CS207: File Processing Term I/2018-19

Lecture 16:

A Pragmatic View of Computer Systems

x86-64 and friends II

Sunsern Cheamanunkul



MUIC: File Pro

Today's Topic

- More C to Assembly
 - Loops
 - The switch-case translation
 - Function calls

Recap: Everything is bits!

- Seen many data types so far:
 - Integers:
 - char/short/int/long (encoding as unsigned or two's complement signed)
 - Letters/punctutation/etc:
 - char (ASCII encoding)
 - Real numbers:
 - float/double (IEEE floating point encoding, didn't discuss)
 - Memory addresses:
 - pointer types (unsigned long encoding)
 - the code itself!
 - Instructions

MUIC: File

LOOPING: FOR, WHILE, DO-WHILE

"Do-While" Loop Example

C Code

```
long poount do
  (unsigned long x) {
  long result = 0;
   result += x & 0x1;
   x >>= 1:
  } while (x);
  return result:
```

Goto Version

```
long poount goto
  (unsigned long x) {
 long result = 0;
 result += x \& 0x1;
 x >>= 1:
 if(x) goto loop;
 return result;
```

MUIC: File Pro.

- Count number of 1's in argument x ("popcount")
- Use conditional branch to either continue looping or to exit loop

General "Do-While" Translation

C Code

```
do
  Body
  while (Test);
```

Body:

```
Statement,;
Statement,;
Statement;
```

Goto Version

```
loop:
  Body
  if (Test)
    goto loop
```

"Do-While" Loop Compilation **Goto Version** long poount goto

```
(unsigned long x) {
long result = 0:
result += x \& 0x1;
x >>= 1;
if(x) goto loop;
return result:
```

```
Use(s)
Register
%rdi
              Argument x
              result
%rax
```

```
# result = 0
        $0, %eax
movl
                   # loop:
movq
        %rdi, %rdx
andl
       $1, %edx
                   # t = x & 0x1
addq
       %rdx, %rax #
                      result += t
shrq
       %rdi
                   # x >>= 1
        . ь2
                   # if (x) goto loop
ine
rep; ret
```

General "While" Translation #1

- "Jump-to-middle" translation
- Used with **–Og** (optimized debugging experience) Goto Version

While version

```
while (Test)
  Body
```



```
goto test;
loop:
  Body
test:
  if (Test)
    goto loop;
done:
```

MUIC: File Pro.

While Loop Example #1

C Code

```
long pcount_while
  (unsigned long x) {
  long result = 0;
  while (x) {
    result += x & 0x1;
    x >>= 1;
  }
  return result;
}
```

Jump to Middle

```
long poount_goto_jtm
  (unsigned long x) {
  long result = 0;
  goto test;
  loop:
    result += x & 0x1;
    x >>= 1;
  test:
    if(x) goto loop;
    return result;
}
```

- Compare to do-while version of function
- Initial goto starts loop at test

While Loop Example #2

C Code

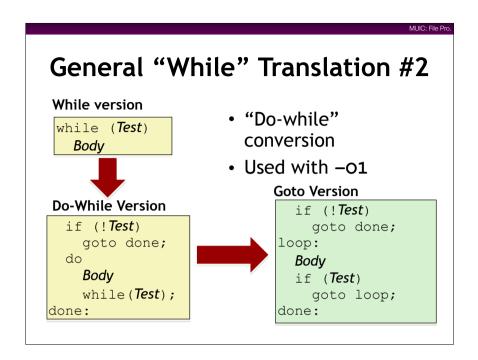
```
long pcount_while
  (unsigned long x) {
  long result = 0;
  while (x) {
    result += x & 0x1;
    x >>= 1;
  }
  return result;
}
```

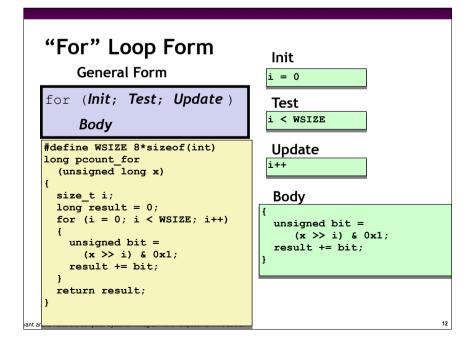
Do-While

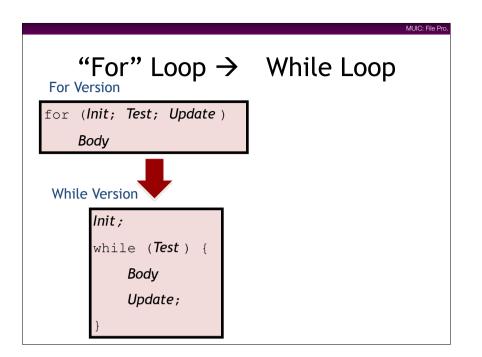
```
long pcount_goto_dw
  (unsigned long x) {
  long result = 0;
  if (!x) goto done;
  loop:
    result += x & 0x1;
    x >>= 1;
  if(x) goto loop;
  done:
    return result;
}
```

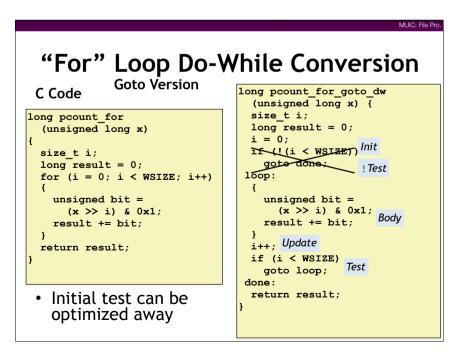
MUIC: File Pro

- Compare to do-while version of function
- Initial conditional guards entrance to loop









For-While Conversion (unsigned long x) Init i = 0 size t i; long result = 0; Test i = 0: while (i < WSIZE) i < WSIZE unsigned bit = Update (x >> i) & 0x1;result += bit: i++ Body return result: unsigned bit = (x >> i) & 0x1;result += bit:

MUIC: File Pro

Summarizing

- C Control
- if-then-else
- do-while
- while, for
- 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

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

```
Switch Statement
Example

    Multiple case

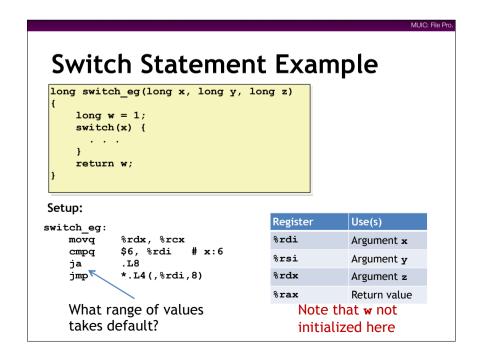
   lahels
   - Here: 5 & 6

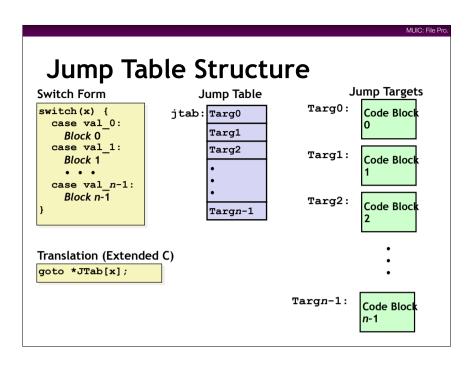
    Fall through cases

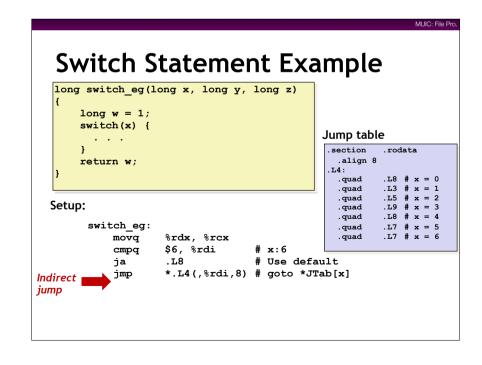
   - Here: 2

    Missing cases

   - Here: 4
```







Assembly Setup Explanation

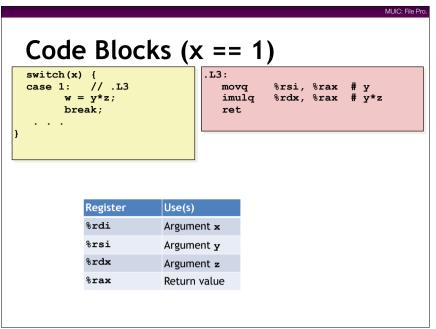
- Table Structure
- Each target requires 8 bytes
- Base address at . L4
- Jumping
- Direct: jmp .L8
- Jump target is denoted by label .L8
- Indirect: jmp *.L4(,%rdi,8)
- Start of jump table: .L4
- Must scale by factor of 8 (addresses are 8 bytes)
- Fetch target from effective Address .L4 + x*8
 - Only for $0 \le \mathbf{x} \le 6$

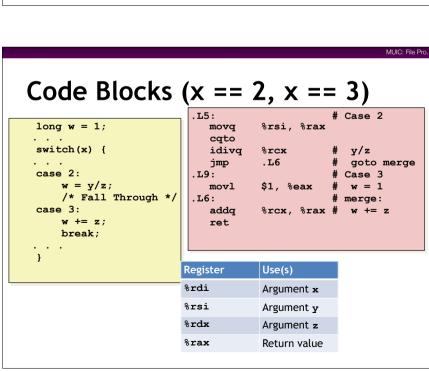
```
Jump table
 .section
            .rodata
  .align 8
 .L4:
            .L8 \# x = 0
   . quad
            .L3 \# x = 1
   . quad
   . quad
            .L5 \# x = 2
   .quad
            .L9 \# x = 3
   . quad
            .L8 \# x = 4
   . quad
            .L7 # x = 5
            .L7 \# x = 6
```

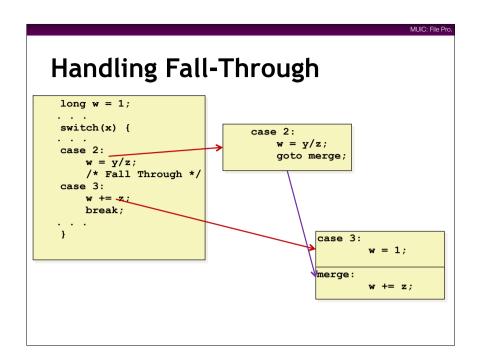
MUIC: File Pro Finding Jump Table in Binary 00000000004005e0 <switch eq>: 48 89 d1 4005e0: %rdx,%rcx 48 83 ff 06 4005e3: \$0x6,%rdi 4005e7: 77 2b 400614 <switch eg+0x34> ff 24 fd f0 07 40 00 4005e9: jmpq *0x4007f0(,%rdi,8) 4005f0: 48 89 f0 %rsi,%rax 4005f3: 48 Of af c2 %rdx,%rax imul 4005f7: 4005f8 · 48 89 f0 %rsi,%rax mov 4005fb: 48 99 cqto 48 f7 f9 4005fd: idiv %rcx eb 05 400600: 400607 <switch eg+0x27> jmp 400602: b8 01 00 00 00 \$0x1, %eax 48 01 c8 400607: add %rcx,%rax 40060a: retq 40060b: b8 01 00 00 00 mov \$0x1,%eax 400610: 48 29 d0 sub %rdx,%rax 400613: retq ъ8 02 00 00 00 400614: mov \$0x2, %eax 400619: retq

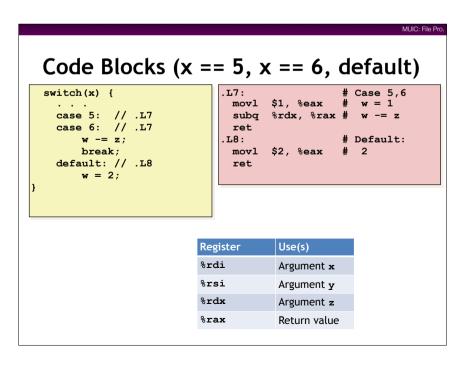
```
Jump Table
Jump table
                                      switch(x) {
                                      case 1:
                                                     // .L3
     .section
               .rodata
                                          w = v*z;
       .align 8
                                          break:
     т.4 •
               .L8 \# x = 0
       . guad
              .L3 # x = 0
                                      case 2:
                                                     // .L5
       .quad
                                          w = y/z;
       .quad
               .L5 \# x = 2
                                          /* Fall Through */
       .quad
               .L9 \# x = 3
       . quad
               .L8 \# x = 4
                                      case 3:
                                                     // .L9
               .L7 # x = 5
       . guad
                                          w += z:
       .quad
               .L7 \# x = 6
                                          break;
                                      case 5:
                                      case 6:
                                                     // .L7
                                          w -= z;
                                          break:
                                      default:
                                                     // .L8
                                          w = 2;
```

```
MUIC: File Pro
 Finding Jump Table in Binary (cont.)
  00000000004005e0 <switch eq>:
 4005e9:
              ff 24 fd f0 07 40 00 jmpq *0x4007f0(,%rdi,8)
 . . .
% qdb switch
(gdb) x /8xg 0x4007f0
0x4007f0:
              0x0000000000400614
                                    0x00000000004005f0
0x400800:
              0x0000000004005f8
                                    0x0000000000400602
0x400810:
              0x0000000000400614
                                    0x000000000040060b
0x400820:
              0x000000000040060b
                                    0x2c646c25203d2078
(gdb)
```









MUIC: File Pr

MUIC: File Pro

Techniques for Switches

- •Large switch statements use jump tables
- •Sparse switch statements may use decision trees (if-else..if-else..if-else)

FUNCTION CALLS

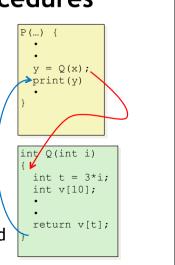
Mechanisms in Procedures

- · 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

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

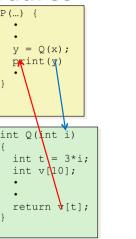
Mechanisms in Procedures

- 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



Mechanisms in Procedures

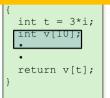
- Passing control
 - To beginning of procedure code
 - Back to return point
- Passing data
 - Procedure arguments
 - Return value
- Memory management
 - Allocate during procedure execution
 - Deallocate upon return
- Mechanisms all implemented with machine instructions



Mechanisms in Procedures

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

- Memory management
 - Allocate during procedure execution
 - Deallocate upon return
- Mechanisms all implemented with machine instructions



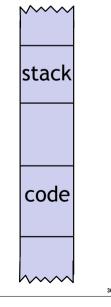
Today

■Procedures

- Stack Structure
- Calling Conventions
 - Passing control
 - Passing data
 - Managing local data

x86-64 Stack

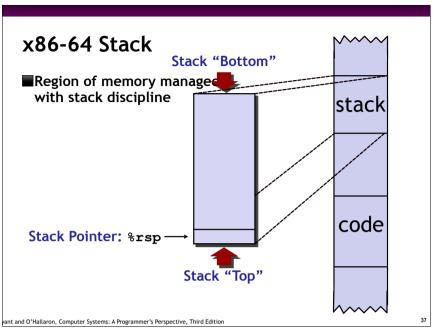
- Region of memory managed with stack discipline
 - Memory viewed as array of bytes.
 - Different regions have different purposes.
 - · (Like ABI, a policy decision)

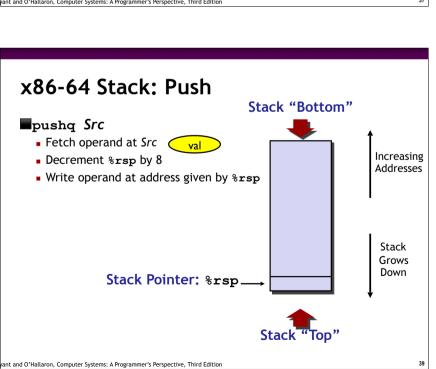


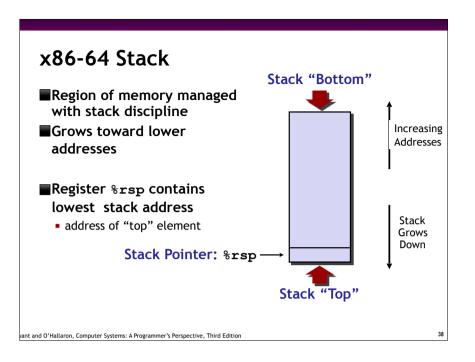
ant and O'Hallaron, Computer Systems: A Programmer's Perspective, Third Edition

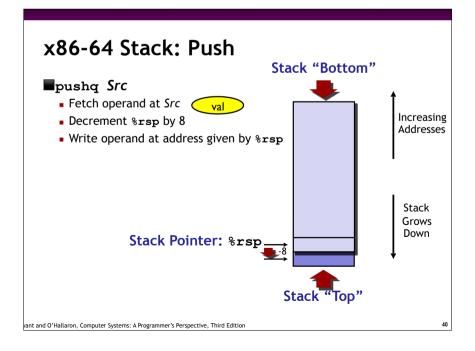
35

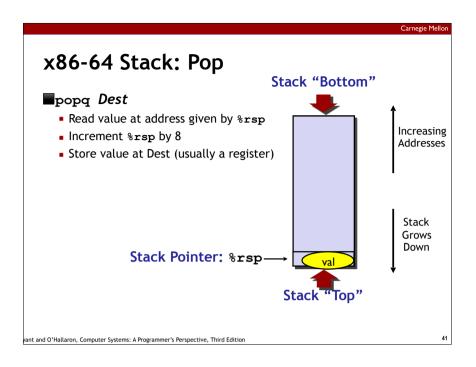
ant and O'Hallaron, Computer Systems: A Programmer's Perspective, Third Edition

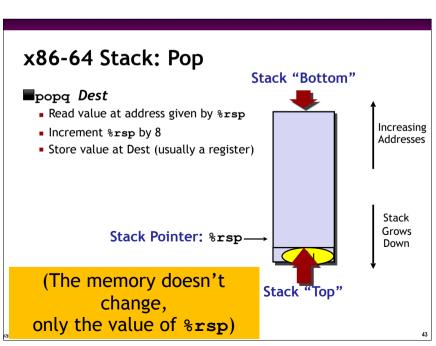


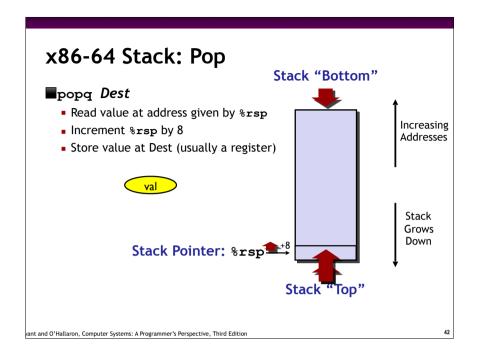












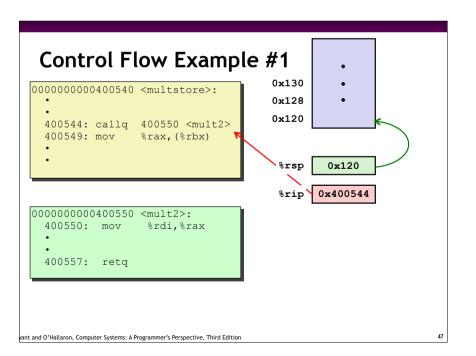
Today

■Procedures

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

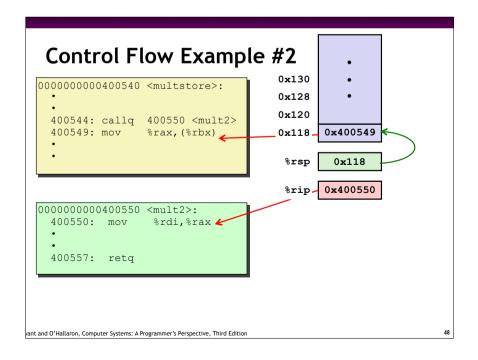
ant and O'Hallaron, Computer Systems: A Programmer's Perspective, Third Edition

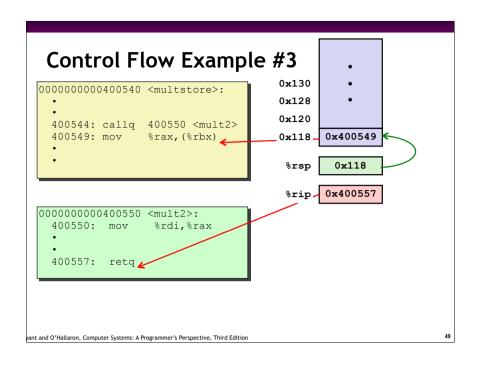
```
Code
void multstore (long x, long y, long
                                             Examples
 *dest)
    long t = mult2(x, y);
                0000000000400540 <multstore>:
                  400540: push %rbx
                                                   # Save %rbx
                  400541: mov %rdx,%rbx
                                                   # Save dest
                  400544: callq 400550 <mult2> # mult2(x,y)
                  400549: mov %rax, (%rbx)
                                                   # Save at dest
                  40054c: pop %rbx
                                                   # Restore %rbx
                  40054d: retg
                                                   # Return
long mult2(long a, long b)
                      00000000000400550 <mult2>:
  long s = a * b;
                        400550: mov
                                         %rdi,%rax
                                                         # a
  return s;
                        400553: imul
                                       %rsi,%rax
                                                         # a * b
                        400557: retq
                                                         # Return
vant and O'Hallaron, Computer Systems: A Programmer's Perspective, Third Edition
```

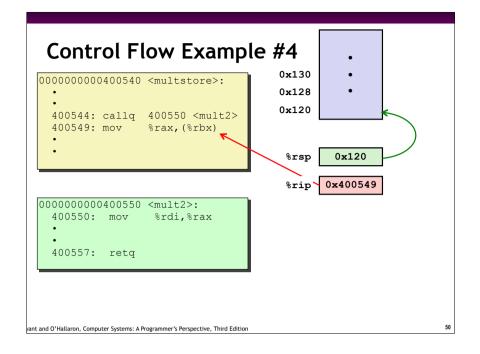


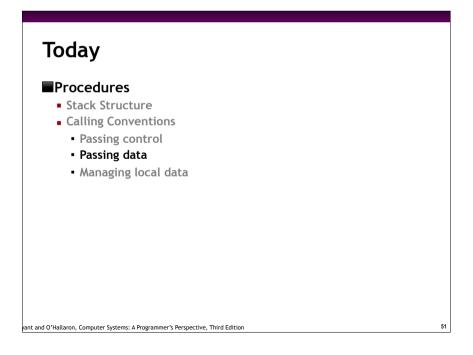
Procedure Control Flow Use stack to support procedure call and return Procedure call: call label Push return address on stack Jump to label Return address: Address of the next instruction right after call Procedure return: ret Pop address from stack Jump to address

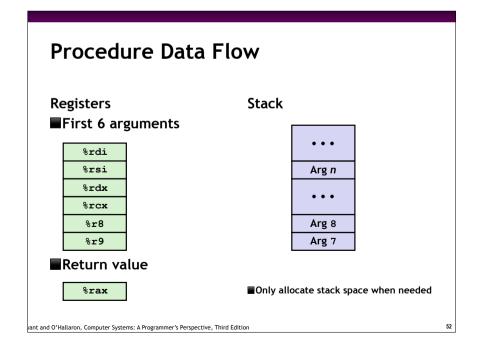
vant and O'Hallaron, Computer Systems: A Programmer's Perspective, Third Edition











```
void multstore
  Data Flow
                            (long x, long y, long *dest)
  Examples
                               long t = mult2(x, y);
                               *dest = t;
        0000000000400540 <multstore>:
          # x in %rdi, y in %rsi, dest in %rdx
          400541: mov
                          %rdx,%rbx
                                            # Save dest
          400544: callq 400550 <mult2> # mult2(x,y)
          # t in %rax
          400549: mov
                          %rax,(%rbx)
                                            # Save at dest
long mult2
                       0000000000400550 <mult2>:
  (long a, long b)
                         # a in %rdi, b in %rsi
                                          %rdi,%rax
                         400550: mov
  long s = a * b;
                                                           # a * b
                         400553: imul
                                          %rsi,%rax
                         # s in %rax
  return s;
                         400557: retq
                                                           # Return
vant and O'Hallaron, Computer Systems: A Programmer's Perspective, Third Edition
```

Stack-Based Languages

■Languages that support recursion

- e.g., C, 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*

vant and O'Hallaron, Computer Systems: A Programmer's Perspective, Third Edition

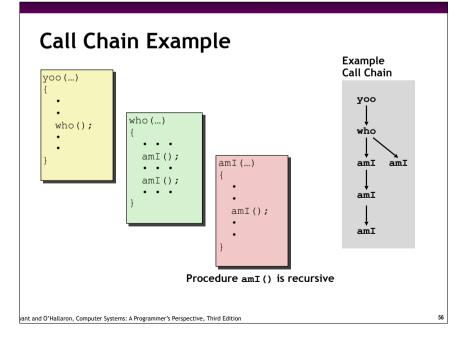
state for single procedure instantiation

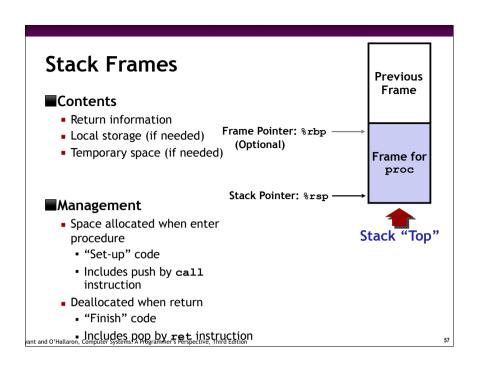
Today

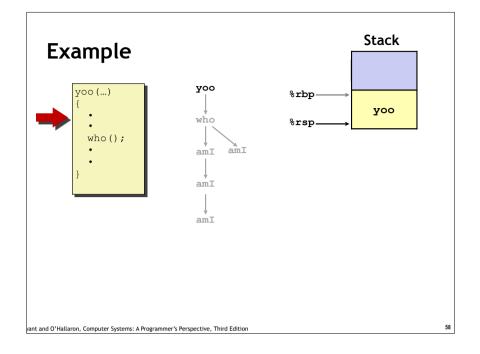
■Procedures

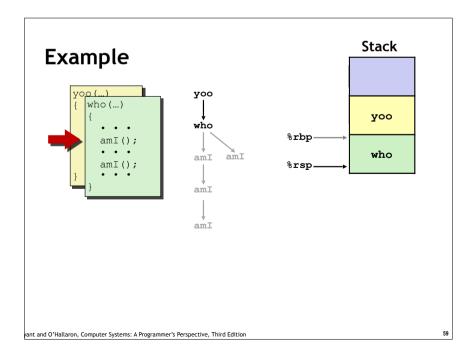
- Stack Structure
- Calling Conventions
 - Passing control
 - Passing data
 - Managing local data

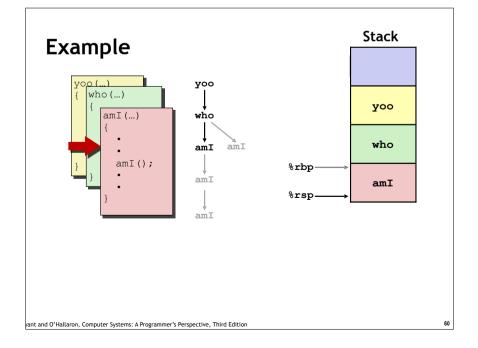
vant and O'Hallaron, Computer Systems: A Programmer's Perspective, Third Edition

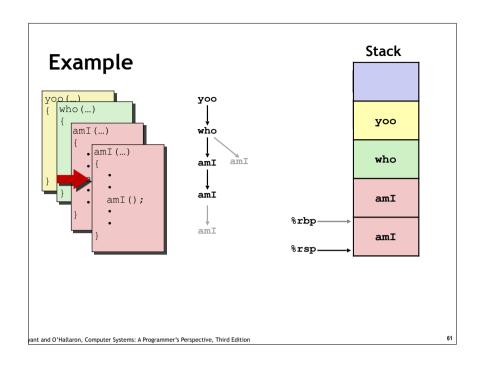


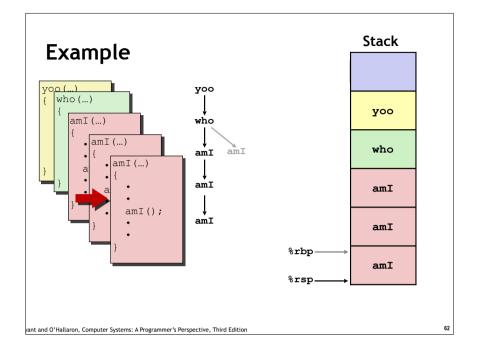


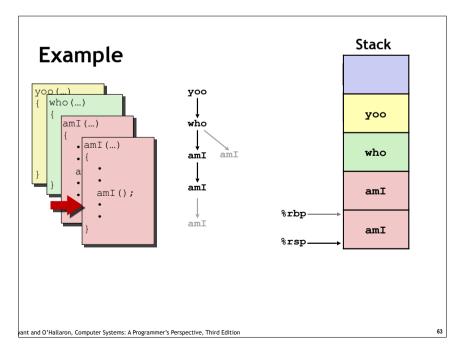


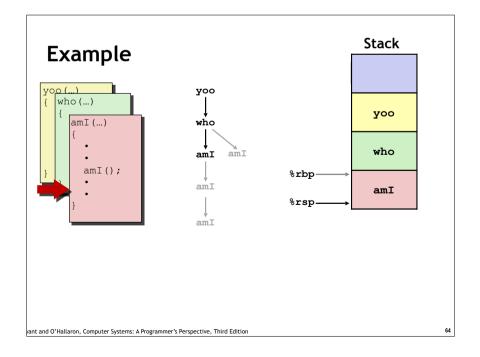


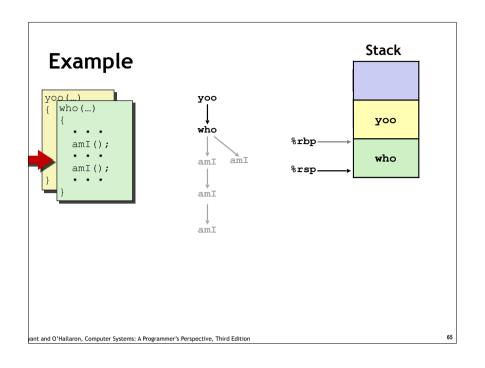


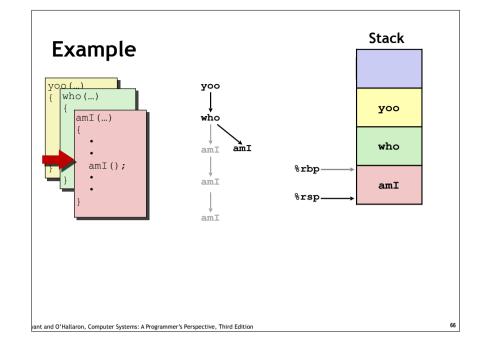


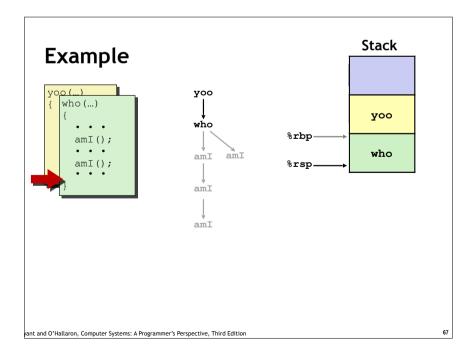


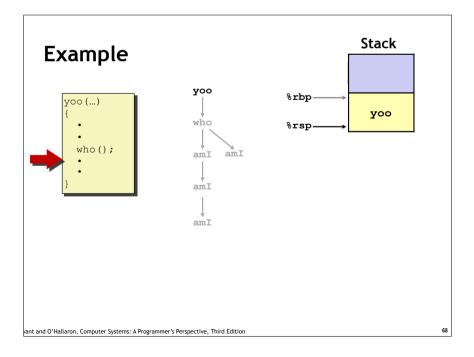


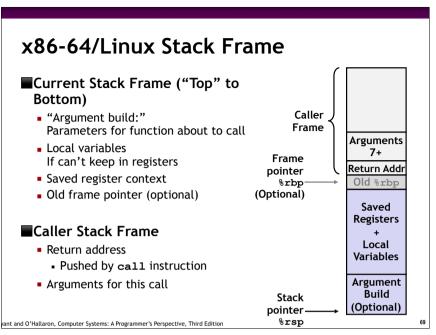


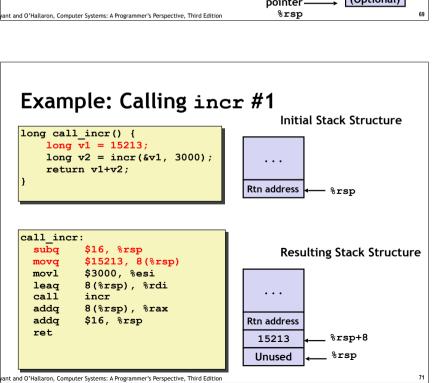


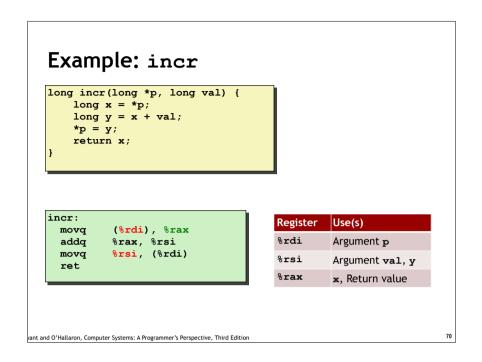


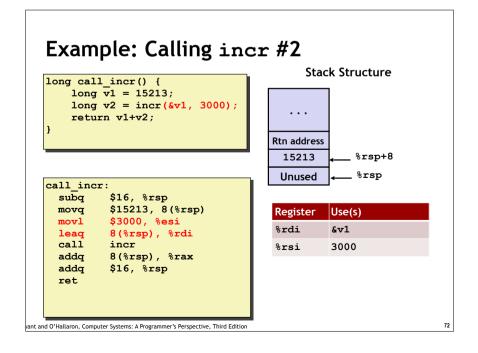


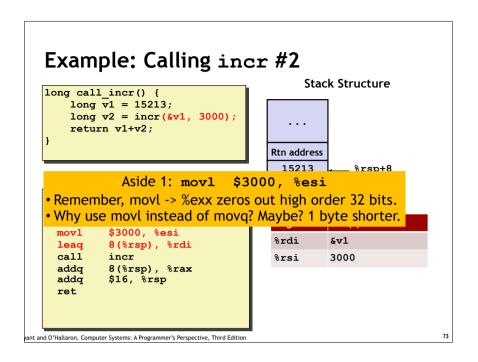


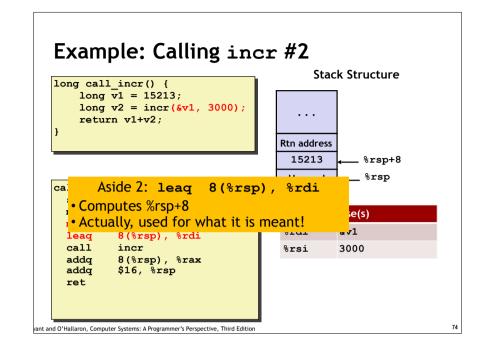


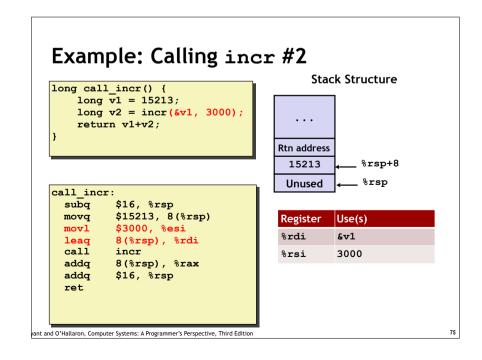


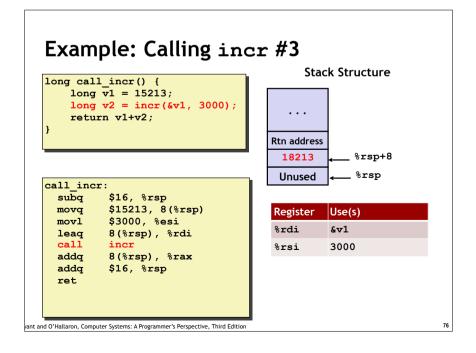


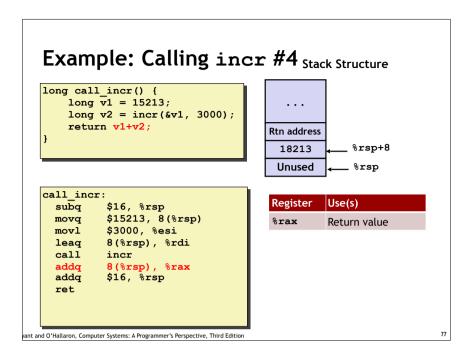


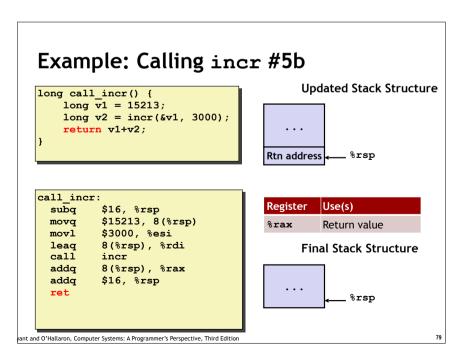


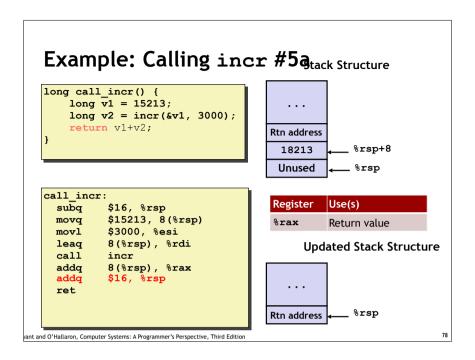












Today

- Procedures
 - Stack Structure
 - Calling Conventions
 - Passing control
 - Passing data
 - Managing local data

ant and O'Hallaron, Computer Systems: A Programmer's Perspective, Third Edition

