

Assembly Language

Alan L. Cox
alc@rice.edu

Some slides adapted from CMU 15.213 slides

Objectives

Be able to read simple x86-64 assembly language programs

Why Learn Assembly Language?

You'll probably never write a program in assembly

- ♦ **Compilers are much better and more patient than you are**

But, understanding assembly is key to understanding the machine-level execution model

- ♦ **Behavior of programs in presence of bugs**
 - High-level language model breaks down
- ♦ **Tuning program performance**
 - Understanding sources of program inefficiency
- ♦ **Implementing system software**
 - Compiler has machine code as target
 - Operating systems must manage process state

Assembly Language

One assembly instruction

- ♦ **Straightforward translation to a group of machine language bits that describe one instruction**

What do these instructions do?

- ♦ **Same kinds of things as high-level languages!**
 - **Arithmetic & logic**
 - Core computation
 - **Data transfer**
 - Copying data between memory locations and/or registers
 - **Control transfer**
 - Changing which instruction is next

Assembly Language (cont.)

But, assembly language has additional features:

- ♦ **Distinguishes instructions & data**
- ♦ **Labels = names for program control points**
- ♦ **Pseudo-instructions = special directives to the assembler**
- ♦ **Macros = user-definable abbreviations for code & constants**

Example C Program

main.c:

```
#include <stdio.h>

void
hello(char *name, int hour, int min)
{
    printf("Hello, %s, it's %d:%02d.",
           name, hour, min);
}

int
main(void)
{
    hello("Alan", 2, 55);
    return (0);
}
```

Run the command:

UNIX% clang -S main.c



Output a file named main.s
containing the assembly
code for main.c

C Compiler's Output

```
.file "main.c"
.section .rodata
.LC0:
.string "Hello, %s, it's %d:%02d."
.text
.globl hello
.type hello, @function
hello:
.LFB2:
    pushq    %rbp
.LCFI0:
    movq     %rsp, %rbp
.LCFI1:
    subq     $16, %rsp
.LCFI2:
    movq     %rdi, -8(%rbp)
    movl     %esi, -12(%rbp)
    movl     %edx, -16(%rbp)
    movl     -16(%rbp), %ecx
    movl     -12(%rbp), %edx
    movq     -8(%rbp), %rsi
    movl     $.LC0, %edi
    movl     $0, %eax
```

```
    call     printf
    leave
    ret
.LFE2:
.size hello, .-hello
.section .rodata
.LC1:
.string "Alan"
.text
.globl main
.type main, @function
main:
.LFB3:
    pushq    %rbp
.LCFI3:
    movq     %rsp, %rbp
.LCFI4:
    movl     $55, %edx
    movl     $2, %esi
    movl     $.LC1, %edi
    call     hello
    movl     $0, %eax
<..snip..>
```

A Breakdown of the Output

```
.file    "main.c"
.section .rodata
.LC0:
.string  "Hello, %s, it's %d:%02d."
.text
.globl  hello
.type   hello, @function
hello:
.LFB2:
        pushq   %rbp
.LCFI0:
        movq    %rsp, %rbp
.LCFI1:
        subq    $16, %rsp
.LCFI2:
        movq    %rdi, -8(%rbp)
        movl    %esi, -12(%rbp)
        movl    %edx, -16(%rbp)
        movl    -16(%rbp), %ecx
        movl    -12(%rbp), %edx
        movq    -8(%rbp), %rsi
        movl    $.LC0, %edi
        movl    $0, %eax
```

```
        call    printf
        leave
        ret
.LFE2:
        .size   hello, .-hello
<..snip..>
```

Instructions,
Pseudo-Instructions,
& Label Definitions

Instructions: Opcodes

```
.file "main.c"
.section .rodata
.LC0:
.string "Hello, %s, it's %d:%02d."
.text
.globl hello
.type hello, @function
hello:
.LFB2:
    pushq    %rbp
.LCFI0:
    movq     %rsp, %rbp
.LCFI1:
    subq     $16, %rsp
.LCFI2:
    movq     %rdi, -8(%rbp)
    movl     %esi, -12(%rbp)
    movl     %edx, -16(%rbp)
    movl     -16(%rbp), %ecx
    movl     -12(%rbp), %edx
    movq     -8(%rbp), %rsi
    movl     $.LC0, %edi
    movl     $0, %eax
```

```
    call     printf
    leave
    ret
.LFE2:
    .size    hello, .-hello
<..snip..>
```

Arithmetic,
data transfer, &
control transfer

Instructions: Operands

```
.file "main.c"
.section .rodata
.LC0:
.string "Hello, %s, it's %d:%02d."
.text
.globl hello
.type hello, @function
hello:
.LFB2:
pushq %rbp
.LCFI0:
movq %rsp, %rbp
.LCFI1:
subq $16, %rsp
.LCFI2:
movq %rdi, -8(%rbp)
movl %esi, -12(%rbp)
movl %edx, -16(%rbp)
movl -16(%rbp), %ecx
movl -12(%rbp), %edx
movq -8(%rbp), %rsi
movl $.LC0, %edi
movl $0, %eax
```

```
call printf
leave
ret
.LFE2:
.size hello, .-hello
<..snip..>
```

Registers,
constants, &
labels

Instruction Set Architecture

Contract between programmer and the hardware

- ♦ Defines visible state of the system
- ♦ Defines how state changes in response to instructions

Assembly Programmer (compiler)

- ♦ ISA is model of how a program will execute

Hardware Designer

- ♦ ISA is formal definition of the correct way to execute a program

Architecture vs. Implementation

Instruction Set Architecture

- ♦ Defines what a computer system does in response to a program and a set of data
- ♦ Programmer visible elements of computer system

Implementation

- ♦ Defines how a computer does it
- ♦ Sequence of steps to complete operations
- ♦ Time to execute each operation
- ♦ Hidden “bookkeeping” functions

Often Many Implementations of an ISA

| ISA | Implementations |
|---------|-----------------|
| x86-64 | Intel Core i7 |
| | AMD FX-83XX |
| | VIA Nano |
| ARMv7-A | Apple A6/A6X |
| | Qualcomm Krait |

Why separate architecture and implementation?

Compatibility

- ♦ **VAX architecture: mainframe \Rightarrow single chip**
- ♦ **ARM: 20x performance range**
 - high vs. low performance, power, price

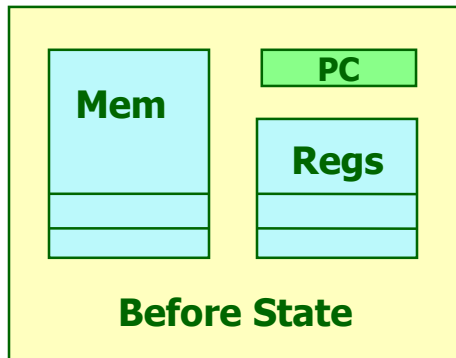
Longevity

- ♦ **20-30 years of ISA**
- ♦ **x86/x86-64 in 10th generation of implementations (architecture families)**
- ♦ **Retain software investment**
- ♦ **Amortize development costs over multiple markets**

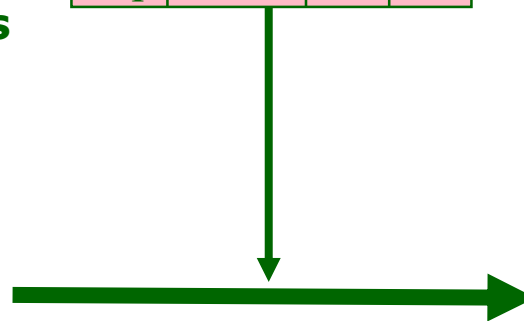
Instruction Set Basics

Instruction formats
Instruction types
Addressing modes

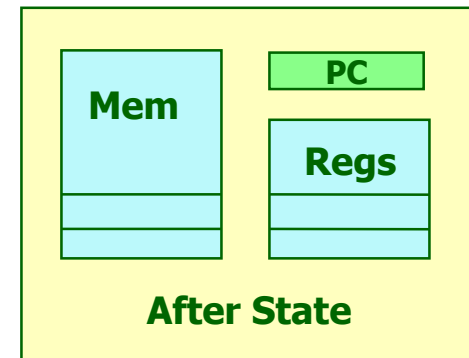
instruction



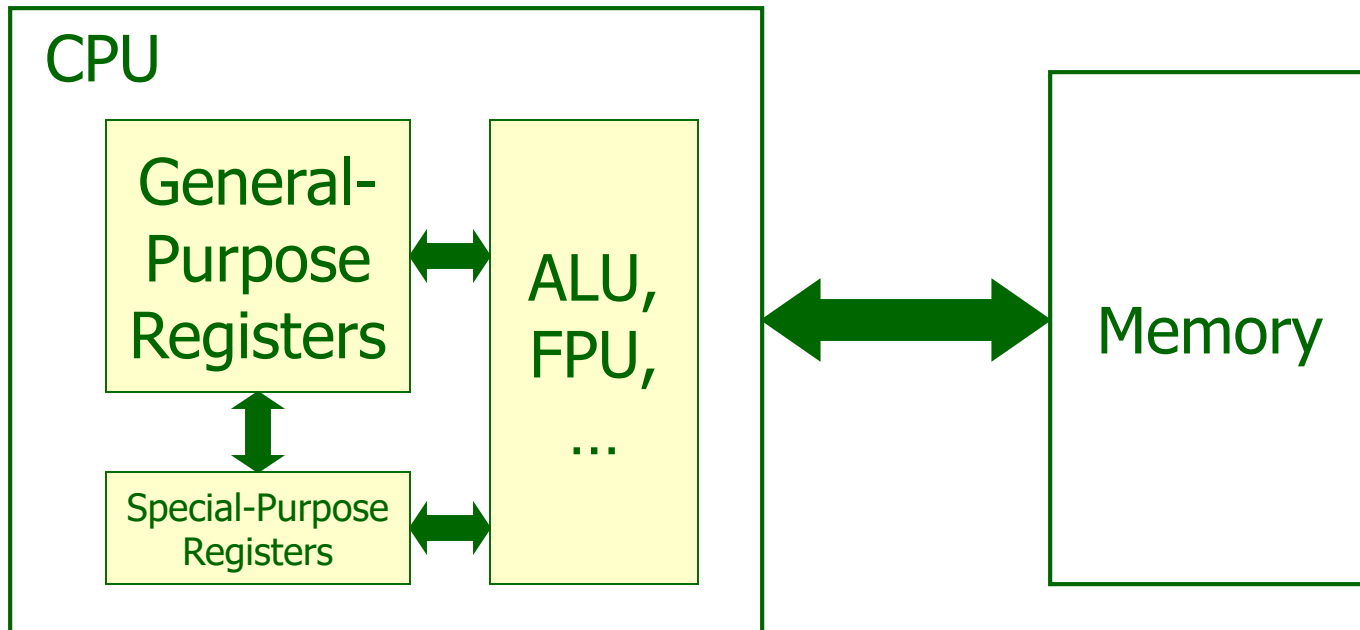
Machine state
Memory organization
Register organization



Data types
Operations



Typical Machine State



Integer Registers (IA32)

Origin
(mostly obsolete)

| | | | | | |
|-----------------|-------------|------------|------------|------------|--------------------------|
| general purpose | %eax | %ax | %ah | %al | <i>accumulate</i> |
| | %ecx | %cx | %ch | %cl | <i>counter</i> |
| | %edx | %dx | %dh | %dl | <i>data</i> |
| | %ebx | %bx | %bh | %bl | <i>base</i> |
| | %esi | %si | | | <i>source index</i> |
| | %edi | %di | | | <i>destination index</i> |
| | %esp | %sp | | | <i>stack pointer</i> |
| | %ebp | %bp | | | <i>base pointer</i> |

16-bit virtual registers
(backwards compatibility)

Moving Data: IA32

Moving Data

`movl Source, Dest:`

Operand Types

- ♦ **Immediate: Constant integer data**
 - **Example:** `$0x400`, `$-533`
 - Like C constant, but prefixed with ‘\$’
 - Encoded with 1, 2, or 4 bytes
- ♦ **Register: One of 8 integer registers**
 - **Example:** `%eax`, `%edx`
 - But `%esp` and `%ebp` reserved for special use
 - Others have special uses for particular instructions
- ♦ **Memory: 4 consecutive bytes of memory at address given by register**
 - **Simplest example:** `(%eax)`
 - Various other “address modes”

`%eax`

`%ecx`

`%edx`

`%ebx`

`%esi`

`%edi`

`%esp`

`%ebp`

movl Operand Combinations

| | Source | Dest | Src, Dest | C Analog |
|------|--------|------|---------------------|----------------|
| movl | Imm | Reg | movl \$0x4, %eax | temp = 0x4; |
| | | Mem | movl \$-147, (%eax) | *p = -147; |
| | Reg | Reg | movl %eax, %edx | temp2 = temp1; |
| | | Mem | movl %eax, (%edx) | *p = temp; |
| | Mem | Reg | movl (%eax), %edx | temp = *p; |

Cannot do memory-memory transfer with a single instruction

Simple Memory Addressing Modes

Normal (R) Mem[Reg[R]]

- ♦ Register R specifies memory address

```
movl (%ecx), %eax
```

Displacement D(R) Mem[Reg[R]+D]

- ♦ Register R specifies start of memory region
- ♦ Constant displacement D specifies offset

```
movl 8(%ebp), %edx
```

Using Simple Addressing Modes

```
void
swap(int *xp, int *yp)
{
    int t0 = *xp;
    int t1 = *yp;

    *xp = t1;
    *yp = t0;
}
```

swap:

```
pushl %ebp
movl  %esp, %ebp
pushl %ebx
```

} Set
Up

```
movl  8(%ebp), %edx
movl  12(%ebp), %ecx
movl  (%edx), %ebx
movl  (%ecx), %eax
movl  %eax, (%edx)
movl  %ebx, (%ecx)
```

} Body

```
popl  %ebx
popl  %ebp
ret
```

} Finish

Using Simple Addressing Modes

```
void
swap(int *xp, int *yp)
{
    int t0 = *xp;
    int t1 = *yp;

    *xp = t1;
    *yp = t0;
}
```

swap:

```
pushl %ebp
movl  %esp, %ebp
pushl %ebx
```

} Set Up

```
movl  8(%ebp), %edx
movl  12(%ebp), %ecx
movl  (%edx), %ebx
movl  (%ecx), %eax
movl  %eax, (%edx)
movl  %ebx, (%ecx)
```

} Body

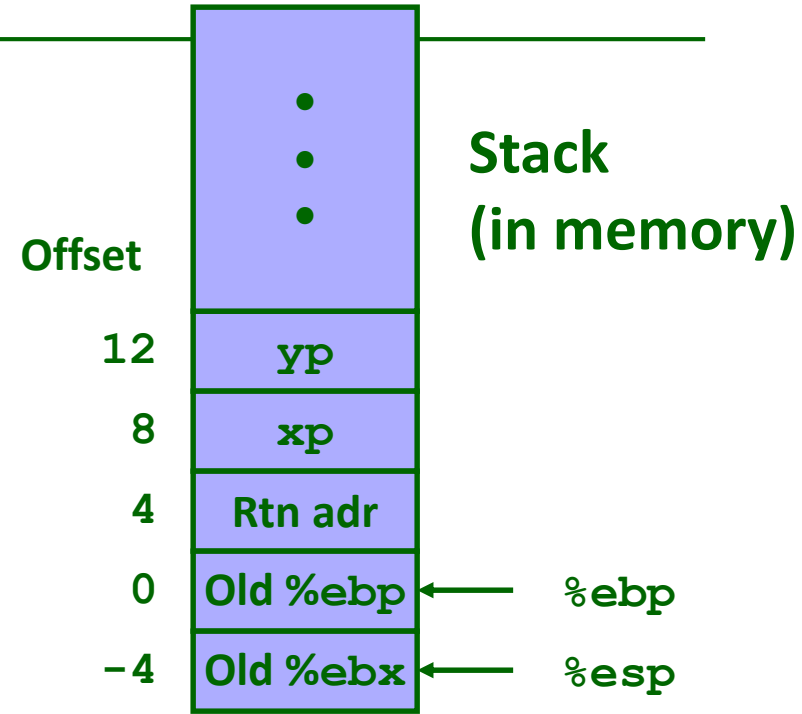
```
popl  %ebx
popl  %ebp
ret
```

} Finish

Understanding Swap

```
void
swap(int *xp, int *yp)
{
    int t0 = *xp;
    int t1 = *yp;

    *xp = t1;
    *yp = t0;
}
```



| Register | Value |
|----------|-------|
| %edx | xp |
| %ecx | yp |
| %ebx | t0 |
| %eax | t1 |

```
movl 8(%ebp), %edx    # edx = xp
movl 12(%ebp), %ecx   # ecx = yp
movl (%edx), %ebx     # ebx = *xp (t0)
movl (%ecx), %eax     # eax = *yp (t1)
movl %eax, (%edx)     # *xp = t1
movl %ebx, (%ecx)     # *yp = t0
```

Understanding Swap

| | |
|------|-------|
| %eax | |
| %edx | |
| %ecx | |
| %ebx | |
| %esi | |
| %edi | |
| %esp | |
| %ebp | 0x104 |

| | | Address |
|------|----|---------|
| | | 0x124 |
| | | 0x120 |
| | | 0x11c |
| | | 0x118 |
| | | 0x114 |
| yp | 12 | 0x120 |
| xp | 8 | 0x124 |
| | 4 | Rtn adr |
| %ebp | 0 | 0x108 |
| | -4 | 0x104 |
| | | 0x100 |

```

movl 8(%ebp), %edx    # edx = xp
movl 12(%ebp), %ecx   # ecx = yp
movl (%edx), %ebx     # ebx = *xp (t0)
movl (%ecx), %eax     # eax = *yp (t1)
movl %eax, (%edx)     # *xp = t1
movl %ebx, (%ecx)     # *yp = t0
    
```


Understanding Swap

| | |
|------|-------|
| %eax | |
| %edx | 0x124 |
| %ecx | |
| %ebx | |
| %esi | |
| %edi | |
| %esp | |
| %ebp | 0x104 |

| | | Address |
|------|----|---------|
| | | 0x124 |
| | | 0x120 |
| | | 0x11c |
| | | 0x118 |
| | | 0x114 |
| yp | 12 | 0x120 |
| xp | 8 | 0x124 |
| | 4 | Rtn adr |
| %ebp | 0 | 0x108 |
| | -4 | 0x104 |
| | | 0x100 |

```

movl 8(%ebp), %edx    # edx = xp
movl 12(%ebp), %ecx   # ecx = yp
movl (%edx), %ebx     # ebx = *xp (t0)
movl (%ecx), %eax     # eax = *yp (t1)
movl %eax, (%edx)     # *xp = t1
movl %ebx, (%ecx)     # *yp = t0
    
```

Understanding Swap

| | |
|------|-------|
| %eax | |
| %edx | 0x124 |
| %ecx | 0x120 |
| %ebx | |
| %esi | |
| %edi | |
| %esp | |
| %ebp | 0x104 |

| | | Address |
|------|----|---------|
| | | 0x124 |
| | | 0x120 |
| | | 0x11c |
| | | 0x118 |
| | | 0x114 |
| yp | 12 | 0x120 |
| xp | 8 | 0x124 |
| | 4 | Rtn adr |
| %ebp | 0 | 0x108 |
| | -4 | 0x104 |
| | | 0x100 |

```

movl 8(%ebp), %edx    # edx = xp
movl 12(%ebp), %ecx   # ecx = yp
movl (%edx), %ebx     # ebx = *xp (t0)
movl (%ecx), %eax     # eax = *yp (t1)
movl %eax, (%edx)     # *xp = t1
movl %ebx, (%ecx)     # *yp = t0
    
```

Understanding Swap

| | |
|------|-------|
| %eax | |
| %edx | 0x124 |
| %ecx | 0x120 |
| %ebx | 123 |
| %esi | |
| %edi | |
| %esp | |
| %ebp | 0x104 |

| | | Address |
|------|-----|---------|
| | 123 | 0x124 |
| | 456 | 0x120 |
| | | 0x11c |
| | | 0x118 |
| | | 0x114 |
| yp | 12 | 0x120 |
| xp | 8 | 0x124 |
| | 4 | Rtn adr |
| %ebp | 0 | |
| | -4 | |
| | | 0x104 |
| | | 0x100 |

```

movl 8(%ebp), %edx    # edx = xp
movl 12(%ebp), %ecx   # ecx = yp
movl (%edx), %ebx     # ebx = *xp (t0)
movl (%ecx), %eax     # eax = *yp (t1)
movl %eax, (%edx)     # *xp = t1
movl %ebx, (%ecx)     # *yp = t0
    
```

Understanding Swap

| | |
|------|-------|
| %eax | 456 |
| %edx | 0x124 |
| %ecx | 0x120 |
| %ebx | 123 |
| %esi | |
| %edi | |
| %esp | |
| %ebp | 0x104 |

| | | Address |
|------|----|---------|
| | | 0x124 |
| | | 0x120 |
| | | 0x11c |
| | | 0x118 |
| | | 0x114 |
| yp | 12 | 0x120 |
| xp | 8 | 0x124 |
| | 4 | Rtn adr |
| %ebp | 0 | 0x108 |
| | -4 | 0x104 |
| | | 0x100 |

```

movl 8(%ebp), %edx    # edx = xp
movl 12(%ebp), %ecx   # ecx = yp
movl (%edx), %ebx     # ebx = *xp (t0)
movl (%ecx), %eax    # eax = *yp (t1)
movl %eax, (%edx)     # *xp = t1
movl %ebx, (%ecx)     # *yp = t0
    
```

Understanding Swap

| | |
|------|-------|
| %eax | 456 |
| %edx | 0x124 |
| %ecx | 0x120 |
| %ebx | 123 |
| %esi | |
| %edi | |
| %esp | |
| %ebp | 0x104 |

| | | Address |
|------|-----|---------|
| | 456 | 0x124 |
| | 456 | 0x120 |
| | | 0x11c |
| | | 0x118 |
| | | 0x114 |
| yp | 12 | 0x120 |
| xp | 8 | 0x124 |
| | 4 | Rtn adr |
| %ebp | 0 | 0x108 |
| | | 0x104 |
| | -4 | 0x100 |

```

movl 8(%ebp), %edx    # edx = xp
movl 12(%ebp), %ecx   # ecx = yp
movl (%edx), %ebx     # ebx = *xp (t0)
movl (%ecx), %eax     # eax = *yp (t1)
movl %eax, (%edx)    # *xp = t1
movl %ebx, (%ecx)     # *yp = t0
    
```

Understanding Swap

| | |
|------|-------|
| %eax | 456 |
| %edx | 0x124 |
| %ecx | 0x120 |
| %ebx | 123 |
| %esi | |
| %edi | |
| %esp | |
| %ebp | 0x104 |

| | | Address |
|------|----|---------|
| | | 0x124 |
| | | 0x120 |
| | | 0x11c |
| | | 0x118 |
| | | 0x114 |
| yp | 12 | 0x120 |
| xp | 8 | 0x124 |
| | 4 | Rtn adr |
| %ebp | 0 | 0x108 |
| | -4 | 0x104 |
| | | 0x100 |

```

movl 8(%ebp), %edx    # edx = xp
movl 12(%ebp), %ecx   # ecx = yp
movl (%edx), %ebx     # ebx = *xp (t0)
movl (%ecx), %eax     # eax = *yp (t1)
movl %eax, (%edx)     # *xp = t1
movl %ebx, (%ecx)   # *yp = t0
  
```

Complete Memory Addressing Modes

Most General Form

$D(Rb, Ri, S)$ $Mem[Reg[Rb] + S * Reg[Ri] + D]$

- ♦ **D:** Constant “displacement” 1, 2, or 4 bytes
- ♦ **Rb:** Base register: Any of 8 integer registers
- ♦ **Ri:** Index register: Any, except for `%esp`
 - Unlikely you’d use `%ebp`, either
- ♦ **S:** Scale: 1, 2, 4, or 8 (*why these numbers?*)

Special Cases

| | |
|---------------|------------------------------|
| (Rb, Ri) | $Mem[Reg[Rb] + Reg[Ri]]$ |
| $D(Rb, Ri)$ | $Mem[Reg[Rb] + Reg[Ri] + D]$ |
| (Rb, Ri, S) | $Mem[Reg[Rb] + S * Reg[Ri]]$ |

x86-64 Integer Registers

| | |
|-------------------|-------------------|
| <code>%rax</code> | <code>%eax</code> |
| <code>%rbx</code> | <code>%ebx</code> |
| <code>%rcx</code> | <code>%ecx</code> |
| <code>%rdx</code> | <code>%edx</code> |
| <code>%rsi</code> | <code>%esi</code> |
| <code>%rdi</code> | <code>%edi</code> |
| <code>%rsp</code> | <code>%esp</code> |
| <code>%rbp</code> | <code>%ebp</code> |

| | |
|-------------------|--------------------|
| <code>%r8</code> | <code>%r8d</code> |
| <code>%r9</code> | <code>%r9d</code> |
| <code>%r10</code> | <code>%r10d</code> |
| <code>%r11</code> | <code>%r11d</code> |
| <code>%r12</code> | <code>%r12d</code> |
| <code>%r13</code> | <code>%r13d</code> |
| <code>%r14</code> | <code>%r14d</code> |
| <code>%r15</code> | <code>%r15d</code> |

- ◆ **Extend existing registers. Add 8 new ones.**
- ◆ **Make `%ebp/%rbp` general purpose**

Instructions

Long word l (4 Bytes) \leftrightarrow Quad word q (8 Bytes)

New instructions:

- ♦ `movl` \rightarrow `movq`
- ♦ `addl` \rightarrow `addq`
- ♦ `sall` \rightarrow `salq`
- ♦ **etc.**

32-bit instructions that generate 32-bit results

- ♦ **Set higher order bits of destination register to 0**
- ♦ **Example:** `addl`

32-bit code for swap

```
void
swap(int *xp, int *yp)
{
    int t0 = *xp;
    int t1 = *yp;

    *xp = t1;
    *yp = t0;
}
```

swap:

```
pushl %ebp
movl  %esp,%ebp
pushl %ebx
```

} Set
Up

```
movl  8(%ebp), %edx
movl  12(%ebp), %ecx
movl  (%edx), %ebx
movl  (%ecx), %eax
movl  %eax, (%edx)
movl  %ebx, (%ecx)
```

} Body

```
popl  %ebx
popl  %ebp
ret
```

} Finish

64-bit code for swap

```
void
swap(int *xp, int *yp)
{
    int t0 = *xp;
    int t1 = *yp;

    *xp = t1;
    *yp = t0;
}
```

swap:

```
movl    (%rdi), %edx
movl    (%rsi), %eax
movl    %eax, (%rdi)
movl    %edx, (%rsi)
```

ret

} Set Up

} Body

} Finish

Operands passed in registers (why useful?)

- ♦ First (xp) in %rdi, second (yp) in %rsi
- ♦ 64-bit pointers

No stack operations required

32-bit data

- ♦ Data held in registers %eax and %edx

Cox ♦ **movl operation**

Assembly

64-bit code for long int swap

```
void
swap_1(long *xp, long *yp)
{
    long t0 = *xp;
    long t1 = *yp;

    *xp = t1;
    *yp = t0;
}
```

swap_1:

```
movq    (%rdi), %rdx
movq    (%rsi), %rax
movq    %rax, (%rdi)
movq    %rdx, (%rsi)
```

ret

} Set
Up

} Body

} Finish

64-bit data

- ♦ Data held in registers `%rax` and `%rdx`
- ♦ `movq` operation
 - “q” stands for quad-word

Application Binary Interface (ABI)

Standardizes the use of memory and registers by C compilers

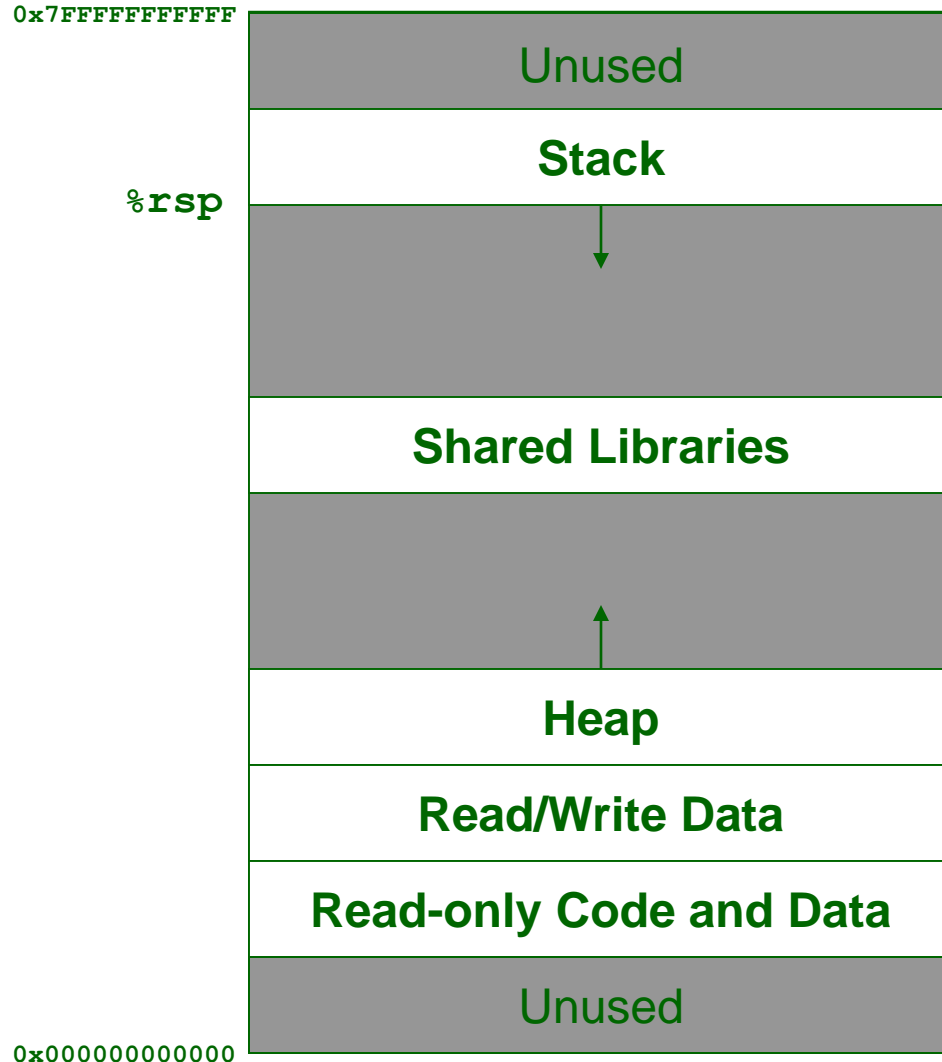
- ♦ **Enables interoperability of code compiled by different C compilers**
 - E.g., a program compiled with Intel's optimizing C compiler can call a library function that was compiled by the GNU C compiler
- ♦ **Sets the size of built-in data types**
 - E.g., int, long, etc.
- ♦ **Dictates the implementation of function calls**
 - E.g., how parameters and return values are passed

Register Usage

The x86-64 ABI specifies that registers are used as follows

- ♦ **Temporary (callee can change these)**
`%rax, %r10, %r11`
- ♦ **Parameters to function calls**
`%rdi, %rsi, %rdx, %rcx, %r8, %r9`
- ♦ **Callee saves (callee can only change these after saving their current value)**
`%rbx, %rbp, %r12-%r15`
 - `%rbp` is typically used as the “frame” pointer to the current function’s local variables
- ♦ **Return values**
`%rax, %rdx`

Procedure Calls and the Stack



Where are local variables stored?

- ♦ Registers (only 16)
- ♦ *Stack*

Stack provides as much local storage as necessary

- ♦ Until memory exhausted
- ♦ Each procedure allocates its own space on the stack

Referencing the stack

- ♦ `%rsp` points to the bottom of the stack in x86-64

Control Flow: Function Calls

What must assembly/machine language do?

| Caller | Callee |
|---|---|
| <ol style="list-style-type: none">1. Save function arguments2. Branch to function body | <ol style="list-style-type: none">3. Execute body<ul style="list-style-type: none">• May allocate memory• May call functions4. Save function result5. Branch to where called |

1. Use registers to pass arguments, save return location on stack)

Program Stack

Figure 3.3: Stack Frame with Base Pointer

| Position | Contents | Frame |
|-------------------------------|----------------------------------|----------|
| $8n+16$ (<code>%rbp</code>) | memory argument eightbyte n | Previous |
| | ... | |
| 16 (<code>%rbp</code>) | memory argument eightbyte 0 | Current |
| 8 (<code>%rbp</code>) | return address | |
| 0 (<code>%rbp</code>) | previous <code>%rbp</code> value | |
| -8 (<code>%rbp</code>) | unspecified | |
| | ... | |
| 0 (<code>%rsp</code>) | variable size | |
| -128 (<code>%rsp</code>) | red zone | |

Figure 3.3 is reproduced from the AMD64 ABI Draft 0.99.5 by Matz et al.

What are Pseudo-Instructions?

Assembler directives, with various purposes

Data & instruction encoding:

- ♦ **Separate instructions & data into sections**
- ♦ **Reserve memory with initial data values**
- ♦ **Reserve memory w/o initial data values**
- ♦ **Align instructions & data**

Provide information useful to linker or debugger

- ♦ **Correlate source code with assembly/machine**
- ♦ **...**

Instructions & Pseudo-Instructions

```
.file "main.c"
.section .rodata
.LC0:
.string "Hello, %s, it's %d:%02d."
.text
.globl hello
.type hello, @function
hello:
.LFB2:
    pushq    %rbp
.LCFI0:
    movq     %rsp, %rbp
.LCFI1:
    subq     $16, %rsp
.LCFI2:
    movq     %rdi, -8(%rbp)
    movl     %esi, -12(%rbp)
    movl     %edx, -16(%rbp)
    movl     -16(%rbp), %ecx
    movl     -12(%rbp), %edx
    movq     -8(%rbp), %rsi
    movl     $.LC0, %edi
    movl     $0, %eax
```

```
    call     printf
    leave
    ret
.LFE2:
    .size    hello, .-hello
<..snip..>
```

Instructions,
Pseudo-Instructions,
& Label Definitions

Separate instructions & data

Instructions & Pseudo-Instructions

```
.file "main.c"
.section .rodata
.LC0:
.string 'Hello, %s, it's %d:%02d.'
.text
.globl hello
.type hello, @function
hello:
.LFB2:
    pushq    %rbp
.LCFI0:
    movq     %rsp, %rbp
.LCFI1:
    subq     $16, %rsp
.LCFI2:
    movq     %rdi, -8(%rbp)
    movl     %esi, -12(%rbp)
    movl     %edx, -16(%rbp)
    movl     -16(%rbp), %ecx
    movl     -12(%rbp), %edx
    movq     -8(%rbp), %rsi
    movl     $.LC0, %edi
    movl     $0, %eax
```

```
    call     printf
    leave
    ret
.LFE2:
    .size    hello, .-hello
<..snip..>
```

Instructions,
Pseudo-Instructions,
& Label Definitions

Reserve memory with
initial data values

Instructions & Pseudo-Instructions

```
.file "main.c"
.section .rodata
.LC0:
.string "Hello, %s, it's %d:%02d."
.text
.globl hello
.type hello, @function
hello:
.LFB2:
    pushq    %rbp
.LCFI0:
    movq     %rsp, %rbp
.LCFI1:
    subq     $16, %rsp
.LCFI2:
    movq     %rdi, -8(%rbp)
    movl     %esi, -12(%rbp)
    movl     %edx, -16(%rbp)
    movl     -16(%rbp), %ecx
    movl     -12(%rbp), %edx
    movq     -8(%rbp), %rsi
    movl     $.LC0, %edi
    movl     $0, %eax
```

```
    call     printf
    leave
    ret
.LFE2:
.size hello, .-hello
<..snip..>
```

Instructions,
Pseudo-Instructions,
& Label Definitions

Correlate source code with
assembly/machine

Label Types

```
.file "main.c"
.section .rodata
.LC0:
.string "Hello, %s, it's %d:%02d."
.text
.globl hello
.type hello, @function
hello:
.LFB2:
pushq %rbp
.LCFI0:
movq %rsp, %rbp
.LCFI1:
subq $16, %rsp
.LCFI2:
movq %rdi, -8(%rbp)
movl %esi, -12(%rbp)
movl %edx, -16(%rbp)
movl -16(%rbp), %ecx
movl -12(%rbp), %edx
movq -8(%rbp), %rsi
movl $.LC0, %edi
movl $0, %eax
```

```
call printf
leave
ret
.LFE2:
.size hello, .-hello
<..snip..>
```

Definitions,
internal references,
& external references

The label's value is the
address of the subsequent
instruction or pseudo-
instruction

Assembly/Machine Language – Semantics

Basic model of execution

- ♦ **Fetch instruction, from memory @ PC**
- ♦ **Increment PC**
- ♦ **Decode instruction**
- ♦ **Fetch operands, from registers or memory**
- ♦ **Execute operation**
- ♦ **Store result(s), in registers or memory**

Simulate Program Execution

```
.file    "main.c"
.section .rodata
.LC0:
.string "Hello, %s, it's %d:%02d."
.text
.globl hello
.type    hello, @function
hello:
.LFB2:
    pushq   %rbp
.LCFI0:
    movq    %rsp, %rbp
.LCFI1:
    subq    $16, %rsp
.LCFI2:
    movq    %rdi, -8(%rbp)
    movl    %esi, -12(%rbp)
    movl    %edx, -16(%rbp)
    movl    -16(%rbp), %ecx
    movl    -12(%rbp), %edx
    movq    -8(%rbp), %rsi
    movl    $.LC0, %edi
    movl    $0, %eax
```

```
    call    printf
    leave
    ret
.LFE2:
.size     hello, .-hello
.section .rodata
.LC1:
.string "Alan"
.text
.globl main
.type     main, @function
main:
.LFB3:
    pushq   %rbp
.LCFI3:
    movq    %rsp, %rbp
.LCFI4:
    movl    $55, %edx
    movl    $2, %esi
    movl    $.LC1, %edi
    call    hello
    movl    $0, %eax
<..next slide..>
```


Simulate Program ... (cont.)

```
        movl    $0, %eax
        leave
        ret
.LFE3:
        .size   main, .-main
<..snip..>
```

Exercise

```
0x7000000000000088
0x7000000000000084
0x7000000000000080
0x700000000000007c
0x7000000000000078
0x7000000000000074
0x7000000000000070
0x700000000000006c
0x7000000000000068
0x7000000000000064
0x7000000000000060
0x700000000000005c
0x7000000000000058
0x7000000000000054
0x7000000000000050
0x700000000000004c
0x7000000000000048
0x7000000000000044
0x7000000000000040
0x700000000000003c
0x7000000000000038
0x7000000000000034
0x7000000000000030
0x700000000000002c
0x7000000000000028
0x7000000000000024
0x7000000000000020
0x700000000000001c
0x7000000000000018
0x7000000000000014
0x7000000000000010
0x700000000000000c
0x7000000000000008
0x7000000000000004
0x7000000000000000
```

initial values:

```
%rbp 0x70000000000088
```

```
%rsp    0x7000000000006c
```

```
.LC0 0x408280
```

```
.LC1 0x408400
```

```
&"movl $0, %eax" in main() == 0x400220
```

| | |
|------|--|
| %rbp | |
|------|--|

| | |
|------|--|
| %rsp | |
|------|--|

```
%rdi
```

```
%rsi
```

| | |
|------|--|
| %rax | |
|------|--|

$\frac{9}{10}rcx$

[illegible]

More x86-64 Assembly

Chapter 3 of the textbook explains x86 and x86-64 assembly in greater detail

- ♦ **More examples translated from C**
- ♦ **More detail than you're expected to know for this course**

Some code sequences generated by the compiler can still be confusing

- ♦ **Usually not important for this class (web is helpful if you are still curious)**

Next Time

Program Linking