A friendly user guide to ELF micropatching with Shiva (alpha v1)

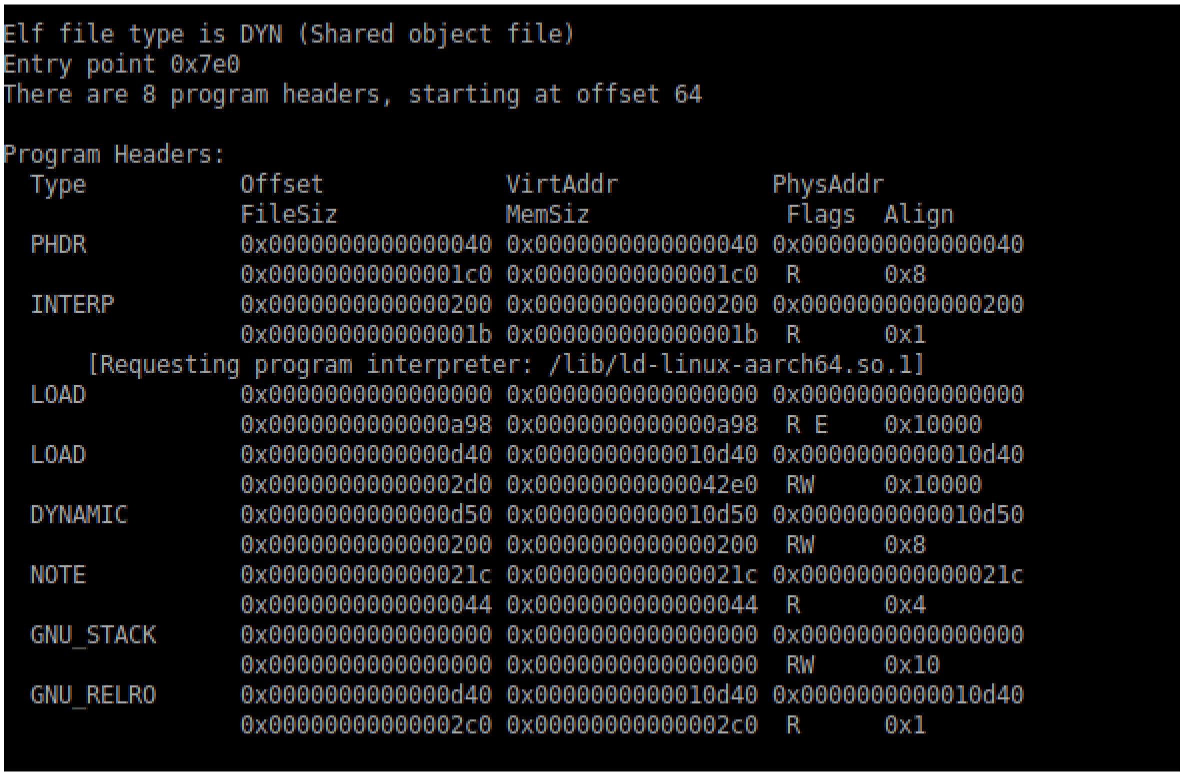
Shiva Manual

1 Introduction to Shiva

ELF program interpreters

Shiva is an ELF microcode-patching technology that works similarly to the well known (*but little understood*) ELF dynamic linker "/lib/ld-linux.so". Every ELF program that is dynamically linked has a program header segment type called PT_INTERP that describes the path to the *program interpreter.* The program interpreter is a separate ELF program loaded by the kernel that is responsible for helping to build the process runtime image by loading and linking extra modules containing code and data. Traditionally these modules are called shared libraries, and the program interpreter is "/lib/ld-linux.so". On aarch64 it is "/lib/ld-linux-aarch64.so.1"

Image 1.0: Requesting program interpreter



```
Elf file type is DYN (Shared object file)
Entry point 0x7e0
There are 8 program headers, starting at offset 64

Program Headers:
  Type           Offset             VirtAddr           PhysAddr
                 FileSiz            MemSiz              Flags  Align
  PHDR           0x0000000000000040 0x0000000000000040 0x0000000000000040
                 0x00000000000001c0 0x00000000000001c0  R      0x8
  INTERP         0x0000000000000200 0x0000000000000200 0x0000000000000200
                 0x000000000000001b 0x000000000000001b  R      0x1
      [Requesting program interpreter: /lib/ld-linux-aarch64.so.1]
  LOAD           0x0000000000000000 0x0000000000000000 0x0000000000000000
                 0x0000000000000a98 0x0000000000000a98  R E    0x10000
  LOAD           0x0000000000000d40 0x0000000000010d40 0x0000000000010d40
                 0x00000000000002d0 0x00000000000042e0  RW     0x10000
  DYNAMIC        0x0000000000000d50 0x0000000000010d50 0x0000000000010d50
                 0x0000000000000200 0x0000000000000200  RW     0x8
  NOTE           0x000000000000021c 0x000000000000021c 0x000000000000021c
                 0x0000000000000044 0x0000000000000044  R      0x4
  GNU_STACK      0x0000000000000000 0x0000000000000000 0x0000000000000000
                 0x0000000000000000 0x0000000000000000  RW     0x10
  GNU_RELRO      0x0000000000000d40 0x0000000000010d40 0x0000000000010d40
                 0x00000000000002c0 0x00000000000002c0  R      0x1
```

Just like any other interpreter, such as Python, or Java, the ELF program interpreter is responsible for performing various types of loading, patching and transformation.

Shiva is an ELF program interpreter

Shiva is a newly designed program interpreter who's sole purpose is to load ELF microcode patches that are written in C and compiled into Shiva modules (ELF relocatable objects). Shiva modules are relocatable objects and by nature already contain the ELF meta-data *(sections, symbols relocations)* necessary to re-build much of the process image. Additionally Shiva has introduced some new concepts such as *Transformations* which are an extension to traditional ELF relocations, allowing for complex patching operations such as function splicing.

The core purpose of Shiva is to create a microcode patching system that seamlessly fits into the existing ELF ABI toolchain of compilers, linkers, and loaders. Shiva aims to load and link patches that are written 100% in C, in a way that is natural for developers who may not be reverse engineers but are experts in C. Shiva is in-itself a linker and a loader that works alongside and in conjunction with the existing Dynamic linker "ld-linux.so". To understand more technical details see "shiva_final_desing.pdf" within the documentation directory.

Systems and arch's that Shiva supports

Currently Shiva's AMP tailored-microcode patching system runs on **Linux AArch64**, and supports **AArch64 Linux ELF PIE binaries**. Future support for ARM32, x86_64/x86_32 and other requested architectures. At the present moment Shiva doesn't support ET_EXEC ELF binaries (non-pie) due to time constraints, but it's on the road map.

2 Installing Shiva-AArch64

Download Source code

```
git clone git@github.com:advanced-microcode-patching/shiva
git clone github.com/elfmaster/libelfmaster
```

Build source code

Build and install libelfmaster

Shiva requires a custom libelfmaster branch that has not been merged into main yet. Specifically make sure to use the *aarch64_support* branch.

```
cd libelfmaster
git checkout aarch64_support
cd ./src
sudo ./make.sh
```

Build and install Shiva

Commands:

```
cd shiva
make
make shiva-ld
make patches
sudo make install
```

Build results:

Shiva is a static ELF executable and installed to "/lib/shiva". The Shiva modules are typically stored in "/opt/shiva/modules". The example patches all live within "/git/shiva/modules/aarch64_patches/".

Build and install example patches

Please see the "shiva/modules/aarch64_patches/" directory. All of the example patches live here. Let us illustrate an example of building and testing one such patch on a program called 'test_rodata'.

Commands:

```
cd shiva/modules/aarch64_patches/rodata_interposing
make
sudo make install
```

Test the patch by setting the patch path and running "/lib/shiva" manually.

```
$ SHIVA_MODULE_PATH=./ro_patch.o /lib/shiva ./test_rodata
```

Notice that the software prints the new string "The great arcanum", per the patch. Yet when the program is executed without Shiva it executes with its old behavior.

```
$ ./test_rodata
```

Install the patch permanently

You may also install the patch permanently with the Shiva prelinking tool. This tool updates the PT_INTERP segment from "/lib/ld-linux.so" to "/lib/shiva", and updates the dynamic segment with several new tags describing the path to the patch that should be loaded and linked at runtime: "/opt/shiva/modules/ro_patch.o".

```
$ /bin/shiva-ld -e test_rodata -p ro_patch.o -i /lib/shiva -s /opt/shiva/modules -o test_rodata.patched
```

Shiva now installs the patch in memory everytime it's executed

```
$ ./test_rodata.patched
```

The only discernible difference between test_rodata.patched and test_rodata are the PT_INTERP segment modification and the PT_DYNAMIC addition of two new entries describing the patch path. No actual code or data has changed within the ELF binary. Shiva performs all of the patching at runtime.

3 Writing patches with Shiva

The Shiva workflow

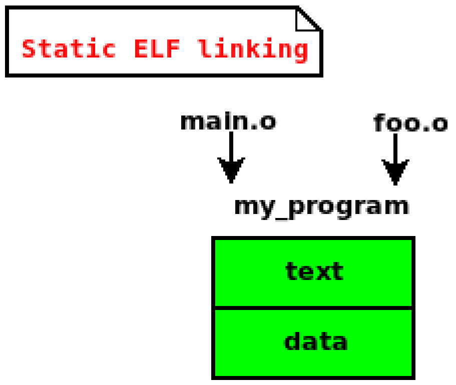
Before delving in to the use of Shiva, let us first explain the workflow of what the ELF compilation and linking process looks like before and after Shiva.

ELF Program compilation and linking example

Standard program compilation begins with a compiler like gcc; the source code files (I.e. main.o and foo.o) get linked into a final output ELF executable my_program.

Static linking of objects

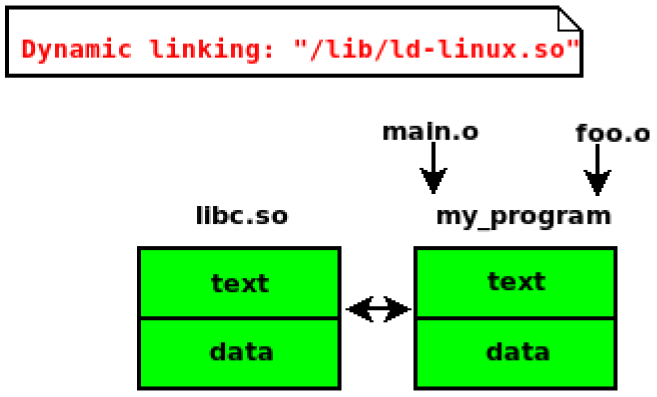
Image 3.0: static linking of two ELF relocatable objects



Additionally these ELF relocatable object files contain calls to functions that live within ELF shared libraries. Once the object files (foo.o and main.o) are linked into an ELF executable (my_program), additional Dynamic linking meta-data is created and stored into what is known as the PT_DYNAMIC segment of the program header table. The PT_DYNAMIC segment contains additional meta-data that is to be passed along to the Program interpreter at runtime. The program interpreter (aka: dynamic linker) is specified as a file path in the PT_INTERP segment: "/lib/ld-linux.so". When the executable is run, the kernel loads and executes the interpreter program first: "/lib/ld-linux.so". The interpreter loads and links the needed shared libraries (i.e. libc.so). The program has now been compiled, and linked twice: once statically, and again dynamically.

Dynamically linked executable

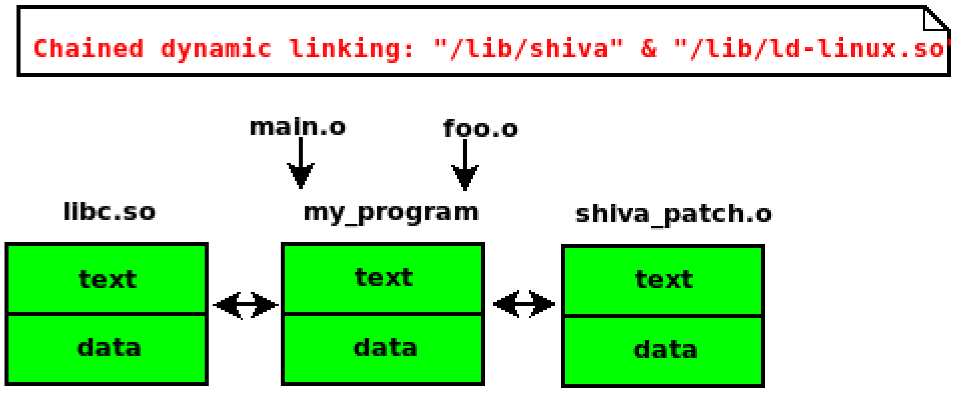
Image 3.1: Dynamically link libc.so in at runtime



A Shiva patch can be written in C and compiled as an ELF relocatable object file. In the same sense that "ld-linux.so" loads and links shared libraries (ELF ET_DYN files), Shiva will load and link patch objects (ELF ET_REL files) to re-write the program in memory just before execution.

Dynamically patched executable

Image 3.2: Chained linking with Shiva



ELF Symbol interposition

The ELF format uses symbols to denote code and data. Program functions generally live in the .text section, and global variables live within the .data, .rodata, and .bss sections. Shiva makes use of the ELF relocation, section, and symbol data to allow patch developers to *interpose* existing symbols; in other words re-link functions and data that have symbols associated with them. As patch developers we will find this as a natural and beautifully convenient way to patch code and data.

Citation 2.0: From "Oracle linkers and libraries guide"

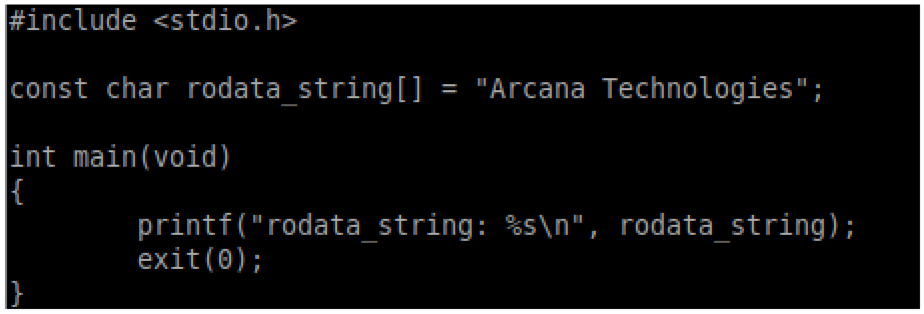
"Interposition can occur when multiple instances of a symbol, having the same name, exist in different dynamic objects that have been loaded into a process. Under the default search model, symbol references are bound to the first definition that is found in the series of dependencies that have been loaded. This first symbol is said to interpose on the other symbols of the same name."

Patching an .rodata string (Simple first patch)

This patching exercise lives in "modules/aarch64_patches/rodata_interposing".
In the following example we are going to look at a simple program that prints a constant string "Arcana Technologies". Constant data is read-only and is usually stored in the .rodata section of a program. The .rodata section is generally put into text segment, or

a read-only segment that is adjacent to the text segment. A shiva module has it's own .rodata section within it's mapped text segment at runtime. Our goal with the following patch is to replace a read-only string with another by simply re-defining it in C.

Image 2.0: Source code of program rodata_test.c



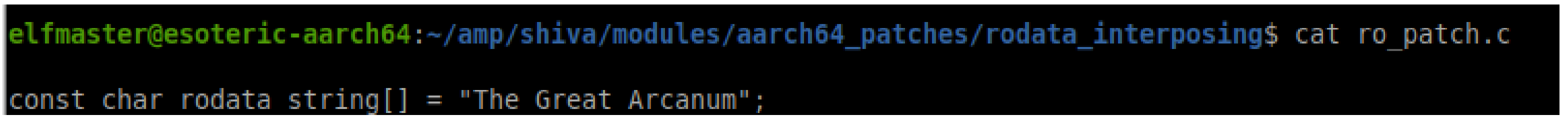
```
#include <stdio.h>

const char rodata_string[] = "Arcana Technologies";

int main(void)
{
        printf("rodata_string: %s\n", rodata_string);
        exit(0);
}
```

NOTE: The source code to rodata_test.c is only being shown for the sake of illustrating what the program does that we are patching. In may workflows the source code wouldn't even exist, and Shiva doesn't depend on or use the source in any way.

Image 2.1: Source code of patch object: ro_patch.c



```
elfmaster@esoteric-aarch64:~/amp/shiva/modules/aarch64_patches/rodata_interposing$ cat ro_patch.c

const char rodata string[] = "The Great Arcanum";
```

This simple patch is re-defining a variable named *rodata_string[]*. At runtime Shiva will build a module process image which will contain the patch .rodata section with the updated version of *rodata_string[]*. Shiva will re-link the executable to use the new version of *rodata_string[]* that lives within the loaded modules .rodata section This is the concept of symbol interposition, one global symbol taking precedence over another with the same symbol name and type. Furthermore Shiva allows the user to interpose symbols of different types as well (including STB_LOCAL symbols) as seen in later examples.

Compile this patch as an AArch64 Linux ELF relocatable object using a large code model. Each patch example, such as "modules/aarch64_patches/rodata_interposing", contains a Makefile that illustrates how to build a patch object. Generally speaking, a Shiva patch should contain the include path to "modules/include" which contain the "shiva_module.h" header file, necessary to build mo

Image 2.2: "rodata_interposing" Makefile with comments

Building the patch

The patch source code "modules/aarch64_patches/rodata_interposing/ro_patch.c" is a single line patch, redefining the symbol "rodata_string[]". To build this patch we can simply run the following gcc command to build the patch:

```
$ gcc -mcmodel=large -fno-pic -fno-stack-protector -c ro_patch.c -o ro_patch.o
```

The ELF ET_REL object: ro_patch.o describes a symbol with the same exact name as the symbol **rodata_string** in the target executable. Shiva will see this when linking and it will link in the version of the symbol from our patch. This perfectly demonstrates the concept of symbol interposition in Shiva.

Testing our patch: ro_patch.o

In this example we are patching the AArch64 ELF PIE executable "modules/aarch64_patches/rodata_interposing/test_rodata" Let's first run the program un-patched.

```
$ ./test_rodata
rodata_string: Arcana Technologies
$
```

The program simply prints the string "rodata_string: Arcana technologies". To see if our patch worked we can invoke Shiva directly to load and link the executable with the patch at runtime. We must specify the path to our patch object with the SHIVA_MODULE_PATH environment variable.

```
$ SHIVA_MODULE_PATH=./ro_patch.o /lib/shiva ./test_rodata
rodata_string: The Great Arcanum
$
```

When executing ./test_rodata with "/lib/shiva" as the primary loader it installs the specified patch, and the new rodata_string[] value "The Great Arcanum" is printed to stdout. When we run ./test_rodata without "/lib/shiva" it does not install the patch and the original string "Arcana Technologies" is printed.

Using shiva-ld to prelink the patch

The reality is that a user doesn't want to invoke "/lib/shiva" every single time they want to run a program that needs a patch. This is why "/lib/shiva" is designed to be an ELF interpreter like "ld-linux.so". We can use the Shiva prelinker "/bin/shiva-ld" to install the appropriate meta-data into the program we are patching:

- Update PT_INTERP with the path to "/lib/shiva" (Replacing "/lib/ld-linux.so").
- Add several custom entries to PT_DYNAMIC describing the patch basename: "ro_patch.o" and the module search path "/opt/shiva/modules".

The following command will install the Shiva interpreter path and patch information into the ./test_rodata binary.

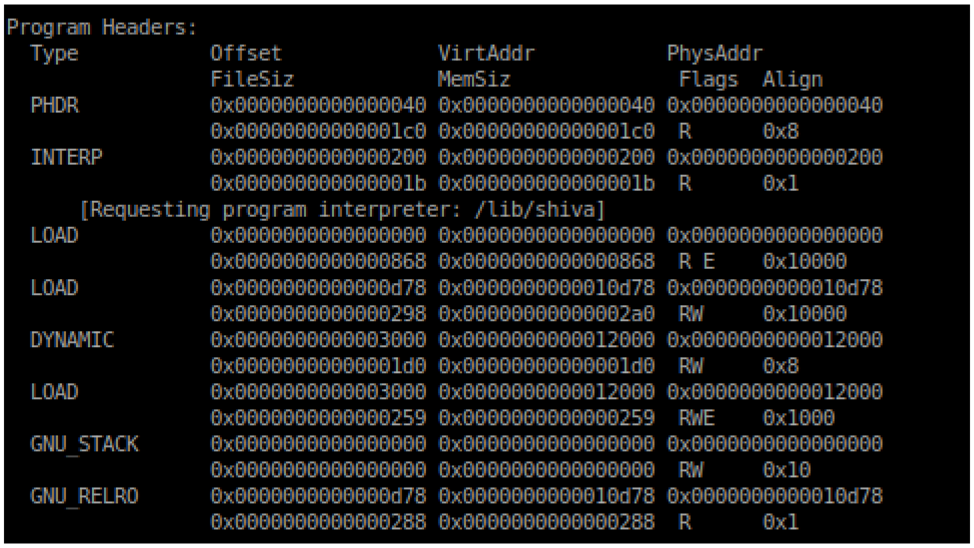
```
$ shiva-ld  -e test_rodata -p ro_patch.o -i /lib/shiva -s /opt/shiva/modules
```

This next command copies the patch file into the correct search path: "/opt/shiva/modules"

```
$ sudo cp ro_patch.o /opt/shiva/modules
```

Before we run ./test_rodata, let's take a quick look to see what modifications shiva-ld made to the executable:

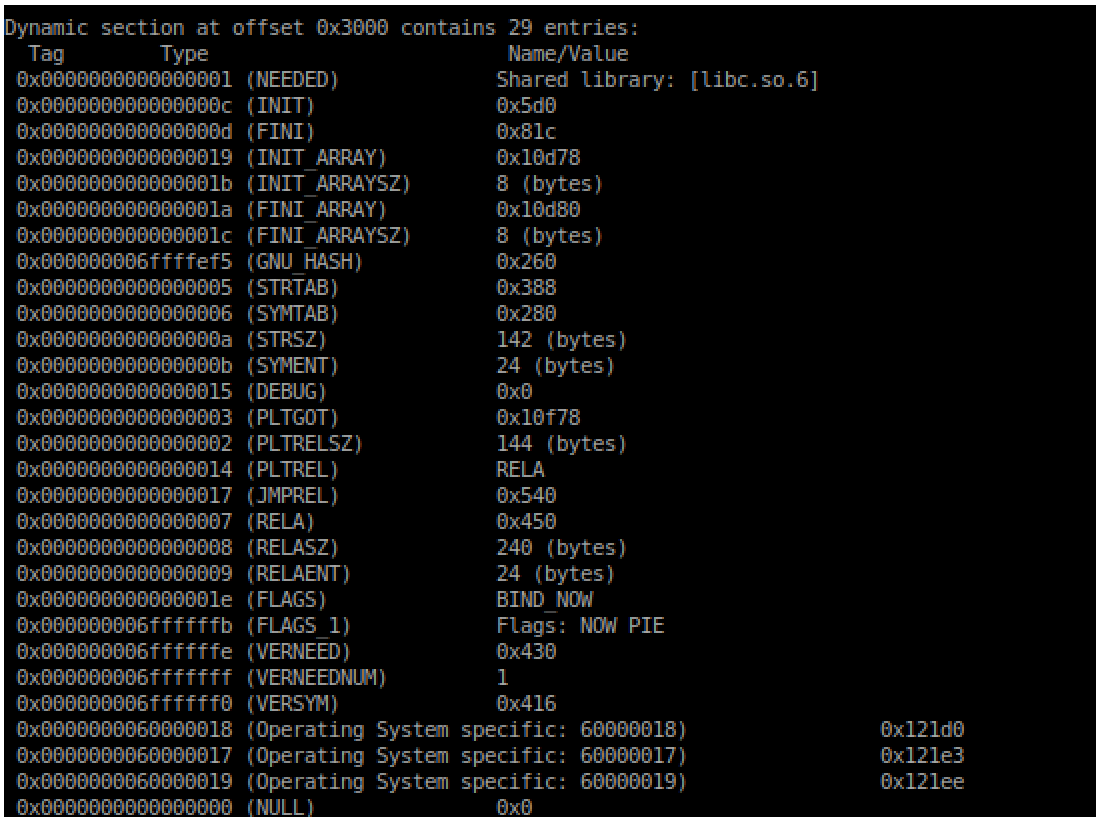
Image 2.2: Requesting program interpreter is now "/lib/shiva"



```
Program Headers:
  Type           Offset             VirtAddr           PhysAddr
                 FileSiz            MemSiz              Flags  Align
  PHDR           0x0000000000000040 0x0000000000000040 0x0000000000000040
                 0x00000000000001c0 0x00000000000001c0  R      0x8
  INTERP         0x0000000000000200 0x0000000000000200 0x0000000000000200
                 0x000000000000001b 0x000000000000001b  R      0x1
      [Requesting program interpreter: /lib/shiva]
  LOAD           0x0000000000000000 0x0000000000000000 0x0000000000000000
                 0x0000000000000868 0x0000000000000868  R E    0x10000
  LOAD           0x0000000000000d78 0x0000000000010d78 0x0000000000010d78
                 0x0000000000000298 0x00000000000002a0  RW     0x10000
  DYNAMIC        0x0000000000003000 0x0000000000012000 0x0000000000012000
                 0x00000000000001d0 0x00000000000001d0  RW     0x8
  LOAD           0x0000000000003000 0x0000000000012000 0x0000000000012000
                 0x0000000000000259 0x0000000000000259  RWE    0x1000
  GNU_STACK      0x0000000000000000 0x0000000000000000 0x0000000000000000
                 0x0000000000000000 0x0000000000000000  RW     0x10
  GNU_RELRO      0x0000000000000d78 0x0000000000010d78 0x0000000000010d78
                 0x0000000000000288 0x0000000000000288  R      0x1
```

In image 2.2 we can see that "/lib/shiva" is set as the new Interpreter path. The astute reader will also notice that there is a third PT_LOAD segment marked RWE. The shiva-ld tool creates a new dynamic segment (As shown below in Image 2.3). In order to make room for this new and larger dynamic segment shiva-ld creates a new PT_LOAD segment and updates the PT_DYNAMIC segment program header so that it points into the new PT_LOAD segment at 0x3000. The updated PT_DYNAMIC segment is moved and stored in this new segment

Image 2.3: New dynamic segment contains 3 custom entries



```
Dynamic section at offset 0x3000 contains 29 entries:
  Tag        Type                         Name/Value
 0x0000000000000001 (NEEDED)             Shared library: [libc.so.6]
 0x000000000000000c (INIT)               0x5d0
 0x000000000000000d (FINI)               0x81c
 0x0000000000000019 (INIT_ARRAY)         0x10d78
 0x000000000000001b (INIT_ARRAYSZ)       8 (bytes)
 0x000000000000001a (FINI_ARRAY)         0x10d80
 0x000000000000001c (FINI_ARRAYSZ)       8 (bytes)
 0x000000006ffffef5 (GNU_HASH)           0x260
 0x0000000000000005 (STRTAB)             0x388
 0x0000000000000006 (SYMTAB)             0x280
 0x000000000000000a (STRSZ)              142 (bytes)
 0x000000000000000b (SYMENT)             24 (bytes)
 0x0000000000000015 (DEBUG)              0x0
 0x0000000000000003 (PLTGOT)             0x10f78
 0x0000000000000002 (PLTRELSZ)           144 (bytes)
 0x0000000000000014 (PLTREL)             RELA
 0x0000000000000017 (JMPREL)             0x540
 0x0000000000000007 (RELA)               0x450
 0x0000000000000008 (RELASZ)             240 (bytes)
 0x0000000000000009 (RELAENT)            24 (bytes)
 0x000000000000001e (FLAGS)              BIND_NOW
 0x000000006ffffffb (FLAGS_1)            Flags: NOW PIE
 0x000000006ffffffe (VERNEED)            0x430
 0x000000006fffffff (VERNEEDNUM)         1
 0x000000006ffffff0 (VERSYM)             0x416
 0x0000000060000018 (Operating System specific: 60000018)                0x121d0
 0x0000000060000017 (Operating System specific: 60000017)                0x121e3
 0x0000000060000019 (Operating System specific: 60000019)                0x121ee
 0x0000000000000000 (NULL)               0x0
```

The above image 2.3 shows the dynamic segment of the ELF binary: test_rodata, after we ran the shiva-ld tool on it. Notice the three *Operating System specific* entries at the tail end. The readelf utility doesn't know how to parse these custom entries. Here is what the actual entry types and values are:

```
(SHIVA_NEEDED)          Patch object: [ro_patch.o]
(SHIVA_SEARCH)          Search path: [/opt/shiva/modules]
(SHIVA_ORIG_INTERP)     Original RTLD: [/lib/ld-linux-aarch64.so.1]
```

This data is needed by "/lib/shiva" at runtime so that it can locate and load the patch object, and the original dynamic linker (See Chained Linking in section x.x).

Final outcome of patching test_rodata

Now every time that the test_rodata executable is ran it will invoke the Shiva interpreter which will in turn load the patch object "/opt/shiva/modules/ro_patch.o".

```
$ ./test_rodata
rodata_string: The Great Arcanum
$
```

We can observe that the program test_rodata is now printing the patched version of *rodata_string []* when it is ran. This patch may appear permanent, but in effect the patch is actually being installed at runtime by "/lib/shiva" everytime the program runs.

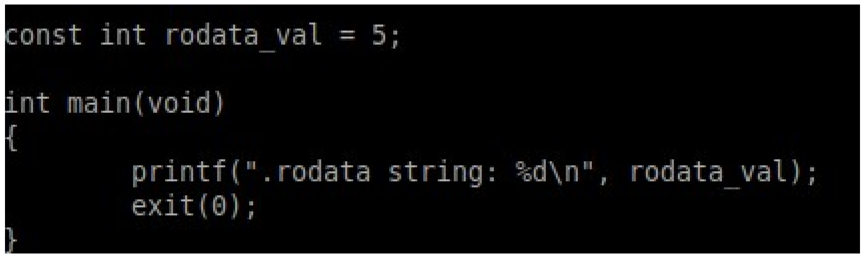
Current limitations of patching const data

There are patching scenarios with *const* data that fail with symbol interposition due to code optimizations with read-only data. In the previous example we illustrated how to patch a read-only global variable defined in C as:

const char rodata_string[] = “Arcana Technologies”;

The target program gets re-linked to use the patch version of *rodata_string*. Using symbol interposition to overwrite *.rodata* variables works perfectly unless the read-only value is optimized out of memory and stored only in a register. Let’s look at the following example:

Image 2.3: C code illustrating a read-only global variable

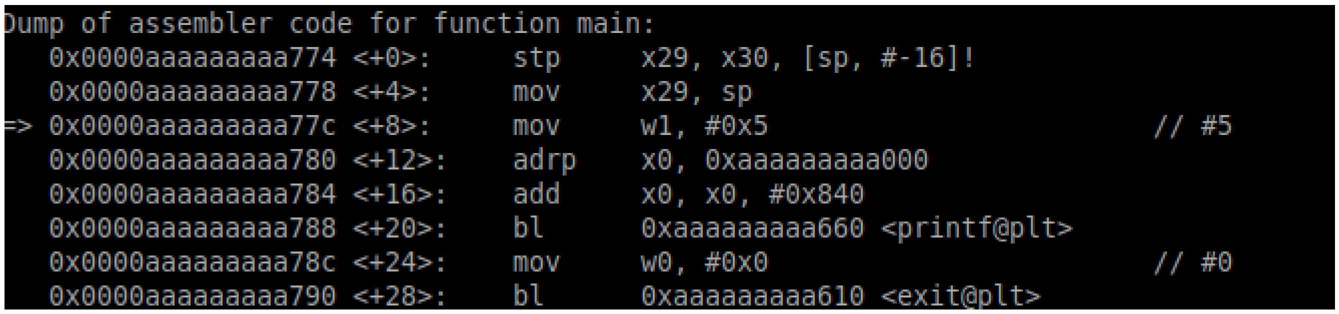


```
const int rodata_val = 5;

int main(void)
{
        printf(".rodata string: %d\n", rodata_val);
        exit(0);
}
```

After compiling the code above with gcc, even after all optimizations are disabled, you will get code similar to the following AArch64 assembly.

Image 2.4: Code optimization of read-only variable in aarch64 assembly



```
Dump of assembler code for function main:
   0x0000aaaaaaaaa774 <+0>:     stp     x29, x30, [sp, #-16]!
   0x0000aaaaaaaaa778 <+4>:     mov     x29, sp
=> 0x0000aaaaaaaaa77c <+8>:     mov     w1, #0x5                    // #5
   0x0000aaaaaaaaa780 <+12>:    adrp    x0, 0xaaaaaaaaa000
   0x0000aaaaaaaaa784 <+16>:    add     x0, x0, #0x840
   0x0000aaaaaaaaa788 <+20>:    bl      0xaaaaaaaaa660 <printf@plt>
   0x0000aaaaaaaaa78c <+24>:    mov     w0, #0x0                    // #0
   0x0000aaaaaaaaa790 <+28>:    bl      0xaaaaaaaaa610 <exit@plt>
```

In the code above we can see that the value 5 is being copied as an immediate value into a register. This is an optimization that removes the need for an adrp/add/ldr instruction trio, thus eliminating two instructions and the need for memory access. This optimization is inauspicious for symbol interposition based patching since the actual *.rodata* variables memory is not being accessed, and therefore there is no linking code to update. Instead there is only an instruction: **"mov w1, #0x5"**. Clearly this instruction can be re-encoded rather simply, but what would a patch look like for this? For the curious reader, move ahead to the section titled ***"ELF Transformations & Function splicing".***