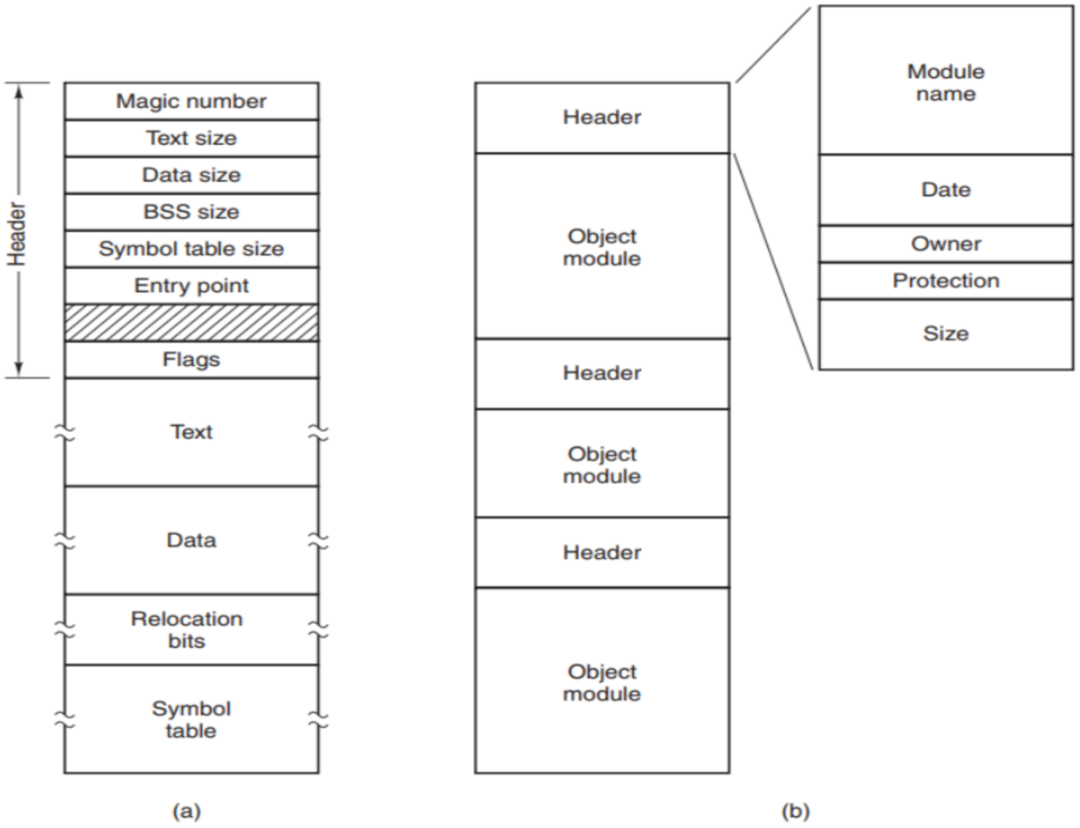


Figure IV-3. *(a) An executable file. (b) An archive (1)*

- The file is in ASCII code accessible by any text editor. Meanwhile, binary files can only be accessed by the respective programs. For example, the entity file is accessed by the respective programs. For example, the executable file (.exe) is only accessible by the Windows operating system; text files (.doc) can only be accessed by editors such as MSWord (on Windows) or OpenOffice (on Linux). A linked file is a file that does not contain data but links to normal files or directories. For example, shortcut file (.lnk) on Windows and hard link file, symbolic link on UNIX/Linux. In addition, on some systems (such as UNIX/Linux) special files are also supported. These special files are used to interact with I/O devices. For example, a special sequence file for interacting with serial devices such as keyboards, mice, modems, etc.; and special block files to interact with block devices such as disks, etc.

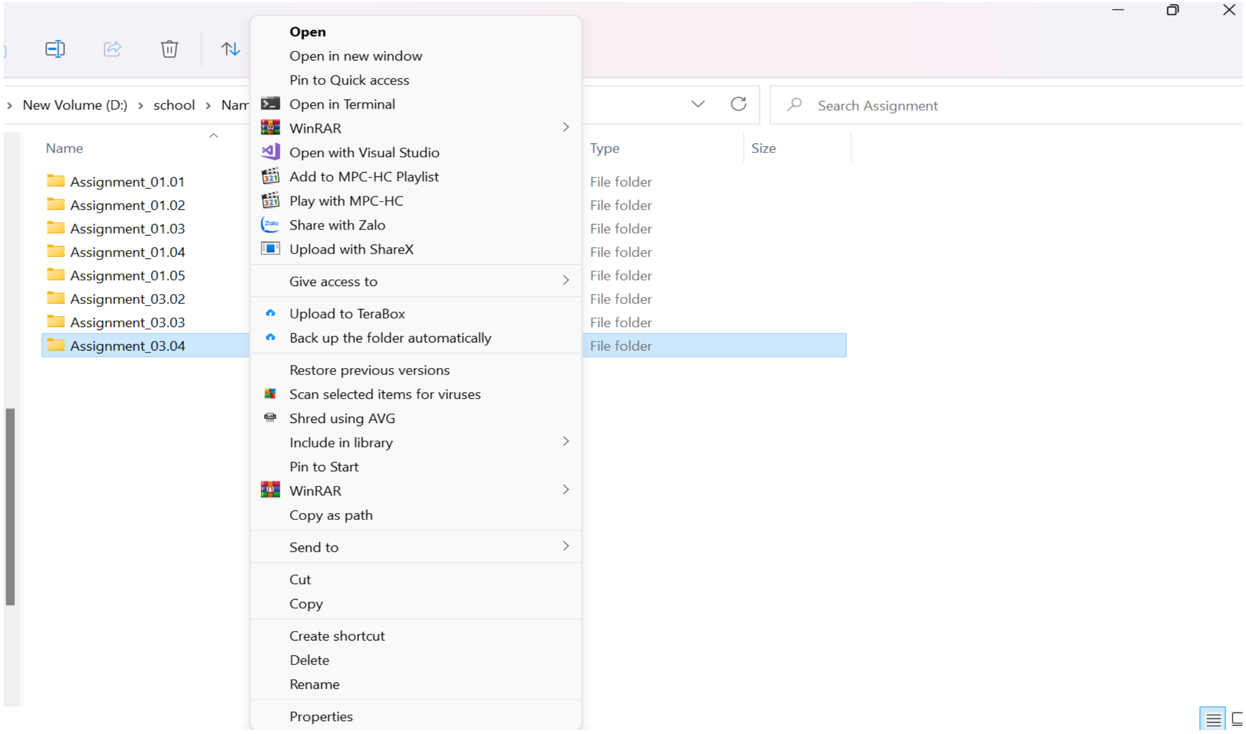
IV.4. File Access (4)

In the early days of operating systems, file access was limited to sequential access, where a process could only read data in order from the beginning of a file. Sequential files could be rewound and read multiple times, which was useful for storing data on magnetic tape. However, with the introduction of disks, random access files became possible, which allowed for reading data out of order or accessing records based on a key. Random access files are necessary for many applications, such as database systems, where specific records need to be accessed quickly. Two methods for specifying where to start reading in random access files are either providing a position in the file or using a special operation called seek to set the current position. The seek method is commonly used in UNIX and Windows. This method allows to move immediately to the location in the requested file.

IV.5. File Operations (2)

- The purpose of files is to store information and enable retrieval of that information later. Different computer systems provide various operations for file storage and retrieval. To make interacting with files easy, operating systems often provide operations such as creating, deleting, reading, writing, copying, renaming files, etc., as shown in Figure IV-5. The following are some of the most common system calls associated with files:

1. Create - creates a new file with no data and sets some of its attributes.
2. Delete - removes the file and frees up disk space.
3. Open - retrieves the attributes and list of disk addresses for the file and stores them in memory for later access.
4. Close - frees up internal table space and writes the file's last block to disk.
5. Read - retrieves data from the file, typically starting from the current position.
6. Write - writes data to the file, typically at the current position. If the current position is at the end of the file, the file's size increases.
7. Append - adds data only to the end of the file.
8. Seek - repositions the file pointer to a specific place in the file for random-access files.
9. Get attributes - retrieves file attributes, such as modification times.
10. Set attributes - changes user-settable file attributes, such as protection mode information.
11. Rename - changes the name of an existing file.



Name	Type	Size
Assignment_01.01	File folder	
Assignment_01.02	File folder	
Assignment_01.03	File folder	
Assignment_01.04	File folder	
Assignment_01.05	File folder	
Assignment_03.02	File folder	
Assignment_03.03	File folder	
Assignment_03.04	File folder	

Figure IV-5. *Some operations on files of the Windows operating system*

V. References

(1): https://csc-knu.github.io/sys-prog/books/Andrew%20S.%20Tanenbaum%20-%20Modern%20Operating%20Systems.pdf

(2): https://www.javatpoint.com/operating-system

(3): https://ecomputernotes.com/fundamental/disk-operating-system/what-is-files-types-of-files-types-of-file-operations

(4): https://www.geeksforgeeks.org/file-access-methods-in-operating-system/