FILE SYSTEMS

File Systems

- File organization
- Access methods
- Directory Organization: single, two-level, hierarchy
- File system and directory implementation
- Allocation schemes : contiguous, linked, indexed

File system

In a computer, a file system (sometimes written filesystem) is the way in which files are named and where they are placed logically for storage and retrieval.

The logical unit within a file system is the file

- logical files are mapped into physical entities by the OS
- in user's view, file is the smallest unit that can be saved to disk

- A file system defines the structure and the rules used to read, write, and maintain information stored on a disk.
- Which system used is determined by;
  - Hardware
  - Software
  - Security needs
  - Need for a dual-booting system

File Attributes

- *name : provides handle for reference*

  • DOS (8 chars + 3 char extension), Windows (unlimited? length)

  • UNIX (spaces tricky, no extension needed)

- *type : indicates how the file should be treated*

  • DOS/Windows rely on extension, can map extensions to programs

  • Mac associates creator attribute with each file (OS X?)

  • UNIX uses "magic number", first few bytes of file specify file type

protection : *permissions, access control information*
• UNIX utilizes permission string: chmod 644 foo.txt •
-rw-r--r—
owner & group: chown, chgrp
• Windows utilizes file properties/attributes: NoAccess,
List, Read, Read & Add, ...

Location & Size

Accounting Information

e.g. MS-DOS filename

- Filename: 1 to 8 characters
- File extension: 1 to 3 characters
- DOS only uses uppercase
- You cannot use spaces
- A single period separates the file extension from the filename
- You cannot use these characters:

    / \ : * ? " < > | + = ; , [ ]

The Full Path

- What does it mean?
  - The complete path from the root of the files system to the desired object.
  - Use \   (vs. /)
  - C:\ - the root

  For Example;

  E:\BSIT\6th Semester\System Administration\File System.ppt

File Names continued

- Filenames consist of two parts
    - Main part of filename
        - Identifies the contents of the file
    - File extension
        - Identifies the file type
- The last period separates the file extension from the main part of the filename.
- Windows typically associates a file extension with an application installed on a computer

Folder Names

- Folder names follow the same guidelines for naming files

- Folder names usually do not have a file extension

File Operations

- *create : find space on disk and make entry in directory*
- *write : write to file, requires positioning within the file*
- *read : read from file, involves positioning within the file*
- *delete : delete directory entry, reclaim disk space*
- *reposition : move read/write position*

the OS must maintain information about all open files

- *file pointer : the current position of the read/write pointer in the file*

- *disk location : the location of the file on the disk*

- *file open count : keep track of number of processes currently accessing the file*

such a table of information allows the OS to enforce policies such as only one process can write to a file at a given time

File Structure

files can be stored physically as

- Bytes
- Lines
- records

whatever entity is stored, OS must map into a disk sector

because on a physical disk, sectors are smallest writeable unit

access methods:

- *sequential : information in the file is accessed from first to last*

- readNext , writeNext , reset

- *direct : possible to reposition read/write pointer to any position*

- such files are generally made up of fixed-length records

- readRecord N, writeRecord N, positionAt N, reset

- *indexed : built on top of direct access, but accesses records in file using a key*

- each record has a key associated with it, an index of keys is stored with the file

- readRecord KEY, writeRecord KEY, positionAt KEY, reset

FAT???

file allocation table – where the OS records how the disk space is used

- locates the file allocation table near the beginning of the volume

- the location of the FAT is specified in the boot sector (BIOS Parameter Block)

- actually, 2 copies of the FAT are stored for redundancy

- the FAT number refers to the number of bits per table entry

Boot Sector	Reserved Sector	FAT 1	FAT 2 (Duplicate)	Root Folder	Other Folders and All Files

File Systems

- **FAT12:** The earliest version the file system, FAT12 allows a partition to contain up to 4096MB/ (212)clusters.

- **FAT 16:** oldest, created for DOS, supported by most OS's, cannot be installed on partitions larger than 2 GB, or on hard drives larger than 4GB.

- **FAT 32:** supports disks from 512MB to 2TB, compatible with Windows 98 and up

Contd..

- the FAT number refers to the number of bits per table entry
- FAT12 -> 2¹² = 4M different clusters can be addressed (used for floppy disks)
- FAT16 -> 2¹⁶ = 64M different clusters can be addressed (MS-DOS compatible)
- FAT32 -> 2²⁸ = 256G different clusters (4 bits are reserved)

Advantages of FAT File System

- The FAT file system is best for drives and/or partitions under approximately 200 MB

- It is better to format system partition as FAT

Disadvantage of FAT File System

- It is not better to use FAT on partitions that are greater than 200 megabytes. FAT partitions are limited in size to a maximum of 4 Gigabytes (GB) under Windows NT and 2 GB in MS-DOS.

- Any FAT partitions that use DOS-based disk compression (such as DriveSpace) will not have readable files when running Windows NT

- The FAT file system is also prone to fragmentation

.

NTFS

- New technology file system
- Better file security (Encrypting File System),
- Disk compression- can compress a file/folder, any Windows app will automatically expand as needed
- Disk Quota features- Can enforce quotas on disk usage

Contd..

- Recovery features-each file operation broken down into atomic transactions.

maintains a **transaction log** – updates disk after each transaction

if **failure** occurs during a transaction, info is sufficient to complete or rollback

if a **bad sector** is found when writing, will automatically map to a different sector

- NTFS volumes can not be accessed by DOS,
  or Windows 95 or Windows 98.

Advantages of NTFS

- **Large disks and large files:** *NTFS is best for use on volumes of about 400 MB or more*

- **Recoverability:** *The recoverability designed into NTFS is such that a user should never have to run any sort of disk repair utility on an NTFS partition.*

- **Security:** *NTFS uses the Windows NT object model to enforce security. An open file is implemented as a file object with a security descriptor that defines its security attributes.*

- **General indexing facility:** *NTFS associates a collection of attributes with each file. The set of files in the file management system is organized as a relational database, so that files can be indexed by any attribute.*

Disadvantages of NTFS

- It is not recommended to use NTFS on a volume that is smaller than approximately 400 MB
- Currently, there is no file encryption built into NTFS – encryption is available as an external option .
- It is not possible to format a floppy disk with the NTFS file system

File Systems

➡ The operating system keeps track of data (documents, pictures, etc.) by placing it into a file.

➡ To store and retrieve files:

    ➡ Disk divided into tracks

    ➡ Tracks are divided into sectors

    ➡ Sectors grouped into clusters

       ➡ Number of sectors in a cluster is determined by

         ➡ Size of the hard drive

         ➡ File allocation system – FAT, FAT32, NTFS

Sectors within a Track
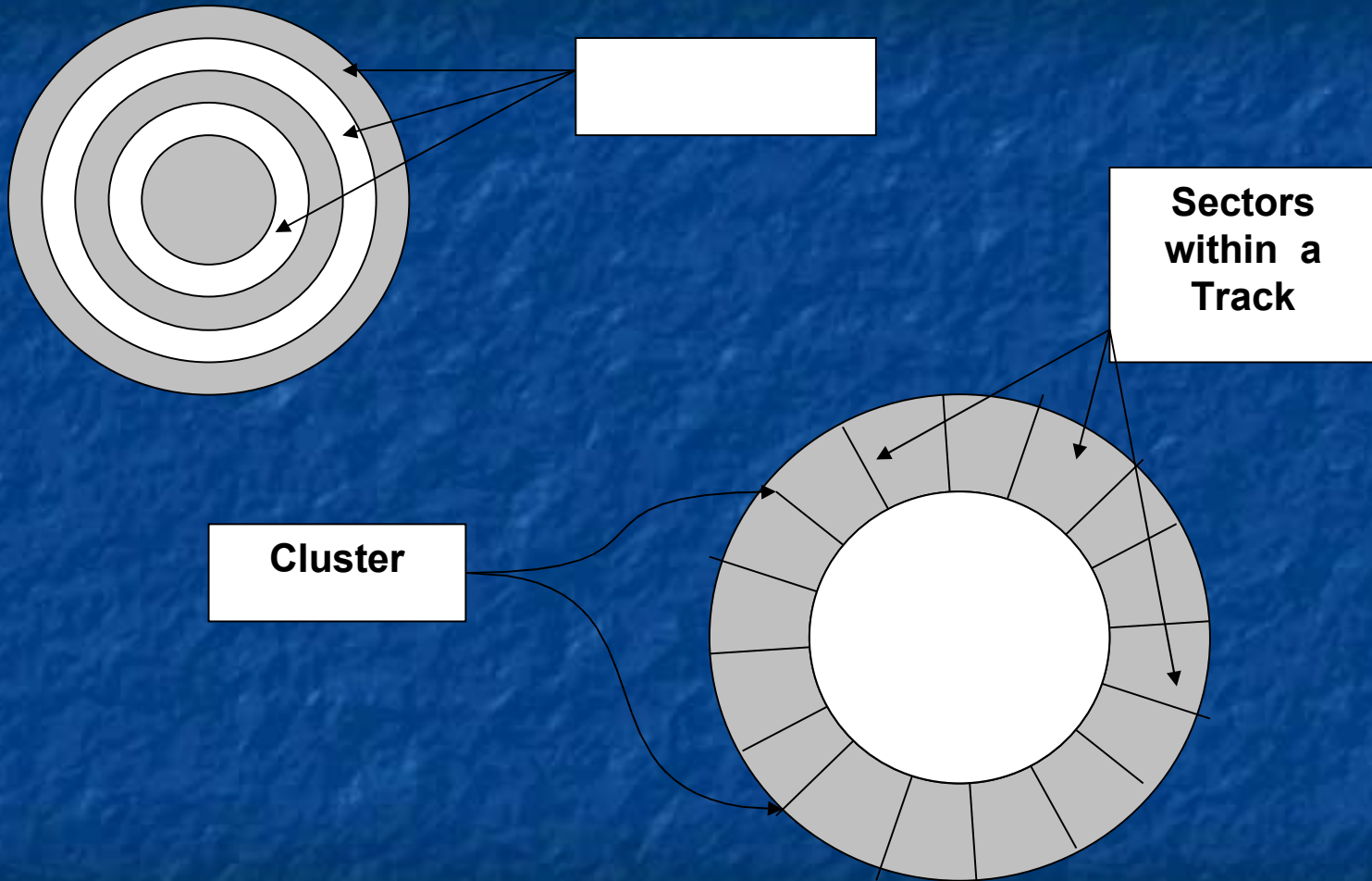
Cluster

Clusters?

- A **cluster**, *also known as an allocation unit,* consists of one or more sectors of storage space, and represents the minimum amount of space that an operating system allocates when saving the contents of a file to a disk.

- The number of sectors per cluster is dependent on
  - Type of disk (floppy disk, hard disk)
  - Version of operating systems
  - Size of disk

- Every sector contains 512 bytes. (NTFS does allow you to change this number.)

- The number of clusters per disk is determined by the filing system (FAT 16, FAT 32 or NTFS).

DRIVE SIZE	FAT 16 Cluster Size	FAT 32 Cluster Size	NTFS Cluster Size
260 to 511 MB	8 KB	Not Supported	512 bytes
512 to 1023 MB	16 KB	4 KB	1KB
1024 MB to 2 GB	32 KB	4 KB	2 KB
2 to 4 GB	64 KB	4 KB	4 KB
4 to 8 GB	Not Supported	4 KB	4 KB
8 to 16 GB	Not Supported	8 KB	4 KB
16 to 32 GB	Not Supported	16 KB	4 KB
>32 GB *(up to 2 TB)*	Not Supported	32 KB	4 KB

OS and File System Compatibility

Operating System	FAT16	FAT32	NTFS
Windows XP	☑	☑	☑
Windows 2000	☑	☑	☑
Windows NT	☑		☑
Windows 95, 98, ME	☑	☑	
Windows 95	☑		
MS-DOS	☑		

What happens during the Formatting Process?

- OS creates four tables in the 1st sectors
  - **Boot Record** – the name & version number of the OS, info. on the physical characteristics of the disk
  - **Master File Table #1-** keeps track of
    - Available clusters
    - Clusters that contain data
    - Clusters that are defective
    - Clusters that contain OS files
  - **Master File Table #2** – copy of MFT #1
  - **Directory Table** – top level folder and file information

Formatting Process, cont.

- Full Format
  - lays down new tracks and sectors
  - Verifies the integrity of each sector
    - By doing a surface scan – OS will put dummy data into sectors and then try to read the sector
- Quick Format
  - Removes files
  - Does not check for defective sectors

Formatting Process, cont.

Should you format brand new preformatted disk?

YES

Registered Files

- A file that is associated with an application on your computer via its file extension.

- Where is this information kept?

- THE REGISTRY – a database that consists of a set of files where Windows XP stores your computer's hardware, software, network, security,user settings or profiles, and property settings for folders and programs.

Compare FAT and NTFS File System

- ## Compatibility

  Only Windows NT supports NTFS partitions , Windows NT and Windows 95 supports both NTFS and FAT

- ## Volume size

  FAT supports partition sizes only up to 2 GB , the maximum size of an NTFS partition is 16 EB

Fault tolerance

Windows NT offers software support for several alternate disk-access methods that increase speed and/or fault tolerance, While FAT does maintain two copies of the file-allocation table, in case one copy is damaged, it's incapable of automatically fixing errors)

File compression

NTFS has its native support for file compression, It offers you the chance to compress individual files and directories of your choice

The system partition

A better solution is to format your system partition as FAT because NTFS partitions are accessible only via Windows NT. If you have a fatal error with Windows NT, you can't simply boot a system disk to a command prompt and fix a problem on an NTFS partition

Converting to NTFS

In MS-DOS type command: CONVERT drive: /FS:NTFS

Security

NTFS has a built-in security system , FAT has no local protection, it only has the share permission (protect the file from network

Ext3 third extended file system

- Stephen Tweedie
- November,2001
- Journaled file system commonly used by linux.
- (journaling capability means no worrying about metadata corruption. What is most noticeable is that you can switch back and forth between ext2 and ext3 on a partition without any problem: it is just a matter of giving the mount command the right file system type.)

Advantages

(over ext2 file system)

- Simple
- Have the facility of backup and restore data
- journaling improves reliability

- indexing for larger directories.

Size limits

Block size	Max file size	Max filesystem size
1 KB	16 GB	2 TB
2 KB	256 GB	8 TB
4 KB	2 TB	16 TB
8 KB	2 TB	32 TB

Journaling levels

There are three levels.

- Journal (lowest risk)
- Ordered (medium risk)
- Write back (highest risk)

- Both metadata and file contents are written to the journal before being committed to the main file system.

- Only metadata is journaled; file contents are not, but it's guaranteed that file contents are written to disk before associated metadata is marked as committed in the journal.

- Only metadata is journaled; file contents are not. The contents might be written before or after the journal is updated.

Thanks!!!!!!!!!!