Assembly Programmer’s Execution Model
Accessing Information
Registers
Memory
Arithmetic operations

BTW: We’re through with Y86 for a while, and starting the x86. We’ll come back to the Y86 later for pipelining.

x86 processors totally dominate the laptop/desktop/server market.

Evolutionary Design
- Starting in 1978 with 8086
- Added more features over time.

Complex Instruction Set Computer (CISC)
- Still support many old, now obsolete, features.
- There are many different instructions with many different formats, but only a small subset are encountered with Linux programs.
- Hard to match performance of Reduced Instruction Set Computers (RISC), though Intel has done just that!

Intel x86 Processors
x86 Evolution: Programmer’s View

<table>
<thead>
<tr>
<th>Model</th>
<th>Date</th>
<th>Trans.</th>
</tr>
</thead>
<tbody>
<tr>
<td>386</td>
<td>1985</td>
<td>0.3M</td>
</tr>
<tr>
<td>Pentium</td>
<td>1993</td>
<td>3.1M</td>
</tr>
<tr>
<td>Pentium/MMX</td>
<td>1997</td>
<td>4.5M</td>
</tr>
<tr>
<td>Pentium Pro</td>
<td>1995</td>
<td>6.5M</td>
</tr>
<tr>
<td>Pentium III</td>
<td>1999</td>
<td>8.2M</td>
</tr>
<tr>
<td>Pentium 4</td>
<td>2001</td>
<td>42M</td>
</tr>
<tr>
<td>Core 2 Duo</td>
<td>2006</td>
<td>291M</td>
</tr>
<tr>
<td>Core i7</td>
<td>2008</td>
<td>731M</td>
</tr>
</tbody>
</table>

Added Features
- Instructions to support multimedia operations
- Instructions to enable more efficient conditional operations
- Transition from 32 to 64 bits
- More cores
2015 State of the Art

Core i7 Broadwell 2015

Desktop Model
- 4 cores
- Integrated graphics
- 3.3–3.8 GHz
- 65W

Server Model
- 8 cores
- Integrated I/O
- 2–2.6 GHz
- 45W

Historically
- AMD has followed behind Intel
- A little bit slower, a lot cheaper

Then
- Recruited top circuit designers from Digital Equipment Corp. and other downward trending companies
- Built Opteron: tough competitor to Pentium 4
- Developed x86-64, their own extension to 64 bits

Recent Years
- Intel got its act together; leads the world in semiconductor technology
- AMD has fallen behind; relies on external semiconductor manufacturers

x86 Evolution: Clones

Transmeta
- Radically different approach to implementation.
  - Translate x86 code into “very long instruction word” (VLIW) code.
  - Very high degree of parallelism.

Centaur / Via
- Continued evolution from Cyrix, the 3rd x86 vendor. Low power, design team in Austin.
  - 32-bit processor family.
    - At 2 GHz, around 2 watts; at 600 MHz around 0.5 watt.
  - 64-bit processor family, used by HP, Lenovo, OLPC, IBM.
    - Very low power, only a few watts at 1.2 GHz.
    - Full virtualization and SSE support.

Intel’s 64-Bit History

- **2001**: Intel attempts radical shift from IA32 to IA64
  - Totally different architecture (Itanium)
  - Executes IA32 code only as legacy
  - Performance disappointing

- **2003**: AMD steps in with evolutionary solution (x86-64, now called AMD64)
  - Intel felt obligated to focus on IA64; hard to admit mistake or that AMD is better

- **2004**: Intel announces EM64T extension to IA32
  - Extended Memory 64-bit technology
  - Almost identical to AMD’s x86-64
  - All but low-end x86 processors support x86-64
  - But lots of code still runs in 32-bit mode.
Definitions:

**Architecture:** (also ISA or instruction set architecture). The parts of a processor design one needs in order to understand or write assembly/machine code.
- Examples: instruction set specification, registers

**Microarchitecture:** implementation of the architecture.
- Examples: cache sizes and core frequency

**Code Forms:**
- Machine code: the byte-level programs that a processor executes
- Assembly code: a textual representation of machine code

**Example ISAs:**
- Intel: x86, IA32, Itanium, x86-64
- ARM: used in almost all mobile phones

---

Assembly Programmer’s View

**Programmer Visible State**
- PC (Program Counter): address of next instruction. Called %rip in x86-64.
- Register file: heavily used program data.
- Condition codes:
  - Store status info about most recent arithmetic operation.
  - Used for conditional branching.

**Memory**
- Byte addressable array.
- Code, user data, (some) OS data.
- Includes stack.

---

ISA Principles

- Contract between programmer and the hardware.
  - Defines visible state of the system.
  - Defines how state changes in response to instructions.

For Programmer: ISA is model of how a program will execute.
For Hardware Designer: ISA is formal definition of the correct way to execute a program.
- With a stable ISA, SW doesn’t care what the HW looks like under the hood.
- Hardware implementations can change drastically.
- As long as the HW implements the same ISA, all prior SW should still run.
- Example: x86 ISA has spanned many chips; instructions have been added but the SW for prior chips still runs.

**ISA specification:** the binary encoding of the instruction set.
ISA Basics

- Instruction formats
- Instruction types
- Addressing modes

Op | Mode | Ra | Rb
---|---|---|---

Architecture vs. Implementation

**Architecture**: defines what a computer system does in response to a program and set of data.
- Programmer visible elements of computer system.

**Implementation (microarchitecture)**: defines how a computer does it.
- Sequence of steps to complete operations.
- Time to execute each operation.
- Hidden “bookkeeping” function.

If the architecture changes, some programs may no longer run or return the same answer. If the implementation changes, some programs may run faster/slower/better, but the answers won’t change.

Examples

Which of the following are part of the architecture and which are part of the implementation?
- Number of general purpose registers
- Width of memory bus
- Binary representation of each instruction
- Number of cycles to execute a FP instruction
- Condition code bits set by a move instruction
- Size of the instruction cache
- Type of FP format

Turning C into Object Code

- Code in files: p1.c, p2.c
- For minimal optimization, compile with command:
  gcc -Og p1.c p2.c -o p
- Use optimization (-Og); new to recent versions of gcc
- Put resulting binary in file p

![Diagram of C into Object Code process]
Compiling into Assembly

C Code (sum.c):

```c
long plus(long x, long y);

void sumstore(long x, long y, long *dest) {
    long t = plus(x, y);
    *dest = t;
}
```

Run command: gcc -Og -S sum.c produces file sum.s.

Assembler

```
sumstore:
    pushq %rbx
    movq %rdx, %rbx
    call plus
    movq %rax, (%rbx)
    popq %rbx
    ret
```

Warning: you may get different results due to variations in gcc and compiler settings.

Object Code

<table>
<thead>
<tr>
<th>Address</th>
<th>Instruction</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x4000595</td>
<td>0x53 0x48 0x89 0xd3 0xe8 0xf2 0xff 0xff 0xff</td>
<td>Total of 14 bytes</td>
</tr>
<tr>
<td>0x400059e</td>
<td>48 89 03</td>
<td>Store value t where designated by dest</td>
</tr>
</tbody>
</table>

Machine Instruction Example

```
*dest = t;
```

C Code

```
void sumstore(long x, long y, long *dest) {
    long t = plus(x, y);
    *dest = t;
}
```

Assembler

```
sumstore:
    pushq %rbx
    movq %rdx, %rbx
    call plus
    movq %rax, (%rbx)
    popq %rbx
    ret
```

Object Code

```
0x4000595: 0x53 0x48 0x89 0xd3 0xe8 0xf2 0xff 0xff 0xff # Total of 14 bytes
```

Minimal Data Types

- “Integer” data of 1, 2, 4 or 8 bytes
- Addresses (untyped pointers)
- Floating point data of 4, 8 or 10 bytes
- No aggregate types such as arrays or structures
- Just contiguously allocated bytes in memory

Primitive Operations

- Perform arithmetic functions on register or memory data
- Transfer data between memory and register
  - Load data from memory into register
  - Store register data into memory
- Transfer control
  - Unconditional jumps to/from procedures
  - Conditional branches

Assembly

```
movq %rax, (%rbx)
```

C Code

```
void sumstore(long x, long y, long *dest) {
    long t = plus(x, y);
    *dest = t;
}
```

Object Code

```
0x400059e: 48 89 03 # 3-byte instruction
```

Store value t where designated by dest

M 0x0400595: 0x53 0x48 0x89 0xd3 0xe8 0xf2 0xff 0xff 0xff # Total of 14 bytes

M 0x400059e: 48 89 03 # 3-byte instruction

Store at address 0x40059e

Operands:
- t: Register %rax
- dest: Register %rbx
- *dest: Memory M[%rbx]
Disassembling Object Code

Disassembled

<table>
<thead>
<tr>
<th>Address</th>
<th>OP 1</th>
<th>OP 2</th>
<th>OP 3</th>
<th>OP 4</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0400595:</td>
<td>push</td>
<td>%rbx</td>
<td></td>
<td></td>
<td>53</td>
</tr>
<tr>
<td>0400596:</td>
<td>mov</td>
<td>%rdx</td>
<td>%rbx</td>
<td></td>
<td>48 89 d3</td>
</tr>
<tr>
<td>0400599:</td>
<td>callq</td>
<td>400590</td>
<td></td>
<td>&lt;plus&gt;</td>
<td></td>
</tr>
<tr>
<td>04005a1:</td>
<td>push</td>
<td>%rax</td>
<td>%rbx</td>
<td></td>
<td>48 89 03</td>
</tr>
<tr>
<td>04005a2:</td>
<td>pop</td>
<td>%rbx</td>
<td></td>
<td></td>
<td>5b</td>
</tr>
<tr>
<td></td>
<td>ret</td>
<td></td>
<td></td>
<td></td>
<td>c3</td>
</tr>
</tbody>
</table>

Disassembler

- objdump -d sum
- Useful tool for examining object code
- Analyzes bit pattern of series of instructions
- Produces approximate rendition of assembly code
- Can be run on either a.out (complete executable) or .o file

Alternate Disassembly

Object code:

<table>
<thead>
<tr>
<th>Address</th>
<th>OP 1</th>
<th>OP 2</th>
<th>OP 3</th>
<th>OP 4</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x53</td>
<td>push</td>
<td>%rbx</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0x48</td>
<td>mov</td>
<td>%rdx</td>
<td>%rbx</td>
<td></td>
<td>48 89</td>
</tr>
<tr>
<td>0x89</td>
<td>callq</td>
<td>0x400590</td>
<td></td>
<td>&lt;plus&gt;</td>
<td></td>
</tr>
<tr>
<td>0xd3</td>
<td>mov</td>
<td>%rax</td>
<td>%rbx</td>
<td></td>
<td>48 89</td>
</tr>
<tr>
<td>0xe8</td>
<td>pop</td>
<td>%rbx</td>
<td></td>
<td></td>
<td>5b</td>
</tr>
<tr>
<td>0xf2</td>
<td>retq</td>
<td></td>
<td></td>
<td></td>
<td>c3</td>
</tr>
</tbody>
</table>

Within gdb debugger:

```
gdb sum
disassemble sumstore
x/14xb sumstore
```

Examine the 14 bytes starting at sumstore.

What Can be Disassembled?

- Anything that can be interpreted as executable code.
- Disassembler examines bytes and reconstructs assembly source.

```
% objdump -d WINWORD.EXE

WINWORD.EXE: file format pei-i386

No symbols in "WINWORD.EXE".

Disassembly of section .text:

30001000 <.text >:
30001000: 55     push %ebp
30001001: 8b      mov %esp, %ebp
30001003: 6a      push $0xffffffff
30001005: 68 90 10 00 30 push $0x3001090
3000100a: 68 69 6c 4c 30 push $0x304cdc91
```

Whose Assembler?

Intel/Microsoft Format

```
lea rax, [%rcx+%rcx*4]
sub rsp, 8
cmp quad ptr [ebp-8], 0
mov rax, quad ptr [rax+4+10h]
```

GAS/Gnu Format

```
lea (%rcx,%rcx,4), %rax
subq $8,%rsp
cmpq $0,-(,%rbp)
movq $0x10(,%rax,4),%rax
```

Intel/Microsoft Differs from GAS

- Operands are listed in opposite order:
  - mov Dest, Src
  - movq Src, Dest
- Constants not preceded by $'; denote hex with 'h' at end.
  - $0x10
- Operand size indicated by operands rather than operator suffix.
  - sub
  - subq
- Addressing format shows effective address computation.
  - [rax*4+10h]

From now on we'll always use GAS assembler format.
x86-64 Integer Registers

<table>
<thead>
<tr>
<th>Reg.</th>
<th>low bytes</th>
<th>Reg.</th>
<th>low bytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>%rax</td>
<td>%eax</td>
<td>%r8</td>
<td>%r8d</td>
</tr>
<tr>
<td>%rbx</td>
<td>%ebx</td>
<td>%r9</td>
<td>%r9d</td>
</tr>
<tr>
<td>%rcx</td>
<td>%ecx</td>
<td>%r10</td>
<td>%r10d</td>
</tr>
<tr>
<td>%rdx</td>
<td>%edx</td>
<td>%r11</td>
<td>%r11d</td>
</tr>
<tr>
<td>%rsi</td>
<td>%esi</td>
<td>%r12</td>
<td>%r12d</td>
</tr>
<tr>
<td>%rdi</td>
<td>%edi</td>
<td>%r13</td>
<td>%r13d</td>
</tr>
<tr>
<td>%r8p</td>
<td>%esp</td>
<td>%r14</td>
<td>%r14d</td>
</tr>
<tr>
<td>%rbp</td>
<td>%ebp</td>
<td>%r15</td>
<td>%r15d</td>
</tr>
</tbody>
</table>

Can reference low-order 4 bytes (also low order 1, 2 bytes)

Some History: IA32 Registers

<table>
<thead>
<tr>
<th>Use</th>
<th>32-bit reg</th>
<th>16-bit reg</th>
<th>8-bit reg</th>
<th>8-bit Reg</th>
</tr>
</thead>
<tbody>
<tr>
<td>accumulator</td>
<td>%eax</td>
<td>%ax</td>
<td>%ah</td>
<td>%al</td>
</tr>
<tr>
<td>counter</td>
<td>%ecx</td>
<td>%cx</td>
<td>%ch</td>
<td>%cl</td>
</tr>
<tr>
<td>data</td>
<td>%edx</td>
<td>%dx</td>
<td>%dh</td>
<td>%dl</td>
</tr>
<tr>
<td>base</td>
<td>%ebx</td>
<td>%bx</td>
<td>%bh</td>
<td>%bl</td>
</tr>
<tr>
<td>source index</td>
<td>%esi</td>
<td>%si</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dest. index</td>
<td>%edi</td>
<td>%di</td>
<td></td>
<td></td>
</tr>
<tr>
<td>stack pointer</td>
<td>%esp</td>
<td>%sp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>base pointer</td>
<td>%ebp</td>
<td>%bp</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Moving Data

Moving Data:
- Form: movq Source, Dest
- Move 8-byte “long” word
- Lots of these in typical code

Operand Types

- **Immediate**: Constant integer data
  - Like C constant, but prefixed with '$'
  - E.g., $0x400, $-533
  - Encoded with 1, 2, or 4 bytes
- **Register**: One of 16 integer registers
  - Example: %rax, %r13
  - But %rsp is reserved for special use
  - Others have special uses for particular instructions
- **Memory**: source/dest is first address of block
  - Example: (%rax)
  - Various “addressing modes”

movq Operand Combinations

Unlike the Y86, we don’t distinguish the operator depending on the operand addressing modes.

<table>
<thead>
<tr>
<th>Source</th>
<th>Dest.</th>
<th>Assembler</th>
<th>C Analog</th>
</tr>
</thead>
</table>
| Immediate| Register| movq $0x4,%rax | temp = 0x4;
| Immediate| Memory  | movq $-147,(%rax) | *p = -147;
| Register | Register| movq %rax,%rdx  | temp2 = temp1;
| Register | Memory  | movq %rax,(%rdx) | *p = temp;  
| Memory   | Register| movq (%rax),%rdx| temp = *p  

Memory-memory transfers are not allowed within a single instruction.
Simple Addressing Modes

- **Immediate**: value
  - `movq $0xab, %rbx`

- **Register**: Reg[R]
  - `movq %rcx, %rbx`

- **Normal (R)**: Mem[Reg[R]]
  - Register R specifies memory address.
  - This is often called *indirect* addressing.
  - Aha! Pointer dereferencing in C
    - `movq (%rcx), %rax`

- **Displacement D(R)**: Mem[Reg[R] + D]
  - Register R specifies start of memory region.
  - Constant displacement D specifies offset
    - `movq 8(%rcb), %rdx`

Addresses and Pointers in C

- **C programming model is close to machine language.**
  - Machine language manipulates memory addresses.
  - For address computation;
  - To store addresses in registers or memory.
  - C employs pointers, which are just addresses of primitive data elements or data structures.

- **Examples of operators * and &:**
  - `int a, b; /* declare integers a and b */`
  - `int *a_ptr; /* a is a pointer to an integer */`
  - `a_ptr = a; /* illegal, types don’t match*/`
  - `a_ptr = &a; /* a_ptr holds address of a */`
  - `b = *a_ptr; /* dereference a_ptr and assign value to b */`

Using Simple Addressing Modes

**void** swap( **long** *xp, **long** *yp) {
  **long** t0 = *xp;
  **long** t1 = *yp;
  *xp = t1;
  *yp = t0;
}

Understanding Swap (1)

```c
void swap( long *xp, long *yp) {
  long t0 = *xp;
  long t1 = *yp;
  *xp = t1;
  *yp = t0;
}
```

<table>
<thead>
<tr>
<th>Register</th>
<th>Value</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>%rdi</td>
<td>xp</td>
<td>points into memory</td>
</tr>
<tr>
<td>%rsi</td>
<td>yp</td>
<td>points into memory</td>
</tr>
<tr>
<td>%rax</td>
<td>t0</td>
<td>temporary storage</td>
</tr>
<tr>
<td>%rdx</td>
<td>t1</td>
<td>temporary storage</td>
</tr>
</tbody>
</table>
Understanding Swap (2)

\[
\text{swap:} \\
\text{movq } (%\ rdi), \ %rax \quad \# \ t0 = *xp \\
\text{movq } (%\ rsi), \ %rdx \quad \# \ t1 = *yp \\
\text{movq } %\ rdx, \ (%\ rdi) \quad \# \ *xp = t1 \\
\text{movq } %\ rax, \ (%\ rsi) \quad \# \ *yp = t0 \\
\text{ret}
\]

Initial State:

<table>
<thead>
<tr>
<th>Registers</th>
<th>Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>%rdi 0x120</td>
<td>123 0x120</td>
</tr>
<tr>
<td>%rsi 0x100</td>
<td>0x118</td>
</tr>
<tr>
<td>%rax</td>
<td>0x110</td>
</tr>
<tr>
<td>%rdx</td>
<td>0x108</td>
</tr>
<tr>
<td></td>
<td>456 0x100</td>
</tr>
</tbody>
</table>

Understanding Swap (3)

\[
\text{swap:} \\
\text{movq } (%\ rdi), \ %rax \quad \# \ t0 = *xp, \leftarrow \ PC \ here \\
\text{movq } (%\ rsi), \ %rdx \quad \# \ t1 = *yp \\
\text{movq } %\ rdx, \ (%\ rdi) \quad \# \ *xp = t1 \\
\text{movq } %\ rax, \ (%\ rsi) \quad \# \ *yp = t0 \\
\text{ret}
\]

Understanding Swap (4)

\[
\text{swap:} \\
\text{movq } (%\ rdi), \ %rax \quad \# \ t0 = *xp \\
\text{movq } (%\ rsi), \ %rdx \quad \# \ t1 = *yp, \leftarrow \ PC \ here \\
\text{movq } %\ rdx, \ (%\ rdi) \quad \# \ *xp = t1 \\
\text{movq } %\ rax, \ (%\ rsi) \quad \# \ *yp = t0 \\
\text{ret}
\]

Understanding Swap (5)

\[
\text{swap:} \\
\text{movq } (%\ rdi), \ %rax \quad \# \ t0 = *xp \\
\text{movq } (%\ rsi), \ %rdx \quad \# \ t1 = *yp \\
\text{movq } %\ rdx, \ (%\ rdi) \quad \# \ *xp = t1, \leftarrow \ PC \ here \\
\text{movq } %\ rax, \ (%\ rsi) \quad \# \ *yp = t0 \\
\text{ret}
\]

<table>
<thead>
<tr>
<th>Registers</th>
<th>Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>%rdi 0x120</td>
<td>123 0x120</td>
</tr>
<tr>
<td>%rsi 0x100</td>
<td>0x118</td>
</tr>
<tr>
<td>%rax 123</td>
<td>0x110</td>
</tr>
<tr>
<td>%rdx 456</td>
<td>0x108</td>
</tr>
<tr>
<td></td>
<td>456 0x100</td>
</tr>
</tbody>
</table>
Understanding Swap (6)

```
swap:
  movq (%rdi), %rax  # t0 = *xp
  movq (%rsi), %rdx  # t1 = *yp
  movq %rdx, (%rdi)  # *xp = t1
  movq %rax, (%rsi)  # *yp = t0, <-- PC here
  ret
```

**Simple Addressing Modes**

- **Immediate**: value
  ```
  movq $0xab, %rbx
  ```

- **Register**: Reg[R]
  ```
  movq %rcx, %rbx
  ```

- **Normal (R)**: Mem[Reg[R]]
  - Register R specifies memory address.
  - This is often called *indirect* addressing.
  - Aha! Pointer dereferencing in C
  ```
  movq (%rcx), %rax
  ```

- **Displacement D(R)**: Mem[Reg[R] + D]
  - Register R specifies start of memory region.
  - Constant displacement D specifies offset
  ```
  movq 8(%rcx), %rdx
  ```

**Indexed Addressing Modes**

**Most General Form**:

\[ D(E_b, E_i, s) = M[E_b + E_i \cdot s] \]

- D: Constant “displacement” of 1, 2 or 4 bytes
- Rb: Base register, any of the 16 integer registers
- Ri: Index register, any except %rsp (and probably not %rbp)
- S: Scale, one of 1, 2, 4 or 8.

**Special Cases**:

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

- Immediate
  ```
  movq $0xab, %rbx
  ```

- Register
  ```
  movq %rcx, %rbx
  ```

- Normal (R)
  ```
  movq (%rcx), %rax
  ```

- Displacement D(R)
  ```
  movq 8(%rcx), %rdx
  ```

The scaling factor \(s\) must be either 1, 2, 4, or 8.
### Address Computation Example

<table>
<thead>
<tr>
<th>%rdx</th>
<th>0xf000</th>
</tr>
</thead>
<tbody>
<tr>
<td>%rcx</td>
<td>0x100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Expression</th>
<th>Computation</th>
<th>Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x8(%rdx)</td>
<td>0xf000 + 0x8</td>
<td>0xf008</td>
</tr>
<tr>
<td>(%rdx, %rcx)</td>
<td>0f000 + 0x100</td>
<td>0xf100</td>
</tr>
<tr>
<td>(%rdx, %rcx, 4)</td>
<td>0xf000 + 4*0x100</td>
<td>0xf400</td>
</tr>
<tr>
<td>0x80(%rdx, 2)</td>
<td>2*0xf000 + 0x80</td>
<td>0x1e080</td>
</tr>
<tr>
<td>0x80(%rdx, 2)</td>
<td>illegal</td>
<td></td>
</tr>
<tr>
<td>0x80(%rdx, 3)</td>
<td>illegal</td>
<td></td>
</tr>
</tbody>
</table>

### Some Arithmetic Operations

**Two operand instructions:**

<table>
<thead>
<tr>
<th>Format</th>
<th>Computation</th>
</tr>
</thead>
<tbody>
<tr>
<td>addq Src, Dest</td>
<td>Dest = Dest + Src</td>
</tr>
<tr>
<td>subq Src, Dest</td>
<td>Dest = Dest - Src</td>
</tr>
<tr>
<td>imulq Src, Dest</td>
<td>Dest = Dest * Src</td>
</tr>
<tr>
<td>salq Src, Dest</td>
<td>Dest = Dest &lt;&lt; Src</td>
</tr>
<tr>
<td>sarq Src, Dest</td>
<td>Dest = Dest &gt;&gt; Src</td>
</tr>
<tr>
<td>shrq Src, Dest</td>
<td>Dest = Dest &gt;&gt;&gt; Src</td>
</tr>
<tr>
<td>xorq Src, Dest</td>
<td>Dest = Dest ^ Src</td>
</tr>
<tr>
<td>andq Src, Dest</td>
<td>Dest = Dest &amp; Src</td>
</tr>
<tr>
<td>orq Src, Dest</td>
<td>Dest = Dest</td>
</tr>
</tbody>
</table>

- Watch out for argument order!
- There’s no distinction between signed and unsigned. Why?

**One operand instructions:**

<table>
<thead>
<tr>
<th>Format</th>
<th>Computation</th>
</tr>
</thead>
<tbody>
<tr>
<td>incq Dest</td>
<td>Dest = Dest + 1</td>
</tr>
<tr>
<td>decq Dest</td>
<td>Dest = Dest - 1</td>
</tr>
<tr>
<td>negq Dest</td>
<td>Dest = ~Dest</td>
</tr>
<tr>
<td>notq Dest</td>
<td>Dest = ~Dest</td>
</tr>
</tbody>
</table>

More instructions in the book.

### Address Computation Instruction

**Form:** leaq Src, Dest

- Src is address mode expression.
- Sets Dest to address denoted by the expression

LEA stands for “load effective address.”

After the effective address computation, place the address, not the contents of the address, into the destination.
Consider the following computation:

<table>
<thead>
<tr>
<th>Reg.</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>%rax</td>
<td>0x100</td>
</tr>
<tr>
<td>%rbx</td>
<td>0x200</td>
</tr>
</tbody>
</table>

After this sequence,
- %rcx will contain the contents of location 0x610;
- %rdx will contain the number (address) 0x610.

What should the following do?

```
leaq %rbx, %rdx
```

It really shouldn’t be legal since %rbx doesn’t have an address. However, the semantics makes it equal to `movq %rbx, %rdx`.

---

The `leaq` instruction is widely used for address computations and for some general arithmetic computations.

**Uses:**
- Computing address without doing a memory reference:
  - E.g., translation of `p = &x[i];`
- Computing arithmetic expressions of the form `x + k \times y`, where `k \in \{1, 2, 4, 8\}`

**Example:**

```c
long m12(long x) {
    return x*12;
}
```

**Converted to ASM by compiler:**

```asm
leaq (%rdi,%rdi,2),%rax       # t <- x+x*2
salq $2,%rax                  # ret. t<<2
```

**Interesting instructions:**
- `leaq`: address computation
- `salq`: shift
- `imulq`: multiplication, but only used once
Understanding our Arithmetic Expression Example

```c
long arith
(long x, long y, long z)
{
    long t1 = x + y;
    long t2 = z + t1;
    long t3 = x + 4;
    long t4 = y * 48;
    long t5 = t3 + t4;
    long rval = t2 * t5;
    return rval;
}
```

<table>
<thead>
<tr>
<th>Register</th>
<th>Use(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>%rdi</td>
<td>Argument x</td>
</tr>
<tr>
<td>%rsi</td>
<td>Argument y</td>
</tr>
<tr>
<td>%rdx</td>
<td>Argument z</td>
</tr>
<tr>
<td>%rax</td>
<td>t1, t2, rval</td>
</tr>
<tr>
<td>%rdx</td>
<td>t4</td>
</tr>
<tr>
<td>%rcx</td>
<td>t5</td>
</tr>
</tbody>
</table>

History of Intel processors and architectures
- Evolutionary design leads to many quirks and artifacts

C, assembly, machine code
- New forms of visible state: program counter, registers, etc.
- Compiler must transform statements, expressions, procedures into low-level instruction sequences

Assembly Basics: Registers, operands, move
- The x86-64 move instructions cover a wide range of data movement forms

Arithmetic
- C compiler will figure out different instruction combinations to carry out computation