Operating System Fundamentals

Module 4:

Files Systems

- Description
- Functions
- General Operation
- Utilities
- Common File Systems
- Files and Directories (Folders) in Windows

Agenda

- Mechanism in an OS to store, update and retrieve data in an organized way
- Efficient and reliable, often tuned specifically for the OS
- Generally uses permanent storage
  - Exception is a RAM disk
  - Common media: floppy, hard drive, optical drive, flash (memory stick), tape, cloud

Description

- Local - generally block oriented
- Network – generally stream oriented
  - UNC names widely used (\\servername\share\filepath)
  - Could be abstracted to look like Local
- Cloud – transparent with little folder navigation
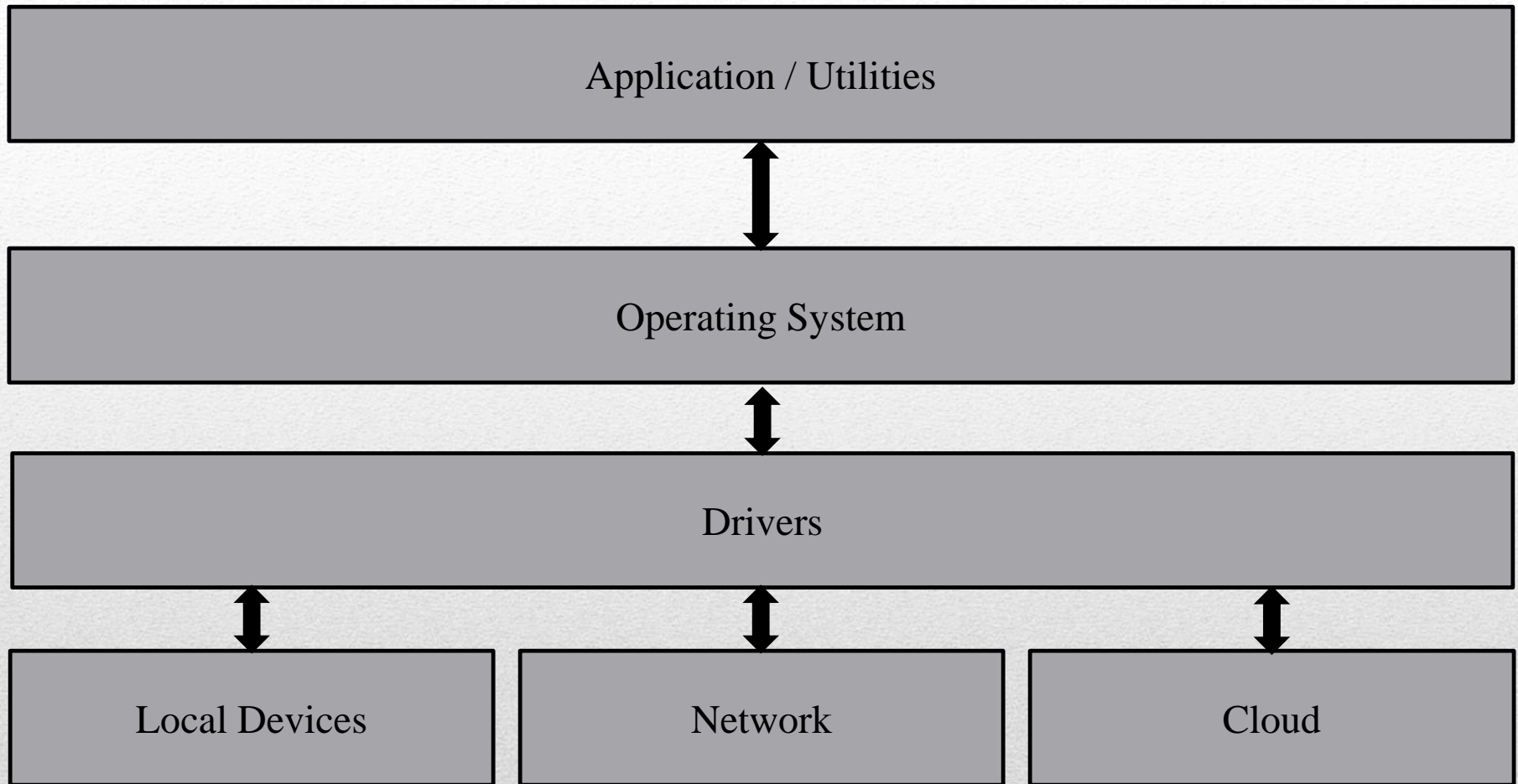 - Could be abstracted to look like Network or Local

Description

- Files accessed using a pathname or fully qualified file name
- Folder hierarchy used to organize files
- Information tracked (usually in directory)
  - Filename
  - Date of modification (and creation)
  - File attributes

Description

- Formatting
  - Puts important data structures on the volume related to the file system
    - Master Boot Record
      - Starts the boot process; transfers control to Boot Sector
    - Partition table
      - Information about the file system
    - File Allocation Table
    - Optionally initialize all data in sectors

Description

```
┌─────────────────────────────────────────────────────┐
│                Application / Utilities                │
└─────────────────────────────────────────────────────┘
                         ↕
┌─────────────────────────────────────────────────────┐
│                  Operating System                     │
└─────────────────────────────────────────────────────┘
                         ↕
┌─────────────────────────────────────────────────────┐
│                      Drivers                          │
└─────────────────────────────────────────────────────┘
         ↕                   ↕                   ↕
┌──────────────┐    ┌──────────────┐    ┌──────────────┐
│ Local Devices│    │   Network    │    │    Cloud     │
└──────────────┘    └──────────────┘    └──────────────┘
```

General Hierarchy

- Good references:
  - Wikipedia (Technical Detail): http://en.wikipedia.org/wiki/File_Allocation_Table
  - Microsoft Technet http://technet.microsoft.com/en-us/library/cc776720(v=WS.10).aspx

File Allocation Table (FAT): General Operation

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

Source: Technet

FAT Volume

- Sector is physical unit of storage on a disk (512 bytes)
- Cluster is smallest amount of disk space allocated to a file
  - Also known as Allocation Unit
- FAT16 – 16-bit
  - Total of 2^{16} clusters in the FAT
- FAT32 – 32-bit
  - Total of 2^{28} clusters (theoretically)
- Number of clusters determine the minimum space a file can occupy
 - FAT32 allows more efficient use of disk space

FAT: Clusters

Volume Size	FAT16 Cluster Size	FAT32 Cluster Size
7 megabytes (MB)–16 MB	2 KB	Not supported
17 MB–32 MB	512 bytes	Not supported
33 MB–64 MB	1 KB	512 bytes
65 MB–128 MB	2 KB	1 KB
129 MB–256 MB	4 KB	2 KB
257 MB–512 MB	8 KB	4 KB
513 MB–1,024 MB	16 KB	4 KB
1,025 MB–2 gigabytes (GB)	32 KB	4 KB
2 GB–4 GB	64 KB	4 KB
4 GB–8 GB	Not supported	4 KB
8 GB–16 GB	Not supported	8 KB
16 GB–32 GB	Not supported	16 KB
32 GB–2 terabytes	Not supported	Not supported

Source: Technet

FAT Cluster Sizes

- File Directory has file/folder information
- Root Directory indicates starting cluster of file
- File allocation table is a map of clusters, indicating where a file is stored
- Each block in the table below represents a cluster on the volume:
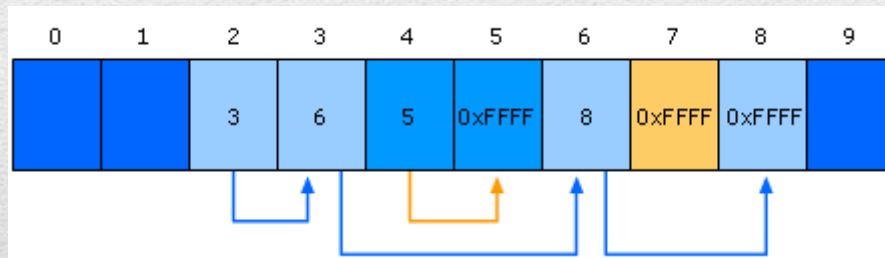
FAT File Processing

- Directory
  - Create
  - Delete
  - Rename
  - List Directory
- File
  - Create
  - Update
  - Delete

Functions

- Create/Delete/Rename/List directory
- Create/Move/Copy/Delete/Undelete/Purge files
- Disk Integrity Check
- Defragmentation
- Backup/Restore

Utilities

- Apple (OS X)
  - HFS and HFS Plus
    - Forked files
- Linux
  - ext2, ext3, ext4 amongst many others
- Windows
  - FAT32, NTFS amongst many other proprietary ones
  - Drive letter oriented
- See list: http://en.wikipedia.org/wiki/List_of_file_systems

Common File Systems

- Files organized in directories/folders in a hierarchy
- Path defines the location of a folder (e.g. C:\temp or \\mycomputer\sharedfolder\anothertemp)
- "Under the covers", file names are added to the path to create pathname (e.g. C:\temp\file.txt)
  - Enables the use of the same file name in different directories, effectively creating a unique name for each
    - e.g. C:\temp\file.txt and C:\nm\file.txt are different pathnames, allowing the repeated use of file.txt, even though they may be different files altogether)

Files and Directories (Folders) in Windows

- Universal Naming Convention (UNC)
  - Standardized approach to accessing resources on a network
  - Platform independent (not just Windows)
  - Format: *ServerName**ShareName**pathname*
    - *ServerName* is the shared network computer
    - *ShareName* is shared computer's reference to a shared folder
      - Does not need to be the same name as the actual shared folder
      - "Hides" the actual location of the shared folder
    - *pathname* is directory structure/file name based on the shared folder as the root
  - Highly recommended instead of using drive letters
  - If desired, can map drive letter to the UNC

Files and Directories (Folders) in Windows

- Useful commands in Command Prompt:
  - Change drive letter: *d:* where *d* is the desired drive
  - Make Directory: mkdir *or* md
  - Change Directory: chdir *or* cd
  - Remove Directory: rmdir *or* rd
  - Copy File(s): copy *or* xcopy
  - Move File(s): move
  - Delete File(s): erase *or* del
  - Rename Directory or File: rename *or* ren

Files and Directories (Folders) in Windows

- Fully qualified paths and pathnames (starting with the drive letter) leave no ambiguity
  - When moved between computers, could cause problems if the directories do not exist
- Relative paths and pathnames start without the drive letter
  - Generally concatenated to the current path
    - (e.g. If you are in *C:\temp*, reference *dir1\test.txt,* the system references *C:\temp\dir1\test.txt*.)
  - Good for portability and reduced typing
- Shortcuts:
  - The "double dot" (..) moves you up a level in the directory structure
  - The "single dot" (.) refers to the current directory

Files and Directories (Folders) in Windows

- Use relative references in your code/applications to allow portability (if appropriate)
- UNC names instead of drive letters
  - Particularly useful in Virtual Private Networks (VPN)
- Encrypt folders or drives on laptops
- Compress folders/files to save disk space
  - Very useful for less frequently use files
  - Not effective on already compressed formats (e.g. video, audio, "zipped" folders

Some considerations