Machine-Level Programming III: Procedures

15-213/14-513/15-513: Introduction to Computer Systems 6th Lecture, May 25, 2023

Instructors:

Brian Railing

Reminder: Condition Codes

■Single bit registers

- CF Carry Flag (for unsigned)SF Sign Flag (for signed)
- ZF Zero Flag OF Overflow Flag (for signed)

■jX and SetX isntructions

jΧ	Condition	Description
jmp	1	Unconditional
jе	ZF	Equal / Zero
jne	~ZF	Not Equal / Not Zero
js	SF	Negative
jns	~SF	Nonnegative
jg	~ (SF^OF) &~ZF	Greater (Signed)
jge	~(SF^OF)	Greater or Equal (Signed)
j1	(SF^OF)	Less (Signed)
jle	(SF^OF) ZF	Less or Equal (Signed)
ja	~CF&~ZF	Above (unsigned)
jb	CF	Below (unsigned)

SetX	Condition	Description
sete	ZF	Equal / Zero
setne	~ZF	Not Equal / Not Zero
sets	SF	Negative
setns	~SF	Nonnegative
setg	~ (SF^OF) &~ZF	Greater (Signed)
setge	~ (SF^OF)	Greater or Equal (Signed)
setl	(SF^OF)	Less (Signed)
setle	(SF^OF) ZF	Less or Equal (Signed)
seta	~CF&~ZF	Above (unsigned)
setb	CF	Below (unsigned)

Machine Level Programming – Control

C Control

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

Assembler Control

- Conditional jump
- Conditional move
- Indirect jump (via jump tables)
- Compiler generates code sequence to implement more complex control

Standard Techniques

- Loops converted to do-while or jump-to-middle form
- Large switch statements use jump tables
- Sparse switch statements may use decision trees (if-elseif-else)

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(...) {
      O(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(...) {
    = O(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];
}
```

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

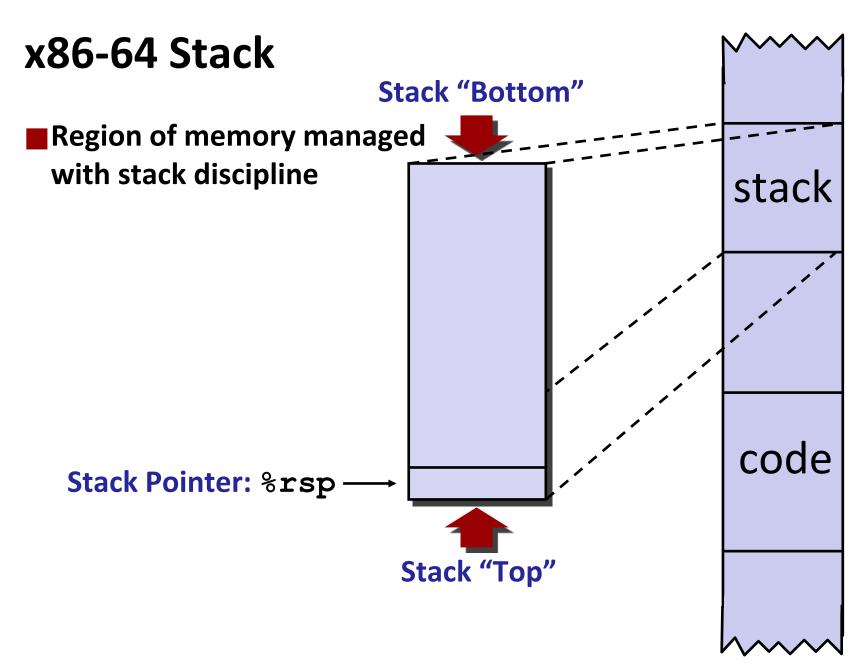
- Procedures
 - Stack Structure
 - Calling Conventions
 - Passing control
 - Passing data
 - Managing local data
 - Activity
 - If we have time: illustration of recursion

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)



code

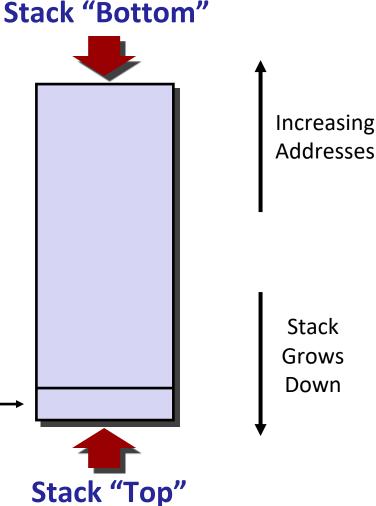


x86-64 Stack

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

Stack Pointer: %rsp →

Stack "Tor



x86-64 Stack: Push

Stack "Bottom" pushq Src Fetch operand at Src val Increasing Decrement %rsp by 8 **Addresses** Write operand at address given by %rsp Stack Grows Down **Stack Pointer:** %rsp

Stack "Top"

x86-64 Stack: Push

pushq Src

- Fetch operand at Src
- Decrement %rsp by 8
- Write operand at address given by %rsp

val Increasing **Addresses** Stack Grows Down **Stack Pointer:** Stack "Top"

Stack "Bottom"

x86-64 Stack: Pop

■popq Dest

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

Value is **copied**; it remains in memory at old %**rsp**

Stack Pointer:
%rsp

Stack "Top"

Stack "Bottom"

Increasing Addresses

Stack Grows Down

Today

Procedures

- Stack Structure
- Calling Conventions
 - Passing control
 - Passing data
 - Managing local data
- Activity
- If we have time: illustration of recursion

Code Examples

```
void multstore(long x, long y, long *dest)
   long t = mult2(x, y);
   *dest = t;
              0000000000400540 <multstore>:
                400540: push %rbx
                                              # Save %rbx
                400541: mov %rdx,%rbx
                                              # Save dest
                400544: call 400550 <mult2>
                                              # mult2(x,y)
                400549: mov %rax, (%rbx)
                                              # Save at dest
                40054c: pop %rbx
                                              # Restore %rbx
                40054d: ret
                                              # Return
```

```
long mult2(long a, long b)
                    0000000000400550 <mult2>:
 long s = a * b;
                      400550: mov %rdi,%rax
                                                     # a
 return s;
                      400553: imul %rsi,%rax
                                                     # a * b
                      400557: ret
                                                     # 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

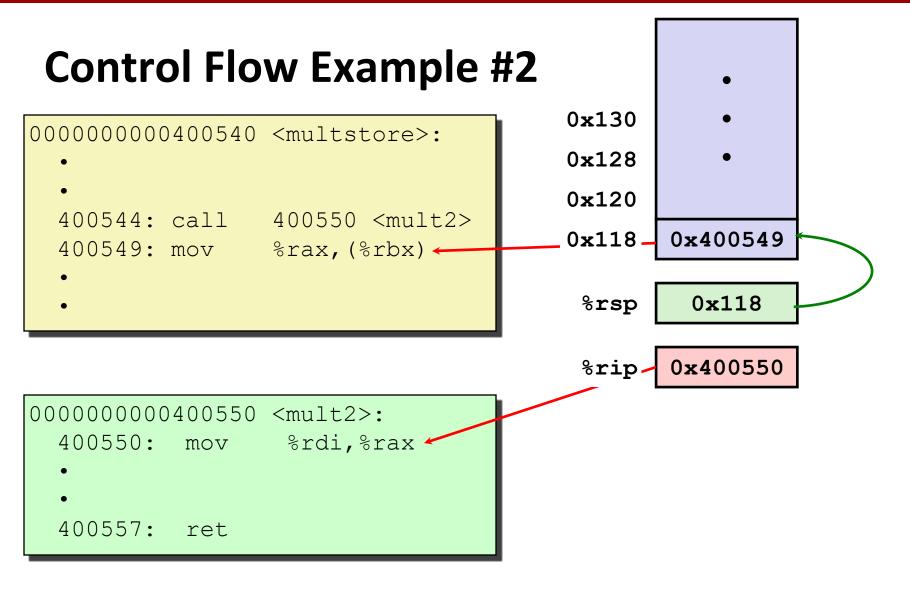
These instructions are sometimes printed with a q suffix

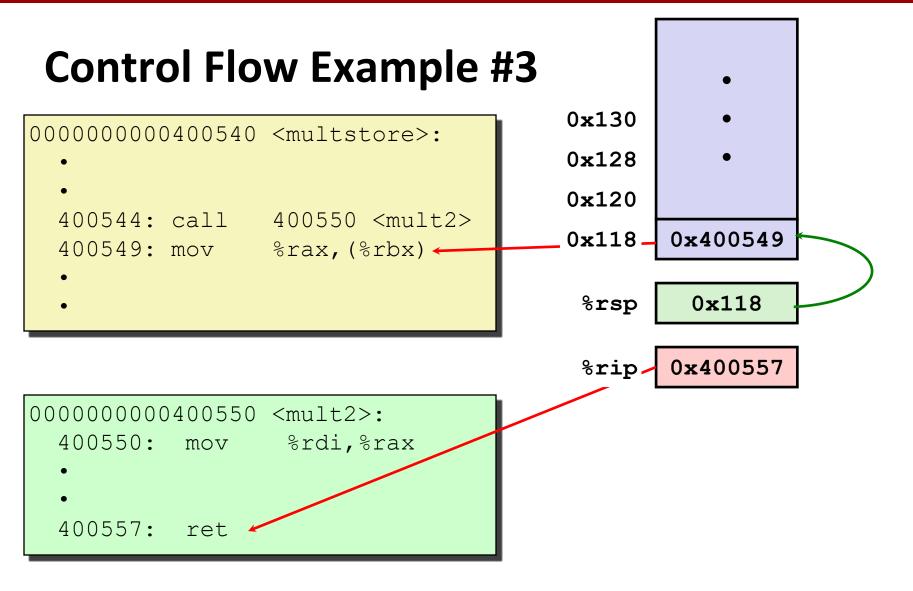
This is just to remind you that you're looking at 64-bit code

Control Flow Example #1

```
0x130
0x128
0x120
%rsp 0x120
%rip 0x400544
```

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





Control Flow Example #4

```
0x130

0x128

0x120

%rsp 0x120

%rip 0x400549
```

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

Today

Procedures

- Stack Structure
- Calling Conventions
 - Passing control
 - Passing data
 - Managing local data
- Activity
- If we have time: illustration of recursion

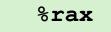
Procedure Data Flow

Registers

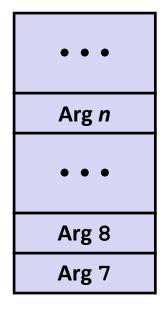
First 6 arguments



■ 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: ret # Return
```

Today

Procedures

- Stack Structure
- Calling Conventions
 - Passing control
 - Passing data
 - Managing local data
- Activity
- If we have time: 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*

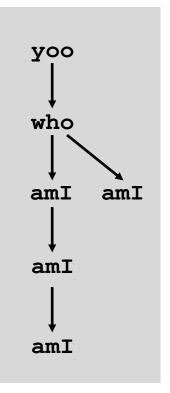
state for single procedure instantiation

Call Chain Example

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

Procedure amI () is recursive

Example Call Chain



Stack Frames

Contents

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

Frame Pointer: %rbp ——
(Optional)

Stack Pointer: %rsp

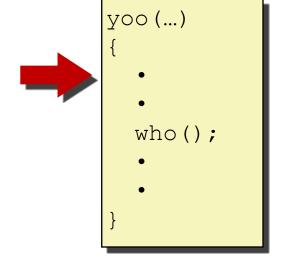
Previous Frame

Frame for proc

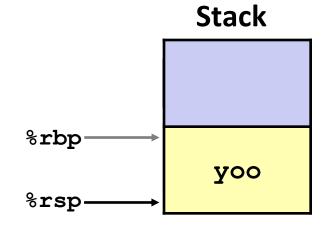
■ Management

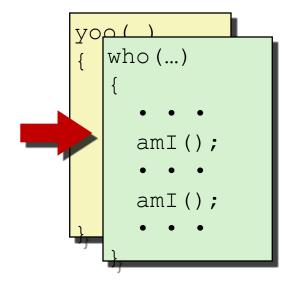
- Space allocated when enter procedure
 - "Set-up" code
 - Includes push by call instruction
- Deallocated when return
 - "Finish" code
 - Includes pop by ret instruction



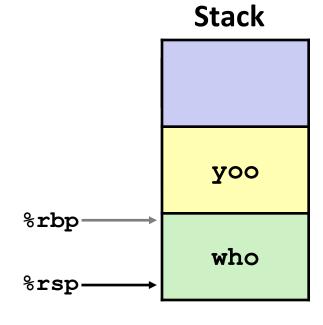


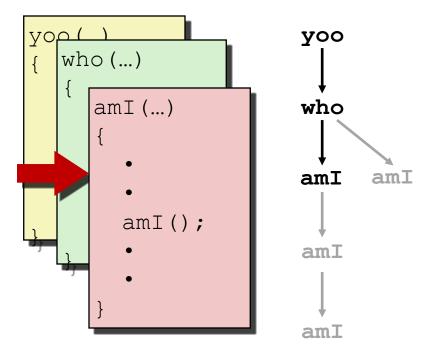


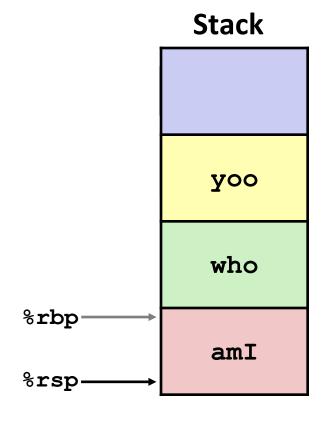


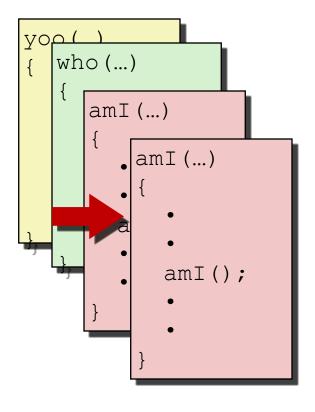


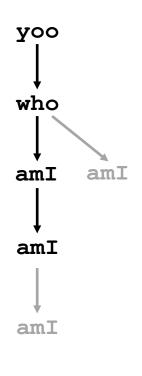


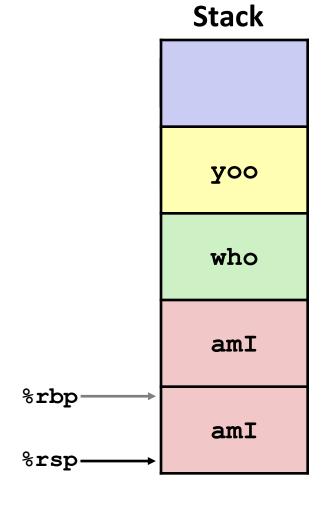


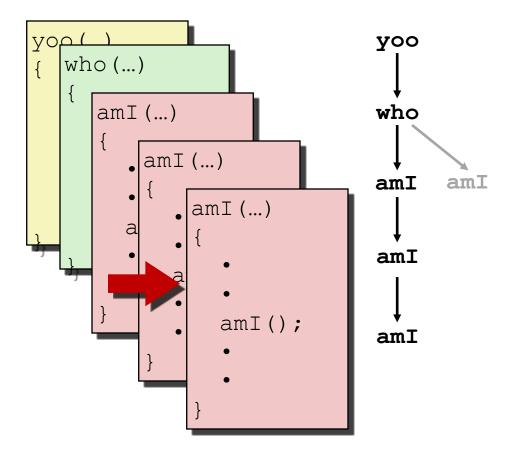


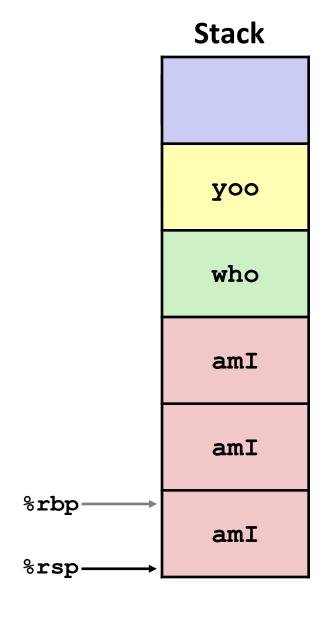


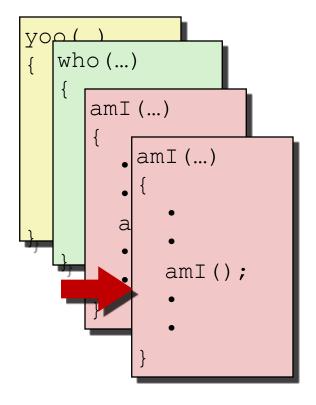


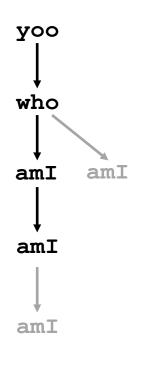


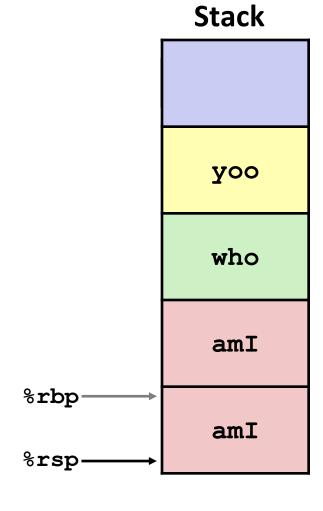


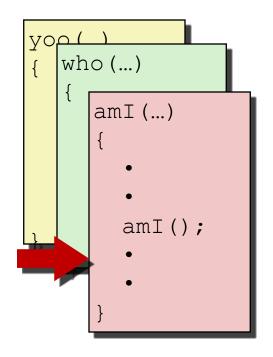


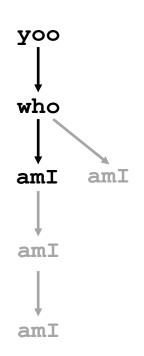


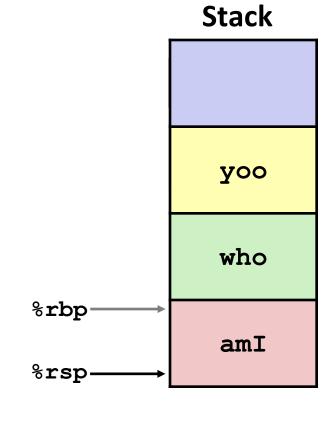


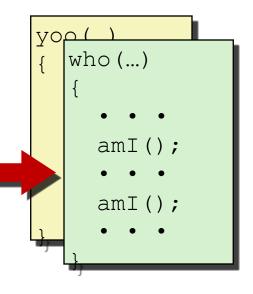




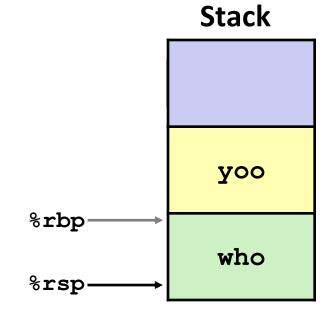


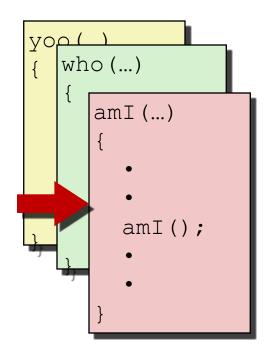


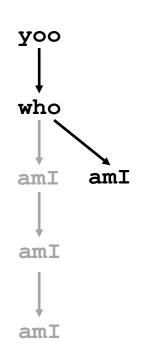


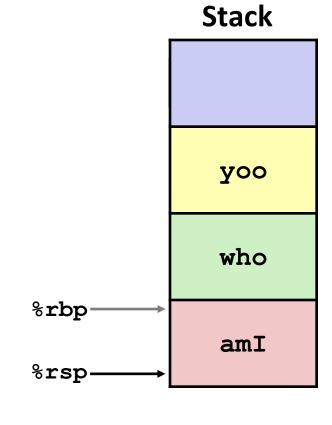


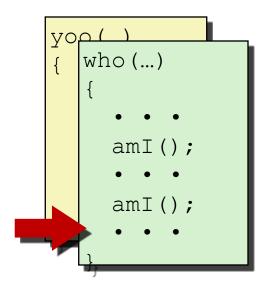


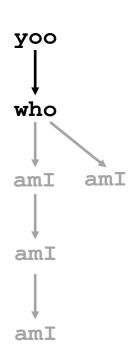


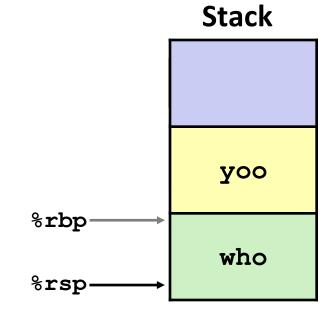


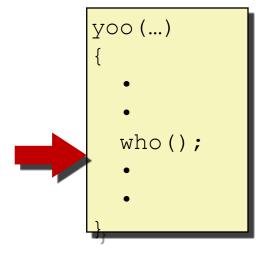




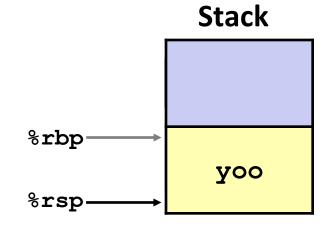












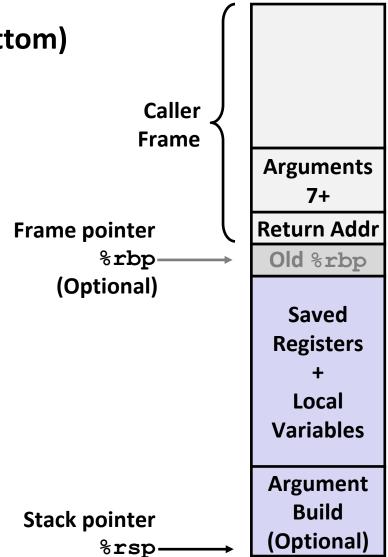
x86-64/Linux Stack Frame

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 p
%rsi	Argument val , y
%rax	x, Return value

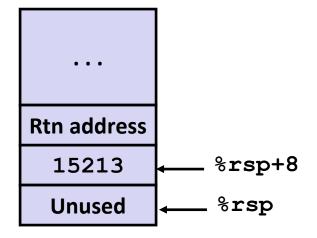
long call_incr() { long v1 = 15213; long v2 = incr(&v1, 3000); return v1+v2; }

Initial Stack Structure

```
Rtn address ← %rsp
```

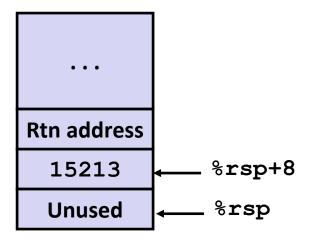
```
call_incr:
    subq $16, %rsp
    movq $15213, 8(%rsp)
    movl $3000, %esi
    leaq 8(%rsp), %rdi
    call incr
    addq 8(%rsp), %rax
    addq $16, %rsp
    ret
```

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
15213 %rsp+8
```

Aside 1: movl \$3000, %esi

- Remember, movl -> %exx zeros out high order 32 bits.
 - Why use movl instead of movq? 1 byte shorter.

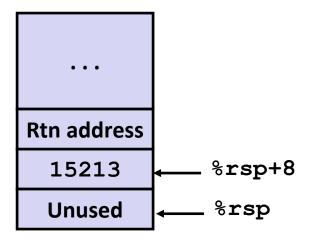
```
movl $3000, %esi
leaq 8(%rsp), %rdi
call incr
addq 8(%rsp), %rax
addq $16, %rsp
ret
```

%rdi	&v1
%rsi	3000

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

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

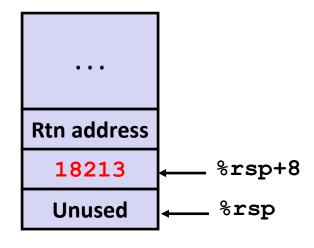
```
call_incr:
    subq    $16, %rsp
    movq    $15213, 8(%rsp)
    movl    $3000, %esi
    leaq    8(%rsp), %rdi
    call    incr
    addq    8(%rsp), %rax
    addq    $16, %rsp
    ret
```



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

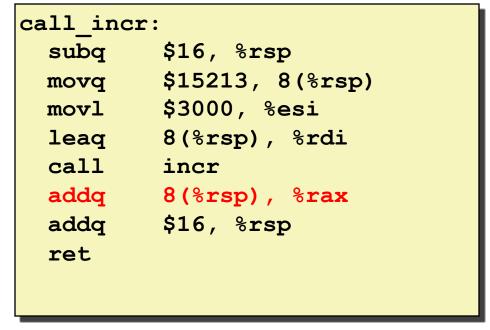
```
long call_incr() {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return v1+v2;
}
```

```
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

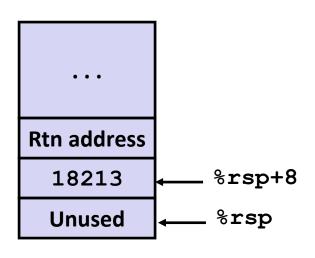
```
long call_incr() {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return v1+v2;
}
```



Register	Use(s)
%rax	Return value

Stack Structure

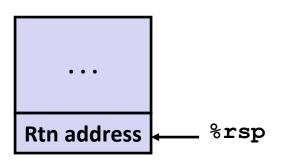
```
long call_incr() {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return v1+v2;
}
```



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	

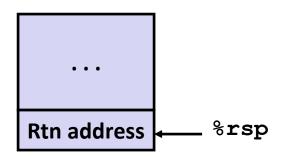
Register	Use(s)
%rax	Return value

Updated Stack Structure



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

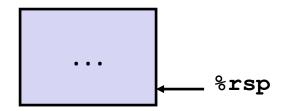
Updated Stack Structure



```
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

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

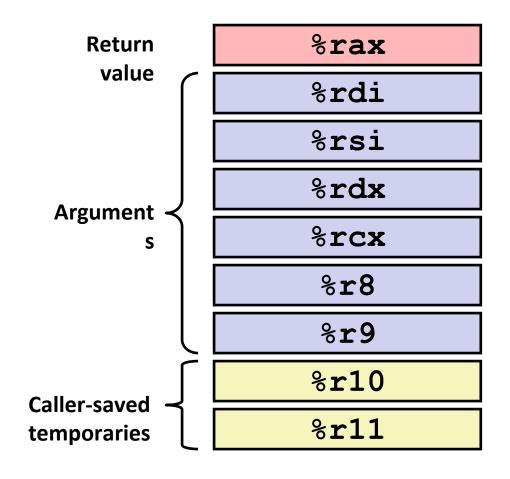
Register Saving Conventions

- ■When procedure yoo calls who:
 - yoo is the caller
 - who is the callee
- Can register be used for temporary storage?
- Conventions
 - "Caller Saved" (aka "Call-Clobbered")
 - Caller saves temporary values in its frame before the call
 - "Callee Saved" (aka "Call-Preserved")
 - 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

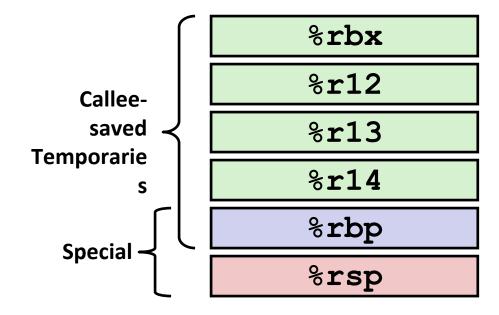
- Callee-saved
- Callee must save & restore

■%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



Activity Time!

Today

Procedures

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

Recursive Function

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

Recursive Function Terminal Case

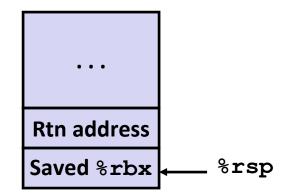
Register	Use(s)	Туре
%rdi	ж	Argument
%rax	Return value	Return value

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

Register	Use(s)	Туре
%rdi	x	Argument



Recursive Function Call Setup

Register	Use(s)	Туре
%rdi	x >> 1	Rec. 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
 call
        pcount r
        %rbx, %rax
 addq
        %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
 call
        pcount r
        %rbx, %rax
 addq
        %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
 call
        pcount r
 addq %rbx, %rax
        %rbx
 popq
.L6:
 rep; ret
```

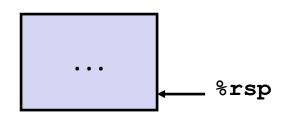
Recursive Function Completion

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

```
movl
        $0, %eax
        %rdi, %rdi
 testq
 iе
        .L6
 pushq %rbx
 movq %rdi, %rbx
 andl
        $1, %ebx
 shrq %rdi
 call
        pcount r
 addq
        %rbx, %rax
        %rbx
 popq
.L6:
```

pcount r:

Register	Use(s)	Туре
%rax	Return value	Return value



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

