Introduction to OSs

- What is an Operating System?
- Architectural Support for Operating Systems
- System Calls
- Basic Organization of an Operating System
What is an operating system?

- What an operating system is not:
  - An o.s. is not a language or a compiler
  - An o.s. is not a command interpreter / window system
  - An o.s. is not a library of commands
  - An o.s. is not a set of utilities

A Short Historical Tour

  - Single user: writes program, operates computer through console or card reader / printer
  - Absolute machine language
  - I/O devices
  - Development of libraries; device drivers
  - Compilers, linkers, loaders
  - Relocatable code
Programming Early Machines

Wiring the ENIAC with a new program
(U.S. Army photo, from archives of the ARL Technical Library)


- Problems: scheduling, setup time
- Automation of Load/Translate/Load/Execute
  - Batch systems
  - Monitor programs
- Job Control Language
- Advent of operators: computers as input/output box
- Problem: Resource management and I/O still under control of programmer
  - Memory protection
  - Timers
  - Privileged instructions

Monitor device drivers job sequencer / loader control card interpreter user program area
Example: IBM Punch Card System

Card Punch

Card Verifier

Card Sorter

(Computer Museum of America)

Batching Program Execution

Job

$FTN
$JOB
$END
... Data ...
$RUN
$LOAD...
... Program ...
$FTN
$JOB
Overlapping CPU and I/O Operations

Traditional Batch Operation:

- Card reader → CPU → Line printer

Off-Line Processing:

- Card readers → CPU → Line printers

Spooling: I/O Channels:

- Card reader → Disk → CPU → Line printer

Off-Line vs. Pure Batch

- **batch**

  - Card reader → CPU → Printer

- **off-line (single set of card reader/printer)**

  - Card reader → CPU → Tape reader → Printer
    - Rewind → Setup

- **off-line (multiple sets of card reader/printer)**

  - Card reader → CPU → Tape reader → Printer
    - Rewind → Setup
Off-Line vs. Pure Batch (II)

- **batch**
  
  - card reader
  - CPU
  - printer

- **off-line** *(multiple of card readers/ printers)*
  
  - card reader
  - tape reader
  - CPU
  - tape reader
  - printer


- Problem with batching: one-job-at-a-time
  
  - sequential:
    - CPU
    - I/O
  
  - better:
    - CPU
    - I/O

- Solution: **Multiprogramming**
  
  - Job pools: have several programs ready to execute
  - Keep several programs in memory

- New issues:
  
  - Job scheduling
  - Memory management
  - Protection
Time Sharing (mid 1960s on)

- OS interleaves execution of multiple user programs with time quantum
  - CTSS (1961): time quantum 0.2 sec

- User returns to own the machine

- New aspects and issues:
  - On-line file systems
  - Resource protection
  - Virtual memory
  - Sophisticated process scheduling

- Advent of systematic techniques for designing and analyzing OSs.

The Recent Past

- Personal computers and Computing as Utility
  - History repeats itself

- Parallel systems
  - Resource management
  - Fault tolerance

- Real-Time Systems

- Distributed Systems
  - Communication
  - Resource sharing
  - Network operating systems
  - Distributed operating systems

- Secure Systems
The Future?

- The “Invisible Computer”
- Computing-in-the-ultra-small
- Speed vs. Power vs. Heat
- Breaking up the layered design

What, then, is an Operating System?

- Controls and coordinates the use of system resources.
  - **Primary goal**: Provide a *convenient* environment for a user to access the available resources (CPU, memory, I/O)
    - Provide appropriate abstractions (files, processes, ...)
    - “virtual machine”
  - **Secondary goal**: *Efficient* operation of the computer system.
- **Resource Management**
  - **Transforming**: Create virtual substitutes that are easier to use.
  - **Multiplexing**: Create the illusion of multiple resources from a single resource
  - **Scheduling**: “Who gets the resource when?”
Resources

- Disks and other Devices
- Timers / Clocks
- Locks
- CPU
- Memory
- Power / Heat
- I/O Controllers

The OS as Servant to Two Masters

- Devices
- Clocks & Timers
- Locks
- Memory
- Heat & Power
- I/O Controllers
- CPUs

- Performance
- Plug & Play
- Security
- Predictability
- Convenience
- Fault-Tolerance
- Power-Effectiveness
Introduction to OSs

• What is an Operating System?

• Architectural Support for Operating Systems

• System Calls

• Basic Organization of an Operating System

Architectural Support for OS’s

• Dealing with Asynchronous Events: Exceptions, Interrupts
  - Modern OS’s are interrupt-driven (some still are not!).
  - Simple interrupt handling vs. exception handling MIPS-style.

• Hardware Protection
  - Privilege Levels (e.g. user/kernel/ supervisor, etc.)
  - Privileged instructions: typically CPU control instructions
  - I/O Protection
  - Memory Protection

• Support for Address Spaces

• Timers
Modern OS’s are Interrupt-Driven

Interrupts / Exceptions

- When an interrupt occurs, CPU stops, saves state, typically changes into supervisor mode, and immediately jumps to predefined location.
- Appropriate interrupt service routine is found through the interrupt vector.
- Return-from-interrupt automatically restores state.

- Interrupts/Exceptions can be invoked by asynchronous events (I/O devices, timers, various errors) or can be software-generated (system calls).
Exceptions, MIPS-Style

- MIPS CPU deals with exceptions.
  - Interrupts are just a special case of exceptions.

- The MIPS Architecture has no interrupt-vector table!
  - All exceptions trigger a jump to the same location, and de-multiplexing happens in the exception handler, after looking up the reason for the exception in the cause register.

MIPS Exception Handler (low-level)

```
xceptlow_handler
  set up exception frame on stack
  save enough registers to get by
  save rest of registers
  call C exception handler
  restore registers
  return from exception
```
Hardware Protection

- Originally: User owned the machine, no monitor. No protection necessary.
- Resident monitor, resource sharing: One program can adversely affect the execution of others.
- Examples
  - `halt` and other instructions
  - modify data or code in other programs or monitor itself
  - access/modify data on storage devices
  - refuse to relinquish processor
- Benign (bug) vs. malicious (virus)

Hardware Protection (2)

- Dual-mode operation
  - `user mode` vs. `supervisor mode`
  - e.g. `halt` instruction is privileged.
- I/O Protection
  - define all I/O operations to be privileged
- Memory Protection
  - protect interrupt vector, interrupt service routines
  - determine legal address ranges

![Diagram showing CPU, base, base + limit, and memory with conditions and trap to operating system!]}
Timers

- Timers can be set, and a trap occurs when the timer expires. (And OS acquires control over the CPU.)

- Other uses of timers:
  - time sharing
  - time-of-day

Introduction to OSs

- What is an Operating System?
- Architectural Support for Operating Systems
- System Calls
- Basic Organization of an Operating System
### External Structure of an OS

The outsider’s view of the OS.

![Diagram of the external structure of an OS]

- Applications programs/processes
- System call interface
- Kernel
- Device drivers
- Hardware

### System Calls

Provide the interface between a process and the OS.

#### Example: vanilla copy:

```c
int copy(char *.fname1, *fname2) {
    FILE *f, *g;
    char c;
    f = fopen(fname1, "r");
    g = fopen(fname2, "w");
    while (read(f, &c, 1) > 0) {
        write(g, &c, 1);
        fclose(f);
        fclose(g);
    }
}
```
System Call Implementation: Linux on x86

- Example: `syscall(int, setuid, uid_t, uid)`
- expands to:
  ```
  setuid:
  subl $4,esp
  pushl $esp
  movl 12($esp),eax
  movl $0x0,4($esp)
  movl $23,leax    System Call number (setuid = 23)
  movl 4($esp),ebp
  int $0x80
  movl leax,esp
  testl leax,leax
  jpe L2
  negl leax
  movl leax,errno
  movl 5-1,leax
  popl leax
  addl $4,leax
  retL2:
  movl leax,leax
  popl leax
  addl $4,leax
  ret
  ```

Why Interrupts?

**Reason 1:** Can load user program into memory without knowing exact address of system procedures

**Reason 2:** Separation of address space, including stacks: user stack and kernel stack.

**Reason 3:** Automatic change to supervisor mode.

**Reason 4:** Can control access to kernel by masking interrupts.
Reason 2: Buffer Overrun Attacks (Silberschatz et al)

```c
#include <stdio.h>
#define BUFFER_SIZE 256
int main(int argc, char *argv[]) {
    char buffer[BUFFER_SIZE];
    if (argc < 2)
        return -1;
    else {
        strcpy(buffer, argv[1]);
        return 0;
    }
}
```

Stack Separation sufficient?

- Buffer overruns in kernel code?
- Device drivers?
**Reason 4: Mutual Exclusion in Kernel**

1. User process 1
   - User space
   - System call
   - Trap
   - Process 1 executing in kernel
   - Interrupts are masked

2. User process 2
   - Trap
   - Process 2 can not enter kernel because of masked interrupts

3. Unmask interrupts and return

**Introduction to OSs**

- What is an Operating System?
- Architectural Support for Operating Systems
- System Calls
- Basic Organization of an Operating System
External Structure of an OS

The outsider's view of the OS.

- applications programs/processes
- system call interface
- kernel
- device drivers
- hardware

Internal Structure: Layered Services

The insider's view of the OS.
Example: XINU [Comer 1984]

- user programs
- file system
- intermachine network communication
- device manager and device drivers
- real-time clock manager
- interprocess communication
- process coordinator
- process manager
- memory manager
- hardware
Internal Structure: µ-Kernels

- Layered Kernels vs. Microkernels

Hierarchical decomposition. Interaction only between adjacent layers.

Kernel has only core operating system functions (memory management, IPC, I/O, interrupts)
Other functions run in server processes in user space.

Operations in a µ-Kernel

- Non-kernel components of the OS are implemented as server processes.
- Communication between user and servers using messages through kernel.
- “client-server architecture within a single computer”
- Examples: Mach, Windows NT, Chorus, L4, ...
Benefits of $\mu$-Kernels

- **Extensibility:**
  - New services can be added by adding server processes.
- **Flexibility:**
  - Services can be customized.
- **Portability:**
  - Kernel small, with well-defined interface.
- **Distributed System Support:**
  - Interface between users and services is message-based.
μ-Kernels: Performance is Problem

- Request traverses user/kernel boundary twice, same for reply.
- Solutions:
  - Move critical services back into the kernel ("make kernel bigger")
  - Make kernel "smaller"

Why are OSs so Slow?

<table>
<thead>
<tr>
<th>Hardware</th>
<th>Abbreviation</th>
<th>RISC/CISC</th>
<th>MIPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIPS M2000</td>
<td>M2000</td>
<td>RISC</td>
<td>20</td>
</tr>
<tr>
<td>DECstation 3100</td>
<td>DS3100</td>
<td>RISC</td>
<td>13</td>
</tr>
<tr>
<td>Sun-4/280</td>
<td>Sun4</td>
<td>RISC</td>
<td>9</td>
</tr>
<tr>
<td>VAX 8800</td>
<td>8800</td>
<td>CISC</td>
<td>6</td>
</tr>
<tr>
<td>Sun-3/75</td>
<td>Sun3</td>
<td>CISC</td>
<td>1.8</td>
</tr>
<tr>
<td>Microvax II</td>
<td>MVAX2</td>
<td>CISC</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Table 1: Hardware Platforms

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Time (microseconds)</th>
<th>MIPS-Relative Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>M2000 RISCos 4.0</td>
<td>18</td>
<td>0.54</td>
</tr>
<tr>
<td>DS3100 Sprite</td>
<td>26</td>
<td>0.49</td>
</tr>
<tr>
<td>DS3100 Ultrix 3.1</td>
<td>25</td>
<td>0.60</td>
</tr>
<tr>
<td>8800 Ultrix 3.0</td>
<td>28</td>
<td>1.15</td>
</tr>
<tr>
<td>Sun4 SunOS 4.0</td>
<td>32</td>
<td>0.68</td>
</tr>
<tr>
<td>Sun4 Sprite</td>
<td>32</td>
<td>0.58</td>
</tr>
<tr>
<td>Sun3 Sprite</td>
<td>92</td>
<td>1.9</td>
</tr>
<tr>
<td>Sun3 SunOS 3.5</td>
<td>108</td>
<td>1.0</td>
</tr>
<tr>
<td>MVAX2 Ultrix 3.0</td>
<td>207</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Table 2: Getpwd kernel call time
## Why are OSs so Slow? (2)

<table>
<thead>
<tr>
<th>Hardware</th>
<th>Abbreviation</th>
<th>RISC/CISC</th>
<th>MIPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIPS M2000</td>
<td>M2000</td>
<td>RISC</td>
<td>20</td>
</tr>
<tr>
<td>DECstation 3100</td>
<td>DS3100</td>
<td>RISC</td>
<td>13</td>
</tr>
<tr>
<td>Sun-4/280</td>
<td>Sun4</td>
<td>RISC</td>
<td>9</td>
</tr>
<tr>
<td>VAX 8800</td>
<td>8800</td>
<td>CISC</td>
<td>6</td>
</tr>
<tr>
<td>Sun-3/75</td>
<td>Sun3</td>
<td>CISC</td>
<td>1.8</td>
</tr>
<tr>
<td>Microvax II</td>
<td>MVAX2</td>
<td>CISC</td>
<td>0.9</td>
</tr>
</tbody>
</table>

**Table 1: Hardware Platforms**

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Time (ms)</th>
<th>MIPS-Relative Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>M2000 RISC/cis 4.0</td>
<td>0.30</td>
<td>0.71</td>
</tr>
<tr>
<td>DS3100 Ultras 3.1</td>
<td>0.34</td>
<td>0.96</td>
</tr>
<tr>
<td>DS3100 Sprite</td>
<td>0.51</td>
<td>0.65</td>
</tr>
<tr>
<td>8800 Ultras 3.0</td>
<td>0.70</td>
<td>1.0</td>
</tr>
<tr>
<td>Sun4 SunOS 4.0</td>
<td>1.02</td>
<td>0.47</td>
</tr>
<tr>
<td>Sun4 Sprite</td>
<td>1.17</td>
<td>0.41</td>
</tr>
<tr>
<td>Sun3 SunOS 3.5</td>
<td>2.36</td>
<td>1.0</td>
</tr>
<tr>
<td>Sun3 Sprite</td>
<td>2.41</td>
<td>1.0</td>
</tr>
<tr>
<td>MVAX2 Ultras 3.0</td>
<td>3.66</td>
<td>1.3</td>
</tr>
</tbody>
</table>

**Table 2: C (slow) echo one byte between processes using pipes.**