Tutorial and User Manual

Abstract

ELFIO is a header-only C++ library intended for reading and generating files in the ELF binary format

Serge Lamikhov-Center

to_serge@users.sourceforge.net

1 TABLE OF CONTENTS

<u>2</u>	INTRODUCTION	2
<u>3</u>	GETTING STARTED WITH ELFIO	2
3.1	ELF FILE READER	2
3.2	ELF SECTION DATA ACCESSORS	5
3.3	ELFDUMP UTILITY	6
3.4	ELF FILE WRITER	6
<u>4</u>	ELFIO LIBRARY CLASSES	10
4.1	ELFIO	10
4.2	SECTION	12
4.3	SEGMENT	14
4.4	STRING_SECTION_ACCESSOR	15
4.5	SYMBOL_SECTION_ACCESSOR	15
4.6	RELOCATION_SECTION_ACCESSOR	17
4.7	NOTE_SECTION_ACCESSOR	18

2 Introduction

ELFIO is a header-only C++ library intended for reading and generating files in the ELF binary format. It is used as a standalone library - it is not dependant on any other product or project. Adhering to ISO C++, it compiles on a wide variety of architectures and compilers.

While the library is easy to use, some basic knowledge of the ELF binary format is required. Such Information can easily be found on the Web.

The full text of this tutorial comes together with ELFIO library distribution

3 GETTING STARTED WITH ELFIO

3.1 ELF FILE READER

The ELFIO library is just normal C++ header files. In order to use all its classes and types, simply include the main header file "elfio.hpp". All ELFIO library declarations reside in a namespace called "ELFIO". This can be seen in the following example:

```
#include <iostream>
#include <elfio/elfio.hpp>

using namespace ELFIO

int main( int argc, char** argv )
{
   if ( argc != 2 ) {
      std::cout << "Usage: tutorial <elf_file>" << std::endl;
   return 1;
}</pre>
```

- 1 Include elfio.hpp header file
- 2 The ELFIO namespace usage

This section of the tutorial will explain how to work with the reader portion of the ELFIO library.

The first step would be creating an elfio class instance. The elfio constructor has no parameters. The creation is normally followed by invoking the 'load' member method, passing it an ELF file name as a parameter:

```
// Create elfio reader
elfio reader;

// Load ELF data
if ( !reader.load( argv[1] ) ) {
    std::cout << "Can't find or process ELF file " << argv[1] << std::endl;
    return 2;
}</pre>
```

- **1** Create elfio class instance
- **2** Initialize the instance by loading ELF file. The function load returns 'true' if the ELF file was found and processed successfully. It returns 'false' otherwise

The load() method returns 'true' if the corresponding file was found and processed successfully.

All the ELF file header properties such as encoding, machine type and entry point are accessible now. To get the class and the encoding of the file use:

```
// Print ELF file properties
std::cout << "ELF file class : ";
if ( reader.get_class() == ELFCLASS32 )
    std::cout << "ELF32" << std::endl;
else
    std::cout << "ELF64" << std::endl;

std::cout << "ELF file encoding : ";
if ( reader.get_encoding() == ELFDATA2LSB )
    std::cout << "Little endian" << std::endl;
else
    std::cout << "Big endian" << std::endl;</pre>
```

- **1** Member function get_class() returns ELF file class. Possible return values are: ELFCLASS32 or ELFCLASS64
- ② Member function get_encoding() returns ELF file format encoding. Possible values are: ELFDATA2LSB or ELFDATA2MSB standing for little- and big-endianess correspondingly

Note:

Standard ELF types, flags and constants are defined in the elf_types.hpp header file. This file is included automatically into the project. For example: ELFCLASS32, ELFCLASS64 constants define values for 32/64 bit architectures. Constants ELFDATA2LSB and ELFDATA2MSB define values for little- and big-endian encoding.

ELF binary files consist of sections and segments. Each section has its own responsibility: some contains executable code, others –program's data, some are symbol tables and so on. See ELF binary format documentation for purpose and content description of sections and segments.

The following code demonstrates how to find out the amount of sections the ELF file contains. The code also presents how to access particular section properties like names and sizes:

- Retrieve the number of sections
- Use operator[] to access a section by its number or symbolic name.
- **3** get_name(), get_size() and get_data() are member functions of 'section' class

The 'sections' data member of ELFIO's 'reader' object permits obtaining the number of sections inside a given ELF file. It also serves for getting access to individual sections by using operator[], which returns a pointer to the corresponding section's interface.

Similarly, for executables, the segments of the ELF file can be processed:

```
// Print ELF file segments info
                                                    0
Elf Half seg num = reader.segments.size();
std::cout << "Number of segments: " << seg num << std::endl;</pre>
for ( int i = 0; i < seg_num; ++i ) {
                                                    0
    const segment* pseg = reader.segments[i];
    std::cout << " [" << i << "] 0x" << std::hex
                                                    0
              << pseg->get flags()
              << "\t0x"
              << pseg->get_virtual_address()
              << "\t0x"
              << pseg->get file size()
              << "\t0x"
              << pseg->get memory size()
              << std::endl;
    // Access segments's data
    const char* p = reader.segments[i]->get data(); 6
```

- Retrieve the number of segments
- Use operator[] to access a segment by its number
- **3** get_flags(), get_virtual_address(), get_file_size(), get_memory_size() and get_data() are member methods of 'segment' class

In this case, the segments' attributes and data are obtained by using the 'segments' data member of ELFIO's 'reader' class.

3.2 ELF Section Data Accessors

To simplify creation and interpretation of specific ELF sections, the ELFIO library provides accessor classes. Currently, the following classes are available:

- String section accessor
- Symbol section accessor
- Relocation section accessor
- Note section accessor
- Dynamic section accessor

More accessors may be implemented in future versions of the library.

Let's see how the accessors can be used with the previous ELF file reader example. The following example prints out all symbols in a section that turns out to be a symbol section:

- Check section's type
- Build symbol section accessor
- 6 Get the number of symbols by using the symbol section accessor
- Get particular symbol properties its name, value, etc.

First we create a 'symbol_section_accessor' class instance. Usually, accessors's constructors receive references to both the elfio and a 'section' objects as parameters. The get_symbol() method is used for retrieving particular entries in the symbol table.

3.3 ELFDUMP UTILITY

The source code for the ELF Dump Utility can be found in the "examples" directory. It heavily relies on dump facilities provided by the auxiliary header file <elfio_dump.hpp>. This header file demonstrates more accessor's usage examples.

3.4 ELF FILE WRITER

In this chapter we will create a simple ELF executable file that prints out the classical "Hello, World!" message. The executable will be created and run on i386 Linux OS platform. It is supposed to run well on both 32 and 64-bit Linux platforms. The file will be created without invoking the compiler or assembler tools in the usual way (i.e. translating high level source code that makes use of the standard library functions). Instead, using the ELFIO writer, all the necessary sections and segments of the file will be created and filled explicitly, each, with its appropriate data. The physical file would then be created by the ELFIO library.

Before starting, two implementation choices of elfio that users should be aware of are:

- 1. The ELF standard does not require that executables will contain any ELF sections only presence of ELF segments is mandatory. The elfic library, however, requires that all data will belong to sections. It means that in order to put data in a segment, a section should be created first. Sections are associated with segments by invoking the segment's member function add section index().
- 2. The elfio writer class, while constructing, creates a string table section automatically.

Our usage of the library API will consist of several steps:

- Creating an empty elfio object
- Setting-up ELF file properties
- Creating code section and data content for it
- Creating data section and its content
- Addition of both sections to corresponding ELF file segments
- Setting-up the program's entry point
- Dumping the elfio object to an executable ELF file

```
#include <elfio/elfio.hpp>
using namespace ELFIO;
int main ( void )
   elfio writer;
   writer.create( ELFCLASS32, ELFDATA2LSB );
   writer.set os abi( ELFOSABI LINUX );
    writer.set type( ET EXEC );
    writer.set machine( EM 386 );
   section* text sec = writer.sections.add( ".text" );
   text sec->set type( SHT PROGBITS );
   text sec->set flags( SHF ALLOC | SHF EXECINSTR );
    text sec->set addr align( 0x10 );
   // mov eax, 4
                                                             // mov ebx, 1
                                                             // mov ecx, msg
                                                             // mov edx, 14
                    '\xCD', '\x80',
                                                              // int 0x80
                    '\xB8', '\x01', '\x00', '\x00', '\x00',
                                                             // mov eax, 1
                    '\xCD', '\x80' };
                                                              // int 0x80
                                                              0
    text_sec->set_data( text, sizeof( text ) );
                                                              0
    segment* text seg = writer.segments.add();
                                                              0
    text_seg->set_type( PT LOAD );
   text_seg->set_virtual_address( 0x08048000 );
   text_seg->set_physical_address( 0x08048000 );
    text_seg->set_flags( PF_X | PF_R );
    text_seg->set_align( 0x1000 );
    text_seg->add_section_index( text_sec->get_index(),
                                 text sec->get addr align() );
    section* data sec = writer.sections.add( ".data" );
    data_sec->set_type( SHT_PROGBITS );
    data_sec->set_flags( SHF_ALLOC | SHF_WRITE );
    data sec->set addr align( 0x4 );
    char data[] = { '\x48', '\x65', '\x6C', '\x6C', '\x6F', '\x2C', '\x2O', '\x57', '\x6F', '\x72',
                                                              // "Hello, World!\n"
                    '\x6C', '\x64', '\x21', '\x0A' };
                                                              0
    data sec->set data( data, sizeof( data ) );
                                                              0
    segment* data_seg = writer.segments.add();
                                                              0
    data seg->set type( PT LOAD );
    data_seg->set_virtual_address( 0x08048020 );
    data_seg->set_physical_address( 0x08048020 );
    data_seg->set_flags( PF_W | PF_R );
    data_seg->set_align( 0x10 );
    data seg->add section index( data sec->get index(),
                                 data_sec->get_addr_align() );
                                                              0
    writer.set_entry( 0x08048000 );
                                                              o
    writer.save( "hello i386 32" );
    return 0;
```

- Initialize empty 'elfio' object. This should be done as the first step when creating a new 'elfio' object as other API is relying on parameters provided ELF file 32-bits/64-bits and little/big endianness
- ② Other attributes of the file. Linux OS loader does not require full set of the attributes, but they are provided when a regular linker used for creation of ELF files
- **3** Create a new section, set section's attributes. Section type, flags and alignment have a big significance and controls how this section is treated by a linker or OS loader
 - 4 Add section's data
 - 6 Create new segment
 - **6** Set attributes and properties for the segment
 - **7** Associate a section with segment containing it
 - 8 Setup entry point for your program
 - Oreate ELF binary file on disk

Let's compile the example above (put into a source file named 'writer.cpp') into an executable file (named 'writer'). Invoking 'writer' will create the executable file "hello_i386_32" that prints the "Hello, World!" message. We'll change the permission attributes of this file, and finally, run it:

```
> ls
writer.cpp
> g++ writer.cpp -o writer
> ls
writer writer.cpp
> ./writer
> ls
hello_i386_32 writer writer.cpp
> chmod +x ./hello_i386_32
> ./hello_i386_32
Hello, World!
```

In case you already compiled the 'elfdump' utility, you can inspect the properties of the produced executable file (the '.note' section was not discussed in this tutorial, but it is produced by the sample file writer.cpp located in the 'examples' folder of the library distribution):

```
./elfdump hello i386 32
ELF Header
  Class: ELF32
Encoding: Little endian
  ELFVersion: Current
  Type: Executable file Machine: Intel 80386
Version: Current Entry: 0x8048000
Flags: 0x0
[ Nr ] Type Addr Size ES Flg Lk Inf Al Name [ 0] NULL 00000000 00000000 00 0 0 0 0 0 [ 1] STRTAB 00000000 0000001d 00 0 0 0 0 shstr [ 2] PROGBITS 08048000 0000001d 00 AX 0 0 16 .text [ 3] PROGBITS 08048020 00000000 00 WA 0 0 4 .data [ 4] NOTE 00000000 00000044 00 0 0 1 .note
Section Headers:
                                                                         0 0 0 .shstrtab
Key to Flags: W (write), A (alloc), X (execute)
Segment headers:
Note section (.note)
    No Type
                    Name
   [ 0] 00000001 Created by ELFIO
 [ 1] 00000001 Never easier!
```

Note:

The elfic library takes care of the resulting binary file layout calculation. It does this on base of the provided memory image addresses and sizes. It is the user's responsibility to provide correct values for these parameters. Please refer to your OS (other execution environment or loader) manual for specific requirements related to executable ELF file attributes and/or mapping.

Similarly to the 'reader' example, you may use provided accessor classes to interpret and modify content of section's data.

4 ELFIO LIBRARY CLASSES

This section contains detailed description of classes provided by elfio library

4.1 ELFIO

4.1.1 Data members

The ELFIO library's main class is 'elfio'. The class contains two public data members:

Data member	Description
sections	The container stores ELFIO library section instances. Implements operator[], add() and size(). operator[] permits access to individual ELF file section according to its index.
segments	The container stores ELFIO library segment instances. Implements operator[], add() and size(). operator[] permits access to individual ELF file segment according to its index.

4.1.2 Member functions

Here is the list of elfio public member functions. The functions permit to retrieve or set ELF file properties.

Member Function	Description
elfio()	The constructor.
~elfio()	The destructor.
<pre>void create(unsigned char file_class, unsigned char encoding)</pre>	Cleans and/or initializes elfio object. file_class is either ELFCLASS32 or ELFCLASS64. file_class is either ELFDATA2LSB or ELFDATA2MSB.
<pre>bool load(const std::string& file_name)</pre>	Initializes elfio object by loading data from ELF binary file. File name provided in file_name. Returns true if the file was processed successfully.
<pre>bool save(const std::string& file_name)</pre>	Creates a file in ELF binary format. File name provided in file_name. Returns true if the file was created successfully.

<pre>unsigned char get_class()</pre>	Returns ELF file class. Possible values are ELFCLASS32 or ELFCLASS64.	
<pre>unsigned char get_elf_version()</pre>	Returns ELF file format version.	
unsigned char get_encoding()	Returns ELF file format encoding. Possible values are ELFDATA2LSB and ELFDATA2MSB.	
<pre>Elf_Word get_version()</pre>	Identifies the object file version.	
<pre>Elf_Half get_header_size()</pre>	Returns the ELF header's size in bytes.	
<pre>Elf_Half get_section_entry_size()</pre>	Returns a section's entry size in ELF file header section table.	
<pre>Elf_Half get_segment_entry_size()</pre>	Returns a segment's entry size in ELF file header program table.	
unsigned char get_os_abi()	Returns operating system ABI identification.	
<pre>void set_os_abi(unsigned char value)</pre>	Sets operating system ABI identification.	
<pre>unsigned char get_abi_version();</pre>	Returns ABI version.	
<pre>void set_abi_version(unsigned char value)</pre>	Sets ABI version.	
<pre>Elf_Half get_type()</pre>	Returns the object file type.	
void set_type(Elf_Half value)	Sets the object file type.	
<pre>Elf_Half get_machine()</pre>	Returns the object file's architecture.	
void set_machine(Elf_Half value)	Sets the object file's architecture.	
Elf_Word get_flags ()	Returns processor-specific flags associated with the file.	
<pre>void set_flags(Elf_Word value)</pre>	Sets processor-specific flags associated with the file.	

Elf64_Addr get_entry()	Returns the virtual address to which the system first transfers control.
<pre>void set_entry(Elf64_Addr value)</pre>	Sets the virtual address to which the system first transfers control.
Elf64_Off get_sections_offset()	Returns the section header table's file offset in bytes.
<pre>void set_sections_offset(Elf64_Off value)</pre>	Sets the section header table's file offset. Attention! The value can be overridden by the library, when it creates new ELF file layout.
<pre>Elf64_Off get_segments_offset()</pre>	Returns the program header table's file offset.
<pre>void set_segments_offset(Elf64_Off value)</pre>	Sets the program header table's file offset. Attention! The value can be overridden by the library, when it creates new ELF file layout.
<pre>Elf_Half get_section_name_str_index()</pre>	Returns the section header table index of the entry associated with the section name string table.
<pre>void set_section_name_str_index(Elf_Half value)</pre>	Sets the section header table index of the entry associated with the section name string table.
<pre>endianess_convertor& get_convertor()</pre>	Returns endianess convertor reference for the specific elfic object instance.
<pre>Elf_Xword get_default_entry_size(Elf_Word section_type)</pre>	Returns default entry size for known section types having different values on 32 and 64 bit architectures. At the moment, only SHT_RELA, SHT_REL, SHT_SYMTAB and SHT_DYNAMIC are 'known' section types. The function returns 0 for other section types.

4.2 SECTION

Class 'section' has no public data members.

4.2.1 Member functions

 ${\tt section}$ public member functions listed in the table below. These functions permit to retrieve or set ELF file section properties

Member Fund	tion Description
section()	The default constructor. No section class instances are created manually. Usually, 'add' method is used for 'sections' data member of 'elfio' object
~section()	The destructor.
<pre>Elf_Half get_index()</pre>	Returns section index. Sometimes, this index is passed to another section for inter-referencing between the sections. Section's index is also passed to 'segment' for segment/section association
Set functions:	Sets attributes for the section
<pre>void set_name(std::string void set_type(Elf_Word) void set_flags(Elf_Xword void set_info(Elf_Word) void set_link(Elf_Word) void set_addr_align(Elf_X void set_entry_size(Elf_X void set_address(Elf64_Ad void set_size(Elf_Xword) void set_name_string_offse</pre>	word) word) dr)
Get functions:	Returns section attributes
std::string get_name() Elf_Word get_type() Elf_Xword get_flags() Elf_Word get_info() Elf_Word get_link() Elf_Xword get_addr_align Elf_Xword get_entry_size Elf64_Addr get_address() Elf_Xword get_size() Elf_Word get_name_strin	
Data manipulation function	Manages section data
const char* get_data()	
<pre>void set_data(const char* pData, Elf_Word size)</pre>	
void set_data(

```
const std::string& data )

void     append_data(
     const char* pData,
     Elf_Word size )

void     append_data(
     const std::string& data )
```

4.3 SEGMENT

Class 'segment' has no public data members.

4.3.1 Member functions

segment public member functions listed in the table below. These functions permit to retrieve or set ELF file segment properties

Member Function	Description
segment()	The default constructor. No segment class instances are created manually. Usually, 'add' method is used for 'segments' data member of 'elfio' object
~segment()	The destructor.
<pre>Elf_Half get_index()</pre>	Returns segment's index
Set functions:	Sets attributes for the segment
<pre>void set_type(Elf_Word) void set_flags(Elf_Word) void set_align(Elf_Xword) void set_virtual_address(Elf64_Addr) void set_physical_address(Elf64_Addr) void set_file_size(Elf_Xword) void set_memory_size(Elf_Xword)</pre>	
Get functions:	Returns segment attributes
<pre>Elf_Word get_type() Elf_Word get_flags() Elf_Xword get_align() Elf64_Addr get_virtual_address() Elf64_Addr get_physical_address() Elf_Xword get_file_size()</pre>	

```
Elf_Xword get_memory_size()

Elf_Half
add_section_index(
   Elf_Half index,
   Elf_Xword addr_align )

Elf_Half
get_sections_num()

Elf_Half
get_section_index_at(
   Elf_Half num )

Manages segment-section association
```

4.4 STRING_SECTION_ACCESSOR

4.4.1 Member functions

Member Function	Description
<pre>string_section_accessor(section* section_)</pre>	The constructor
<pre>const char* get_string(Elf_Word index)</pre>	Retrieves string by its offset (index) in the section
<pre>Elf_Word add_string(const char* str)</pre>	Appends section data with new string. Returns position (index) of the new record
<pre>Elf_Word add_string(const std::string& str)</pre>	

4.5 SYMBOL_SECTION_ACCESSOR

4.5.1 Member functions

Member Function	Description
symbol_section_accessor(The constructor
const elfio& elf_file,	
section* symbols_section)	

```
Elf Half
                                                 Returns segment's index
get index()
Elf Xword
                                                 Returns number of symbols in the section
get_symbols_num()
Get functions:
                                                 Retrieves symbol properties by symbol index
                                                 or name
bool
get symbol(
 Elf Xword index,
 std::string& name,
 Elf64_Addr& value,
Elf_Xword& size,
 unsigned char& bind,
 unsigned char& type,
  Elf Half& section index,
  unsigned char& other )
bool
get symbol(
 const std::string& name,
Elf64_Addr& value,

Elf_Xword& size,

unsigned char& bind,

unsigned char& type,

Elf_Half& section_index,

unsigned char& other)
Elf Word
                                                 Adds symbol to the symbol table updating
add symbol(
                                                 corresponding string section if required
 Elf Word name,
 Elf64 Addr value,
 Elf Xword size,
 unsigned char info,
  unsigned char other,
  Elf Half shndx )
Elf Word
add symbol (
 Elf Word name,
 Elf64_Addr value,
 Elf Xword size,
 unsigned char bind,
  unsigned char type,
  unsigned char other,
  Elf Half shndx )
Elf Word
add symbol (
  string_section_accessor& pStrWriter,
```

16

```
const char*
                       str,
 Elf64 Addr
                       value,
 Elf Xword
                       size,
 unsigned char unsigned char
                       info,
                       other,
 Elf Half
                       shndx )
Elf Word
add_symbol(
 string section accessor& pStrWriter,
 const char*
               str,
                     value,
 Elf64 Addr
 Elf Xword
                      size,
                      bind,
 unsigned char
 unsigned char
                       type,
 unsigned char
                       other,
                       shndx )
 Elf Half
```

4.6 RELOCATION_SECTION_ACCESSOR

4.6.1 Member functions

Member Function	Description
<pre>relocation_section_accessor(elfio& elf_file_, section* section_)</pre>	The constructor
<pre>Elf_Xword get_entries_num()</pre>	Retrieves number of relocation entries in the section
<pre>bool get_entry(Elf_Xword index, Elf64_Addr& offset, Elf_Word& symbol, Elf_Word& type, Elf_Sxword& addend)</pre>	Retrieves properties for relocation entry by its index. Calculated value in the second flavor of this function may not work for all architectures
<pre>bool get_entry(Elf_Xword index, Elf64_Addr& offset, Elf64_Addr& symbolValue, std::string& symbolName, Elf_Word& type, Elf Sxword& addend,</pre>	

Elf_Sxword& calcValue)		
void		Adds new relocation entry. The last function
add_entry(in this set is capable to add relocation entry
Elf64_Addr offset,		for a symbol, automatically updating symbol
Elf_Xword info)		
		and string tables for this symbol
void		
add_entry(
Elf64_Addr offset,		
Elf_Word symbol,		
unsigned char type)		
void		
add_entry(
Elf64_Addr offset,		
Elf_Xword info,		
Elf_Sxword addend)		
void		
add entry(
Elf64 Addr offset,		
Elf Word symbol,		
unsigned char type,		
Elf_Sxword addend)		
void		
add entry(
string_section_accessor s	str_writer,	
	str,	
symbol_section_accessor s	sym_writer,	
Elf64_Addr v	alue,	
Elf_Word s	size,	
unsigned char s	sym_info,	
unsigned char o	other,	
Elf_Half s	shndx,	
Elf64_Addr o	offset,	
unsigned char t	cype)	

4.7 NOTE_SECTION_ACCESSOR

4.7.1 Member functions

Member Function	Description
-----------------	-------------

note_section_accessor const elfio& elf_fi section* section	ile_,	The constructor
<pre>Elf_Word get_notes_num()</pre>		Retrieves number of note entries in the section
bool		Retrieves particular note by its index
get_note(
Elf_Word index,		
Elf_Word& type,		
std::string& name,		
void*& desc,		
Elf_Word& descSi	ize)	
void		Appends the section with a new note
add_note(
Elf_Word	type,	
const std::string&	name,	
const void*	desc,	
Elf_Word	descSize)	