x86-64 Programming III
CSE 351 Winter 2024

Instructor:
Justin Hsia

Teaching Assistants:
Adithi Raghavan
Aman Mohammed
Connie Chen
Eyoel Gebre
Jiawei Huang
Malak Zaki
Naama Amiel
Nathan Khuat
Nikolas McNamee
Pedro Amarante
Will Robertson

http://xkcd.com/571/
Relevant Course Information

- Lab 1a regrade requests open on Gradescope
- Lab 1b submissions close tonight
- Lab 2 due next Friday (2/2)
  - Section tomorrow on to help prep you for Lab 2 – use the midterm reference sheet & bring your laptop!
  - Optional GDB Tutorial in Ed Lessons

- Midterm (take home, 2/8–2/10)
  - Make notes and use the \textit{midterm reference sheet}
  - Form study groups and look at past exams!
Extra Credit

❖ All labs starting with Lab 2 have extra credit portions
  ▪ These are meant to be fun extensions to the labs

❖ Extra credit points *don't* affect your lab grades
  ▪ From the course policies: “they will be accumulated over the course and will be used to bump up borderline grades at the end of the quarter.”
  ▪ Make sure you finish the rest of the lab before attempting any extra credit
x86-64 Programming III
Lesson Summary (1/3)

❖ **Labels** *(e.g., main, .L0)* refer to an instruction address and used as jump targets in assembly

❖ Control flow in x86 determined by Condition Codes

- **Set instructions** *(set*)
  read out flag values (0/1)

- **Jump instructions** *(j*)
  use flag values to determine next instruction to execute

- Result of 1st instruction gets compared against 0 in a way determined by 2nd instruction:

<table>
<thead>
<tr>
<th>Instruction</th>
<th>Condition</th>
<th>Test Code</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>je sete</strong></td>
<td>“Equal”</td>
<td>d (op) s == 0 b-a == 0 b&amp;a == 0</td>
</tr>
<tr>
<td><strong>jne setne</strong></td>
<td>“Not equal”</td>
<td>d (op) s != 0 b-a != 0 b&amp;a != 0</td>
</tr>
<tr>
<td><strong>js sets</strong></td>
<td>“Signed” (negative)</td>
<td>d (op) s &lt; 0 b-a &lt; 0 b&amp;a &lt; 0</td>
</tr>
<tr>
<td><strong>jns setns</strong></td>
<td>“Not signed” (nonnegative)</td>
<td>d (op) s &gt;= 0 b-a &gt;= 0 b&amp;a &gt;= 0</td>
</tr>
<tr>
<td><strong>jg setg</strong></td>
<td>“Greater”</td>
<td>d (op) s &gt; 0 b-a &gt; 0 b&amp;a &gt; 0</td>
</tr>
<tr>
<td><strong>jge setge</strong></td>
<td>“Greater or equal”</td>
<td>d (op) s &gt;= 0 b-a &gt;= 0 b&amp;a &gt;= 0</td>
</tr>
<tr>
<td><strong>jl setl</strong></td>
<td>“Less”</td>
<td>d (op) s &lt; 0 b-a &lt; 0 b&amp;a &lt; 0</td>
</tr>
<tr>
<td><strong>jle setle</strong></td>
<td>“Less or equal”</td>
<td>d (op) s &lt;= 0 b-a &lt;= 0 b&amp;a &lt;= 0</td>
</tr>
<tr>
<td><strong>ja seta</strong></td>
<td>“Above” (unsigned &gt;)</td>
<td>d (op) s &gt; 0U b &gt;U a b&amp;a &gt; 0U</td>
</tr>
<tr>
<td><strong>jb setb</strong></td>
<td>“Below” (unsigned &lt;)</td>
<td>d (op) s &lt; 0U b &lt;U a b&amp;a &lt; 0U</td>
</tr>
</tbody>
</table>
Lesson Summary (2/3)

❖ Most control flow constructs (e.g., if-else, for-loop, while-loop) can be implemented in assembly using combinations of conditional and unconditional jumps

- Differences come from placement of jumps and whether they jump forward or backwards in code

```assembly
if (<test>)
    <then>
else
    <else>

!test
  j* .L2
  <then>
  jmp done
L2:
  <else>
  jmp done
done:
```

```assembly
do {
    <body>
} while (<test>)
```

```assembly
while (<test>) {
    <body>
}
```

```assembly
<init>
while (<test>) {
    <loop body>
    <update>
}
```

```assembly
<test>
  j* .L2
  <then>
  jmp done
L2:
  <else>
  jmp done
done:
```

```assembly
top: <body>
  <test>
  j* top
done:
```

```assembly
top: !test
  j* done
  <body>
  jmp top
done:
```

```assembly
top: !test
  j* done
  <body>
  <test>
  j* top
done:
```
Lesson Summary (3/3)

❖ Switch statements can be implemented using *jump tables* and *indirect jump instructions*
  ▪ Jump tables are arrays of pointers to code blocks
  ▪ Indirect jump jumps to address stored somewhere in memory instead of target specified in instruction

```
switch (x) {
  case 1: <code> break;
  case 2: <code>
  case 3: <code> break;
  case 5:
  case 6: <code> break;
  case 7: <code> break;
  default: <code>
}
```

```
jump table

.section .rodata
.align 8
.L4:
.quad .L9 # x = 0
.quad .L8 # x = 1
.quad .L7 # x = 2
.quad .L10 # x = 3
.quad .L9 # x = 4
.quad .L5 # x = 5
.quad .L5 # x = 6
.quad .L3 # x = 7
```

```
cmpq $7, %rdi
ja .L9 # default
jmp *.L4(%rdi,8)
```
Lesson Q&A

❖ Learning Objectives:
   ▪ Without executing, describe the overall purpose of snippets of x86-64 assembly code containing arithmetic, if-else statements, and/or loops.

❖ What lingering questions do you have from the lesson?
   ▪ Chat with your neighbors about the lesson for a few minutes to come up with questions
x86-64 Programming III – Practice
Polling Question (1/2)

<table>
<thead>
<tr>
<th>Register</th>
<th>Use(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>%rdi</td>
<td>1st argument (x)</td>
</tr>
<tr>
<td>%rsi</td>
<td>2nd argument (y)</td>
</tr>
<tr>
<td>%rax</td>
<td>return value</td>
</tr>
</tbody>
</table>

A. `cmpq %rsi, %rdi`  
`jle .L4`  

B. `cmpq %rsi, %rdi`  
`jg .L4`  

C. `testq %rsi, %rdi`  
`jle .L4`  

D. `testq %rsi, %rdi`  
`jg .L4`  

```c
long absdiff(long x, long y) {
    long result;
    if (x > y)
        result = x - y;
    else
        result = y - x;
    return result;
}
```

absdiff:

```
movq %rdi, %rax
subq %rsi, %rax
ret

.L4:
```

# x > y:
```c
# x <= y:
```
```c
movq %rsi, %rax
subq %rdi, %rax
ret
```

less than or equal to (le)
Polling Question (2/2)

❖ The following is assembly code for a for-loop; identify the corresponding parts (Init, Test, Update)

- \( i \rightarrow \%eax, \ x \rightarrow \%rdi, \ y \rightarrow \%esi \)

<table>
<thead>
<tr>
<th>Line</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><code>movl $0, %eax</code></td>
</tr>
<tr>
<td>2</td>
<td><code>.L2: cmpl %esi, %eax</code></td>
</tr>
<tr>
<td>3</td>
<td><code>jge .L4</code></td>
</tr>
<tr>
<td>4</td>
<td><code>movslq %eax, %rdx</code></td>
</tr>
<tr>
<td>5</td>
<td><code>leaq (%rdi,%rdx,4), %rcx</code></td>
</tr>
<tr>
<td>6</td>
<td><code>movl (%rcx), %edx</code></td>
</tr>
<tr>
<td>7</td>
<td><code>addl $1, %edx</code></td>
</tr>
<tr>
<td>8</td>
<td><code>movl %edx, (%rcx)</code></td>
</tr>
<tr>
<td>9</td>
<td><code>addl $1, %eax</code></td>
</tr>
<tr>
<td>10</td>
<td><code>jmp .L2</code></td>
</tr>
<tr>
<td>11</td>
<td><code>.L4:</code></td>
</tr>
</tbody>
</table>

```
for( int i = 0 ; i < y ; i++ )
```

- **Init**: Line 1
- **Test**: Lines 2-3
- **Update**: Lines 9

```
i-y >= 0
i >= y
```
x86-64 Programming III – Context
Labels & Jumps in C (goto)

- C allows goto as means of transferring control (jump)
  - Closer to assembly programming style
  - Generally considered bad coding style

```c
long absdiff(long x, long y) {
    long result;
    if (x > y)
        result = x - y;
    else
        result = y - x;
    return result;
}
```

```c
long absdiff_j(long x, long y) {
    long result;
    int ntest = (x <= y);
    if (ntest) goto Else;
    result = x - y;
    goto Done;
Else:
    result = y - x;
Done:
    return result;
}
```
Labels & Jumps in C (goto)

- C allows goto as means of transferring control (jump)
  - Closer to assembly programming style
  - Generally considered bad coding style… listen to Kernighan & Ritchie:

> 3.8 Goto and Labels

C provides the infinitely-abusable goto statement, and labels to branch to. Formally, the goto is never necessary, and in practice it is almost always easy to write code without it. We have not used goto in this book.

Nevertheless, there are a few situations where gotos may find a place. The most common is to abandon processing in some deeply nested structure, such as breaking out of two or more loops at once. The break statement cannot be used directly since it only exits from the innermost loop. Thus:
Mainstream ISAs, Revisited

Macbooks & PCs (Core i3, i5, i7, M)  
x86-64 Instruction Set

Smartphone-like devices (iPhone, iPad, Raspberry Pi)  
ARM Instruction Set

Mostly research (some traction in embedded)  
RISC-V Instruction Set

<table>
<thead>
<tr>
<th>ISA</th>
<th>Designer</th>
<th>Bits</th>
<th>Introduced</th>
<th>Design</th>
<th>Type</th>
<th>Encoding</th>
<th>Branching</th>
<th>Endianness</th>
</tr>
</thead>
<tbody>
<tr>
<td>x86</td>
<td>Intel, AMD</td>
<td>16-bit, 32-bit and 64-bit</td>
<td>1978 (16-bit), 1985 (32-bit), 2003 (64-bit)</td>
<td>CISC</td>
<td>Register-memory</td>
<td>Variable (1 to 15 bytes)</td>
<td>Condition code</td>
<td>Little</td>
</tr>
<tr>
<td>ARM</td>
<td>Arm Holdings</td>
<td>32-bit, 64-bit</td>
<td>1985</td>
<td>RISC</td>
<td>Register-Register</td>
<td>AArch64/A64 and AArch32/A32 use 32-bit instructions, T32 (Thumb-2) uses mixed 16- and 32-bit instructions; ARMv7 user-space compatibility</td>
<td>Condition code, compare and branch</td>
<td>Little[1][3]</td>
</tr>
<tr>
<td>RISC-V</td>
<td>University of California, Berkeley</td>
<td>32-bit, 64-bit</td>
<td>2010</td>
<td>RISC</td>
<td>Load-store</td>
<td>Variable</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[1]: User-space compatibility
[3]:Little
Discussion Question

- Discuss the following question(s) in groups of 3-4 students
  - I will call on a few groups afterwards so please be prepared to share out
  - Be respectful of others’ opinions and experiences

- We taught you assembly using x86-64; you didn’t have a choice
  - What are some of the advantages and drawbacks of this choice?
  - What are some possible assumptions we are making about our students or values we are forcing on our students with this choice?
Group Work Time

- During this time, you are encouraged to work on the following:
  1) If desired, continue your discussion
  2) Work on the homework problems
  3) Work on the lab (if applicable)

- Resources:
  - You can revisit the lesson material
  - Work together in groups and help each other out
  - Course staff will circle around to provide support