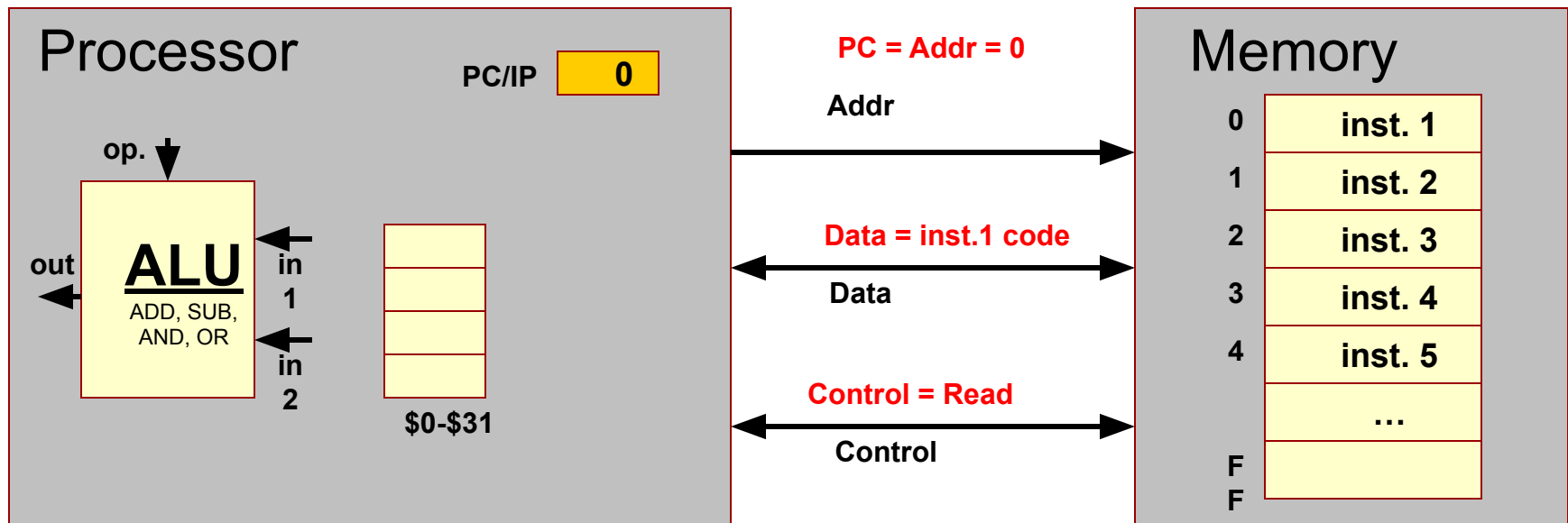


# CS356 Unit 6

x86 Procedures  
Basic Stack Frames

# Review of Program Counter (IP register)

- PC/IP is used to fetch an instruction
  - PC/IP contains the address of the next instruction
  - The value in the PC/IP is placed on the **address bus** and the memory is told to read with a signal on the **control bus**
  - **PC/IP is incremented**
  - The process is repeated for the next instruction



# Procedures (Subroutines)

CS:APP 3.7.1

- Procedures (aka subroutines or functions) are reusable sections of code that we can call from some location, execute that procedure, and then **return to where we left off**

C code:

```
int main() {  
  
    ...  
    x = 8;  
    res = avg(x,4);  
    printf("%d\n", res);  
}  
  
int avg(int a, int b){  
    return (a+b)/2;  
}
```

A procedure to  
calculate the average  
of 2 numbers

We call the  
procedure to  
calculate the average  
and when it is  
finished it will return  
to where we left off

# Procedures

- Procedure calls are similar to 'jump' instructions where we go to a new location in the code

C code:

```
int main() {  
  
    ...  
    x = 8;  
    res = avg(x,4);  
    printf("%d\n", res);  
}  
  
int avg(int a, int b){  
    return (a+b)/2;  
}
```

① Call "avg" procedure will require us to jump to that code

# Normal Jumps vs. Procedures

- Difference between normal jumps and procedure calls is that **with procedures we have to return to where we left off**
- We need to **leave a link** to the return location before we jump to the procedure...

C code:

```
int main() {  
  
    ...  
    x = 8;  
    res = avg(x,4);  
    printf("%d\n", res);  
}  
  
int avg(int a, int b){  
    return (a+b)/2;  
}
```

After procedure completes, return to the statement in the main code where we left off ②

Call "avg" procedure will require us to jump to that code ①

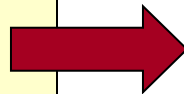
# Implementing Procedures

- To implement procedures in assembly we need to be able to:
  - Jump to the procedure code, leaving a "return link" (i.e. return address) to know where to return
  - Find the return address and go back to that location

C code:

```
...  
Call res = avg(x,4);  
...  
  
Definition  
int avg(int a, int b)  
{ return (a+b)/2; }
```

Desired return  
location

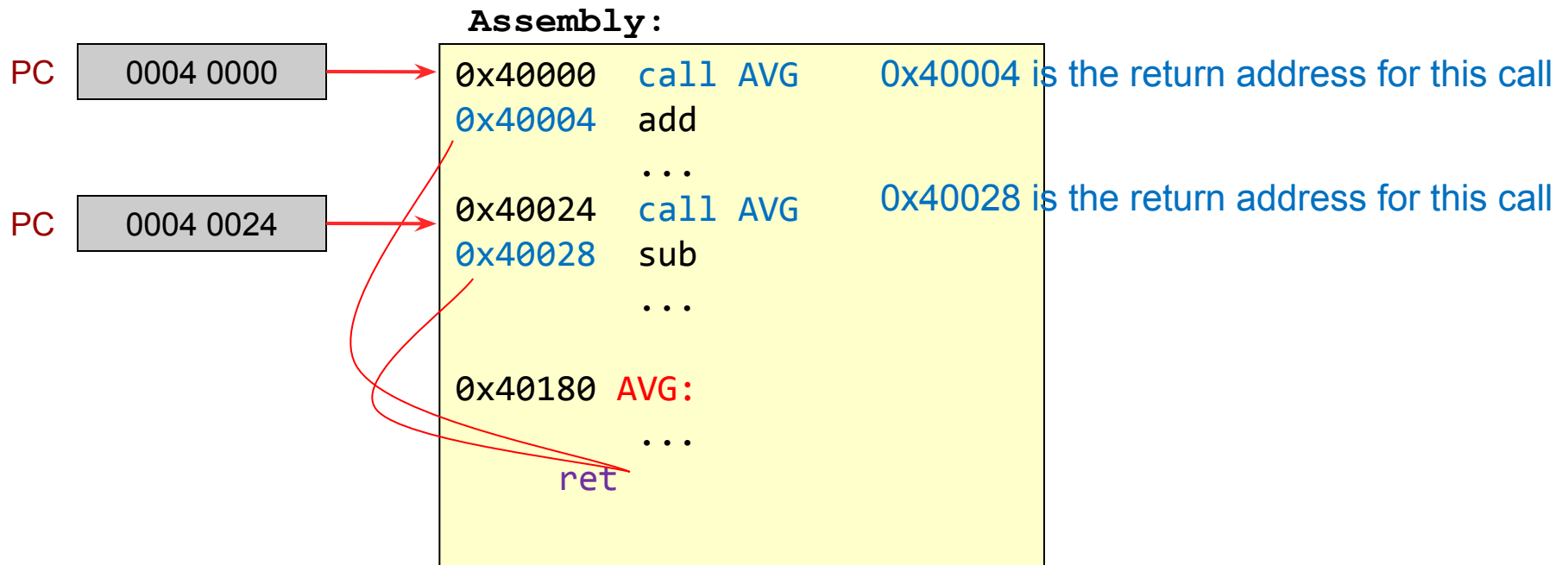


Assembly:

```
113b    callq avg    # save a link  
1140    next inst. # to next instruc.  
  
avg:  
1125    addl %edi,%esi  
1127    movl %esi,%eax  
1129    shr1 $0x1f,%eax  
112c    addl %esi,%eax  
112e    sarl %eax  
1130    retl
```

# Return Addresses

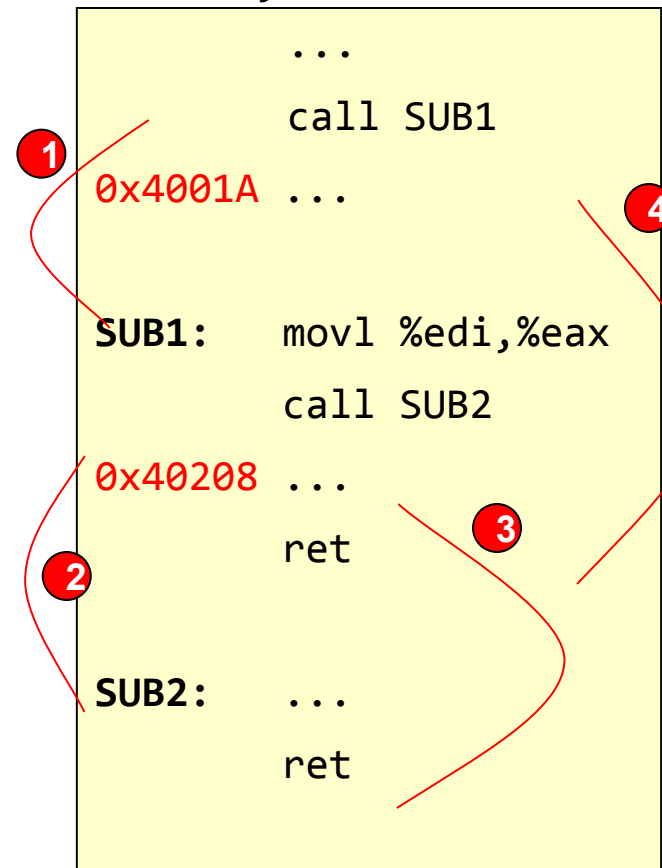
- When **calling** a procedure, the address to jump to is ALWAYS the same
- The location where a procedure **returns** will vary
  - Always the address of the instruction after the 'call'



# Return Addresses

- A further (very common) complication is nested procedure calls
  - One procedure calls another
- Example: Main routine calls SUB1 which calls SUB2
- Must store both return addresses but where?
  - Registers? No...very limited number
  - Memory? Yes...usually enough memory for deep levels of nesting

Assembly:

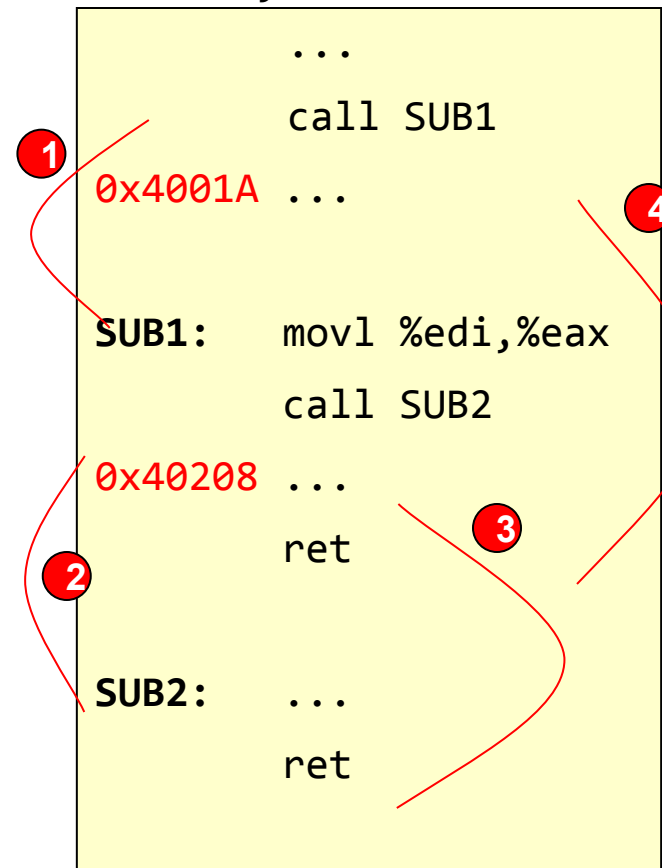




# Return Addresses and Stacks

- Note: Return addresses will be accessed in reverse order as they are stored
  - 0x40208 is the second RA to be stored but should be the first one used to return
- A stack structure is appropriate!
- The system stack will be a place where we can store
  - Return addresses and other saved register values
  - Local variables of a function
  - Arguments for procedures

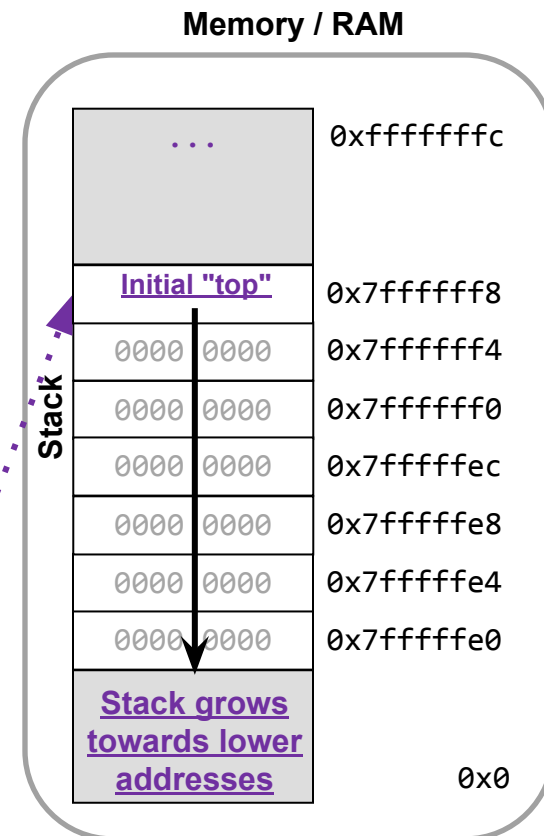
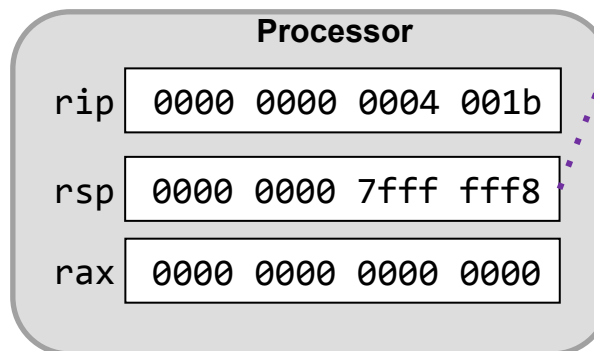
Assembly:



# System Stack

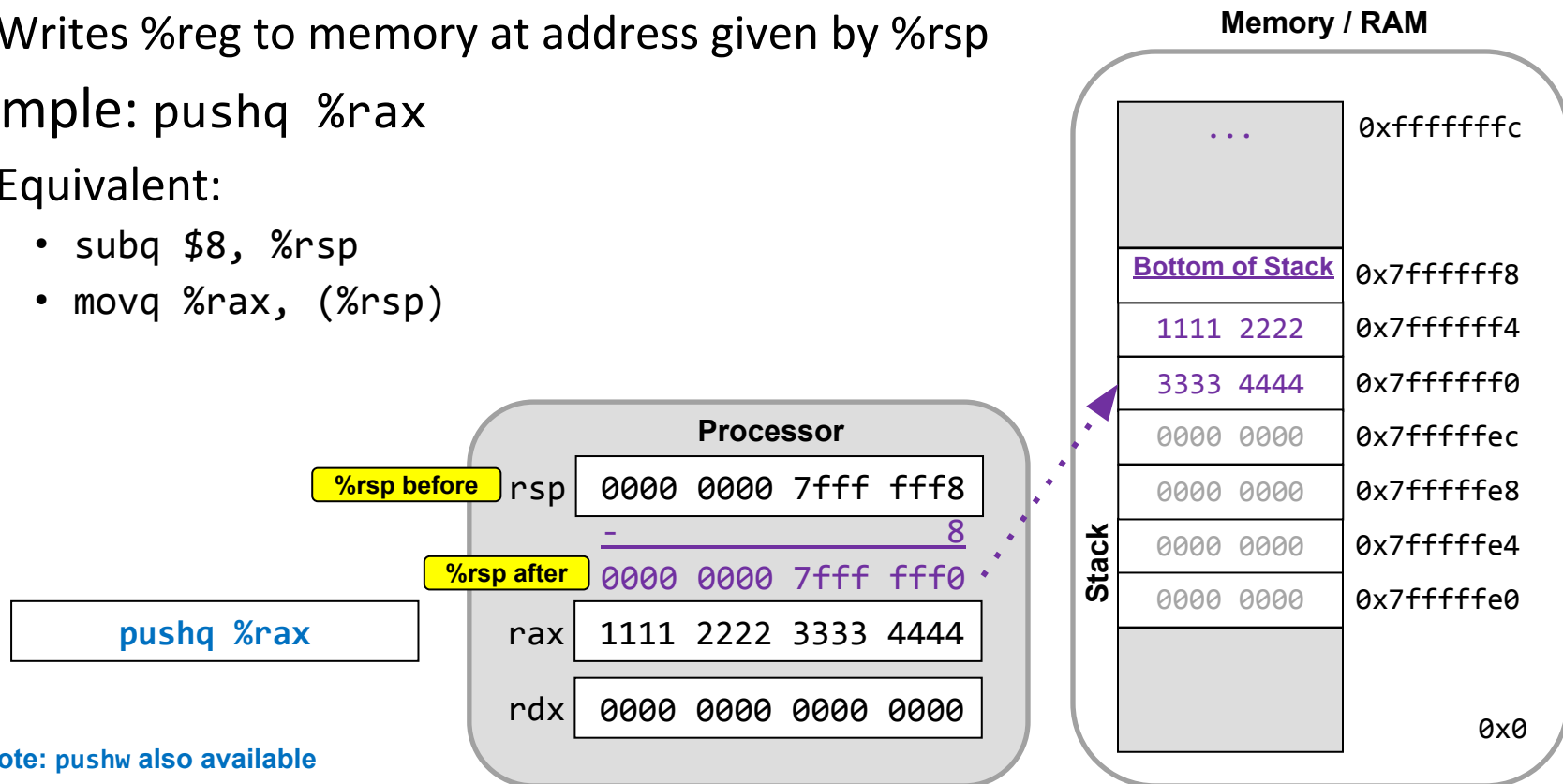
- Stack is a data structure where data is accessed in reverse order as it is stored (a.k.a. LIFO = Last-in First-out)
- Use a stack to store the return addresses and other data
- System stack defined as growing towards smaller addresses
  - Usually starts around  $\frac{1}{2}$  to  $\frac{3}{4}$  of the way through the address space (i.e. for a 32-bit somewhere around 0x7fff... or 0xbfff...)
- Top of stack is accessed and maintained using %rsp (stack pointer) register
  - %rsp points at top **occupied** location of the stack

Stack Pointer  
Always points to top occupied element of the stack



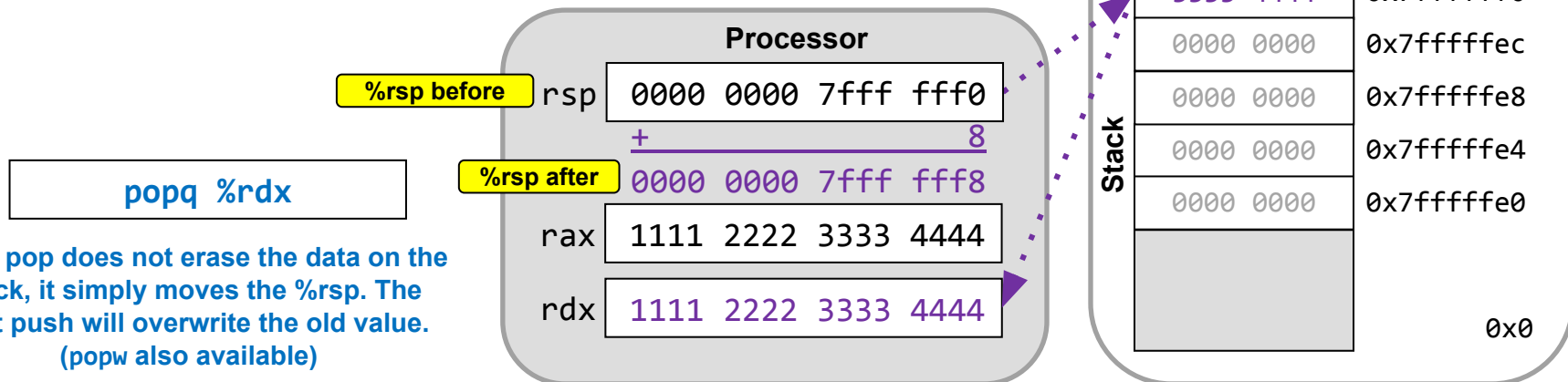
# Push Operation and Instruction

- Push operation adds data to system stack
- Format: **pushq %reg**
  - Decrements %rsp by 8
  - Writes %reg to memory at address given by %rsp
- Example: **pushq %rax**
  - Equivalent:
    - `subq $8, %rsp`
    - `movq %rax, (%rsp)`



# Pop Operation and Instruction

- Pop operation removes data from system stack
- Format: `popq %reg`
  - Reads memory at address given by `%rsp` and places value into `%reg`
  - Increments `%rsp` by 8
- Example: `popq %rdx`
  - Equivalent:
    - `movq (%rsp), %rdx`
    - `addq $8, %rsp`



# Jumping to a Procedure

CS:APP 3.7.2

- Format:
  - `call label`
  - `call *operand [e.g. call (%rax)]`
- Operations:
  - Pushes the address of next instruction (i.e. return address (RA) ) onto the stack
    - Implicitly performs `subq $8,%rsp` and `movq %rip,(%rsp)`
  - Updates the PC to go to the start of the desired procedure [i.e. `PC = addr`]
    - `addr` is the address you want to branch to (*usually specified as a label*)

# Returning From a Procedure

- Format:
  - `ret`
- Operations:
  - Pops the return address from the stack into `%rip`  
[i.e. `PC = return-address`]
  - Implicitly performs `movq (%rsp), %rip` and `addq $8, %rsp`

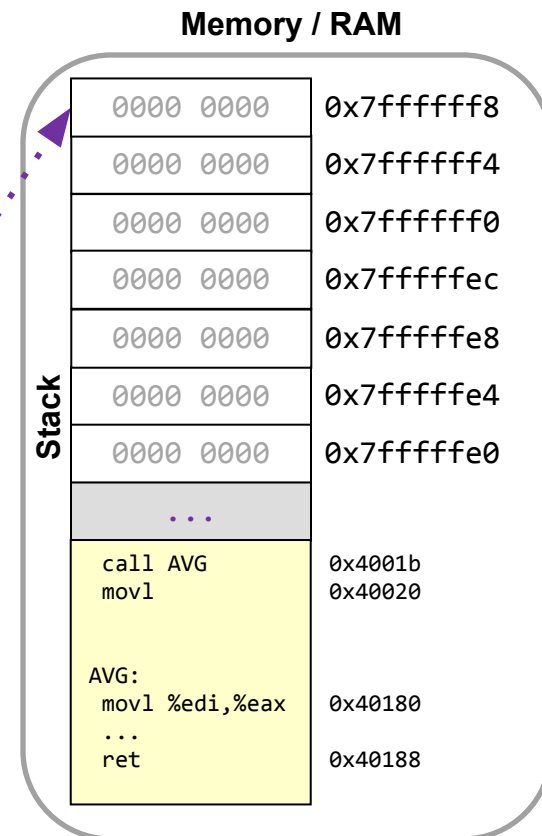
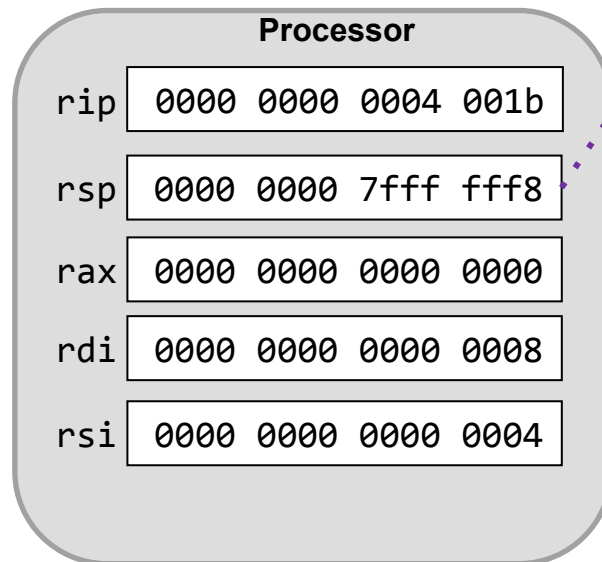
# Procedure Call Sequence 1a

- Initial conditions
  - About to execute the 'call' instruction
  - Current top of stack is at 0x7fffffff8

```

...
call AVG
movl %eax, (%rbp)
...

AVG:
movl %edi, %eax
...
ret
    
```



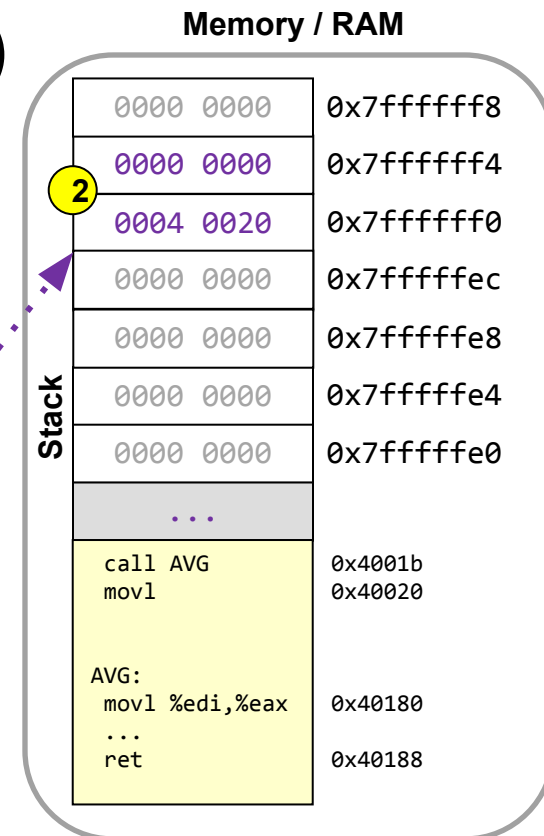
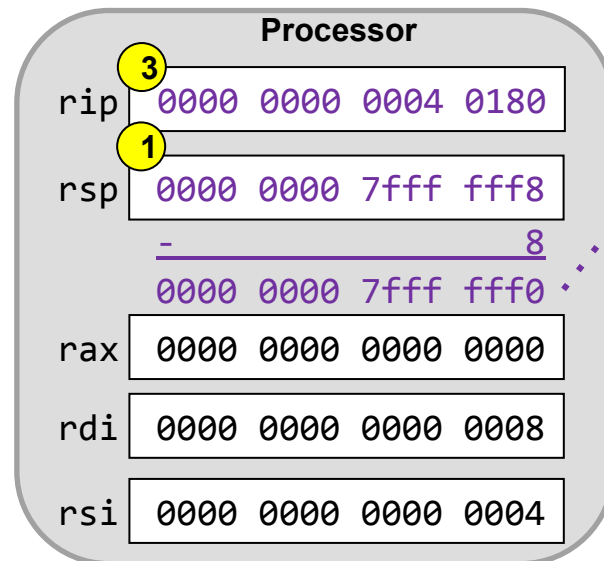
# Procedure Call Sequence 1b

- call Operation (i.e. push return address) & jump
  - Decrement stack pointer (\$rsp) and push RA (0x40020) onto stack (as 64-bit address)
  - Update PC to start of procedure (0x40180)

```

...
call AVG
movl %eax, (%rbp)
...

AVG:
movl %edi, %eax
...
ret
    
```





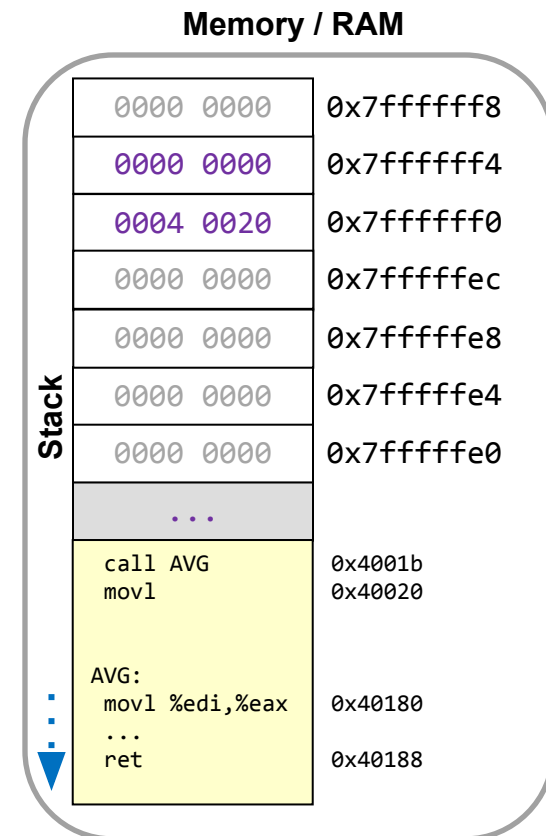
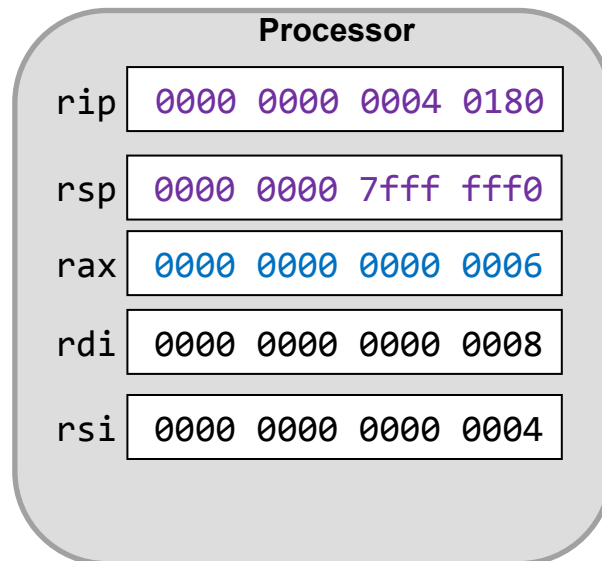
# Procedure Call Sequence 1c

- Execute the code for the procedure
- Return value should be in %rax/%eax

```

...
call AVG
movl %eax, (%rbp)
...

AVG:
movl %edi,%eax
...
ret
    
```



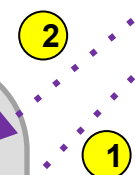
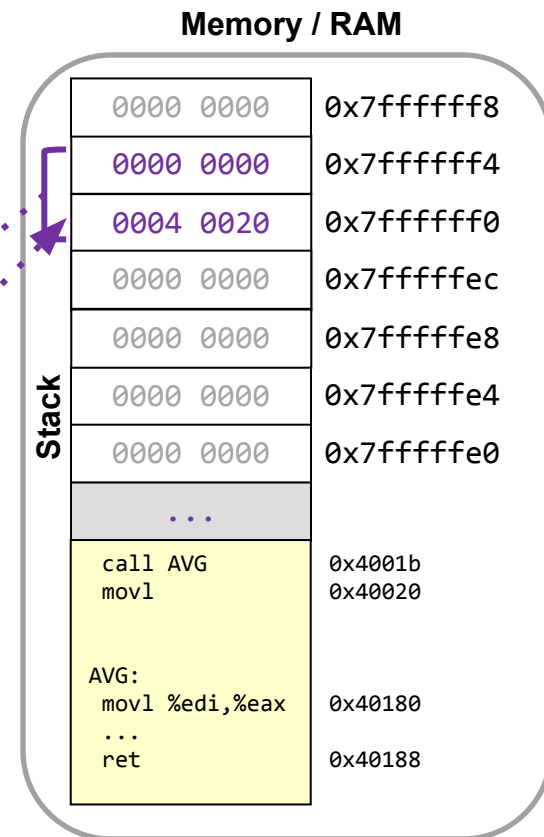
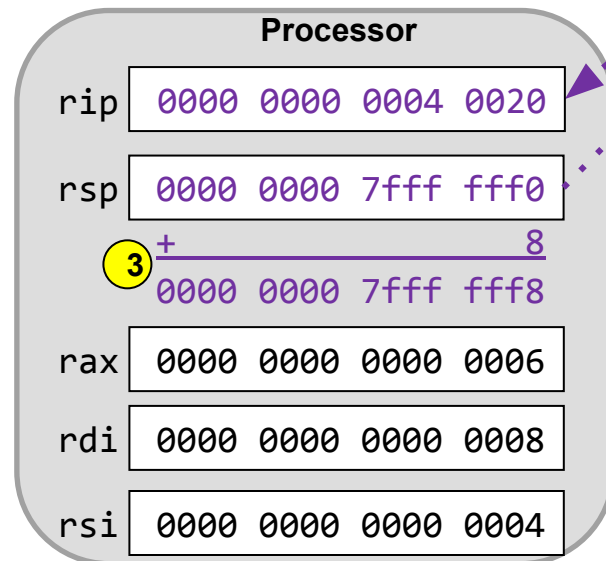
# Procedure Call Sequence 1d

- ret Operation (i.e. pop return address)
  - Retrieve RA (0x40020) from stack
  - Put it in the PC
  - Increment the stack pointer (%rsp)

```

...
call AVG
movl %eax, (%rbp)
...

AVG:
movl %edi, %eax
...
ret
    
```



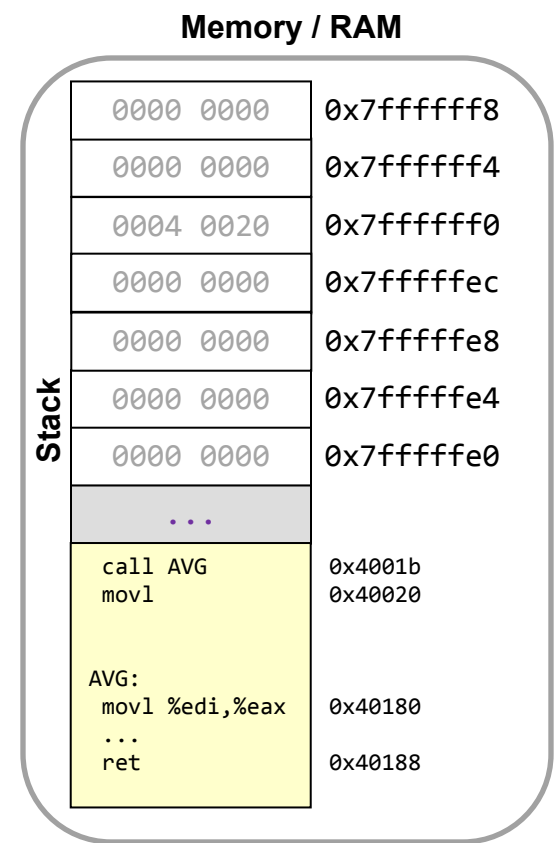
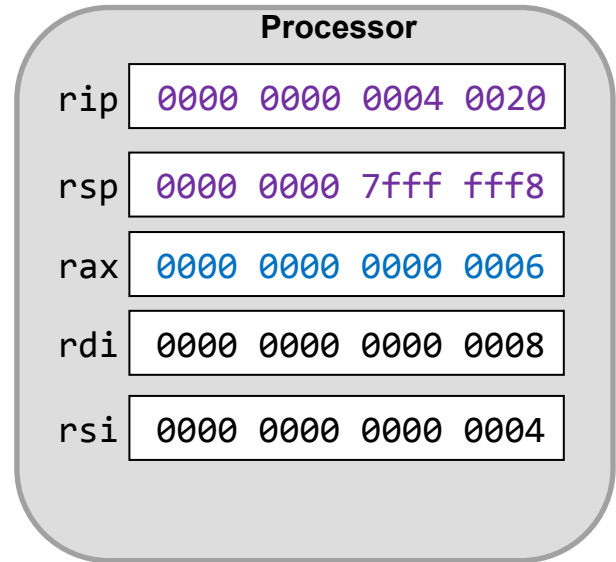
# Procedure Call Sequence 1e

- Execution resumes after the procedure call

```

...
call AVG
1 movl %eax, (%rbp)
...

AVG:
movl %edi,%eax
...
ret
    
```



# Procedure Call Sequence 2

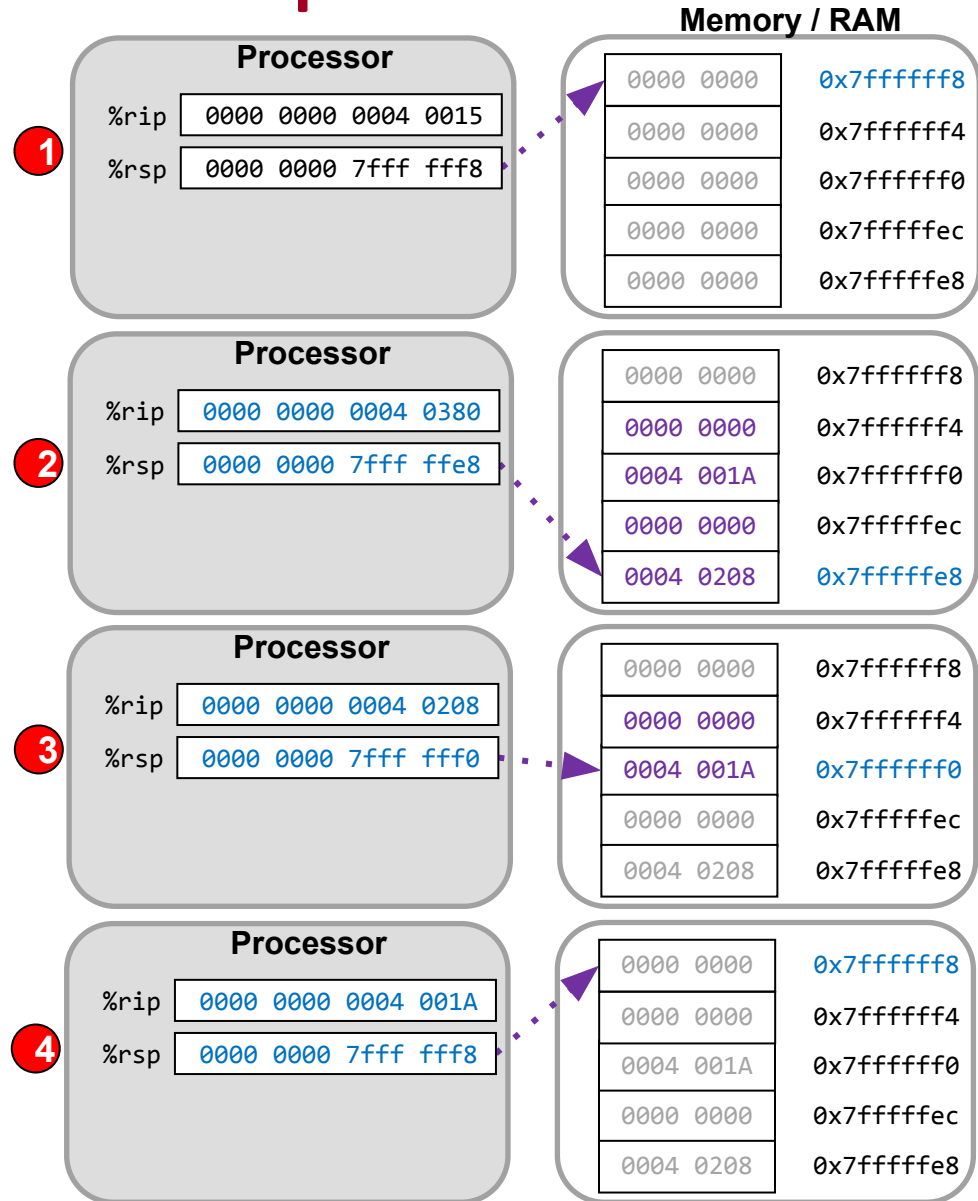
- Show the values of the stack, %rsp, and %rip at the various timestamps for the following code

```

    ...
1 0x40015 call SUB1
4 0x4001A ...

0x40200
SUB1:  movl %edi,%eax
      call SUB2
3 0x40208 ...
      ret

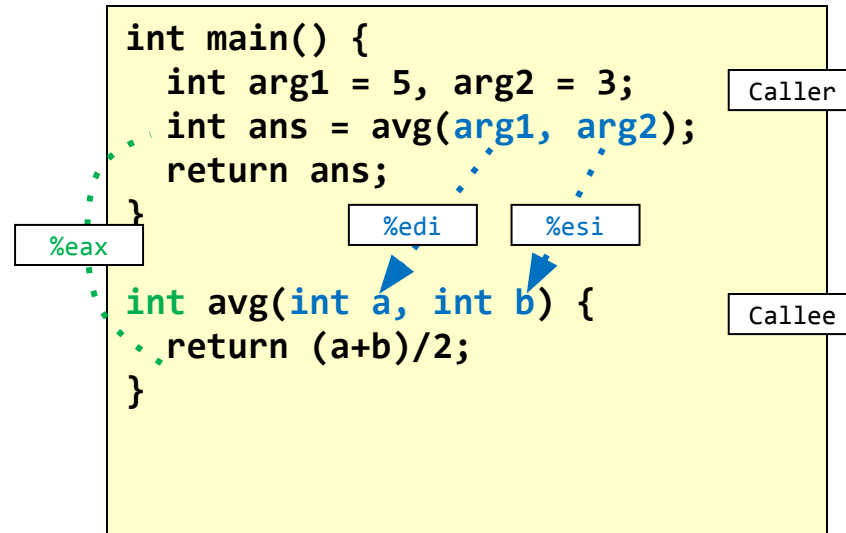
0x40380
2 SUB2:  ...
      ret
    
```



# Arguments and Return Values

CS:APP 3.7.3

- Most procedure calls pass arguments/parameters to the procedure and it often produces return values
- To implement this, there must be locations agreed upon by caller and callee for where this information will be found
- x86-64 convention is to use certain registers for this task (**see table**)



1 <sup>st</sup> Argument	%rdi
2 <sup>nd</sup> Argument	%rsi
3 <sup>rd</sup> Argument	%rdx
4 <sup>th</sup> Argument	%rcx
5 <sup>th</sup> Argument	%r8
6 <sup>th</sup> Argument	%r9
Additional arguments	Pass on stack
Return value	%rax

# Passing Arguments and Return Values

```

int avg(int a, int b) {
    return (a+b)/2;
}

int main()
{
    int arg1 = 5;
    int arg2 = 3;
    int ans = avg(arg1, arg2);
    return ans;
}
    
```

**C Code**

```

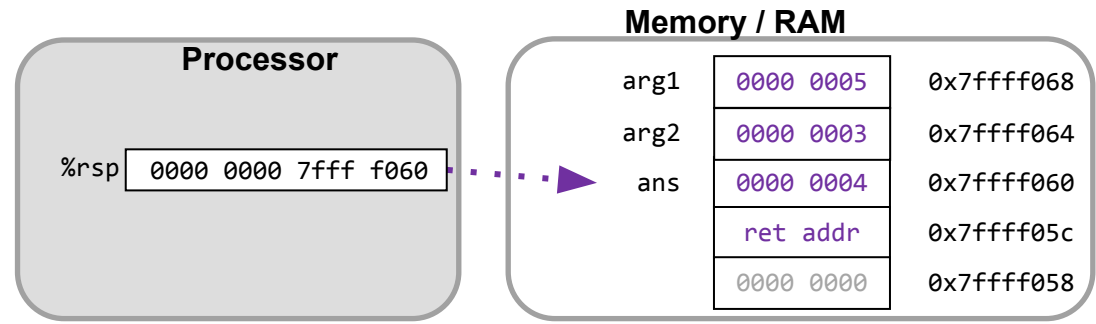
.text
.globl avg

avg:
    addl    %edi, %esi
    movl    %esi, %eax
    shrl    $31, %eax
    addl    %esi, %eax
    sarl    %eax
    ret

.globl main

main:
    movl    $3, %esi
    movl    $5, %edi
    call    avg
    ret
    
```

**Assembly**



# Compiler Handling of Procedures

- When coding in an high level language & using a compiler, certain conventions are followed that may lead to heavier usage of the stack
  - We have to be careful not to **overwrite** registers that have useful data
- High level languages (HLL) use the stack:
  - to **save register values** including the return address
  - for storage of **local variables** declared in the procedure
  - to pass **arguments** to a procedure
- Compilers usually put data on the stack in a certain order, which we call a **stack frame**

# Stack Frames

- Frame = **Def**: All data on stack belonging to a procedure / function
  - Space for saved registers
  - Space for local variables (those declared in a function)
  - Space for arguments

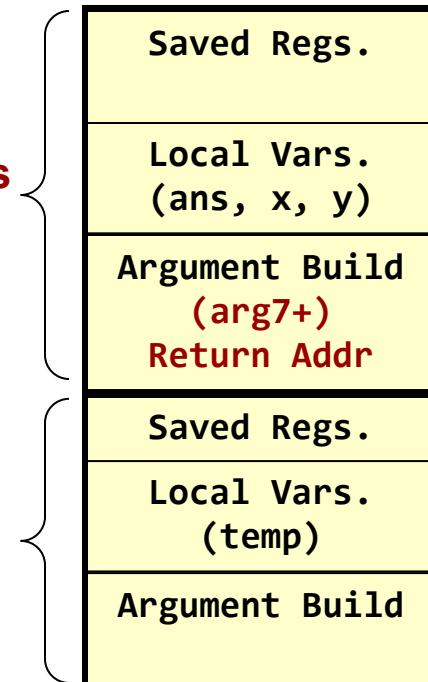
```

void main() {
    int ans, x, y;
    ans = avg(x, y);
    ...
}
int avg(int a, int b) {
    int temp=1; // local vars
    ...
}
    
```

Main Routine's Stack Frame

AVG's Stack Frame

Stack Growth

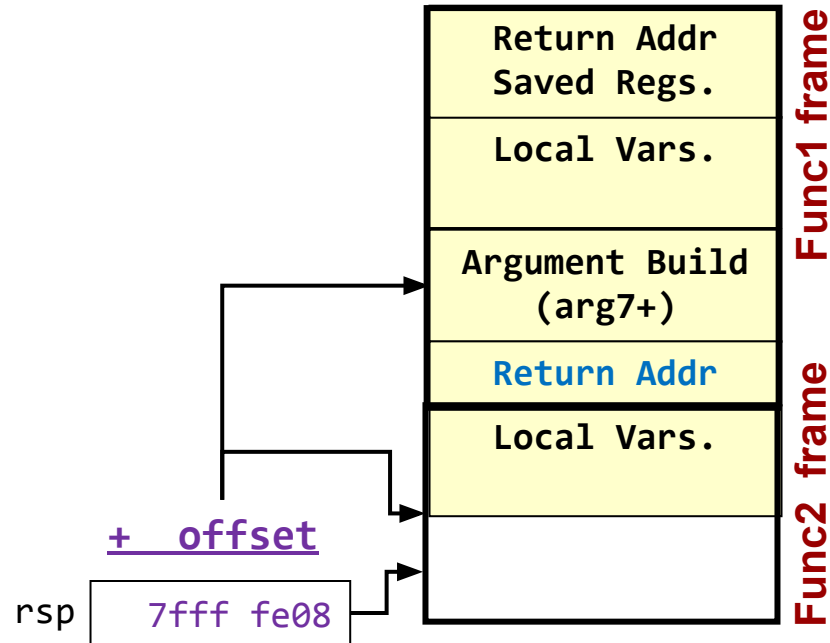


Stack Frame Organization



# Accessing Values on the Stack

- Stack pointer (%rsp) is usually used to access only the top value on the stack
- To access arguments and local variables, we need to access values buried in the stack
  - We can simply use an offset from %rsp [ e.g. 8(%rsp) ]



To access parameters we could try to use some displacement [i.e. `d(%rsp)` ]

# Many Arguments Examples

- Examine the following C code and corresponding assembly
- Assume initially %rsp = 0x7ffffff8
- Note how the 7<sup>th</sup> and 8<sup>th</sup> arguments are passed via the stack

```

caller:
    pushq    $8
    pushq    $7
    movl    $6, %r9d
    movl    $5, %r8d
    movl    $4, %ecx
    movl    $3, %edx
    movl    $2, %esi
    movl    $1, %edi
    call    f1
    addq    $16, %rsp
    ret
    
```

1

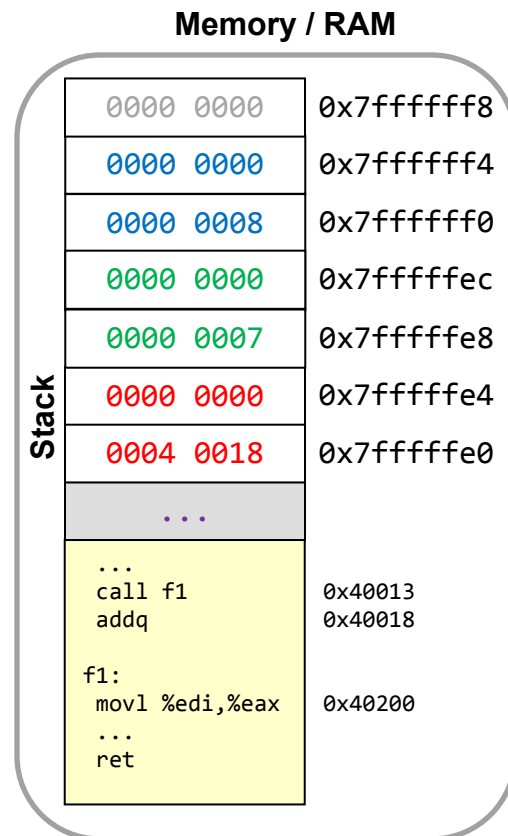
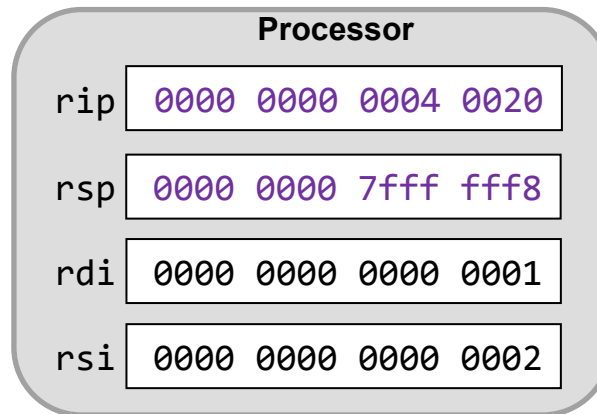
```

f1:    # 0x40200
    addl    %edi, %esi
    addl    %esi, %edx
    addl    %edx, %ecx
    addl    %ecx, %r8d
    addl    %r8d, %r9d
    movl    %r9d, %eax
    addl    8(%rsp), %eax
    addl    16(%rsp), %eax
    ret
    
```

```

int caller()
{
    int sum = f1(1, 2, 3, 4, 5, 6, 7, 8);
    return sum;
}

int f1(int a1, int a2, int a3, int a4,
        int a5, int a6, int a7, int a8)
{
    return a1+a2+a3+a4+a5+a6+a7+a8;
}
    
```



# Local Variables

CS:APP 3.7.4

- For simple integer/pointers the compiler can optimize code by **using a register** rather than allocating the variable on the stack
- Local variables need to be allocated on the **stack** if:
  - No free registers (too many locals)
  - The **&** operator is used and thus we need to be able to generate an address
  - Arrays or structs are used

# Local Variables Example

Memory / RAM

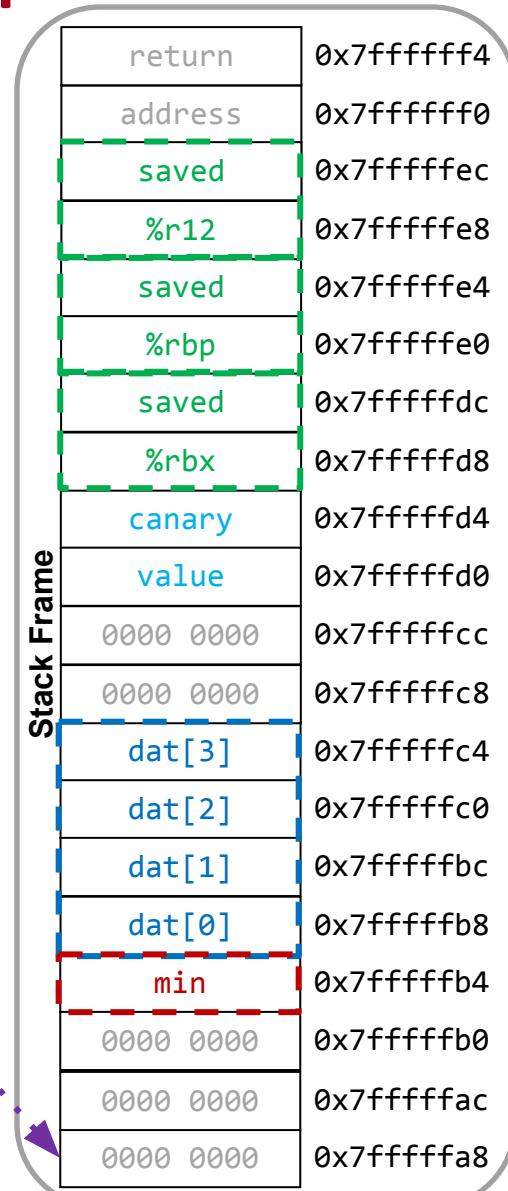
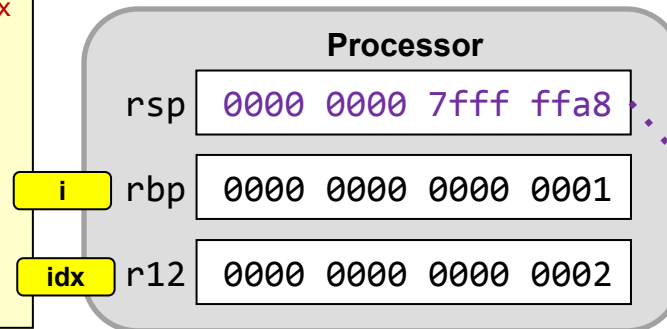
```

f2: ① pushq %r12
    pushq %rbp
    pushq %rbx
    ① subq $0x30, %rsp
    movl %edi, %r12d
    movq %fs:0x28, %rax
    movq %rax, 0x28(%rsp)
    xorl %eax, %eax
    ② leaq 0xc(%rsp), %rdi
    call getInt
    ③ movl $0, %ebx
    jmp .L4
.L6: movslq %ebx, %rbp
    ⑤ leaq 0x10(%rsp,%rbp,4), %rdi
    call getInt
    ⑥ movl 0x10(%rsp,%rbp,4), %eax
    cmpl 0xc(%rsp), %eax
    jge .L5
    ⑦ movl %eax, 0xc(%rsp)
.L5: addl $1, %ebx
.L4: ④ cmpl $3, %ebx
    jle .L6
    ⑧ movslq %r12d, %r12
    movl 0xc(%rsp), %eax
    addl 0x10(%rsp,%r12,4), %eax
    ⑨ movq 0x28(%rsp), %rdx
    xorq %fs:0x28, %rdx
    je .L7
    call __stack_chk_fail
.L7: addq $0x30, %rsp
    popq %rbx
    popq %rbp
    popq %r12
    ret
    
```

```

void getInt(int* ptr);
int f2(int idx)
① {
    ① int dat[4], min;
    ② getInt(&min);
    ③ for(int i=0; i < 4; i++){
        ⑤ getInt(&dat[i]);
        ⑥ if(dat[i] < min) min = dat[i];
        ⑦
    }
    ⑧ return dat[idx] + min;
    ⑨ }
    
```

- %rdi = %r12 = idx
- %rbp = %ebx = int i
- Notice %rdi must be reused from idx to the arguments for getInt(), thus the use of %r12 to hold idx



# Saved Register Problem

CS:APP 3.7.5

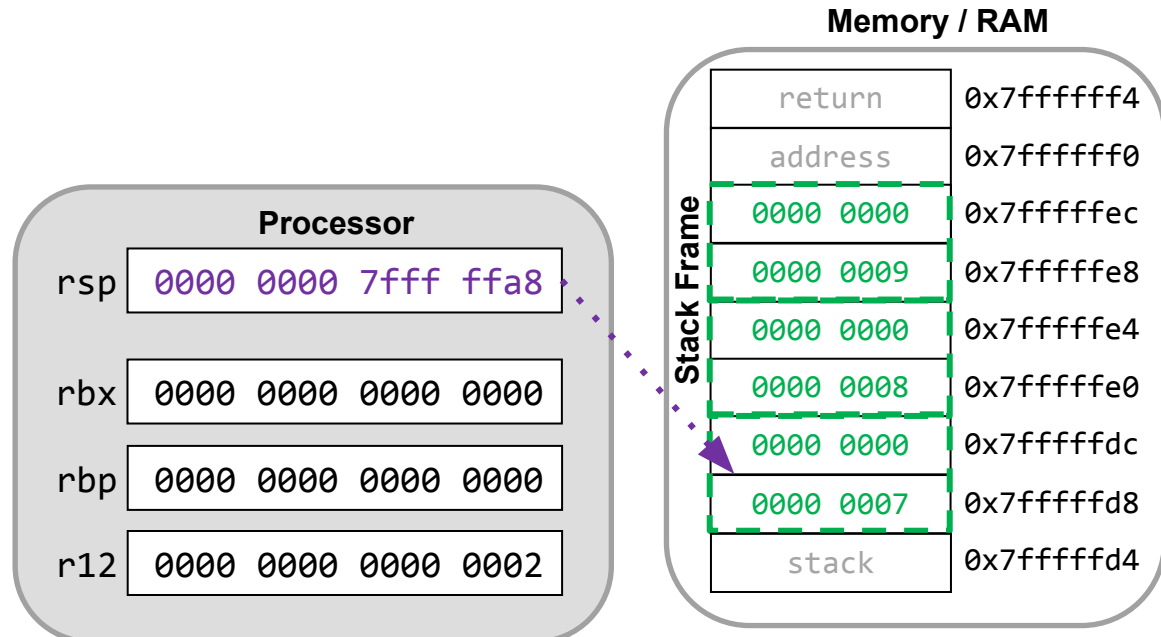
- Procedures are generally compiled separately
- The compiler will use registers for some temporaries and local variables
- What could go wrong?

```

f2:   pushq  %r12
      pushq  %rbp
      pushq  %rbx
      subq   $0x30, %rsp
      movl   %edi, %r12d
      ...
      movl   $0, %ebx
      ...
      movslq %ebx, %rbp
      leaq   0x10(%rsp,%rbp,4), %rdi
      ...
      popq   %rbx
      popq   %rbp
      popq   %r12
      ret

f1:   ...
      movl   $7, %ebx
      movl   $8, %ebp
      movq   $9, %r12
      movl   $2, %rdi
      call  f2
      ...
      add    %ebx, %ebp
      subq   $1, %r12
      ...
    
```

} Why are these needed?



# Saved Register Problem

- One procedure might overwrite a register value needed by the caller
- If f1() had values in %rbx, %rbp, and %r12 before calling f2() and then needed those values upon return, f2() may accidentally overwrite them

```

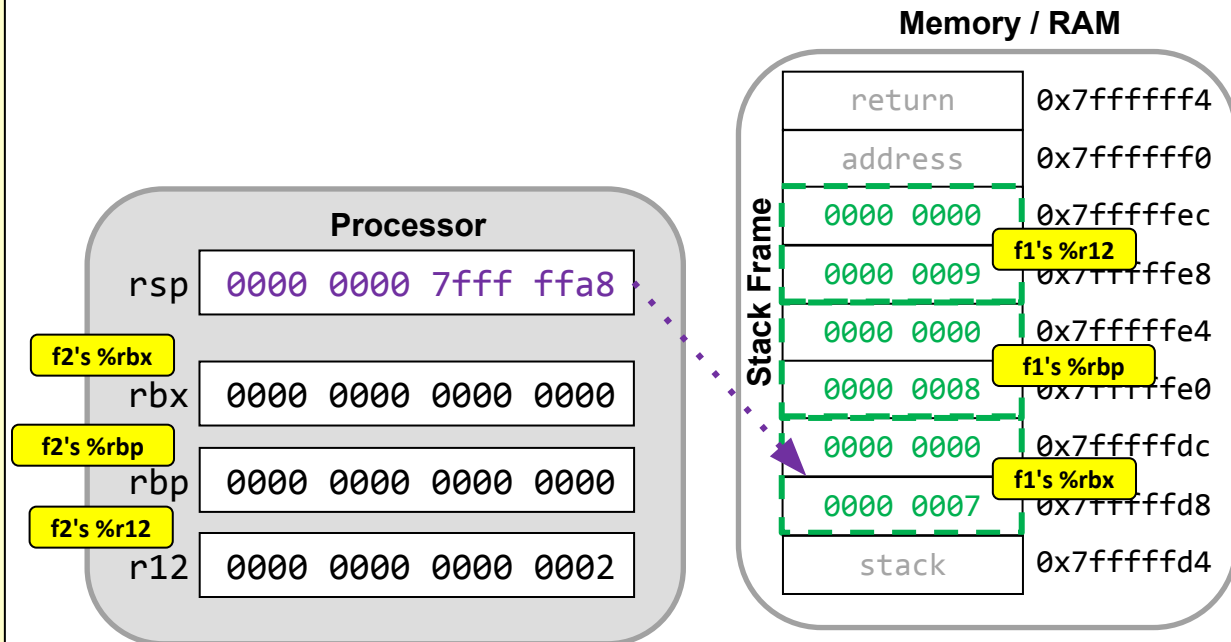
f2:  pushq  %r12
     pushq  %rbp
     pushq  %rbx
     subq   $0x30, %rsp
     movl   %edi, %r12d
     ...
     movl   $0, %ebx
     ...
     movslq %ebx, %rbp
     leaq   0x10(%rsp,%rbp,4), %rdi
     ...
     popq   %rbx
     popq   %rbp
     popq   %r12
     ret

f1:  ...
     movl   $7, %ebx
     movl   $8, %ebp
     movq   $9, %r12
     movl   $2, %rdi
     call  f2
     ...
     add    %ebx, %ebp
     subq   $1, %r12
     ...
    
```

**Why are these needed?**

Solution: Save/restore registers to/from the stack before overwriting it

- Which ones? Any register?



# Caller & Callee-Saved Convention

- Having to always play it safe and save a register to the stack before using it can decrease performance
- To increase performance, a standard is set to indicate which registers must be **preserved (callee-saved)** and which ones can be **overwritten freely (caller-saved)**
  - Callee Saved: Push values before overwriting them; restore before returning
  - Caller Saved: Push if the register value is needed after the function call; callee can freely overwrite; caller will restore upon return

<b>Callee-saved</b> (Callee must ensure the value is not modified)	%rbp, %rbx, %r12-%r15, %rsp*
<b>Caller-saved</b> (Caller must save the value if it wants to preserve it across a function call)	All other registers

\*%rsp need not be saved to the stack but should have the same value upon return as it did when the call was made

# Caller vs. Callee Saved

- One procedure might overwrite a register value needed by the caller
- If f1() had values in %rbx, %rbp, and %r12 before calling f2() and then needed those values upon return, f2() may accidentally overwrite them

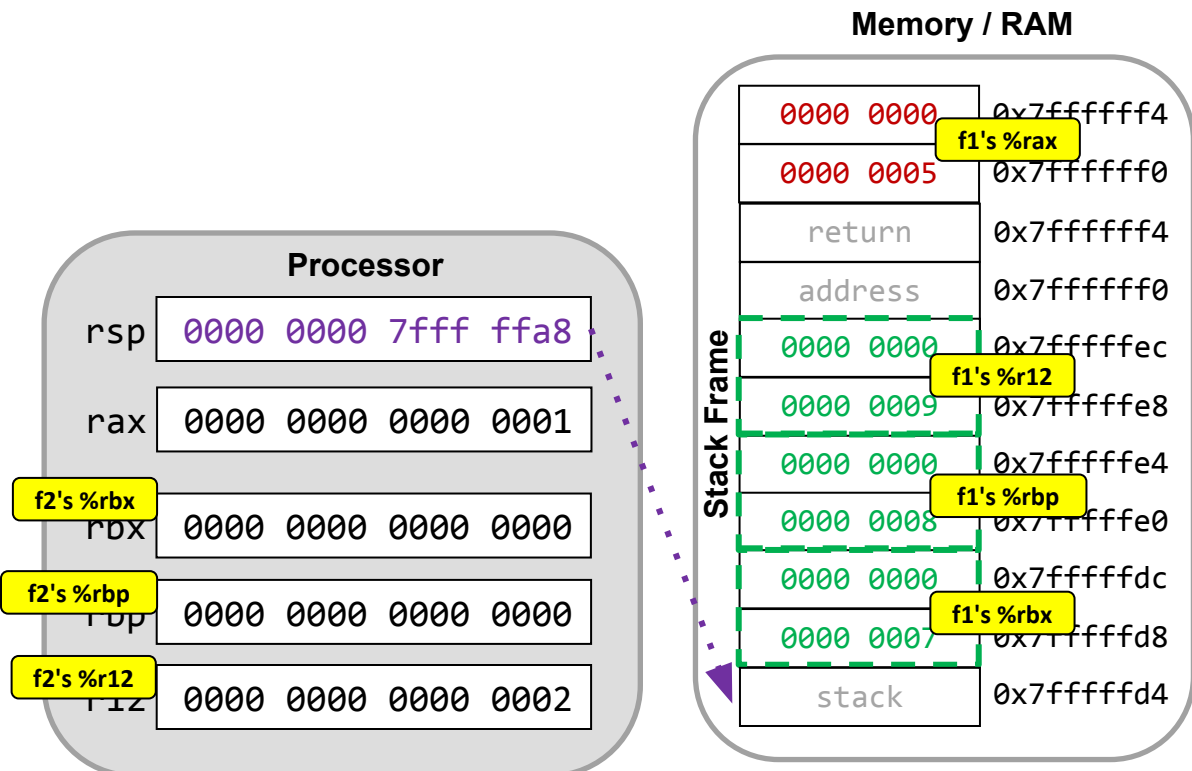
```

f2:  pushq  %r12
     pushq  %rbp
     pushq  %rbx
     subq   $0x30, %rsp
     movl   %edi, %r12d
     movl   $0, %ebx
     movl   $1, %eax
     movslq %ebx, %rbp
     leaq   0x10(%rsp,%rbp,4), %rdi
     popq   %rbx
     popq   %rbp
     popq   %r12
     ret

f1:  ...
     movl   $7, %ebx
     movl   $8, %ebp
     movq   $9, %r12
     movq   $5, %rax
     pushq  %rax
     movl   $2, %rdi
     call   f2
     popq   %rax
     addl   %ebx, %ebp
     subq   $1, %r12
     ...
    
```

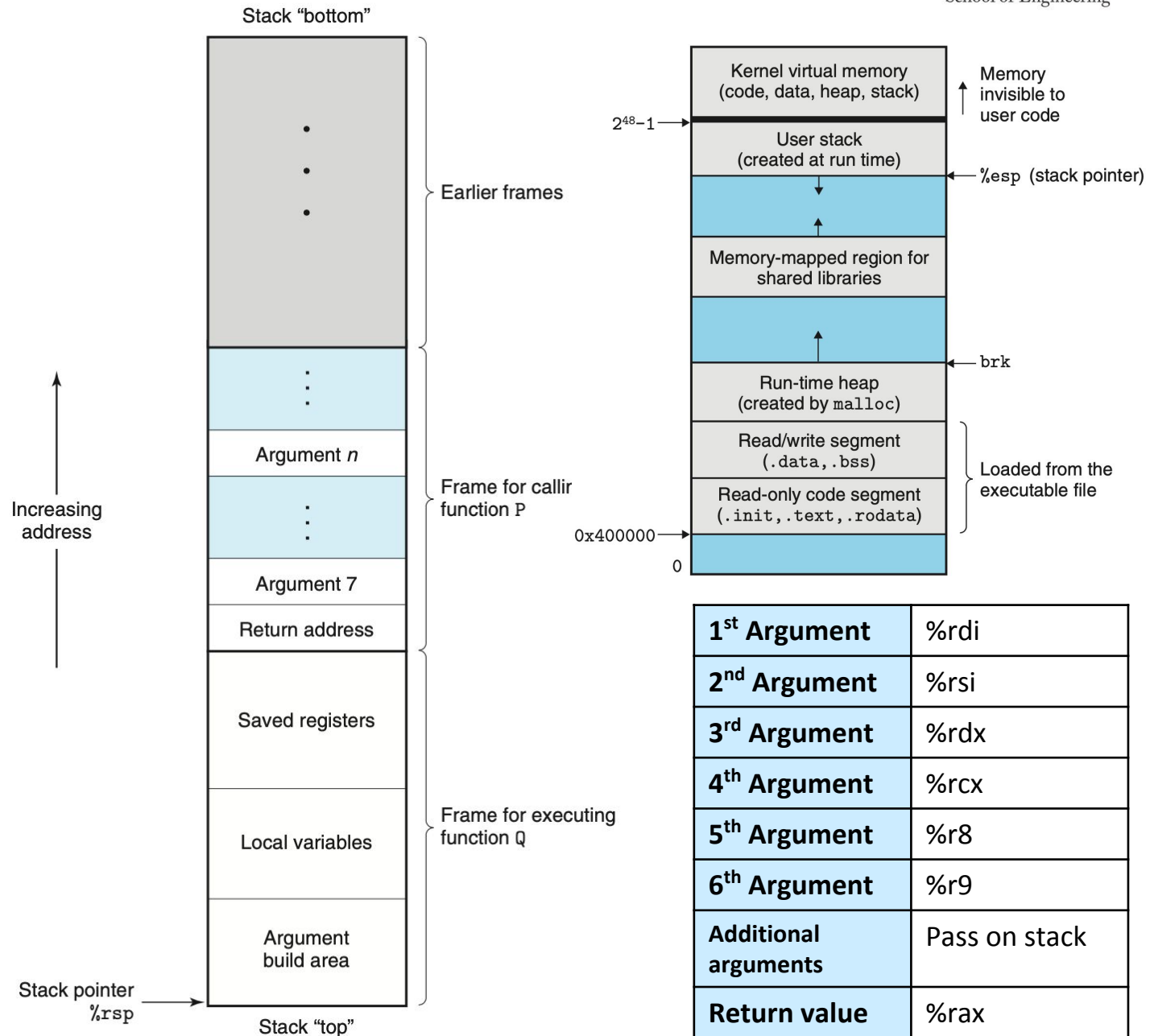
**Callee Saved**

**Caller Saved**





**Figure 3.25**  
**General stack frame structure.** The stack can be used for passing arguments, for storing return information, for saving registers, and for local storage. Portions may be omitted when not needed.



1 <sup>st</sup> Argument	<code>%rdi</code>
2 <sup>nd</sup> Argument	<code>%rsi</code>
3 <sup>rd</sup> Argument	<code>%rdx</code>
4 <sup>th</sup> Argument	<code>%rcx</code>
5 <sup>th</sup> Argument	<code>%r8</code>
6 <sup>th</sup> Argument	<code>%r9</code>
Additional arguments	Pass on stack
Return value	<code>%rax</code>

Recap

# Summary

- To support subroutines we need to save the return address on the stack
  - call and ret perform this implicitly
- There must be agreed upon locations where arguments and return values can be communicated
- The stack is a common memory location to allocate space for saved values and local variables