

# **Machine-Level Programming III: (Switch Statements) and IA32 Procedures**

Lecture 4 - 2015  
Mads Chr. Olesen

Credits to Alexandre David (AAU) and  
Randy Bryant & Dave O'Hallaron (CMU)

# Today

- **Switch statements \***
- **IA 32 Procedures**
  - Stack Structure
  - Calling Conventions
  - Illustrations of Recursion & Pointers
- Arrays
- Structs

```

long switch_eg
(long x, long y, long z)
{
    long w = 1;
    switch(x) {
    case 1:
        w = y*z;
        break;
    case 2:
        w = y/z;
        /* Fall Through */
    case 3:
        w += z;
        break;
    case 5:
    case 6:
        w -= z;
        break;
    default:
        w = 2;
    }
    return w;
}

```

# Switch Statement Example

- Multiple case labels
  - Here: 5 & 6
- Fall through cases
  - Here: 2
- Missing cases
  - Here: 4

# Jump Table Structure

## Switch Form

```
switch(x) {  
  case val_0:  
    Block 0  
  case val_1:  
    Block 1  
    ...  
  case val_n-1:  
    Block n-1  
}
```

## Approximate Translation

```
target = JTab[x];  
goto *target;
```

## Jump Table

jtab:	Targ0
	Targ1
	Targ2
	•
	•
	•
	Targn-1

## Jump Targets

Targ0:

Code Block  
0

Targ1:

Code Block  
1

Targ2:

Code Block  
2

•  
•  
•

Targn-1:

Code Block  
n-1

# Switch Statement Example (IA32)

```
long switch_eg(long x, long y, long z)
{
    long w = 1;
    switch(x) {
        . . .
    }
    return w;
}
```

What range of values takes default?

Setup:

```
switch_eg:
    pushl    %ebp                # Setup
    movl     %esp, %ebp         # Setup
    movl     8(%ebp), %eax       # %eax = x
    cmpl     $6, %eax           # Compare x:6
    ja       .L2                # If unsigned > goto default
    jmp      *.L7(, %eax, 4)      # Goto *JTab[x]
```

Note that **w** not initialized here

# Switch Statement Example (IA32)

```
long switch_eg(long x, long y, long z)
{
    long w = 1;
    switch(x) {
        . . .
    }
    return w;
}
```

Jump table

.section	.rodata
.align 4	
.L7:	
.long	.L2 # x = 0
.long	.L3 # x = 1
.long	.L4 # x = 2
.long	.L5 # x = 3
.long	.L2 # x = 4
.long	.L6 # x = 5
.long	.L6 # x = 6

Setup:

switch\_eg:

pushl	%ebp	# Setup
movl	%esp, %ebp	# Setup
movl	8(%ebp), %eax	# eax = x
cmpl	\$6, %eax	# Compare x:6
ja	.L2	# If unsigned > goto default
Indirect jump	→ jmp	*.L7(, %eax, 4) # Goto *JTab[x]

# Assembly Setup Explanation

## ■ Table Structure

- Each target requires 4 bytes
- Base address at .L7

## ■ Jumping

- **Direct:** `jmp .L2`
- Jump target is denoted by label .L2
- **Indirect:** `jmp *.L7(, %eax, 4)`
- Start of jump table: .L7
- Must scale by factor of 4 (labels have 32-bits = 4 Bytes on IA32)
- Fetch target from effective Address `.L7 + eax*4`
  - Only for  $0 \leq x \leq 6$

Jump table

```
.section      .rodata
    .align 4
.L7:
    .long     .L2 # x = 0
    .long     .L3 # x = 1
    .long     .L4 # x = 2
    .long     .L5 # x = 3
    .long     .L2 # x = 4
    .long     .L6 # x = 5
    .long     .L6 # x = 6
```

Try it yourself!

# IA32 Object Code

## ■ Setup

- Label .L2 becomes address 0x8048422
- Label .L7 becomes address 0x8048660

## Assembly Code

```
switch_eg:
    . . .
    ja      .L2          # If unsigned > goto default
    jmp     *.L7(, %eax, 4) # Goto *JTab[x]
```

## Disassembled Object Code

```
08048410 <switch_eg>:
    . . .
    8048419: 77 07                ja      8048422 <switch_eg+0x12>
    804841b: ff 24 85 60 86 04 08 jmp     *0x8048660(, %eax, 4)
```

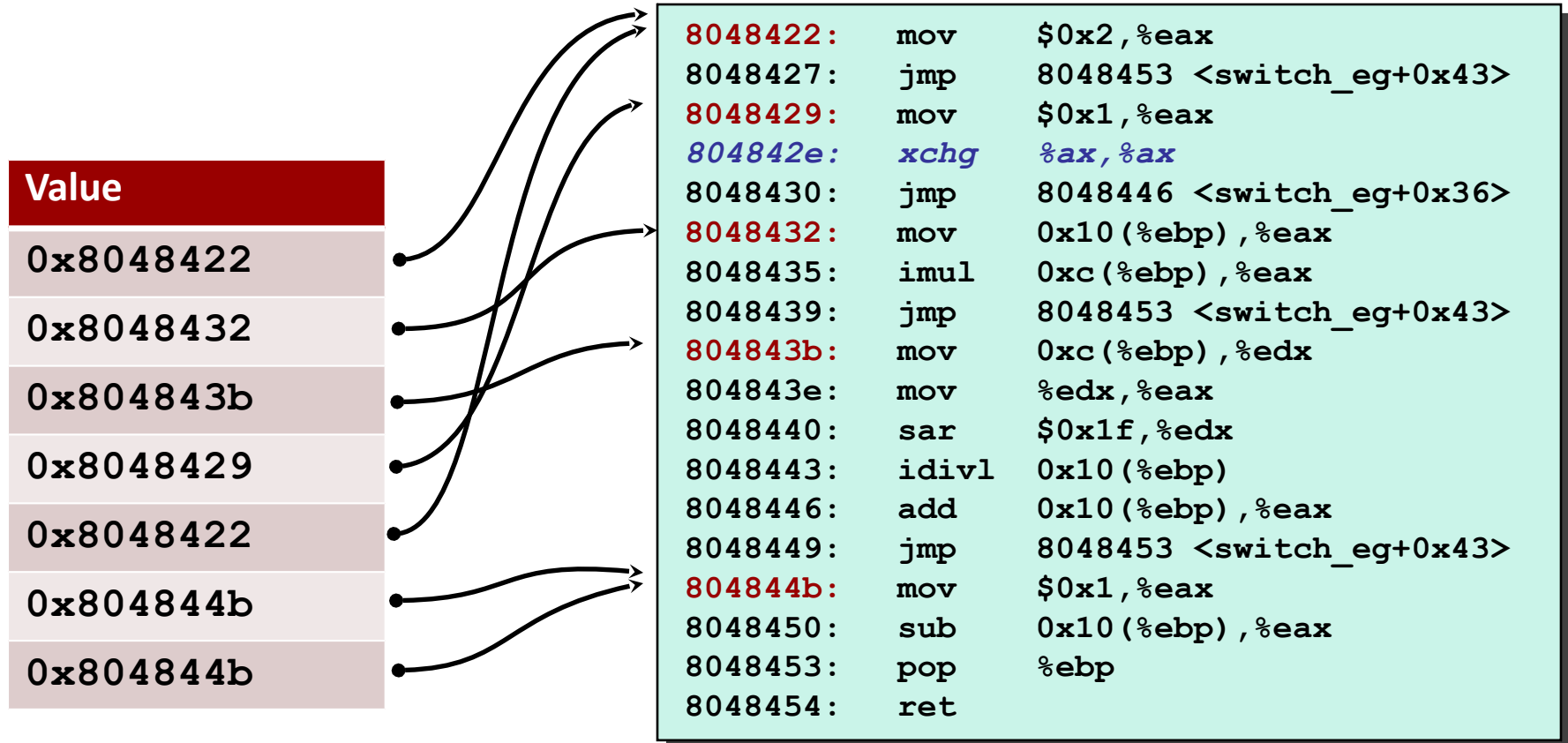


# IA32 Object Code (cont.)

## ■ Jump Table

- Doesn't show up in disassembled code
- Does show up with gcc -S
- Can inspect using GDB
- gdb switch
- (gdb) x/7xw 0x8048660
  - Examine 7 hexadecimal format "words" (4-bytes each)
  - Use command "help x" to get format documentation

# Matching Disassembled Targets



Jump table

# Summarizing

## ■ C Control

- if-then-else
- do-while
- while, for
- switch

## ■ Assembler Control

- Conditional jump
- Conditional move
- Indirect jump
- Compiler generates code sequence to implement more complex control

## ■ Standard Techniques

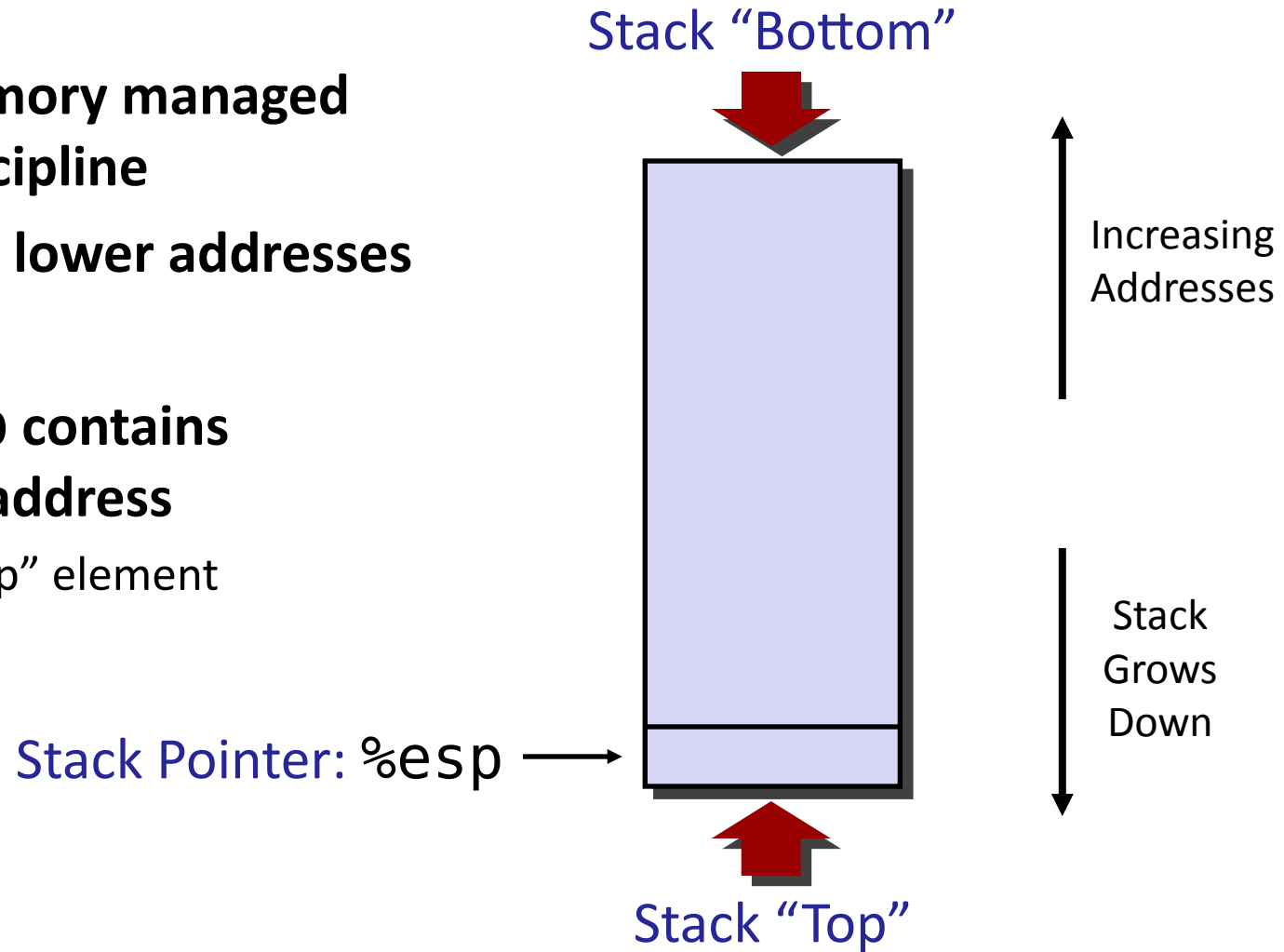
- Loops converted to do-while form
- Large switch statements use jump tables
- Sparse switch statements may use decision trees

# Today

- Switch statements
- **IA 32 Procedures**
  - Stack Structure
  - Calling Conventions
  - Illustrations of Recursion & Pointers
- Arrays
- (Structs)

# IA32 Stack

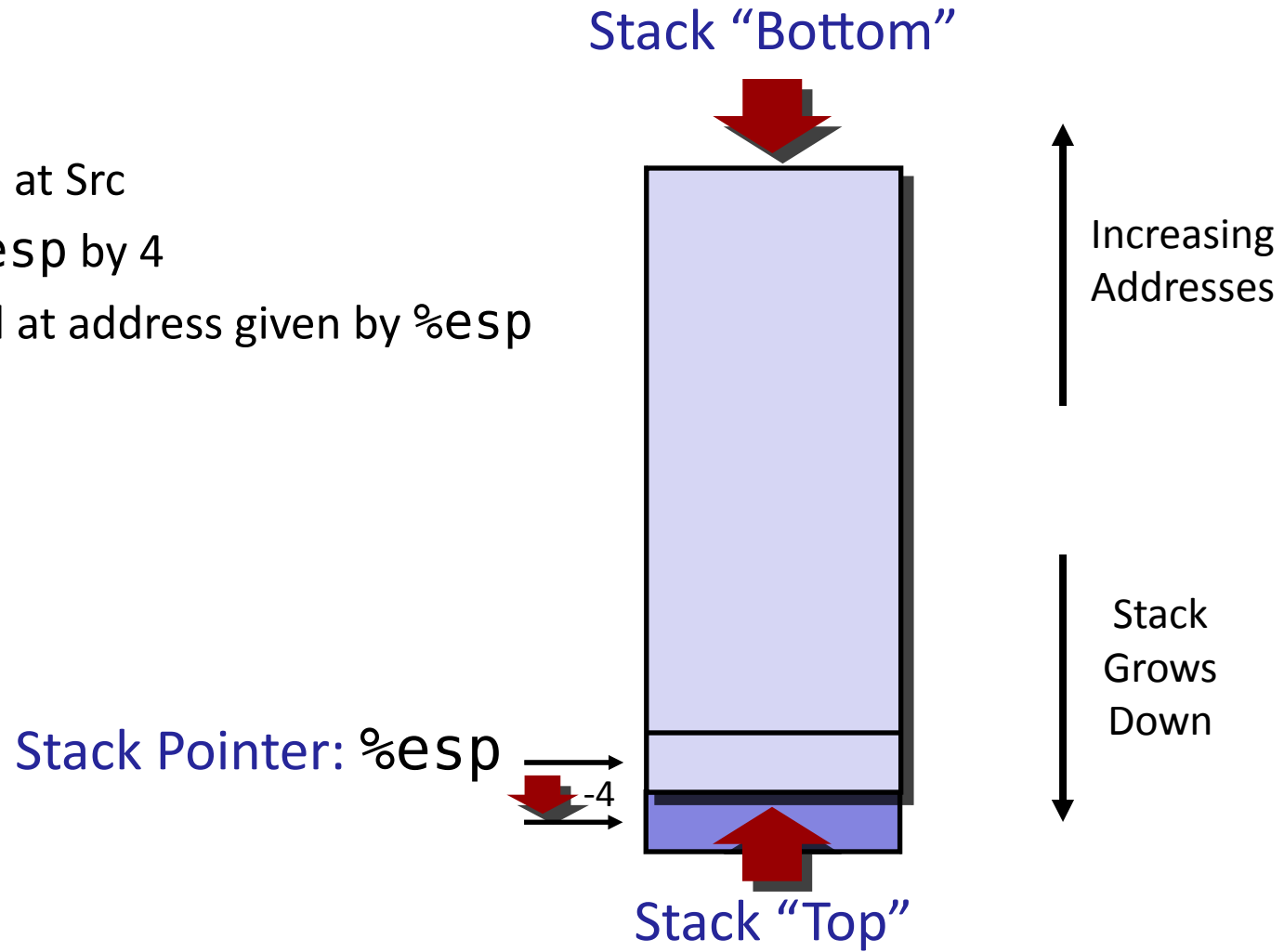
- Region of memory managed with stack discipline
- Grows toward lower addresses
- Register **%esp** contains lowest stack address
  - address of “top” element



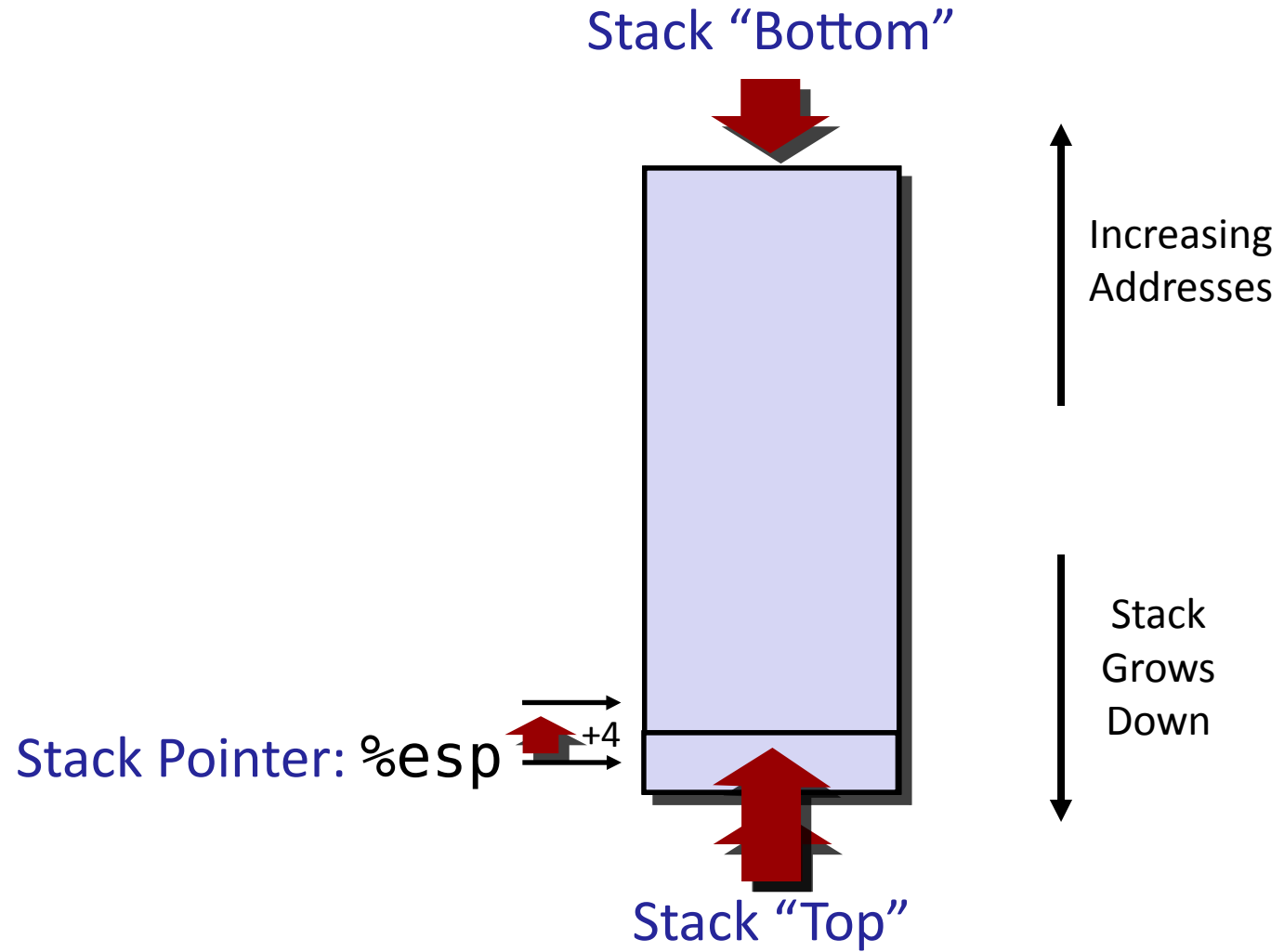
# IA32 Stack: Push

## ■ `pushl Src`

- Fetch operand at Src
- Decrement `%esp` by 4
- Write operand at address given by `%esp`



# IA32 Stack: Pop



# Procedure Control Flow

## ■ Use stack to support procedure call and return

### ■ **Procedure call:** `call label`

- Push return address on stack
- Jump to label

### ■ **Return address:**

- Address of the next instruction right after call
- Example from disassembly

```
804854e:  e8 3d 06 00 00    call    8048b90 <main>
8048553:  50               pushl   %eax
```

- Return address = 0x8048553

### ■ **Procedure return:** `ret`

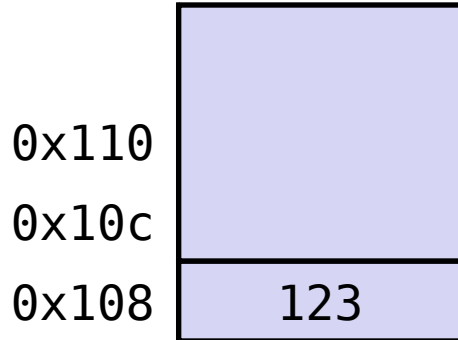
- Pop address from stack
- Jump to address



# Procedure Call Example

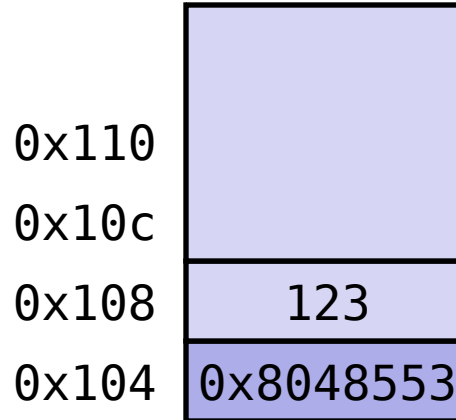
```
804854e:  e8 3d 06 00 00  call  8048b90 <main>
8048553:  50              pushl  %eax
```

call 8048b90



%esp 0x108

%eip 0x804854e



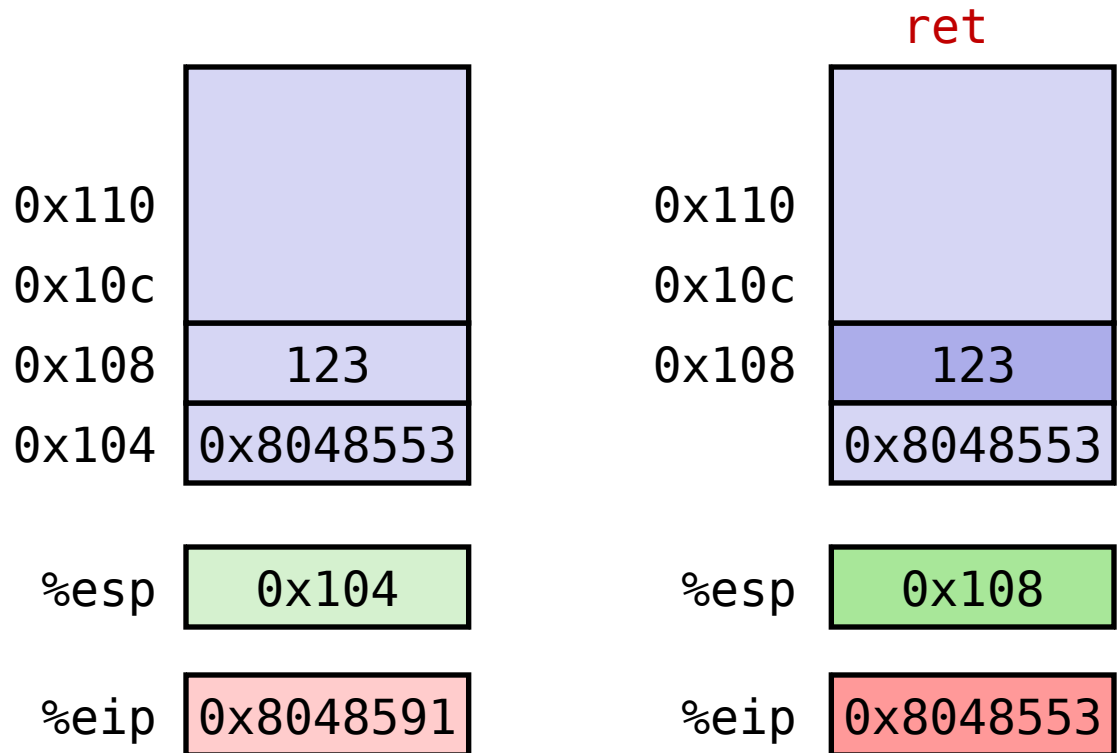
%esp 0x104

%eip 0x8048b90

%eip: program counter

# Procedure Return Example

8048591: c3 ret



`%eip`: program counter

# Stack-Based Languages

## ■ Languages that support recursion

- e.g., C, Pascal, Java
- Code must be “Reentrant”
  - Multiple simultaneous instantiations of single procedure
- Need some place to store state of each instantiation
  - Arguments
  - Local variables
  - Return pointer

## ■ Stack discipline

- State for given procedure needed for limited time
  - From when called to when return
- Callee returns before caller does

## ■ Stack allocated in **Frames**

- state for single procedure instantiation

# Call Chain Example

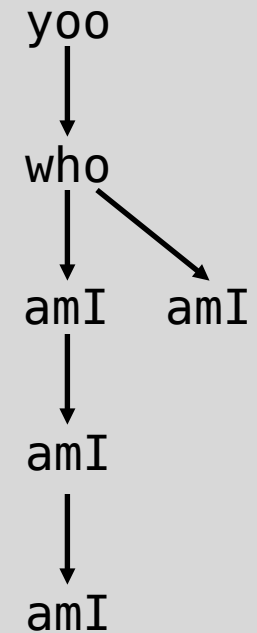
```
yoo (...)  
{  
  .  
  .  
  who ();  
  .  
  .  
}
```

```
who (...)  
{  
  . . .  
  amI ();  
  . . .  
  amI ();  
  . . .  
}
```

```
amI (...)  
{  
  .  
  .  
  amI ();  
  .  
  .  
}
```

Procedure `amI ()` is recursive

Example  
Call Chain



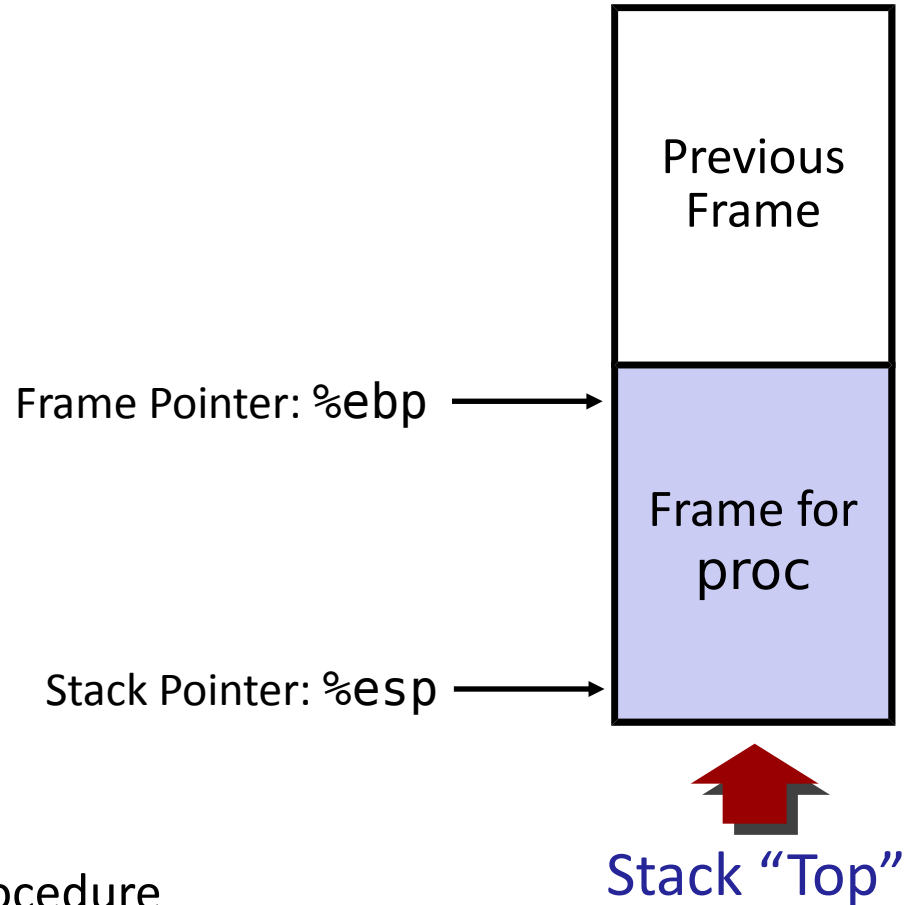
# Stack Frames

## ■ Contents


- Local variables
- Return information
- Temporary space

## ■ Management

- Space allocated when enter procedure
  - “Set-up” code
- Deallocated when return
  - “Finish” code

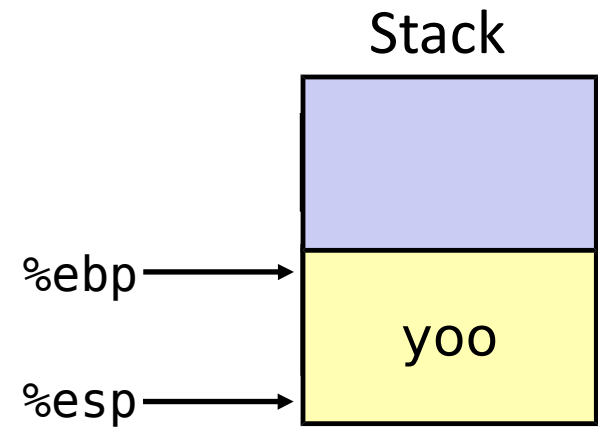


# Example

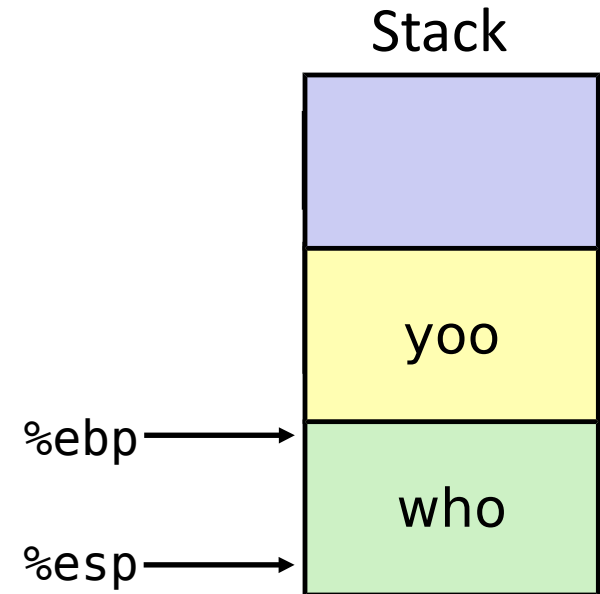
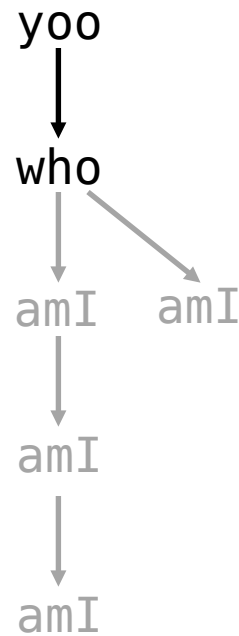
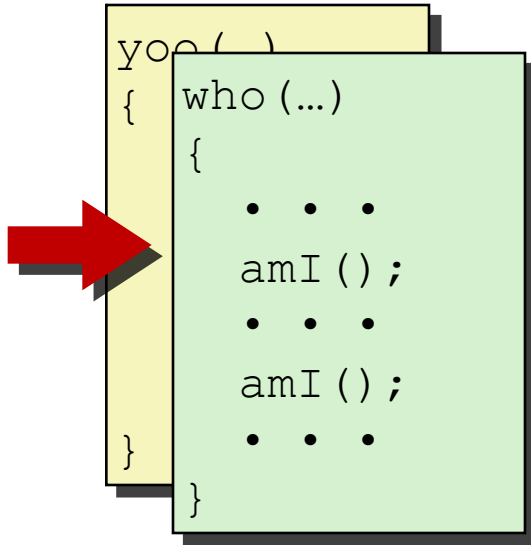


```
yoo (...)  
{  
  .  
  .  
  who ();  
  .  
  .  
}
```

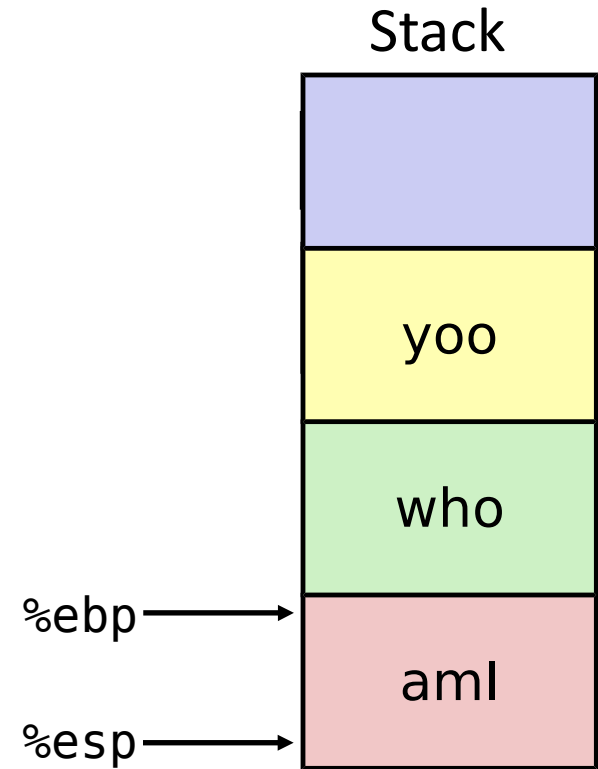
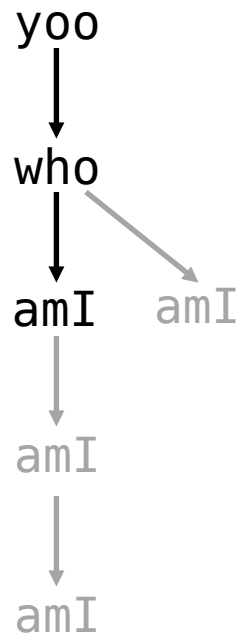
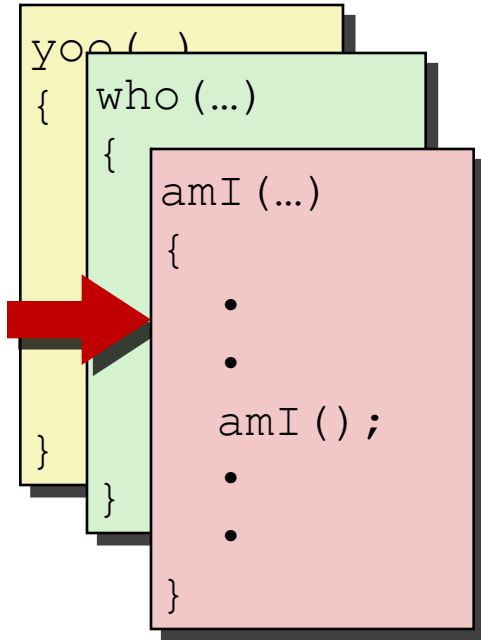
```
yoo  
  ↓  
who  
  ↓  ↘  
amI  amI  
  ↓  
amI  
  ↓  
amI
```



# Example

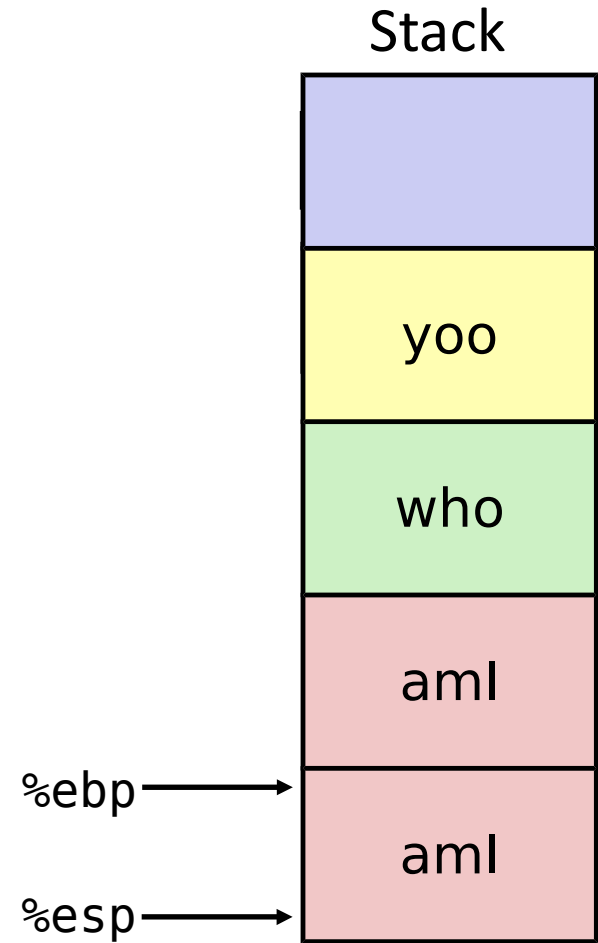
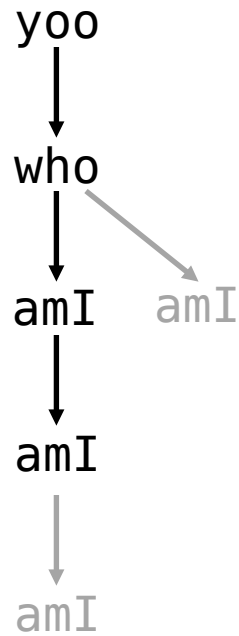
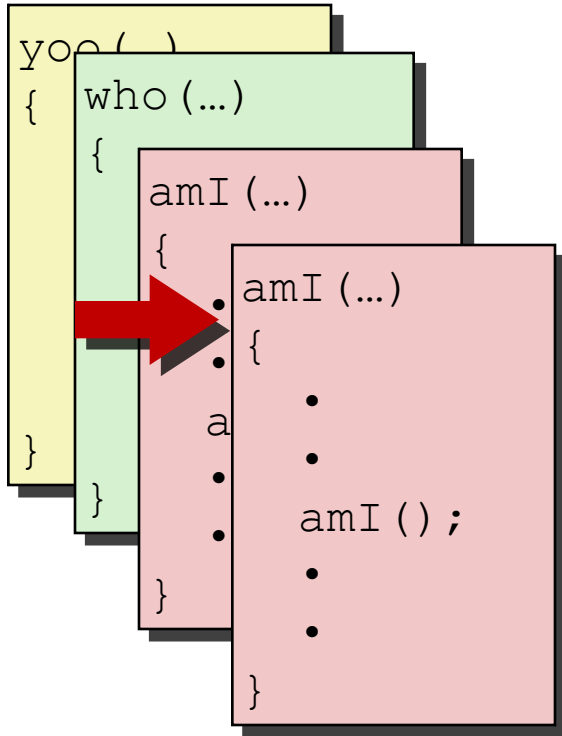


# Example

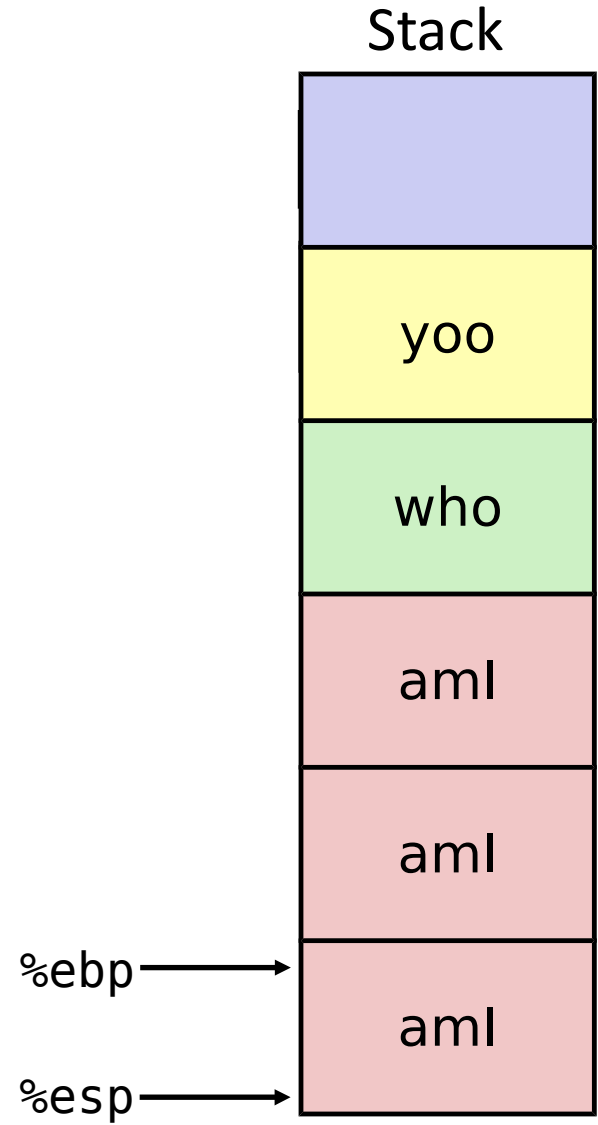
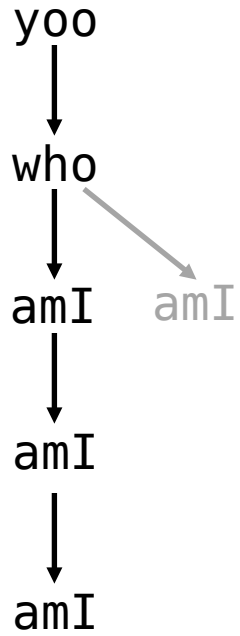
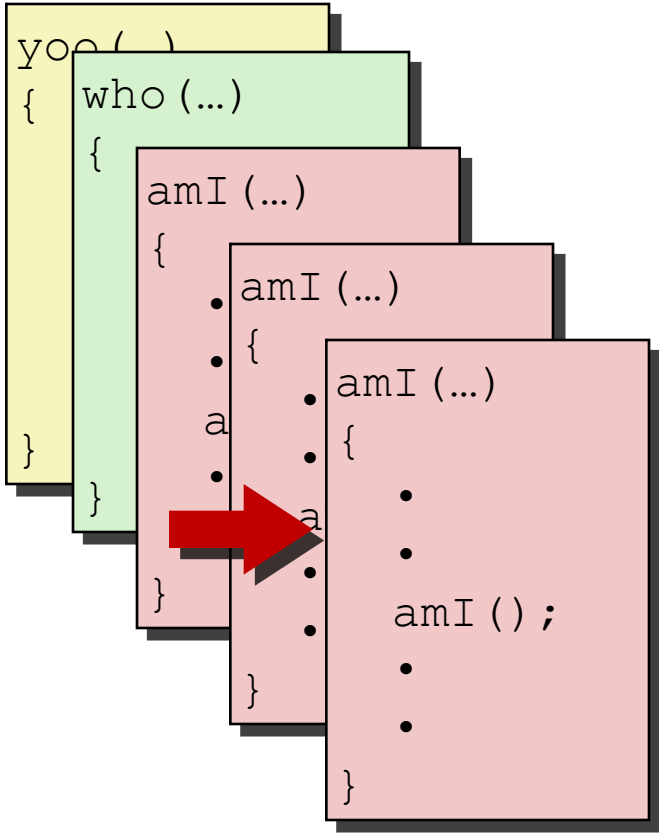




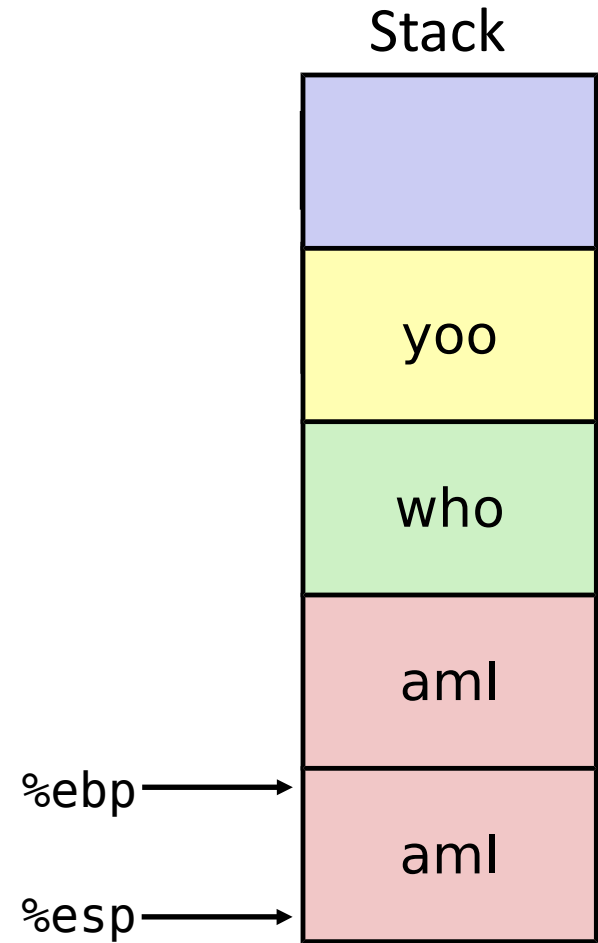
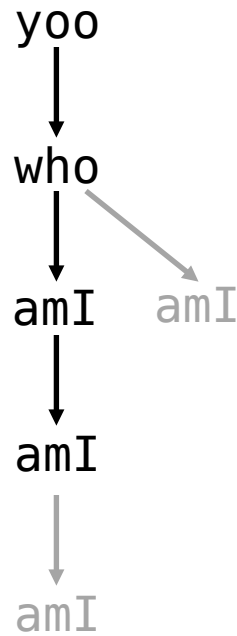
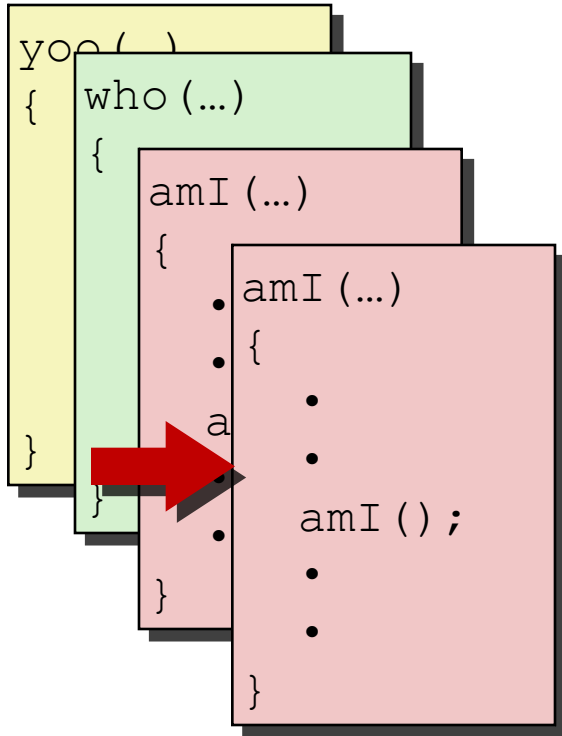
# Example



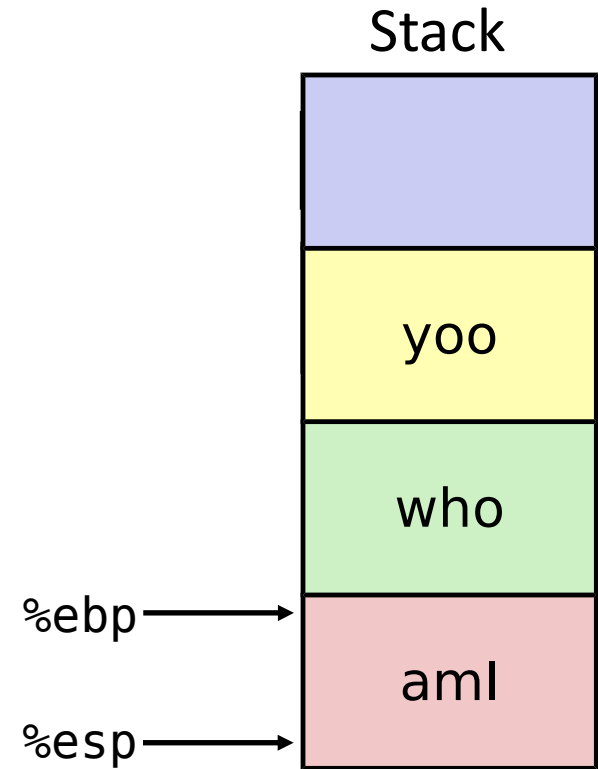
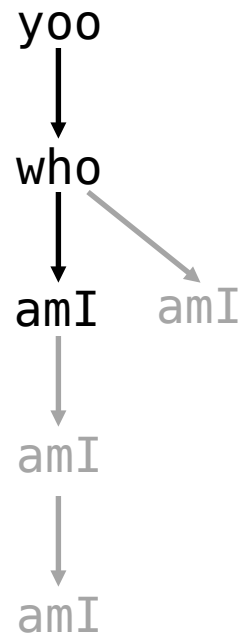
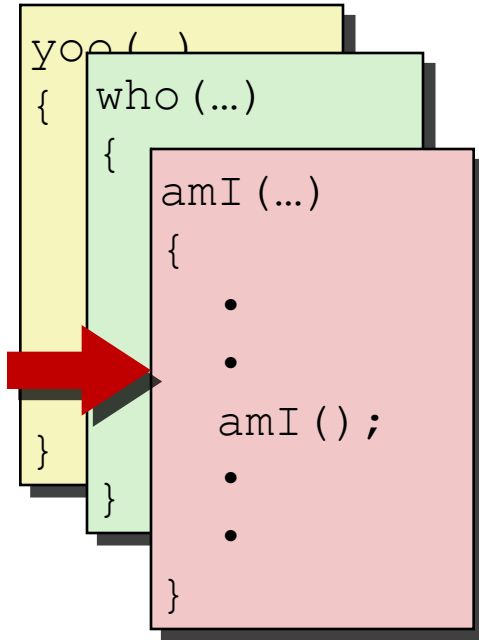
## Example



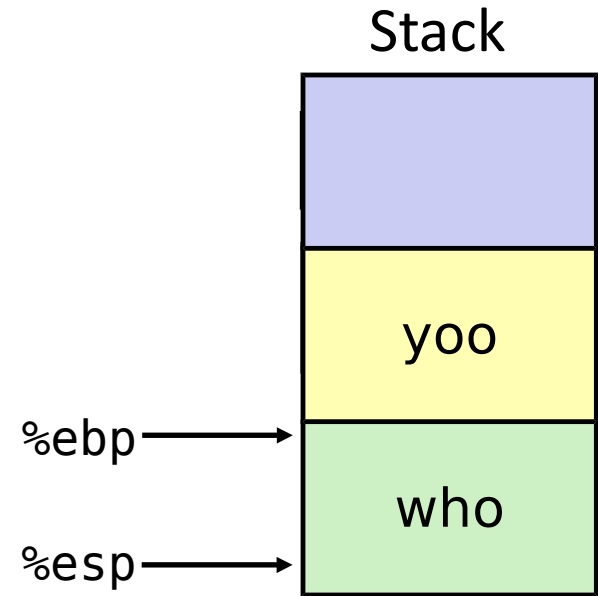
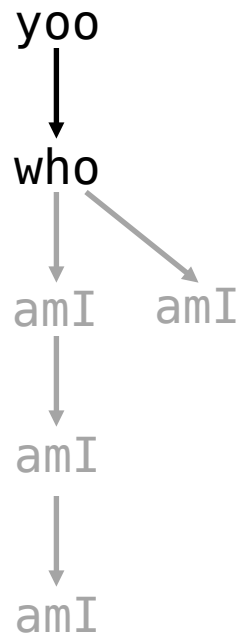
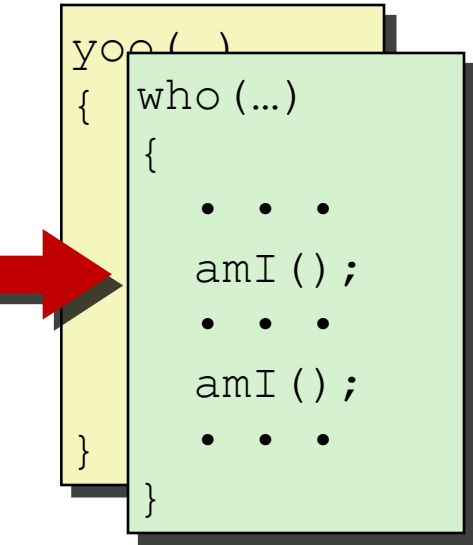
# Example



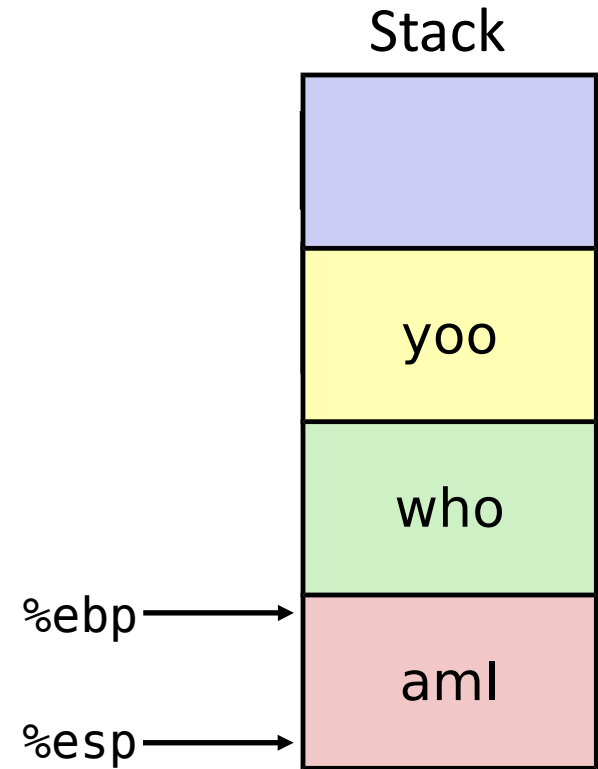
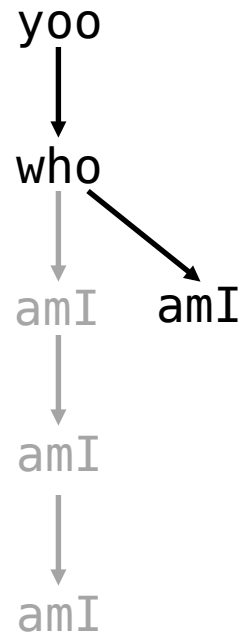
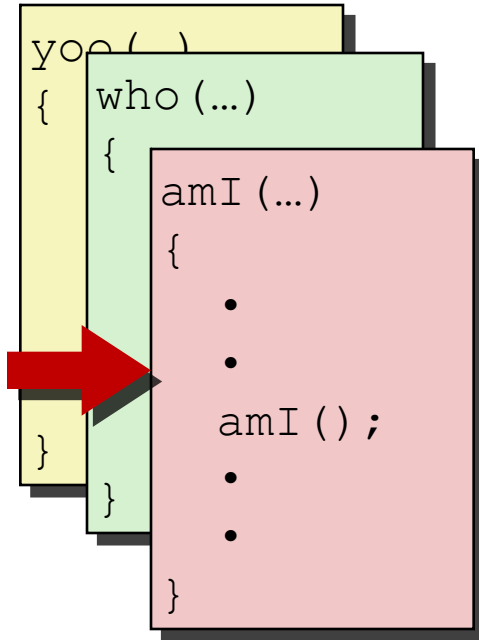
# Example



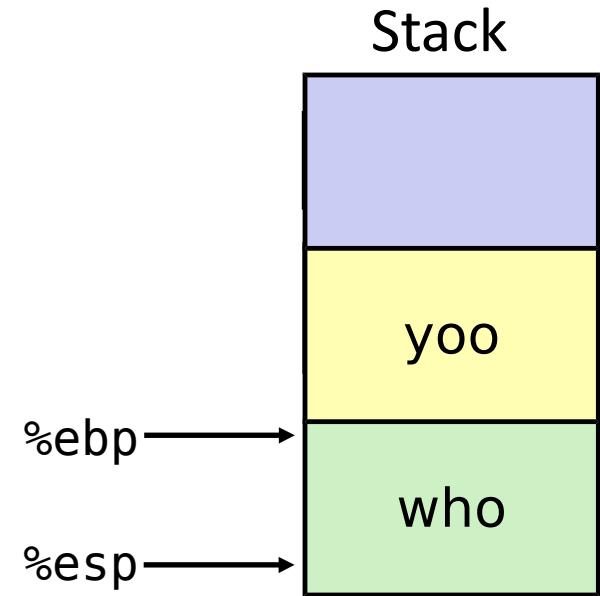
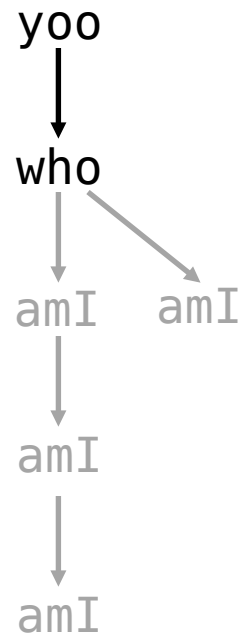
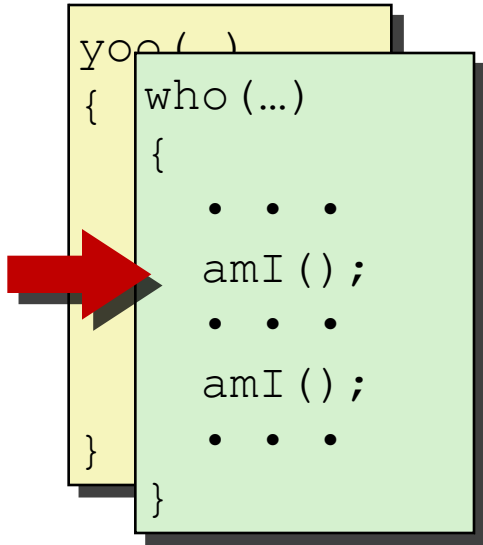
# Example



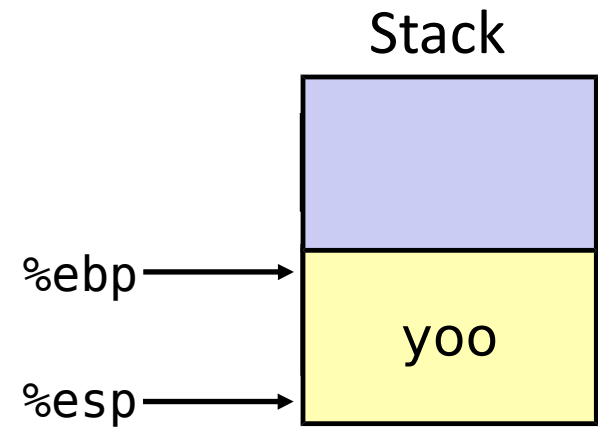
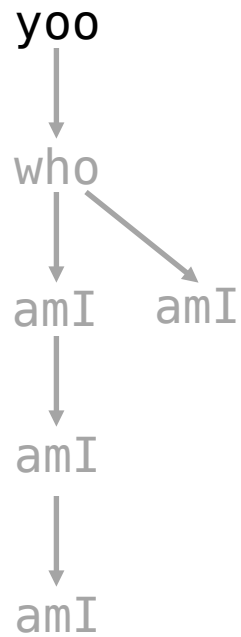
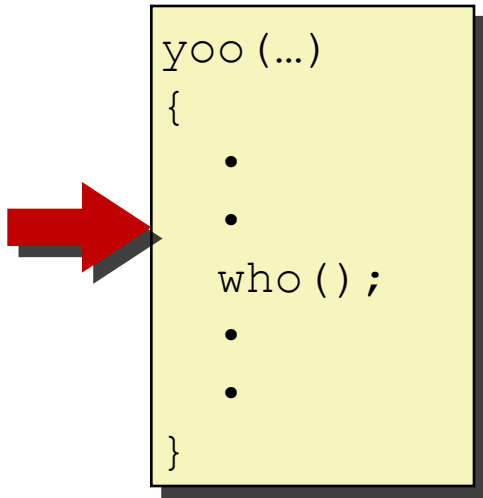
# Example



# Example



# Example





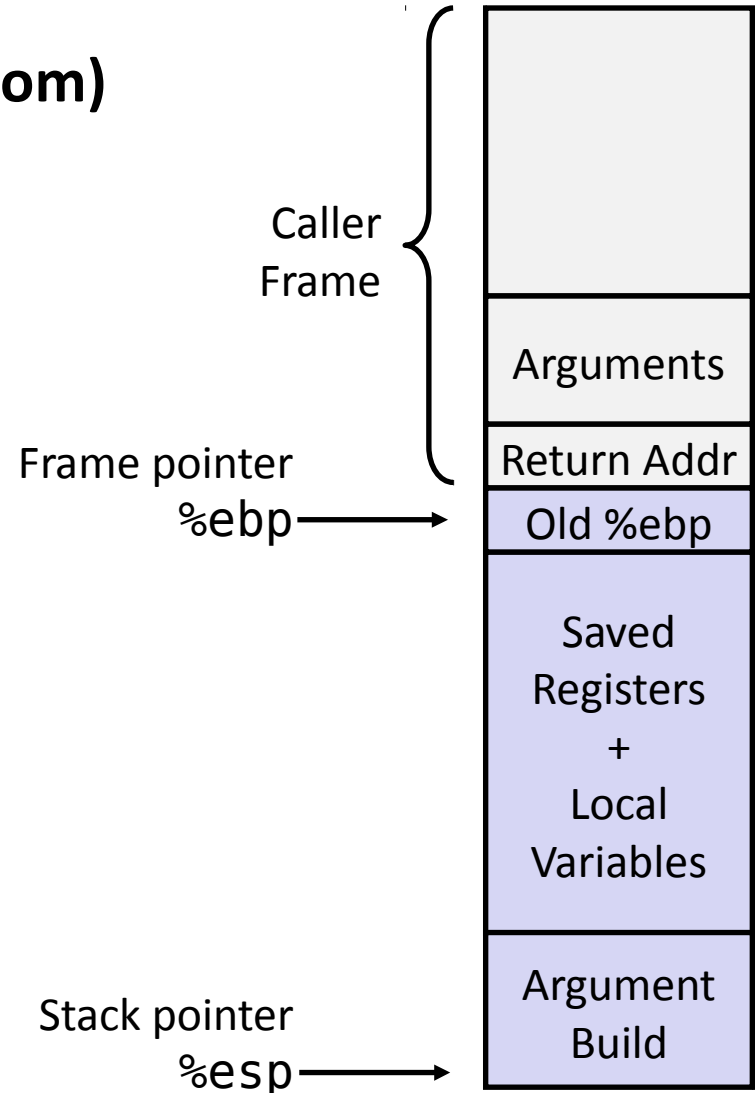
# IA32/Linux Stack Frame

## ■ Current Stack Frame (“Top” to Bottom)

- “Argument build:”  
Parameters for function about to call
- Local variables  
If can’t keep in registers
- Saved register context
- Old frame pointer

## ■ Caller Stack Frame

- Return address
  - Pushed by `call` instruction
- Arguments for this call



# Revisiting swap

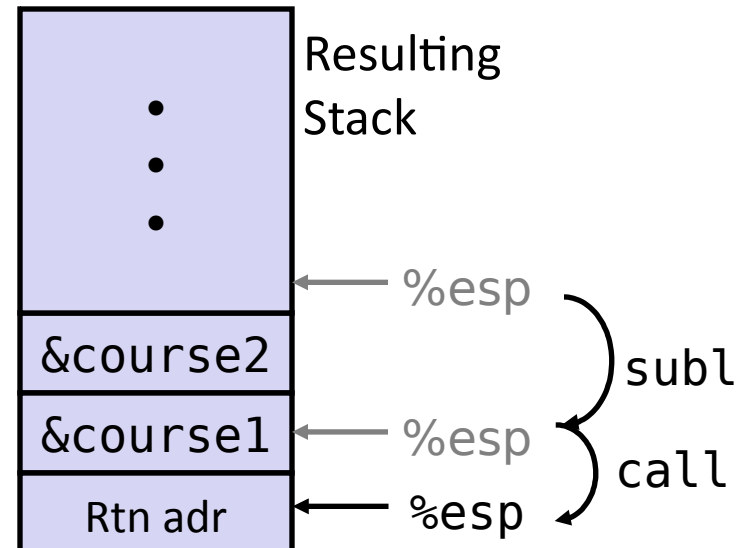
```
int course1 = 15213;
int course2 = 18243;

void call_swap() {
    swap(&course1, &course2);
}
```

```
void swap(int *xp, int *yp)
{
    int t0 = *xp;
    int t1 = *yp;
    *xp = t1;
    *yp = t0;
}
```

Calling swap from call\_swap

```
call_swap:
    . . .
    subl    $8, %esp
    movl    course2, 4(%esp)
    movl    course1, (%esp)
    call    swap
    . . .
```



# Revisiting Swap

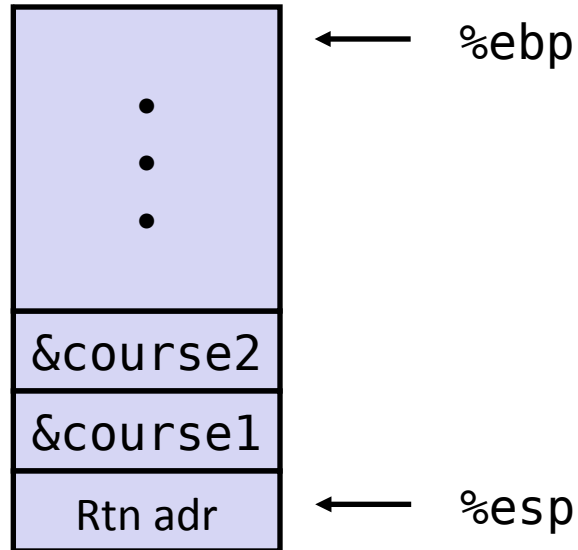
```
void swap(int *xp, int *yp)
{
    int t0 = *xp;
    int t1 = *yp;
    *xp = t1;
    *yp = t0;
}
```

swap:

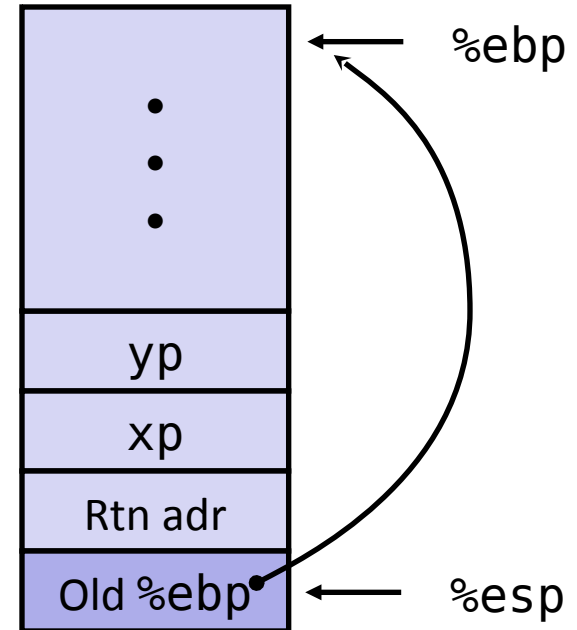
pushl %ebp	}	Set Up
movl %esp, %ebp		
pushl %ebx		
movl 8(%ebp), %edx	}	Body
movl 12(%ebp), %ecx		
movl (%edx), %ebx		
movl (%ecx), %eax		
movl %eax, (%edx)		
movl %ebx, (%ecx)		
popl %ebx	}	Finish
popl %ebp		
ret		

# swap Setup #1

Entering Stack



Resulting Stack



**swap:**

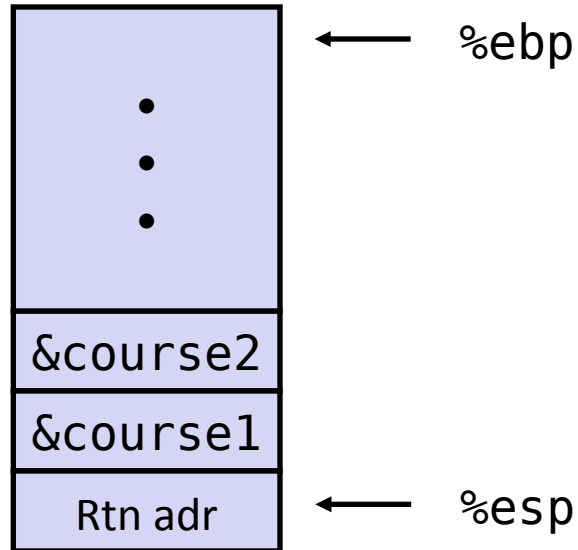
**pushl %ebp**

**movl %esp, %ebp**

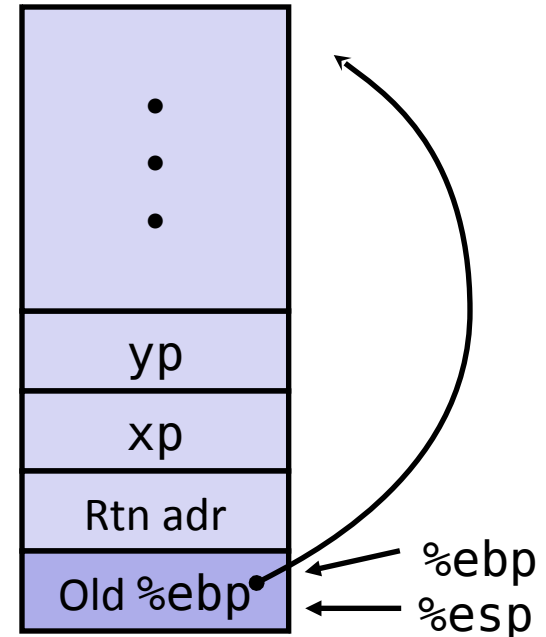
**pushl %ebx**

# swap Setup #2

Entering Stack



Resulting Stack



**swap:**

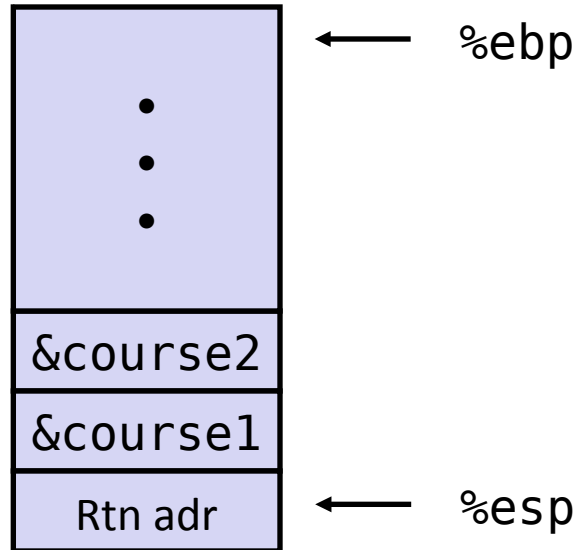
```
pushl %ebp
```

```
movl %esp, %ebp
```

```
pushl %ebx
```

# swap Setup #3

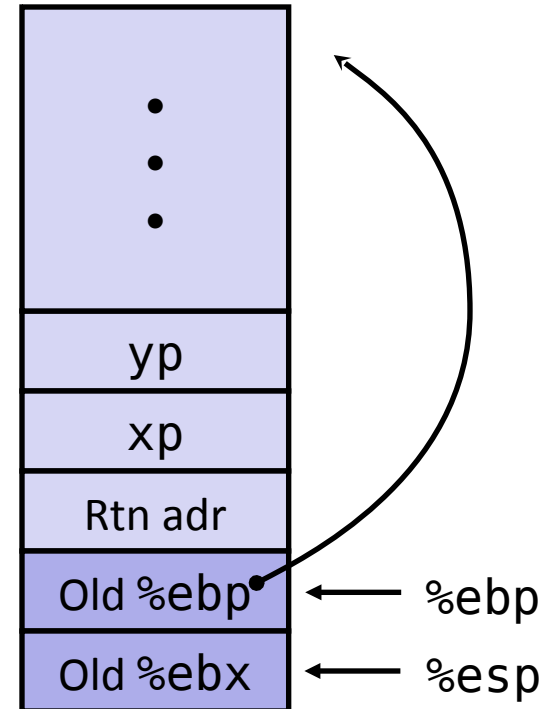
Entering Stack



**swap:**

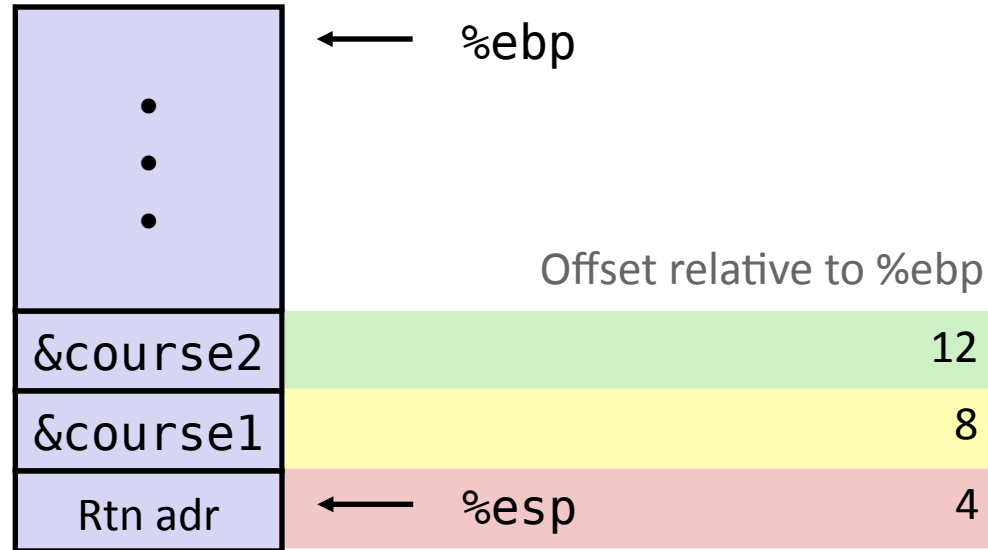
```
pushl %ebp
movl %esp,%ebp
pushl %ebx
```

Resulting Stack

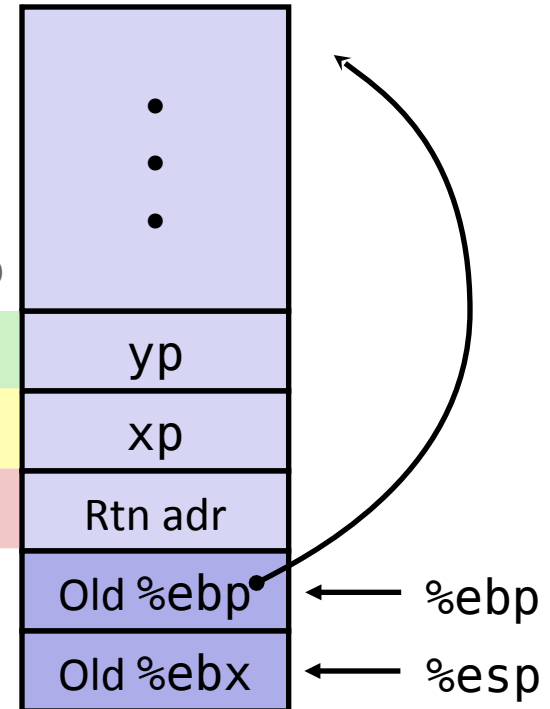


# Swap Body

Entering Stack



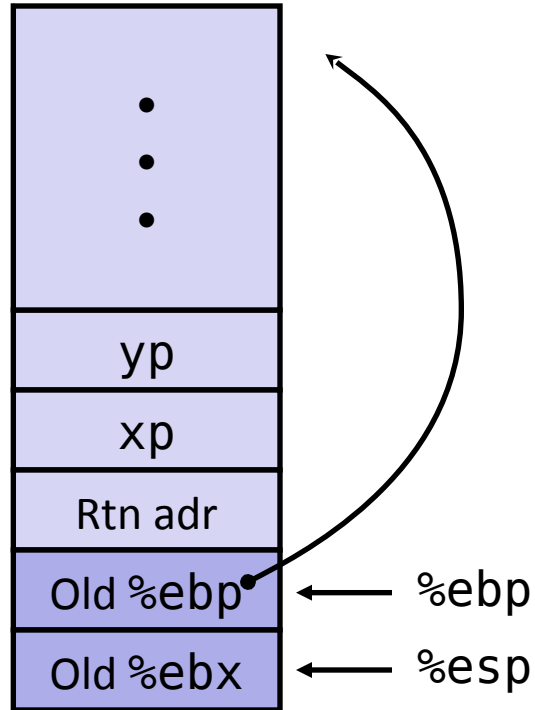
Resulting Stack



```
movl 8(%ebp),%edx    # get xp
movl 12(%ebp),%ecx   # get yp
. . .
```

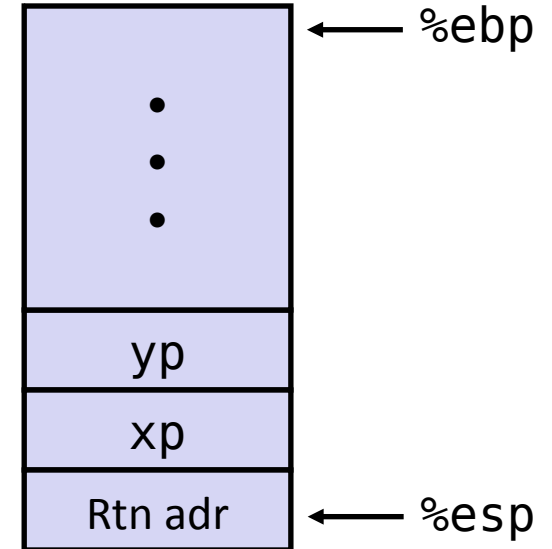
# swap Finish

Stack Before Finish



`popl %ebx`  
`popl %ebp`

Resulting Stack



## ■ Observation

- Saved and restored register %ebx
- Not so for %eax, %ecx, %edx



# Disassembled swap

08048384 <swap>:

<b>8048384:</b>	55	push	%ebp
8048385:	89 e5	mov	%esp, %ebp
8048387:	53	push	%ebx
8048388:	8b 55 08	mov	0x8(%ebp), %edx
804838b:	8b 4d 0c	mov	0xc(%ebp), %ecx
804838e:	8b 1a	mov	(%edx), %ebx
8048390:	8b 01	mov	(%ecx), %eax
8048392:	89 02	mov	%eax, (%edx)
8048394:	89 19	mov	%ebx, (%ecx)
8048396:	5b	pop	%ebx
8048397:	5d	pop	%ebp
8048398:	c3	ret	

## Calling Code

80483b4:	movl	\$0x8049658, 0x4(%esp)	# Copy &course2
80483bc:	movl	\$0x8049654, (%esp)	# Copy &course1
80483c3:	call	<b>8048384</b> <swap>	# Call swap
80483c8:	leave		# Prepare to return
80483c9:	ret		# Return

# Pointer Code

## Generating Pointer

```
/* Compute x + 3 */  
int add3(int x) {  
    int localx = x;  
    incrk(&localx, 3);  
    return localx;  
}
```

## Referencing Pointer

```
/* Increment value by k */  
void incrk(int *ip, int k) {  
    *ip += k;  
}
```

- **add3** creates pointer and passes it to **incrk**

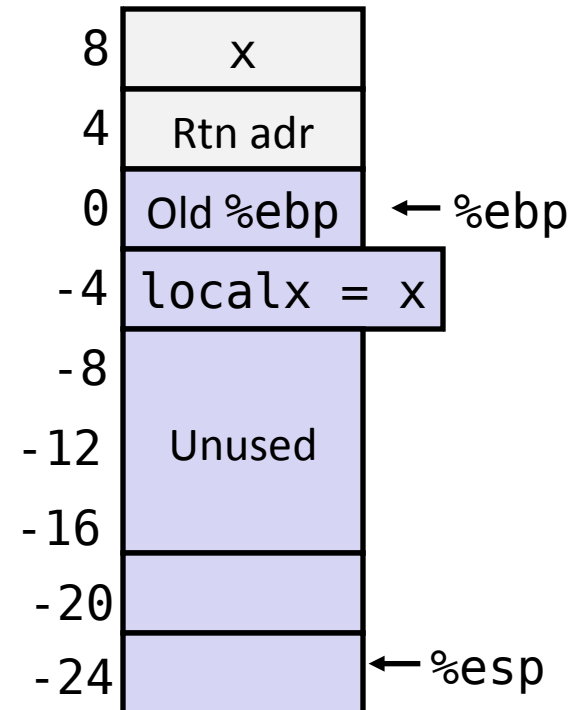
# Creating and Initializing Local Variable

```
int add3(int x) {  
    int localx = x;  
    incr(&localx, 3);  
    return localx;  
}
```

- Variable localx must be stored on stack
  - Because: Need to create pointer to it
- Compute pointer as  $-4(\%ebp)$

First part of add3

```
add3:  
    pushl %ebp  
    movl %esp, %ebp  
    subl $24, %esp      # Alloc. 24 bytes  
    movl 8(%ebp), %eax  
    movl %eax, -4(%ebp) # Set localx to x
```



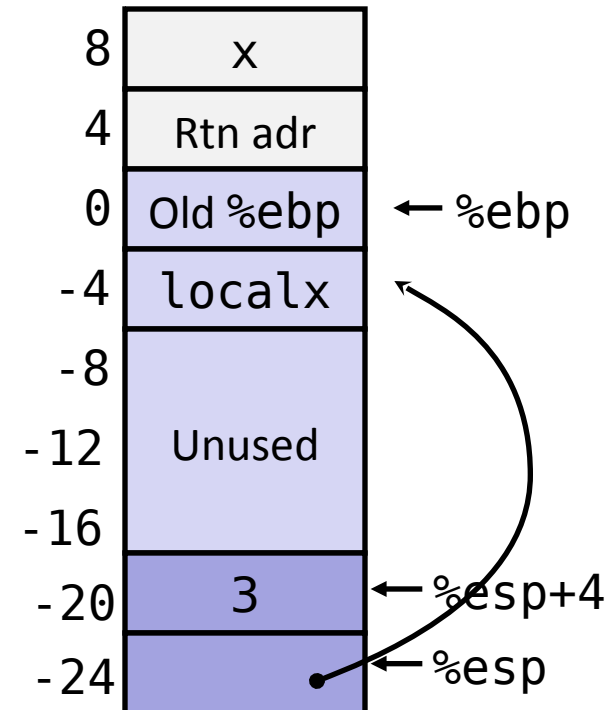
# Creating Pointer as Argument

```
int add3(int x) {  
    int localx = x;  
    incrk(&localx, 3);  
    return localx;  
}
```

- Use leal instruction to compute address of localx

Middle part of add3

```
movl $3, 4(%esp)    # 2nd arg = 3  
leal -4(%ebp), %eax # &localx  
movl %eax, (%esp)   # 1st arg = &localx  
call incrk
```



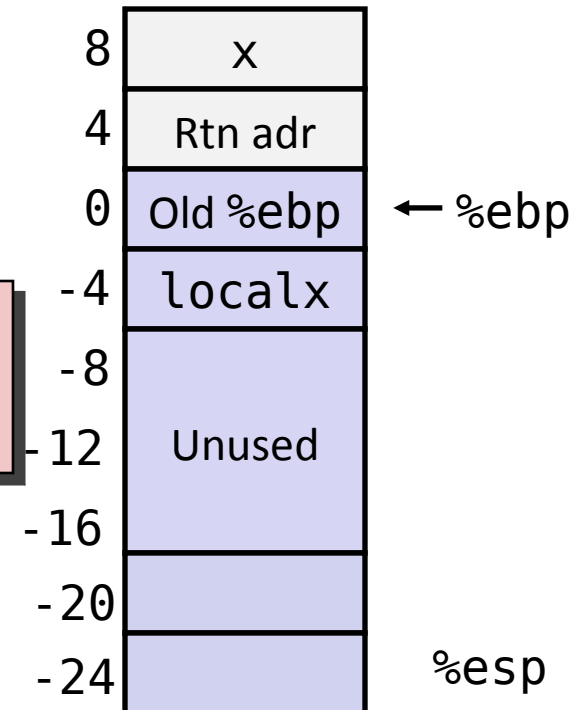
# Retrieving local variable

```
int add3(int x) {  
    int localx = x;  
    incr(&localx, 3);  
    return localx;  
}
```

- Retrieve localx from stack as return value

Final part of add3

```
movl -4(%ebp), %eax # Return val= localx  
leave  
ret
```



# Today

- Switch statements \*
- IA 32 Procedures
  - Stack Structure
  - Calling Conventions
  - Illustrations of Recursion & Pointers
- Arrays
- (Structs)

# Basic Data Types

## ■ Integral

- Stored & operated on in general (integer) registers
- Signed vs. unsigned depends on instructions used

Intel	ASM	Bytes	C
byte	<b>b</b>	1	<b>[unsigned] char</b>
word	<b>w</b>	2	<b>[unsigned] short</b>
double word	<b>l</b>	4	<b>[unsigned] int</b>
quad word	<b>q</b>	8	<b>[unsigned] long int (x86-64)</b>

## ■ Floating Point

- Stored & operated on in floating point registers

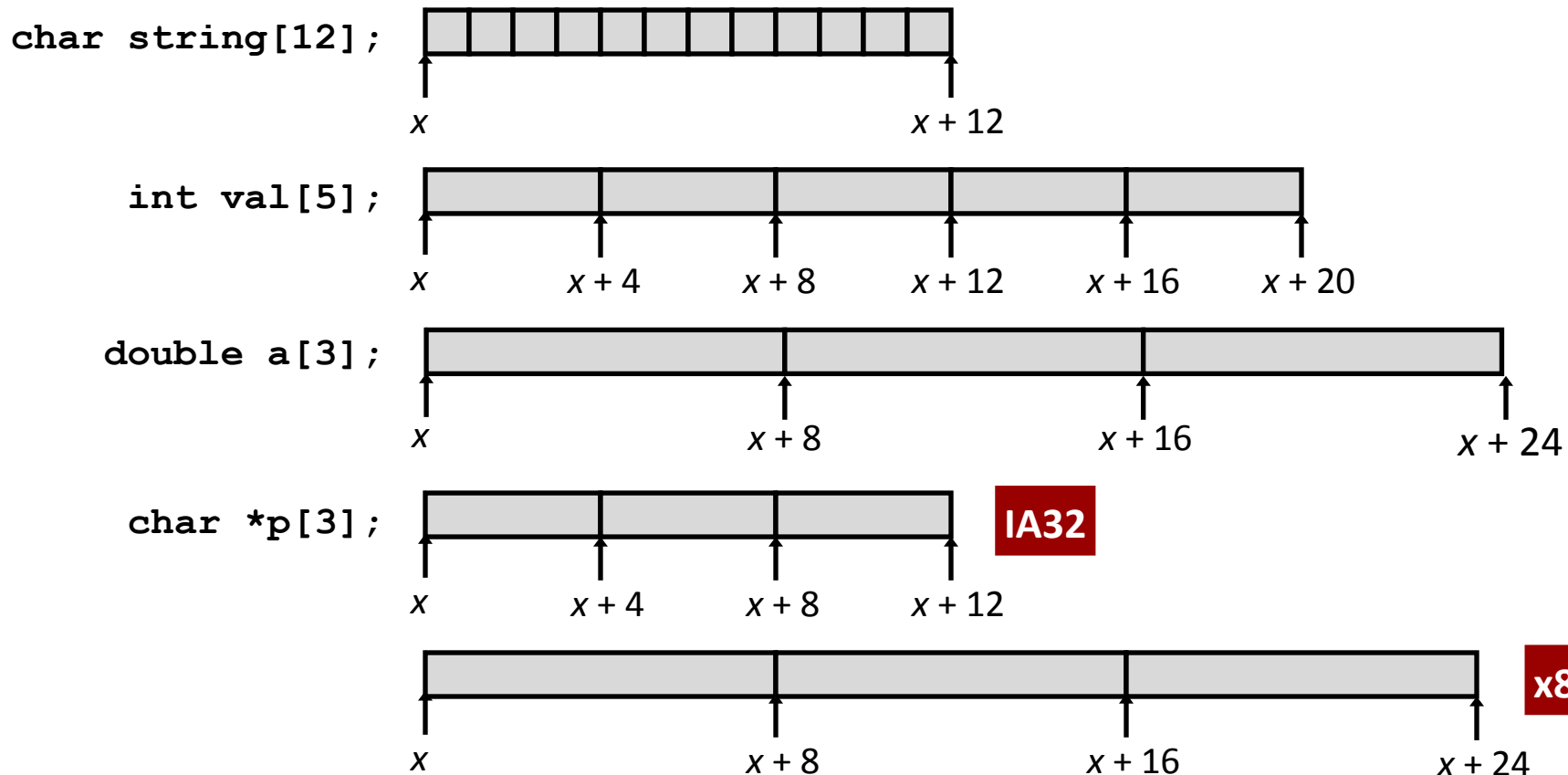
Intel	ASM	Bytes	C
Single	<b>s</b>	4	<b>float</b>
Double	<b>l</b>	8	<b>double</b>
Extended	<b>t</b>	10/12/16	<b>long double</b>

# Array Allocation

## ■ Basic Principle

$T \ A[L];$

- Array of data type  $T$  and length  $L$
- Contiguously allocated region of  $L * \text{sizeof}(T)$  bytes



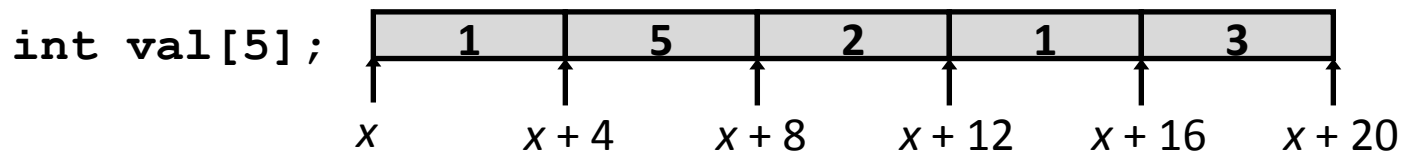


# Array Access

## ■ Basic Principle

$T$  **A**[ $L$ ] ;

- Array of data type  $T$  and length  $L$
- Identifier **A** can be used as a pointer to array element 0: Type  $T^*$



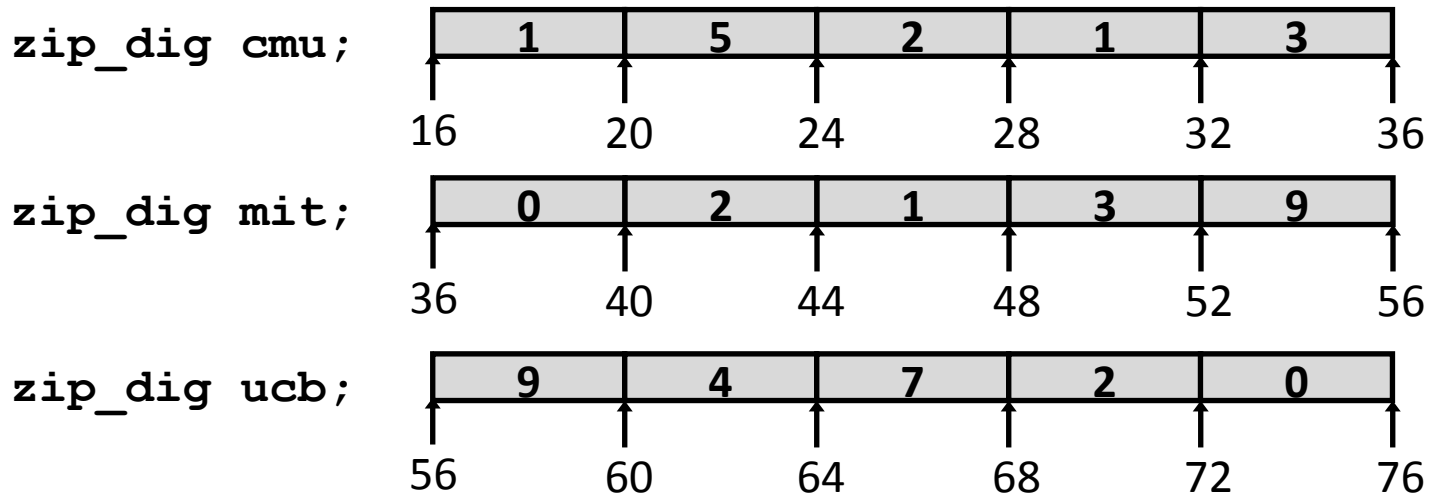
## ■ Reference

	Type	Value
<code>val[4]</code>	<code>int</code>	3
<code>val</code>	<code>int *</code>	$x$
<code>val+1</code>	<code>int *</code>	$x + 4$
<code>&amp;val[2]</code>	<code>int *</code>	$x + 8$
<code>val[5]</code>	<code>int</code>	??
<code>*(val+1)</code>	<code>int</code>	5
<code>val + i</code>	<code>int *</code>	$x + 4 i$

# Array Example

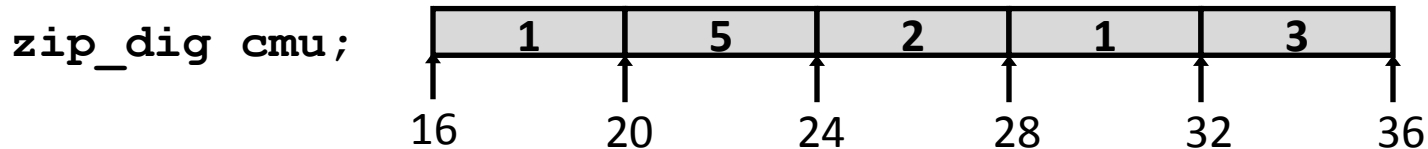
```
#define ZLEN 5
typedef int zip_dig[ZLEN];

zip_dig cmu = { 1, 5, 2, 1, 3 };
zip_dig mit = { 0, 2, 1, 3, 9 };
zip_dig ucb = { 9, 4, 7, 2, 0 };
```



- Declaration “zip\_dig cmu” equivalent to “int cmu[5]”
- Example arrays were allocated in successive 20 byte blocks
  - Not guaranteed to happen in general

# Array Accessing Example



```
int get_digit
  (zip_dig z, int dig)
{
  return z[dig];
}
```

## IA32

```
# %edx = z
# %eax = dig
movl (%edx,%eax,4),%eax # z[dig]
```

- Register `%edx` contains starting address of array
- Register `%eax` contains array index
- Desired digit at  $4 * \%eax + \%edx$
- Use memory reference  $(\%edx, \%eax, 4)$

# Array Loop Example (IA32)

```
void zincr(zip_dig z) {  
    int i;  
    for (i = 0; i < ZLEN; i++)  
        z[i]++;  
}
```

```
# edx = z  
movl    $0, %eax           # %eax = i  
.L4:      # loop:  
addl    $1, (%edx,%eax,4)  # z[i]++  
addl    $1, %eax           # i++  
cmpl    $5, %eax           # i:5  
jne     .L4                # if !=, goto loop
```

# Pointer Loop Example (IA32)

```
void zincr_p(zip_dig z) {  
    int *zend = z+ZLEN;  
    do {  
        (*z)++;  
        z++;  
    } while (z != zend);  
}
```

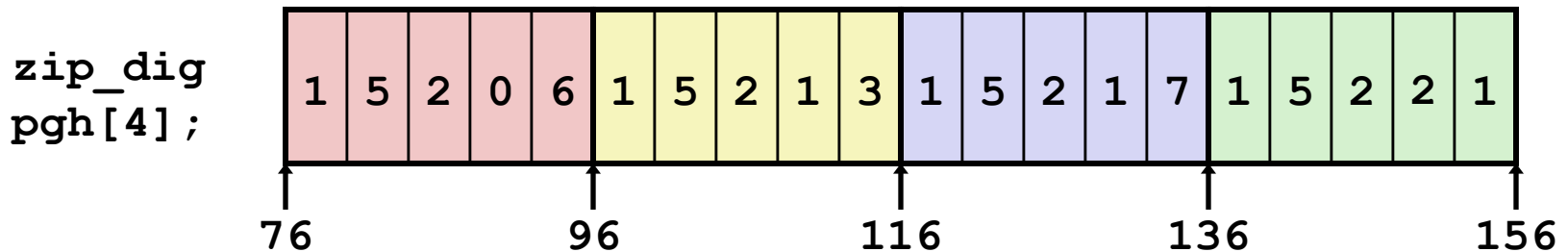


```
void zincr_v(zip_dig z) {  
    void *vz = z;  
    int i = 0;  
    do {  
        (*(int *) (vz+i))++;  
        i += ISIZE;  
    } while (i != ISIZE*ZLEN);  
}
```

# edx = z = vz	
movl \$0, %eax	# i = 0
.L8:	# loop:
addl \$1, (%edx,%eax)	# Increment vz+i
addl \$4, %eax	# i += 4
cmpl \$20, %eax	# Compare i:20
jne .L8	# if !=, goto loop

# Nested Array Example

```
#define PCOUNT 4
zip_dig pgh[PCOUNT] =
    {{1, 5, 2, 0, 6},
     {1, 5, 2, 1, 3 },
     {1, 5, 2, 1, 7 },
     {1, 5, 2, 2, 1 }};
```



- **“zip\_dig pgh[4]” equivalent to “int pgh[4][5]”**
  - Variable **pgh**: array of 4 elements, allocated contiguously
  - Each element is an array of 5 **int**’s, allocated contiguously
- **“Row-Major” ordering of all elements guaranteed**

# Multidimensional (Nested) Arrays

## ■ Declaration

`T A[R][C];`

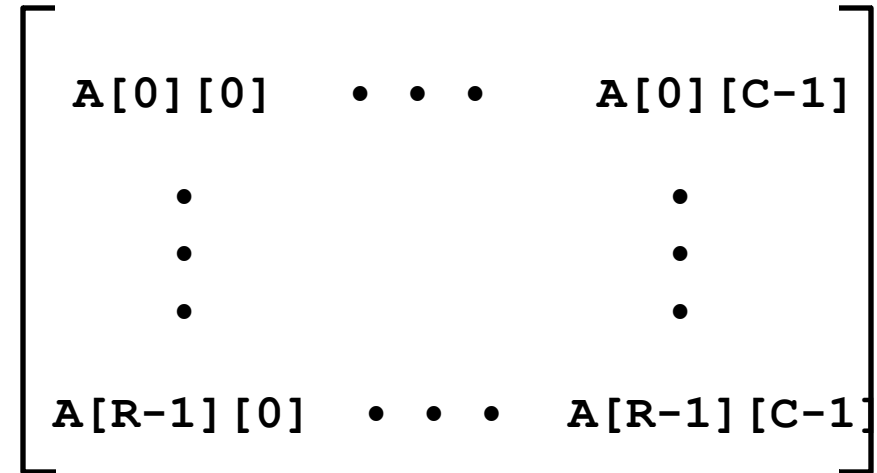
- 2D array of data type  $T$
- $R$  rows,  $C$  columns
- Type  $T$  element requires  $K$  bytes

## ■ Array Size

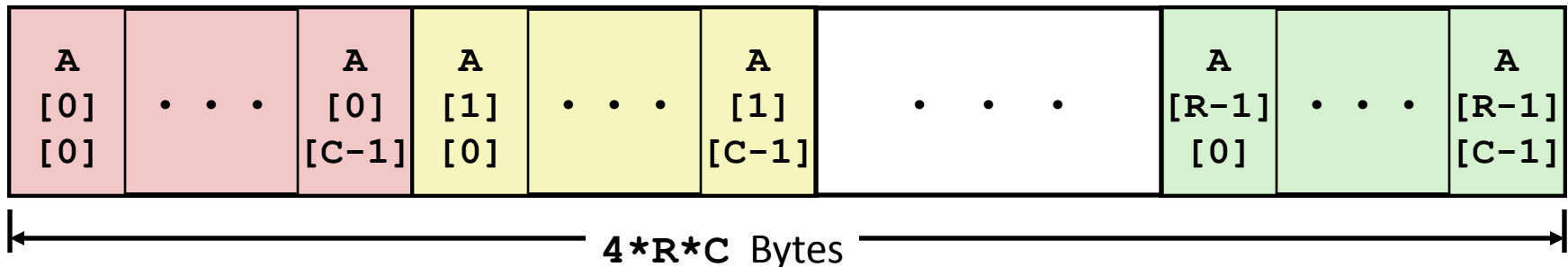
- $R * C * K$  bytes

## ■ Arrangement

- Row-Major Ordering



`int A[R][C];`

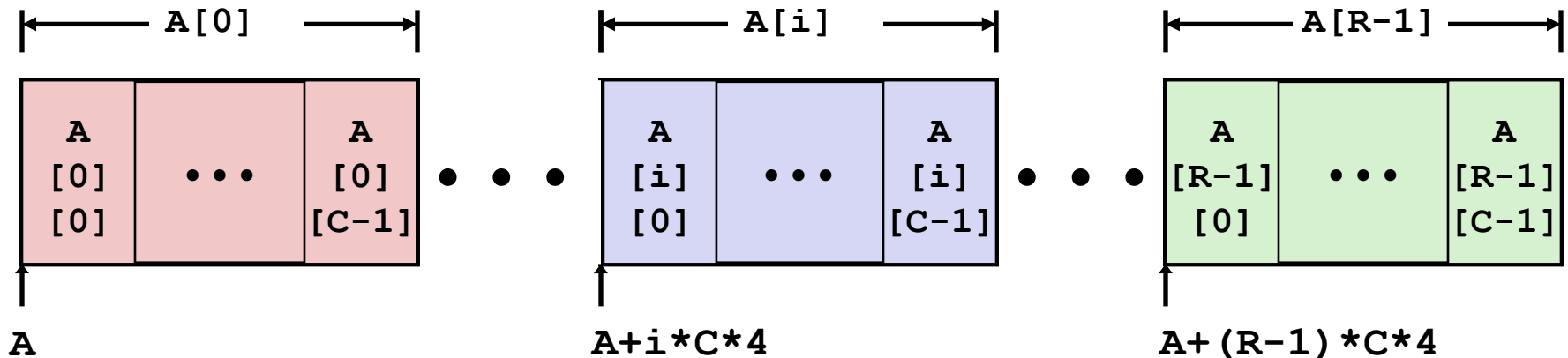


# Nested Array Row Access

## ■ Row Vectors

- $A[i]$  is array of  $C$  elements
- Each element of type  $T$  requires  $K$  bytes
- Starting address  $A + i * (C * K)$

```
int A[R][C];
```





# Nested Array Row Access Code

```
int *get_pgh_zip(int index)
{
    return pgh[index];
}
```

```
#define PCOUNT 4
zip_dig pgh[PCOUNT] =
    {{1, 5, 2, 0, 6},
     {1, 5, 2, 1, 3 },
     {1, 5, 2, 1, 7 },
     {1, 5, 2, 2, 1 }};
```

```
# %eax = index
leal (%eax,%eax,4),%eax # 5 * index
leal pgh(,%eax,4),%eax # pgh + (20 * index)
```

## ■ Row Vector

- `pgh[index]` is array of 5 `int`'s
- Starting address `pgh+20*index`

## ■ IA32 Code

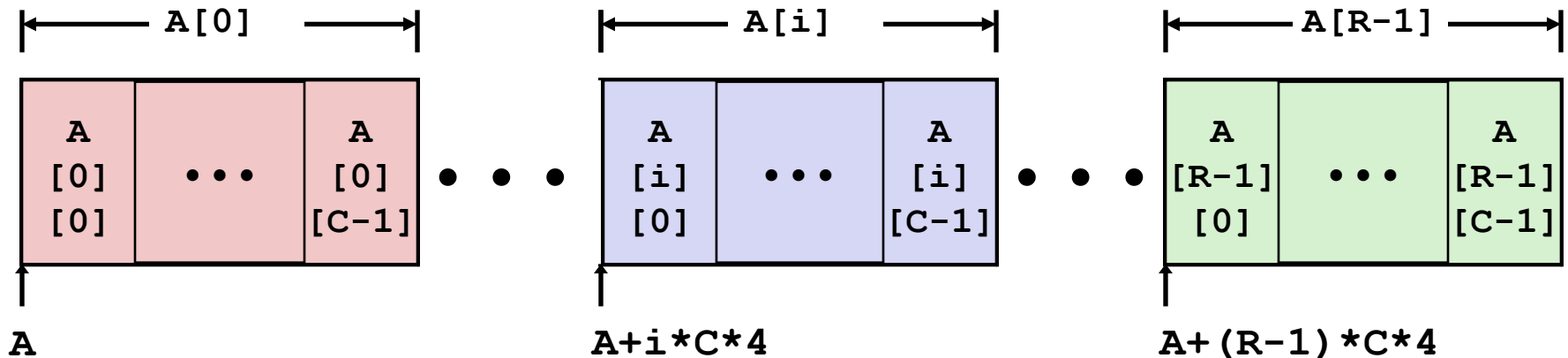
- Computes and returns address
- Compute as `pgh + 4*(index+4*index)`

# Nested Array Row Access

## ■ Row Vectors

- $A[i]$  is array of  $C$  elements
- Each element of type  $T$  requires  $K$  bytes
- Starting address  $A + i * (C * K)$

```
int A[R][C];
```



# Nested Array Element Access Code

```
int get_pgh_digit
(int index, int dig)
{
    return pgh[index][dig];
}
```

```
movl    8(%ebp), %eax        # index
leal    (%eax,%eax,4), %eax   # 5*index
addl    12(%ebp), %eax        # 5*index+dig
movl    pgh(,%eax,4), %eax    # offset 4*(5*index+dig)
```

## ■ Array Elements

- `pgh[index][dig]` is `int`
- Address: `pgh + 20*index + 4*dig`
  - `= pgh + 4*(5*index + dig)`

## ■ IA32 Code

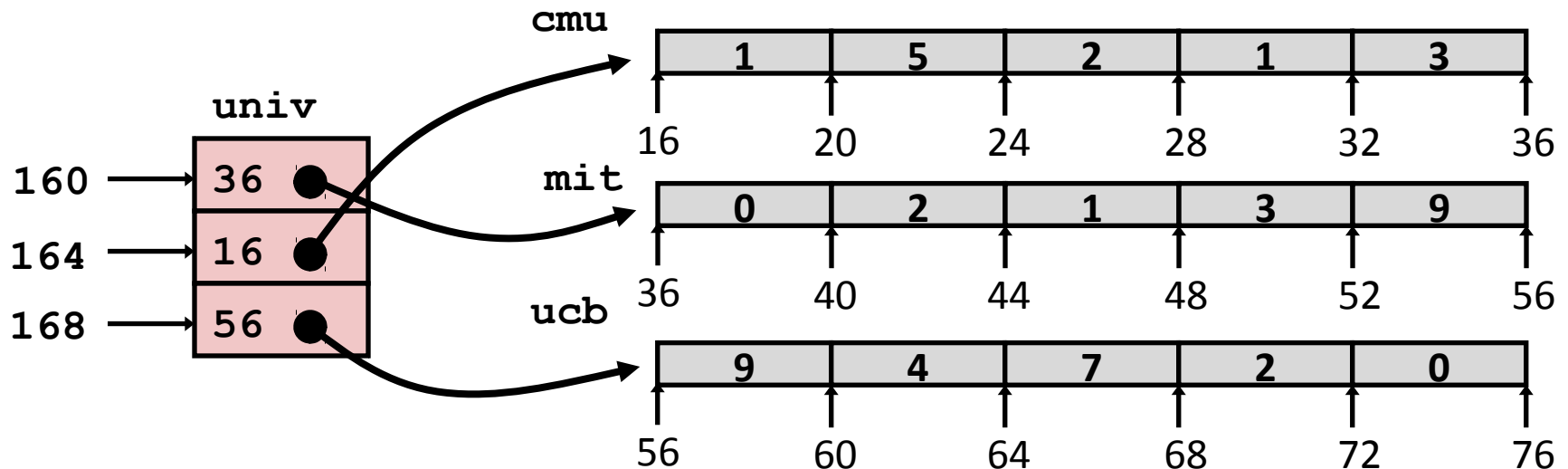
- Computes address `pgh + 4*((index+4*index)+dig)`

# Multi-Level Array Example

```
zip_dig cmu = { 1, 5, 2, 1, 3 };  
zip_dig mit = { 0, 2, 1, 3, 9 };  
zip_dig ucb = { 9, 4, 7, 2, 0 };
```

```
#define UCOUNT 3  
int *univ[UCOUNT] = {mit, cmu, ucb};
```

- Variable `univ` denotes array of 3 elements
- Each element is a pointer
  - 4 bytes
- Each pointer points to array of `int`'s



# Element Access in Multi-Level Array

```
int get_univ_digit
(int index, int dig)
{
    return univ[index][dig];
}
```

```
movl    8(%ebp), %eax        # index
movl    univ(,%eax,4), %edx   # p = univ[index]
movl    12(%ebp), %eax       # dig
movl    (%edx,%eax,4), %eax   # p[dig]
```

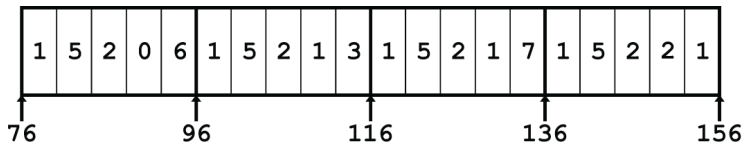
## ■ Computation (IA32)

- Element access **Mem[Mem[univ+4\*index]+4\*dig]**
- Must do two memory reads
  - First get pointer to row array
  - Then access element within array

# Array Element Accesses - Comparison

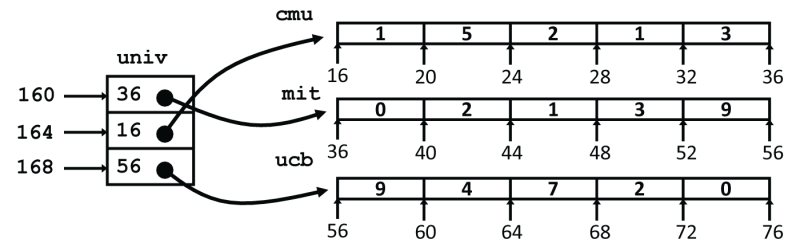
## Nested array

```
int get_pgh_digit
(int index, int dig)
{
    return pgh[index][dig];
}
```



## Multi-level array

```
int get_univ_digit
(int index, int dig)
{
    return univ[index][dig];
}
```



Accesses looks similar in C, but addresses very different:

`Mem[pgh+20*index+4*dig]`

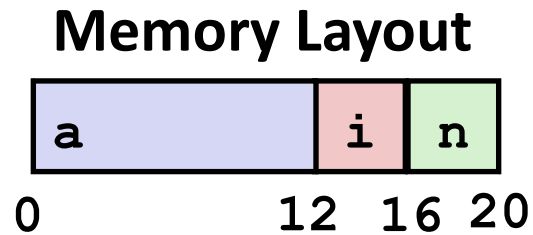
`Mem[Mem[univ+4*index]+4*dig]`

# Today

- Switch statements \*
- IA 32 Procedures
  - Stack Structure
  - Calling Conventions
  - Illustrations of Recursion & Pointers
- Arrays
- (Structs)

# Structure Allocation

```
struct rec {  
    int a[3];  
    int i;  
    struct rec *n;  
};
```



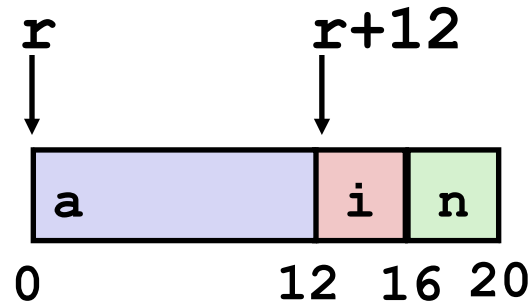
## ■ Concept

- Contiguously-allocated region of memory
- Refer to members within structure by names
- Members may be of different types



# Structure Access

```
struct rec {  
    int a[3];  
    int i;  
    struct rec *n;  
};
```



## ■ Accessing Structure Member

- Pointer indicates first byte of structure
- Access elements with offsets

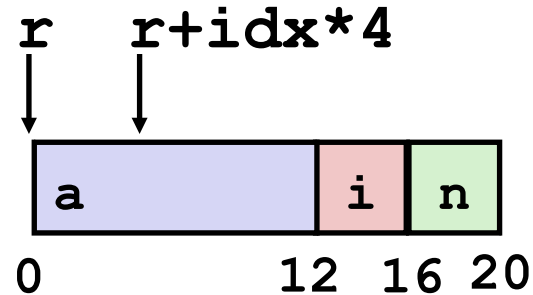
```
void  
set_i(struct rec *r,  
      int val)  
{  
    r->i = val;  
}
```

## IA32 Assembly

```
# %edx = val  
# %eax = r  
movl %edx, 12(%eax) # Mem[r+12] = val
```

# Generating Pointer to Structure Member

```
struct rec {  
    int a[3];  
    int i;  
    struct rec *n;  
};
```



## ■ Generating Pointer to Array Element

- Offset of each structure member determined at compile time
- Arguments
  - `Mem[%ebp+8]: r`
  - `Mem[%ebp+12]: idx`

```
int *get_ap  
(struct rec *r, int idx)  
{  
    return &r->a[idx];  
}
```

```
movl    12(%ebp), %eax    # Get idx  
sall    $2, %eax         # idx*4  
addl    8(%ebp), %eax     # r+idx*4
```

# IA 32 Procedure Summary

## ■ Important Points

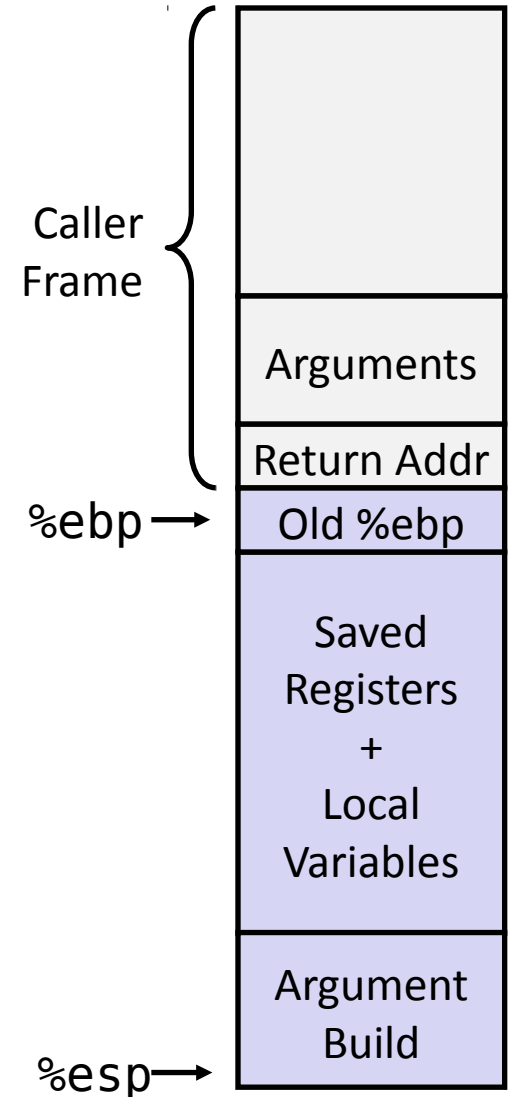
- Stack is the right data structure for procedure call / return
  - If P calls Q, then Q returns before P

## ■ Recursion (& mutual recursion) handled by normal calling conventions

- Can safely store values in local stack frame and in callee-saved registers
- Put function arguments at top of stack
- Result return in %eax

## ■ Pointers are addresses of values

- On stack or global



# Array, Structs Summary

## ■ Arrays

- One-dimensional
- Multi-dimensional (nested)
- Multi-level
- You can count on your compiler!

## ■ Structures

- Allocation
- Access