4 OBJECT FILES

DRAFT COPY
March 18, 1997
File: Cchap4
386:adm.book:sum

Introduction

This chapter describes the object file format, called ELF (Executable and Linking Format). There are three main types of object files.

- A *relocatable file* holds code and data suitable for linking with other object files to create an executable or a shared object file.
- An *executable file* holds a program suitable for execution; the file specifies how the function `exec` creates a program's process image. X
- A *shared object file* holds code and data suitable for linking in two contexts. First, the link editor [see `ld`(SD_CMD)] may process it with other relocatable and shared object files to create another object file. Second, the dynamic linker combines it with an executable file and other shared objects to create a process image.

Created by the assembler and link editor, object files are binary representations of programs intended to execute directly on a processor. Programs that require other abstract machines, such as shell scripts, are excluded.

After the introductory material, this chapter focuses on the file format and how it pertains to building programs. Chapter 5 also describes parts of the object file, concentrating on the information necessary to execute a program.

File Format

Object files participate in program linking (building a program) and program execution (running a program). For convenience and efficiency, the object file format provides parallel views of a file's contents, reflecting the differing needs of these activities. Figure 4-1 shows an object file's organization.

Figure 4-1: Object File Format

Linking View	Execution View
ELF header	ELF header
Program header table *optional*	Program header table
Section 1	Segment 1
. . .	
Section *n*	Segment 2
. . .	
. . .	. . .
Section header table	Section header table *optional*

An *ELF header* resides at the beginning and holds a ''road map'' describing the file's organization. *Sections* hold the bulk of object file information for the linking view: instructions, data, symbol table, relocation information, and so on. Descriptions of special sections appear later in the chapter. Chapter 5 discusses *segments* and the program execution view of the file.

A *program header table*, if present, tells the system how to create a process image. Files used to build a process image (execute a program) must have a program header table; relocatable files do not need one. A *section header table* contains information describing the file's sections. Every section has an entry in the table; each entry gives information such as the section name, the section size, and so on. Files used during linking must have a section header table; other object files may or may not have one.

NOTE Although the figure shows the program header table immediately after the ELF header, and the section header table following the sections, actual files may differ. Moreover, sections and segments have no specified order. Only the ELF header has a fixed position in the file.

Data Representation

As described here, the object file *format* supports various processors with 8-bit bytes and 32-bit architectures. Nevertheless, it is intended to be extensible to larger (or smaller) architectures. Object files therefore represent some control data with a machine-independent format, making it possible to identify object files and interpret their contents in a common way. Remaining data in an object file use the encoding of the target processor, regardless of the machine on which the file was created.

Figure 4-2: 32-Bit Data Types

Name	Size	Alignment	Purpose
`Elf32_Addr`	4	4	Unsigned program address
`Elf32_Half`	2	2	Unsigned medium integer
`Elf32_Off`	4	4	Unsigned file offset
`Elf32_Sword`	4	4	Signed large integer
`Elf32_Word`	4	4	Unsigned large integer
`unsigned char`	1	1	Unsigned small integer

All data structures that the object file format defines follow the ‘‘natural’’ size and alignment guidelines for the relevant class. If necessary, data structures contain explicit padding to ensure 4-byte alignment for 4-byte objects, to force structure sizes to a multiple of 4, and so on. Data also have suitable alignment from the beginning of the file. Thus, for example, a structure containing an `Elf32_Addr` member will be aligned on a 4-byte boundary within the file.

For portability reasons, ELF uses no bit-fields.

ELF Header

Some object file control structures can grow, because the ELF header contains their actual sizes. If the object file format changes, a program may encounter control structures that are larger or smaller than expected. Programs might therefore ignore ‘‘extra’’ information. The treatment of ‘‘missing’’ information depends on context and will be specified when and if extensions are defined.

Figure 4-3: ELF Header

```
#define EI_NIDENT  16

typedef struct {
        unsigned char      e_ident[EI_NIDENT];
        Elf32_Half         e_type;
        Elf32_Half         e_machine;
        Elf32_Word         e_version;
        Elf32_Addr         e_entry;
        Elf32_Off          e_phoff;
        Elf32_Off          e_shoff;
        Elf32_Word         e_flags;
        Elf32_Half         e_ehsize;
        Elf32_Half         e_phentsize;
        Elf32_Half         e_phnum;
        Elf32_Half         e_shentsize;
        Elf32_Half         e_shnum;
        Elf32_Half         e_shstrndx;
} Elf32_Ehdr;
```

`e_ident` The initial bytes mark the file as an object file and provide machine-independent data with which to decode and interpret the file’s contents. Complete descriptions appear below, in ‘‘ELF Identification.’’

e_type This member identifies the object file type.

Name	Value	Meaning
ET_NONE	0	No file type
ET_REL	1	Relocatable file
ET_EXEC	2	Executable file
ET_DYN	3	Shared object file
ET_CORE	4	Core file
ET_LOPROC	0xff00	Processor-specific
ET_HIPROC	0xffff	Processor-specific

Although the core file contents are unspecified, type ET_CORE is reserved to mark the file. Values from ET_LOPROC through ET_HIPROC (inclusive) are reserved for processor-specific semantics. If meanings are specified, the processor supplement explains them. Other values are reserved and will be assigned to new object file types as necessary.

e_machine This member's value specifies the required architecture for an individual file.

Name	Value	Meaning	
EM_NONE	0	No machine	
EM_M32	1	AT&T WE 32100	
EM_SPARC	2	SPARC	
EM_386	3	Intel 80386	
EM_68K	4	Motorola 68000	
EM_88K	5	Motorola 88000	
EM_860	7	Intel 80860	
EM_MIPS	8	MIPS RS3000 Big-Endian	E
EM_MIPS_RS4_BE	10	MIPS RS4000 Big-Endian	E
RESERVED	11-16	Reserved for future use	E

Other values are reserved and will be assigned to new machines as necessary. Processor-specific ELF names use the machine name to distinguish them. For example, the flags mentioned below use the prefix EF_; a flag named WIDGET for the EM_XYZ machine would be called EF_XYZ_WIDGET.

e_version This member identifies the object file version.

Name	Value	Meaning
EV_NONE	0	Invalid version
EV_CURRENT	1	Current version

The value 1 signifies the original file format; extensions will create new versions with higher numbers. The value of EV_CURRENT, though given as 1 above, will change as necessary to reflect the current version number.

e_entry — This member gives the virtual address to which the system first transfers control, thus starting the process. If the file has no associated entry point, this member holds zero.

e_phoff — This member holds the program header table's file offset in bytes. If the file has no program header table, this member holds zero.

e_shoff — This member holds the section header table's file offset in bytes. If the file has no section header table, this member holds zero.

e_flags — This member holds processor-specific flags associated with the file. Flag names take the form EF_*machine_flag*. See ''Machine Information'' in the processor supplement for flag definitions.

e_ehsize — This member holds the ELF header's size in bytes.

e_phentsize — This member holds the size in bytes of one entry in the file's program header table; all entries are the same size.

e_phnum — This member holds the number of entries in the program header table. Thus the product of e_phentsize and e_phnum gives the table's size in bytes. If a file has no program header table, e_phnum holds the value zero.

e_shentsize — This member holds a section header's size in bytes. A section header is one entry in the section header table; all entries are the same size.

e_shnum — This member holds the number of entries in the section header table. Thus the product of e_shentsize and e_shnum gives the section header table's size in bytes. If a file has no section header table, e_shnum holds the value zero.

e_shstrndx — This member holds the section header table index of the entry associated with the section name string table. If the file has no section name string table, this member holds the value SHN_UNDEF. See ''Sections'' and ''String Table'' below for more information.

ELF Identification

As mentioned above, ELF provides an object file framework to support multiple processors, multiple data encodings, and multiple classes of machines. To support this object file family, the initial bytes of the file specify how to interpret the file, independent of the processor on which the inquiry is made and independent of the file's remaining contents.

The initial bytes of an ELF header (and an object file) correspond to the `e_ident` member.

Figure 4-4: `e_ident[ ]` **Identification Indexes**

Name	Value	Purpose
EI_MAG0	0	File identification
EI_MAG1	1	File identification
EI_MAG2	2	File identification
EI_MAG3	3	File identification
EI_CLASS	4	File class
EI_DATA	5	Data encoding
EI_VERSION	6	File version
EI_PAD	7	Start of padding bytes
EI_NIDENT	16	Size of e_ident[]

These indexes access bytes that hold the following values.

`EI_MAG0` to `EI_MAG3`

A file's first 4 bytes hold a ''magic number,'' identifying the file as an ELF object file.

Name	Value	Position
ELFMAG0	0x7f	e_ident[EI_MAG0]
ELFMAG1	'E'	e_ident[EI_MAG1]
ELFMAG2	'L'	e_ident[EI_MAG2]
ELFMAG3	'F'	e_ident[EI_MAG3]

EI_CLASS The next byte, e_ident[EI_CLASS], identifies the file's class, or capacity.

Name	Value	Meaning
ELFCLASSNONE	0	Invalid class
ELFCLASS32	1	32-bit objects
ELFCLASS64	2	64-bit objects

The file format is designed to be portable among machines of various sizes, without imposing the sizes of the largest machine on the smallest. Class ELFCLASS32 supports machines with files and virtual address spaces up to 4 gigabytes; it uses the basic types defined above.

Class ELFCLASS64 is reserved for 64-bit architectures. Its appearance here shows how the object file may change, but the 64-bit format is otherwise unspecified. Other classes will be defined as necessary, with different basic types and sizes for object file data.

EI_DATA Byte e_ident[EI_DATA] specifies the data encoding of the processor-specific data in the object file. The following encodings are currently defined.

Name	Value	Meaning
ELFDATANONE	0	Invalid data encoding
ELFDATA2LSB	1	See below
ELFDATA2MSB	2	See below

More information on these encodings appears below. Other values are reserved and will be assigned to new encodings as necessary.

EI_VERSION Byte e_ident[EI_VERSION] specifies the ELF header version number. Currently, this value must be EV_CURRENT, as explained above for e_version.

EI_PAD This value marks the beginning of the unused bytes in e_ident. These bytes are reserved and set to zero; programs that read object files should ignore them. The value of EI_PAD will change in the future if currently unused bytes are given meanings.

A file's data encoding specifies how to interpret the basic objects in a file. As described above, class ELFCLASS32 files use objects that occupy 1, 2, and 4 bytes. Under the defined encodings, objects are represented as shown below. Byte numbers appear in the upper left corners.

Encoding `ELFDATA2LSB` specifies 2's complement values, with the least significant byte occupying the lowest address.

Figure 4-5: Data Encoding `ELFDATA2LSB`

0x01	^{0} 01			
0x0102	^{0} 02	^{1} 01		
0x01020304	^{0} 04	^{1} 03	^{2} 02	^{3} 01

Encoding `ELFDATA2MSB` specifies 2's complement values, with the most significant byte occupying the lowest address.

Figure 4-6: Data Encoding `ELFDATA2MSB`

0x01	^{0} 01			
0x0102	^{0} 01	^{1} 02		
0x01020304	^{0} 01	^{1} 02	^{2} 03	^{3} 04

Machine Information (Processor-Specific)

NOTE This section requires processor-specific information. The ABI supplement for the desired processor describes the details.

Sections

An object file's section header table lets one locate all the file's sections. The section header table is an array of Elf32_Shdr structures as described below. A section header table index is a subscript into this array. The ELF header's e_shoff member gives the byte offset from the beginning of the file to the section header table; e_shnum tells how many entries the section header table contains; e_shentsize gives the size in bytes of each entry.

Some section header table indexes are reserved; an object file will not have sections for these special indexes.

Figure 4-7: Special Section Indexes

Name	Value
SHN_UNDEF	0
SHN_LORESERVE	0xff00
SHN_LOPROC	0xff00
SHN_HIPROC	0xff1f
SHN_ABS	0xfff1
SHN_COMMON	0xfff2
SHN_HIRESERVE	0xffff

SHN_UNDEF — This value marks an undefined, missing, irrelevant, or otherwise meaningless section reference. For example, a symbol ''defined'' relative to section number SHN_UNDEF is an undefined symbol.

NOTE Although index 0 is reserved as the undefined value, the section header table contains an entry for index 0. That is, if the e_shnum member of the ELF header says a file has 6 entries in the section header table, they have the indexes 0 through 5. The contents of the initial entry are specified later in this section.

SHN_LORESERVE — This value specifies the lower bound of the range of reserved indexes.

SHN_LOPROC through SHN_HIPROC

Values in this inclusive range are reserved for processor-specific semantics. If meanings are specified, the processor supplement explains them.

SHN_ABS	This value specifies absolute values for the corresponding reference. For example, symbols defined relative to section number SHN_ABS have absolute values and are not affected by relocation.
SHN_COMMON	Symbols defined relative to this section are common symbols, such as FORTRAN COMMON or unallocated C external variables.
SHN_HIRESERVE	This value specifies the upper bound of the range of reserved indexes. The system reserves indexes between SHN_LORESERVE and SHN_HIRESERVE, inclusive; the values do not reference the section header table. That is, the section header table does *not* contain entries for the reserved indexes.

Sections contain all information in an object file, except the ELF header, the program header table, and the section header table. Moreover, object files' sections satisfy several conditions.

- Every section in an object file has exactly one section header describing it. Section headers may exist that do not have a section.
- Each section occupies one contiguous (possibly empty) sequence of bytes within a file.
- Sections in a file may not overlap. No byte in a file resides in more than one section.
- An object file may have inactive space. The various headers and the sections might not ''cover'' every byte in an object file. The contents of the inactive data are unspecified.

A section header has the following structure.

Figure 4-8: Section Header

```
typedef struct {
        Elf32_Word      sh_name;
        Elf32_Word      sh_type;
        Elf32_Word      sh_flags;
        Elf32_Addr      sh_addr;
        Elf32_Off       sh_offset;
        Elf32_Word      sh_size;
        Elf32_Word      sh_link;
        Elf32_Word      sh_info;
        Elf32_Word      sh_addralign;
        Elf32_Word      sh_entsize;
} Elf32_Shdr;
```

`sh_name` This member specifies the name of the section. Its value is an index into the section header string table section [see ''String Table'' below], giving the location of a null-terminated string.

`sh_type` This member categorizes the section's contents and semantics. Section types and their descriptions appear below.

`sh_flags` Sections support 1-bit flags that describe miscellaneous attributes. Flag definitions appear below.

`sh_addr` If the section will appear in the memory image of a process, this member gives the address at which the section's first byte should reside. Otherwise, the member contains 0.

`sh_offset` This member's value gives the byte offset from the beginning of the file to the first byte in the section. One section type, `SHT_NOBITS` described below, occupies no space in the file, and its `sh_offset` member locates the conceptual placement in the file.

`sh_size` This member gives the section's size in bytes. Unless the section type is `SHT_NOBITS`, the section occupies `sh_size` bytes in the file. A section of type `SHT_NOBITS` may have a non-zero size, but it occupies no space in the file.

`sh_link` This member holds a section header table index link, whose interpretation depends on the section type. A table below describes the values.

sh_info This member holds extra information, whose interpretation depends on the section type. A table below describes the values.

sh_addralign Some sections have address alignment constraints. For example, if a section holds a doubleword, the system must ensure doubleword alignment for the entire section. That is, the value of sh_addr must be congruent to 0, modulo the value of sh_addralign. Currently, only 0 and positive integral powers of two are allowed. Values 0 and 1 mean the section has no alignment constraints.

sh_entsize Some sections hold a table of fixed-size entries, such as a symbol table. For such a section, this member gives the size in bytes of each entry. The member contains 0 if the section does not hold a table of fixed-size entries.

A section header's sh_type member specifies the section's semantics.

Figure 4-9: Section Types, sh_type

Name	Value
SHT_NULL	0
SHT_PROGBITS	1
SHT_SYMTAB	2
SHT_STRTAB	3
SHT_RELA	4
SHT_HASH	5
SHT_DYNAMIC	6
SHT_NOTE	7
SHT_NOBITS	8
SHT_REL	9
SHT_SHLIB	10
SHT_DYNSYM	11
SHT_LOPROC	0x70000000
SHT_HIPROC	0x7fffffff
SHT_LOUSER	0x80000000
SHT_HIUSER	0xffffffff

SHT_NULL This value marks the section header as inactive; it does not have an associated section. Other members of the section header have undefined values.

SHT_PROGBITS — The section holds information defined by the program, whose format and meaning are determined solely by the program.

SHT_SYMTAB and SHT_DYNSYM — These sections hold a symbol table. Currently, an object file may have only one section of each type, but this restriction may be relaxed in the future. Typically, SHT_SYMTAB provides symbols for link editing, though it may also be used for dynamic linking. As a complete symbol table, it may contain many symbols unnecessary for dynamic linking. Consequently, an object file may also contain a SHT_DYNSYM section, which holds a minimal set of dynamic linking symbols, to save space. See ‘‘Symbol Table’’ below for details.

SHT_STRTAB — The section holds a string table. An object file may have multiple string table sections. See ‘‘String Table’’ below for details.

SHT_RELA — The section holds relocation entries with explicit addends, such as type Elf32_Rela for the 32-bit class of object files. An object file may have multiple relocation sections. See ‘‘Relocation’’ below for details.

SHT_HASH — The section holds a symbol hash table. All objects participating in dynamic linking must contain a symbol hash table. Currently, an object file may have only one hash table, but this restriction may be relaxed in the future. See ‘‘Hash Table’’ in Chapter 5 for details.

SHT_DYNAMIC — The section holds information for dynamic linking. Currently, an object file may have only one dynamic section, but this restriction may be relaxed in the future. See ‘‘Dynamic Section’’ in Chapter 5 for details.

SHT_NOTE — The section holds information that marks the file in some way. See ‘‘Note Section’’ in Chapter 5 for details.

SHT_NOBITS — A section of this type occupies no space in the file but otherwise resembles SHT_PROGBITS. Although this section contains no bytes, the sh_offset member contains the conceptual file offset.

SHT_REL — The section holds relocation entries without explicit addends, such as type Elf32_Rel for the 32-bit class of object files. An object file may have multiple relocation sections. See ‘‘Relocation’’ below for details.

SHT_SHLIB — This section type is reserved but has unspecified semantics. Programs that contain a section of this type do not conform to the ABI.

SHT_LOPROC through SHT_HIPROC — Values in this inclusive range are reserved for processor-specific semantics. If meanings are specified, the processor supplement explains them.

SHT_LOUSER — This value specifies the lower bound of the range of indexes reserved for application programs.

SHT_HIUSER — This value specifies the upper bound of the range of indexes reserved for application programs. Section types between SHT_LOUSER and SHT_HIUSER may be used by the application, without conflicting with current or future system-defined section types.

Other section type values are reserved. As mentioned before, the section header for index 0 (SHN_UNDEF) exists, even though the index marks undefined section references. This entry holds the following.

Figure 4-10: Section Header Table Entry: Index 0

Name	Value	Note
sh_name	0	No name
sh_type	SHT_NULL	Inactive
sh_flags	0	No flags
sh_addr	0	No address
sh_offset	0	No file offset
sh_size	0	No size
sh_link	SHN_UNDEF	No link information
sh_info	0	No auxiliary information
sh_addralign	0	No alignment
sh_entsize	0	No entries

A section header's sh_flags member holds 1-bit flags that describe the section's attributes. Defined values appear below; other values are reserved.

Figure 4-11: Section Attribute Flags, `sh_flags`

Name	Value
SHF_WRITE	0x1
SHF_ALLOC	0x2
SHF_EXECINSTR	0x4
SHF_MASKPROC	0xf0000000

If a flag bit is set in `sh_flags`, the attribute is ‘‘on’’ for the section. Otherwise, the attribute is ‘‘off’’ or does not apply. Undefined attributes are set to zero.

`SHF_WRITE` The section contains data that should be writable during process execution.

`SHF_ALLOC` The section occupies memory during process execution. Some control sections do not reside in the memory image of an object file; this attribute is off for those sections.

`SHF_EXECINSTR` The section contains executable machine instructions.

`SHF_MASKPROC` All bits included in this mask are reserved for processor-specific semantics. If meanings are specified, the processor supplement explains them.

Two members in the section header, `sh_link` and `sh_info`, hold special information, depending on section type.

Figure 4-12: `sh_link` **and** `sh_info` **Interpretation**

sh_type	sh_link	sh_info
SHT_DYNAMIC	The section header index of the string table used by entries in the section.	0
SHT_HASH	The section header index of the symbol table to which the hash table applies.	0
SHT_REL SHT_RELA	The section header index of the associated symbol table.	The section header index of the section to which the relocation applies.
SHT_SYMTAB SHT_DYNSYM	The section header index of the associated string table.	One greater than the symbol table index of the last local symbol (binding STB_LOCAL).
other	SHN_UNDEF	0

Special Sections

Various sections hold program and control information. Sections in the list below are used by the system and have the indicated types and attributes.

Figure 4-13: Special Sections

Name	Type	Attributes
.bss	SHT_NOBITS	SHF_ALLOC + SHF_WRITE
.comment	SHT_PROGBITS	none
.data	SHT_PROGBITS	SHF_ALLOC + SHF_WRITE
.data1	SHT_PROGBITS	SHF_ALLOC + SHF_WRITE
.debug	SHT_PROGBITS	none
.dynamic	SHT_DYNAMIC	see below
.dynstr	SHT_STRTAB	SHF_ALLOC
.dynsym	SHT_DYNSYM	SHF_ALLOC
.fini	SHT_PROGBITS	SHF_ALLOC + SHF_EXECINSTR
.got	SHT_PROGBITS	see below

Figure 4-13: Special Sections (continued)

.hash	SHT_HASH	SHF_ALLOC
.init	SHT_PROGBITS	SHF_ALLOC + SHF_EXECINSTR
.interp	SHT_PROGBITS	see below
.line	SHT_PROGBITS	none
.note	SHT_NOTE	none
.plt	SHT_PROGBITS	see below
.rel*name*	SHT_REL	see below
.rela*name*	SHT_RELA	see below
.rodata	SHT_PROGBITS	SHF_ALLOC
.rodata1	SHT_PROGBITS	SHF_ALLOC
.shstrtab	SHT_STRTAB	none
.strtab	SHT_STRTAB	see below
.symtab	SHT_SYMTAB	see below
.text	SHT_PROGBITS	SHF_ALLOC + SHF_EXECINSTR

`.bss` This section holds uninitialized data that contribute to the program's memory image. By definition, the system initializes the data with zeros when the program begins to run. The section occupies no file space, as indicated by the section type, `SHT_NOBITS`.

`.comment` This section holds version control information.

`.data` and `.data1`
These sections hold initialized data that contribute to the program's memory image.

`.debug` This section holds information for symbolic debugging. The contents are unspecified. All section names with the prefix `.debug` are reserved for future use in the ABI. E

`.dynamic` This section holds dynamic linking information. The section's attributes will include the `SHF_ALLOC` bit. Whether the `SHF_WRITE` bit is set is processor specific. See Chapter 5 for more information.

`.dynstr` This section holds strings needed for dynamic linking, most commonly the strings that represent the names associated with symbol table entries. See Chapter 5 for more information.

`.dynsym` This section holds the dynamic linking symbol table, as ''Symbol Table'' describes. See Chapter 5 for more information.

`.fini` This section holds executable instructions that contribute to the process termination code. That is, when a program exits normally, the system arranges to execute the code in this section.

`.got` This section holds the global offset table. See ''Coding Examples'' in Chapter 3, ''Special Sections'' in Chapter 4, and ''Global Offset Table'' in Chapter 5 of the processor supplement for more information.

`.hash` This section holds a symbol hash table. See ''Hash Table'' in Chapter 5 for more information.

`.init` This section holds executable instructions that contribute to the process initialization code. That is, when a program starts to run, the system arranges to execute the code in this section before calling the main program entry point (called `main` for C programs).

`.interp` This section holds the path name of a program interpreter. If the file has a loadable segment that includes the section, the section's attributes will include the `SHF_ALLOC` bit; otherwise, that bit will be off. See Chapter 5 for more information.

`.line` This section holds line number information for symbolic debugging, which describes the correspondence between the source program and the machine code. The contents are unspecified.

`.note` This section holds information in the format that ''Note Section'' in Chapter 5 describes.

`.plt` This section holds the procedure linkage table. See ''Special Sections'' in Chapter 4 and ''Procedure Linkage Table'' in Chapter 5 of the processor supplement for more information.

`.rel`*name* and `.rela`*name*
These sections hold relocation information, as ''Relocation'' below describes. If the file has a loadable segment that includes relocation, the sections' attributes will include the `SHF_ALLOC` bit; otherwise, that bit will be off. Conventionally, *name* is supplied by the section to which the relocations apply. Thus a relocation section for `.text` normally would have the name `.rel.text` or `.rela.text`.

`.rodata` and `.rodata1`
These sections hold read-only data that typically contribute to a non-writable segment in the process image. See ''Program Header'' in Chapter 5 for more information.

`.shstrtab` This section holds section names.

`.strtab` This section holds strings, most commonly the strings that represent the names associated with symbol table entries. If the file has a loadable segment that includes the symbol string table, the section's attributes will include the `SHF_ALLOC` bit; otherwise, that bit will be off.

`.symtab` This section holds a symbol table, as ''Symbol Table'' in this chapter describes. If the file has a loadable segment that includes the symbol table, the section's attributes will include the `SHF_ALLOC` bit; otherwise, that bit will be off.

`.text` This section holds the ''text,'' or executable instructions, of a program.

Section names with a dot (`.`) prefix are reserved for the system, although applications may use these sections if their existing meanings are satisfactory. Applications may use names without the prefix to avoid conflicts with system sections. The object file format lets one define sections not in the list above. An object file may have more than one section with the same name.

Section names reserved for a processor architecture are formed by placing an abbreviation of the architecture name ahead of the section name. The name should be taken from the architecture names used for `e_machine`. For instance .FOO.psect is the psect section defined by the FOO architecture. Existing extensions are called by their historical names.

Pre-existing Extensions	
`.sdata`	`.tdesc`
`.sbss`	`.lit4`
`.lit8`	`.reginfo`
`.gptab`	`.liblist`
`.conflict`	

NOTE For information on processor-specific sections, see the ABI supplement for the desired processor.

String Table

String table sections hold null-terminated character sequences, commonly called strings. The object file uses these strings to represent symbol and section names. One references a string as an index into the string table section. The first byte, which is index zero, is defined to hold a null character. Likewise, a string table's last byte is defined to hold a null character, ensuring null termination for all strings. A string whose index is zero specifies either no name or a null name, depending on the context. An empty string table section is permitted; its section header's `sh_size` member would contain zero. Non-zero indexes are invalid for an empty string table.

A section header's `sh_name` member holds an index into the section header string table section, as designated by the `e_shstrndx` member of the ELF header. The following figures show a string table with 25 bytes and the strings associated with various indexes.

Index	+0	+1	+2	+3	+4	+5	+6	+7	+8	+9
0	`\0`	`n`	`a`	`m`	`e`	`.`	`\0`	`V`	`a`	`r`
10	`i`	`a`	`b`	`l`	`e`	`\0`	`a`	`b`	`l`	`e`
20	`\0`	`\0`	`x`	`x`	`\0`					

Figure 4-14: String Table Indexes

Index	String
0	*none*
1	`name.`
7	`Variable`
11	`able`
16	`able`
24	*null string*

As the example shows, a string table index may refer to any byte in the section. A string may appear more than once; references to substrings may exist; and a single string may be referenced multiple times. Unreferenced strings also are allowed.

Symbol Table

An object file's symbol table holds information needed to locate and relocate a program's symbolic definitions and references. A symbol table index is a subscript into this array. Index 0 both designates the first entry in the table and serves as the undefined symbol index. The contents of the initial entry are specified later in this section.

Name	Value
STN_UNDEF	0

A symbol table entry has the following format.

Figure 4-15: Symbol Table Entry

```
typedef struct {
        Elf32_Word        st_name;
        Elf32_Addr        st_value;
        Elf32_Word        st_size;
        unsigned char     st_info;
        unsigned char     st_other;
        Elf32_Half        st_shndx;
} Elf32_Sym;
```

`st_name` This member holds an index into the object file's symbol string table, which holds the character representations of the symbol names. If the value is non-zero, it represents a string table index that gives the symbol name. Otherwise, the symbol table entry has no name.

NOTE External C symbols have the same names in C and object files' symbol tables.

`st_value` This member gives the value of the associated symbol. Depending on the context, this may be an absolute value, an address, and so on; details appear below.

st_size — Many symbols have associated sizes. For example, a data object's size is the number of bytes contained in the object. This member holds 0 if the symbol has no size or an unknown size.

st_info — This member specifies the symbol's type and binding attributes. A list of the values and meanings appears below. The following code shows how to manipulate the values.

```
#define ELF32_ST_BIND(i)    ((i)>>4)
#define ELF32_ST_TYPE(i)    ((i)&0xf)
#define ELF32_ST_INFO(b,t)  (((b)<<4)+((t)&0xf))
```

st_other — This member currently holds 0 and has no defined meaning.

st_shndx — Every symbol table entry is ''defined'' in relation to some section; this member holds the relevant section header table index. As Figure 4-7 and the related text describe, some section indexes indicate special meanings.

A symbol's binding determines the linkage visibility and behavior.

Figure 4-16: Symbol Binding, `ELF32_ST_BIND`

Name	Value
STB_LOCAL	0
STB_GLOBAL	1
STB_WEAK	2
STB_LOPROC	13
STB_HIPROC	15

STB_LOCAL — Local symbols are not visible outside the object file containing their definition. Local symbols of the same name may exist in multiple files without interfering with each other.

STB_GLOBAL — Global symbols are visible to all object files being combined. One file's definition of a global symbol will satisfy another file's undefined reference to the same global symbol.

STB_WEAK Weak symbols resemble global symbols, but their definitions have lower precedence.

STB_LOPROC through STB_HIPROC
Values in this inclusive range are reserved for processor-specific semantics. If meanings are specified, the processor supplement explains them.

Global and weak symbols differ in two major ways.

- When the link editor combines several relocatable object files, it does not allow multiple definitions of STB_GLOBAL symbols with the same name. On the other hand, if a defined global symbol exists, the appearance of a weak symbol with the same name will not cause an error. The link editor honors the global definition and ignores the weak ones. Similarly, if a common symbol exists (that is, a symbol whose st_shndx field holds SHN_COMMON), the appearance of a weak symbol with the same name will not cause an error. The link editor honors the common definition and ignores the weak ones.
- When the link editor searches archive libraries [see ''Archive File'' in Chapter 7], it extracts archive members that contain definitions of undefined global symbols. The member's definition may be either a global or a weak symbol. The link editor does *not* extract archive members to resolve undefined weak symbols. Unresolved weak symbols have a zero value.

In each symbol table, all symbols with STB_LOCAL binding precede the weak and global symbols. As ''Sections'' above describes, a symbol table section's sh_info section header member holds the symbol table index for the first non-local symbol.

A symbol's type provides a general classification for the associated entity.

Figure 4-17: Symbol Types, ELF32_ST_TYPE

Name	Value
STT_NOTYPE	0
STT_OBJECT	1
STT_FUNC	2
STT_SECTION	3
STT_FILE	4
STT_LOPROC	13
STT_HIPROC	15

STT_NOTYPE The symbol's type is not specified.

STT_OBJECT The symbol is associated with a data object, such as a variable, an array, and so on.

STT_FUNC The symbol is associated with a function or other executable code.

STT_SECTION The symbol is associated with a section. Symbol table entries of this type exist primarily for relocation and normally have STB_LOCAL binding.

STT_FILE Conventionally, the symbol's name gives the name of the source file associated with the object file. A file symbol has STB_LOCAL binding, its section index is SHN_ABS, and it precedes the other STB_LOCAL symbols for the file, if it is present.

STT_LOPROC through STT_HIPROC
Values in this inclusive range are reserved for processor-specific semantics. If meanings are specified, the processor supplement explains them.

Function symbols (those with type STT_FUNC) in shared object files have special significance. When another object file references a function from a shared object, the link editor automatically creates a procedure linkage table entry for the referenced symbol. Shared object symbols with types other than STT_FUNC will not be referenced automatically through the procedure linkage table.

If a symbol's value refers to a specific location within a section, its section index member, st_shndx, holds an index into the section header table. As the section moves during relocation, the symbol's value changes as well, and references to the symbol continue to ''point'' to the same location in the program. Some special section index values give other semantics.

SHN_ABS The symbol has an absolute value that will not change because of relocation.

SHN_COMMON The symbol labels a common block that has not yet been allocated. The symbol's value gives alignment constraints, similar to a section's sh_addralign member. That is, the link editor will allocate the storage for the symbol at an address that is a multiple of st_value. The symbol's size tells how many bytes are required.

SHN_UNDEF This section table index means the symbol is undefined. When the link editor combines this object file with another that defines the indicated symbol, this file's references to the symbol will be linked to the actual definition.

As mentioned above, the symbol table entry for index 0 (STN_UNDEF) is reserved; it holds the following.

Figure 4-18: Symbol Table Entry: Index 0

Name	Value	Note
st_name	0	No name
st_value	0	Zero value
st_size	0	No size
st_info	0	No type, local binding
st_other	0	
st_shndx	SHN_UNDEF	No section

Symbol Values

Symbol table entries for different object file types have slightly different interpretations for the st_value member.

- In relocatable files, st_value holds alignment constraints for a symbol whose section index is SHN_COMMON.
- In relocatable files, st_value holds a section offset for a defined symbol. That is, st_value is an offset from the beginning of the section that st_shndx identifies.
- In executable and shared object files, st_value holds a virtual address. To make these files' symbols more useful for the dynamic linker, the section offset (file interpretation) gives way to a virtual address (memory interpretation) for which the section number is irrelevant.

Although the symbol table values have similar meanings for different object files, the data allow efficient access by the appropriate programs.

Relocation

Relocation is the process of connecting symbolic references with symbolic definitions. For example, when a program calls a function, the associated call instruction must transfer control to the proper destination address at execution. In other words, relocatable files must have information that describes how to modify their section contents, thus allowing executable and shared object files to hold the right information for a process's program image. *Relocation entries* are these data.

Figure 4-19: Relocation Entries

```
typedef struct {
        Elf32_Addr        r_offset;
        Elf32_Word        r_info;
} Elf32_Rel;

typedef struct {
        Elf32_Addr        r_offset;
        Elf32_Word        r_info;
        Elf32_Sword       r_addend;
} Elf32_Rela;
```

`r_offset` This member gives the location at which to apply the relocation action. For a relocatable file, the value is the byte offset from the beginning of the section to the storage unit affected by the relocation. For an executable file or a shared object, the value is the virtual address of the storage unit affected by the relocation.

`r_info` This member gives both the symbol table index with respect to which the relocation must be made, and the type of relocation to apply. For example, a call instruction's relocation entry would hold the symbol table index of the function being called. If the index is `STN_UNDEF`, the undefined symbol index, the relocation uses 0 as the ''symbol value.'' Relocation types are processor-specific; descriptions of their behavior appear in the processor supplement. When the text in the processor supplement refers to a relocation entry's relocation type or symbol table index, it means the result of applying `ELF32_R_TYPE` or `ELF32_R_SYM`, respectively, to the entry's `r_info` member.

```
#define ELF32_R_SYM(i)    ((i)>>8)
#define ELF32_R_TYPE(i)   ((unsigned char)(i))
#define ELF32_R_INFO(s,t) (((s)<<8)+(unsigned char)(t))
```

`r_addend` This member specifies a constant addend used to compute the value to be stored into the relocatable field.

As shown above, only `Elf32_Rela` entries contain an explicit addend. Entries of type `Elf32_Rel` store an implicit addend in the location to be modified. Depending on the processor architecture, one form or the other might be necessary or more convenient. Consequently, an implementation for a particular machine may use one form exclusively or either form depending on context.

A relocation section references two other sections: a symbol table and a section to modify. The section header's `sh_info` and `sh_link` members, described in ''Sections'' above, specify these relationships. Relocation entries for different object files have slightly different interpretations for the `r_offset` member.

- In relocatable files, `r_offset` holds a section offset. That is, the relocation section itself describes how to modify another section in the file; relocation offsets designate a storage unit within the second section.
- In executable and shared object files, `r_offset` holds a virtual address. To make these files' relocation entries more useful for the dynamic linker, the section offset (file interpretation) gives way to a virtual address (memory interpretation).

Although the interpretation of `r_offset` changes for different object files to allow efficient access by the relevant programs, the relocation types' meanings stay the same.

Relocation Types (Processor-Specific)

NOTE This section requires processor-specific information. The ABI supplement for the desired processor describes the details.

5 PROGRAM LOADING AND DYNAMIC LINKING

DRAFT COPY
March 18, 1997
File: Cchap5
386:adm.book:sum

Introduction

This chapter describes the object file information and system actions that create running programs. Some information here applies to all systems; information specific to one processor resides in sections marked accordingly.

Executable and shared object files statically represent programs. To execute such programs, the system uses the files to create dynamic program representations, or process images. As section ''Virtual Address Space'' in Chapter 3 of the processor supplement describes, a process image has segments that hold its text, data, stack, and so on. This chapter's major sections discuss the following.

- *Program header.* This section complements Chapter 4, describing object file structures that relate directly to program execution. The primary data structure, a program header table, locates segment images within the file and contains other information necessary to create the memory image for the program.
- *Program loading.* Given an object file, the system must load it into memory for the program to run.
- *Dynamic linking.* After the system loads the program, it must complete the process image by resolving symbolic references among the object files that compose the process.

NOTE The processor supplement defines a naming convention for ELF constants that have processor ranges specified. Names such as DT_, PT_, for processor specific extensions, incorporate the name of the processor: DT_M32_SPECIAL, for example. Pre–existing processor extensions not using this convention will be supported.

Pre-existing Extensions
`DT_JMP_REL`

Program Header

An executable or shared object file's program header table is an array of structures, each describing a segment or other information the system needs to prepare the program for execution. An object file *segment* contains one or more *sections*, as ''Segment Contents'' describes below. Program headers are meaningful only for executable and shared object files. A file specifies its own program header size with the ELF header's `e_phentsize` and `e_phnum` members [see ''ELF Header'' in Chapter 4].

Figure 5-1: Program Header

```
typedef struct {
        Elf32_Word      p_type;
        Elf32_Off       p_offset;
        Elf32_Addr      p_vaddr;
        Elf32_Addr      p_paddr;
        Elf32_Word      p_filesz;
        Elf32_Word      p_memsz;
        Elf32_Word      p_flags;
        Elf32_Word      p_align;
} Elf32_Phdr;
```

`p_type` This member tells what kind of segment this array element describes or how to interpret the array element's information. Type values and their meanings appear below.

`p_offset` This member gives the offset from the beginning of the file at which the first byte of the segment resides.

`p_vaddr` This member gives the virtual address at which the first byte of the segment resides in memory.

`p_paddr` On systems for which physical addressing is relevant, this member is reserved for the segment's physical address. Because System V ignores physical addressing for application programs, this member has unspecified contents for executable files and shared objects.

p_filesz — This member gives the number of bytes in the file image of the segment; it may be zero.

p_memsz — This member gives the number of bytes in the memory image of the segment; it may be zero.

p_flags — This member gives flags relevant to the segment. Defined flag values appear below.

p_align — As ‘‘Program Loading’’ describes in this chapter of the processor supplement, loadable process segments must have congruent values for p_vaddr and p_offset, modulo the page size. This member gives the value to which the segments are aligned in memory and in the file. Values 0 and 1 mean no alignment is required. Otherwise, p_align should be a positive, integral power of 2, and p_vaddr should equal p_offset, modulo p_align.

Some entries describe process segments; others give supplementary information and do not contribute to the process image. Segment entries may appear in any order, except as explicitly noted below. Defined type values follow; other values are reserved for future use.

Figure 5-2: Segment Types, p_type

Name	Value
PT_NULL	0
PT_LOAD	1
PT_DYNAMIC	2
PT_INTERP	3
PT_NOTE	4
PT_SHLIB	5
PT_PHDR	6
PT_LOPROC	0x70000000
PT_HIPROC	0x7fffffff

PT_NULL — The array element is unused; other members’ values are undefined. This type lets the program header table have ignored entries.

PT_LOAD — The array element specifies a loadable segment, described by p_filesz and p_memsz. The bytes from the file are mapped to the beginning of the memory segment. If the segment’s memory size (p_memsz) is larger than the file size (p_filesz), the ‘‘extra’’ bytes are defined to hold the value 0 and to follow the segment's initialized area. The file size may not be larger than the memory size. Loadable segment entries in the program header table appear in ascending order, sorted on the p_vaddr member.

PT_DYNAMIC The array element specifies dynamic linking information. See ''Dynamic Section'' below for more information.

PT_INTERP The array element specifies the location and size of a null-terminated path name to invoke as an interpreter. This segment type is meaningful only for executable files (though it may occur for shared objects); it may not occur more than once in a file. If it is present, it must precede any loadable segment entry. See ''Program Interpreter'' below for further information.

PT_NOTE The array element specifies the location and size of auxiliary information. See ''Note Section'' below for details.

PT_SHLIB This segment type is reserved but has unspecified semantics. Programs that contain an array element of this type do not conform to the ABI.

PT_PHDR The array element, if present, specifies the location and size of the program header table itself, both in the file and in the memory image of the program. This segment type may not occur more than once in a file. Moreover, it may occur only if the program header table is part of the memory image of the program. If it is present, it must precede any loadable segment entry. See ''Program Interpreter'' below for further information.

PT_LOPROC through PT_HIPROC
Values in this inclusive range are reserved for processor-specific semantics. If meanings are specified, the processor supplement explains them.

NOTE Unless specifically required elsewhere, all program header segment types are optional. That is, a file's program header table may contain only those elements relevant to its contents.

Base Address

As ‘‘Program Loading’’ in this chapter of the processor supplement describes, the virtual addresses in the program headers might not represent the actual virtual addresses of the program’s memory image. Executable files typically contain absolute code. To let the process execute correctly, the segments must reside at the virtual addresses used to build the executable file. On the other hand, shared object segments typically contain position-independent code. This lets a segment’s virtual address change from one process to another, without invalidating execution behavior. Though the system chooses virtual addresses for individual processes, it maintains the segments’ *relative positions*. Because position-independent code uses relative addressing between segments, the difference between virtual addresses in memory must match the difference between virtual addresses in the file. The difference between the virtual address of any segment in memory and the corresponding virtual address in the file is thus a single constant value for any one executable or shared object in a given process. This difference is the *base address*. One use of the base address is to relocate the memory image of the program during dynamic linking.

An executable or shared object file’s base address is calculated during execution from three values: the virtual memory load address, the maximum page size, and the lowest virtual address of a program’s loadable segment. To compute the base address, one determines the memory address associated with the lowest `p_vaddr` value for a `PT_LOAD` segment. This address is truncated to the nearest multiple of the maximum page size. The corresponding `p_vaddr` value itself is also truncated to the nearest multiple of the maximum page size. The base address is the difference between the truncated memory address and the truncated `p_vaddr` value.

See this chapter in the processor supplement for more information and examples. ‘‘Operating System Interface’’ of Chapter 3 in the processor supplement contains more information about the virtual address space and page size.

Segment Permissions

A program to be loaded by the system must must have at least one loadable segment (although this is not required by the file format). When the system creates loadable segments’ memory images, it gives access permissions as specified in the `p_flags` member.

Figure 5-3: Segment Flag Bits, `p_flags`

Name	Value	Meaning
PF_X	0x1	Execute
PF_W	0x2	Write
PF_R	0x4	Read
PF_MASKPROC	0xf0000000	Unspecified

All bits included in the PF_MASKPROC mask are reserved for processor-specific semantics. If meanings are specified, the processor supplement explains them.

If a permission bit is 0, that type of access is denied. Actual memory permissions depend on the memory management unit, which may vary from one system to another. Although all flag combinations are valid, the system may grant more access than requested. In no case, however, will a segment have write permission unless it is specified explicitly. The following table shows both the exact flag interpretation and the allowable flag interpretation. ABI-conforming systems may provide either.

Figure 5-4: Segment Permissions

Flags	Value	Exact	Allowable
none	0	All access denied	All access denied
PF_X	1	Execute only	Read, execute
PF_W	2	Write only	Read, write, execute
PF_W+PF_X	3	Write, execute	Read, write, execute
PF_R	4	Read only	Read, execute
PF_R+PF_X	5	Read, execute	Read, execute
PF_R+PF_W	6	Read, write	Read, write, execute
PF_R+PF_W+PF_X	7	Read, write, execute	Read, write, execute

For example, typical text segments have read and execute—but not write—permissions. Data segments normally have read, write, and execute permissions.

Segment Contents

An object file segment comprises one or more sections, though this fact is transparent to the program header. Whether the file segment holds one or many sections also is immaterial to program loading. Nonetheless, various data must be present for program execution, dynamic linking, and so on. The diagrams below illustrate segment contents in general terms. The order and membership of sections within a segment may vary; moreover, processor-specific constraints may alter the examples below. See the processor supplement for details.

Text segments contain read-only instructions and data, typically including the following sections described in Chapter 4. Other sections may also reside in loadable segments; these examples are not meant to give complete and exclusive segment contents.

Figure 5-5: Text Segment

`.text`
`.rodata`
`.hash`
`.dynsym`
`.dynstr`
`.plt`
`.rel.got`

Data segments contain writable data and instructions, typically including the following sections.

Figure 5-6: Data Segment

`.data`
`.dynamic`
`.got`
`.bss`

A `PT_DYNAMIC` program header element points at the `.dynamic` section, explained in ‘‘Dynamic Section’’ below. The `.got` and `.plt` sections also hold information related to position-independent code and dynamic linking. Although the `.plt` appears in a text segment above, it may reside in a text or a data segment, depending on the processor. See ‘‘Global Offset Table’’ and ‘‘Procedure Linkage Table’’ in this chapter of the processor supplement for details.

As ‘‘Sections’’ in Chapter 4 describes, the `.bss` section has the type `SHT_NOBITS`. Although it occupies no space in the file, it contributes to the segment's memory image. Normally, these uninitialized data reside at the end of the segment, thereby making `p_memsz` larger than `p_filesz` in the associated program header element.

Note Section

Sometimes a vendor or system builder needs to mark an object file with special information that other programs will check for conformance, compatibility, etc. Sections of type `SHT_NOTE` and program header elements of type `PT_NOTE` can be used for this purpose. The note information in sections and program header elements holds any number of entries, each of which is an array of 4-byte words in the format of the target processor. Labels appear below to help explain note information organization, but they are not part of the specification.

Figure 5-7: Note Information

`namesz`
`descsz`
`type`
`name` . . .
`desc` . . .

`namesz` and `name`

The first `namesz` bytes in `name` contain a null-terminated character representation of the entry's owner or originator. There is no formal mechanism for avoiding name conflicts. By convention, vendors use their own name, such as ‘‘XYZ Computer Company,’’ as the identifier. If no name is present, `namesz` contains 0. Padding is present, if necessary, to ensure 4-byte alignment for the descriptor. Such padding is not included in `namesz`.

`descsz` and `desc`

The first `descsz` bytes in `desc` hold the note descriptor. The ABI places no constraints on a descriptor's contents. If no descriptor is present, `descsz` contains 0. Padding is present, if necessary, to ensure 4-byte alignment for the next note entry. Such padding is not included in `descsz`.

`type`

This word gives the interpretation of the descriptor. Each originator controls its own types; multiple interpretations of a single type value may exist. Thus, a program must recognize both the name and the type to ''understand'' a descriptor. Types currently must be non-negative. The ABI does not define what descriptors mean.

To illustrate, the following note segment holds two entries.

Figure 5-8: Example Note Segment

	+0	+1	+2	+3	
`namesz`	7				
`descsz`	0				No descriptor
`type`	1				
`name`	`X`	`Y`	`Z`		
	`C`	`o`	`\0`	*pad*	
`namesz`	7				
`descsz`	8				
`type`	3				
`name`	`X`	`Y`	`Z`		
	`C`	`o`	`\0`	*pad*	
`desc`	*word 0*				
	word 1				

NOTE The system reserves note information with no name (`namesz==0`) and with a zero-length name (`name[0]=='\0'`) but currently defines no types. All other names must have at least one non-null character.

NOTE

Note information is optional. The presence of note information does not affect a program's ABI conformance, provided the information does not affect the program's execution behavior. Otherwise, the program does not conform to the ABI and has undefined behavior.

Program Loading (Processor-Specific)

NOTE

This section requires processor-specific information. The ABI supplement for the desired processor describes the details.

Dynamic Linking

Program Interpreter

An executable file that participates in dynamic linking shall have one `PT_INTERP` E
program header element. During the function `exec`, the system retrieves a path X
name from the `PT_INTERP` segment and creates the initial process image from the interpreter file's segments. That is, instead of using the original executable file's segment images, the system composes a memory image for the interpreter. It then is the interpreter's responsibility to receive control from the system and provide an environment for the application program.

As ''Process Initialization'' in Chapter 3 of the processor supplement mentions, the interpreter receives control in one of two ways. First, it may receive a file descriptor to read the executable file, positioned at the beginning. It can use this file descriptor to read and/or map the executable file's segments into memory. Second, depending on the executable file format, the system may load the executable file into memory instead of giving the interpreter an open file descriptor. With the possible exception of the file descriptor, the interpreter's initial process state matches what the executable file would have received. The interpreter itself may not require a second interpreter. An interpreter may be either a shared object or an executable file.

- A shared object (the normal case) is loaded as position-independent, with addresses that may vary from one process to another; the system creates its
  segments in the dynamic segment area used by the function `mmap` and X
  related services [see ''Virtual Address Space'' in Chapter 3 of the processor supplement]. Consequently, a shared object interpreter typically will not conflict with the original executable file's original segment addresses.
- An executable file is loaded at fixed addresses; the system creates its segments using the virtual addresses from the program header table. Consequently, an executable file interpreter's virtual addresses may collide with the first executable file; the interpreter is responsible for resolving conflicts.

Dynamic Linker

When building an executable file that uses dynamic linking, the link editor adds a program header element of type PT_INTERP to an executable file, telling the system to invoke the dynamic linker as the program interpreter.

NOTE The locations of the system provided dynamic linkers are processor–specific.

Exec and the dynamic linker cooperate to create the process image for the program, which entails the following actions:

- Adding the executable file's memory segments to the process image;
- Adding shared object memory segments to the process image;
- Performing relocations for the executable file and its shared objects;
- Closing the file descriptor that was used to read the executable file, if one was given to the dynamic linker;
- Transferring control to the program, making it look as if the program had received control directly from the function exec X

The link editor also constructs various data that assist the dynamic linker for executable and shared object files. As shown above in ''Program Header,'' these data reside in loadable segments, making them available during execution. (Once again, recall the exact segment contents are processor-specific. See the processor supplement for complete information.)

- A .dynamic section with type SHT_DYNAMIC holds various data. The structure residing at the beginning of the section holds the addresses of other dynamic linking information.
- The .hash section with type SHT_HASH holds a symbol hash table.
- The .got and .plt sections with type SHT_PROGBITS hold two separate tables: the global offset table and the procedure linkage table. Chapter 3 discusses how programs use the global offset table for position-independent code. Sections below explain how the dynamic linker uses and changes the tables to create memory images for object files.

Because every ABI-conforming program imports the basic system services from a shared object library [see ‘‘System Library’’ in Chapter 6], the dynamic linker participates in every ABI-conforming program execution.

As ‘‘Program Loading’’ explains in the processor supplement, shared objects may occupy virtual memory addresses that are different from the addresses recorded in the file’s program header table. The dynamic linker relocates the memory image, updating absolute addresses before the application gains control. Although the absolute address values would be correct if the library were loaded at the addresses specified in the program header table, this normally is not the case.

If the process environment [see the function `exec`] contains a variable named `LD_BIND_NOW` with a non-null value, the dynamic linker processes all relocation before transferring control to the program. For example, all the following environment entries would specify this behavior. X

- `LD_BIND_NOW=1`
- `LD_BIND_NOW=on`
- `LD_BIND_NOW=off`

Otherwise, `LD_BIND_NOW` either does not occur in the environment or has a null value. The dynamic linker is permitted to evaluate procedure linkage table entries lazily, thus avoiding symbol resolution and relocation overhead for functions that are not called. See ‘‘Procedure Linkage Table’’ in this chapter of the processor supplement for more information.

Dynamic Section

If an object file participates in dynamic linking, its program header table will have an element of type `PT_DYNAMIC`. This ‘‘segment’’ contains the `.dynamic` section. A special symbol, `_DYNAMIC`, labels the section, which contains an array of the following structures.

Figure 5-9: Dynamic Structure

```
typedef struct {
        Elf32_Sword     d_tag;
        union {
                Elf32_Word      d_val;
                Elf32_Addr      d_ptr;
        } d_un;
} Elf32_Dyn;

extern Elf32_Dyn   _DYNAMIC[];
```

For each object with this type, `d_tag` controls the interpretation of `d_un`.

`d_val` These `Elf32_Word` objects represent integer values with various interpretations.

`d_ptr` These `Elf32_Addr` objects represent program virtual addresses. As mentioned previously, a file's virtual addresses might not match the memory virtual addresses during execution. When interpreting addresses contained in the dynamic structure, the dynamic linker computes actual addresses, based on the original file value and the memory base address. For consistency, files do *not* contain relocation entries to ‘‘correct’’ addresses in the dynamic structure.

The following table summarizes the tag requirements for executable and shared object files. If a tag is marked ‘‘mandatory,’’ then the dynamic linking array for an ABI-conforming file must have an entry of that type. Likewise, ‘‘optional’’ means an entry for the tag may appear but is not required.

Figure 5-10: Dynamic Array Tags, `d_tag`

Name	Value	d_un	Executable	Shared Object
DT_NULL	0	ignored	mandatory	mandatory
DT_NEEDED	1	d_val	optional	optional
DT_PLTRELSZ	2	d_val	optional	optional
DT_PLTGOT	3	d_ptr	optional	optional
DT_HASH	4	d_ptr	mandatory	mandatory
DT_STRTAB	5	d_ptr	mandatory	mandatory
DT_SYMTAB	6	d_ptr	mandatory	mandatory

Figure 5-10: Dynamic Array Tags, `d_tag` (continued)

Name	Value	d_un	Executable	Shared Object
DT_RELA‡	7	d_ptr	mandatory	optional
DT_RELASZ	8	d_val	mandatory	optional
DT_RELAENT	9	d_val	mandatory	optional
DT_STRSZ	10	d_val	mandatory	mandatory
DT_SYMENT	11	d_val	mandatory	mandatory
DT_INIT	12	d_ptr	optional	optional
DT_FINI	13	d_ptr	optional	optional
DT_SONAME	14	d_val	ignored	optional
DT_RPATH	15	d_val	optional	ignored
DT_SYMBOLIC	16	ignored	ignored	optional
DT_REL†	17	d_ptr	mandatory	optional
DT_RELSZ	18	d_val	mandatory	optional
DT_RELENT	19	d_val	mandatory	optional
DT_PLTREL	20	d_val	optional	optional
DT_DEBUG	21	d_ptr	optional	ignored
DT_TEXTREL	22	ignored	optional	optional
DT_JMPREL	23	d_ptr	optional	optional
DT_BIND_NOW	24	ignored	optional	optional M
DT_LOPROC	0x70000000	unspecified	unspecified	unspecified
DT_HIPROC	0x7fffffff	unspecified	unspecified	unspecified

† See the description of DT_RELA and DT_REL below for the relationship between these two tags. M

`DT_NULL` An entry with a `DT_NULL` tag marks the end of the `_DYNAMIC` array.

`DT_NEEDED` This element holds the string table offset of a null-terminated string, giving the name of a needed library. The offset is an index into the table recorded in the `DT_STRTAB` entry. See ‘‘Shared Object Dependencies’’ for more information about these names. The dynamic array may contain multiple entries with this type. These entries’ relative order is significant, though their relation to entries of other types is not.

`DT_PLTRELSZ` This element holds the total size, in bytes, of the relocation entries associated with the procedure linkage table. If an entry of type `DT_JMPREL` is present, a `DT_PLTRELSZ` must accompany it.

DT_PLTGOT — This element holds an address associated with the procedure linkage table and/or the global offset table. See this section in the processor supplement for details.

DT_HASH — This element holds the address of the symbol hash table, described in ''Hash Table.'' This hash table refers to the symbol table referenced by the DT_SYMTAB element.

DT_STRTAB — This element holds the address of the string table, described in Chapter 4. Symbol names, library names, and other strings reside in this table.

DT_SYMTAB — This element holds the address of the symbol table, described in Chapter 4, with Elf32_Sym entries for the 32-bit class of files.

DT_RELA — This element holds the address of a relocation table, described in Chapter 4. Entries in the table have explicit addends, such as Elf32_Rela for the 32-bit file class. An object file may have multiple relocation sections. When building the relocation table for an executable or shared object file, the link editor catenates those sections to form a single table. Although the sections remain independent in the object file, the dynamic linker sees a single table. When the dynamic linker creates the process image for an executable file or adds a shared object to the process image, it reads the relocation table and performs the associated actions. If this element is present, the dynamic structure must also have DT_RELASZ and DT_RELAENT elements. When relocation is ''mandatory'' for a file, either DT_RELA or DT_REL must occur (both are permitted but only one is required). M

DT_RELASZ — This element holds the total size, in bytes, of the DT_RELA relocation table.

DT_RELAENT — This element holds the size, in bytes, of the DT_RELA relocation entry.

DT_STRSZ — This element holds the size, in bytes, of the string table.

DT_SYMENT — This element holds the size, in bytes, of a symbol table entry.

DT_INIT — This element holds the address of the initialization function, discussed in ''Initialization and Termination Functions'' below.

DT_FINI — This element holds the address of the termination function, discussed in ''Initialization and Termination Functions'' below.

DT_SONAME This element holds the string table offset of a null-terminated string, giving the name of the shared object. The offset is an index into the table recorded in the DT_STRTAB entry. See ‘‘Shared Object Dependencies’’ below for more information about these names.

DT_RPATH This element holds the string table offset of a null-terminated search library search path string, discussed in ‘‘Shared Object Dependencies.’’ The offset is an index into the table recorded in the DT_STRTAB entry.

DT_SYMBOLIC This element’s presence in a shared object library alters the dynamic linker’s symbol resolution algorithm for references within the library. Instead of starting a symbol search with the executable file, the dynamic linker starts from the shared object itself. If the shared object fails to supply the referenced symbol, the dynamic linker then searches the executable file and other shared objects as usual.

DT_REL This element is similar to DT_RELA, except its table has implicit addends, such as Elf32_Rel for the 32-bit file class. If this element is present, the dynamic structure must also have DT_RELSZ and DT_RELENT elements.

DT_RELSZ This element holds the total size, in bytes, of the DT_REL relocation table.

DT_RELENT This element holds the size, in bytes, of the DT_REL relocation entry.

DT_PLTREL This member specifies the type of relocation entry to which the procedure linkage table refers. The d_val member holds DT_REL or DT_RELA, as appropriate. All relocations in a procedure linkage table must use the same relocation.

DT_DEBUG This member is used for debugging. Its contents are not specified for the ABI; programs that access this entry are not ABI-conforming.

DT_TEXTREL This member’s absence signifies that no relocation entry should cause a modification to a non-writable segment, as specified by the segment permissions in the program header table. If this member is present, one or more relocation entries might request modifications to a non-writable segment, and the dynamic linker can prepare accordingly.

`DT_JMPREL` If present, this entries' `d_ptr` member holds the address of relocation entries associated solely with the procedure linkage table. Separating these relocation entries lets the dynamic linker ignore them during process initialization, if lazy binding is enabled. If this entry is present, the related entries of types `DT_PLTRELSZ` and `DT_PLTREL` must also be present.

`DT_BIND_NOW` If present in a shared object or executable, this entry instructs the dynamic linker to process all relocations for the object containing this entry before transferring control to the program. The presence of this entry takes precedence over a directive to use lazy binding for this object when specified through the environment or via `dlopen`(BA_LIB).

`DT_LOPROC` through `DT_HIPROC`
Values in this inclusive range are reserved for processor-specific semantics. If meanings are specified, the processor supplement explains them.

Except for the `DT_NULL` element at the end of the array, and the relative order of `DT_NEEDED` elements, entries may appear in any order. Tag values not appearing in the table are reserved.

Shared Object Dependencies

When the link editor processes an archive library, it extracts library members and copies them into the output object file. These statically linked services are available during execution without involving the dynamic linker. Shared objects also provide services, and the dynamic linker must attach the proper shared object files to the process image for execution.

When the dynamic linker creates the memory segments for an object file, the dependencies (recorded in `DT_NEEDED` entries of the dynamic structure) tell what shared objects are needed to supply the program's services. By repeatedly connecting referenced shared objects and their dependencies, the dynamic linker builds a complete process image. When resolving symbolic references, the dynamic linker examines the symbol tables with a breadth-first search. That is, it first looks at the symbol table of the executable program itself, then at the symbol tables of the `DT_NEEDED` entries (in order), then at the second level `DT_NEEDED` entries, and so on. Shared object files must be readable by the process; other permissions are not required.

NOTE Even when a shared object is referenced multiple times in the dependency list, the dynamic linker will connect the object only once to the process.

Names in the dependency list are copies either of the `DT_SONAME` strings or the path names of the shared objects used to build the object file. For example, if the link editor builds an executable file using one shared object with a `DT_SONAME` entry of `lib1` and another shared object library with the path name `/usr/lib/lib2`, the executable file will contain `lib1` and `/usr/lib/lib2` in its dependency list.

If a shared object name has one or more slash (`/`) characters anywhere in the name, such as `/usr/lib/lib2` above or `directory/file`, the dynamic linker uses that string directly as the path name. If the name has no slashes, such as `lib1` above, three facilities specify shared object path searching, with the following precedence.

- First, the dynamic array tag `DT_RPATH` may give a string that holds a list of directories, separated by colons (`:`). For example, the string `/home/dir/lib:/home/dir2/lib:` tells the dynamic linker to search first the directory `/home/dir/lib`, then `/home/dir2/lib`, and then the current directory to find dependencies.
- Second, a variable called `LD_LIBRARY_PATH` in the process environment [see the function `exec`] may hold a list of directories as above, optionally followed by a semicolon (`;`) and another directory list. The following values would be equivalent to the previous example: X
  - `LD_LIBRARY_PATH=/home/dir/lib:/home/dir2/lib:`
  - `LD_LIBRARY_PATH=/home/dir/lib;/home/dir2/lib:`
  - `LD_LIBRARY_PATH=/home/dir/lib:/home/dir2/lib:;`

  All `LD_LIBRARY_PATH` directories are searched after those from `DT_RPATH`. Although some programs (such as the link editor) treat the lists before and after the semicolon differently, the dynamic linker does not. Nevertheless, the dynamic linker accepts the semicolon notation, with the semantics described above.
- Finally, if the other two groups of directories fail to locate the desired library, the dynamic linker searches `/usr/lib`.

NOTE For security, the dynamic linker ignores environmental search specifications (such as `LD_LIBRARY_PATH`) for set-user and set-group ID programs. It does, M
however, search `DT_RPATH` directories and `/usr/lib`. The same restriction M
may be applied to processes that have more than minimal privileges on systems with installed extended security systems.

Global Offset Table (Processor-Specific)

NOTE This section requires processor-specific information. The ABI supplement for the desired processor describes the details.

Procedure Linkage Table (Processor-Specific)

NOTE This section requires processor-specific information. The ABI supplement for the desired processor describes the details.

Hash Table

A hash table of `Elf32_Word` objects supports symbol table access. Labels appear below to help explain the hash table organization, but they are not part of the specification.

Figure 5-11: Symbol Hash Table

`nbucket`
`nchain`

`bucket[0]`
. . .
`bucket[nbucket - 1]`

`chain[0]`
. . .
`chain[nchain - 1]`

The bucket array contains nbucket entries, and the chain array contains nchain entries; indexes start at 0. Both bucket and chain hold symbol table indexes. Chain table entries parallel the symbol table. The number of symbol table entries should equal nchain; so symbol table indexes also select chain table entries. A hashing function (shown below) accepts a symbol name and returns a value that may be used to compute a bucket index. Consequently, if the hashing function returns the value *x* for some name, bucket[*x*%nbucket] gives an index, *y*, into both the symbol table and the chain table. If the symbol table entry is not the one desired, chain[*y*] gives the next symbol table entry with the same hash value. One can follow the chain links until either the selected symbol table entry holds the desired name or the chain entry contains the value STN_UNDEF.

Figure 5-12: Hashing Function

```
unsigned long
elf_hash(const unsigned char *name)
{
        unsigned long       h = 0, g;

        while (*name)
        {
                h = (h << 4) + *name++;
                if (g = h & 0xf0000000)
                        h ^= g >> 24;
                h &= ~g;
        }
        return h;
}
```

Initialization and Termination Functions

After the dynamic linker has built the process image and performed the relocations, each shared object gets the opportunity to execute some initialization code. All shared object initializations happen before the executable file gains control.

Before the initialization code for any object A is called, the initialization code for any other objects that object A depends on are called. For these purposes, an object A depends on another object B, if B appears in A's list of needed objects (recorded in the DT_NEEDED entries of the dynamic structure). The order of initialization for circular dependencies is undefined. M M M M M

The initialization of objects occurs by recursing through the needed entries of each object. The initialization code for an object is invoked after the needed entries for that object have been processed. The order of processing among the entries of a particular list of needed objects is unspecified.

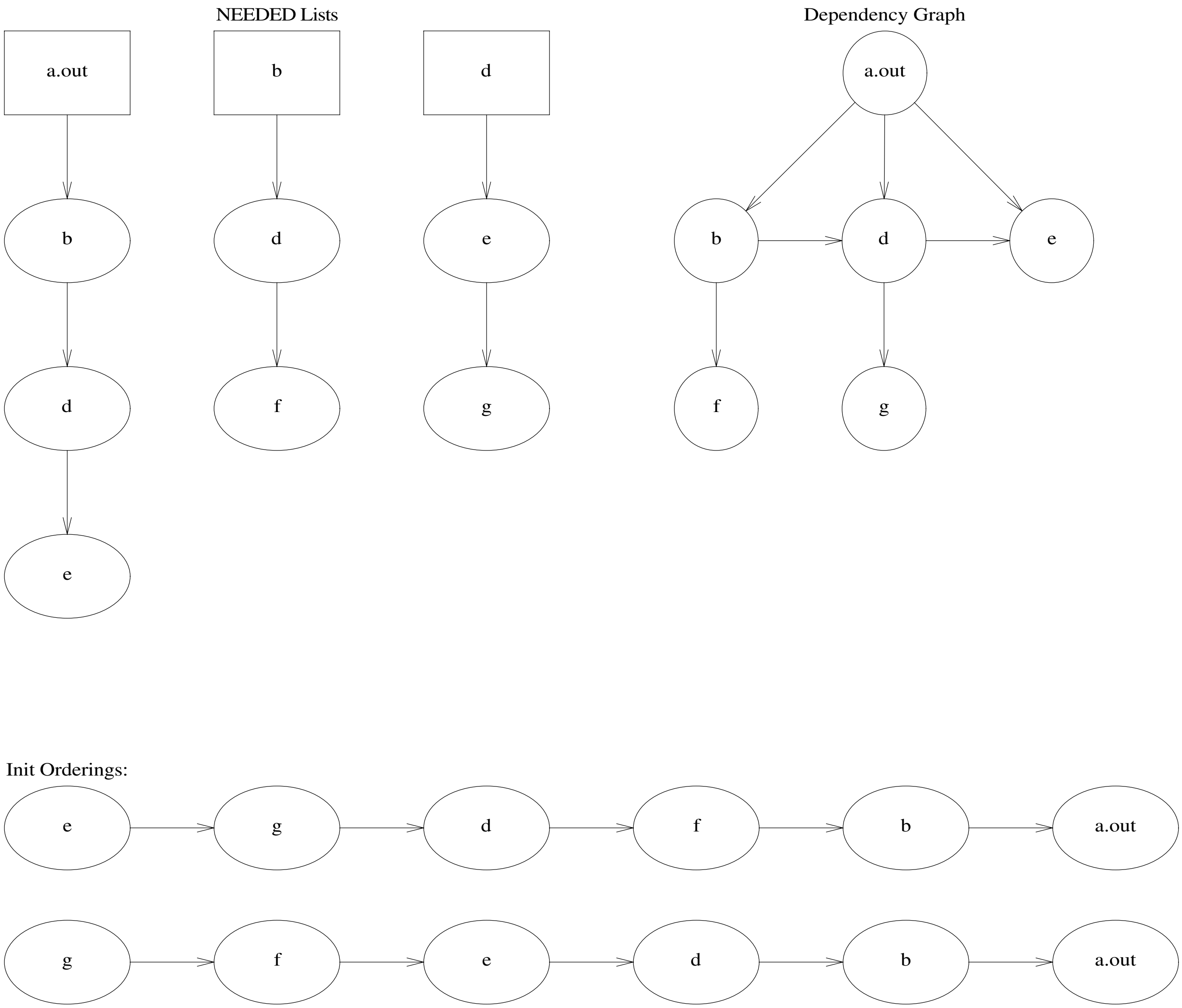
NOTE

Each processor supplement may optionally further restrict the algorithm used to determine the order of initialization. Any such restriction, however, may not conflict with the rules described by this specification.

The following example illustrates two of the possible correct orderings which can be generated for the example NEEDED lists. In this example the a.out is dependent on b, d, and e. b is dependent on d and f, while d is dependent on e and g. From this information a dependency graph can be drawn. The above algorithm on initialization will then allow the following specified initialization orderings among others.

Initialization Ordering Example

Figure 5-13: Initialization Ordering Example



Similarly, shared objects may have termination functions, which are executed with the function `atexit` mechanism after the base process begins its termination sequence. The order in which the dynamic linker calls termination functions is the exact reverse order of their corresponding initialization functions. If a shared object has a termination function, but no initialization function, the termination function will execute in the order it would have as if the shared object's initialization function was present. The dynamic linker ensures that it will not execute any initialization or termination functions more than once.

Shared objects designate their initialization and termination functions through the `DT_INIT` and `DT_FINI` entries in the dynamic structure, described in ‘‘Dynamic Section’’ above. Typically, the code for these functions resides in the `.init` and `.fini` sections, mentioned in ‘‘Sections’’ of Chapter 4.

NOTE Although the function `atexit` termination processing normally will be done, it is not guaranteed to have executed upon process death. In particular, the process will not execute the termination processing if it calls `_exit` [see the function `exit`] or if the process dies because it received a signal that it neither caught nor ignored.

The dynamic linker is not responsible for calling the executable file's `.init` section or registering the executable file's `.fini` section with the function `atexit`. Termination functions specified by users via the `atexit` mechanism must be executed before any termination functions of shared objects.