CS 423
Operating System Design: Far too much information about interrupts

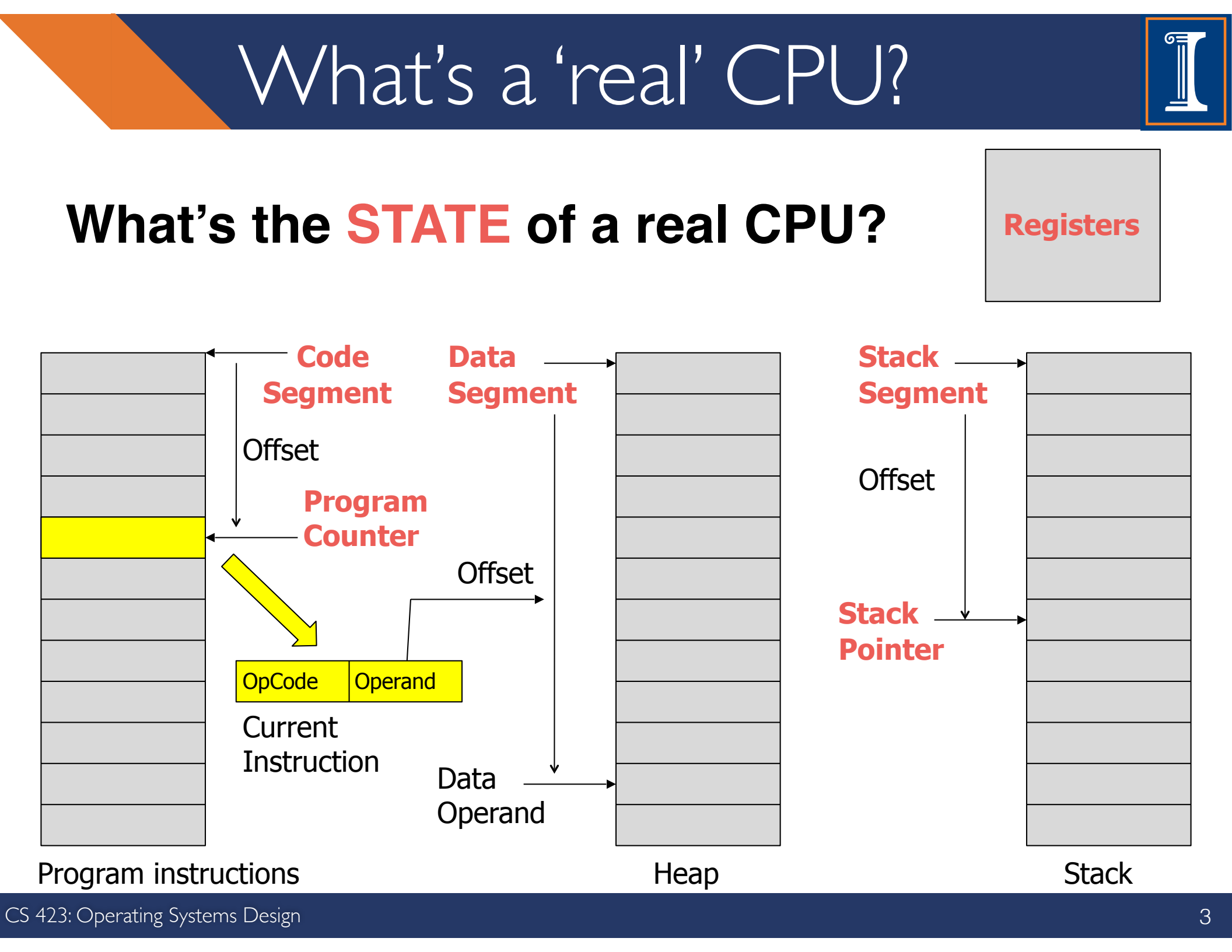
Professor Adam Bates
Fall 2018

Goals for Today

- Learning Objectives:
  - Understand the role and types of of Interrupts
- Announcements:
  - C4 Week 2 Summaries due today! **Week 3 is out**.
  - MP0 is available on Compass! **Due Jan 29**
  - HW0 is available on Compass! **Due Jan 29**

Reminder: Please put away devices at the start of class

What's a 'real' CPU?



What's the STATE of a real CPU?

Registers

Code Segment

Offset

Program Counter

OpCode | Operand

Current Instruction

Data Segment

Offset

Data Operand

Stack Segment

Offset

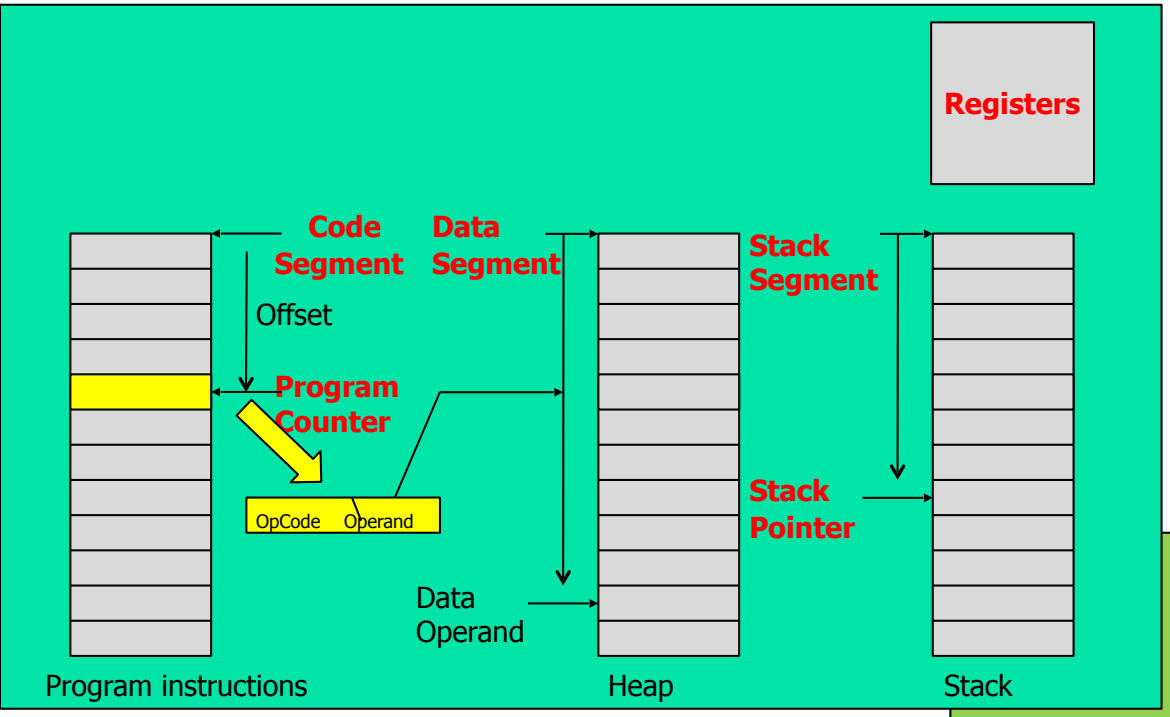
Stack Pointer

Program instructions

Heap

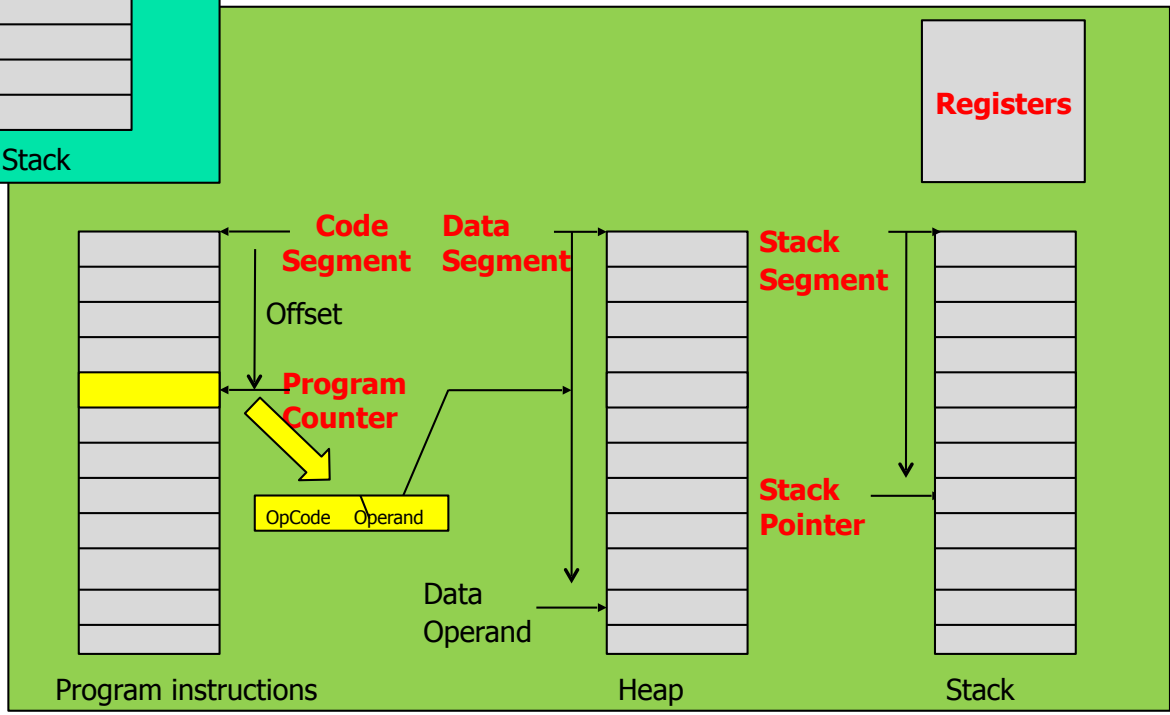
Stack

The Context Switch



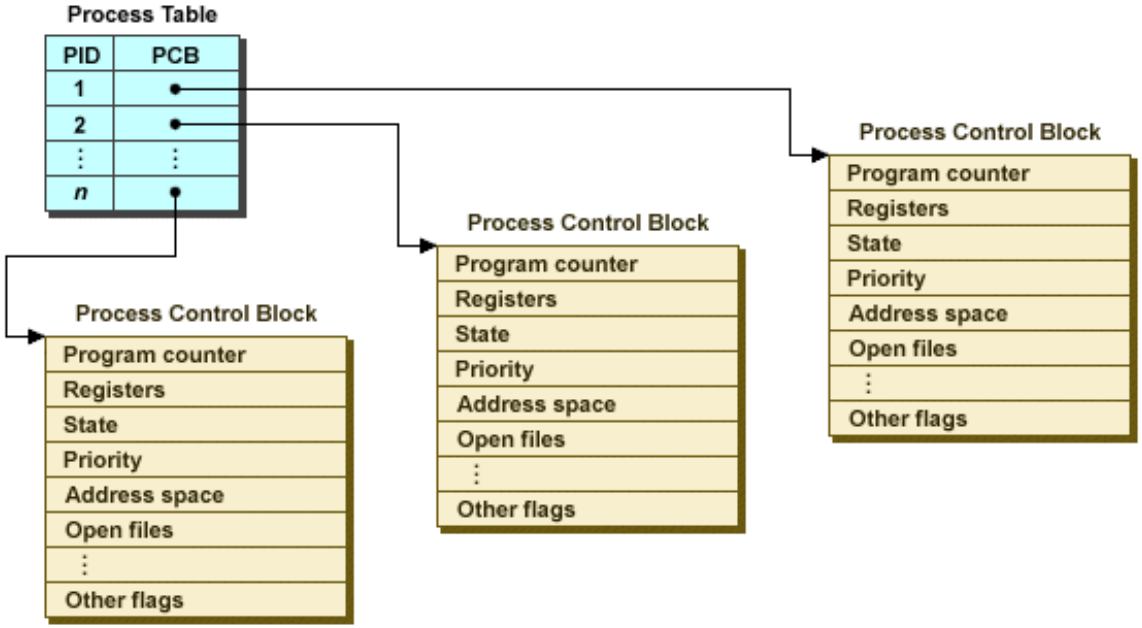
Load State (Context)

Save State (Context)

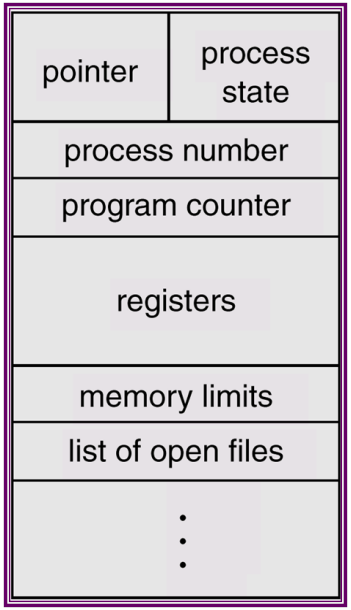


Process Control Block

The state for processes that are not running on the CPU are maintained in the Process Control Block (PCB) data structure



Updated during context switch

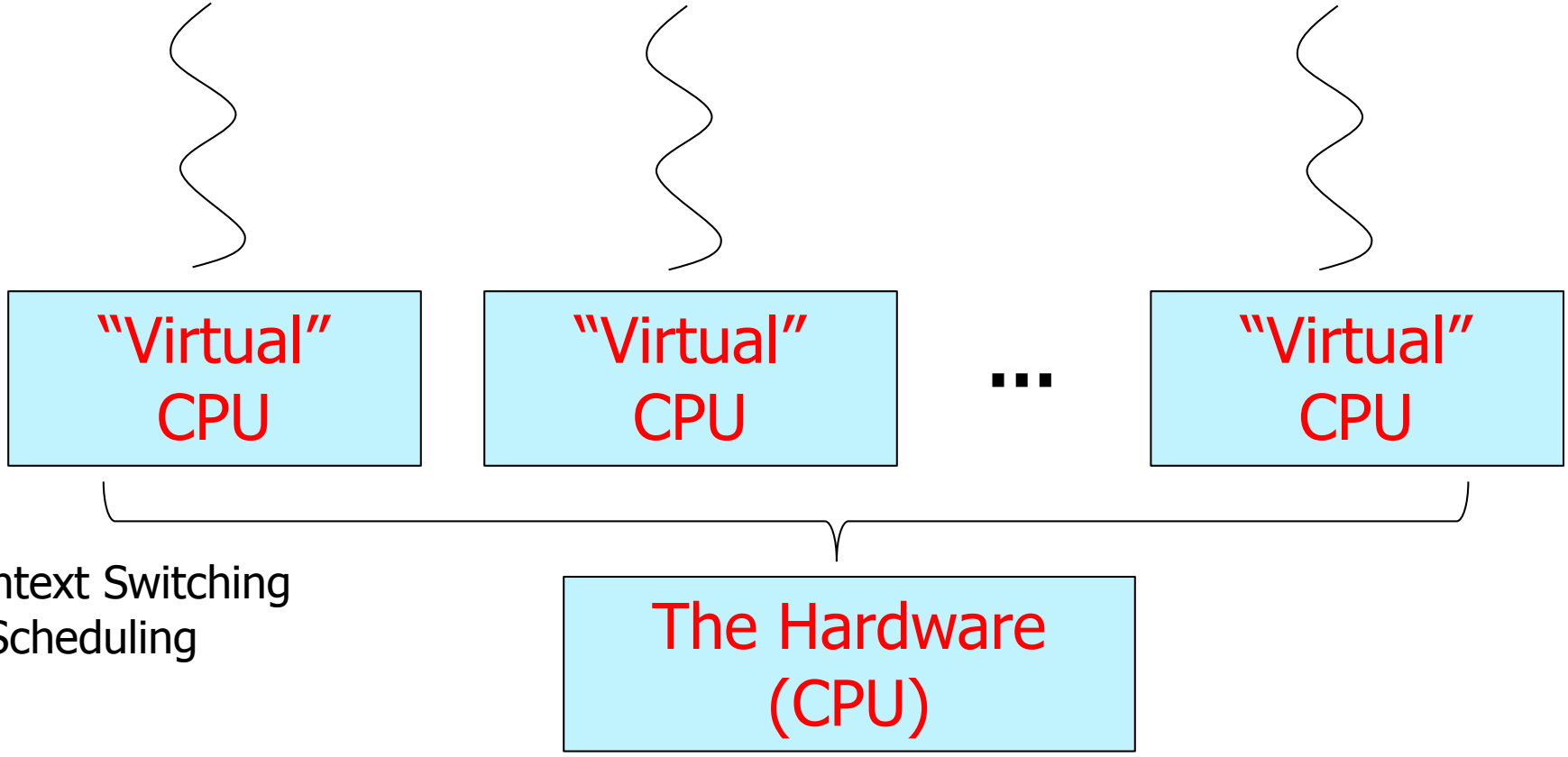


An alternate PCB diagram

Where We Are:

Last class, we discussed how context switches allow a single CPU to handle multiple tasks:

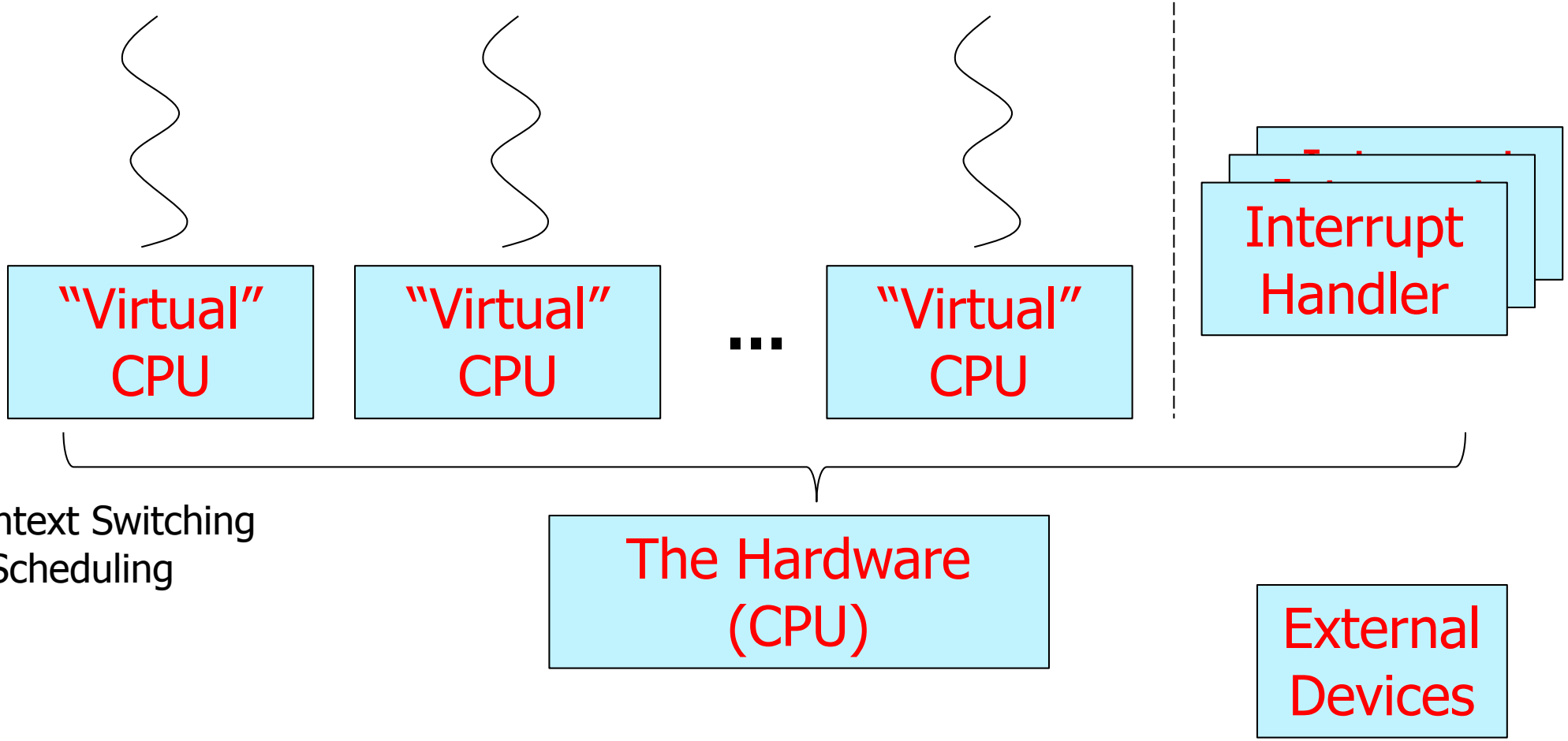
What's missing from this picture?



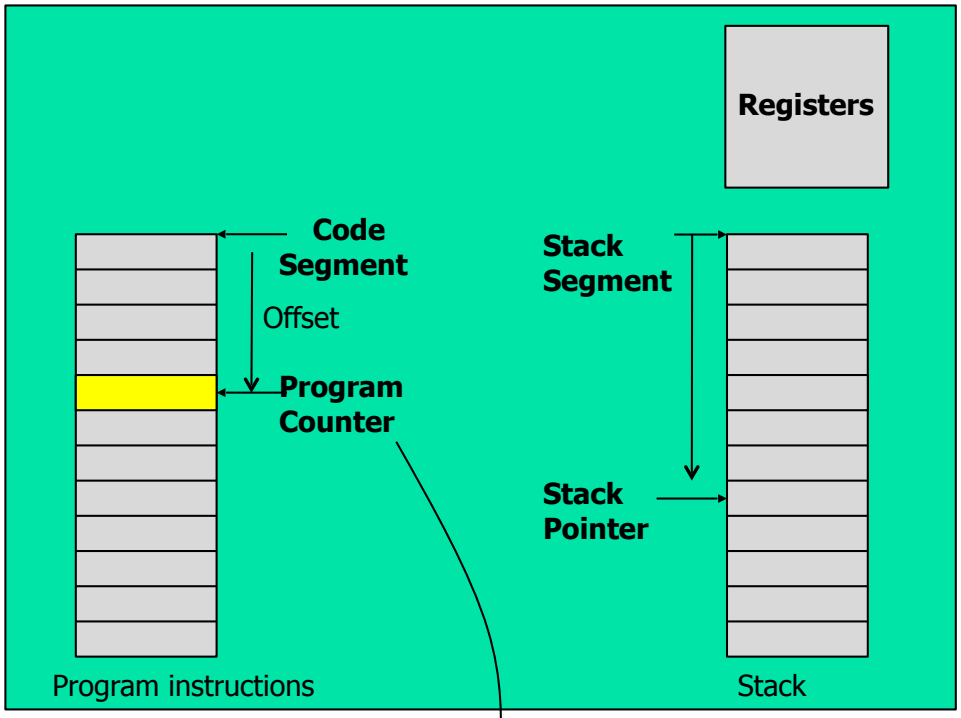
Where We Are:

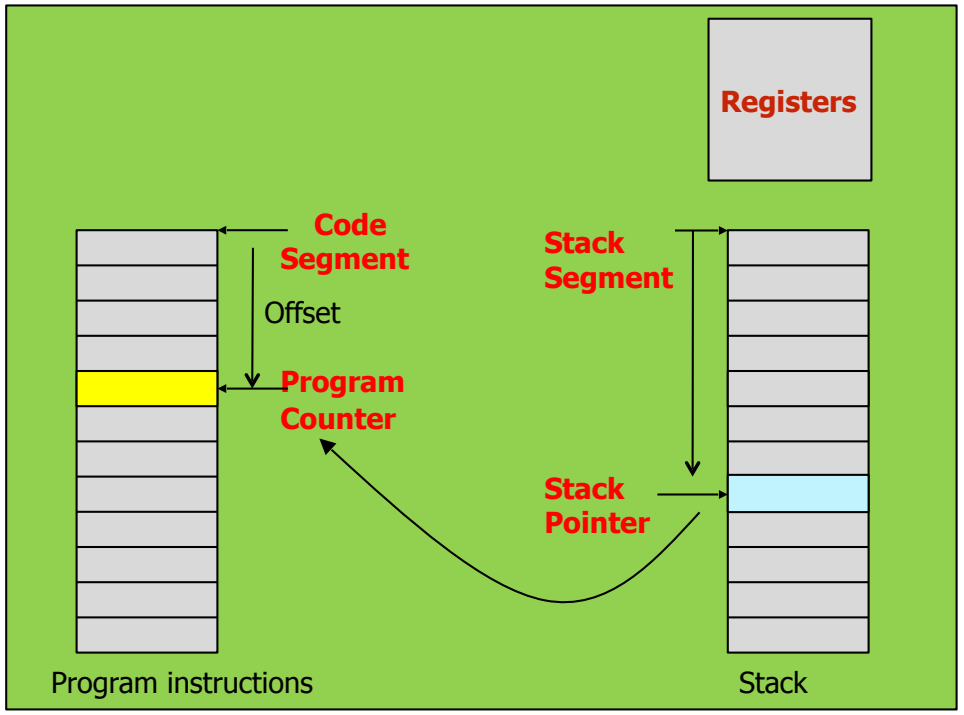
Interrupts to drive scheduling decisions!

Interrupt handlers are also tasks that share the CPU.



CTX Switch: Interrupt





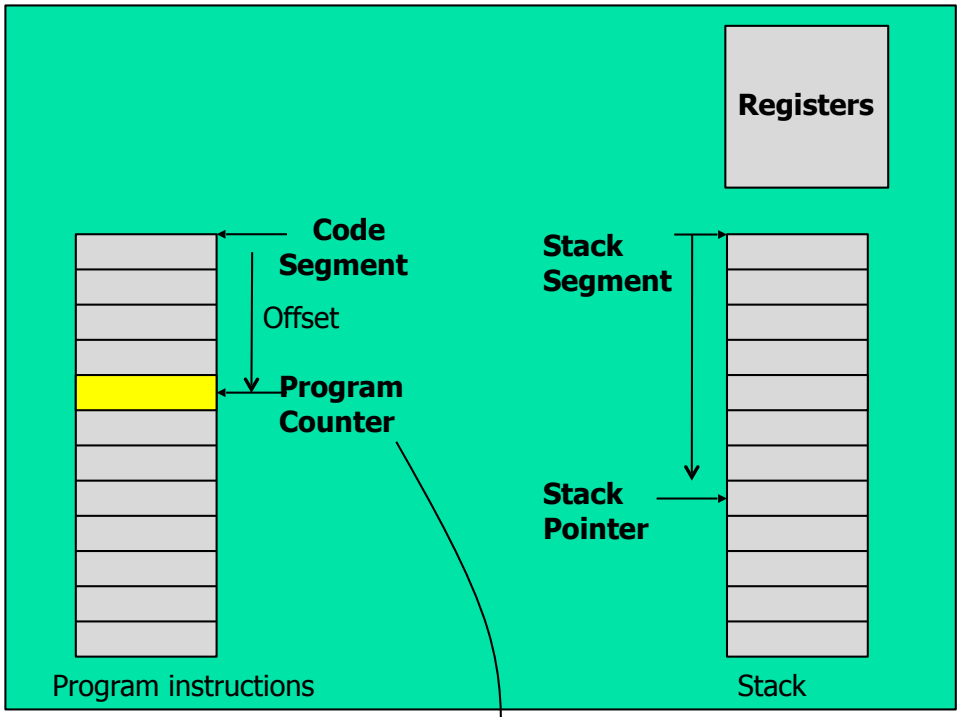
Save PC on thread stack
Jump to Interrupt handler

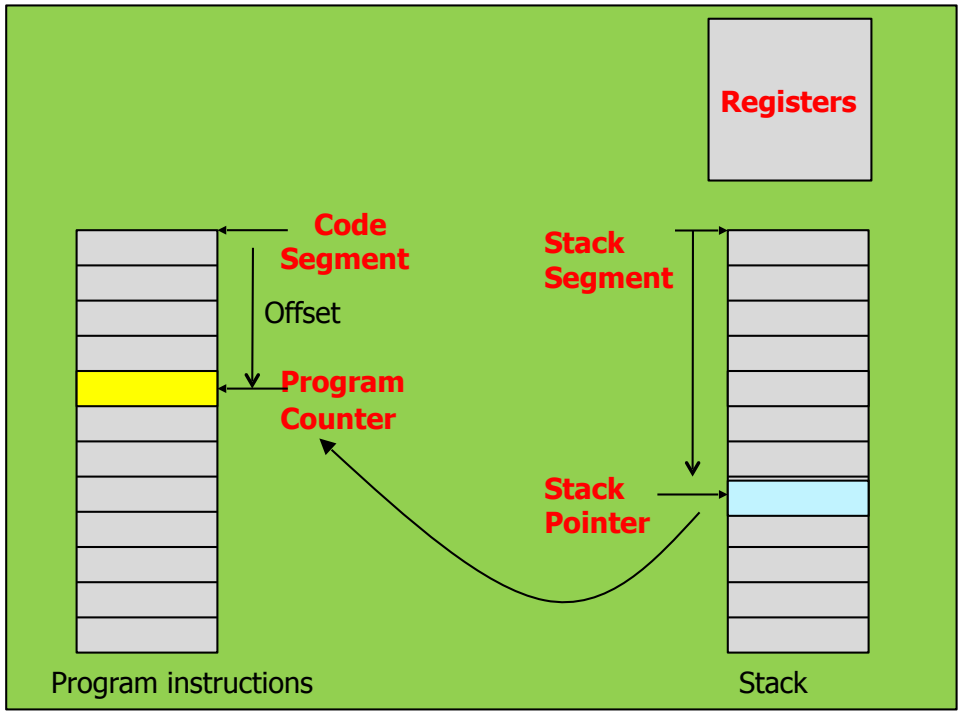
Handler
- Save thread state in thread control block
  (SP, registers, segment pointers, …)
- **Handle Interrupt**
- Choose next thread
- Load thread state from control block
- Pop PC from thread stack (return from handler)
- Resume prior task

Thread Control Block

Thread Control Block

Can also CTX Switch from Yield





Save PC on thread stack
Jump to yield() function

yield()
- Save thread state in thread control block
  (SP, registers, segment pointers, …)
- Choose next thread
- Load thread state from control block
- Pop PC from thread stack (return from handler)

Thread Control Block

Thread Control Block

How do we take interrupts safely??

- **Interrupt vector**
  - Where the processor looks for a handler
  - Limited number of entry points into kernel
- **Atomic transfer of control**
  - Single instruction to change:
    - Program counter
    - Stack pointer
    - Memory protection
    - Kernel/user mode
- **Transparent restartable execution**
  - User program does not know interrupt occurred

Interrupt Vector Table

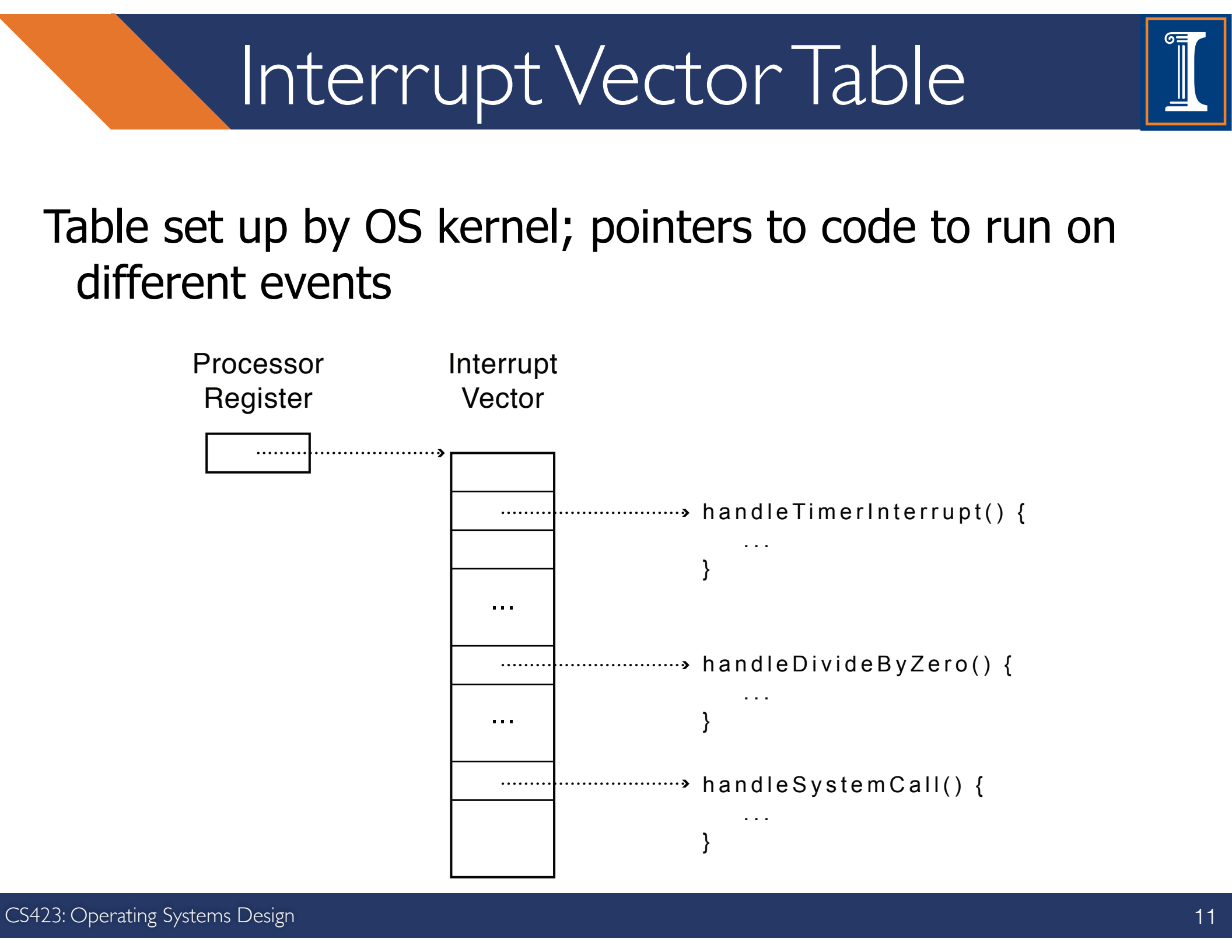


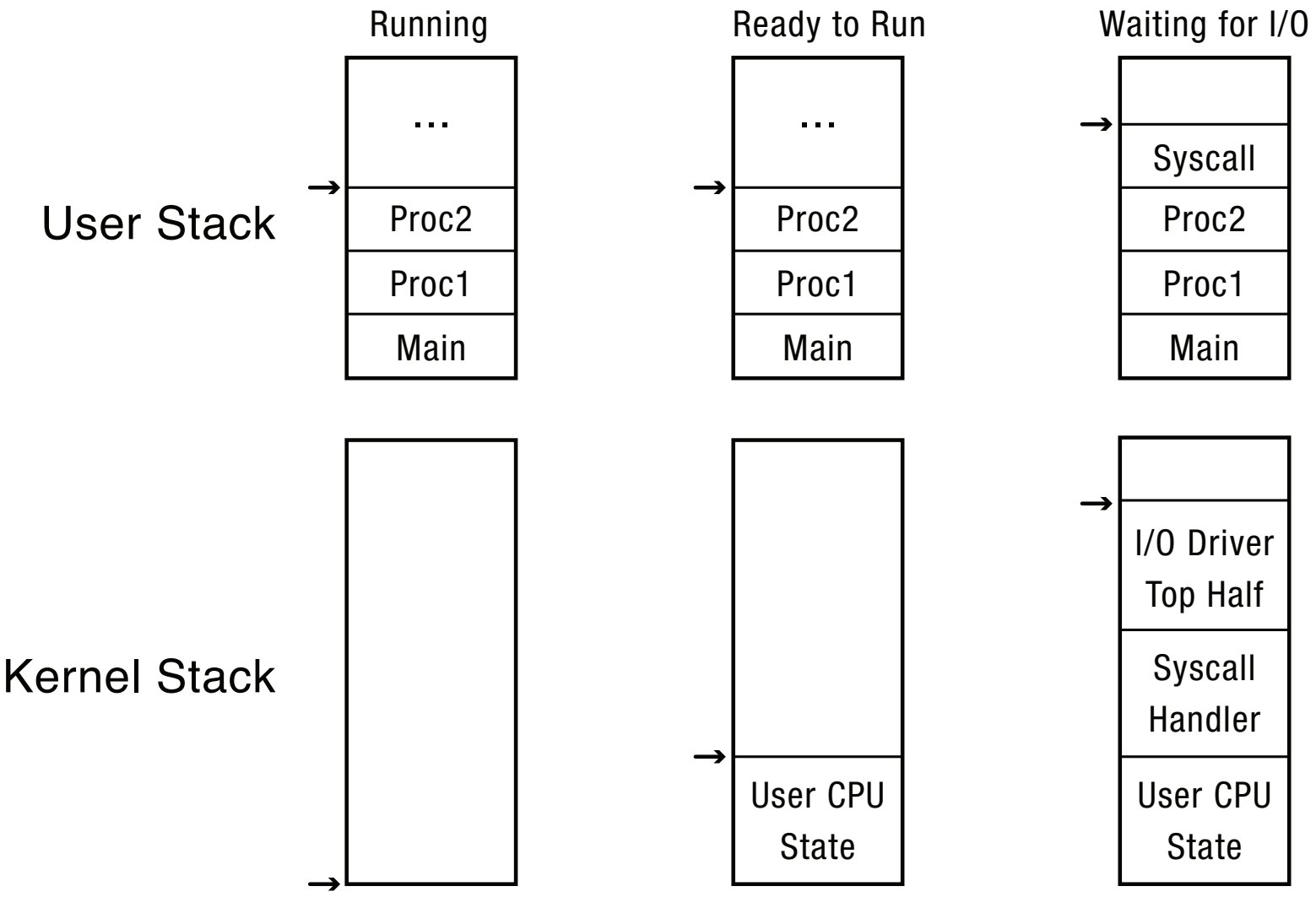
Table set up by OS kernel; pointers to code to run on different events

Interrupt Stack

- Per-processor, located in kernel (not user) memory
  - Fan fact! Usually a process/thread has both a kernel and user stack

- **Why can't the interrupt handler run on the stack of the interrupted user process?**

Interrupt Stack



Hardware Interrupts

- Hardware generated:
  - Different I/O devices are connected to different physical lines (pins) of an "Interrupt controller"
  - Device hardware signals the corresponding line
  - Interrupt controller signals the CPU (by signaling the Interrupt pin and passing an interrupt number)
  - CPU saves return address after next instruction and jumps to corresponding interrupt handler

Why Hardware INTs?

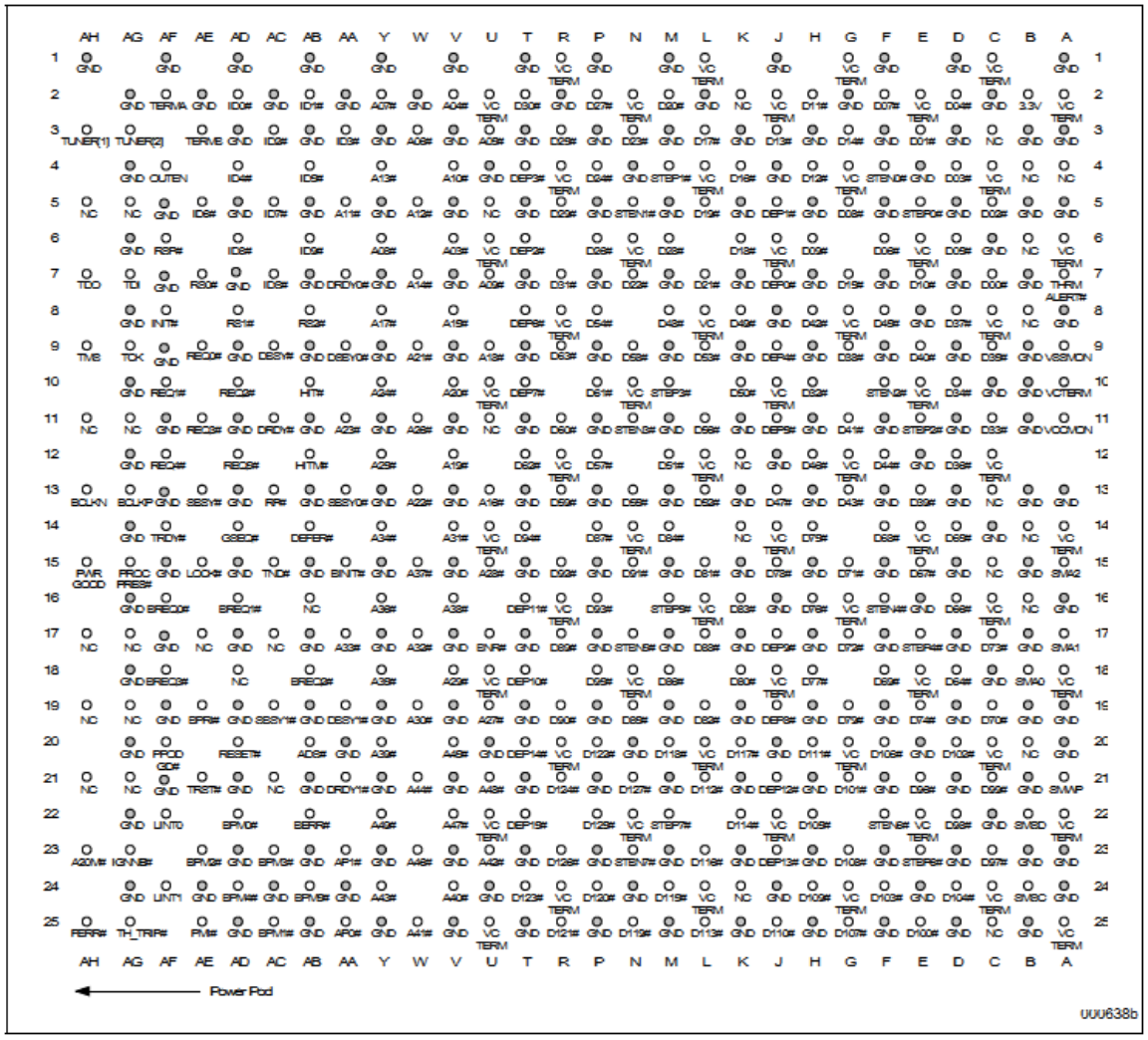
Why Hardware INTs?

- Hardware devices may need asynchronous and immediate service. For example:
  - Timer interrupt: Timers and time-dependent activities need to be updated with the passage of time at precise intervals

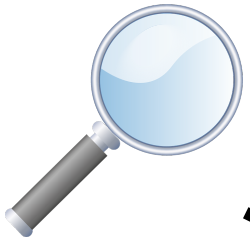
Why Hardware INTs?

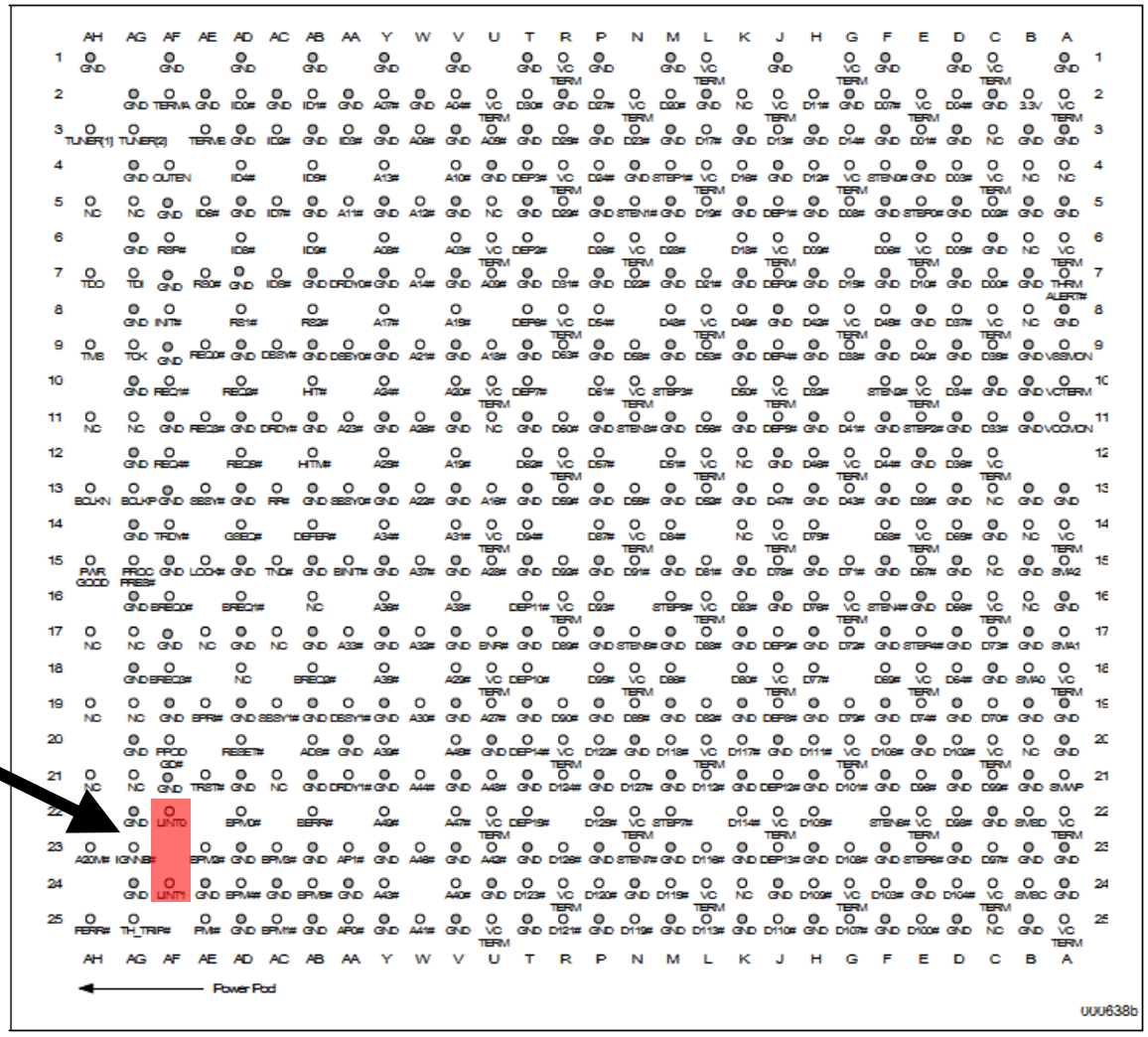
- Hardware devices may need asynchronous and immediate service. For example:
  - Timer interrupt: Timers and time-dependent activities need to be updated with the passage of time at precise intervals
  - Network interrupt: The network card interrupts the CPU when data arrives from the network
  - I/O device interrupt: I/O devices (such as mouse and keyboard) issue hardware interrupts when they have input (e.g., a new character or mouse click)

Ex: Itanium 2 Pinout

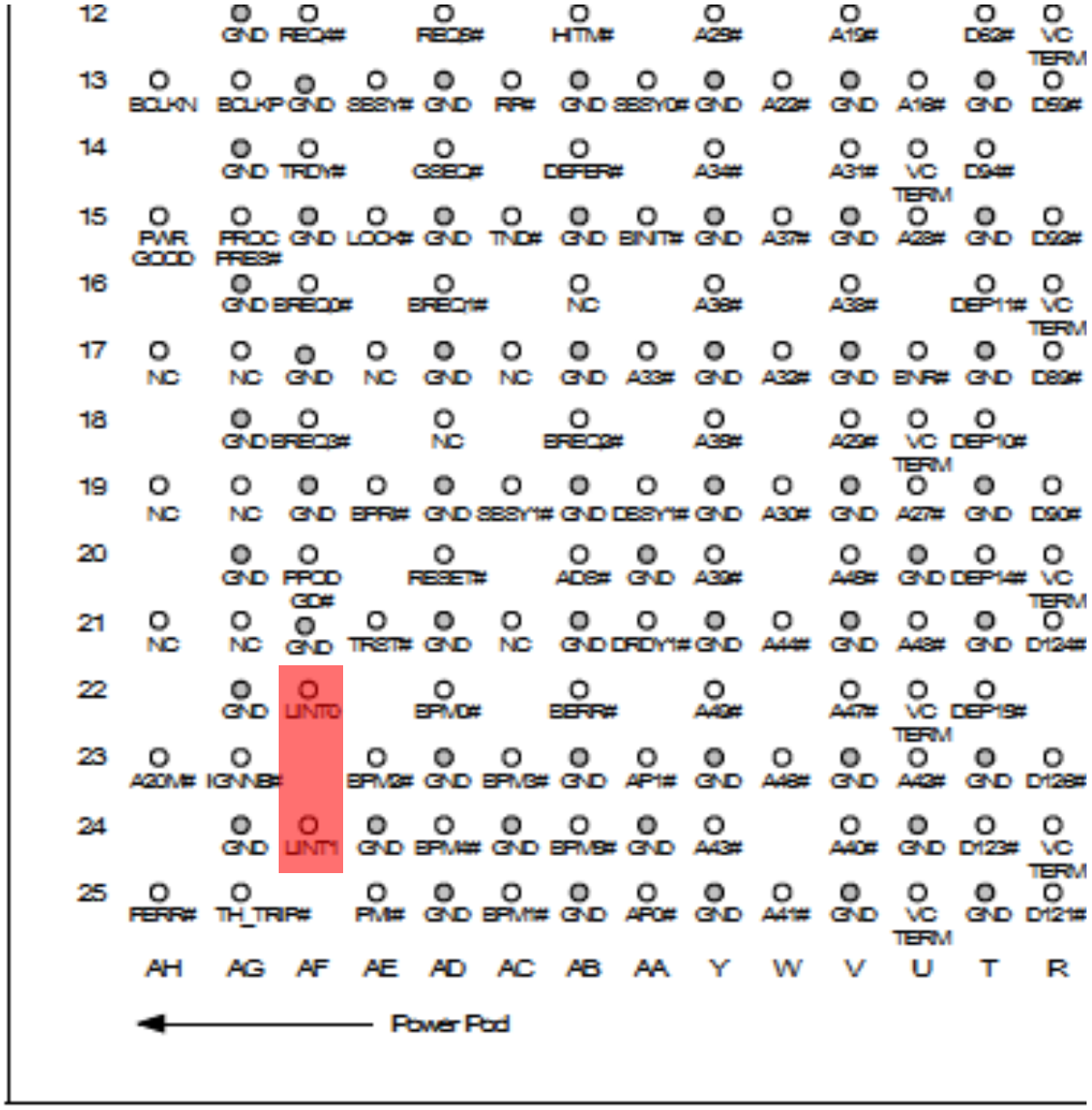


Ex: Itanium 2 Pinout





Ex: Itanium 2 Pinout



LINTx — lines/pins for hardware interrupts.

In this case…

LINT0 — line for unmaskable interrupts

LINT1 — line for maskable interrupts

A Note on Multicore

A Note on Multicore

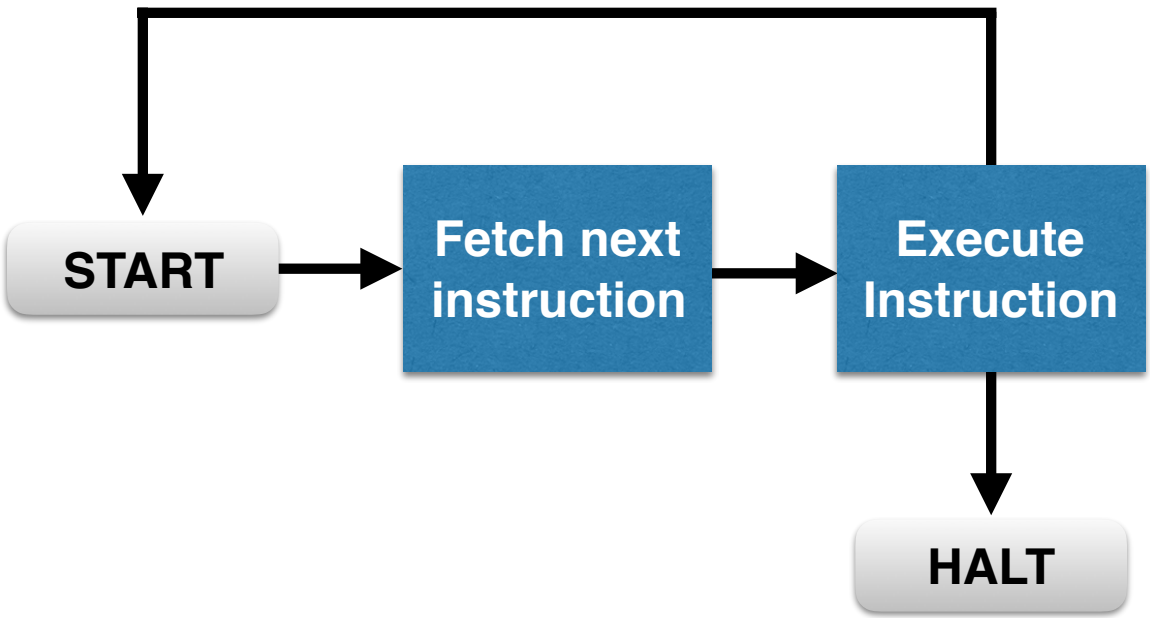
- How are interrupts handled on multicore machines?

A Note on Multicore

- How are interrupts handled on multicore machines?
  - On x86 systems each CPU gets its own local Advanced Programmable Interrupt Controller (APIC). They are wired in a way that allows routing device interrupts to any selected local APIC.
  - The OS can program the APICs to determine which interrupts get routed to which CPUs.
  - The default (unless OS states otherwise) is to route all interrupts to processor 0

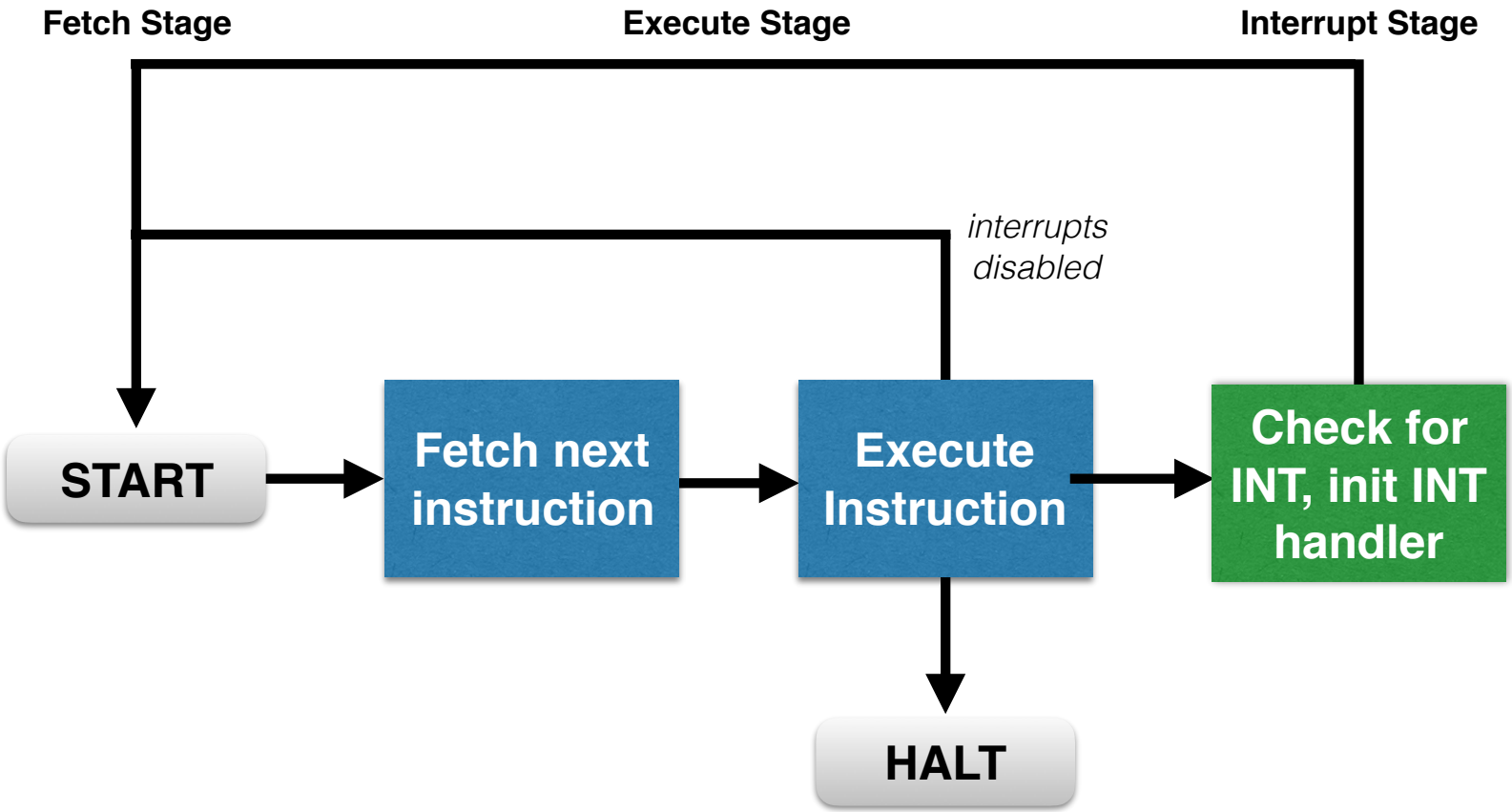
Instruction Cycle

How does interrupt handling change the instruction cycle?



Instruction Cycle w/ INTs

How does interrupt handling change the instruction cycle?



Processing HW INT's

Hardware

- Device controller or other hardware issues an interrupt.
- Processor finishes execution of current instruction.
- Processor signals acknowledgment of interrupt.
- Processor pushes PSW and PC onto control stack.
- Processor loads new PC value based on interrupt.

Software

- Save remainder of state information.
- Process interrupt.
- Restore process state information.
- Restore old PSW and PC.

Program Status Word (PSW) contains interrupt masks, privilege states, etc.

Other Interrupts

- Software Interrupts:
  - Interrupts caused by the execution of a software instruction:
    - `INT <interrupt_number>`
  - Used by the system call `interrupt()`
- Initiated by the running (user level) process
- Cause current processing to be interrupted and transfers control to the corresponding interrupt handler in the kernel

Other Interrupts

- Exceptions
  - Initiated by processor hardware itself
  - Example: divide by zero
- Like a software interrupt, they cause a transfer of control to the kernel to handle the exception

They're all interrupts

- HW -> CPU -> Kernel: Classic HW Interrupt
- User -> Kernel: SW Interrupt
- CPU -> Kernel: Exception
- Interrupt Handlers used in all 3 scenarios

INTs, Priorities, & Blocking

INTs, Priorities, & Blocking

- Interrupts (as the name suggests) have the highest priority (compared to user and kernel threads) and therefore run first
  - What are the implications on regular program execution?

INTs, Priorities, & Blocking

- Interrupts (as the name suggests) have the highest priority (compared to user and kernel threads) and therefore run first
  - What are the implications on regular program execution?
    - Must keep interrupt code short in order not to keep other processing stopped for a long time
    - Cannot block (regular processing does not resume until interrupt returns, so if the interrupt blocks in the middle the system "hangs")

INTs, Priorities, & Blocking

INTs, Priorities, & Blocking

- Can an interrupt handler use malloc()?

INTs, Priorities, & Blocking

- Can an interrupt handler use malloc()?
- Can an interrupt handler write data to disk?

INTs, Priorities, & Blocking

- Can an interrupt handler use malloc()?
- Can an interrupt handler write data to disk?
- Can an interrupt handler use busy wait?
  - E.G. — `while (!event) loop;`

Interrupt Masking

- Interrupt handler runs with interrupts off
  - Re-enabled when interrupt completes
- OS kernel can also turn interrupts off
  - Eg., when determining the next process/thread to run

Interrupt Handlers

Designing an Interrupt Handler (Bottom Half):

- Since the interrupt handler must be minimal, all other processing related to the event that caused the interrupt must be deferred
  - Example:
    - Network interrupt causes packet to be copied from network card
    - Other processing on the packet should be deferred until its time comes
- The deferred portion of interrupt processing is called the "Bottom Half"

soft_irq's

- 32 handlers that must be statically defined in the Linux kernel.
- A hardware interrupt (before returning) uses raise_softirq() to mark that a given soft_irq must execute the bottom half
- At a later time, when scheduling permits, the marked soft_irq handler is executed
  - When a hardware interrupt is finished
  - When a process makes a system call
  - When a new process is scheduled

soft_irq types

- HI_SOFTIRQ
- TIMER_SOFTIRQ
- NET_TX_SOFTRQ
- NET_RX_SOFTIRQ
- BLOCK_SOFTIRQ
- TASKLET_SOFTIRQ
- SCHED_SOFTIRQ
- …

soft_irq types

- HI_SOFTIRQ
- TIMER_SOFTIRQ
- NET_TX_SOFTRQ
- NET_RX_SOFTIRQ
- BLOCK_SOFTIRQ
- TASKLET_SOFTIRQ
- SCHED_SOFTIRQ
- …

Task let

- Bottom halves multiplexed on top of soft_irq's
- Scheduled using
  - tasklet_schedule()
  - tasklet_hi_schedule()
- Same tasklet invocations are serialized
- Tasklets can be created or removed dynamically
- Cannot sleep (cannot save their context)

Work Queues

- Work deferred to its own thread
- Can be scheduled together with other threads according to priorities set by a scheduling policy
- Associated with its thread control block and hence can block (and save context)
  - DECLARE_WORK(name, void (*func)(void *), void *data);
  - INIT_WORK(struct work_struct *work, void (*func)(void *), void *data);
  - schedule_work(&work);