Machine-Level Programming III: Procedures

COMP400727: Introduction to Computer Systems

Danfeng Shan Xi'an Jiaotong University

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

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

```
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

```
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

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

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

Today

Procedures

Stack Structure

Calling Conventions

Passing control

Passing data

Managing local data

If we have time: illustration of recursion

x86-64 Stack

Region of memory managed with stack discipline

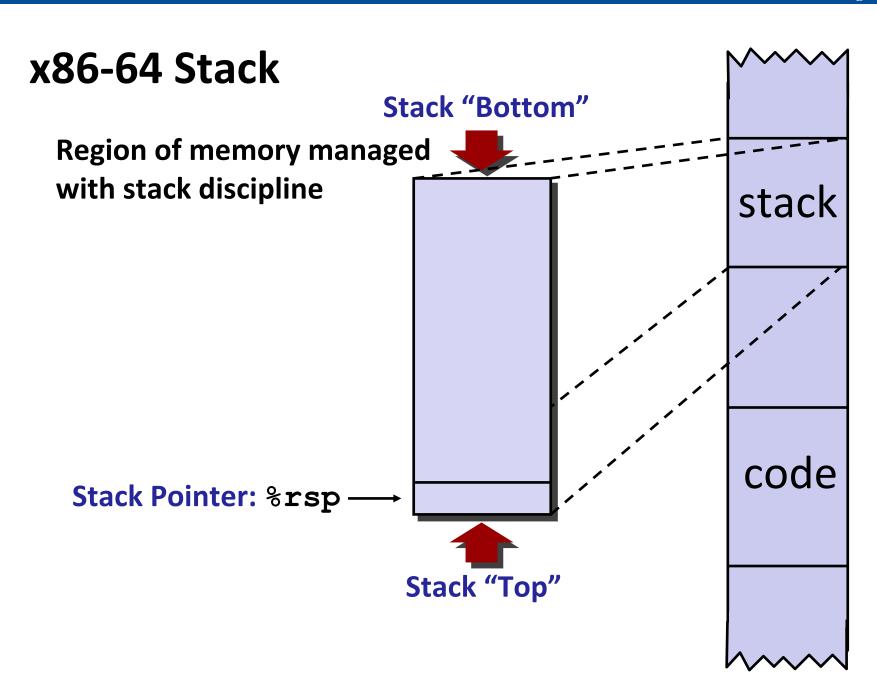
Memory viewed as array of bytes.

Different regions have different purposes.

(Like ABI, a policy decision)

stack

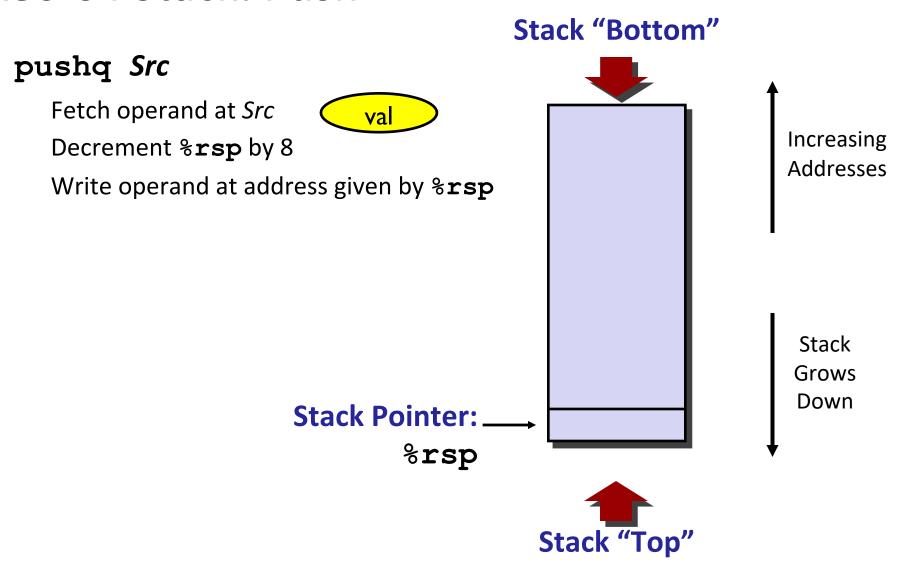
code



x86-64 Stack

Stack "Bottom" Region of memory managed with stack discipline **Increasing Grows toward lower addresses Addresses** Register %rsp contains lowest stack address address of "top" element Stack **Grows** Down Stack Pointer: %rsp Stack "Top"

x86-64 Stack: Push



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:** Stack "Top"

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

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

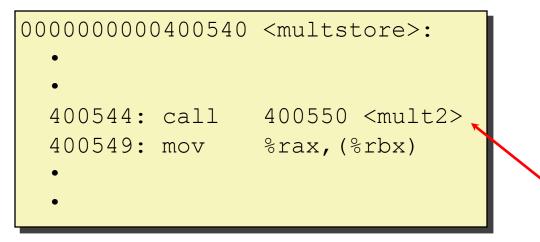
Pop address from stack

Jump to address

These instructions are sometimes printed with a q suffix

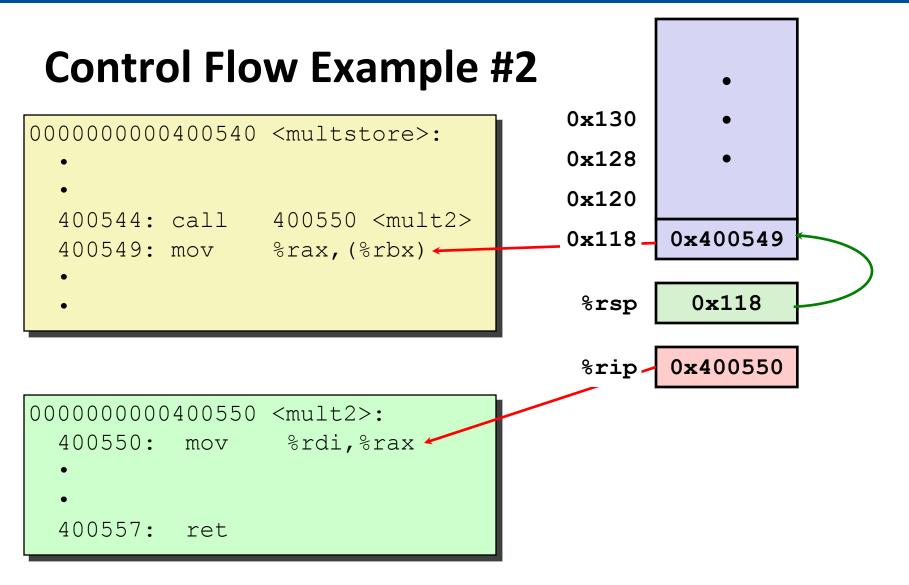
This is just to remind you that you're looking at 64-bit code

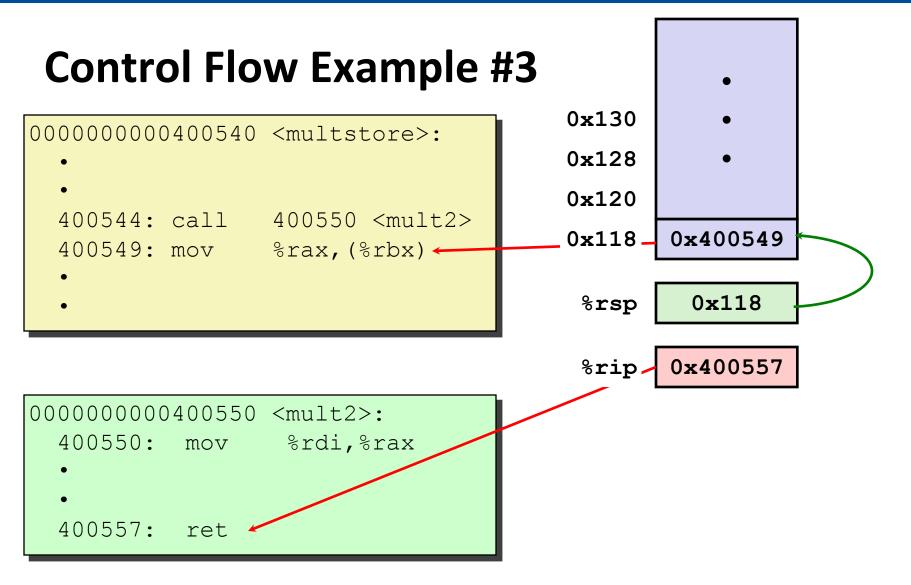
Control Flow Example #1



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

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

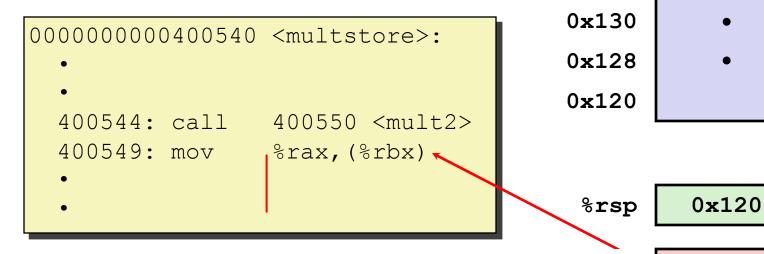




 0×400549

%rip

Control Flow Example #4



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

Today

Procedures

Stack Structure

Calling Conventions

Passing control

Passing data

Managing local data

If we have time: illustration of recursion

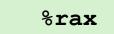
Procedure Data Flow

Registers

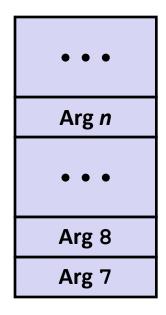
First 6 arguments



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

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*

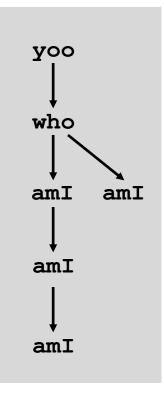
state for single procedure instantiation

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)

Previous Frame

Frame Pointer: %rbp

(Optional)

Frame for proc

Stack Pointer: %rsp

Management

Space allocated when enter procedure

"Set-up" code

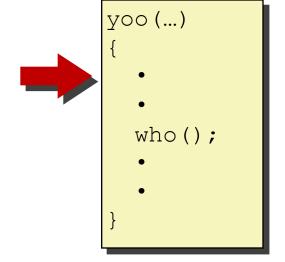
Includes push by call instruction

Deallocated when return

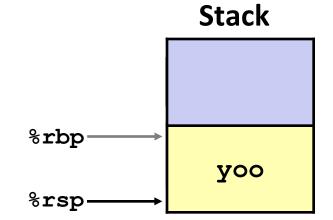
"Finish" code

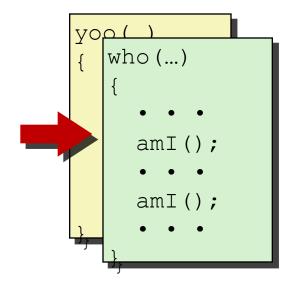
Includes pop by ret instruction



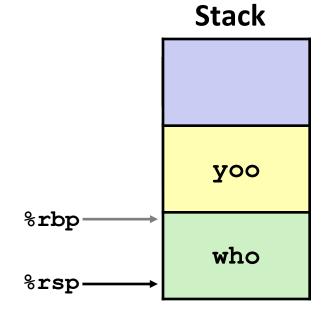


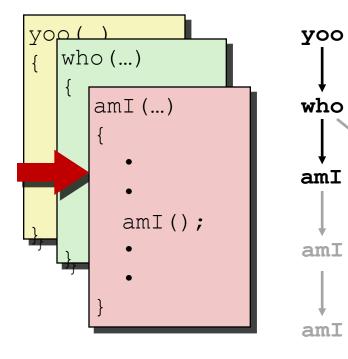




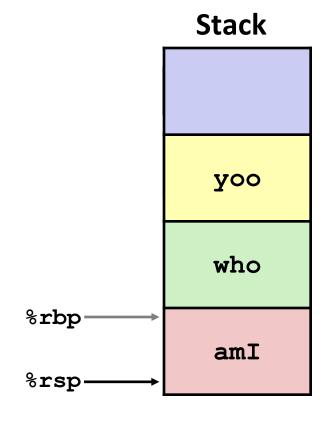


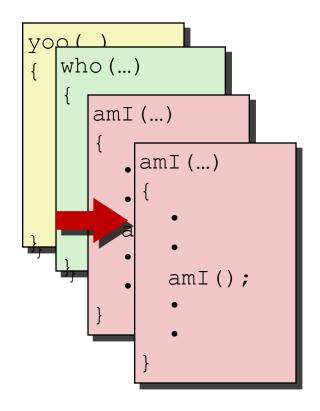


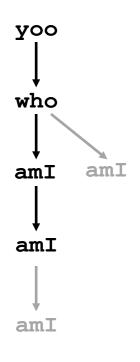


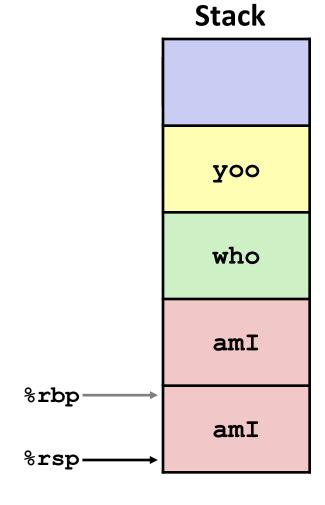


amI

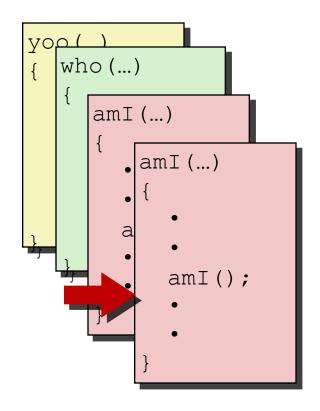


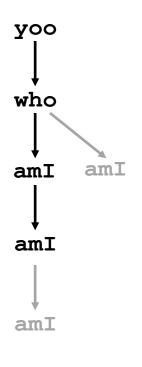


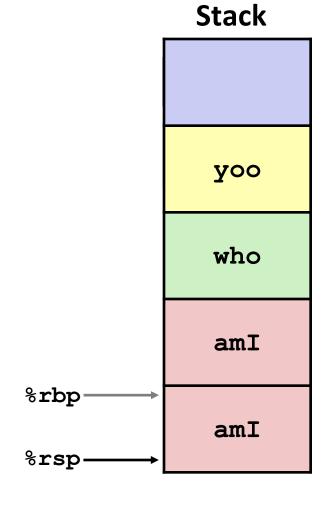




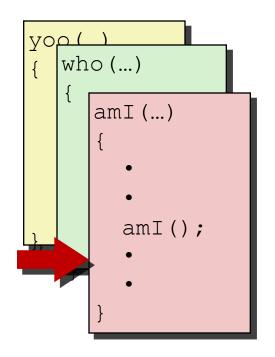
Stack **Example** yop() yoo who (...) yoo amI (...) who • amI (...) who amIamI• amI (...) amIamIamI(); amIamI%rbp amI%rsp

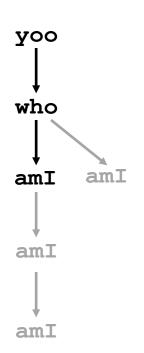


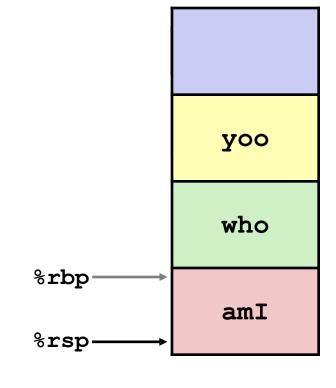


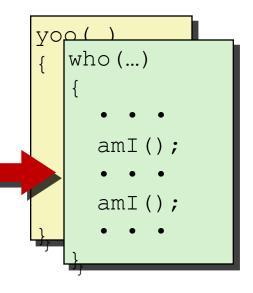


Stack

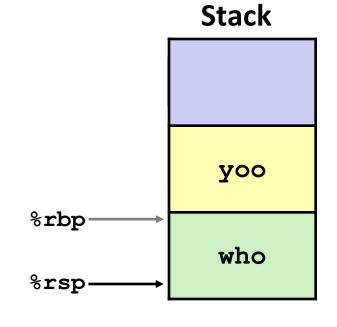


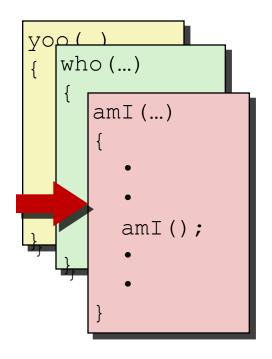


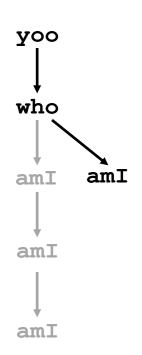


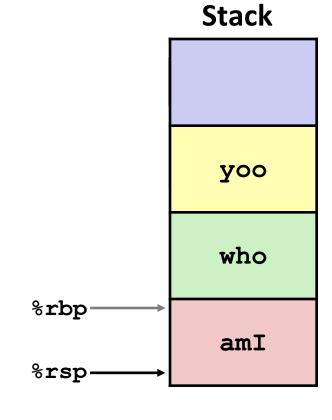




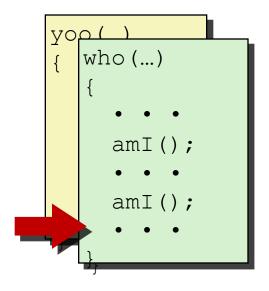


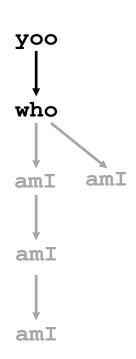


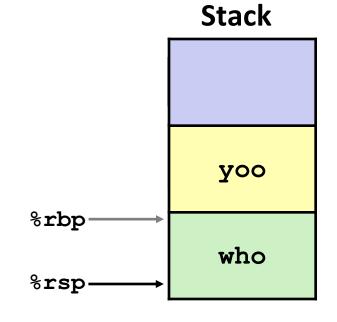




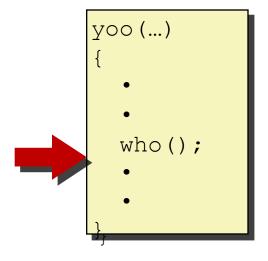
Example



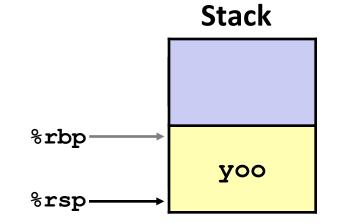




Example







x86-64/Linux Stack Frame

Caller Stack Frame

Arguments for this call

Return address

Pushed by **call** instruction

Current Stack Frame

Old frame pointer (optional)

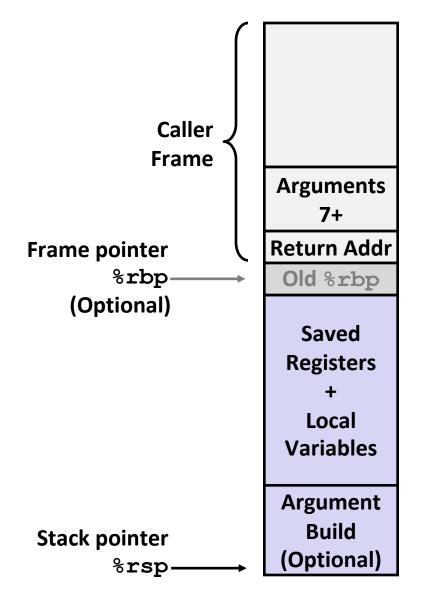
Saved register context

Local variables

If can't keep in registers

"Argument build:"

Parameters for function about to 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	

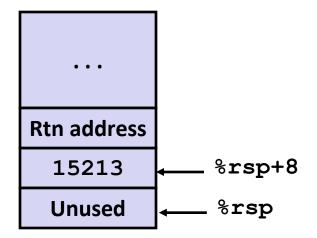
```
long call_incr() {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return v1+v2;
}
```

Initial Stack Structure

```
Rtn address ← %rsp
```

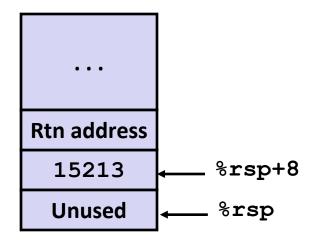
```
call_incr:
    subq    $16, %rsp
    movq    $15213, 8(%rsp)
    movl    $3000, %esi
    leaq    8(%rsp), %rdi
    call    incr
    addq    8(%rsp), %rax
    addq    $16, %rsp
    ret
```

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

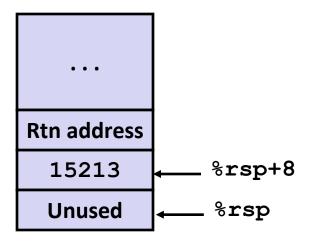
```
Stack Structure
long call incr() {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return v1+v2;
                                    Rtn address
                                                  %rsp+8
                                      15213
                                                  %rsp
       Aside 2: leaq 8(%rsp), %rdi
ca:
  Computes %rsp+8
                                               se(s)

    Actually, used for what it is meant!

  leaq 8(%rsp), %rdi
                                              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;
}
```

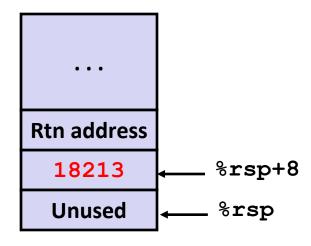
```
call_incr:
    subq    $16, %rsp
    movq    $15213, 8(%rsp)
    movl    $3000, %esi
    leaq    8(%rsp), %rdi
    call    incr
    addq    8(%rsp), %rax
    addq    $16, %rsp
    ret
```



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

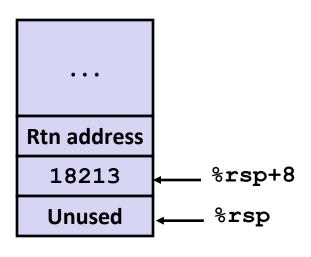
```
long call_incr() {
   long v1 = 15213;
   long v2 = incr(&v1, 3000);
   return v1+v2;
}
```

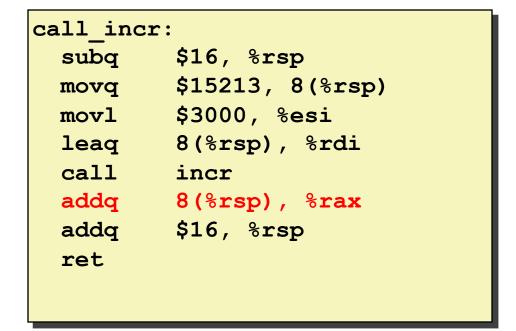
```
call_incr:
    subq    $16, %rsp
    movq    $15213, 8(%rsp)
    movl    $3000, %esi
    leaq    8(%rsp), %rdi
    call    incr
    addq    8(%rsp), %rax
    addq    $16, %rsp
    ret
```



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

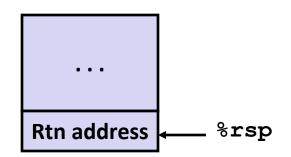
Stack Structure

```
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	

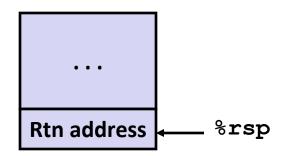
Register	Use(s)
%rax	Return value

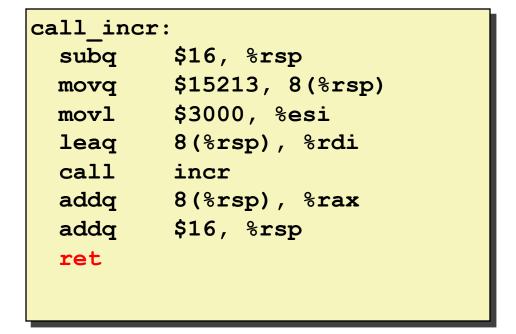
Updated Stack Structure



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

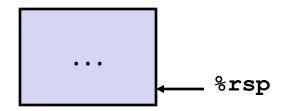
Updated Stack Structure





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

• • •
```

```
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")

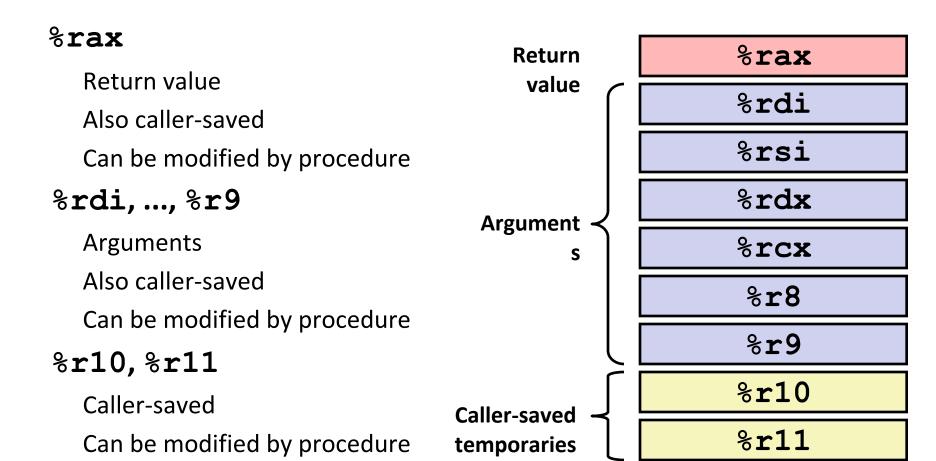
Caller saves temporary values in its frame before the call

"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



x86-64 Linux Register Usage #2

%rbx, %r12, %r13, %r14

Callee-saved

Callee must save & restore

%rbp

Callee-saved

Callee must save & restore

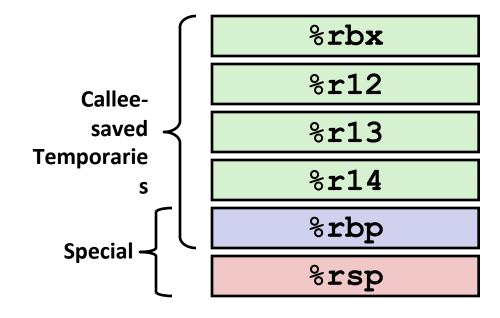
May be used as frame pointer

Can mix & match

%rsp

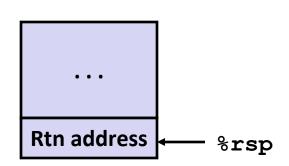
Special form of callee save

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

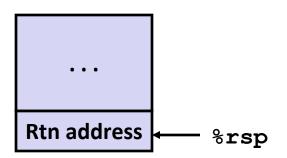


- X comes in register %rdi.
- We need %rdi for the call to incr.
- Where should be put x, so we can use it after the call to incr?

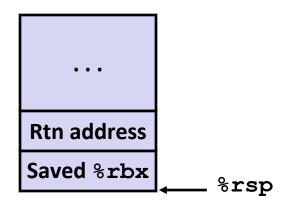
```
long call_incr2(long x) {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return x+v2;
}
```

```
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



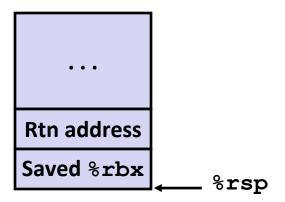
Resulting Stack Structure



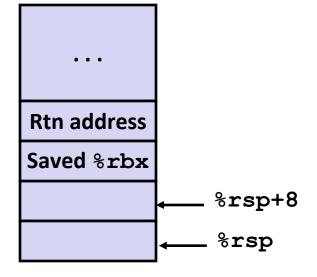
```
long call_incr2(long x) {
    long v1 = 15213;
    long v2 = incr(&v1, 3000);
    return x+v2;
}
```

```
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

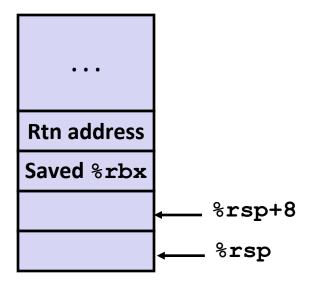


Resulting Stack Structure



```
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
```



- X 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
```

- X 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
```

```
Rtn address
Saved %rbx

18213 ← %rsp+8

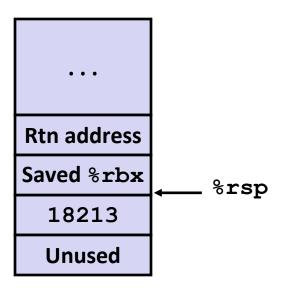
Unused %rsp
```

- X Is safe in %rbx
- Return result in %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

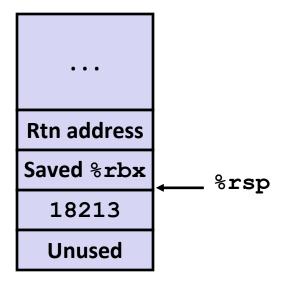


• Return result in %rax

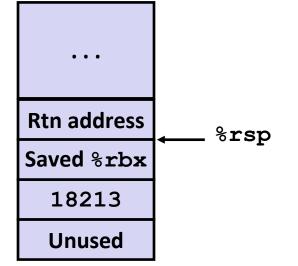
Initial Stack Structure

```
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
```



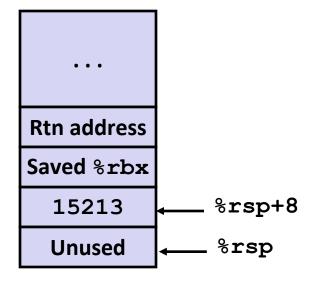
final Stack Structure



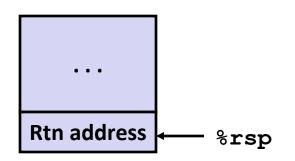
long call_incr2(long x) { long v1 = 15213; long v2 = incr(&v1, 3000); return x+v2; }

```
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
```

Resulting Stack Structure



Pre-return Stack Structure



Today

Procedures

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
 call
        pcount r
 addq
         %rbx, %rax
         %rbx
 popq
.L6:
 rep; ret
```

Recursive Function Terminal Case

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

pcount r:

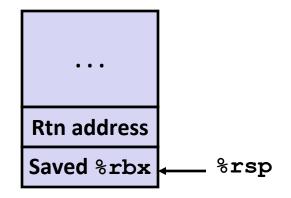
rep; ret

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

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

```
/* 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	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
        %rdi
 shrq
 call
        pcount r
        %rbx, %rax
 addq
        %rbx
 popq
.L6:
 rep; ret
```

Recursive Function Result

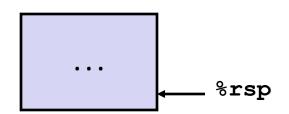
Register	Use(s)	Туре
%rbx	x & 1	Callee-saved
%rax	Return value	

```
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

```
pcount r:
 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:
 rep; ret
```

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



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 Caller **Frame** Recursion (& mutual recursion) handled by **Arguments** normal calling conventions 7+ **Return Addr** Can safely store values in local stack frame and in %rbp→ Old %rbp callee-saved registers (Optional) Put function arguments at top of stack Saved Result return in %rax Registers Pointers are addresses of values Local On stack or global **Variables Argument**

Build

%rsp