ADRIAN PERRIG & TORSTEN HOEFLER

Networks and Operating Systems (252-0062-00)
Chapter 1: Introduction to Operating Systems

If Operating Systems were Airways (~year 2000)

UNIX Airways Everyone brings one piece of the plane along when they come to the airport. They all go out on the runway and put the plane together piece by piece, arguing non-stop about what kind of plane they are supposed to be building.

Air DOS Everybody pushes the airplane until it glides, then they jump on and let the plane coast until it hits the ground again. Then they push again, jump on again, and so on ...

Mac Airlines All the stewards, captains, baggage handlers, and ticket agents look and act exactly the same. Every time you ask questions about details, you are gently but firmly told that you don't need to know, don't want to know, and everything will be done for you without your ever having to know, so just shut up.

Windows Air The terminal is pretty and colorful, with friendly stewards, easy baggage check and boarding, and a smooth take-off. After about 10 minutes in the air, the plane explodes with no warning whatsoever.

Windows NT Air Just like Windows Air, but costs more, uses much bigger planes, and takes out all the other aircraft within a 40-mile radius when it explodes.

Linux Air Disgruntled employees of all the other OS airlines decide to start their own airline. They build the planes, ticket counters, and pave the runways themselves. They charge a small fee to cover the cost of printing the ticket, but you can also download and print the ticket yourself. When you board the plane, you are given a seat, four bolts, a wrench and a copy of the Seat-HOWTO.html. Once settled, the fully adjustable seat is very comfortable, the plane leaves and arrives on time without a single problem, the in-flight meal is wonderful. You try to tell customers of the other airlines about the great trip, but all they can say is, "You had to do what with the seat?“ *(Author unknown)*

Administrivia

- **Two parts:**
  - Operating Systems – Torsten Hoefler
  - Networks – Adrian Perrig
- **Lecture:**
  - Mon 1-3pm, HG G 3
  - Fri 10am-noon, HG E 7
- **Practice sessions**
  - See sign-in sheets!
  - No exercises this (first) week!
- **Go to one of these sessions!**
  - And participate!
  - Well, and **participate in the lecture** as well ☺
- **The official language is English**
  - I do speak German, you may ask in German (I'll repeat in English)

More Administrivia

- **Course webpage (the authoritative information source)**
  - http://spcl.inf.ethz.ch/Teaching/2017-osnet/
  - All slides will be there before the lecture (so you can take notes)
- **Exercises are:**
  - *Theoretical:* Analysis of performance properties
  - *Practical:* Trying out stuff + Programming exercises
- **We assume you know both C and Java.**
  - Exercises start next week!
- **There is a mailing list for questions to the TAs**
  - You are not subscribed but can sign up at (if you want)
  - https://spcl.inf.ethz.ch/cgi-bin/mailman/listinfo/2017-osnet
- **Please register during the break**
  - put your name into lists at front desk of lecture hall

    Watch for resource conflicts!!

Exam

- **(No mid-term.)**
- **Final exam: tbd (in exam session)**
- **Material:**
  - Covered in the lectures, and/or
  - Learned during the lab exercises
- **We will not follow the books closely.**
  - All pieces will be in books though
- ***Optional* extra readings may appear on the web**

- **Review lectures (recorded 2013, pretty much unchanged)**
  http://www.video.ethz.ch/lectures/d-infk/2013/spring/252-0062-00L.html

Course Outline

20.02.: OS Introduction
24.02.: Processes
27.02.: Scheduling
03.03.: Synchronization
06.03.: Memory Management
10.03.: Demand Paging
13.03.: File System Abstractions
17.03.: File System Implementations
20.03.: I/O Subsystem I
24.03.: I/O Subsystem II
27.03.: Virtual Machine Monitors
31.03.: Reliable Storage, Specials

(preliminary – may change)
03.04.: Network Intro / OSI Model
07.04.: Physical Layer
10.04.: Data Link Layer I
28.04.: Data Link Layer II
05.05.: Network Layer I
08.05.: Network Layer II
12.05.: Network Layer III
15.05.: Transport Layer
19.05.: Congestion Control
22.05.: Congestion Control
26.05.: Application Layer
29.05.: TBD
02.06.: TBD

Birds-eye perspective

- **Networks**
  - bridge space

- **Databases**
  - bridge time

- **Networks, Operating Systems, Databases**
  - they all manage resources
  - OS, DB: all resources (storage, computation, communication)
  - Networks: focus on communication

- **OS vs. government**
  - “Like a government, an operating system performs no useful function by itself”

~200 sensors

V

ts elements

Why c

Black Hat USA 2015: The full story of how that Jeep was hacked

August 6, 2015 Alex Drozhzhin Featured Post, News, Security, Threats No comments

Recently we wrote about the now-famous hack of a Jeep Cherokee. At Black Hat USA 2015, a large security conference, researchers Charlie Miller and Chris Valasek finally explained in detail, how exactly that hack happened.

And it has been broken ☹

http://illmatics.com/Remote%20Car%20Hacking.pdf

Today: We start on Operating Systems!

- **Introduction: Why?**
- **Roles of the OS**
  - Referee
  - Illusionist
  - Glue
- **Structure of an OS**

Goals

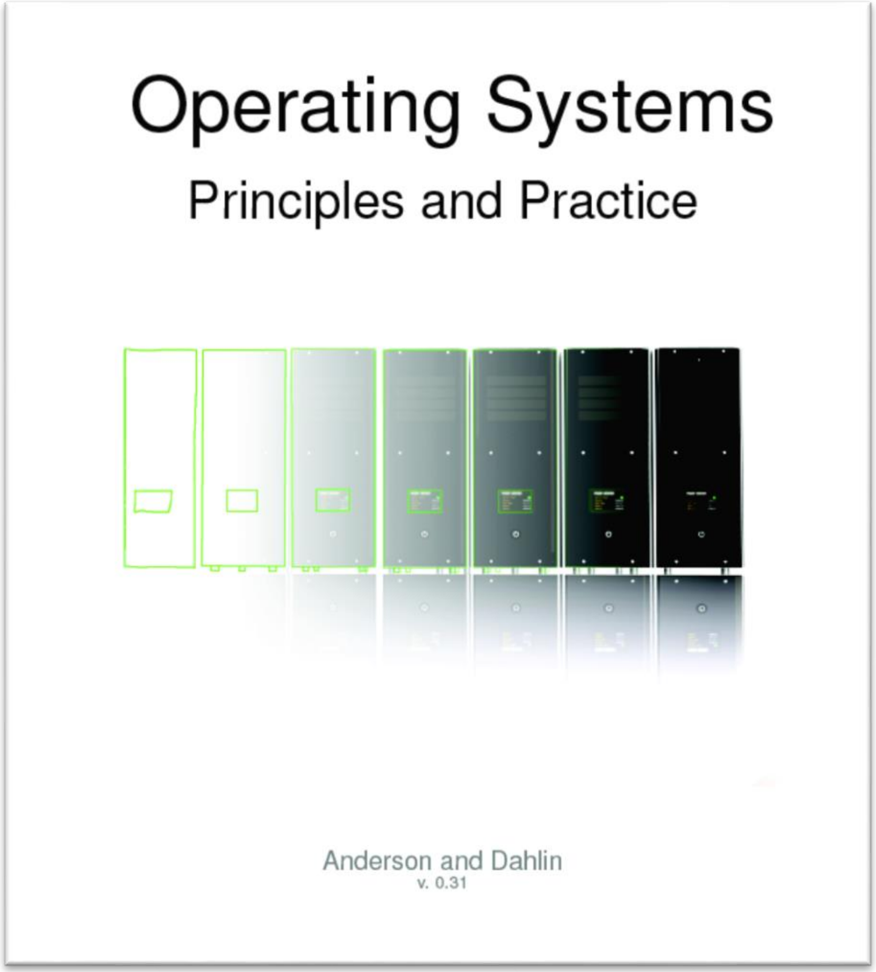
- **Demystify operating systems themselves**
  - What is an OS? What does it do?
  - What is its structure?
  - How do the OS and applications relate to each other?
  - What services does the OS provide?

- **Quintessential "systems" problem**
  - Non-idealizable / non-reducible
  - Scaling, emergent properties
  - Concurrency and asynchrony

The Book

- **On the web:**

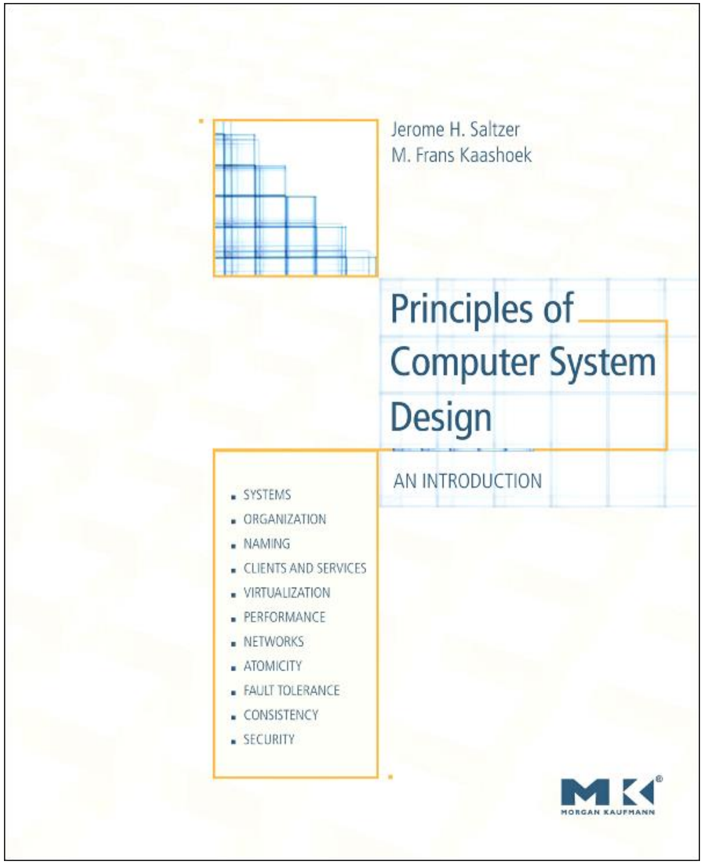
http://ospp.cs.washington.edu/



Also worth a look

- **Jerome H. Saltzer and M. Frans Kaashoek:**
  "Principles of Computer System Design"

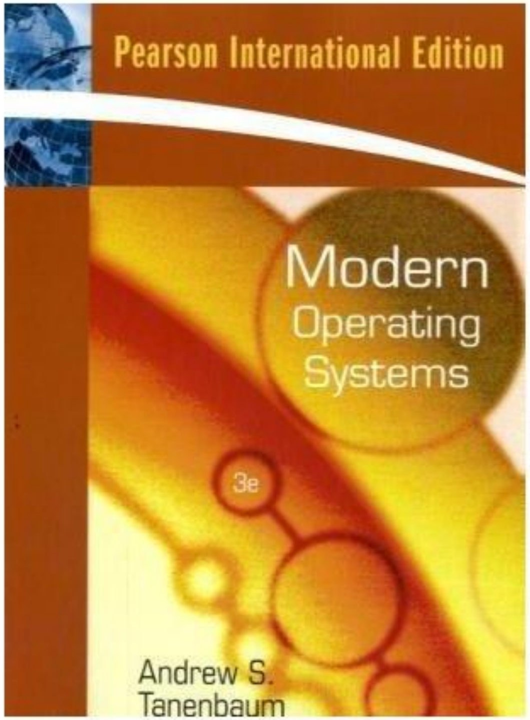
- **Focus on principles, with illustrative examples**



Also worth a look

- **Andrew S. Tanebaum:**
  "Modern Operating Systems"

- ***Must* be at least 3rd Edition!**
- **Very broad – lots of references to recent (2006) research.**



Introduction to Operating Systems

Why learn about Operating Systems?

- **One of the most complex topics in Computer Science!**
  - Very few simplifying assumptions
  - Dealing with the real world
  - Intersection of many areas
- **Mainstream OSes are large:**
  - Windows 7 ~ 40-50 million lines of code
    Average modern high-end car: 100 million [1]
  - Linux rapidly catching up in complexity (~15 million LOC)
- **Most other software systems are a subset**
  - Games, browsers, databases, servers, cloud, etc.

[1]: http://www.informationisbeautiful.net/visualizations/million-lines-of-code/

There are lots of operating systems concepts…

- System calls
- Concurrency and asynchrony
- Processes and threads
- Security, authorization, protection
- Memory, virtual memory, and paging
- Files and file systems, data management
- I/O: Devices, Interrupts, DMA
- Network interfaces and protocol stacks

There are lots of operating systems…

iPhone OS 4

symbian

QNX
QNX SOFTWARE SYSTEMS

VxWorks

android

ETH Oberon

Mac OS X

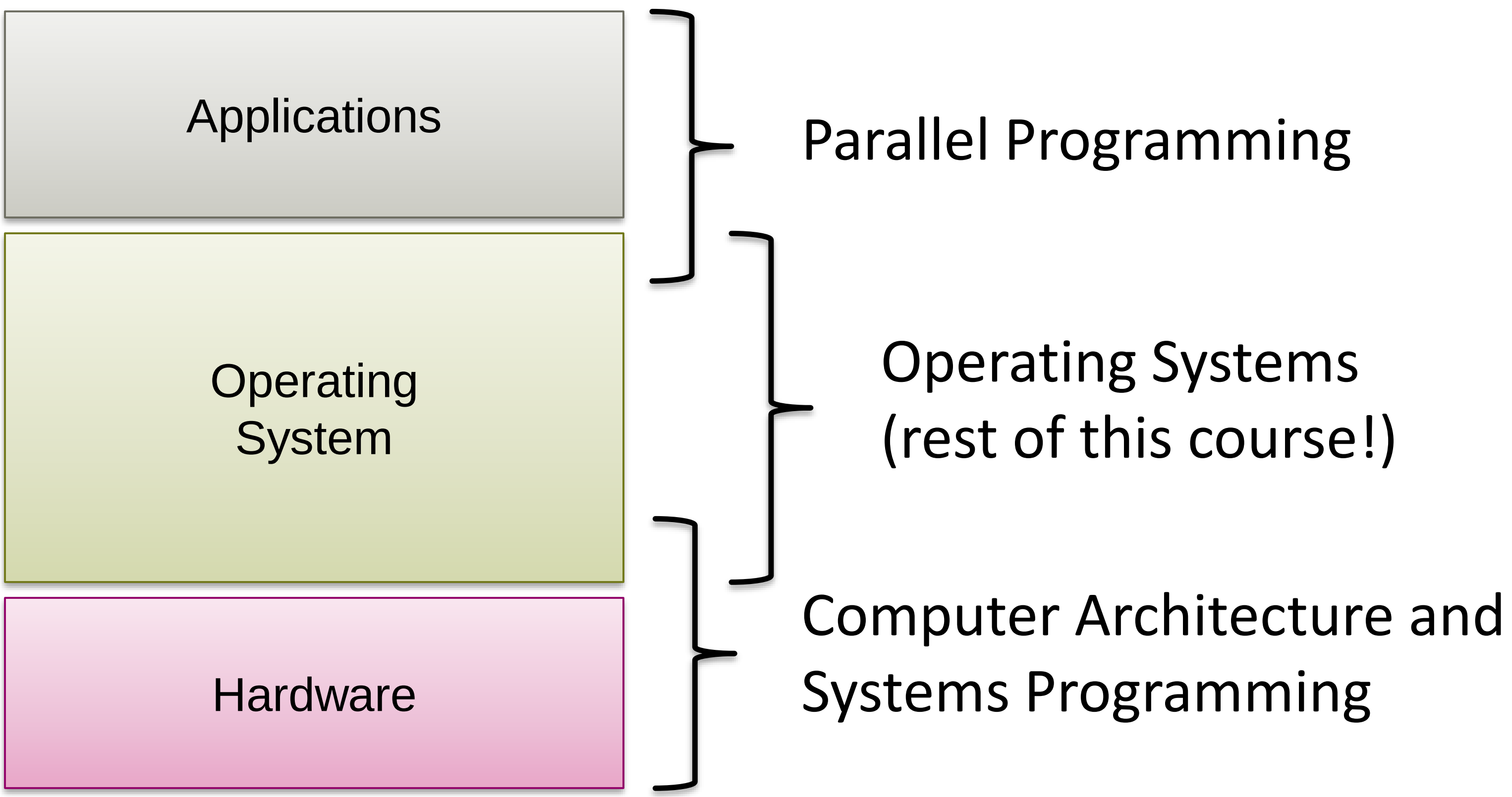
MeeGo

Open Kernel Labs™
Be open. Be safe.

Goals: what makes a good OS?

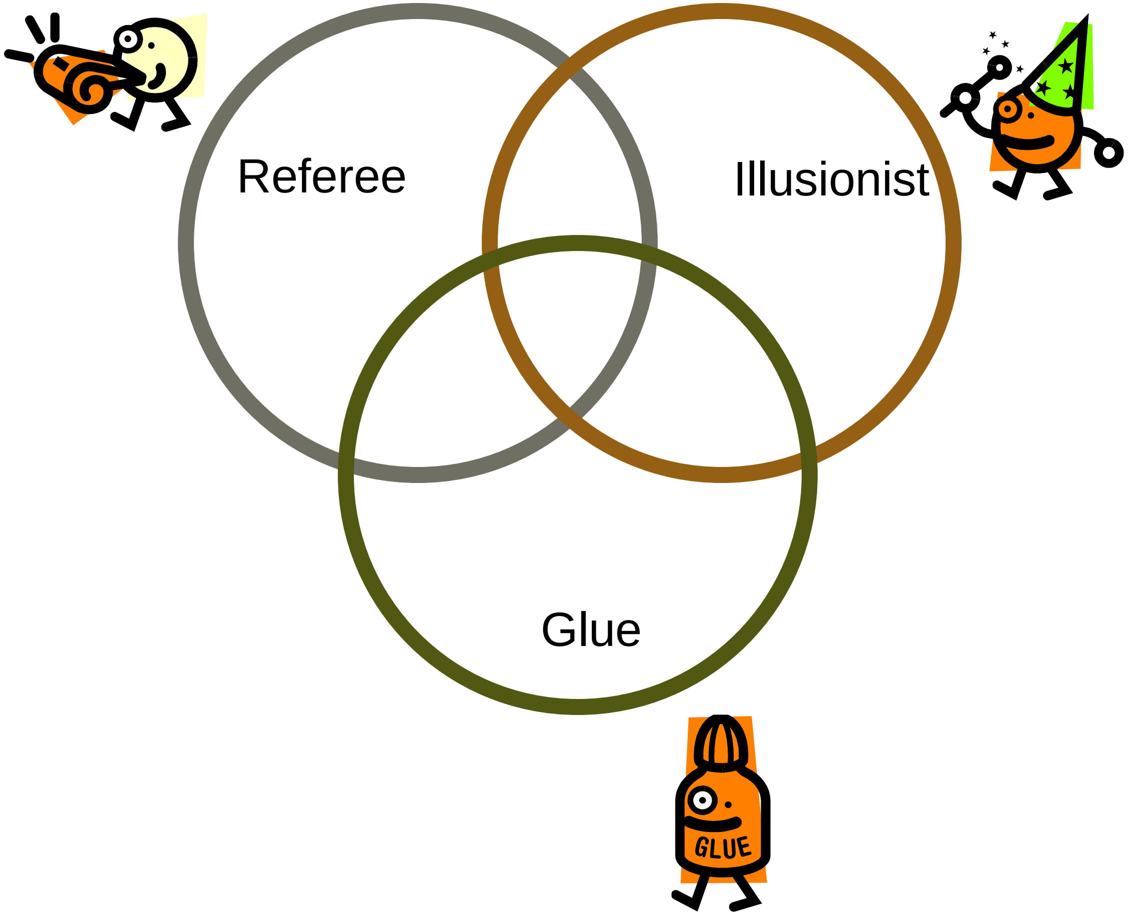
- **Reliability: does it keep working?**
  - And availability
- **Security: can it be compromised?**
  - And isolation: is it fair?
- **Portability: how easily can it be retargeted?**
- **Performance: how fast/cheap/hungry is it?**
- **Adoption: will people use it?**
- **…**

Operating Systems



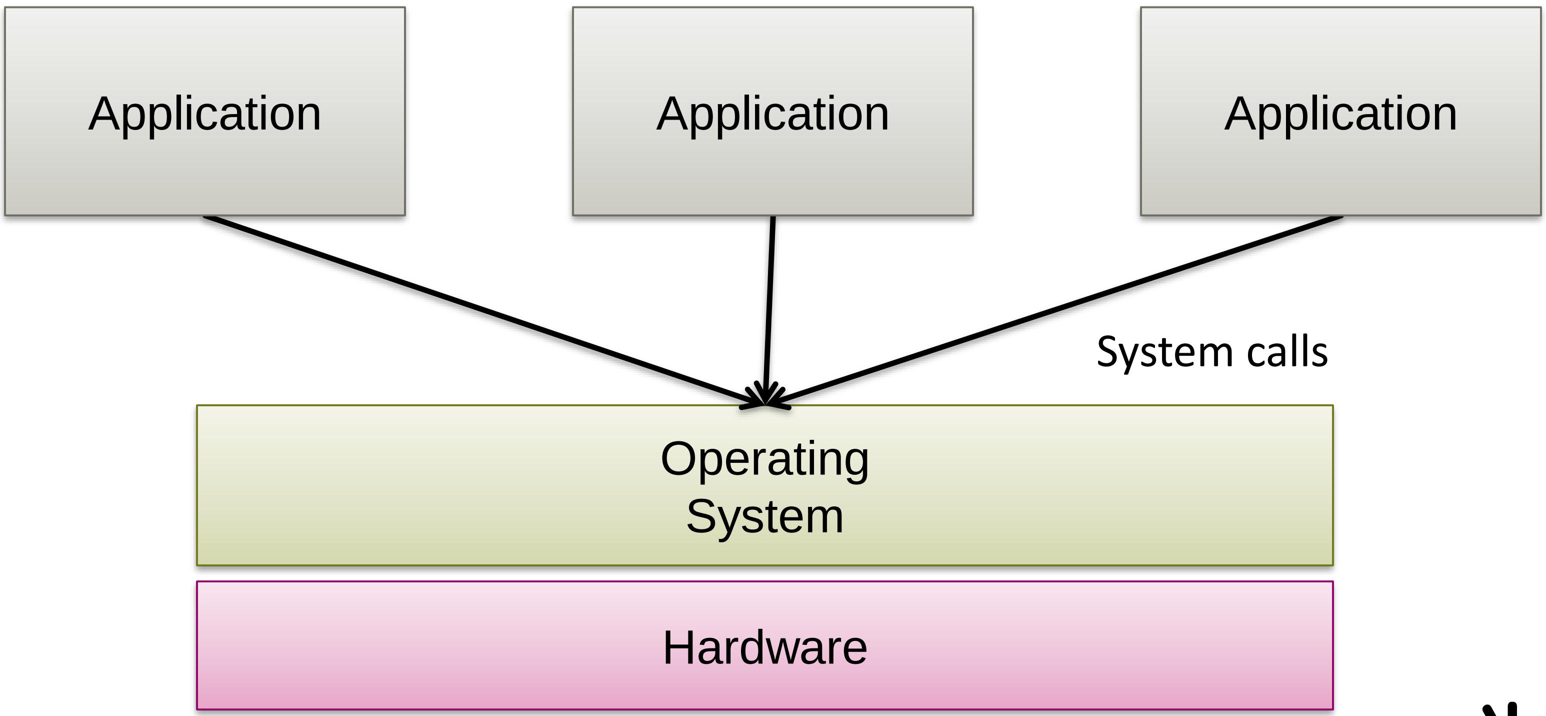
Operating System Roles

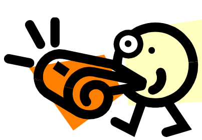
OS roles



The Referee:

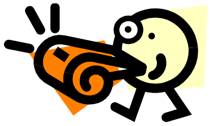
Resource Manager





The OS as Referee

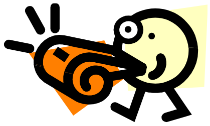
- **Sharing:**
  - Multiplex hardware among applications
    CPU, memory, devices
  - Applications shouldn't need to be aware of each other
- **Protection:**
  - Ensure one application cannot r/w another's data
    In memory, on disk, over network
  - Ensure one application cannot use another's resources
    CPU, storage space, bandwidth, …
- **Communication:**
  - Protected applications must still communicate



Resource management goals

- **Fairness:**
  - No starvation, every application makes progress
- **Efficiency:**
  - Best use of complete machine resources
  - Minimize e.g. power consumpti
- **Predictability:**
  - Guarantee real-time perform

All in mutual contradiction ☹

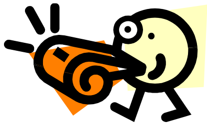


Example: Threads

- **Threads are virtual CPUs**
  - Physical resource: CPUs
  - Virtual resource: Threads

  - Mechanism: pre-emption, timeslicing, context switching, scheduling

- **More on this later in the course…**



The Illusionist

Virtualization:

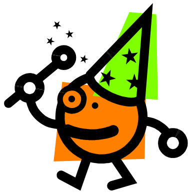
- **OS creates illusion of a "real" resource**
  - Processor, storage, network, links, …
- **The virtual resource looks *a bit* like a physical resource**
- **However, is frequently quite different…**
  - Simpler, larger, better, …

How?

1. **Multiplexing**
   – Divide resources up among clients

2. **Emulation**
   – Create the illusion of a resource using software

3. **Aggregation**
   – Join multiple resources together to create a new one



Why?

1. **Sharing**
   – Enable multiple clients of a single resource

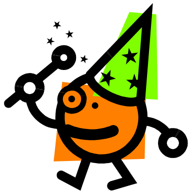
2. **Sandboxing**
   – Prevent a client from accessing other resources

3. **Decoupling**
   – Avoid tying a client to a *particular* instance of a resource

4. **Abstraction**
   – Make a resource easier to use

Example: Virtual memory

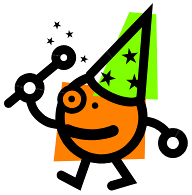
- **Easier memory to manage**
  - Physical resource: RAM
  - Virtual resource: virtual memory
  - Method: multiplexing
  - Mechanism: virtual address translation



Example: Paged virtual memory

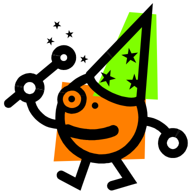
- **More memory than you really have**
  - Physical resource: RAM and disk
  - Virtual resource: paged virtual memory
  - Method: multiplexing and emulation
  - Mechanism: virtual memory + paging to/from disk

- **Much more on this later in the course…**



Example: Virtual machines

- **Quite popular topic commercially right now:**
  - Xen, VMware, HyperV, kvm, etc.
- **Many uses:**
  - Run one OS on another
  - Consolidate servers
  - Migrate running machines around datacenter
  - Run hundreds of “honeypot” machines
  - Deterministic replay of whole machines
  - Etc.



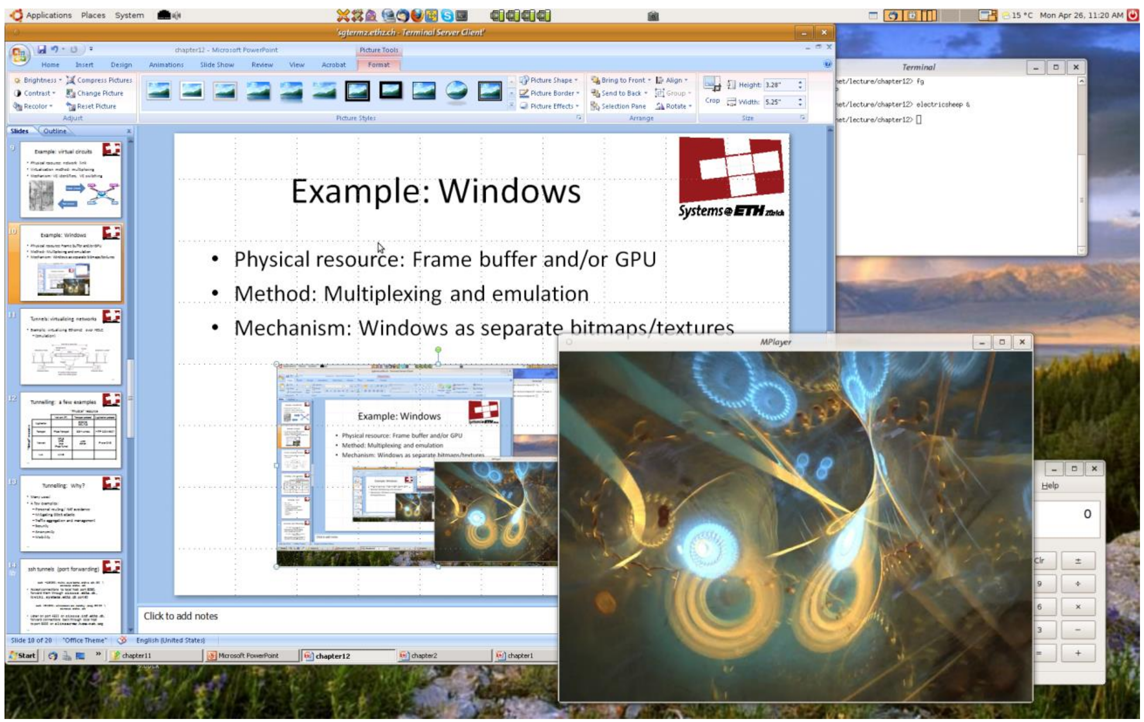
Example: Files (or databases!)

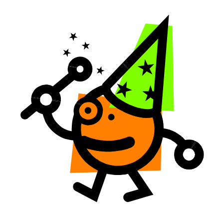
- **Virtual resource: persistent memory**
- **Physical resource: disk**
- **Method: multiplexing, emulation**
- **Mechanism: block allocation, metadata**

- **Again, more later…**

Example: Windows (not the Microsoft OS)

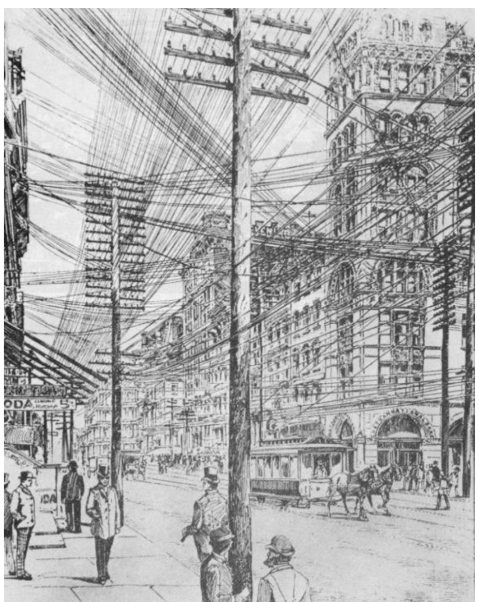
- Physical resource: Frame buffer and/or GPU
- Method: Multiplexing and emulation
- Mechanism: Windows as separate bitmaps/textures

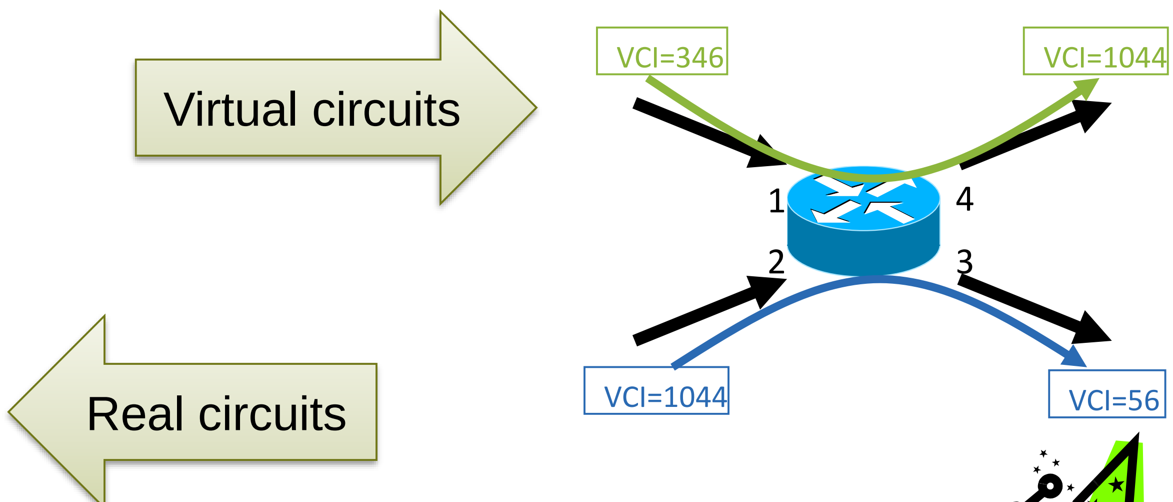


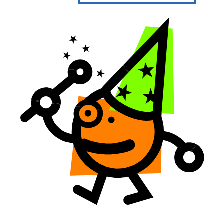


Example: Virtual circuits

- **Physical resource: network link**
- **Virtualization method: multiplexing**
- **Mechanism: VC identifiers, VC switching**

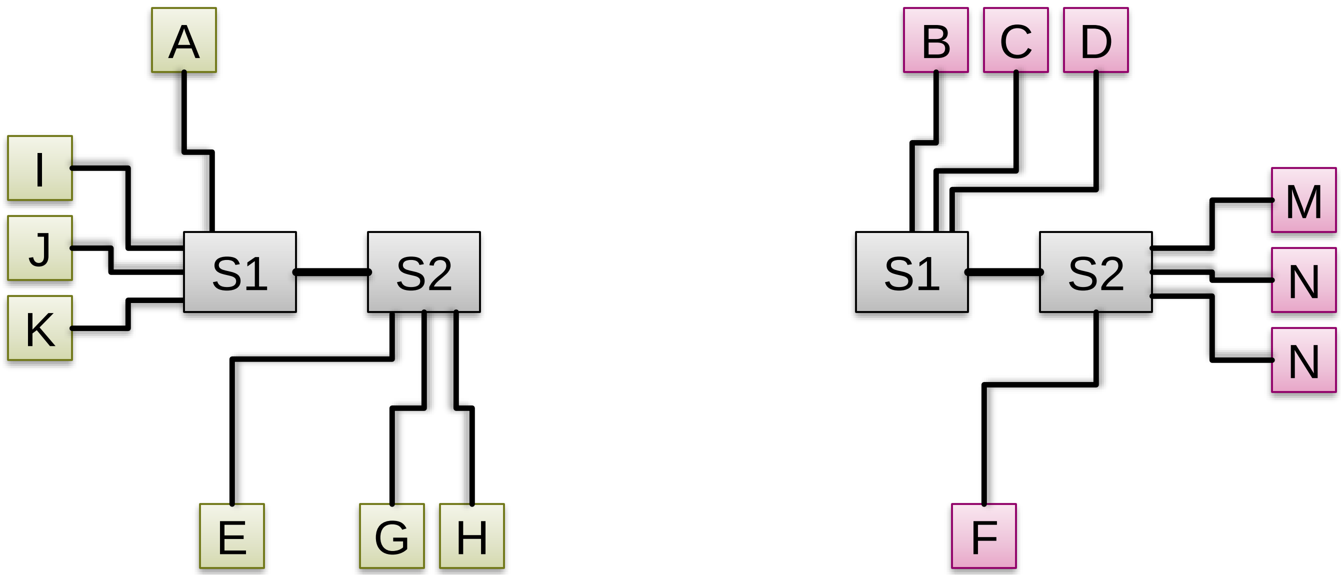




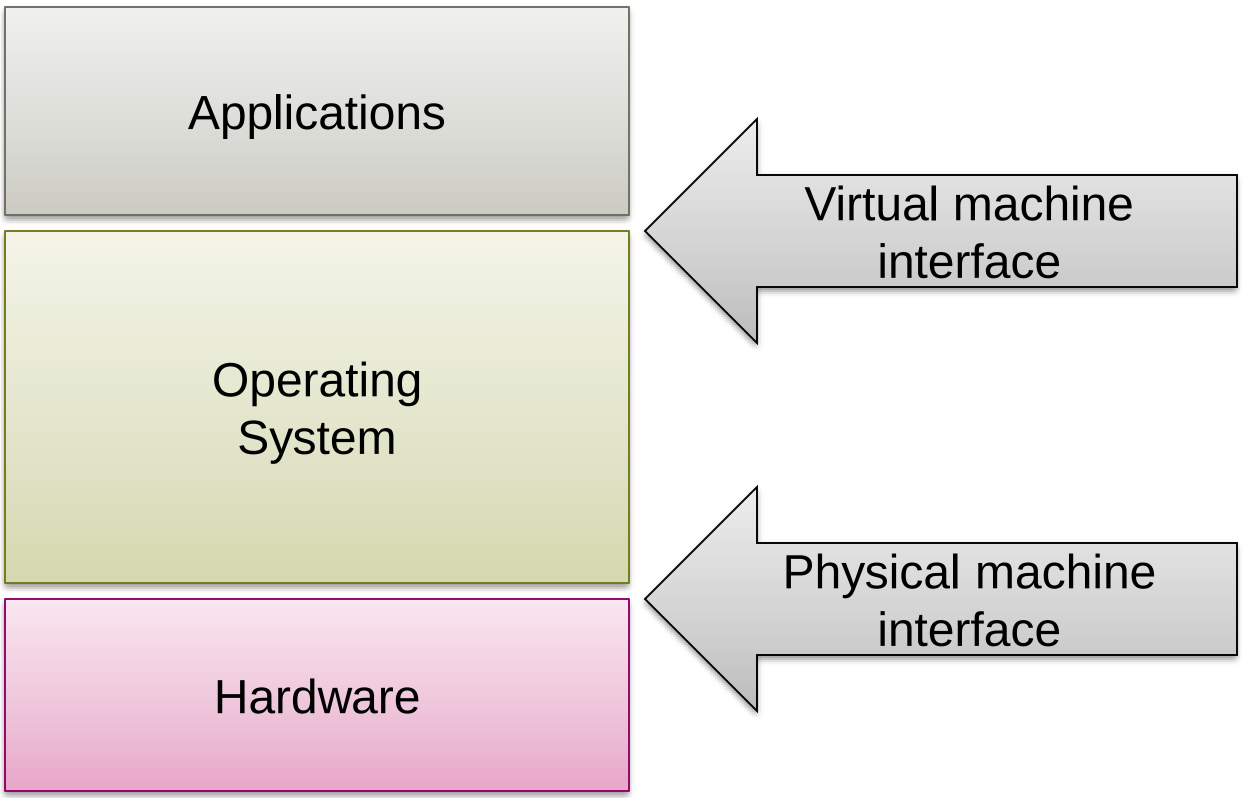


Example: VLANs

- **Methods: multiplexing**
- **Mechanisms: port assignment, tags**



Glue: the OS as Abstract Machine



The OS as Glue

- **Provides high-level abstractions**
  - Easier to program to
  - Shared functionality for all applications
  - Ties together disparate functions and services

- **Extends hardware with added functionality**
  - Direct programming of hardware unnecessary

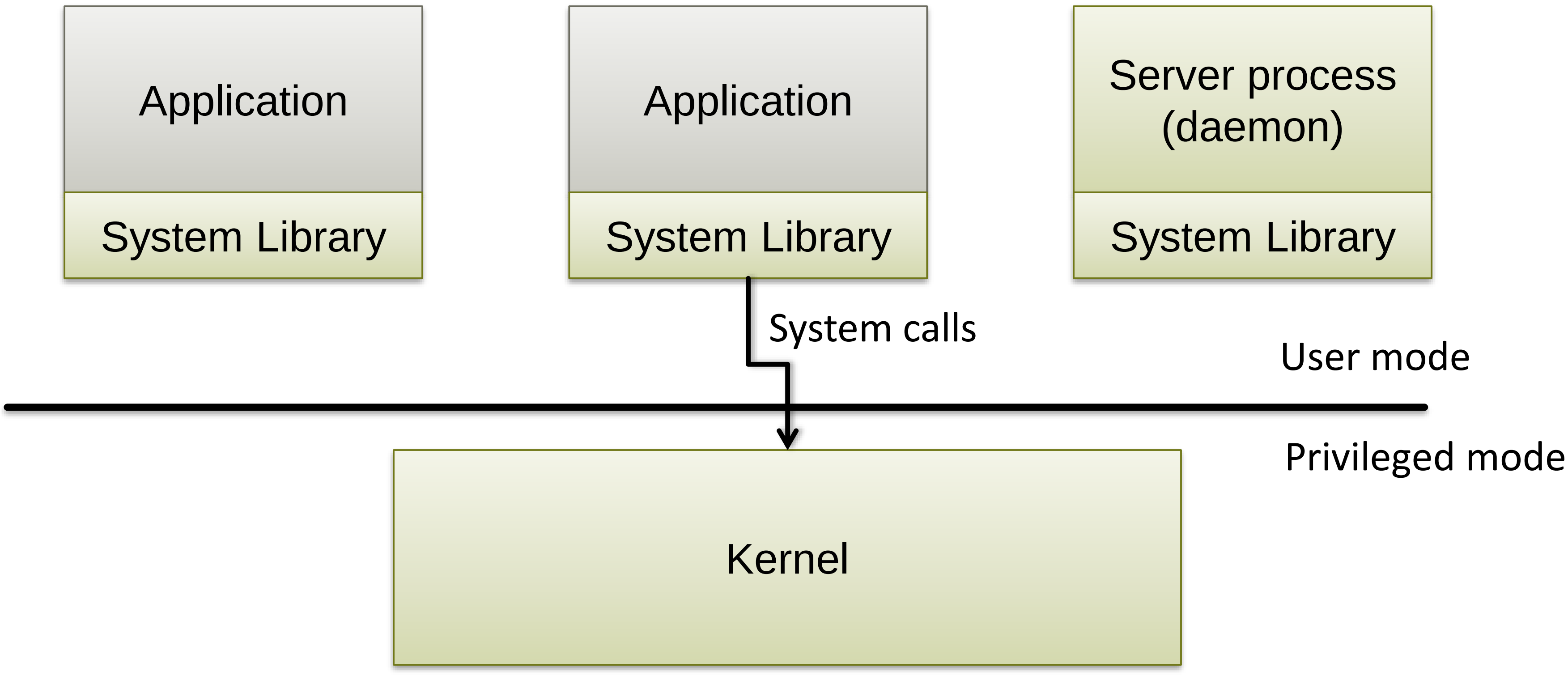
- **Hides details of hardware**
  - Applications decoupled from particular devices

Services provided by an OS

- **Program execution**
  - Load program, execute on 1 or more processors
- **Access to I/O devices**
  - Disk, network, keyboard, screen,…
- **Protection and access control**
  - For files, connections, etc.
- **Error detection and reporting**
  - Trap handling, etc.
- **Accounting and auditing**
  - Statistics, billing, forensics, etc.
- **…**

Operating System Structure

General OS structure

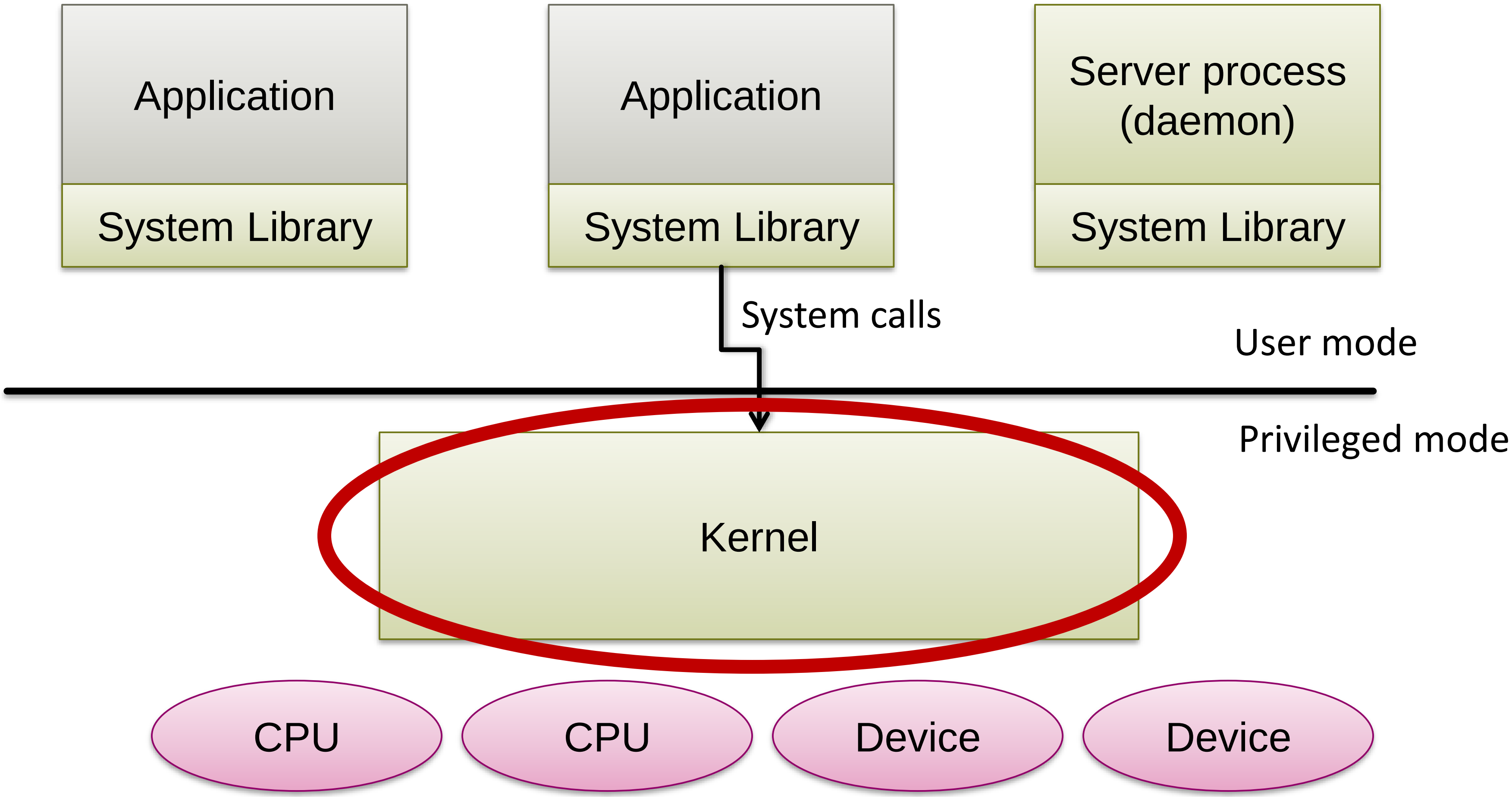


Privileged Mode and User Mode

- **As we saw in Computer Architecture, most CPUs have a "privileged mode":**
  - ia32 protection rings
  - VAX kernel mode
  - Etc.

- **Most Operating Systems use this for protection**
  - In particular, protecting the OS from applications!

General OS structure



Kernel

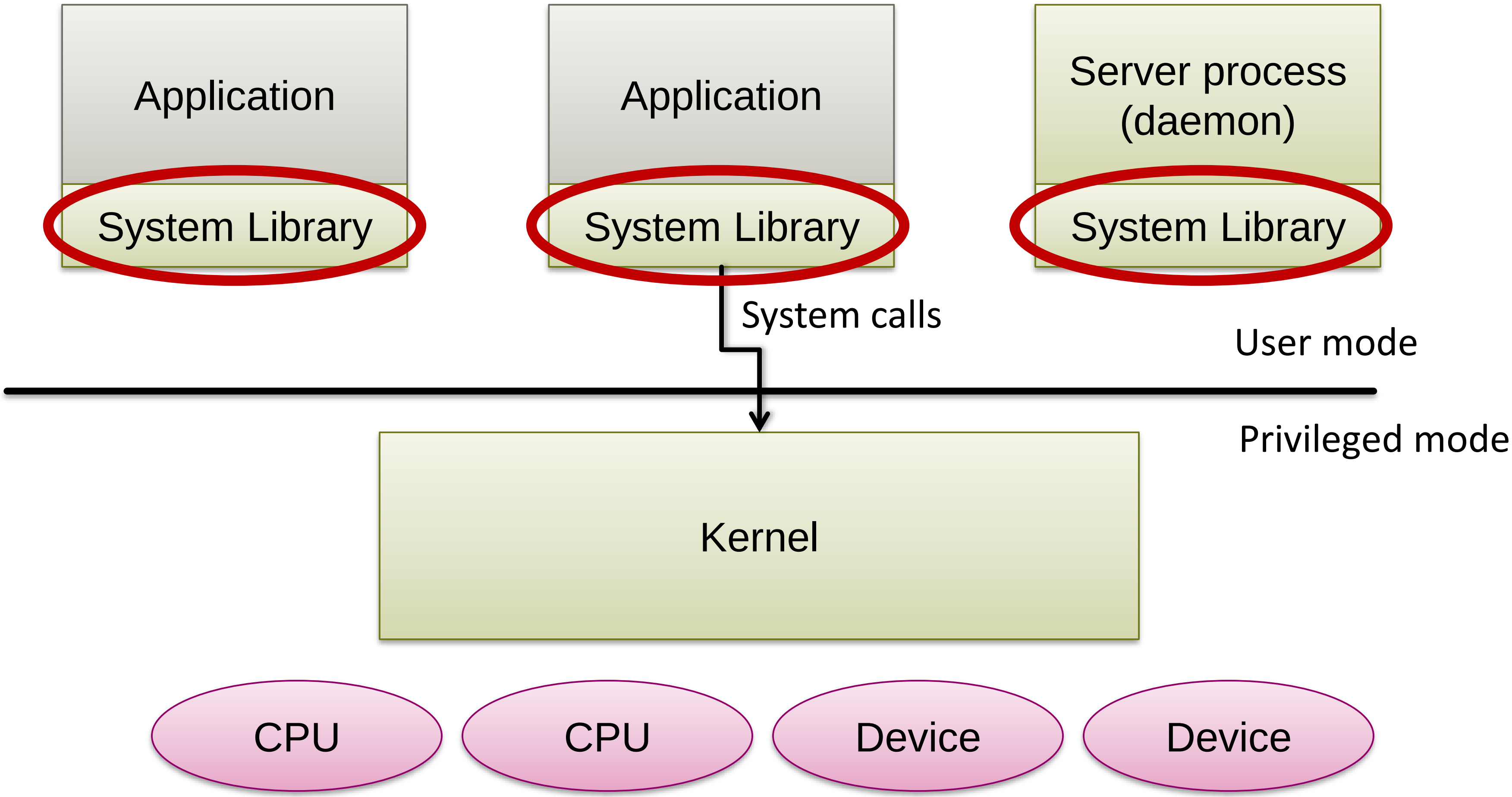
- **That part of the OS which runs in privileged mode**
  - Large part of Unix and Windows (except libraries)
  - Small part of L4, Barrelfish, etc. (microkernels)
  - Does not exist in some embedded systems

- **Also known as:**
  - Nucleus, nub, supervisor, …

The kernel is a program!

- **Kernel is just a (special) computer program.**
- **Typically an event-driven server.**
- **Responds to multiple entry points:**
  - System calls
  - Hardware interrupts
  - Program traps
- **May also include internal threads.**

General OS structure

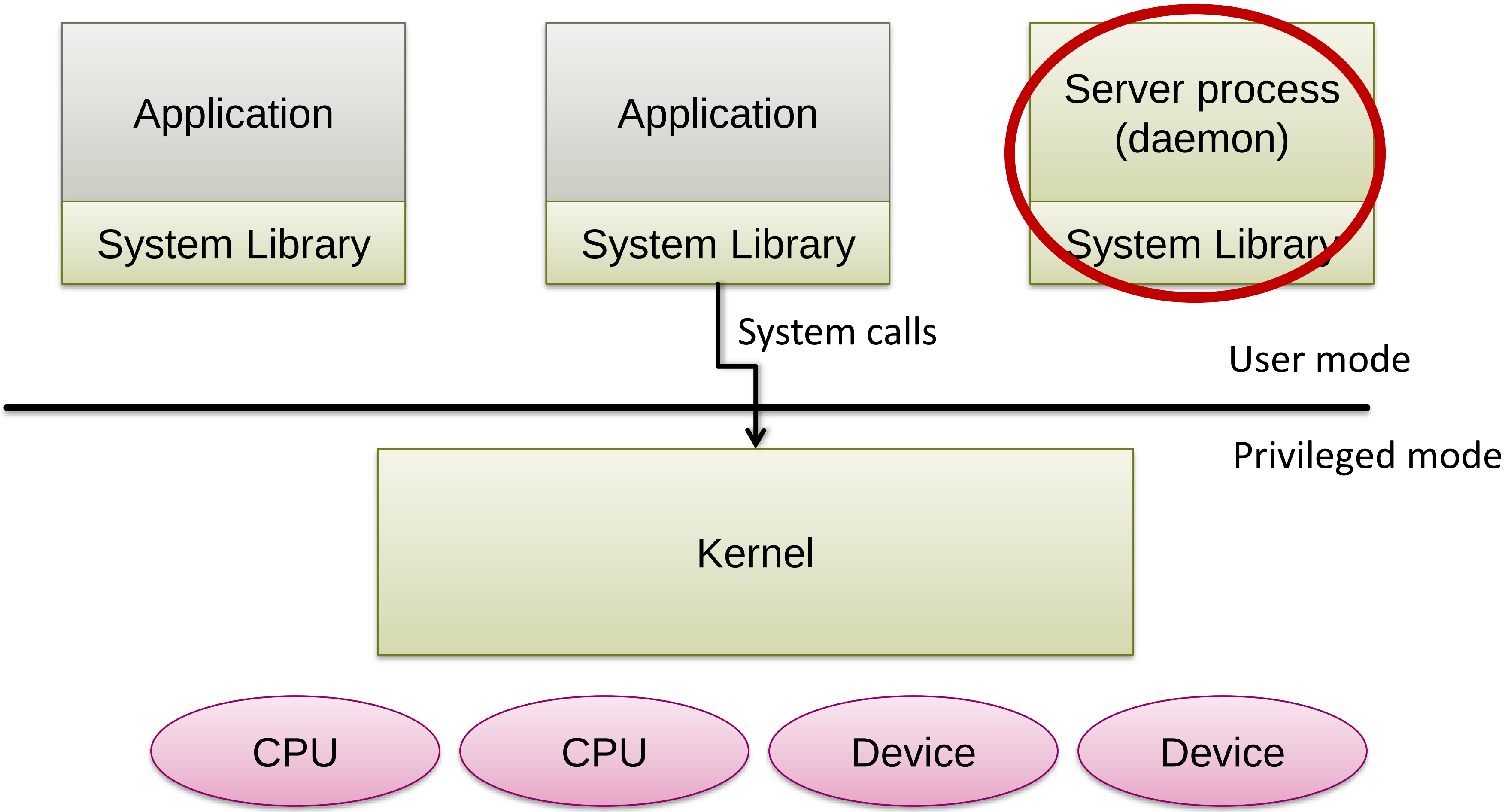


System Libraries

- **Convenience functions**
  - strcmp(), etc.
  - Common functionality

- **System call wrappers**
  - Create and execute system calls from high-level languages
    printf(), etc.
  - See 'man syscalls' on Linux

General OS structure



Daemons

- **Processes which are part of the OS**
  - Microkernels: most of the OS
  - Linux: increasingly large quantity

- **Advantages:**
  - Modularity, fault tolerance
  - Easier to *schedule*…

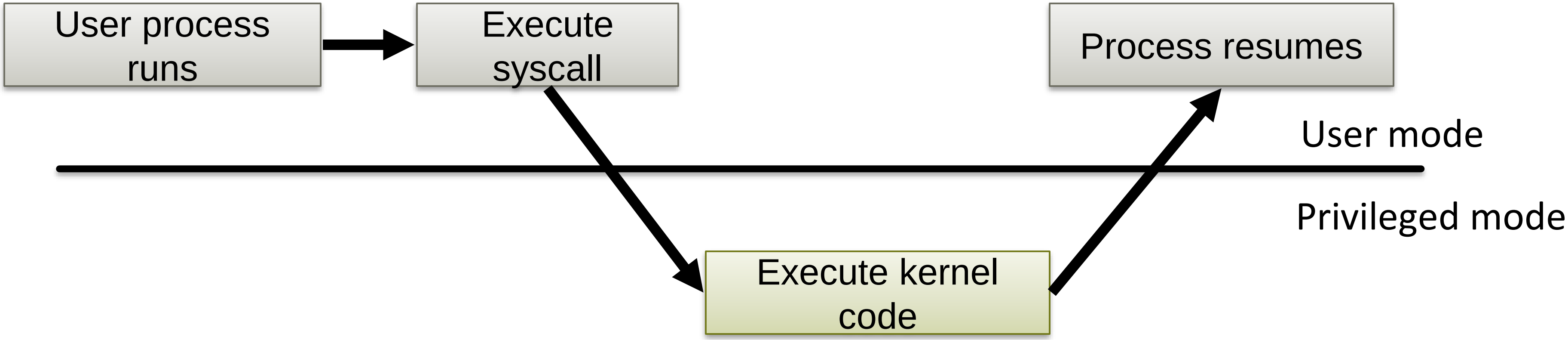
Entering and exiting the kernel

When is the kernel entered?

- Startup
- Interrupt: caused by “something else”
- Exception: caused by user program (also called “trap”)
- System calls

Recall: System Calls

- **RPC to the kernel**
- **Kernel is a series of syscall event handlers**
- **Mechanism is hardware-dependent**



System call arguments

Syscalls are *the* way a program requests services from the kernel.

Implementation varies:

- **Passed in processor registers**
- **Stored in memory (address in register)**
- **Pushed on the stack**

- **System library (libc) wraps as a C function**
- **Kernel code wraps handler as C call**

When is the kernel exited?

- **Creating a new process**
  - Including startup
- **Resuming a process after a trap**
  - Exception, interrupt or system call
- **User-level upcall**
  - Much like an interrupt, but to user-level
- ***Switching to another process***