



VETERANSEC

A Veteran Cyber Security Community

Hacktober CTF 2018 – Binary Analysis – Larry

A basic reverse engineering challenge for a CTF and a mini intro to RE.

By [emtuls](#) in [Binary Analysis](#), [CTF Write-ups](#), [Exploit Development](#), [Reverse Engineering](#) on [October 19, 2018](#)

 6 comments

Hey Everyone! This past week, a couple of us from VetSec participated in the [CTF](#) called [Hacktober](#). I would definitely say that it was a very good CTF for people who are beginner to mediocre at anything related to CTF's, so if you missed out and get a chance to try it next year, give it a shot!

SEARCH

CATEGORIES

- [Binary Analysis](#)
- [Computer Science](#)
- [CTF Write-ups](#)
- [eLearnSecurity](#)
- [Exploit Development](#)
- [Featured](#)
- [Finding a Job](#)
- [Hack The Box Write-ups](#)
- [Hacking Live Streams](#)
- [Reverse Engineering](#)
- [Reviews](#)
- [SANS](#)
- [Steganography](#)



Most of the challenges were very doable for people who were new with no experience, as well as a few challenges that would stump some seasoned players. It involved a very wide variety of challenges, such as the typical Forensics, Steganography, SQL, Binary Analysis, Web Exploitation,

- Uncategorized
- Vulnerability Lab Setup
- Web-Application Hacking
- Wireless Hacking

SUBSCRIBE

Enter your email address to follow this blog and receive notifications of new posts by email.

Join 1,397 other followers

Enter your email address

FOLLOW

AUTHORS



- **Tritonal6**
 - Know thy enemy; part 1/2

Trivia type challenges, as well as a few other uncommon types.

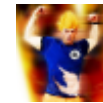
This post will focus on one of the Binary Analysis challenges that I found to be the trickiest of the bunch and it could take some time if attempted solely via **Static Analysis** without a combination of **Dynamic Analysis**. Though there are surely easier ways to solve this challenge, I figured I would turn this into a little bit of an intro to Reverse Engineering of software, since it would be more valuable to learn these concepts going forward.

Takeaways:

1. Basic **Static Analysis** Techniques for Binary Analysis.
2. Basic **Dynamic Analysis** Techniques for Binary Analysis.

Tools to have:

1. A **disassembler** with **x64** (64-bit) and **ELF** capability for **Static Analysis**. I recommend **IDA Pro Free** (free version for this, since the **paid** is **\$\$\$\$\$\$\$**). Another good option that I recommend would be **Binary Ninja** (**\$150**), though you



Cerkoryn

- [Review: SANS Cyber FastTrack 2019](#)



emtuls

- [Vetsec Becoming a Non-Profit!](#)
- [x86 Exploit Development Pt 2 – ELF Files and Memory Segmentation](#)
- [Getting Started Guide for VetSec Wargame Exploit Development Tutorials](#)
- [x86 Exploit Development Pt 1 – Intro to Computer Organization and x86 Instruction Set Architecture Fundamentals](#)
- [Creating VetSecs Wargame Pt. 3: Finishing The Intro Challenges and Reshaping the Makefile](#)

would need the paid version for this challenge, the demo one only supports 32-bit **x86**.

2. A **debugger** for **Dynamic Analysis**. I used **GDB**, a Linux debugger that comes standard on pretty much all distributions of Linux. I chose this one due to familiarity, though I'm sure other debuggers would work just fine.
3. If you plan on following along, you can download the challenge from my **Github**, [here](#).

Disclaimer:

I'll try not to make this a course on **Computer Architecture**, since that would be way too in depth for this challenge, but it would be good to have a light background in how a computers processor functions, but not necessary. Again, this will not be too in depth, just a high level of the useful parts you will need to know for the challenge. I am by no means a professional at any of this, just merely a hobbyist/beginner who is learning as I go and probably wrong about a lot of things.

Where to start:



m4v3r1ck

- **Zero to Hero: Week 9 – NTLM Relay, Token Impersonation, Pass the Hash, PsExec, and more**
- **A Day in the Life of an Ethical Hacker / Penetration Tester**
- **Zero to Hero Pentesting: Episode 8 – Building an AD Lab, LLMNR Poisoning, and NTLMv2 Cracking with Hashcat**
- **Zero to Hero Pentesting: Episode 7 – Exploitation, Shells, and Some Credential Stuffing**
- **Introductory Exploit Development Live Stream – x86 Assembly Primer and SEH Overflows w/ Ruri**



Jace Powell

For most **Binary Reverse Engineering** challenges, I start on Linux and run **'strings'** just to see what hints I might get, which in this case showed a string called **'flag{thisNOTtheflag}'** and another **'flag{NOTtheflag}'**. Of course, I tried these and they were surely *NOT* the flag. I then typically run **'file'** against it, to see what it thinks the file type is, which gave me the following output:

```
Larry.out: ELF 64-bit LSB shared object, x86-64, version :
```

From here, you can see that this is a 64-bit ELF file, with x86-64 architecture. **ELF** is the standard executable file type for Linux, analogous to an **.exe** file for Windows. With this information, I select a disassembler that I have on hand that might help assist me with the binary analysis. A **disassembler** is a tool that will take our compiled **binary** (ELF in this case) and convert it from the 1's and 0's our computer reads, to something that is more human readable. This readable format is known as **assembly language**, and in particular, we will be dealing with the **x86-64 assembly language** in the **Intel syntax**, versus AT&T syntax, for ease of readability.

- [A Veteran's Guide to Making a Career Jump to Information Security](#)



reubadoob

- [“...because I stood on the shoulders of giants”](#)
- [A Year Ago My Life Changed, From Soldier to Cyber](#)



syphon51773

- [Review: SANS VetSuccess Academy](#)

- [RSS – Posts](#)

Now, this is where it could get very confusing, since assembly language is not exactly straight forward. I do not intend for this to be an extensive x86 assembly language tutorial/course, so I will again, limit this to only the information required to solve this challenge.

x86 Instruction Set Architecture:

Reference: [x86 Assembly Tutorial](#)

Some background on [computer architecture](#) that would be useful to know, at a high level, so bear with me:

In a processor, things are all dealt with inside of memory and passed around via something called a '[register](#)' and also sometimes pushed onto a '[stack](#)'. Each processor has their own set of **general purpose registers** as well as proprietary registers that are highly specific, but we will only worry about the main set of general purpose registers, [specifically those in x86](#). These registers are as follows:

x86 Registers:

Resource: [X86_Architecture](#) and [this](#)

32 bit will start with E, which stands for **Extended** (it **extended 16 bit registers**, which follow the same convention, just **drop the E or R at the front**, i.e., AX, BX, CX, DX, SP, BP, etc)

64 bit will start with R instead of E, which has no historical significance. (RAX, RBX, RCX, RDX, RSI, RDI, etc)

NOTE: For this, I will use mainly 32 bit registers, since this challenge mainly used 32 bit registers, though there are a couple of 64 bit registers used, which are denoted with R in the front.

RAX/EAX (64/32 bit Accumulator register): Used in arithmetic operations and also for **return values** from function calls in certain **calling conventions**, such as **cdecl** (**this** might be useful to help understand a bit more), a common convention in x86.

RBX/EBX (64/32 bit Base register): Sometimes used as a pointer to data. No specific uses, but is often set to a commonly used value (such as 0) throughout a function to speed up calculations

RCX/ECX (64/32 bit Counter register): Used in shift/rotate instructions and loops as a loop counter (like `i` in `for` loops).

RDX/EDX (64/32 bit Data register): Used in arithmetic operations and **I/O operations**, and generally used for storing short-term variables within a function.

RSI/ESI (RSI is 64bit, ESI is 32 bit Source Index register): Used as a **pointer** to a source in stream operations (like manipulating **strings**, which are **pointers to char arrays**).

RDI/EDI (RDI is 64bit, EDI is 32 bit Destination Index register): Used as a pointer to a destination in stream operations. (also for manipulating strings, see above register).

EFLAGS – a 32-bit register used as a collection of bits representing Boolean values to store the results of operations and the state of the processor

Memory:

Although it is not needed for this challenge in particular, it would be valuable to know how memory works and the

organization of it, such as **little endian** and **big endian**. In our case, x86 uses **little endian format**.

x86 Instruction Layout:

When dealing with x86 instruction with the **Intel syntax**, instructions are usually written in the form:

```
instruction destination, source
```

for example:

```
mov eax, [ecx]
```

which means, move what is pointed to by the address in the register **ECX** and put it into the register **EAX**. If something is in **brackets** (like how **ECX** is), we use the value that is within these brackets **as a pointer to a memory address** and use the **value** that is **pointed to by this address**. We won't worry about this much right now, since this is not an x86 assembly language course, this was just an example, but if

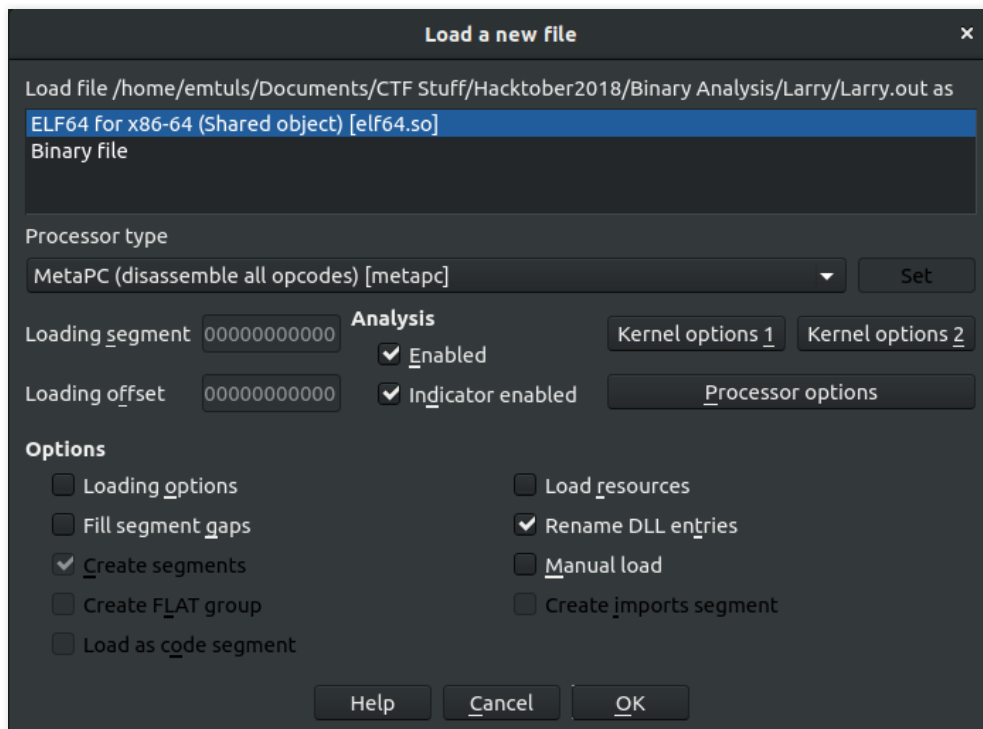
you would like to understand it better, **this** might help. This format is almost always the case for Intel syntax, unless stated otherwise.

Static Analysis: IDA Pro (or your disassembler of choice):



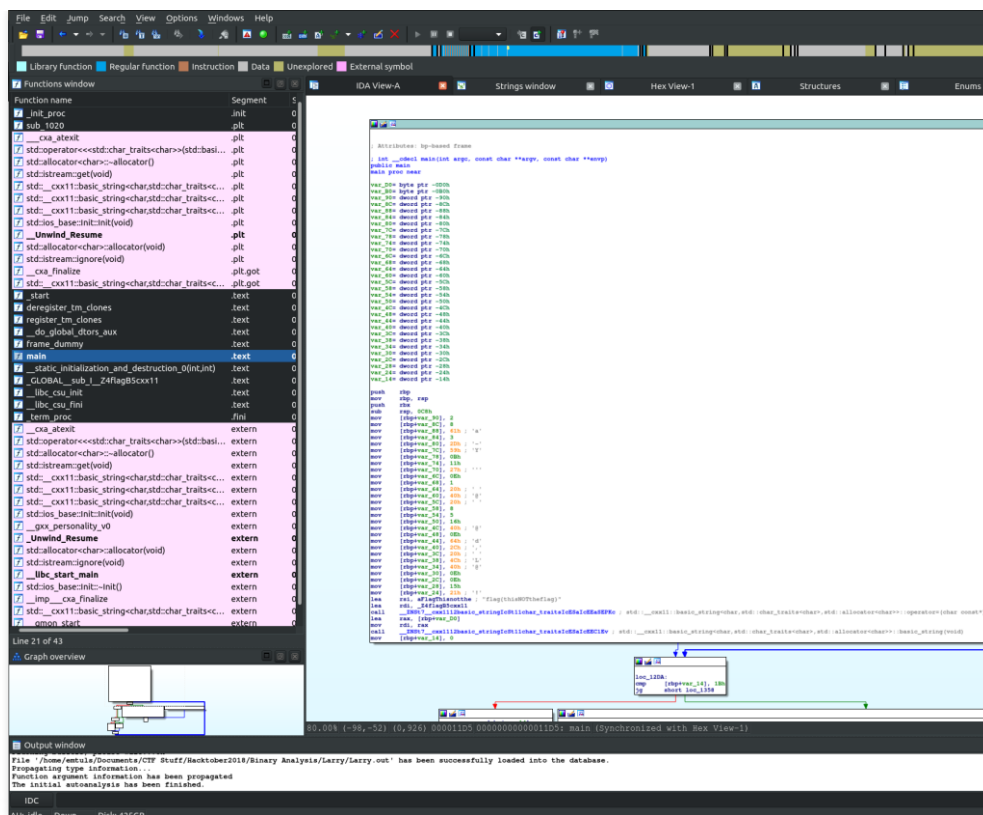
IDA is by far the industry standard when it comes to reverse engineering software. It has a plethora of capabilities and scripts combined with one of the best **decompilers** (sold separately) to date. A **decompiler** attempts to take our compiled binary and turn it into **source code** (what a developer/programmer would write), such as C programming, for instance. They are not perfect, but they can do a great job at assisting with reverse engineering. Though, for this challenge, we will stick with all free tools, so we will just be using the **IDA Pro Free version**, which is just a **disassembler**.

When we first open the file up in **IDA**, we are presented with a screen (shown in the image below) that asks us how to load the file, as in, which file type. IDA is pretty smart and it already recognizes that the file we provided it was an **ELF binary** and is **64 bit**, so we will keep that selected. It then asks us which processor type we intend to use to decode the instructions in this disassembly. I recommend leaving this in **MetaPC**, since IDA is usually smart enough to determine this correctly.



Options when opening file in IDA Pro

After hitting 'OK', IDA attempts to disassemble the program and leaves with a large jumbled mess of instructions and inside of the 'main' function. This isn't an IDA tutorial, so I won't go too in depth as to how to actually use IDA to the full effect. Though if you are interested, you can definitely check out the book on it called, [The IDA Pro Book](#). What you should see on your screen, should look similar to what is in the picture below.



Opening IDA on Larry.out

Getting into the disassembly:

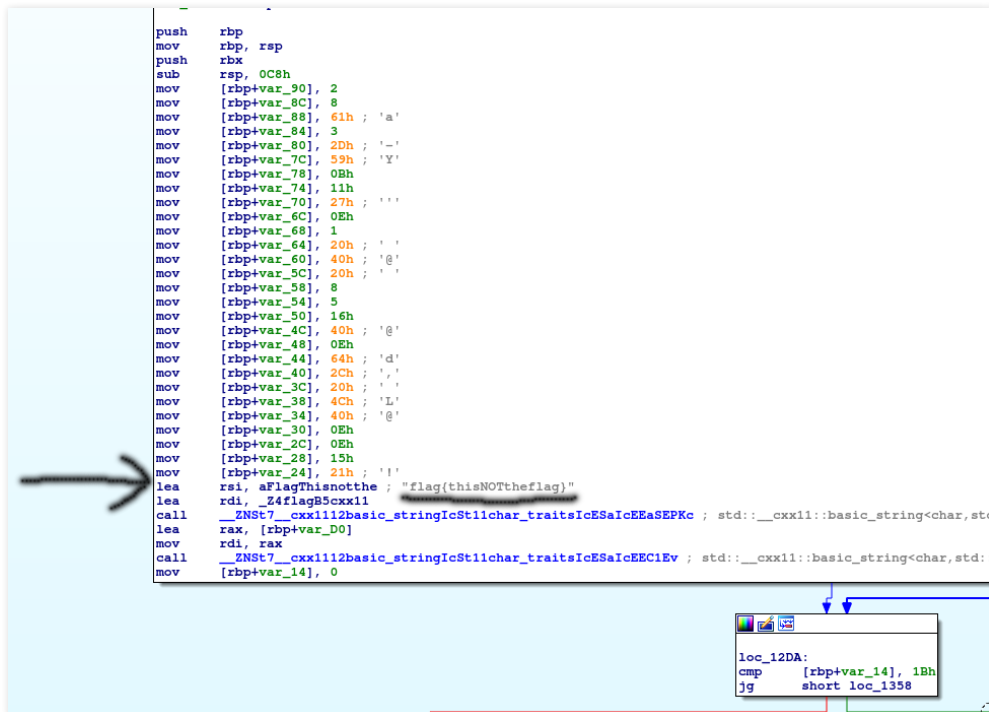
The first screen we start in is the graph view, it allows us to see the flow of the program in a visual perspective, which I find very useful. To navigate on this screen, you can click and hold somewhere outside of one of the boxes and

move the mouse around to shift the screen. You can also scroll up and down, or scroll in or out to view the bigger picture if you would like.

Focus! Don't get too distracted:

Now, the main focus for reverse engineering is to not try and get into the fine details of every last line of assembly in the program (unless that is your actual intention of course), but rather, to find the details that matter to you. For our case, we are trying to complete a challenge, so we would likely want to find a function that handles/returns a **'flag'**, which is sometimes a function labeled as **'win'**, but not for this particular challenge. So if you remember, when I ran the **'strings'** command against the file, I saw a string that said **'flag{thisNOTtheflag}'** and another that said **'flag{NOTtheflag}'**. We can confirm this by also running strings via **IDA Pro**, which can be found in **View -> Open subviews -> Strings** or **'Shift + F12'** as a shortcut. With this in mind, and an attempt to find something that might give me any idea of a **'flag'**, I began searching. What stuck out to me almost immediately was the string we saw before being stored into a variable. IDA is able to determine that the address in the variable **'aFlagThisnotthe'** points to an

ASCII string which comes to be 'flag{thisNOTtheflag}', and displays it for us as a comment automatically, as seen below.



```
push    rbp
mov     rbp, rsp
push    rbx
sub     rsp, 0C8h
mov     [rbp+var_90], 2
mov     [rbp+var_8C], 8
mov     [rbp+var_88], 61h ; 'a'
mov     [rbp+var_84], 3
mov     [rbp+var_80], 2Dh ; '-'
mov     [rbp+var_7C], 59h ; 'Y'
mov     [rbp+var_78], 0Bh
mov     [rbp+var_74], 11h
mov     [rbp+var_70], 27h ; ' '
mov     [rbp+var_6C], 0Eh
mov     [rbp+var_68], 1
mov     [rbp+var_64], 20h ; ' '
mov     [rbp+var_60], 40h ; '@'
mov     [rbp+var_5C], 20h ; ' '
mov     [rbp+var_58], 8
mov     [rbp+var_54], 5
mov     [rbp+var_50], 16h
mov     [rbp+var_4C], 40h ; '@'
mov     [rbp+var_48], 0Eh
mov     [rbp+var_44], 64h ; 'd'
mov     [rbp+var_40], 2Ch ; ' '
mov     [rbp+var_3C], 20h ; ' '
mov     [rbp+var_38], 4Ch ; 'L'
mov     [rbp+var_34], 40h ; '@'
mov     [rbp+var_30], 0Eh
mov     [rbp+var_2C], 0Eh
mov     [rbp+var_28], 15h
mov     [rbp+var_24], 21h ; ' '
lea     rsi, aFlagThisIsNotThe ; "flag{thisNOTtheflag}"
lea     rdi, _Z4flagB5cxx11
call    _ZNSt7_cxx112basic_stringIcSt11char_traitsIcESaIcEEaSEPKc ; std::__cxx11::basic_string<char,std
lea     rax, [rbp+var_D0]
mov     rdi, rax
call    _ZNSt7_cxx112basic_stringIcSt11char_traitsIcESaIcEEC1Ev ; std::__cxx11::basic_string<char,std:
mov     [rbp+var_14], 0

loc_12DA:
cmp     [rbp+var_14], 1Bh
jg      short loc_1358
```

Flag String in IDA

Now, for x86 assembly, the 'lea' instruction means '**load effective address**'. What this does is **store** the address of the **source operand** and put it into the **destination operand**, which if you recall from before, the format is:

instruction destination, source.

Function calls:

From the picture above, at the bottom right, we see `'lea rsi, aFlagThisnotthe'`, which means that we load the address of the variable `'aFlagThisnotthe'` and store it into the `'RSI'` register. Then it looks like we store another variable's address into the `'RDI'` register, from looking at the next instruction, which then leads to a `call` instruction. A `'call'` instruction means that we will be **calling a function** with the arguments given, and how those arguments are set up is based on the type of function being called and it's **calling convention**. So with this knowledge, it looks to be **setting up arguments for a function** to work on them, which I originally had assumed would be a function that manipulated the string in memory. Though this assumption was **incorrect**, I will give a little more detail into this. We know the address for the string, `'flag{thisNOTtheflag}'`, was loaded into `RSI` via an `'lea'` instruction, and another address was loaded into `RDI`, thus, armed with this knowledge, we can assume that the `'call'` instruction located after these `'lea'` instructions, operates on these strings in some manner. It likely is some

sort of memory allocation function for a new string. I won't go into any more detail on that, since again, it is not important for this challenge, and I already determined that this function is not the one we're worried about in the next few discoveries.

After that first function call, there is another set up of some function, which also gets called, but we won't worry about that for now, since it looks like the more interesting stuff comes after it. At the end of the box, we see a `0` being loaded into where `'[rbp+var_14]'` is **pointing** to (recall that if something is in **brackets**, this means we are **not worried about the value** which is evaluated by what is in the brackets, **but rather what that value as an address, points to in memory**), but this was only significant to me after I had looked ahead at the **next** box, which showed that **something was being compared to what was at at that location**, causing me to look back at what might be there. My assumption at this point is that it could be a **counter** that gets used for a **loop**, which is confirmed later, when we notice the large **blue arrow** that **circles** back around into the next box. When we see a **register** (typically **RBP**) has something **added to it** continuously, it typically signifies that we are dealing with a **string/array/struct** and **RBP** is

pointing to the **base of this structure**, with **whatever is being added to it** as the **offset** within this **structure**.

Comparison and Jump Instructions:

Now if we take a look at what the next set of blocks look to be doing, we can see a bunch of **arrows** jumping to other boxes, which symbolizes **code/control flow**. Let's examine this next box a little more. We can see that **2** different directions can be taken, but now we need to figure out **what** determines the path to take. In assembly, when it comes to just **2** paths, the path taken is determined by different **flags** that have been set and stored in the **EFLAGS** register, typically by a **comparison instruction** (such as **<**, **>**, **==**, **<=**, **>=**, etc) and a **jump instruction** (such as **jg**, **jgr**, **jl**, **jle**, **je**, **jz**, etc). Based on this **comparison and jump flag**, you take either the **red** or the **green** arrow. The instruction that sets the flags in this case is the **'cmp [rbp+var_14], 1B'**. If you couldn't guess already, the **'cmp'** instruction means compare, and we're comparing the value **'1B'** as a **hexadecimal** value (which is what the **'h'** means after the **1B**) to the value stored at what is being pointed to by the address located at **'rbp + var_14'** (we can tell because when

something is in **brackets**, it means a **pointer** to an **address**, that address being what is in the brackets).

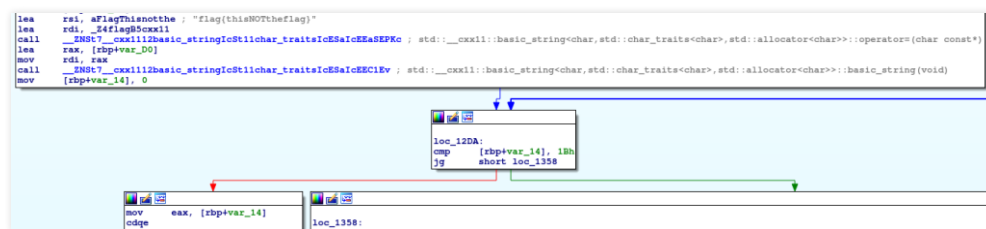
With a **comparison** instruction, certain flags get set in the 'EFLAGS' register, based on the outcome of this comparison. For example, the 'cmp' instruction will do a **subtraction between the two instructions**, the **source** from the **destination** (right side subtracted from left side in Intel syntax) and set **ZF** (Zero Flag) and **SF** (Sign Flag) accordingly. If our subtraction leads to **0**, we will set **ZF** to **1** and the **SF** will be set to **0**. If the subtraction leads to a negative number (the **right** side was **less than** the **left** side), then we will have **SF** be set to **1** and **ZF** be set to **0**.

The next instruction is: 'jg short loc_1358', at the bottom of the box. The 'jg' instruction means 'jump if greater', so it will take either the **red** arrow or the **green** arrow, based on the **flags** set in the prior **comparison instruction** (cmp [rbp+var_14], 1Bh). IDA uses intel syntax, which states that for a **cmp** instruction, we follow:

```
cmp minuend, subtrahend
```

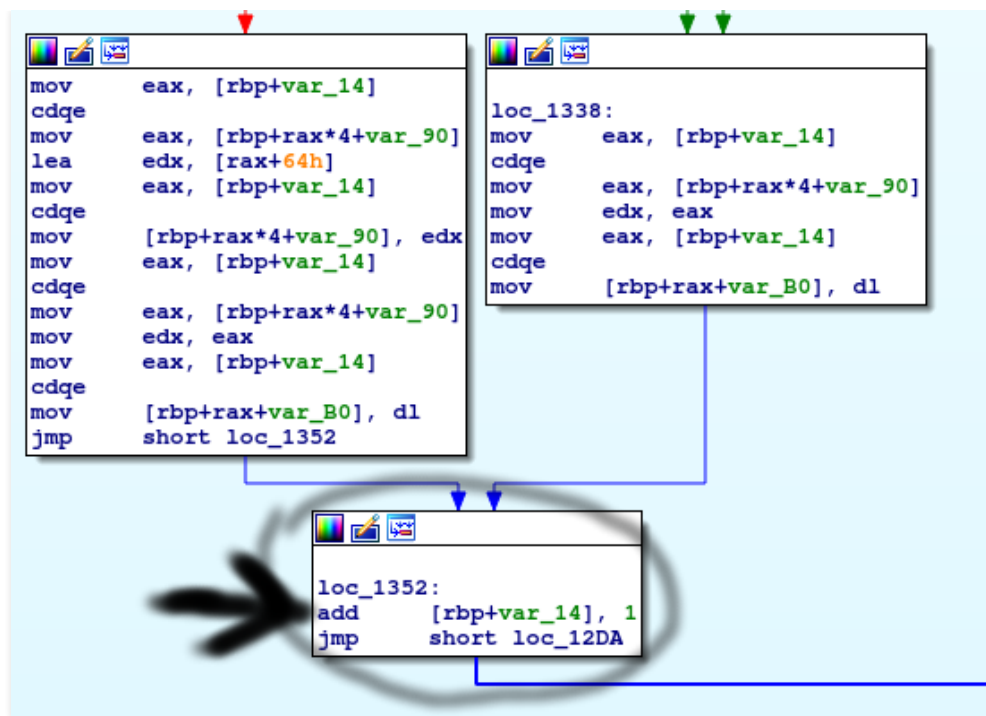
What this does is it performs a comparison between the minuend and subtrahend, and does a **signed subtraction of the subtrahend from the minuend**, or in other words, subtracts the right from the left (minuend – subtrahend). The **jg** instruction checks the **SF** and **ZF** flag to see if either are **0**, meaning, if our subtraction (minuend – subtrahend) gave us **0** or a **negative** number. If we did get **0** (the numbers were the same) the **ZF** flag would be set to **1**, and if we got a **negative** number (right hand side was larger than the left hand side), then the **SF** flag would be set to **1**, and we would take the jump.

So from the picture, we take the **red** arrow (meaning the jump is **NOT** taken) if the comparison showed that the **source** (right side) was **greater** than the **destination** (left side) and the **green** arrow (meaning the jump **IS** taken) if the **source** was **less** than the **destination**. What the arrows look like in IDA can be seen in the picture below.



Comparison Instruction with Jump Instruction

When it comes to reverse engineering **functions**, when you spot an **interesting area** that you want to look into, it's usually best to **start at the end and work backwards**. After the block with the jump and comparison instructions, we see that if we take the **red** arrow side of code flow, we have a **blue** arrow that **loops back** around to the comparison box. This is analogous of a **loop** in code, like a **for loop** for example. Seeing this alongside the fact that we are comparing a number (**1Bh**) to a value that is pointed to by **'rbp+var_14'** (which initially started at **0**) but is **incremented by 1** just before **returning** to the start of the loop (as seen in the picture below) makes it safe to assume that we are **iterating** through something, likely an **array or string** of some sort. This loops continues until this iterator reaches **'1B'** in **hexadecimal** (**28** in **decimal**), in which we take the other path finally.



Iterator being incremented before returning to
start of loop

The Turning Point:

Before we move onto the other path from this original comparison box at the top of this loop, let's reexamine the bottom of the red (left) side and try to gather some more information as to what this might be doing. Again, working backwards, we have something being loaded into a location

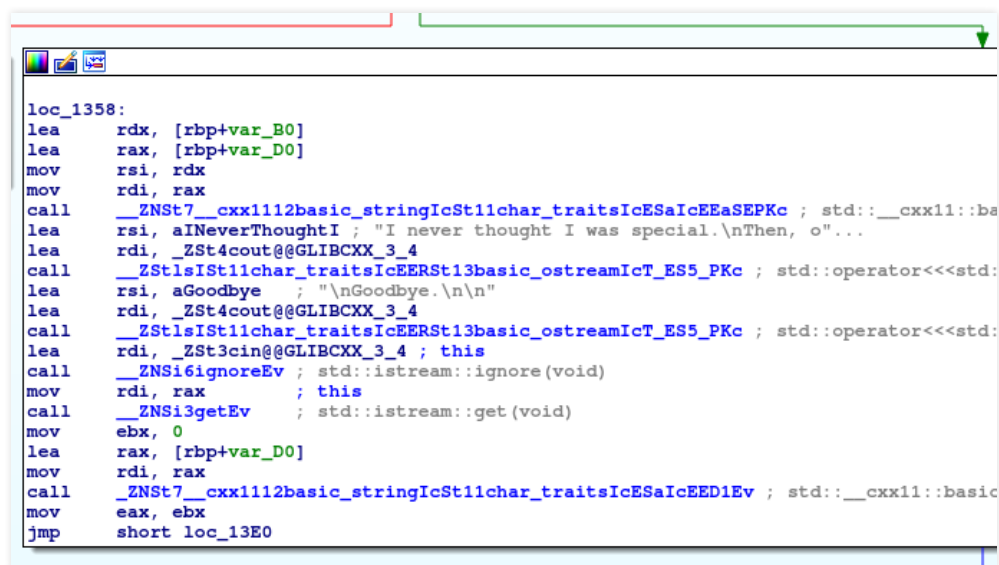
`'[rbp + rax + var_B0]'`. This means that something is being loaded into the offset of `'var_B0'` inside of `'rbp'`, at the iterator in `'rax'`. I stopped here and began my search to find what `'var_B0'` might be (when I reference `'var_B0'` or `'var_90'` by themselves, I am talking with respect to `'rbp + rax'`), since the rest of the box isn't really necessary to understand to complete the challenge, but if you are interested, you can continue to read the next paragraph, otherwise, **you can move on to the next section!**

NOTE: These next few paragraphs **are not needed** to solve challenge, but might be useful to get a better picture of what is going on. Starting at the top of the **bottom left box** in the loop (**left box in above picture**), we see that the **iterator** (`[rbp + var_14]`) is put into `'EAX'`, which is then **multiplied by 4** (via `'RAX'`, which if we recall, is the **64 bit version** of `'EAX'`, so the same register, just **64 bit extended**, rather than 32 bit) and added to `'var_90'` (signifying that we are moving **4 bytes** at a time through whatever is at offset `'var_90'` inside of the structure at location `'rbp'`). Then, whatever is at this location is loaded into the `'EAX'` register. Now that we know that `'var_90'` is some offset in a **structure** that points to some **string/array** potentially, let's move on and see what the end of this box leaves us with.

Continuing the quest to get a better understanding of what is in `'var_B0'`, let's examine the rest of this box. Just before leaving the box, we can see that both boxes (left and right) end with a `'mov [rbp+rax+var_B0], dl'`. Well, what exactly is `'dl'` and what's in it? A `'dl'` is the **8-bit** version of **RDX/EDX** register, specifically, the **lower 8 bits** (hence the **l**), the **upper 8 bits** would be `'dh'`. I know that probably doesn't make a whole lot of sense to you, but again, since it **isn't** exactly necessary to solve this challenge if we go this route, I won't go into detail, but you can reference [this](#) for a better understanding, if you would like. Now, what gets put into `'dl'`? Since `'dl'` is a part of `'EDX'`, we continue going back up and see that `'EAX'` is put into `'EDX'` via the `'mov edx, eax'` instruction. Now we continue going back up and we see that `'[rbp + rax*4 + var_90]'` is put into `'EAX'`. So from this, we can assume that we are **iterating** through a **string** at **offset** `'var_90'`, **4 bytes at a time**, starting at the **base address** in `'RBP'`. From here, we have the option to now figure out what exactly is in `'var_90'` (most of which can be seen in the first box that IDA provided, with the **28 'mov' instructions**), or we can try to just see what we end up with inside of `'var_B0'`, since it looks to be potentially useful to us. I won't continue with the search of `'var_90'`, since it isn't exactly straight forward or needed for this challenge.

Wrapping up the Static Analysis:

With what looks to be the important part of the **red arrow's** code flow, we continue our search for what is inside of (pointed to by) 'var_B0' by checking the **green arrow's** code flow. Below shows the **green arrow** flow after the original comparison box.



```
loc_1358:
lea     rdx, [rbp+var_B0]
lea     rax, [rbp+var_D0]
mov     rsi, rdx
mov     rdi, rax
call    __ZNSt7_cxx1112basic_stringIcSt11char_traitsIcESaIcEEaSEPKc ; std::__cxx11::ba
lea     rsi, aINeverThoughtI ; "I never thought I was special.\nThen, o"...
lea     rdi, _ZSt4cout@@GLIBCXX_3_4
call    __ZStlsISt11char_traitsIcEESt13basic_ostreamIcT_ES5_PKc ; std::operator<<<std:
lea     rsi, aGoodbye ; "\nGoodbye.\n\n"
lea     rdi, _ZSt4cout@@GLIBCXX_3_4
call    __ZStlsISt11char_traitsIcEESt13basic_ostreamIcT_ES5_PKc ; std::operator<<<std:
lea     rdi, _ZSt3cin@@GLIBCXX_3_4 ; this
call    __ZNSi6ignoreEv ; std::istream::ignore(void)
mov     rdi, rax ; this
call    __ZNSi3getEv ; std::istream::get(void)
mov     ebx, 0
lea     rax, [rbp+var_D0]
mov     rdi, rax
call    __ZNSt7_cxx1112basic_stringIcSt11char_traitsIcESaIcEEED1Ev ; std::__cxx11::basic
mov     eax, ebx
jmp     short loc_13E0
```

Green Arrow Code Flow

When we finally get to the **green arrow** code flow, we have passed the check of 'cmp [rbp+var_14], 1Bh', thus the loop

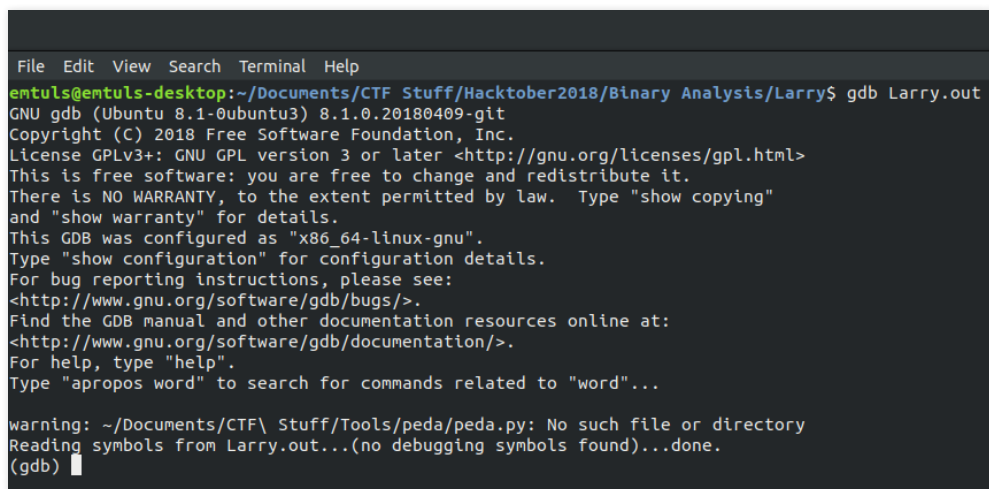
has completed its work, whatever that was. We then see that `'[rbp+var_B0]'` is pushed into `'RDX'` from the `'lea rdx, [rbp + var_B0]'` and `'[rbp+var_D0]'` is pushed into `'RAX'` via the `'lea rax, [rbp+var_D0]'`, which then are moved into `RSI` and `RDI` via `'mov rsi, rdx'` and `'mov rdi, rax'` respectively (the instructions are highlighted because we will need to remember them for the static analysis part). Now, we need to view the contents of either `'RDX'` or `'RSI'` in order to see what was in `'var_B0'`, and to do this, we will finally be pulling out our debugger, for some dynamic analysis, now that we've done enough static analysis!

Whipping out GDB for some Dynamic Analysis:

My debugger of choice for Linux is `GDB`, so this write-up will be written with that context. I enjoy using it for some quick and dirty debugging and it is fairly popular as debugger overall, especially when combined with the `peda` (python exploit development assistance) extension. Though for this challenge, I used vanilla `GDB`.

First, we should start inside a `terminal`/command prompt, inside of the directory with our file, which in this case is

'Larry.out'. From here, we can run the command `'gdb Larry.out'`, which starts our debugger with the file of our choice. Once inside of the debugger, which you should be able to tell you are in because of the terminal displaying (gdb) next to your cursor (shown in the picture below), we can enter a few commands to make this analysis easier. The first command I recommend typing in is `'set disassembly-flavor intel'` to have our assembly language **displayed in Intel syntax**, rather than AT&T syntax (Intel syntax is easier to read, I promise). Then, we can begin to disassemble the program as we see fit. Since we already have an idea of where we want to end up, this will be fairly quick.

A screenshot of a terminal window with a dark background. The terminal shows the command 'gdb Larry.out' being executed. The output includes the GNU gdb version (8.1.0.20180409-git), copyright information (© 2018 Free Software Foundation, Inc.), license information (GPLv3+), and a warning about the file path. The prompt '(gdb) ' is visible at the bottom.

```
File Edit View Search Terminal Help
entuls@entuls-desktop:~/Documents/CTF Stuff/Hacktober2018/Binary Analysis/Larry$ gdb Larry.out
GNU gdb (Ubuntu 8.1-0ubuntu3) 8.1.0.20180409-git
Copyright (C) 2018 Free Software Foundation, Inc.
License GPLv3+: GNU GPL version 3 or later <http://gnu.org/licenses/gpl.html>
This is free software: you are free to change and redistribute it.
There is NO WARRANTY, to the extent permitted by law.  Type "show copying"
and "show warranty" for details.
This GDB was configured as "x86_64-linux-gnu".
Type "show configuration" for configuration details.
For bug reporting instructions, please see:
<http://www.gnu.org/software/gdb/bugs/>.
Find the GDB manual and other documentation resources online at:
<http://www.gnu.org/software/gdb/documentation/>.
For help, type "help".
Type "apropos word" to search for commands related to "word"...
warning: ~/Documents/CTF\ Stuff/Tools/peda/peda.py: No such file or directory
Reading symbols from Larry.out...(no debugging symbols found)...done.
(gdb) 
```

GDB started up with Larry.out

Disassembling Main:

Let's take a look for where the point is that we wanted to reach, so we can tell our disassembler that that is where we want to stop at. We can do this by typing `'disassemble main'` (or `'disas main'` for short), which will spit out a bunch of assembly language. We choose to disassemble the `'main'` function because our entire analysis was only within the main function of the program, as seen from the dynamic analysis. After disassembling this, we get a screen that looks like the picture below.

```
File Edit View Search Terminal Help
Dump of assembler code for function main:
0x00000000000011d5 <+0>:  push    rbp
0x00000000000011d6 <+1>:  mov     rbp, rsp
0x00000000000011d9 <+4>:  push    rbx
0x00000000000011da <+5>:  sub     rsp, 0xc8
0x00000000000011e1 <+12>: mov     DWORD PTR [rbp-0x90], 0x2
0x00000000000011eb <+22>: mov     DWORD PTR [rbp-0x8c], 0x8
0x00000000000011f5 <+32>: mov     DWORD PTR [rbp-0x88], 0x61
0x00000000000011ff <+42>: mov     DWORD PTR [rbp-0x84], 0x3
0x0000000000001209 <+52>: mov     DWORD PTR [rbp-0x80], 0x2d
0x0000000000001210 <+59>: mov     DWORD PTR [rbp-0x7c], 0x59
0x0000000000001217 <+66>: mov     DWORD PTR [rbp-0x78], 0xb
0x000000000000121e <+73>: mov     DWORD PTR [rbp-0x74], 0x11
0x0000000000001225 <+80>: mov     DWORD PTR [rbp-0x70], 0x27
0x000000000000122c <+87>: mov     DWORD PTR [rbp-0x6c], 0xe
0x0000000000001233 <+94>: mov     DWORD PTR [rbp-0x68], 0x1
0x000000000000123a <+101>: mov    DWORD PTR [rbp-0x64], 0x20
0x0000000000001241 <+108>: mov    DWORD PTR [rbp-0x60], 0x40
0x0000000000001248 <+115>: mov    DWORD PTR [rbp-0x5c], 0x20
0x000000000000124f <+122>: mov    DWORD PTR [rbp-0x58], 0x8
0x0000000000001256 <+129>: mov    DWORD PTR [rbp-0x54], 0x5
0x000000000000125d <+136>: mov    DWORD PTR [rbp-0x50], 0x16
0x0000000000001264 <+143>: mov    DWORD PTR [rbp-0x4c], 0x40
0x000000000000126b <+150>: mov    DWORD PTR [rbp-0x48], 0xe
0x0000000000001272 <+157>: mov    DWORD PTR [rbp-0x44], 0x64
0x0000000000001279 <+164>: mov    DWORD PTR [rbp-0x40], 0x2c
0x0000000000001280 <+171>: mov    DWORD PTR [rbp-0x3c], 0x20
0x0000000000001287 <+178>: mov    DWORD PTR [rbp-0x38], 0x4c
0x000000000000128e <+185>: mov    DWORD PTR [rbp-0x34], 0x40
0x0000000000001295 <+192>: mov    DWORD PTR [rbp-0x30], 0xe
0x000000000000129c <+199>: mov    DWORD PTR [rbp-0x2c], 0xe
0x00000000000012a3 <+206>: mov    DWORD PTR [rbp-0x28], 0x15
0x00000000000012aa <+213>: mov    DWORD PTR [rbp-0x24], 0x21
0x00000000000012b1 <+220>: lea     rsi, [rip+0xd51]          # 0x2009
0x00000000000012b8 <+227>: lea     rdi, [rip+0x3021]        # 0x42e0 <_Z4flagB5cxx11>
0x00000000000012bf <+234>: call    0x1070 <_ZNSt7_cxx1112basic_stringIcSt11char_traitsIcEsaIcEEaSEPKc@plt>
0x00000000000012c4 <+239>: lea     rax, [rbp-0xd0]
0x00000000000012cb <+246>: mov     rdi, rax
0x00000000000012ce <+249>: call    0x1090 <_ZNSt7_cxx1112basic_stringIcSt11char_traitsIcEsaIcEEc1Ev@plt>
0x00000000000012d3 <+254>: mov     DWORD PTR [rbp-0x14], 0x0
0x00000000000012da <+261>: cmp     DWORD PTR [rbp-0x14], 0x1b
0x00000000000012de <+265>: jg      0x1358 <main+387>
0x00000000000012e0 <+267>: mov     eax, DWORD PTR [rbp-0x14]
0x00000000000012e3 <+270>: cdq     eax
0x00000000000012e5 <+272>: mov     eax, DWORD PTR [rbp+rax*4-0x90]
0x00000000000012ec <+279>: test    eax, eax
0x00000000000012ee <+281>: jle     0x1338 <main+355>
0x00000000000012f0 <+283>: mov     eax, DWORD PTR [rbp-0x14]
0x00000000000012f3 <+286>: cdq     eax
0x00000000000012f5 <+288>: mov     eax, DWORD PTR [rbp+rax*4-0x90]
0x00000000000012fc <+295>: cmp     eax, 0x16
0x00000000000012ff <+298>: jg      0x1338 <main+355>
0x0000000000001301 <+300>: mov     eax, DWORD PTR [rbp-0x14]
0x0000000000001304 <+303>: cdq     eax
0x0000000000001306 <+305>: mov     eax, DWORD PTR [rbp+rax*4-0x90]
0x000000000000130d <+312>: lea     edx, [rax+0x64]
0x0000000000001310 <+315>: mov     eax, DWORD PTR [rbp-0x14]
0x0000000000001313 <+318>: cdq     eax
0x0000000000001315 <+320>: mov     DWORD PTR [rbp+rax*4-0x90], edx
0x000000000000131c <+327>: mov     eax, DWORD PTR [rbp-0x14]
0x000000000000131f <+330>: cdq     eax
0x0000000000001321 <+332>: mov     eax, DWORD PTR [rbp+rax*4-0x90]
0x0000000000001328 <+339>: mov     edx, eax
0x000000000000132a <+341>: mov     eax, DWORD PTR [rbp-0x14]
0x000000000000132d <+344>: cdq     eax
0x000000000000132f <+346>: mov     BYTE PTR [rbp+rax*1-0xb0], dl
0x0000000000001336 <+353>: jmp     0x1352 <main+381>
0x0000000000001338 <+355>: mov     eax, DWORD PTR [rbp-0x14]
0x000000000000133b <+358>: cdq     eax
0x000000000000133d <+360>: mov     eax, DWORD PTR [rbp+rax*4-0x90]
0x0000000000001344 <+367>: mov     edx, eax
0x0000000000001346 <+369>: mov     eax, DWORD PTR [rbp-0x14]
0x0000000000001349 <+372>: cdq     eax
0x000000000000134b <+374>: mov     BYTE PTR [rbp+rax*1-0xb0], dl
```

```
0x0000000000001352 <+381>:  mov     BYTE PTR [rbp-0xb0],0
0x0000000000001352 <+381>:  add     DWORD PTR [rbp-0x14],0x1
--Type <return> to continue, or q <return> to quit--
```

‘Disassemble main’ output from Larry.out

Now if we recall, we are trying to view what was inside of offset ‘**var_B0**’, which got put into **RSI** from **RDx**. We want to find that point where this happens in the disassembly, so we need to continue a little further down the screen in order to see it, so we need to hit ‘**enter**’ to continue. GDB should reach the ‘**end of assembler dump**’, and we should be able to see ‘**lea rdx, [rbp-0xb0]**’ followed by ‘**lea rax, [rbp-0xd0]**’ and then the ‘**mov rsi, rdx**’ as well as the ‘**mov rdi, rax**’. GDB **does not** automatically name our variables as ‘**var_B0**’ or ‘**var_D0**’ like IDA did, but we can see that they still look fairly similar, since IDA got those variable names from the ‘**0xd0**’ and ‘**0xb0**’ offsets that they were in actuality.

Breakpoints and Solving the Challenge:

Now that we’ve found the location we want to be at for inspection, we need to tell GDB to stop there. In order to do that, we have to use something called a ‘**breakpoint**’. A breakpoint placed at a memory address will **stop just before**

the instruction at that memory address is executed. Let's set a breakpoint right before the `'call'` instruction that comes after the `'mov rdi, rax'` instruction that we might be interested in. To do this, we need to know what the address that the instruction we want to break at is, which if we look to the left, we see `'0x000000000000136c <+407>:'`, and to shorten this, the `'<+407>'` means **+407 bytes** from the **start** of the function we disassembled, which in our case, was `'main'`, so `'break *main+407'`. Since this is a memory address and not a variable or function, we need to **dereference** this address for the instruction, which is what the `*` is for. This should have GDB spit out `'Breakpoint 1 at 0x136c'`, and look like the image below.

```
0x00000000000013e7 <+530>:  pop    rbx
0x00000000000013e8 <+531>:  pop    rbp
0x00000000000013e9 <+532>:  ret
End of assembler dump.
(gdb) break *main+407
Breakpoint 1 at 0x136c
```

Break `*main+407` in `Larry.out`

Next, let's run our program until we hit the breakpoint! To do this, we type `'run'` or `'r'` for short, and gdb should **hit our breakpoint**, which looks like the picture below.


```
(gdb) r
Starting program: /home/emtuls/Documents/CTF Stuff/Hacktober2018/Binary Analysis/Larry/Larry.out
Breakpoint 1, 0x000055555555536c in main ()
(gdb) □
```

Running the program and hitting the breakpoint

This should put us at the location we wanted to be at in order to inspect what is inside of `'var_0B'`, which if we recall, was placed inside of `RDX`, which was then placed inside of `RSI`. We have the option to examine either of them at this point, and they should both give us the same result, which should **hopefully** be a **string**, since we **assumed** previously (in the skipped section if you skipped over it) that since we were **iterating** through this offset, it must be a **string/array/structure**. Let's go with `RSI`. In order to examine this register, we need to use the command `'x/s $rsi'` (be sure `RSI` is **lowercase**), which means, **'examine, as a string, the register rsi'**. The `'x'` command for **examine**, can also be used to view memory addresses as hex bytes, words, quad words, decimals, and other things, but we won't get into that in this blog post, more on that can be found [here](#). Now, once we run this command, we should see the **correct flag**, **"flag-You're @ liz@ard, L@rry!"**, (plus a few extra bytes, `'UU'`, due to how gdb interpreted the bytes as a

string) Your output should look something like the picture below.

```
(gdb) x/s $rsi
0x7fffffffdd70: "flag-You're @ liz@rd, L@rry!UU"
(gdb) █
```

Challenge flag found! (plus an extra UU)

Wrong Assumption turned Good:

During my first assumption, where I thought the loop was manipulating a string in memory, I assumed the flag would still be in **RSI** after this loop, but manipulated to the correct flag, thus I checked '**RSI**' **before the function call** really early on. This was an **incorrect assumption**, but, fortunately for me, the correct flag was stored in RSI at the end anyway, so I was able to get the flag quickly, but that would have made for a boring write up!

Conclusion:

Well, I hope this write up/blog post was educational for you guys! I know I learned a bit doing the challenge again and

trying to break it down in a way that makes sense, since I tend to glisten over facts and run off a lot of assumptions at times. This is my pseudo-introduction to Reverse Engineering, in which I will cover more in later blog posts. Also, check out my other posts on how I am working on a **Wargame** for people to learn **Exploit Development** and **Reverse Engineering**!

If anyone has any trouble with something in the walk-through as far as needing clarification or they are attempting to do it themselves and find I messed up somewhere, please let me know! Thank you.

Where you can learn more about Reverse Engineering:

The binary can be found on my github: https://github.com/emtuls/ctf/tree/master/2018-hacktober.org/Binary_Analysis/binaries -> Larry.out

For anyone that needs resources for learning Reverse Engineering, I can provide you with a baseline that I would recommend starting with. Eventually, I plan on making my own set of tutorials...but that's in the works.

x86 Assembly:

If you don't know assembly language at all, [this list of videos](#) was where I picked up a decent amount of x86 assembly language.

A few good books would be:

- [Hacking: The Art of Exploitation](#) I am a huge advocate for this book. I learned a lot from this and have read it multiple times. It is written very well and teaches someone with no experience how to do C programming and assembly. This is mainly a book for learning exploitation/vulnerability research, but that can play hand and hand with Reverse Engineering. It will show you the assembly language break down of basic exploits and this can help you with RE.
- [Practical Reverse Engineering](#) I read through the beginning of this book and it gave me some good foundations of understanding memory and computer architecture for RE along with assembly of course
- [Secrets of Reverse Engineering](#) This book is a bit in depth, but the beginning gives another good foundation for Comp Architecture and assembly stuff.

- [The IDA Pro Book](#) Haven't personally read this book yet, but I have been told it is the defacto standard for learning IDA Pro, and it has examples you can learn from.

Hands On:

- [Legend of Random](#) Very useful hands on with tutorials. Mainly based on cracking, but that requires reverse engineering. Highly recommend this!
- [Lenas Tutorials](#) Again, another awesome hands on tutorial, mostly based on cracking as well.
- [Crackmes](#) These are more of challenges once you start to have a little understanding down

Courses:

Tons of courses on youtube. I learn well from visual, so I recommend these youtube videos, but there are plenty of other good videos to learn from!:

- [Basic Dynamic Analysis](#)
- [Real World Decompilation](#) There are a few videos to this series and he disassembles a game, definitely nice to learn from.

Beyond that, Google will always be your friend,
and [/r/reverseengineering](#).

Obligatory Ending Statement:

I hope you liked this blog post, and if you have any questions, feel free to contact me in one of my many ways of communication! Thank you!

[@emtuls](#) 

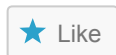
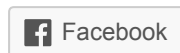
[Github](#)

[Gmail](#)

[LinkedIn](#)

If you're a veteran interested in Cyber Security, consider joining our [Slack](#) channel.

Share this:



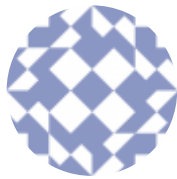
Be the first to like this.

Related

Review: SANS Cyber
FastTrack 2019

Creating VetSecs
Wargame Pt. 2:
Tweaking the VM and
creating the first
challenge
In "Exploit
Development"

VetSec Takes First
in the Hacktober
CTF: Summary &
Steganography
Write-up!
In "CTF Write-ups"



PUBLISHED BY EMTULS

[View all posts by emtuls](#)

[< PREVIOUS POST](#)

[NEXT POST >](#)

VetSec Takes First in the
Hacktober CTF: Summary &
Steganography Write-up!

Hack The Box –
TartarSauce Walkthrough

6 comments on “Hacktober CTF 2018 –
Binary Analysis – Larry”

Pingback: [Hacktober CTF 2018 – Binary Analysis – Larry | | Lowmiller Consulting Group Blog](#)



KOKN3T SAYS:

October 29, 2018 at 6:07 pm

Wow. Very nice for me coz I'm starting for Reversing. Thanks. ❤️



Like

[REPLY](#)



EMTULS SAYS: October 29, 2018 at 6:11 pm

Glad to hear that! I'll be releasing more tutorials in time! 😊



Like

[REPLY](#)



CL SAYS:

November 13, 2018 at 3:02 am

This is a good writeup to me as I am new to that. But may I have some clarification on the `jg` assembly instruction. From your writeup:

The next instruction is: `'jg short loc_1358'`, at the bottom of the box. The `'jg'` instruction means `'jump if greater'`, so it will take either the red arrow or the green arrow, based on the flags set in the prior comparison instruction. It takes the red arrow (meaning the jump is NOT taken) if the comparison showed that the source (right side) was less than the destination (left side) and the green arrow (meaning the jump IS taken) if the source was greater than the destination. What the arrows look like in IDA can be seen in the picture below.

When I read

https://en.wikibooks.org/wiki/X86_Assembly/Control_Flow#Jump_if_Greater, it said the execution would be directed to another place if the destination operand is greater than source operand (that means true), it seems like a

bit different to your writing, do i get something wrong?

Thanks!



REPLY



EMTULS SAYS:

November 13, 2018 at 12:11 pm

Hey there! Thank you for reading!

You're correct in that my writing was wrong. I improperly stated that it would take the red arrow if the comparison showed that the right was less than the left side, when in fact, I meant the exact opposite! I will update the write-up to reflect the correct meaning and also try to clarify it a bit more!

Thank you for the help!

★ Like

REPLY



CL SAYS:

November 15, 2018 at 1:32 am

Thanks for the clarification!

★ Like

LEAVE A REPLY

Enter your comment here...

TO TOP ^

