Machine-Level Representation of Programs: Procedures, Recursion, Data

Yipeng Huang

Rutgers University

April 3, 2023
Table of contents

Announcements

Procedures and function calls
  Memory stack frames

Procedures and function calls: Transferring control
  Procedure call and return: call and ret
  Example in GDB

Procedures and function calls: Transferring data
  Data transferred via registers
  Data transferred via memory

Architecture support for recursive programming
Announcements

Class session plan

- Monday, 4/3: Arrays and data structures in assembly. (Book chapter 3.8)
Table of contents

Announcements

Procedures and function calls
  Memory stack frames

Procedures and function calls: Transferring control
  Procedure call and return: call and ret
  Example in GDB

Procedures and function calls: Transferring data
  Data transferred via registers
  Data transferred via memory

Architecture support for recursive programming
Procedures and function calls

To create the abstraction of functions, need to:

- Transfer control to function and back
- Transfer data to function (parameters)
- Transfer data from function (return type)

Figure: Steps of a C function call. Image credit CS:APP
Structure of stack for currently executing function Q()

- P() calls Q(). P() is the caller function. Q() is the callee function.
Stack instructions: **push** src and **pop** dest

<table>
<thead>
<tr>
<th>Initially</th>
<th>pushq %rax</th>
<th>popq %rdx</th>
</tr>
</thead>
<tbody>
<tr>
<td>%rax</td>
<td>0x123</td>
<td>%rax</td>
</tr>
<tr>
<td>%rdx</td>
<td>0</td>
<td>%rdx</td>
</tr>
<tr>
<td>%rsp</td>
<td>0x108</td>
<td>%rsp</td>
</tr>
</tbody>
</table>

Initially, the stack is as follows:

- %rax: 0x123
- %rdx: 0
- %rsp: 0x108

### After pushq %rax

- %rax: 0x123
- %rdx: 0
- %rsp: 0x100

### After popq %rdx

- %rax: 0x123
- %rdx: 0x123
- %rsp: 0x108

**Figure:** x86-64 offers dedicated instructions to work with stack in memory. In addition to moving data, the updating of %rsp is implied. Image credit: CS:APP.
Table of contents

Announcements

Procedures and function calls
  Memory stack frames

Procedures and function calls: Transferring control
  Procedure call and return: call and ret
  Example in GDB

Procedures and function calls: Transferring data
  Data transferred via registers
  Data transferred via memory

Architecture support for recursive programming
CPU and memory state in support of procedures and functions

Assembly/Machine Code View

Programmer-Visible State

- **PC**: Program counter
  - Address of next instruction
  - Called “RIP” (x86-64)
- **Register file**
  - Heavily used program data
- **Condition codes**
  - Store status information about most recent arithmetic or logical operation
  - Used for conditional branching

Memory

- **Memory**
  - Byte addressable array
  - Code and user data
  - Stack to support procedures

Relevant state in CPU:

- `%rip` register / instruction pointer / program counter
- `%rsp` register / stack pointer

Relevant state in Memory:

- Stack
Procedure call and return: \textit{call} and \textit{ret}

\begin{center}
\begin{tabular}{|c|c|}
\hline
\textbf{%rip} & 0x400563 \\
\hline
\textbf{%rsp} & 0x7fffffffe840 \\
\hline
\end{tabular}
\begin{tabular}{|c|c|}
\hline
\textbf{%rip} & 0x400540 \\
\hline
\textbf{%rsp} & 0x7fffffffe838 \\
\hline
\end{tabular}
\begin{tabular}{|c|c|}
\hline
\textbf{%rip} & 0x400568 \\
\hline
\textbf{%rsp} & 0x7fffffffe840 \\
\hline
\end{tabular}
\end{center}

\begin{enumerate}
\item[(a)] Executing \textit{call}
\item[(b)] After \textit{call}
\item[(c)] After \textit{ret}
\end{enumerate}

\textit{Figure:} Effect of call 0x400540 instruction and subsequent return. call and ret instructions update the instruction pointer, the stack pointer, and the stack to create the procedure / function call abstraction. Image credit: CS:APP.
Example in GDB

```c
#include <stdio.h>

int return_neg_one() {
    return -1;
}

int main() {
    int num = return_neg_one();
    printf("%d", num);
    return 0;
}
```

Compile, and then run it in GDB:
gdb return

In GDB, see evolution of %rip, %rsp, and stack:

- (gdb) layout split
- (gdb) break return_neg_one
- (gdb) info stack
- (gdb) print /a $rip
- (gdb) print /a $rsp
- (gdb) x /a $rsp

Step past return instruction, and inspect again:

- (gdb) stepi
- (gdb) info stack
Table of contents

Announcements

Procedures and function calls
  Memory stack frames

Procedures and function calls: Transferring control
  Procedure call and return: call and ret
  Example in GDB

Procedures and function calls: Transferring data
  Data transferred via registers
  Data transferred via memory

Architecture support for recursive programming
Procedures and function calls: Transferring data

For purposes of this class, the Bomb Lab, and the CS:APP textbook, we study the x86-64 Linux Application Binary Interface (ABI). Would be different on ARM or in Windows. So, don’t memorize this, but it is helpful for PA4 Lab.

Passing parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Register / stack</th>
<th>Subset registers</th>
<th>Mnemonic¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>%rdi</td>
<td>%edi, %di</td>
<td>Diane’s</td>
</tr>
<tr>
<td>2nd</td>
<td>%rsi</td>
<td>%esi, %si</td>
<td>silk</td>
</tr>
<tr>
<td>3rd</td>
<td>%rdx</td>
<td>%edx, %dx, %dl</td>
<td>dress</td>
</tr>
<tr>
<td>4th</td>
<td>%rcx</td>
<td>%ecx, %cx, %cl</td>
<td>cost</td>
</tr>
<tr>
<td>5th</td>
<td>%r8</td>
<td>%r8d</td>
<td>$8</td>
</tr>
<tr>
<td>6th</td>
<td>%r9</td>
<td>%r9d</td>
<td>9</td>
</tr>
<tr>
<td>7th and beyond</td>
<td>Stack</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PA4 Defusing a Binary Bomb: `sscanf()` ;

```c
int scanf (  
    const char *str, // 1st arg, %rdi  
    const char *format, // 2nd arg, %rsi  
    ...  
)
```
Procedures and function calls: Transferring data

Passing function return data
Function return data is passed via:
- the 64-bit %rax register
- the 32-bit subset %eax register

Example from textbook slides on assembly procedures
Slides 33 through 38.
Structure of stack for currently executing function Q()

- P() calls Q(). P() is the caller function. Q() is the callee function.

Example from textbook slides on assembly procedures

Slides 40 through 44.
Table of contents

Announcements

Procedures and function calls
Memory stack frames

Procedures and function calls: Transferring control
Procedure call and return: call and ret
Example in GDB

Procedures and function calls: Transferring data
Data transferred via registers
Data transferred via memory

Architecture support for recursive programming
Discussion points

- Use info stack, info args in GDB to see recursion depth
- Difference between compiling with and without -g for debugging information.
- Memory costs of recursion.
- Compilers can recognize tail recursive calls to reduce memory use. Enabled with -foptimize-sibling-calls, -O2, -O3, and -Os.