

FIRST EDITION - CHAPTER 7 REV 1

Kevin Thomas Copyright © 2023 My Techno Talent

Forward

This book is not associated or endorsed by the Rust Foundation. This is purely a FREE educational resource going step-by-step on how to build and reverse engineer Rust binaries.

Rust redefines how a binary is compiled. It does not operate with similar patterns like we see in older languages and is simply the most difficult language to reverse engineer in history.

Many of the C-style decompilers within the popular reversing tools are literally rendered useless with Rust.

Malware Authors are going the route of Rust as there are few tools that exist yet to properly statically analyze the compiled binaries which coupled with its speed and memory safety, it is very difficult to understand what it is actually doing.

The aim of this book is to teach basic Rust and step-by-step reverse engineer each simple binary to understand what is going on under the hood.

We will develop within the Windows architecture (Intel x64 CISC) as most malware targets this platform by orders of magnitude.

In later chapters we will within a Raspberry Pi 64-bit ARM OS so that you can get a perspective of what hacking that architecture looks like in Golang at the binary level.

Let's begin...

Table Of Contents

Chapter 1: Hello Rust

Chapter 2: Debugging Hello Rust Chapter 3: Hacking Hello Rust Chapter 4: Scalar Data Types

Chapter 5: Debugging Scalar Data Types Chapter 6: Hacking Scalar Data Types

Chapter 7: Compound Data Types

Chapter 1: Hello Rust

We begin our journey with developing a simple hello world program in Rust on a Windows 64-bit OS.

We will then reverse engineer the binary in IDA Free.

Let's first download Rust for Windows.

https://www.rust-lang.org/tools/install

Let's download IDA Free.

https://hex-rays.com/ida-free/#download

Let's download Visual Studio Code which we will use as our integrated development environment.

https://code.visualstudio.com/

Once installed, let's add the Rust extension within VS Code.

https://marketplace.visualstudio.com/items?itemName=rust-lang.rustanalyzer

It is important to update Rust so I would encourage you to run this before every project.

rustup update

Let's create a new project and get started by following the below steps.

cargo new one_hello

Now let's populate our main.rs file with the following.

```
fn main() {
   println!("Hello, world!");
}
```

Let's open up the terminal by click CTRL+SHIFT+` and type the following.

cargo build

Let's run the binary!

.\target\debug\one_hello.exe

Output...

Hello, world!

Congratulations! You just created your first hello world code in Rust. Time for cake!

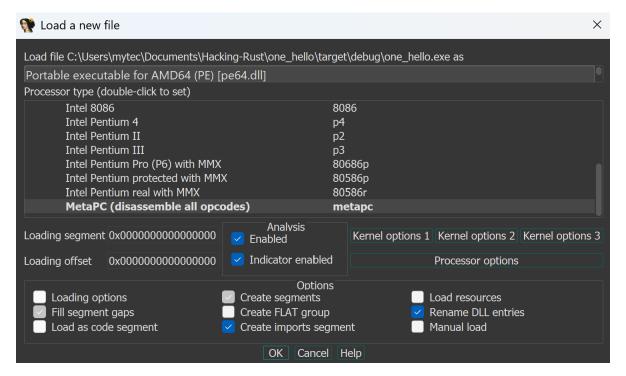
We simply created a hello world style example to get us started.

In our next lesson we will debug this in IDA Free!

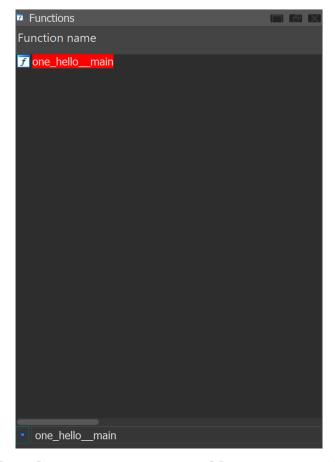
Chapter 2: Debugging Hello Rust

Let's debug our app within IDA Free.

Open IDA Free and we see the load screen. We can keep all the defaults and simply click OK.



In Rust at the assembler level we will need to search for the entry point of our app. This is the *one_hello__main* function. You can use CTRL+F to search.



Now we can double-click on the one_hello__main to launch the focus to this function and graph.

```
one_hello__main proc near
var_38= qword ptr -38h
var_30= byte ptr -30h
sub
      rcx, [rsp+58h+var_30]
rdx, off_14001E3A0; "Hello, world!\n"
lea
mov
      r9, aInvalidArgs ; "invalid args"
lea
      xor
mov
call
lea
      _ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5
call
nop
add
retn
one_hello__main_endp
```

We can see in the box our "Hello, world!\n" text.

If we double-click on $off_14001E3A0$ it will take us to a new window where the string lives within the binary.

```
    .rdata:000000014001E3A0 off_14001E3A0 dq offset aHelloWorld ; DATA XREF: one_hello__main+91o .rdata:000000014001E3A0 ; "Hello, world!\n"
    .rdata:000000014001E3A8 db 0Eh
```

Similar to Go, we can see the strings are in a string pool however we do see a OAh, O so it contains the newline and null terminator as Go handles it slightly different.

These lessons are designed to be short and digestible so that you can code and hack along.

In our next lesson we will learn how to hack this string and force the binary to print something else to the terminal of our choosing.

This will give us the first taste on hacking Rust!

Chapter 3: Hacking Hello Rust

Let's hack our app with IDA Free.

Let's load up IDA and revisit the binary.

```
one_hello__main proc near
var_38= qword ptr -38h
var_30= byte ptr -30h
sub
lea
        rdx, off_14001E3AO; "Hello, world!\n"
lea
mov
        r9, aInvalidArgs ; "invalid args"
lea
xor
mov
        _ZN4core3fmt9Arguments6new_v117ha9b71491ca997d8aE ; core::fmt::Arguments::new_v1::ha9b71491ca997d8a
call
lea |
        _ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5
call
nop
add
retn
one_hello__main_endp
```

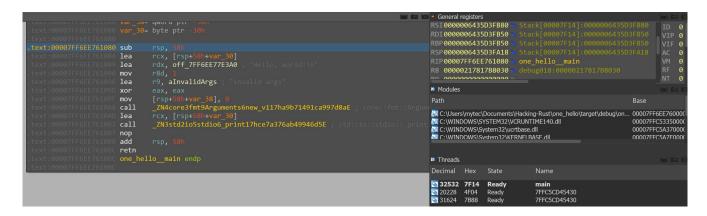
We start off with subtracting 0x58 bytes from RSP. We do not see the C-style prologue here. What we are seeing is what the Rust compiler chose to do at compile time.

This is a VERY simple example but I want to take the time to step through this step-by-step.

Let's set a breakpoint at the entry.

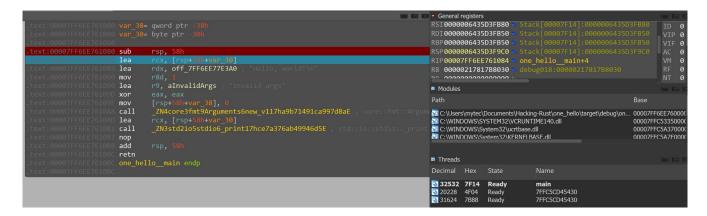
```
one_hello__main proc near
sub
lea
                                                      rdx, off_14001E3A0; "Hello, world!\n"
lea
                                                    r8d, 1
r9, aInvalidArgs; "invalid args"
 mov
lea
xor
mov
call
lea
                                                      \underline{\textbf{ZN4core3fmt9Arguments6new\_v117ha9b71491ca997d8aE}} \ ; \ core::fmt::Arguments::new\_v1::ha9b71491ca997d8aE} \ ; \ core::fmt::Arguments::ha9b71491ca997d8aE} \ ; \ core::fmt::ha9b71491ca
                                                    rcx, [rsp+58h+var_30]
_ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5
call
nop
add
retn
  one_hello__main endp
```

Lets run.



Lets take note where RSP is 0x0000006435d3fa18.

Lets step once.



We now see the value of RSP at 0x0000006435df9c0. This makes sense as 0x0000006435d3fa18 - 0x58 = 0x0000006435d3f9c0.

Next we see lea rcx, [rsp+58h+var_30] so let's break this down.

LEA stands for "Load Effective Address." It is a versatile instruction that can be used to perform arithmetic calculations on addresses without actually accessing memory.

RCX is the destination register where the effective address will be stored.

[rsp+58h+var_30] is the source operand, representing the address that needs to be calculated. It consists of multiple parts:

- * RSP is the stack pointer register.
- * 58h is a hexadecimal constant value, which represents an offset from the stack pointer.
- * var_30 refers to a variable or memory location. In this case, it is -30h, indicating a byte-sized variable or memory location located at an offset of -30h (or -48 in decimal) from the stack pointer.

Therefore, lea rcx, [rsp+58h+var_30] calculates the effective address by adding the stack pointer (RSP) with an offset of 58h and the variable var_30 located at an offset of -30h from the stack pointer. The resulting effective address is stored in the RCX register.

Next we see an offset value being loaded into RDX.

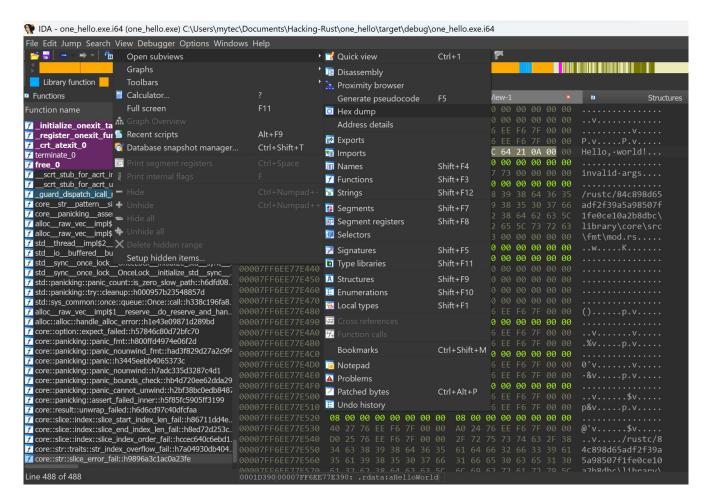
```
FF6EF761080 one hello main proc near
text:00007FF6EE761080 sub
                                          rsp, 58
                                          rcx, [rsp+58h+var_30]
    t:00007FF6EE761084 lea
                                           rdx, off_7FF6EE77E3A0
                               mov
                                           r9, aInvalidArgs; "invalid args"
                               lea
                                          [rsp+58h+var_38], 0
    ZN4core3fmt9Arguments6new_v117ha9b71491ca997d8aE ; core::fmt::Argu
rcx, [rsp+58h+var_30]
    ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print17hce7a376ab49946d5E ;
                               mov
                               call
        0007FF6EE7610AD lea
                              call
                               nop
                               add
                               retn
                               one_hello__main endp
```

This is our string. Let's double-click on the offset.

```
    .rdata:00007FF6EE77E3A0 off_7FF6EE77E3A0 dq offset aHelloWorld ; DATA XREF: one_hello__main+9îo
    .rdata:00007FF6EE77E3A0 ; "Hello, world!\n"
    .rdata:00007FF6EE77E3A8 db 0Eh
```

We see 13 chars and the null terminator therefore 0x0e.

Let's stop the debug and click on the "Hello, world!\n" text and view the Hex View-1.



Here we see the hex view.

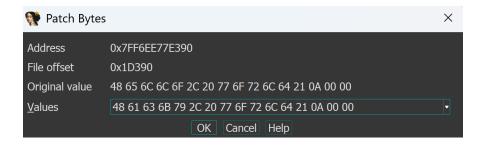
```
00007FF6EE77E390 48 65 6C 6C 6F 2C 20 77 6F 72 6C 64 21 0A 00 00 Hello, world!...
00007FF6EE77E3A0 90 E3 77 EE F6 7F 00 00 0E 00 00 00 00 00 00 00 .....
```

Click edit - Patch program - Change byte...



Let's hack our binary to say "Hacky, world!\n" instead!

```
'H' -> 0x48
'a' -> 0x61
'c' -> 0x63
'k' -> 0x68
'y' -> 0x79
',' -> 0x20
',' -> 0x20
'w' -> 0x77
'o' -> 0x6F
'r' -> 0x6C
'd' -> 0x64
'!' -> 0x21
```



Click OK.

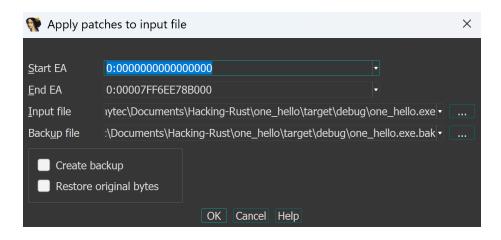
```
      00007FF6EE77E390
      48 61 63 6B 79 2C 20 77 6F 72 6C 64 21 0A 00 00 Hacky, world!...

      00007FF6EE77E3A0
      90 E3 77 EE F6 7F 00 00 0E 00 00 00 00 00 00 00 00 ......
```

SUCCESS!

Press the red stop button to stop debugging.

Click edit - Patch program - Apply patches to input file...



Let's run again...

```
.text:00007FF628C21080 var_38= qword ptr -38h
.text:00007FF628C21080 var_30= byte ptr -30h

.text:00007FF628C21080 sub rsp, 58h
.text:00007FF628C21084 lea rcx, [rsp+58h+var_30]

.text:00007FF628C21089 lea rdx, off_7FF628C3E3A0 ; "Hacky, world[\n"

.text:00007FF628C21090 mov r8d, 1

.text:00007FF628C21090 mov r8d, 1

.text:00007FF628C21090 xor eax, eax

.text:00007FF628C21090 mov [rsp+58h+var_38], 0

.text:00007FF628C21090 mov [rsp+58h+var_38], 0

.text:00007FF628C21082 call _ZM4core3fmt9Arguments6new_v117ha9b71491ca997d8aE ; core::fmt::Arguments::new_v1::ha9b71491ca997d8a

.text:00007FF628C21082 call _ZM3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5

.text:00007FF628C21088 add rsp, 58h

.text:00007FF628C2108C one_hello_main endp
```

Let's F8, step over, until the NOP.

Observe the command window.



SUCCESS! We have successfully hacked our first Rust binary!

The rest of the Assembler is trivial as we are just calling std::io::stdio::_print to echo the line to STDOUT.

Stay tuned as this will get significantly more challenging. With all things we start small and build!

In our next chapter we will discuss the Scalar Data Types.

Chapter 4: Scalar Data Types

We continue our journey with a scalar data types example program in Rust on a Windows 64-bit OS.

Let's create a new project and get started by following the below steps.

cargo new two_scalar-data-types

Now let's populate our main.rs file with the following.

```
fn main() {
    let my_integer: i32 = 42;
    println!("integer: {}", my_integer);

    let my_float: f64 = 3.14;
    println!("float: {}", my_float);

    let my_char: char = 'A';
    println!("character: {}", my_char);
}
```

Let's open up the terminal by click CTRL+SHIFT+` and type the following.

cargo build

Let's run the binary!

.\target\debug\two_scalar-data-types.exe

Output...

integer: 42
float: 3.14
character: A

Congratulations! You just created another program in Rust. Time for cake!

We simply created an example of each of the scalar data types which are int, float and char.

In our next lesson we will debug this in IDA Free!

Chapter 5: Debugging Scalar Data Types

Let's debug our app within IDA Free.

Open IDA Free and we see the load screen. We can keep all the defaults and simply click OK.

On the left-hand side we see the Function name window and we want to search for two_scalar_data_types__main. Double-click it and we will see the function.

```
<u>u</u> 🚄 🚟
 two_scalar_data_types__main proc near
 /ar_E8= qword ptr -0E8h
  result= core::fmt::ArgumentV1 ptr -0D4h
  var_A0= qword ptr -0A0h
var_98= qword ptr -98h
  ar_90= core::fmt::ArgumentV1 ptr -90h
  var_58= qword ptr -58h
var_50= qword ptr -50h
  ar_44= core::fmt::ArgumentV1 ptr -44h
  ar_10= qword ptr -10
 /ar_8= qword ptr -8
sub
sub    rsp, 108h
mov    dword ptr [rsp+108h+result.value], 2Ah ; '*'
lea    rcx, [rsp+108h+result] ; result
call    _ZN4core3fmt10ArgumentV111new_display17hda29a79969d0b5cbE ; core::fmt::ArgumentV1::new_display::hda29a79969d0b5cb
mov    [rsp+108h+var_A0], rax
mov    [rsp+108h+var_98], rdx
lea    rcx, [rsp+108h+result.value+4]
lea    rdx, off_140023430 ; "integer: \n"
mov
lea
           [rsp+108h+var_E8], 1
_ZN4core3fmt9Arguments6new_v117h8807bc63cce0ab3bE ; core::fmt::Arguments::new_v1::h8807bc63cce0ab3b
rcx, [rsp+108h+result.value+4]
mov
call
```

```
ZN4core3fmt10ArgumentV111new_display17h0ac0f3d1bf4c49e3E ; core::fmt::ArgumentV1::new_display::h0ac0f3d1bf4
call
        mov
mov
lea
lea
mov
lea
call
         _ZN4core3fmt9Arguments6new_v117h8807bc63cce0ab3bE ; core::fmt::Arguments::new_v1::h8807bc63cce0ab3b
        rcx, [rsp+108h+var_90.formatter]
_ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5
lea
call
        dword ptr [rsp+108h+var_44.value], 41h; 'Arcx, [rsp+108h+var_44]; result
mov
        _ZN4core3fmt10ArgumentV111new_display17h2cdf30f7c1587727E ; core::fmt::ArgumentV1::new_display::h2cdf30f7c1587727

[rsp+108h+var_10], rax

[rsp+108h+var_8], rdx

rcx, [rsp+108h+var_4.value+4]
call
mov
lea
        rdx, off_140023488 ; "character:
lea
mov
lea
        r9, [rsp+108h+var_10]
mov
call
         _ZN4core3fmt9Arguments6new_v117h8807bc63cce0ab3bE ; core::fmt::Arguments::new_v1::h8807bc63cce0ab3b
lea
        _ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5
call
nop
add
retn
two_scalar_data_types__main endp
```

We see the below instruction.

```
mov dword ptr [rsp+108h+result.value], 2Ah; '*'
```

The value of 2Ah is 42 so we know that the literal value is being stored inside RSP + the offset value above.

We then see a call to the println! macro.

```
call ZN3std2io5stdio6 print17hce7a376ab49946d5E; std::io::stdio:: print::hce7a376ab49946d5
```

Bare with me as I explain the floating point number. Let's first review the actual code we are going to statically reverse.

```
let my_float: f64 = 3.14;
println!("float: {}", my_float);
```

The first instruction loads the floating-point value 3.14 from memory into the XMM0 register. The XMM registers are 128-bit registers that are used to store floating-point values and vectors.

movsd xmm0, cs:__real@40091eb851eb851f

The next instruction stores the value in XMMO to the memory location rsp+108h+var_90.value. This memory location is part of the stack frame that is allocated for the *println!()* function call.

movsd [rsp+108h+var_90.value], xmm0

The next instruction sets the register rcx to point to the memory location $rsp+108h+var_90$. This memory location contains the *ArgumentV1* struct that will be passed to the new_display function.

lea rcx, [rsp+108h+var 90]; result

The next instruction calls the new_display function. The new_display function is responsible for converting the ArgumentV1 struct into a string representation.

call ZN4core3fmt10ArgumentV111new display17h0ac0f3d1bf4c49e3E; core::fmt::ArgumentV1::new display::h0ac0f3d1bf4c49e3E

The new_display function returns the string representation in the RAX and RDX registers. The next instruction stores the value in RAX to the memory location rsp+108h+var_58. This memory location contains the Formatter struct that will be passed to the println!() function.

mov [rsp+108h+var 58], rax

The next instruction stores the value in RDX to the memory location $rsp+108h+var_50$. This memory location contains the slice of bytes that represents the string representation of the ArgumentV1 struct.

mov [rsp+108h+var 50], rdx

The next instruction sets the register RCX to point to the memory location $rsp+108h+var_90.formatter$. This memory location contains the Formatter struct that will be passed to the println!() function.

lea rcx, [rsp+108h+var_90.formatter]

The next instruction sets the register RDX to point to the string literal "float: ".

lea rdx, off 140023458; "float: "

Finally we see the *println* macro print the value.

call ZN3std2io5stdio6 print17hce7a376ab49946d5E ; std::io::stdio:: print::hce7a376ab49946d5

The last item was the char 'A' which is simply moved into RSP and an offset below.

mov dword ptr [rsp+108h+var 44.value], 41h ; 'A'

Then we print.

call _ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5

In our next lesson we will hack this!

Chapter 6: Hacking Scalar Data Types

Let's hack our app within IDA Free.

Open IDA Free and we see the load screen. We can keep all the defaults and simply click OK.

On the left-hand side we see the Function name window and we want to search for two_scalar_data_types__main. Double-click it and we will see the function.

```
<u>I</u> 🗹 📴
two_scalar_data_types__main proc near
  esult= core::fmt::ArgumentV1 ptr -0D4h
 var_AO= qword ptr -0AOh
var_98= qword ptr -98h
var_90= core::fmt::ArgumentV1 ptr -90h
 ar_58= qword ptr -58h
ar_50= qword ptr -50h
 ar 44= core::fmt::ArgumentV1 ptr -44h
 var_10= qword ptr -10
var_8= qword ptr -8
sub rsp, 108h
mov dword ptr [rsp+108h+result.value], 2Ah; '*'
       rcx, [rsp+108h+result]; result
          _ZN4core3fmt10ArgumentV111new_display17hda29a79969d0b5cbE; core::fmt::ArgumentV1::new_display::hda29a79969d0b5cb

[rsp+108h+var_A0], rax

[rsp+108h+var_98], rdx

rcx, [rsp+108h+result_value+4]
call
mov
mov
lea
lea
          rdx, off_140023430; "integer: \n"
mov
lea
          [rsp+108h+var_E8], 1
_ZN4core3fmt9Arguments6new_v117h8807bc63cce0ab3bE ; core::fmt::Arguments::new_v1::h8807bc63cce0ab3b
rcx, [rsp+108h+result.value+4]
mov
call
```

```
call
         ZN4core3fmt10ArgumentV111new_display17h0ac0f3d1bf4c49e3E ; core::fmt::ArgumentV1::new_display::h0ac0f3d1bf4c
        [rsp+108h+var_58], rax
[rsp+108h+var_50], rdx
rcx, [rsp+108h+var_90.formatter]
rdx, off_140023458; "float: "
mov
mov
lea
lea
mov
lea
call
        _ZN4core3fmt9Arguments6new_v117h8807bc63cce0ab3bE ; core::fmt::Arguments::new_v1::h8807bc63cce0ab3b
        rcx, [rsp+108h+var_90.formatter]
_ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5
lea
call
        dword ptr [rsp+108h+var_44.value], 41h; 'Arcx, [rsp+108h+var_44]; result
mov
        lea
call
mov
mov
lea
        rdx, off_140023488 ; "character: "
lea
mov
lea
        r9, [rsp+108h+var_10]
mov
        _ZN4core3fmt9Arguments6new_v117h8807bc63cce0ab3bE ; core::fmt::Arguments::new_v1::h8807bc63cce0ab3b rcx, [rsp+108h+var_44.value+4]
call
lea
        _ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5
call
nop
add
retn
two_scalar_data_types__main endp
```

We see the below instruction.

```
mov dword ptr [rsp+108h+result.value], 2Ah ; '*'
```

The value of 2Ah is 42 so we know that the literal value is being stored inside RSP + the offset value above.

Let's set a breakpoint by highlighting the line and pressing F2.

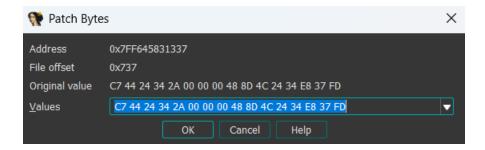
```
mov dword ptr [rsp+108h+result.value], 2Ah; '*'
```

Press play to debug.

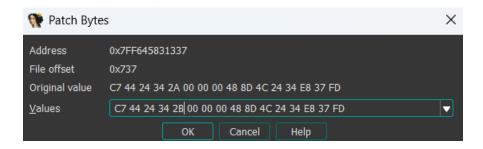
We now see the our breakpoint triggered.

```
.text:00007FF645831337 mov dword ptr [rsp+108h+result.value], 2Ah
```

Click edit - Patch program - Change byte...



Change the 2A to 2B.

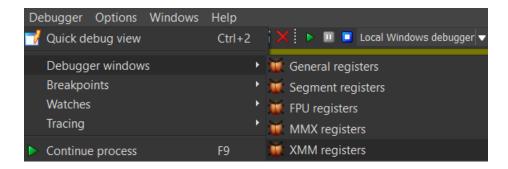


Click ok. We have now hacked our 42 value to 43! Let's put a breakpoint on the below line.



Lets press F9 to run to it.

Click on Debugger - Debugger windows - XMM registers



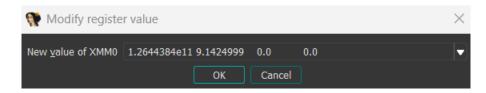
It will load the XMM registers.



Double-click on the 2.1424999.



Change the value to 9 instead of 2.



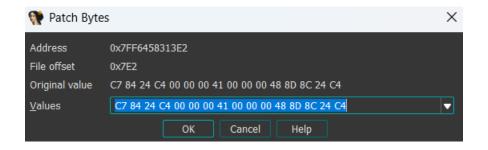
Press OK.

We have successfully hacked 3.14 to 299499.58!

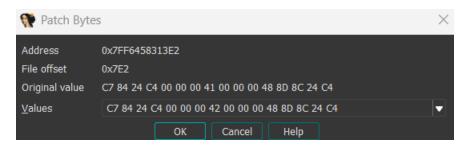
Put a breakpoint on the below line and press F9.



Click edit - Patch program - Change byte...



Change 41 to 42.



Press the red stop button to stop debugging.

Click edit - Patch program - Apply patches to input file...

Place a breakpoint on the NOP under the final call to println.

```
call __ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5
nop
```

BOOM!



In our next lesson we will handle Compound Data Types.

Chapter 7: Compound Data Types

We continue our journey with a compound data types example program in Rust on a Windows 64-bit OS.

Let's create a new project and get started by following the below steps.

cargo new three_compound-data-types

Now let's populate our main.rs file with the following.

```
fn main() {
    let tup = (1337, 3.14, 42);
    let (x, y, z) = tup;
    println!("x: {x}");
    println!("y: {y}");
    println!("z: {z}");

let x = [1, 2, 3];
    let one = x[0];
    let two = x[1];
    let three = x[2];
    println!("one: {one}");
    println!("two: {two}");
    println!("three: {three}");
}
```

Let's open up the terminal by click CTRL+SHIFT+` and type the following.

cargo build

Let's run the binary!

.\target\debug\three_compound-data-types.exe

Output...

```
x: 1337
y: 3.14
z: 42
one: 1
two: 2
three: 3
```

Congratulations! You just created another program in Rust. Time for cake!

We simply created an example of each of the compound data types which are tuple and array.

In our next lesson we will debug this in IDA Free!