#### **SEGURIDAD EN SISTEMAS OPERATIVOS**

# 4º Grado en Informática - Complementos de Ing. del Software Curso 2018-19

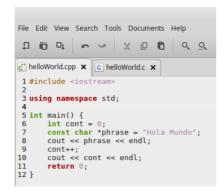
Práctica 2. Ingeniería inversa y vulnerabilidades

Sesión 1. El formato ELF (Executable and Linkable Format) en Linux

Autor<sup>1</sup>: Víctor García Carrera

#### Ejercicio 1.

Para conocer la estructura de los archivos ejecutables ELF y las herramientas básicas para ver su contenido, comenzamos creando un programa que imprime por pantalla un mensaje almacenado en una variable de tipo char\* *phrase* e imprime por pantalla el valor de una variable de tipo entero *cont* inicializada a 0 la cual aumentamos en 1. Desarrollamos dos versiones del mismo en dos lenguajes de programación diferentes, C++ y C, de nombres *helloWorld.cpp* y *helloWorld.c* respectivamente.



Compilamos como se muestra a continuación cada una de las versiones con los compiladores g++ y gcc, y vemos que la ejecución es la misma.

A la hora de observar la estructura y secciones de ambos ejectuables ELF, visualizamos con el comando *file* el tipo de archivos que son, ambos archivos ELF para una arquitectura de 64 bits y little endian (LSB) sin información de depuración (*not stripped*).

```
victor@VKCOMPUTRON ~/Documents/UGR/seguridadSO/p2/p2s1/c $ file HelloWorld
HelloWorld: ELF 64-bit LSB executable, x86-64, version 1 (SYSV), dynamically lin
ked, interpreter /lib64/ld-linux-x86-64, so. 2, for GNU/Linux 2.6.32, BuildID[shal
]=ae185bc2460a34f2848d610699e03ac94ba28059, not stripped
victor@VKCOMPUTRON ~/Documents/UGR/seguridadSO/p2/p2s1/c $ cd ..
victor@VKCOMPUTRON ~/Documents/UGR/seguridadSO/p2/p2s1 $ cd c++/
victor@VKCOMPUTRON ~/Documents/UGR/seguridadSO/p2/p2s1/c++ $ file HelloWorld++
HelloWorld++: ELF 64-bit LSB executable, x86-64, version 1 (SYSV), dynamically l
inked, interpreter /lib64/ld-linux-x86-64, so. 2, for GNU/Linux 2.6.32, BuildID[sh
al]=389dcd3b4f43dbf363d978a4d332c784dfba9741, not stripped
```

<sup>1</sup> Como autor declaro que los contenidos del presente documento son originales y elaborados por mi. De no cumplir con este compromiso, soy consciente de que, de acuerdo con la "Normativa de evaluación y de calificaciones de los estudiantes de la Universidad de Granada" esto "conllevará la calificación numérica de cero ... independientemente del resto de calificaciones que el estudiante hubiera obtenido ..."

Acto seguido, visualizamos con el comando *readelf -h* la cabecera de los archivos ELF, de nuevo sin cambios remarcables entre ambas versiones a excepción de la dirección de inicio de las cabeceras de sección, superior en la versión de C++:

```
File Edit View Search Terminal Help
victor@VKCOMPUTRON ~/Documents/UGR/seguridadS0/p2/p2s1/c $ readelf -h HelloWorld
ELF Header:
                                                         :
7f 45 4c 46 02 01 01 00 00 00 00 00 00 00 00
ELF64
2's complement, little endian
1 (current)
UNIX - System V
           ABI Version:
                                                                                                                                                                                                         EXEC (Executable file)
Advanced Micro Devices X86-64
0x1
0x400470
                                                                                                                                                                                                         64 (bytes into file)
6680 (bytes into file)
        Start of Section headers: 9x8
Size of this header: 64
Size of program headers: 56
Number of program headers: 9
Size of section headers: 64
Number of section headers: 31
Section header string table index: 28
Section header string table index: 28
                                                                                                                                                                                                            9
64 (bytes)
                                                                 PPUTRON -/Documents/UGR/seguridadSO/p2/p2s1/c $ cd ..
MPUTRON -/Documents/UGR/seguridadSO/p2/p2s1 $ cd c++/
MPUTRON -/Documents/UGR/seguridadSO/p2/p2s1/c++ $ readelf -h HelloWorld++
                                                        :
7f 45 4c 46 02 01 01 00 00 00 00 00 00 00 00
ELF64
2's complement, little endian
1 (current)
UNIX - System V
                                                                                                                                                                                                         0
EXEC (Executable file)
Advanced Micro Devices X86-64
           ABI Version:
           Type:
Machine:
Version:
                                                                                                                                                                                                         0x1
0x1
0x4007a0
64 (bytes into file)
7288 (bytes into file)
           Entry point address:
Start of program headers:
Start of section headers:
       Start of Program headers: 56 Number of program headers: 9 Size of section headers: 64 Number of section headers: 31 Section header string table index: 28 Start of St
                                                                                                                                                                                                         64 (bytes)
56 (bytes)
9
64 (bytes)
                                                                                                                                                                                                             guridadSO/p2/p2s1/c++ $
```

A continuación, vamos a observar las secciones presentes en los archivos ELF, comentando algunas de ellas y comparando las dos versiones.

Una sección de interés es la de .interp pues almacena la ruta al intérprete utilizado (basicamente su nombre), en este caso /lib64/ld-linux-x86-64-.so.2.

Otra sección interesante es .got, la cual contiene la dirección de la Global Offset Table o GOT. Esta tabla de desplazamientos globales se utiliza para garantizar el funcionamiento correcto de aquel código que sea independiente de su direccionamiento (position-independent code o PIC). El código PIC no puede por lo general almacenar direcciones virtuales absolutas. Esta tabla GOT almacena direcciones absolutas en datos privados para permitir que cuando se utilizan referencias independientes de la posición, éstas sean traducidas a direcciones absolutas. La sección .got.ptl contiene aquella porción de la tabla GOT que es de sólo lectura o read-only.

La sección .ctors contiene una lista de punteros a función de los constructores globales, mientras que la sección .dtors contiene una lista de punteros a función de los destructores globales.

Al listar las secciones de ambas versiones, podemos comprobar que **no existe ninguna diferencia relevante, excepto las direcciones** de muchas de las secciones excepto *.interp, .note.ABI-tag, .note.gnu.build-i, .gnu.hash, .got* y *.got.plt*, además de un valor distino del campo Info o Align en las secciones *.gnu.version\_r* o *.bss.* 

	@VKCOMPUTRON ~/Doc are 31 section hea			readelf -S	HelloWorld
	n Headers: Name	Туре	Address Openione	0ffset	
[ 0]	Size 0000000000000000000	EntSize NULL 00000000000000000000	Flags Link Info 000000000000000000 0 0	00000000	
[ 1]	.interp 0000000000000000001c	PROGBITS 0000000000000000000	0000000000400238 A 0 0	00000238	
Thumber	.note.ABI-tag 000000000000000020	NOTE 000000000000000000000000000000000000	00000000000400254 A 0 0 00000000000400274	00000254 4 00000274	
[4]	.note.gnu.build-i 00000000000000024 .gnu.hash	00000000000000000000000000000000000000	A 0 0	4 00000298	
[ 5]	0000000000000001c .dynsym	00000000000000000000000000000000000000	A 5 0	8 000002b8	
[ 6]	00000000000000078 .dynstr 000000000000000044	00000000000000018 STRTAB 0000000000000000000	A 6 1 0000000000400330 A 0 0	8 00000330 1	
Danie I and	.gnu.version 00000000000000000	VERSYM 0000000000000000002	0000000000400374 A 5 0	00000374	
	.gnu.version_r 0000000000000000000000 .rela.dyn	VERNEED 00000000000000000000000000000000000	0000000000400380 A 6 1 000000000004003a0	00000380 8 000003a0	
	00000000000000018 .rela.plt	00000000000000018 RELA	A 5 0	8dE00000	
[11]	00000000000000048 .init 000000000000000001a	00000000000000018 PROGBITS 0000000000000000000	AI 5 24 00000000000400400 AX 0 0	8 00000400 4	
	.plt 000000000000000040	PROGBITS 0000000000000000000010	0000000000400420 AX 0 0	00000420 16	
[13]	.plt.got 000000000000000000 .text	PROGBITS 00000000000000000000 PROGBITS	0000000000400460 AX 0 0 0000000000400470	00000460 8 00000470	
	00000000000001b2 .fini	00000000000000000 PROGBITS	AX 0 0 00000000000400624	16 00000624	
[16]	00000000000000000000000000000000000000	00000000000000000 PROGBITS 0000000000000000000	AX 0 0 0000000000400630 A 0 0	4 00000630 4	
	.eh_frame_hdr 00000000000000034	PROGBITS 0000000000000000000	0000000000400644 A 0 0	00000644 4	
	.eh_frame 000000000000000f4 .init_array	PROGBITS 0000000000000000000 INIT_ARRAY	00000000000400678 A 0 0 00000000000600e10	00000678 8 00000e10	

[19]	.init array	INIT ARRAY	00000000000600e10	00000e10
	800000000000000	000000000000000	WA 0 0	= 6 8/man/h
[20]	.fini array	FINI ARRAY	0000000000600e18	00000e18
	800000000000000	0000000000000000	WA 0 0	8
[21]	.jcr	PROGBITS	0000000000600e20	00000e20
	8000000000000008	0000000000000000	WA 0 0	mbre8
[22]	.dynamic	DYNAMIC	0000000000600e28	00000e28
	00000000000001d0	0000000000000010	WA 6 0	8
[23]	.got	PROGBITS	0000000000600ff8	00000ff8
	800000000000000	800000000000000	WA 0 0	ote 8.BI-tag
[24]	.got.plt	PROGBITS	0000000000601000	00001000
	0000000000000000	800000000000000	WA 0 0	
[25]	.data	PROGBITS	0000000000601030	00001030
	00000000000000010	0000000000000000	WA 0 0	8
[26]	.bss	NOBITS	0000000000601040	00001040
	800000000000000	0000000000000000	WA 0 0	
[27]	.comment	PROGBITS	0000000000000000	00001040
	0000000000000035	00000000000000001	MS 0 0	
[28]	.shstrtab	STRTAB	0000000000000000	00001905
	000000000000010c	0000000000000000	0 0	1
[29]	.symtab	SYMTAB	0000000000000000	00001078
	0000000000000660	0000000000000018	30 47	8
[30]	.strtab	STRTAB	0000000000000000	000016d8
	000000000000022d	0000000000000000	0 0	1
ey to	Flags:			
W (wi	rite), A (alloc), :	X (execute), M (me	erge), S (strings)	, l (large)
I (in	nfo), L (link orde	r), G (group), T	(TLS), E (exclude)	, x (unknown)
	ctra OS processing			
	VKCOMPUTRON ~/Doc			

			uridadSO/p2/p2s1/c++	<pre>\$ readelf</pre>	-S Hello	World++
There	are 31 section head	ders, starting	at offset 0x1c78:			
Coctio	n Headers:					
	Name	Туре	Address Opcion	Offset		
[ [NI ]	Size	EntSize	Flags Link Inf			
[ 01		NULL	00000000000000000	00000000		
	0000000000000000	00000000000000		0 0		
[ 1]	.interp	PROGBITS	0000000000400238	00000238		
- 0	000000000000001c	00000000000000		0 1		
[ 2]	.note.ABI-tag	NOTE	0000000000400254	00000254		
	00000000000000020	00000000000000		0 4		
[ 3]	.note.gnu.build-i 000000000000000024	NOTE 0000000000000000	0000000000400274	00000274 0 4		
T 41	.qnu.hash	GNU HASH	0000000000400298	00000298		
1 41	00000000000000000	00000000000000		0 8		
[ 51	.dynsym	DYNSYM	00000000004002c8	000002c8		
Revie	00000000000000150	00000000000000	018 A 6	1 8		
[ 6]	.dynstr	STRTAB	0000000000400418	00000418		
	0000000000000172	00000000000000	000 A 0	0 1		
[7]	.gnu.version	VERSYM	000000000040058a	0000058a		
Benles	0000000000000001c	00000000000000		0 2		
[8]	.gnu.version_r	VERNEED	00000000004005a8	000005a8 2 8		
F 0.1	00000000000000040 .rela.dyn	00000000000000000000000000000000000000	000 A 6 00000000004005e8	2 8 000005e8		
[ 9]	000000000000000000000000000000000000000	000000000000000		0 8		
[10]	.rela.plt	RELA	0000000000400618			
	000000000000000000000	00000000000000				
[11]	.init	PROGBITS	00000000004006d8	000006d8		
	000000000000001a	00000000000000	000 AX 000	9 4		
[12]	.plt	PROGBITS	0000000000400700	00000700		
	0000000000000000	00000000000000		0 16		
[13]	.plt.got	PROGBITS	0000000000400790	00000790 9 8		
[14]	00000000000000008 .text	00000000000000000000000000000000000000	000 AX 0 00000000004007a0	0 8 000007a0		
[14]	00000000000000222	00000000000000		n 16		
[15]	.fini	PROGBITS	00000000004009c4	000009c4		
,,,,,	00000000000000000	00000000000000		0 4		
[16]	.rodata	PROGBITS	00000000004009d0	000009d0		
	000000000000000f	00000000000000	000 A 0	0 4		
[17]	.eh_frame_hdr	PROGBITS	0000000004009e0	000009e0		
	0000000000000044	00000000000000		0 4		
[18]	.eh_frame	PROGBITS	0000000000400a28	00000a28		
[10]	0000000000000134	00000000000000000000000000000000000000	000 A 0 0000000000600df8	0 8 00000df8		
[19]	.init_array	INIT_ARRAY	810000000000000	8100000		

	.init_array 000000000000000010			Θ	8	
[20]	.fini_array		000000000066	00e08	00000e08	
	800000000000000	0000000000000000	WA 0		8	
[21]	.jcr	PROGBITS	000000000066	00e10	00000e10	
	800000000000000	0000000000000000	WA 0	0	mbre8	
[22]	.dynamic	DYNAMIC	000000000066	00e18	00000e18	
	0000000000001e0	000000000000000000000000000000000000000				
[23]	.got	PROGBITS	000000000066	0ff8	00000ff8	
	800000000000000	8000000000000000	WA 0	2 0	ote 8 BI-	
[24]	.got.plt	PROGBITS	000000000066			
	0000000000000058	8000000000000000			8	
[25]	.data	PROGBITS	000000000066	1058	00001058	
	0000000000000010					
[26]	.bss	NOBITS	000000000066	1080	00001068	
	000000000000118				32	
[27]	.comment	PROGBITS			00001068	
	0000000000000035					
[28]	.shstrtab	STRTAB	00000000000	0000	00001b6c	
	00000000000010c		0	0	1	
[29]	.symtab	SYMTAB	00000000000		000010a0	
	0000000000000738		30	50	8	
[30]		STRTAB	00000000000		000017d8	
	0000000000000394	0000000000000000			1	
	Flags:					
	rite), A (alloc), )					
	nfo), L (link orde					
	ktra OS processing @VKCOMPUTRON ~/Docu					ific)

Finalmente, con el comando *readelf -r* podemos ver las secciones de reubicación de nuestros archivos ELF. Los binarios ELF contienen una tabla de reubicación que especifica las reubicaciones necesarias para la ejecución del programa. Estas reubicaciones suelen corresponder a símbolos, donde el enlazador dinámico resuelve un símbolo por su nombre y escribe su dirección en el lugar indicado en su entrada de reubicación para poder acceder posteriormente a él. Los tipos de reubicaciones dependen de la arquitectura, siendo las más comunes *R\_386\_COPY*, que indica básicamente "sólo copia la dirección del símbolo en esa dirección", y *R\_386\_JUMP\_SLOT*, utilizado por la llamada a la función PLT/GOT.

Este último tipo suele aparecer con símbolos como *printf, strcpy, getpid...*, reflejado con el caso particular de *printf* en la sección de reubicación *rela.plt* de la versión del programa en C, que utiliza la llamada a esta función para imprimir por pantalla.

El contenido de las secciones de reubicación es, por ende, aquellos símbolos o palabras empleados en el programa junto con sus entradas de reubicación necesarias (con valores de desplazamiento entre otros) para obtener sus direcciones y poder acceder a ellas, permitiendo la correcta ejecución del programa.

### Ejercicio 2.

En este ejercicio, investigamos el comando *objdump*. Leyendo su entrada en el manual (*man objdump*), nos informa que se trata de un comando que muestra información acerca de un fichero, configurando qué información mostrar en función de diversos parámetros. Su diferencia fundamental con *readelf* es que, mientras que ésta vuelca toda la información de un fichero ELF, *objdump* vuelca información genérica acerca de casi cualquier tipo de archivo (incluyendo ELF). A continuación los diversos parámetros posibles a utilizar con *objdump* para **obtener información similar a la obtenida con** *readelf*, junto con otras opciones que nos proporcionan información valiosa desde el punto de vista de la ingeniería inversa:

La opción -a muestra la cabecera del archivo.

Si queremos volcar el contenido de las secciones, podemos especificar la opción -adjust-vma=offset para establecer el offset a sumar a cada una de las secciones (a su dirección de inicio).

La opción -C permite decodificar símbolos de bajo nivel a nombres de alto nivel, útil para hacer legibles los nombres de las funciones en C++.

La opción -g muestra información de depuración. Intenta parsear la información de debug en formato STABS o IEEE almacenado en el archivo, o si falla trata de imprimir cualquier información DWARF del mismo.

La opción -d está pensada para desensamblar aquellas secciones que se espera que contengan instrucciones. La opción -D lo hace de todas las secciones.

La opción -f muestra información acerca de la cabecera general del archivo.

La opción -h muestra información acerca de las cabeceras de secciones del archivo.

La opción -i name muestra información sólo de la sección de nombre name.

La opción -r imprime las entradas de reubicación. -R las de reubicación dinámica.

La opción -s muestra el contenido completo de cualquier sección especificada (si no se especifica ninguna, lo hace de todas las secciones no vacías).

La opción -S source muestra código fuente mezclado con ensamblador si es posible.

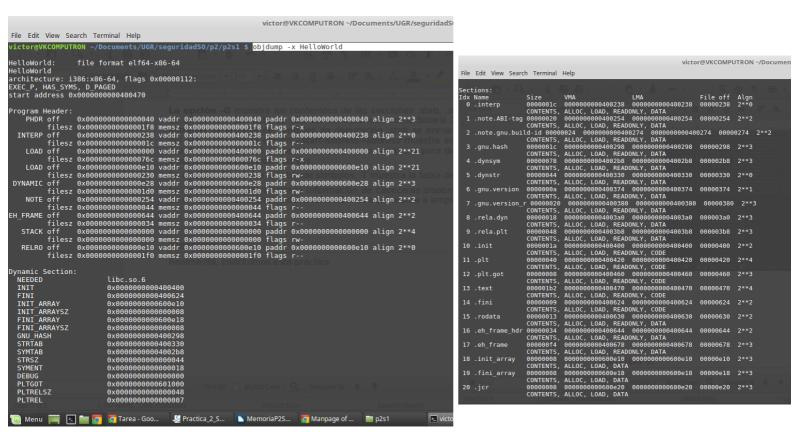
La opción -S --show-raw-ins permite, a la hora de desensamblar instrucciones, imprimirla tanto en hexadecimal como en su forma simbólica.

La opción -W muestra el contenido de las secciones de debug del fichero.

**La opción -G** muestra los contenidos de las secciones .stab, .stab.index y .stab.excl de un fichero ELF. Esto es sólo útil en sistemas (como Solaris 2.0) en los cuales las entradas de la tabla de símbolos de depuración .stab se encuentran en una sección del ELF. Junto con la opción --start-address=address muestra información a partir de esa dirección address. La opción --stop-address=address para de mostrar información a partir de esa dirección address.

La opción -t muestra la tabla de símbolos. -T muestra la tabla de símbolos dinámicos.

**La opción -x** muestra toda la información de cabeceras disponible, incluyendo tabla de símbolos y las entradas de reubicación. Es equivalente a emplear -a -f -h -r -t.



## Ejercicio 3.

Finalmente, concluimos esta práctica