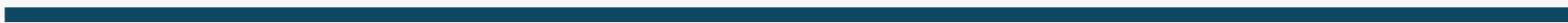




Assembly Code Analysis

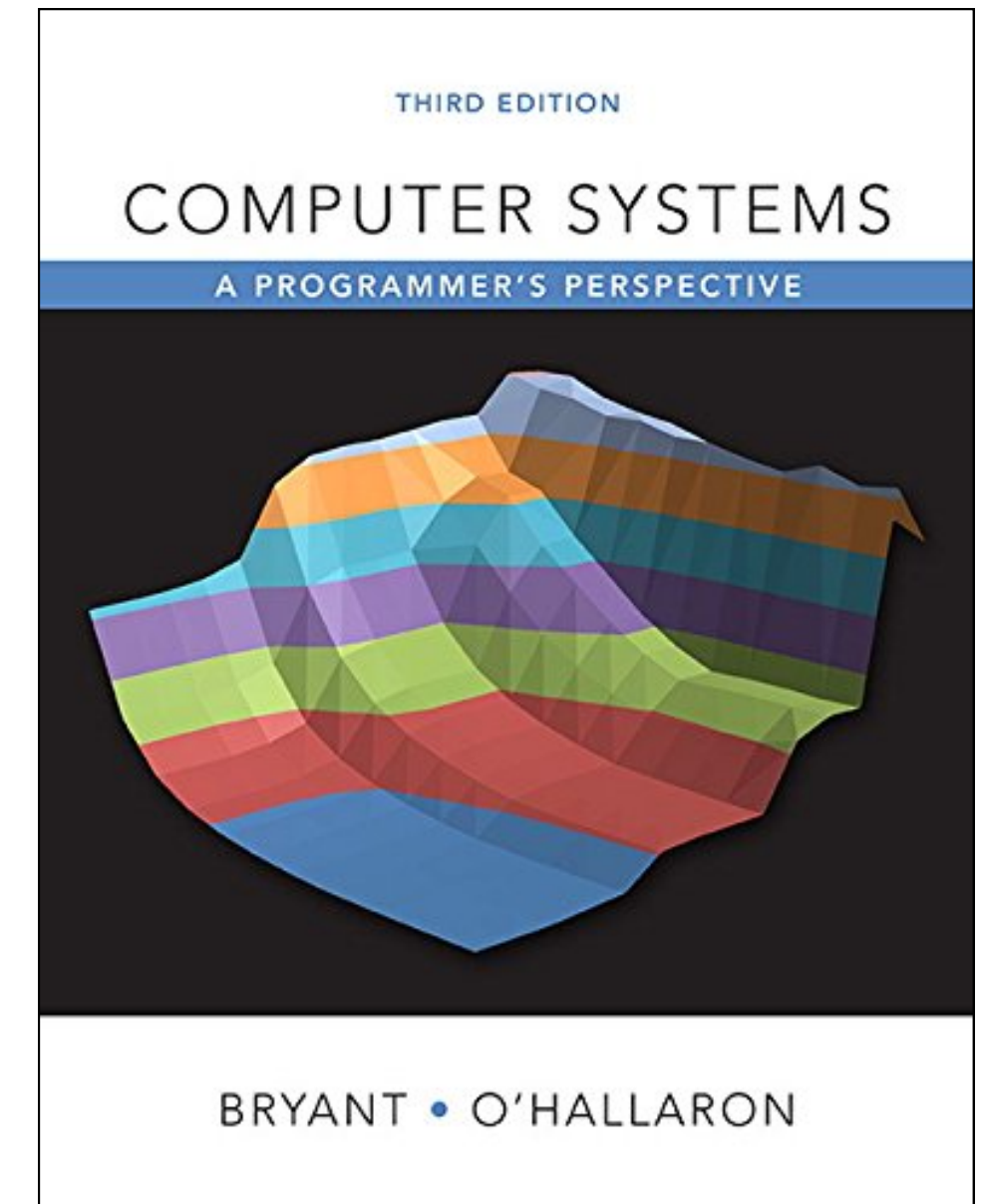


Motivation

Motivation

由於前陣子看了一本名為 Computer Systems A Programmer's Perspective (CS:APP), 讀到關於 x86-64 assembly code 的說明, 以及在 function calling 下, stack 與 registers 會如何變化。

因此想藉此期中報告的實作, 觀察 assembly code 與 register 來驗證書中所說明是否為真, 也能加強自身對於計算機領域的基本知識理解。



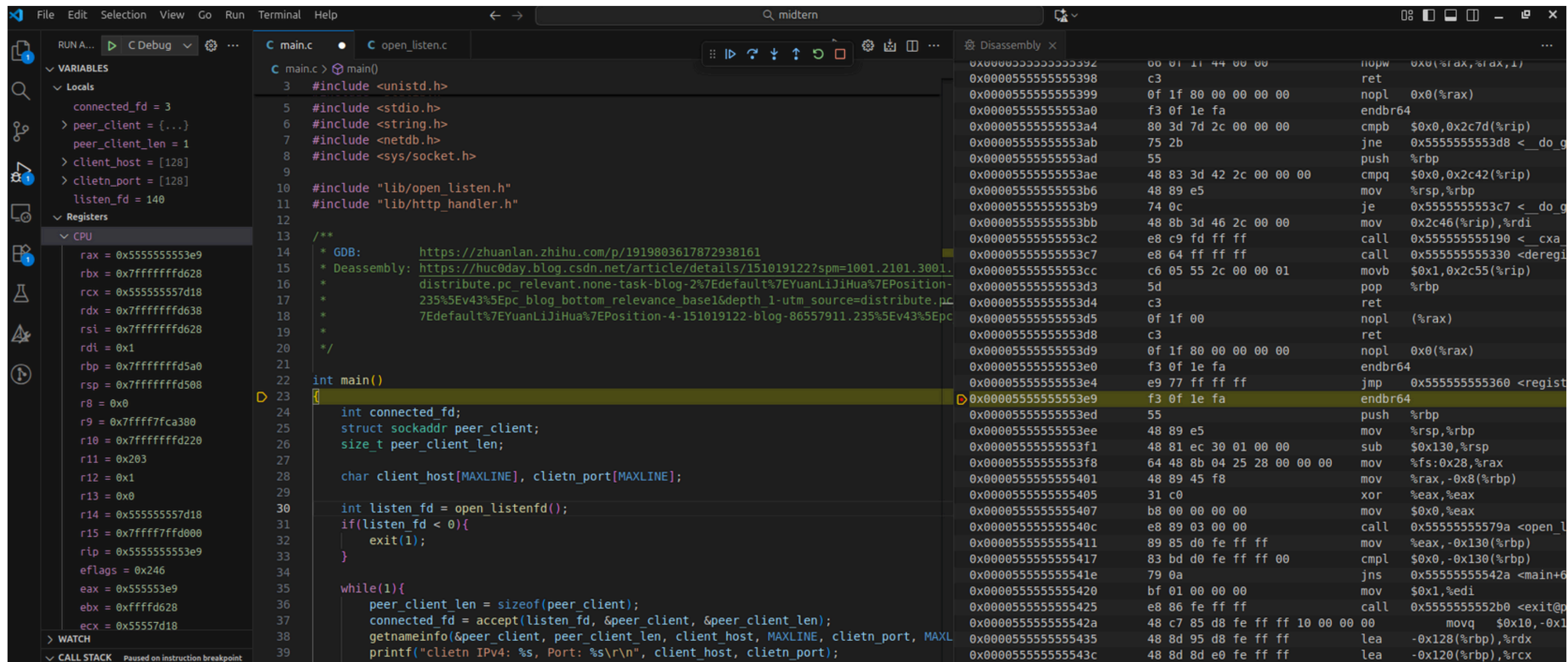
<https://www.tenlong.com.tw/products/9780134092669>

https://www.cs.sfu.ca/~ashriram/Courses/CS295/assets/books/CSAPP_2016.pdf

Method

Method

藉由 VScode 做調適 (Debug), 並對照可執行檔 (executable file) 的 deassembly (解彙編) 逐行執行並觀察，
藉由理解 assembly code 的執行與猜測 register value 的變化並藉由 debug mode 提供逐行執行驗證猜測。



Trace

Trace

雖然在 x86-64 中有許多通用 register，但只需要注意幾個重要的即可。

- %rax → 儲存 return value
- %rsp → 儲存棧頂部
- %rbp → 儲存當前 function 的起始棧位置
- %rip → 儲存 process 執行位置 (PC)
- %rdi, rsi, rdx, rcx, r8, r9 → 儲存 function 參數



Figure 3.2 Integer registers. The low-order portions of all 16 registers can be accessed as byte, word (16-bit), double word (32-bit), and quad word (64-bit) quantities.

Trace

首先來看 main function

13ed → 將 %rbp 內容儲存到 stack 中 (push 會將 %rsp - 0x08 再將 %rbp 內容儲存到 stack)

13ee → 將 %rsp 的值複製到 %rbp 中

13f1 → 將 %rsp 減去 0x130，用於分配 stack space 給 main function

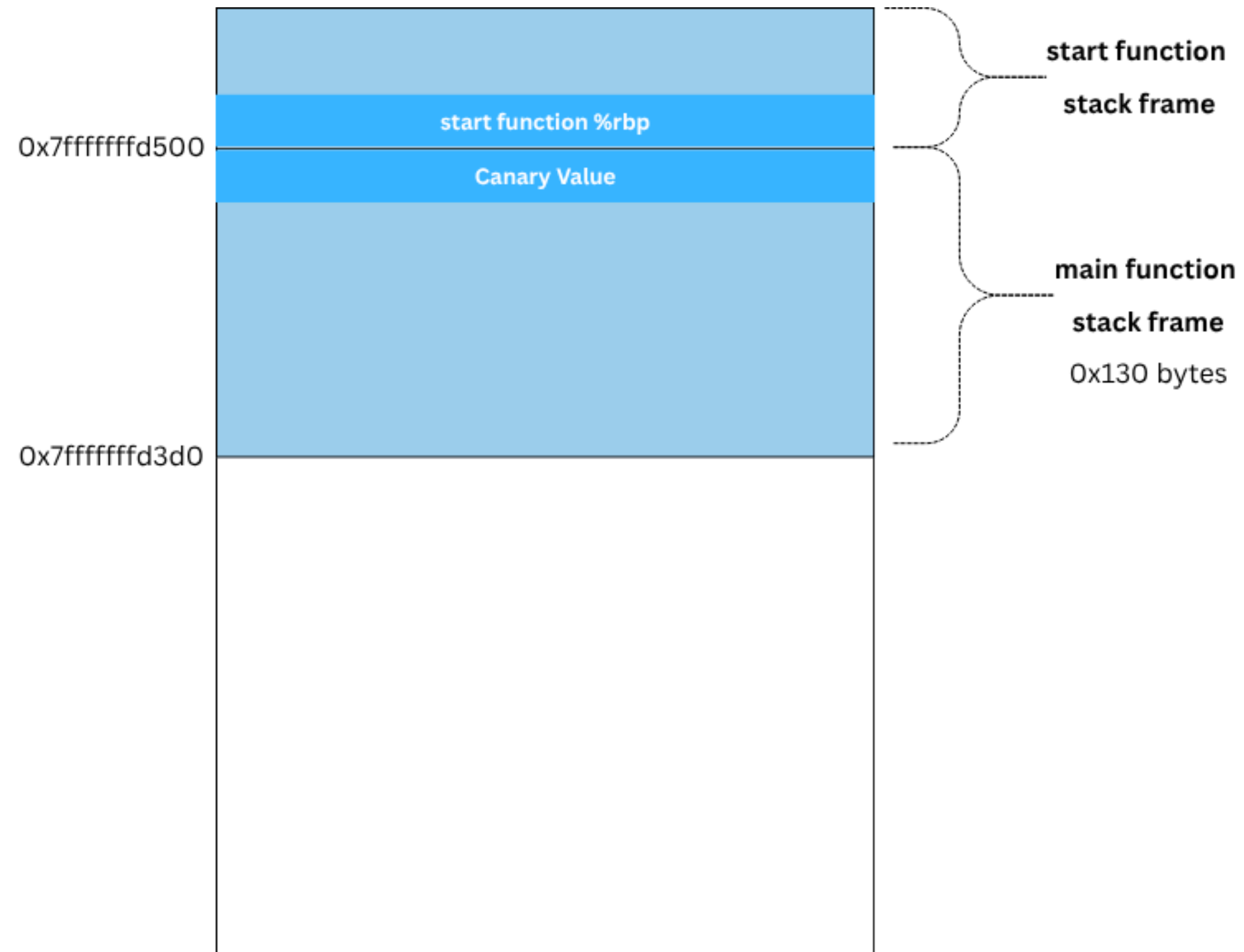
13f8 → 將 Canary value 複製到 %rax 中

1401 → 將 %rax 值複製到 %rbp-0x8 的位置中

1405 → 將 %rax 清除為 0

```
int main()
{
    13e9:      f3 0f 1e fa      endbr64
    13ed:      55          push    %rbp
    13ee:      48 89 e5      mov     %rsp,%rbp
    13f1:      48 81 ec 30 01 00 00  sub     $0x130,%rsp
    13f8:      64 48 8b 04 25 28 00  mov     %fs:0x28,%rax
    13ff:      00 00
    1401:      48 89 45 f8      mov     %rax,-0x8(%rbp)
    1405:      31 c0      xor     %eax,%eax
    1407:      5d          pop     %rbp
    1408:      f3 0f 1e fa      endbr64
    140b:      90          ret     0
}
```


Trace



Trace

The screenshot shows the GDB debugger interface with the following components:

- Menu Bar:** File, Edit, Selection, View, Go, Run, Terminal, Help.
- Toolbar:** Includes icons for running, stepping, and other debugging actions.
- Variables Panel (Left):**
 - Locals:** listen_fd = 140.
 - Registers (CPU):**
 - rax = 0x0
 - rbx = 0x7fffffff628
 - rcx = 0x55555557d18
 - rdx = 0x7fffffff638
 - rsi = 0x7fffffff628
 - rdi = 0x1
 - rbp = 0x7fffffff500
 - rsp = 0x7fffffff3d0
 - r8 = 0x0
 - r9 = 0x7ffff7fca380
 - r10 = 0x7fffffff220
 - r11 = 0x203
 - r12 = 0x1
 - r13 = 0x0
 - r14 = 0x55555557d18
 - r15 = 0x7ffff7ffd000
 - rip = 0x55555555407
 - eflags = 0x246
 - eax = 0x0
 - ebx = 0xffffd628
 - ecx = 0x55557d18
 - edx = 0xffffd638
 - esi = 0xffffd628
 - edi = 0x1
 - ebp = 0xffffd500
- Source Code Panel (Middle):** main.c, open_listen.c. The main() function is visible, starting with #include directives and a while loop for listening.
- Disassembly Panel (Right):** Shows the assembly code for the current instruction, including push, cmpq, mov, je, call, movb, pop, ret, nopl, endbr64, jmp, and xor instructions.

Trace

The screenshot displays the GDB debugger interface with the following components:

- VARIABLES Panel (Left):** Shows local variables and registers. Local variables include `peer_client_len = 1`, `client_host = [128]`, `clietn_port = [128]`, and `listen_fd = 140`. Registers show values for `rax` through `esi`.
- main.c Editor (Center):** Displays the source code of `main.c`. The code includes headers like `<unistd.h>`, `<stdio.h>`, `<string.h>`, `<netdb.h>`, and `<sys/socket.h>`. It defines `MAXLINE` and `MAXBUF`, and implements a `main()` function that sets up a listener socket and enters a loop to accept connections.
- Disassembly Panel (Right):** Shows the assembly code corresponding to the current line in the source code. It includes instructions like `jne`, `push`, `cmpq`, `mov`, `je`, `call`, `movb`, `pop`, `ret`, `nopl`, `endbr64`, `jmp`, `push`, `mov`, `sub`, `mov`, `xor`, `mov`, `call`, `mov`, `cmpl`, `jns`, `mov`, `call`, `movq`, `lea`, `mov`, `mov`, `call`, `mov`, and `mov`.

Trace

The screenshot displays a debugger's trace window for the `open_listenfd()` function in `open_listen.c`. The interface is divided into three main sections:

- VARIABLES (Left):** Shows the state of registers and local variables. Key registers include `rax` (0x0), `rbx` (0x7fffffff628), `rcx` (0x55555557d18), `rdx` (0x7fffffff638), `rsi` (0x7fffffff628), `rdi` (0x1), `rbp` (0x7fffffff500), `rsp` (0x7fffffff3c8), `r8` (0x0), `r9` (0x7ffff7fca380), `r10` (0x7fffffff220), `r11` (0x203), `r12` (0x1), `r13` (0x0), `r14` (0x55555557d18), `r15` (0x7ffff7ffd000), and `rip` (0x5555555579a). Local variables include `eflags` (0x246), `eax` (0x0), `ebx` (0xffffd628), `ecx` (0x55557d18), `edx` (0xffffd638), `esi` (0xffffd628), `edi` (0x1), `ebp` (0xffffd500), and `esp` (0xffffd3c8).
- Code Window (Center):** Shows the C source code for `open_listenfd()`. The function includes headers for `netdb.h`, `unistd.h`, `sys/types.h`, `sys/socket.h`, and `string.h`. It defines a `hints` struct, initializes it with `memset`, and attempts to get address information using `getaddrinfo`. If successful, it prints the IPv4 address and port, then creates a socket using `socket`.
- Instruction Trace (Right):** Provides a detailed view of the assembly instructions being executed. Each line shows a memory address, a hex value, and the assembly instruction. For example, the trace starts with `0x000055555555779 48 89 ce mov %rcx,%rsi` and continues through various instructions like `mov %eax,%edi`, `call 0x555555554db <response>`, `nop`, `mov -0x8(%rbp),%rax`, `sub %fs:0x28,%rax`, `je 0x55555555798 <request>`, `call 0x555555551d0 <__stack_chk_guard>`, `leave`, `ret`, `endbr64`, `push %rbp`, `mov %rsp,%rbp`, `sub $0xd0,%rsp`, `mov %fs:0x28,%rax`, `mov %rax,-0x8(%rbp)`, `xor %eax,%eax`, `lea -0xb0(%rbp),%rax`, `mov $0x30,%edx`, `mov $0x0,%esi`, `mov %rax,%rdi`, `call 0x55555555210 <memset@plt>`, `movl $0x2,-0xac(%rbp)`, `movl $0x1,-0xb0(%rbp)`, `movl $0x1,-0xa8(%rbp)`, `lea -0xc0(%rbp),%rdx`, `lea -0xb0(%rbp),%rax`, `mov %rdx,%rcx`, `mov %rax,%rdx`, `lea 0x8b6(%rip),%rax`, `mov %rax,%rsi`, `mov $0x0,%edi`, `call 0x555555552d0 <getaddrinfo@plt>`, `mov %eax,-0xc4(%rbp)`, `cmpl $0x0,-0xc4(%rbp)`, `je 0x5555555585e <open_listenfd+0x100>`, `mov -0xc4(%rbp),%eax`, `mov %eax,%edi`, `call 0x55555555200 <gai_strerror@plt>`, `mov %rax,%rdx`, and `mov 0x27e3(%rip),%rax`.

Trace

首先來看 open_listenfd function

179e → 將 %rbp 內容儲存到 stack 中 (push 會將 %rsp - 0x08 再將 %rbp 內容儲存到 stack)

179f → 將 %rsp 的值複製到 %rbp 中

17a2→ 將 %rsp 減去 0xd0，用於分配 stack space 給 open_listenfd function

17a9 → 將 Canary value 複製到 %rax 中

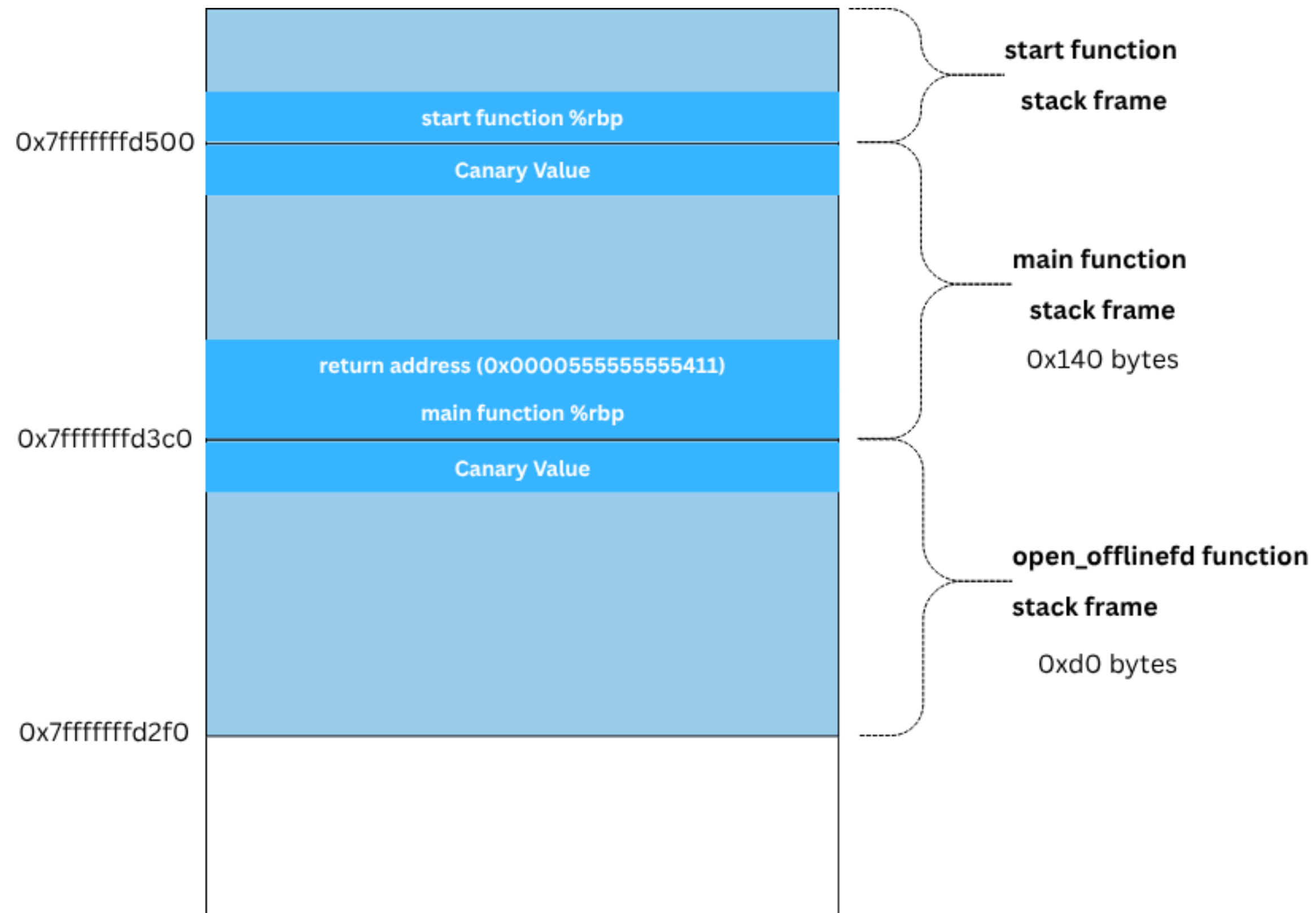
17b0 → 將 %rax 值複製到 %rbp-0x8 的位置中

17b2 → 將 %rax 清除為 0

```
000000000000179a <open_listenfd>:
#include <sys/types.h>
#include <sys/socket.h>
#include <string.h>

int open_listenfd()
{
    179a:    f3 0f 1e fa                endbr64
    179e:    55                        push    %rbp
    179f:    48 89 e5                  mov     %rsp,%rbp
    17a2:    48 81 ec d0 00 00 00      sub     $0xd0,%rsp
    17a9:    64 48 8b 04 25 28 00      mov     %fs:0x28,%rax
    17b0:    00 00
    17b2:    48 89 45 f8              mov     %rax,-0x8(%rbp)
    17b6:    31 c0                    xor     %eax,%eax
    17ba:    41 51 65 65 65 65 65 65  mov     %eax,%eax
    17c0:    5d                        pop     %rbp
    17c1:    c3                        ret
}
```

Trace



Trace

再來看 open_listenfd function 呼叫 memset() function 是如何傳遞參數的

17ba → 將 %rbp 減去 0xb0，並複製到 %rax

17bf → 將 0x30 複製到 %edx 中

17c4 → 將 0x0 值複製到 %esi

17c9 → 將 %rax 的值複製到 %rdi

17cc → 呼叫 function memset()，並將下一個 instruction 的 address 壓入到 stack (%rsp -= 8)，並設置 %rip 為 memset() 的位置。

```
memset(&hints, 0, sizeof(hints));  
17b8:      48 8d 85 50 ff ff ff      lea    -0xb0(%rbp),%rax  
17bf:      ba 30 00 00 00          mov     $0x30,%edx  
17c4:      be 00 00 00 00          mov     $0x0,%esi  
17c9:      48 89 c7                mov     %rax,%rdi  
17cc:      e8 3f fa ff ff          call    1210 <memset@plt>
```


Trace

VARIABLES

- Locals
 - rax = 0x0
 - rbx = 0x7fffffff628
 - rcx = 0x55555557d18
 - rdx = 0x7fffffff638
 - rsi = 0x7fffffff628
 - rdi = 0x1
 - rbp = 0x7fffffff3c0
 - rsp = 0x7fffffff2f0
 - r8 = 0x0
 - r9 = 0x7ffff7fca380
 - r10 = 0x7fffffff220
 - r11 = 0x203
 - r12 = 0x1
 - r13 = 0x0
 - r14 = 0x55555557d18
 - r15 = 0x7ffff7ffd000
 - rip = 0x55555557b8
 - eflags = 0x246
 - eax = 0x0
 - ebx = 0xffffd628
 - ecx = 0x55557d18
 - edx = 0xffffd638
 - esi = 0xffffd628
 - edi = 0x1
 - ebp = 0xffffd3c0
 - esp = 0xffffd2f0
- WATCH
 - open_listenfd() open_listen.c
 - main() main.c 30:1
- CALL STACK
 - open_listenfd() open_listen.c
 - main() main.c 30:1

main.c

```
lib > C open_listen.c > open_listenfd()
1  #include "open_listen.h"
5  #include <netdb.h>
6  #include <unistd.h>
7  #include <sys/types.h>
8  #include <sys/socket.h>
9  #include <string.h>
10
11 int open_listenfd()
12 {
13     struct addrinfo hints;
14     struct addrinfo *list_rp, *rp;
15     char host_buf[HOST_MAXLEN];
16     char service_buf[SERVICE_MAXLEN];
17     int socket_fd;
18
19     int result;
20
21     memset(&hints, 0, sizeof(hints));
22     hints.ai_family = AF_INET;           // IPv4
23     hints.ai_flags = AI_PASSIVE;        // Server
24     hints.ai_socktype = SOCK_STREAM;    // TCP
25
26     result = getaddrinfo(NULL, LISTEN_PORT, &hints, &list_rp);
27     if(result != 0){
28         fprintf(stderr, "[ERROR]-getaddrinfo error: %s\r\n", gai_strerror(result));
29         return -1;
30     }
31
32     /**
33      * getaddrinfo() returns a list of address structures.
34      * Try each address until we successfully bind().
35      * If socket() or bind() fails, we close the socket and try the next address.
36      */
37     for(rp = list_rp; rp != NULL ; rp = rp->ai_next){
38         getnameinfo(rp->ai_addr, rp->ai_addrlen, host_buf, HOST_MAXLEN, service_buf, SERVICE_MAXLEN, 0, 0, 0);
39         printf("IPv4: %s, Port: %s\r\n", host_buf, service_buf);
40
41         socket_fd = socket(rp->ai_family, rp->ai_socktype, rp->ai_protocol);
42         if(socket_fd < 0){
```

Disassembly

Address	Disassembly
0x00005555555777c	call 0x55555557400 <response>
0x000055555557783	90 nop
0x000055555557784	48 8b 45 f8 mov -0x8(%rbp),%rax
0x000055555557788	64 48 2b 04 25 28 00 00 00 sub %fs:0x28,%rax
0x000055555557791	74 05 je 0x55555557798 <request>
0x000055555557793	e8 38 fa ff ff call 0x555555571d0 <_stack_>
0x000055555557798	c9 leave
0x000055555557799	c3 ret
0x00005555555779a	f3 0f 1e fa endbr64
0x00005555555779e	55 push %rbp
0x00005555555779f	48 89 e5 mov %rsp,%rbp
0x0000555555577a2	48 81 ec d0 00 00 00 sub \$0xd0,%rsp
0x0000555555577a9	64 48 8b 04 25 28 00 00 00 mov %fs:0x28,%rax
0x0000555555577b2	48 89 45 f8 mov %rax,-0x8(%rbp)
0x0000555555577b6	31 c0 xor %eax,%eax
0x0000555555577b8	48 8d 85 50 ff ff ff lea -0xb0(%rbp),%rax
0x0000555555577bf	ba 30 00 00 00 mov \$0x30,%edx
0x0000555555577c4	be 00 00 00 00 mov \$0x0,%esi
0x0000555555577c9	48 89 c7 mov %rax,%rdi
0x0000555555577cc	e8 3f fa ff ff call 0x55555557210 <memset@libc>
0x0000555555577d1	c7 85 54 ff ff ff 02 00 00 00 movl \$0x2,-0xac(%rbp)
0x0000555555577db	c7 85 50 ff ff ff 01 00 00 00 movl \$0x1,-0xb0(%rbp)
0x0000555555577e5	c7 85 58 ff ff ff 01 00 00 00 movl \$0x1,-0xa8(%rbp)
0x0000555555577ef	48 8d 95 40 ff ff ff lea -0xc0(%rbp),%rdx
0x0000555555577f6	48 8d 85 50 ff ff ff lea -0xb0(%rbp),%rax
0x0000555555577fd	48 89 d1 mov %rdx,%rcx
0x000055555557800	48 89 c2 mov %rax,%rdx
0x000055555557803	48 8d 05 b6 08 00 00 lea 0x8b6(%rip),%rax
0x00005555555780a	48 89 c6 mov %rax,%rsi
0x00005555555780d	bf 00 00 00 00 mov \$0x0,%edi
0x000055555557812	e8 b9 fa ff ff call 0x555555575d0 <getaddrinfo@libc>
0x000055555557817	89 85 3c ff ff ff mov %eax,-0xc4(%rbp)
0x00005555555781d	83 bd 3c ff ff ff 00 cmpl \$0x0,-0xc4(%rbp)
0x000055555557824	74 38 je 0x5555555785e <open_listenfd@libc>
0x000055555557826	8b 85 3c ff ff ff mov -0xc4(%rbp),%eax
0x00005555555782c	89 c7 mov %eax,%edi
0x00005555555782e	e8 cd f9 ff ff call 0x55555557520 <gai_strerror@libc>
0x000055555557833	48 89 c2 mov %rax,%rdx
0x000055555557836	48 8b 05 e3 27 00 00 mov 0x27e3(%rip),%rax
0x00005555555783d	48 8d 0d 84 08 00 00 lea 0x884(%rip),%rcx
0x000055555557844	48 89 ce mov %rcx,%rsi

Trace

The screenshot displays a debugger window with the following components:

- Top Bar:** Includes a 'RUN A...' button, a 'C Debug' dropdown, and various execution icons (play, step, etc.).
- Left Sidebar:** Contains icons for file explorer, search, and other debugging tools.
- Registers Panel:** Lists CPU registers and their values, such as `rax = 0x7fffffff310`, `rbx = 0x7fffffff628`, and `rcx = 0x55555557d18`.
- Source Code Panel:** Shows the C code for `open_listen.c`. The function `open_listenfd()` is visible, starting with include statements and a `memset` call. A yellow highlight is on line 21: `memset(&hints, 0, sizeof(hints));`.
- Disassembly Panel:** Shows the assembly code corresponding to the source code. The instruction `call 0x55555555210 <memset@libc.so.6>` is highlighted, corresponding to the `memset` call in the source code.



Thank you

