OS History and UNIX

1. What is an Operating System?

| Various computer programs  
e.g., text editor, compiler, games, etc. |
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating System</td>
</tr>
<tr>
<td>Computer Hardware Resources</td>
</tr>
</tbody>
</table>

Figure 1.1 Conceptual Layers of Computer System

“An operating system is a program that **controls** and **coordinates** the use of hardware **resources** among various computer programs for various users.”

- **Resources**: CPU, memory, I/O devices (secondary & tertiary storage, modem, monitor, etc.)

- **Control (Control Program)**: OS controls the execution of various programs to ensure proper use of the hardware resources and to prevent errors.

- **Coordination (Resource allocator)**: OS acts as a manager of underlying resources and allocates/de-allocates resources efficiently and fairly.

2. Why We Need an OS? (The Goal of an OS)

Convenient and Efficient use of computer hardware resources.

3. Type of OS

*The first generation (1945-1955) Vacuum Tubes and Plugboards based Computers: Human operator (No Operating System), on-line*

’44 MARK I by Howard Aiken, Harvard Univ. – Electromechanical Computer

’45 ENIAC by J.W. Mauchly and J.P. Eckert – Electrical Numerical Integrator and Computer

’49 EDSAC by Alan Turing – 1st library of subroutines
The second generation (1955-1965) Transistor based Computers: Developed Resident Monitor (FMS, IBSYS) and Off-line system to utilize CPU.

- '55 TRIDAC – 1st transistor based computer
- '55 – '64 IBM 1401 (Printing, Card reading, commercial) + IBM 7094 (heavy computing, scientific), Fortran Monitor System (FMS), IBM’s OS for 7094 (IBSYS)

→ Transistor made the computers much more stable
Big companies and government departments started buying computer systems. Because the system price was very high, the owners wanted to maximize the system utilization. As a result, batch system was developed.

→ Off-line Batch system increased the CPU utilization

![On-Line Processing](image)

Main Computer must wait the slow Card-reader and Line printer. Cannot I/O while Main Computer is working. Interactive computing environment

![Off-Line Processing](image)

Main computer was not constrained by the speed of slow card-readers and line printers, but was limited by only the speed of the much faster tape drives. Less interactive environment

Dr. Byunggu Yu

'59 Integrated Circuit (IC) by Noyce and Moore
'61 PDP-1 by DEC
'62 CTSS (Compatible Time-Sharing System) by Corbato at MIT – 1st time-sharing OS
'64 IBM 360 by IBM – 1st IC based Computer
'64 PDP-8 by DEC – 1st mass-produced minicomputer
'66 OS/360 – Multiprogramming equipped OS
'66 MULTICS (MUltiplexed Information and Computing Service) by MIT, Bell Lab, and General Electric
'69 UNIX by Thompson and Ritchie at AT&T Bell Lab.
'72 PDP-11 by DEC – The climax of DEC’s PDP series

→ IC reduced the price of computer systems.

IBM produced general-purpose computer IBM 360 that can substitute for both 1401 (for I/O-bound jobs) and 7094 (for CPU-bound jobs).

→ Multiprogramming

When the current job is waiting for an I/O operation to complete, the CPU simply sits idle. With CPU-bound scientific calculating jobs, I/O is infrequent, so this wait time is not significant. However, with I/O-bound commercial data processing jobs, the I/O wait time is significant. Because general-purpose computers, like IBM 360, do both kinds of jobs (I/O-bound and CPU-bound), something had to be done about it. People started modifying/developing operating systems, like OS 360, to load several jobs into the memory so that, while one job is waiting for an I/O to complete, another job can be using the CPU. IBM 360 and several systems were equipped with a special hardware to protect each job against snooping and mischief by the other jobs.

→ Spooling (Simultaneous Peripheral Operation On Line)

Random access storage devices (i.e., Disk storage) made the following possible: the computer systems can read (or write) jobs (or outputs) while other devices are reading or writing data on the storage. As a result, as soon as input data or program is ready, the data or program could be written on the disk. Also, the computer could write output data onto the disk as soon as they are available. Spooling eliminates the delays in tape system (tape delivery, unloading/mounting, tape rewinding, and the delay until enough other jobs or data fill the current tape).

Job pool and Job Scheduling

→ Time-sharing (Multitasking)

Batch systems are designed to maximize the utilization of system resources. However, the users cannot interact with their jobs. Thus, running multi-step programs, whose next step depends on the result of the current step, and debugging programs were very difficult. The desire for quick response time gave birth to time-sharing (a logical
extension of multiprogramming. In time-sharing, the system switches rapidly and frequently from one job to another. As a result, each job continues to run by using many tiny slices of CPU time.

Note, Multiprogramming and Time-sharing are the central themes of modern operating systems.

**The Fourth generation (1980-Present) LSI (Large Scale Integrated) and VLSI (Very Large Scale Integrated) circuit based Computers**

- People started thinking about more convenient ways of computing
  - My computer is more convenient to use than shared computer: Personal Computer
  - Graphic User Interface (GUI) is more convenient than text based command interpreter: Mac OS, Windows series, X-windows

- Now, how to share information and resources
  - Network becomes more popular, Internet was born
  - Distributed Operating Systems become hot issue

- At the same time, I want protection and security
  - Information Security
  - Resource Protection

- My PC is too slow to solve my problems
  - Parallel Systems
  - Super Computers
  - Well organized distributed systems and clustered systems

**Special Type Operating Systems**

- Two Types of Multiprocessor System
  - Tightly coupled: Processors share computer bus, clock, memory, and I/O devices. Processors are in close communication.
    - Symmetric Multiprocessing: Each processor self-schedule its processes (more complex)
    - Asymmetric Multiprocessing: Each processor is assigned a specific task. (e.g., Master/Slave structure, less complex)
  - Loosely coupled: Each processor has its own memory, I/O devices, and clock. Processors do not share computer bus. Processors communicate through communication lines (e.g., LAN, phone lines)

- OS for tightly coupled multiprocessor systems (Parallel OS)
  - Increase throughput
  - Fault tolerant / Reliable
  - Increased complexity
OS for loosely coupled multiprocessor systems (Distributed OS or OS for clustered systems)
+ Increase throughput
+ Fault tolerant / Reliable
+ Resource Sharing
- Increased complexity

Real-time systems (Real-time OS)
- Soft Real-time system: A real-time job gets higher priority and maintains it until it completes. (e.g., multimedia and virtual reality applications) – can implement on general computer systems.
- Hard Real-time system: Guarantees that real-time jobs are completed on time. (e.g., missile control and robotics) – need special hardware to implement

Next Generation?
Let’s think about Moore’s law, Quantum Computing, Biological Computing, and DNA Computing.

4. UNIX

Dawn of UNIX
1) Growth of minicomputers (DEC PDP-1 ’61 – DEC PDP-11 ’72)

2) AT&T and Bell Lab developed Multics, a multi-user OS known as the “Multiplexed Information and Computing Service”

3) One of Multics developers Ken Thompson wrote a simple single-user version Multics for PDP-7 w/ 9Kbyte memory: This is the first Unix OS.

4) Brian Kernigham called this OS “Unics” (Uniplexed Information and Computing Service)

5) UNIX was ported onto PDP-11 w/ 24Kbyte memory in 1971 (with a word-processing system).

6) Dennis Ritchie and Ken Thompson at Bell Lab rewrote the UNIX in C (became a portable OS)

7) AT&T Bell Lab licensed UNIX to universities almost free

8) UNIX spread rapidly through Interdata 7/32, VAX, and Motorola 68000
Mac OS X is based on BSD UNIX

Dr. Byunggu Yu
5. UNIX Standardization

→ In 1983, IEEE (The Institute of Electrical and Electronics Engineers) start developing Portable Operating System Interface (POSIX), defining the interface/commands and behaviors of OS


→ In 1999, IEEE and the Open Group announced a revised POSIX, which was formally adopted by the International Organization for Standardization (ISO) in 2003.

“UNIX” Trademark owner:


6. UNIX GUI and Windows-based User Interface

X Window System → a variety of Windows Managers (e.g., Motif Window Manager, etc) based on X Window System

Miscellaneous Terms

Bit: binary digit
Byte: a group of eight bits
Word: the number of bytes/bits the CPU can manipulate at a time (register size)

Kilo: $2^{10}$ (1024)
Mega: $2^{20}$
Giga: $2^{30}$

Hertz: clock speed (the number of pulses per sec.)
Flops: floating-point operations per sec.
MIPS: millions of instructions per sec.