# Machine-Level Programming III: Procedures

Introduction to Computer Systems 6<sup>th</sup> Lecture, Sep. 26, 2022

#### **Instructors:**

Class 1: Chen Xiangqun, Liu Xianhua

Class 2: Guan Xuetao

Class 3: Lu Junlin

# **Objectives**

- Basic functionality of the pairs: push / pop and call / ret
- Students should be able to identify the different components of a stack (return address, arguments, saved registers, local variables)
- Explain the difference between callee and caller save registers
- Explain how a stack permits functions to be called recursively / re-entrant

# **Today**

- Procedures
  - Mechanisms
  - Stack Structure
  - Calling Conventions
    - Passing control
    - Passing data
    - Managing local data
  - Illustration of Recursion

#### Passing control

- To beginning of procedure code
- Back to return point

### Passing data

- Procedure arguments
- Return value

#### Memory management

- Allocate during procedure execution
- Deallocate upon return
- Mechanisms all implemented with machine instructions
- x86-64 implementation of a procedure uses only those mechanisms required

```
int Q(int i)
{
  int t = 3*i;
  int v[10];
  .
  return v[t];
}
```

- Passing control
  - To beginning of procedure code
  - Back to return point
- Passing data
  - Procedure arguments
  - Return value
- Memory management
  - Allocate during procedure execution
  - Deallocate upon return
- Mechanisms all implemented with machine instructions
- x86-64 implementation of a procedure uses only those mechanisms required

```
P(...) {
      Q(x);
  print(y)
    Q(int i)
  int t = 3*i;
  int v[10];
  return v[t];
```

- Passing control
  - To beginning of procedure code
  - Back to return point
- Passing data
  - Procedure arguments
  - Return value
- Memory management
  - Allocate during procedure execution
  - Deallocate upon return
- Mechanisms all implemented with machine instructions
- x86-64 implementation of a procedure uses only those mechanisms required

```
P(...) {
    = Q(x);
  print(y)
int Q(\int i)
  int t
         = 3*i;
  int v[10];
  return v[t];
```

#### Passing control

- To beginning of procedure code
- Back to return point

#### Passing data

- Procedure arguments
- Return value

#### Memory management

- Allocate during procedure execution
- Deallocate upon return
- Mechanisms all implemented with machine instructions
- x86-64 implementation of a procedure uses only those mechanisms required

```
int Q(int i)
{
   int t = 3*i;
   int v[10];
   return v[t];
}
```

P(...) {

Machine instructions implement the mechanisms, but the choices are determined by designers. These choices make up the **Application Binary Interface** (ABI).

- Deallocate upon return
- Mechanisms all implemented with machine instructions
- x86-64 implementation of a procedure uses only those mechanisms required

```
int v[10];
.
.
return v[t];
}
```

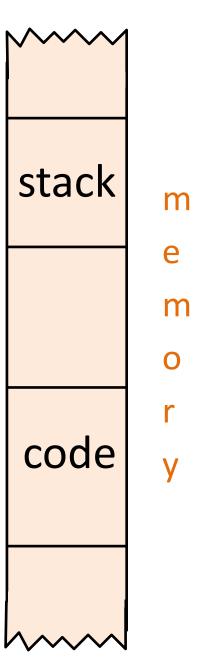
# **Today**

#### Procedures

- Mechanisms
- Stack Structure
- Calling Conventions
  - Passing control
  - Passing data
  - Managing local data
- Illustration of Recursion

### x86-64 Stack

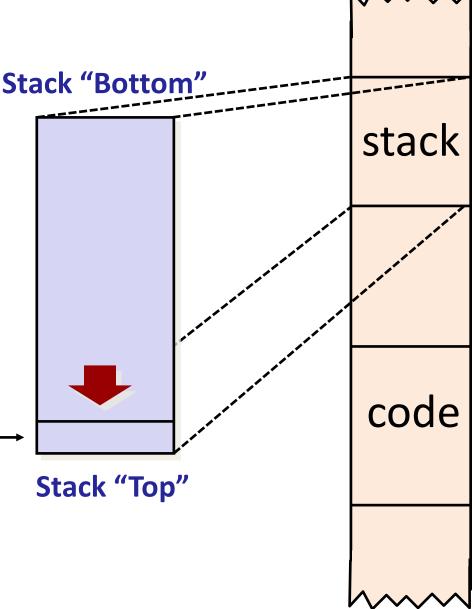
- Region of memory managed with stack discipline
  - Memory viewed as array of bytes.
  - Different regions have different purposes.
  - (Like ABI, a policy decision)



### **x86-64 Stack**

Region of memory managed with stack discipline

Stack Pointer: %rsp



11

### x86-64 Stack

- Region of memory managed with stack discipline
- Grows toward lower addresses
- Register %rsp contains lowest stack address
  - address of "top" element

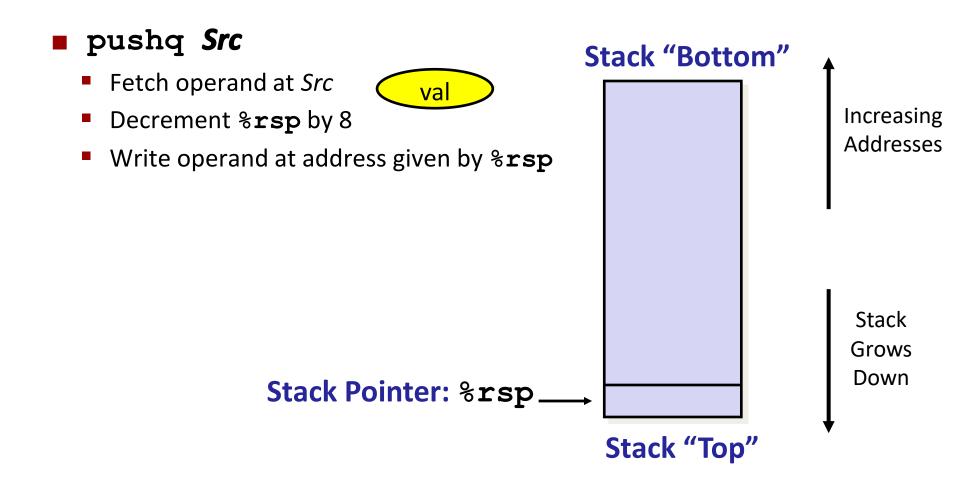
sp contains
k address
top" element

Stack
Grows
Down

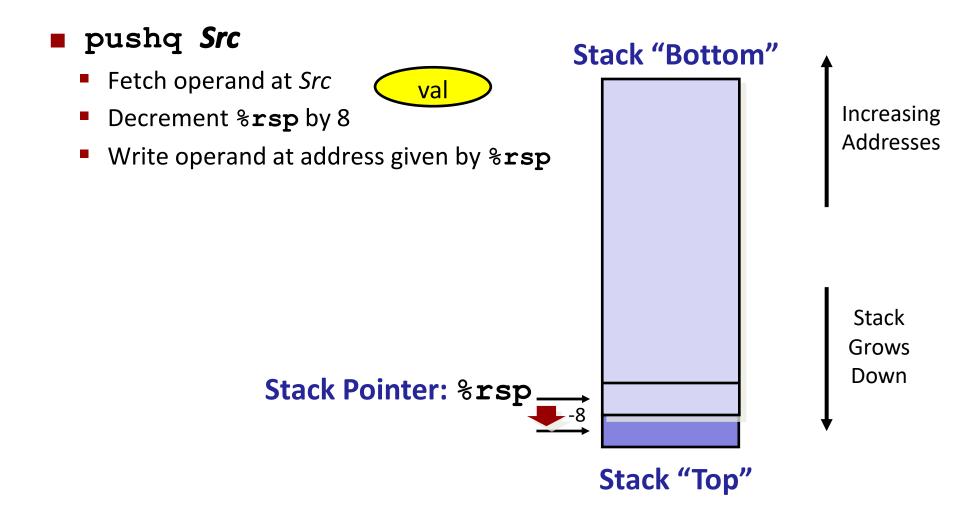
Stack "Top"

Stack "Bottom"

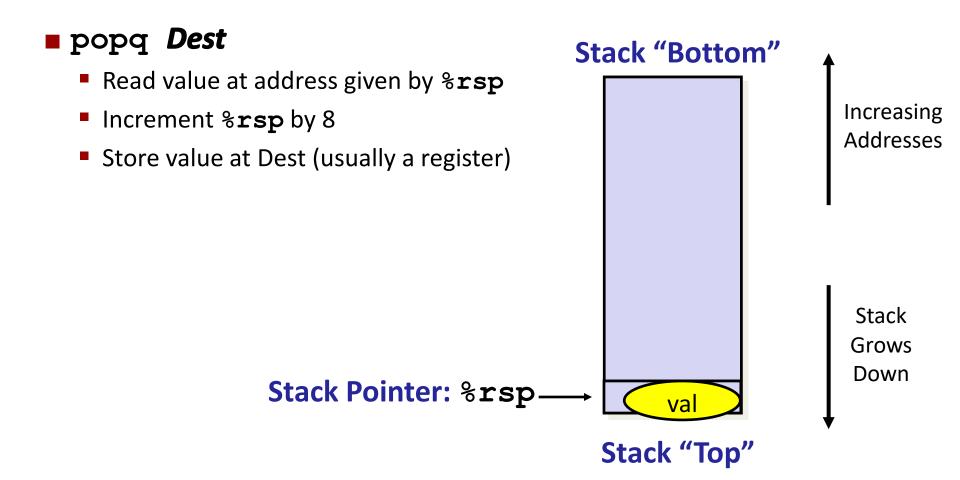
### x86-64 Stack: Push



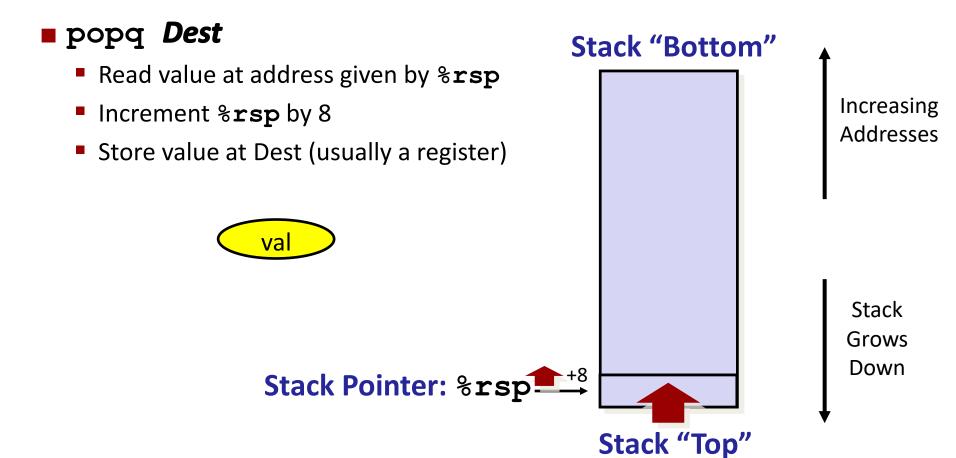
### x86-64 Stack: Push



# x86-64 Stack: Pop



# x86-64 Stack: Pop



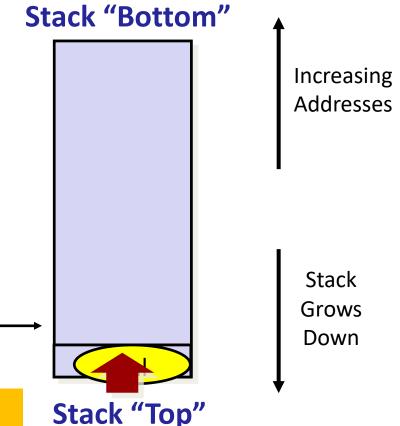
# x86-64 Stack: Pop

#### ■ popq *Dest*

- Read value at address given by %rsp
- Increment %rsp by 8
- Store value at Dest (usually a register)

Stack Pointer: %rsp----

(The memory doesn't change, only the value of %rsp)



# **Today**

#### Procedures

- Mechanisms
- Stack Structure
- Calling Conventions
  - Passing control
  - Passing data
  - Managing local data
- Illustration of Recursion

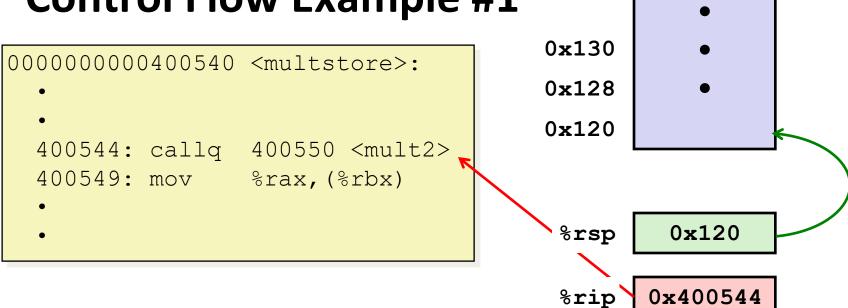
```
long mult2 (long a, long b)
{
  long s = a * b;
  return s;
}

400550: mov %rdi,%rax # a
  400553: imul %rsi,%rax # a * b
  400557: retq # Return
```

### **Procedure Control Flow**

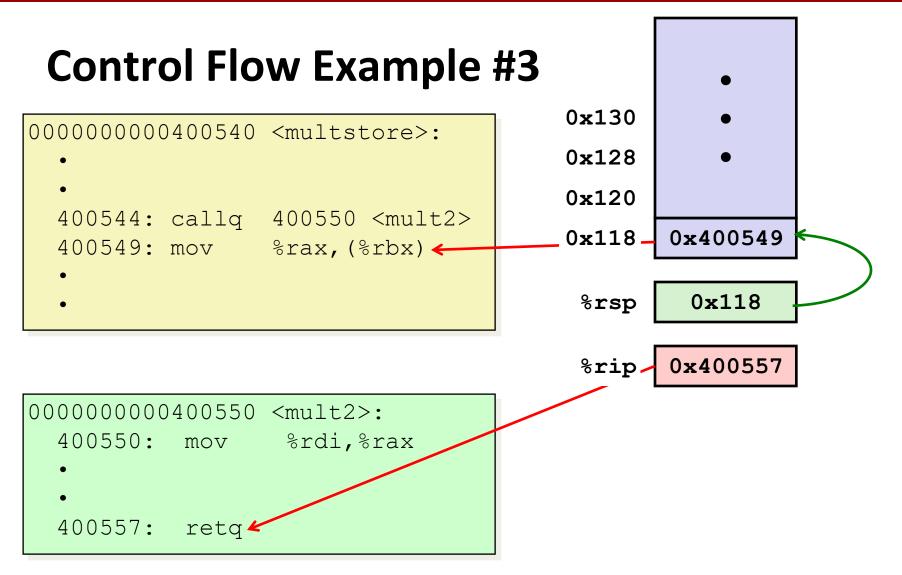
- Use stack to support procedure call and return
- Procedure call: call label
  - Push return address on stack
  - Jump to label
- Return address:
  - Address of the next instruction right after call
  - Example from disassembly
- Procedure return: ret
  - Pop address from stack
  - Jump to address

# **Control Flow Example #1**

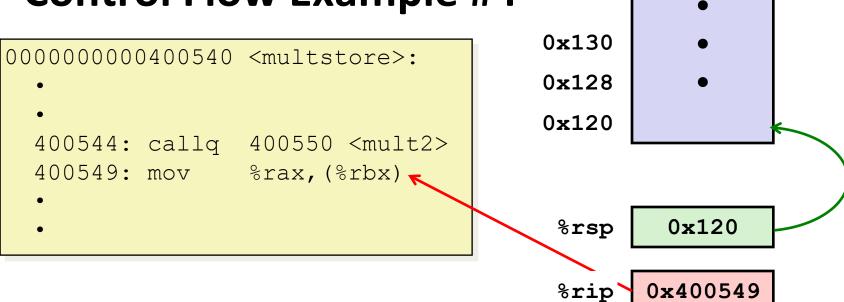


```
0000000000400550 <mult2>:
   400550: mov %rdi,%rax
   •
   400557: retq
```

### **Control Flow Example #2** 0x1300000000000400540 <multstore>: 0x1280x120400544: callq 400550 <mult2> 0x118. 0x400549400549: mov %rax, (%rbx) ← 0x118 %rsp $0 \times 400550$ %rip. 0000000000400550 <mult2>: 400550: mov %rdi,%rax < 400557: retq



# **Control Flow Example #4**



```
0000000000400550 <mult2>:
   400550: mov %rdi,%rax
   •
   400557: retq
```

# **Today**

#### Procedures

- Mechanisms
- Stack Structure
- Calling Conventions
  - Passing control
  - Passing data
  - Managing local data
- Illustrations of Recursion & Pointers

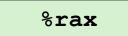
### **Procedure Data Flow**

#### Registers

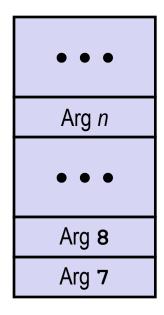
First 6 arguments



Return value



#### Stack



Only allocate stack space when needed

# Data Flow Examples

```
void multstore
  (long x, long y, long *dest)
{
    long t = mult2(x, y);
    *dest = t;
}
```

```
long mult2
  (long a, long b)
{
  long s = a * b;
  return s;
}
```

```
000000000000400550 <mult2>:
    # a in %rdi, b in %rsi
400550: mov %rdi,%rax # a
400553: imul %rsi,%rax # a * b
# s in %rax
400557: retq # Return
```

# **Today**

#### Procedures

- Mechanisms
- Stack Structure
- Calling Conventions
  - Passing control
  - Passing data
  - Managing local data
- Illustration of Recursion

# **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

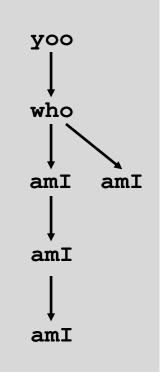
#### Stack allocated in Frames

state for single procedure instantiation

# **Call Chain Example**

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

Example
Call Chain



Procedure amI() is recursive

## **Stack Frames**

#### **Contents**

- Return information
- Local storage (if needed)
- Temporary space (if needed)

**Previous** Frame

Frame Pointer: %rbp

(Optional)

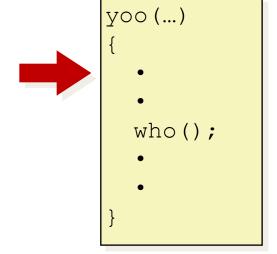
Stack Pointer: %rsp

Frame for proc

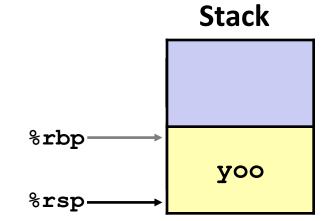
# Stack "Top"

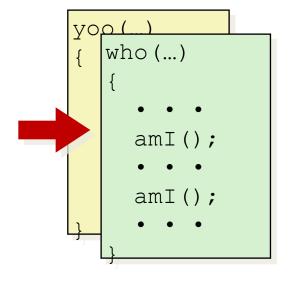
#### Management

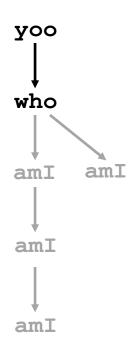
- Space allocated when enter procedure
  - "Set-up" code
  - Includes push by call instruction
- Deallocated when return
  - "Finish" code
  - Includes pop by ret instruction

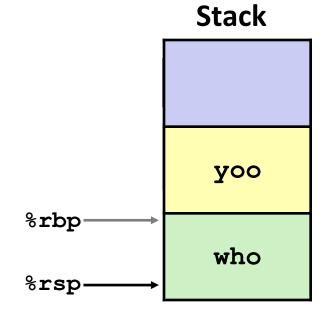


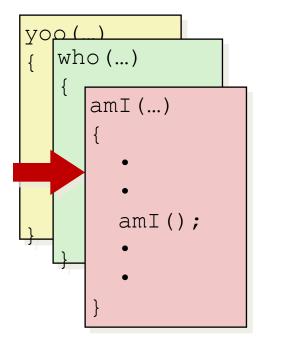


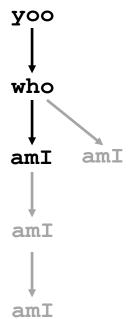


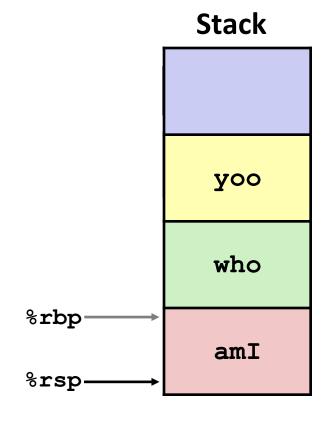


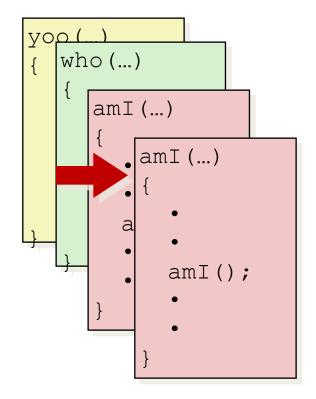


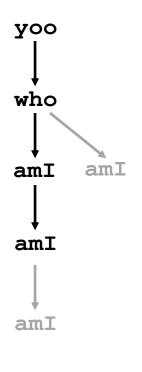


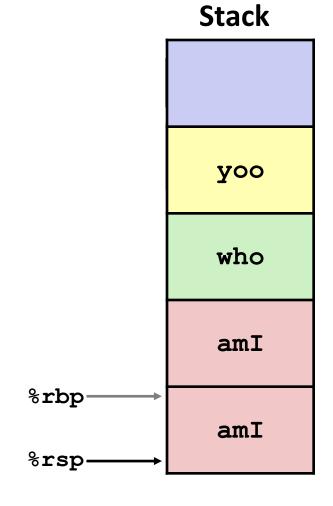


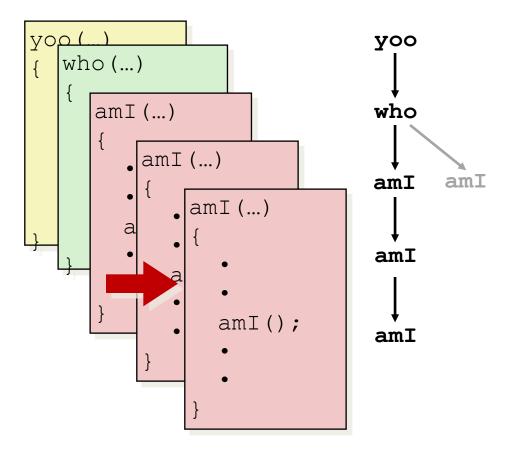


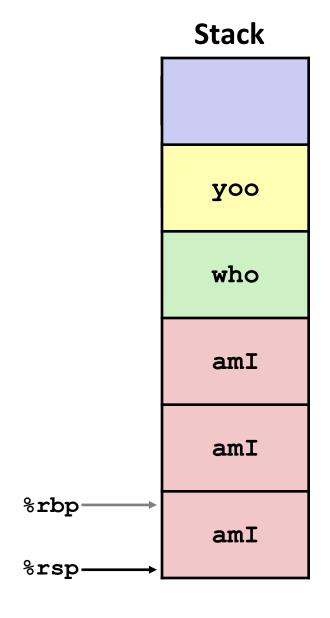


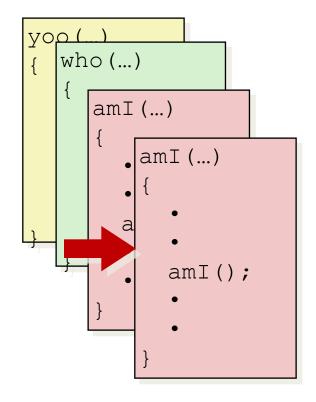


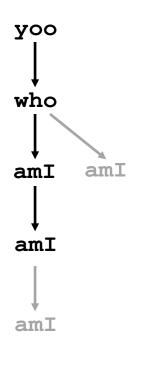


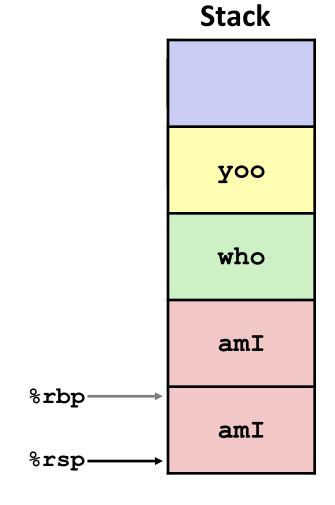


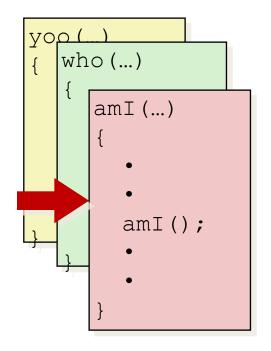


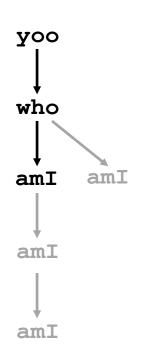


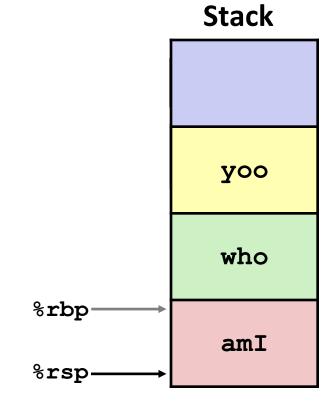


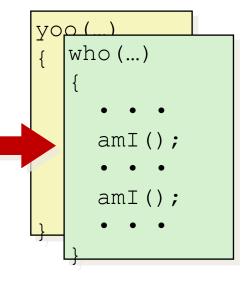




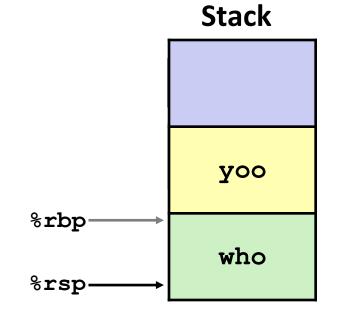


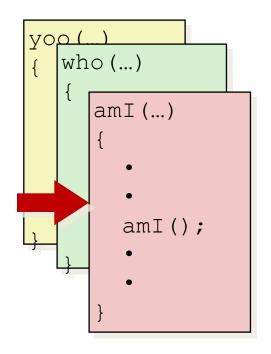


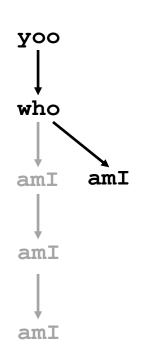


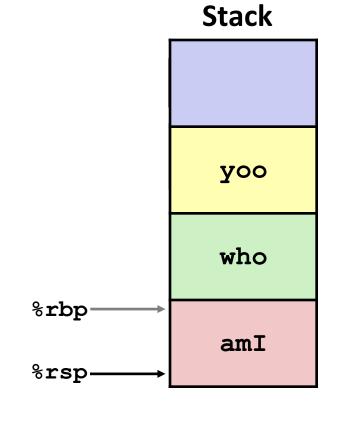


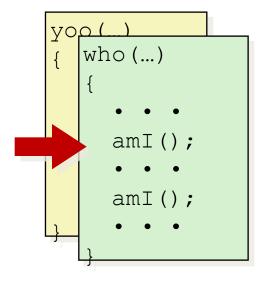


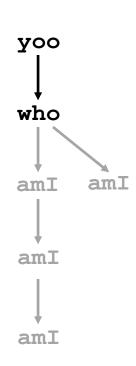


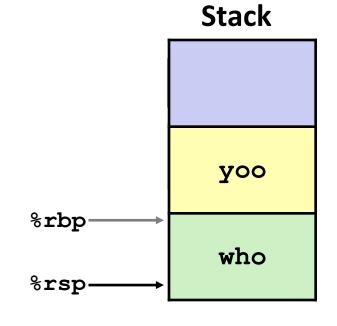


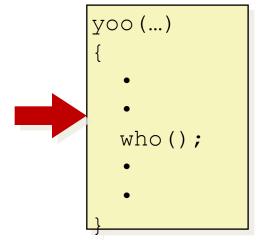




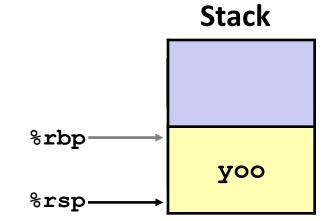












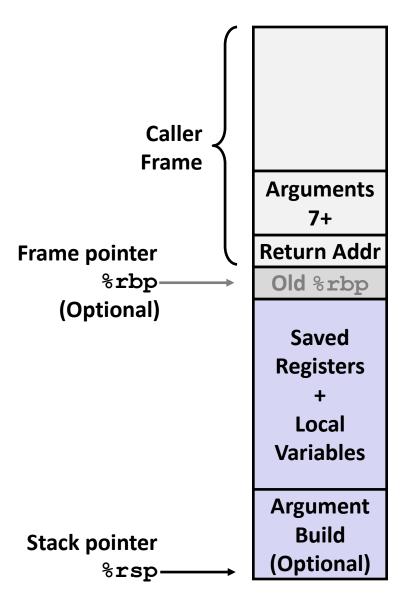
# x86-64/Linux Stack Frame

### Current Stack Frame ("Top" to Bottom)

- "Argument build:"
   Parameters for function about to call
- Local variablesIf can't keep in registers
- Saved register context
- Old frame pointer (optional)

#### Caller Stack Frame

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



## Example: incr

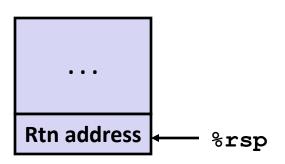
```
long incr(long *p, long val) {
   long x = *p;
   long y = x + val;
   *p = y;
   return x;
}
```

```
incr:
  movq (%rdi), %rax
  addq %rax, %rsi
  movq %rsi, (%rdi)
  ret
```

Register	Use(s)
%rdi	Argument <b>p</b>
%rsi	Argument val, y
%rax	<b>x</b> , Return value

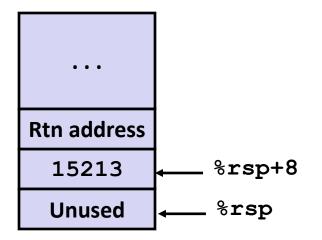
# long call\_incr() { long v1 = 15213; long v2 = incr(&v1, 3000); return v1+v2; }

#### **Initial Stack Structure**



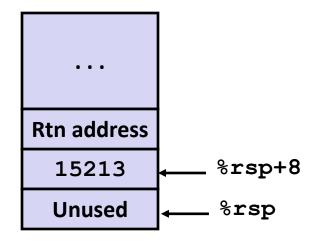
```
call_incr:
    subq    $16, %rsp
    movq    $15213, 8(%rsp)
    movl    $3000, %esi
    leaq    8(%rsp), %rdi
    call    incr
    addq    8(%rsp), %rax
    addq    $16, %rsp
    ret
```

#### **Resulting Stack Structure**



```
long call_incr() {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return v1+v2;
}
```

```
call_incr:
   subq $16, %rsp
   movq $15213, 8(%rsp)
   movl $3000, %esi
   leaq 8(%rsp), %rdi
   call incr
   addq 8(%rsp), %rax
   addq $16, %rsp
   ret
```



Register	Use(s)
%rdi	&v1
%rsi	3000

```
long call_incr() {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return v1+v2;
}
```

#### **Stack Structure**

```
Rtn address
```

```
Aside 1: movl $3000, %esi
```

• Note: movl -> %exx zeros out high order 32 bits.

Why use movl instead of movq? 1 byte shorter.

```
movl $3000, %esi
leaq 8(%rsp), %rdi
call incr
addq 8(%rsp), %rax
addq $16, %rsp
ret
```

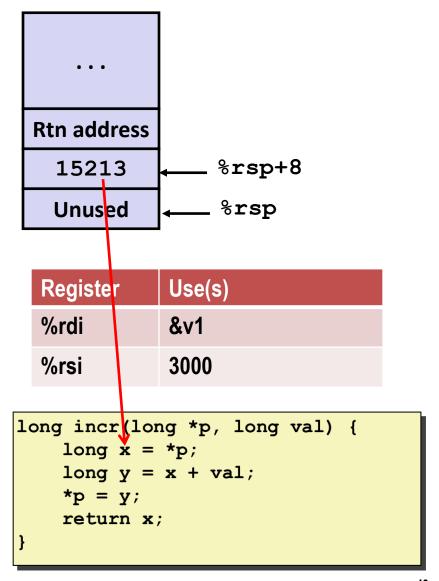
mov

%rdi	&v1
%rsi	3000

```
Stack Structure
long call incr() {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
   return v1+v2;
                                    Rtn address
                                                 %rsp+8
                                     15213
                                                 %rsp
       Aside 2: leaq 8(%rsp), %rdi
ca:
  Computes %rsp+8
                                               se(s)
  Actually, used for what it is meant!
 leaq 8(%rsp), %rdi
                                    %rsi
                                             3000
 call incr
 addq 8(%rsp), %rax
 addq $16, %rsp
 ret
```

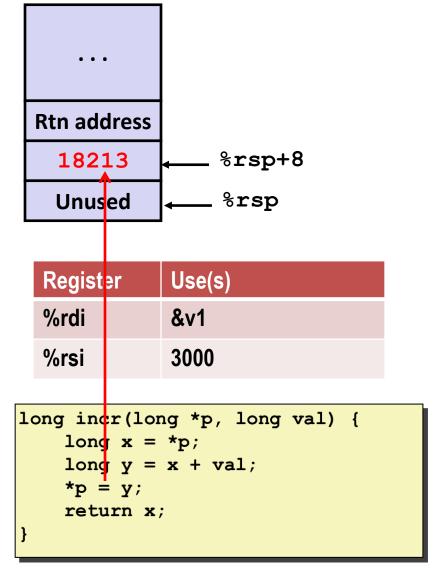
```
long call_incr() {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return v1+v2;
}
```

```
call_incr:
    subq $16, %rsp
    movq $15213, 8(%rsp)
    movl $3000, %esi
    leaq 8(%rsp), %rdi
    call incr
    addq 8(%rsp), %rax
    addq $16, %rsp
    ret
```



```
long call_incr() {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return v1+v2;
}
```

```
call_incr:
    subq    $16, %rsp
    movq    $15213, 8(%rsp)
    movl    $3000, %esi
    leaq    8(%rsp), %rdi
    call    incr
    addq    8(%rsp), %rax
    addq    $16, %rsp
    ret
```



```
long call_incr() {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return v1+v2;
}
```

```
...

Rtn address

18213 ← %rsp+8

Unused %rsp
```

```
call_incr:
    subq $16, %rsp
    movq $15213, 8(%rsp)
    movl $3000, %esi
    leaq 8(%rsp), %rdi
    call incr
    addq 8(%rsp), %rax
    addq $16, %rsp
    ret
```

	Register	Use(s)	
	%rax	Return value, 15213	
10	long incr(long *p, long val) {		
	$\begin{array}{ccc} long & x & = \\ long & y & = \end{array}$	- · /	
	*p = y; return x;		
}			

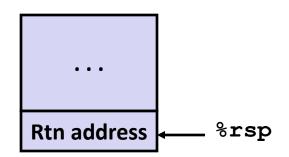
#### **Stack Structure**

```
long call_incr() {
   long v1 = 15213;
   long v2 = incr(&v1, 3000);
   return v1+v2;
}
```

call_incr	:
subq	\$16, %rsp
movq	\$15213, 8(%rsp)
movl	\$3000, %esi
leaq	8(%rsp), %rdi
call	incr
addq	8(%rsp), %rax
addq	\$16, %rsp
ret	

Register	Use(s)
%rax	Return value

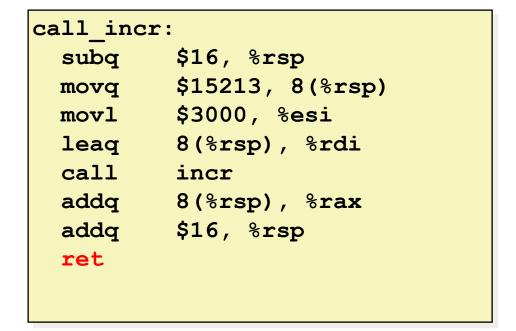
#### **Updated Stack Structure**



```
long call_incr() {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return v1+v2;
}
```

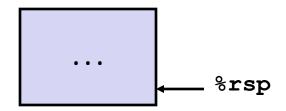
#### **Updated Stack Structure**

```
Rtn address & %rsp
```



Register	Use(s)
%rax	Return value

#### **Final Stack Structure**



## **Register Saving Conventions**

- When procedure yoo calls who:
  - yoo is the caller
  - who is the callee
- Can register be used for temporary storage?

```
yoo:

movq $15213, %rdx
call who
addq %rdx, %rax

ret
```

```
who:

• • •

subq $18213, %rdx

• • •

ret
```

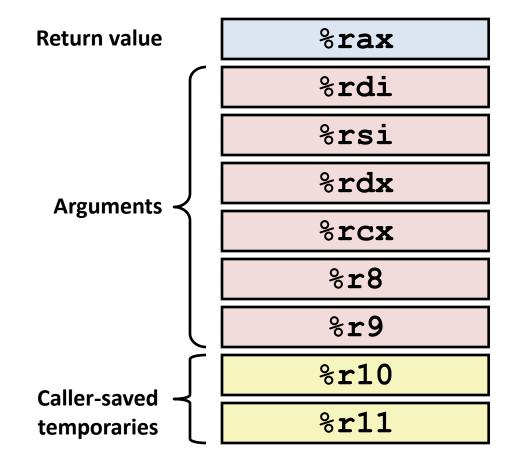
- Contents of register %rdx overwritten by who
- This could be trouble → something should be done!
  - Need some coordination

## **Register Saving Conventions**

- When procedure yoo calls who:
  - yoo is the caller
  - who is the callee
- Can register be used for temporary storage?
- Conventions
  - "Caller Saved"
    - Caller saves temporary values in its frame before the call
  - "Callee Saved"
    - Callee saves temporary values in its frame before using
    - Callee restores them before returning to caller

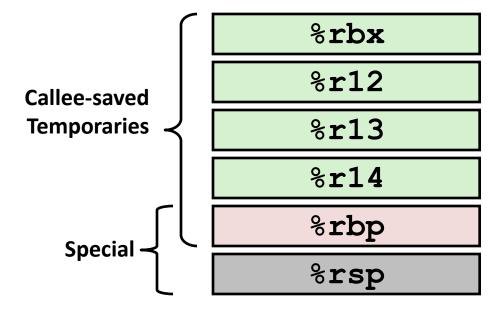
## x86-64 Linux Register Usage #1

- %rax
  - Return value
  - Also caller-saved
  - Can be modified by procedure
- %rdi, ..., %r9
  - Arguments
  - Also caller-saved
  - Can be modified by procedure
- %r10, %r11
  - Caller-saved
  - Can be modified by procedure



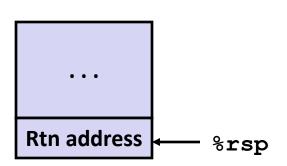
## x86-64 Linux Register Usage #2

- %rbx, %r12, %r13, %r14
  - Callee-saved
  - Callee must save & restore
- %rbp
  - Callee-saved
  - Callee must save & restore
  - May be used as frame pointer
  - Can mix & match
- %rsp
  - Special form of callee save
  - Restored to original value upon exit from procedure



# long call\_incr2(long x) { long v1 = 15213; long v2 = incr(&v1, 3000); return x+v2; }

#### **Initial Stack Structure**



- x comes in register %rdi.
- We need %rdi for the call to incr.
- Where should be put x, so we can use it after the call to incr?

```
long call_incr2(long x) {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return x+v2;
}
```

```
call_incr2:

pushq %rbx

subq $16, %rsp

movq %rdi, %rbx

movq $15213, 8(%rsp)

movl $3000, %esi

leaq 8(%rsp), %rdi

call incr
```

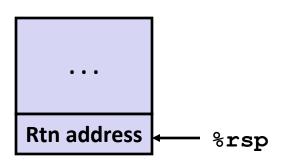
addq %rbx, %rax

addq \$16, %rsp

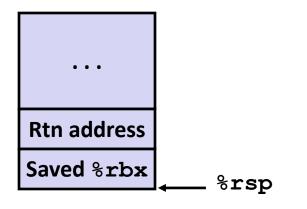
popq %rbx

ret

#### **Initial Stack Structure**



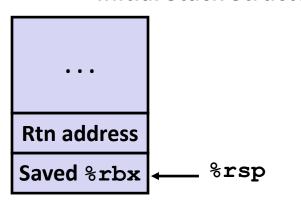
#### **Resulting Stack Structure**



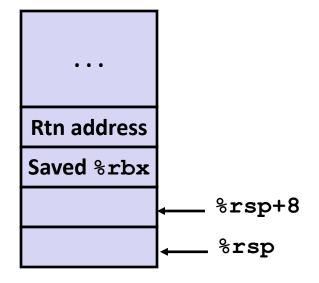
```
long call_incr2(long x) {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return x+v2;
}
```

```
call incr2:
 pushq %rbx
 subq $16, %rsp
 movq %rdi, %rbx
 movq $15213, 8(%rsp)
 movl $3000, %esi
 leaq 8(%rsp), %rdi
 call incr
 addq %rbx, %rax
 addq $16, %rsp
 popq %rbx
 ret
```

#### **Initial Stack Structure**



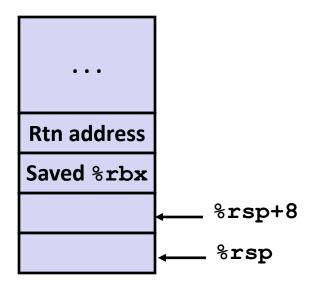
#### **Resulting Stack Structure**



```
long call_incr2(long x) {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return x+v2;
}
```

```
call incr2:
 pushq %rbx
 subq $16, %rsp
 movq %rdi, %rbx
 movq $15213, 8(%rsp)
 movl $3000, %esi
 leaq 8(%rsp), %rdi
 call
        incr
 addq %rbx, %rax
 addq $16, %rsp
 popq %rbx
 ret
```

#### **Stack Structure**

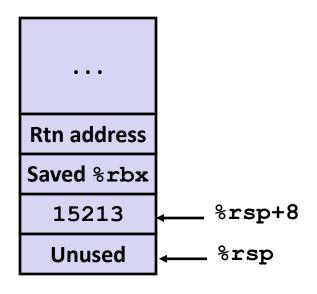


x is saved in %rbx,
 a callee saved register

```
long call_incr2(long x) {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return x+v2;
}
```

```
call incr2:
 pushq %rbx
 subq $16, %rsp
 movq %rdi, %rbx
 movq $15213, 8(%rsp)
 movl $3000, %esi
 leaq 8(%rsp), %rdi
 call
        incr
 addq %rbx, %rax
 addq $16, %rsp
 popq %rbx
 ret
```

#### **Stack Structure**

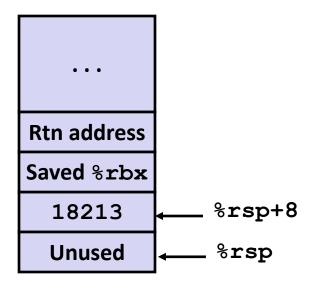


x is saved in %rbx,
 a callee saved register

# long call\_incr2(long x) { long v1 = 15213; long v2 = incr(&v1, 3000); return x+v2; }

```
call_incr2:
  pushq %rbx
  subq $16, %rsp
  movq %rdi, %rbx
  movq $15213, 8(%rsp)
  movl $3000, %esi
  leaq 8(%rsp), %rdi
  call incr
  addq %rbx, %rax
  addq $16, %rsp
  popq %rbx
  ret
```

#### **Stack Structure**



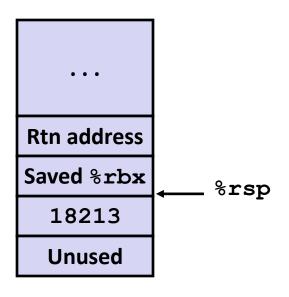
### Upon return from incr:

- x safe in %rbx
- Return val v2 in %rax
- Compute x+v2:addq %rbx, %rax

# long call\_incr2(long x) { long v1 = 15213; long v2 = incr(&v1, 3000); return x+v2; }

```
call incr2:
 pushq %rbx
 subq $16, %rsp
 movq %rdi, %rbx
 movq $15213, 8(%rsp)
 movl $3000, %esi
 leaq 8(%rsp), %rdi
 call incr
 addq %rbx, %rax
 addq $16, %rsp
 popq %rbx
 ret
```

#### **Stack Structure**

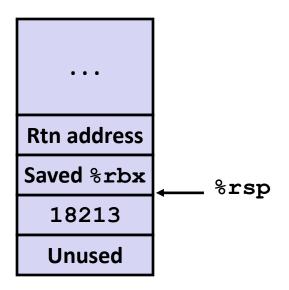


• Return result in %rax

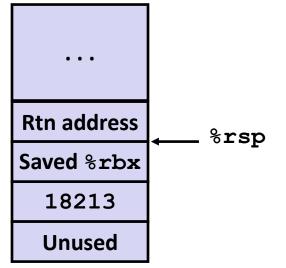
#### **Initial Stack Structure**

```
long call_incr2(long x) {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return x+v2;
}
```

```
call_incr2:
  pushq %rbx
  subq $16, %rsp
  movq %rdi, %rbx
  movq $15213, 8(%rsp)
  movl $3000, %esi
  leaq 8(%rsp), %rdi
  call incr
  addq %rbx, %rax
  addq $16, %rsp
  popq %rbx
  ret
```



#### final Stack Structure



## **Today**

#### Procedures

- Mechanisms
- Stack Structure
- Calling Conventions
  - Passing control
  - Passing data
  - Managing local data
- Illustration of Recursion

## **Recursive Function**

```
pcount r:
 movl $0, %eax
 testq %rdi, %rdi
        .L6
 jе
 pushq %rbx
 movq %rdi, %rbx
 andl $1, %ebx
 shrq %rdi # (by 1)
 call
        pcount r
 addq
        %rbx, %rax
        %rbx
 popq
.L6:
 rep; ret
```

## **Recursive Function Terminal Case**

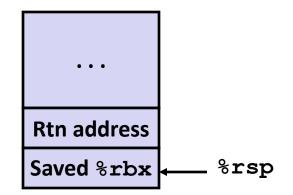
Register	Use(s)	Туре
%rdi	x	Argument
%rax	Return value	Return value

```
pcount r:
 movl $0, %eax
 testq %rdi, %rdi
 je .L6
 pushq %rbx
 movq %rdi, %rbx
 andl $1, %ebx
 shrq %rdi # (by 1)
 call
        pcount r
 addq %rbx, %rax
        %rbx
 popq
.L6:
 rep; ret
```

## **Recursive Function Register Save**

```
pcount r:
 movl $0, %eax
        %rdi, %rdi
 testq
 je .L6
 pushq %rbx
 movq %rdi, %rbx
 andl $1, %ebx
 shrq %rdi # (by 1)
 call
        pcount r
 addq %rbx, %rax
        %rbx
 popq
.L6:
 rep; ret
```

Register	Use(s)	Туре
%rdi	x	Argument



## **Recursive Function Call Setup**

```
/* Recursive popcount */
long pcount r(unsigned long x) {
  if (x == 0)
    return 0;
 else
    return (x & 1)
           + pcount r(x >> 1);
```

Register	Use(s)	Туре
%rdi	x >> 1	Recursive argument
%rbx	x & 1	Callee-saved

```
pcount r:
 movl $0, %eax
 testq %rdi, %rdi
 je .L6
 pushq %rbx
 movq %rdi, %rbx
 andl $1, %ebx
 shrq %rdi # (by 1)
 call
        pcount r
 addq %rbx, %rax
        %rbx
 popq
.L6:
 rep; ret
```

## **Recursive Function Call**

Register	Use(s)	Туре
%rbx	x & 1	Callee-saved
%rax	Recursive call return value	

```
pcount r:
 movl $0, %eax
 testq %rdi, %rdi
 je .L6
 pushq %rbx
 movq %rdi, %rbx
 andl $1, %ebx
 shrq %rdi # (by 1)
 call pcount r
 addq %rbx, %rax
        %rbx
 popq
.L6:
 rep; ret
```

## **Recursive Function Result**

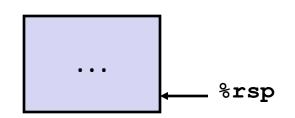
Register	Use(s)	Туре
%rbx	x & 1	Callee-saved
%rax	Return value	

```
pcount r:
 movl $0, %eax
 testq %rdi, %rdi
 je .L6
 pushq %rbx
 movq %rdi, %rbx
 andl $1, %ebx
 shrq %rdi # (by 1)
 call
        pcount r
 addq %rbx, %rax
        %rbx
 popq
.L6:
 rep; ret
```

## **Recursive Function Completion**

```
Register Use(s) Type
%rax Return value Return value
```

```
pcount r:
 movl $0, %eax
        %rdi, %rdi
 testq
 je
        .L6
 pushq %rbx
 movq %rdi, %rbx
 andl $1, %ebx
        %rdi # (by 1)
 shrq
 call
        pcount r
 addq
        %rbx, %rax
        %rbx
 popq
.L6:
 rep; ret
```



## **Observations About Recursion**

### Handled Without Special Consideration

- Stack frames mean that each function call has private storage
  - Saved registers & local variables
  - Saved return pointer
- Register saving conventions prevent one function call from corrupting another's data
  - Unless the C code explicitly does so (e.g., buffer overflow in Lecture
     9)
- Stack discipline follows call / return pattern
  - If P calls Q, then Q returns before P
  - Last-In, First-Out

#### Also works for mutual recursion

P calls Q; Q calls P

## x86-64 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 %rax
- Pointers are addresses of values
  - On stack or global

