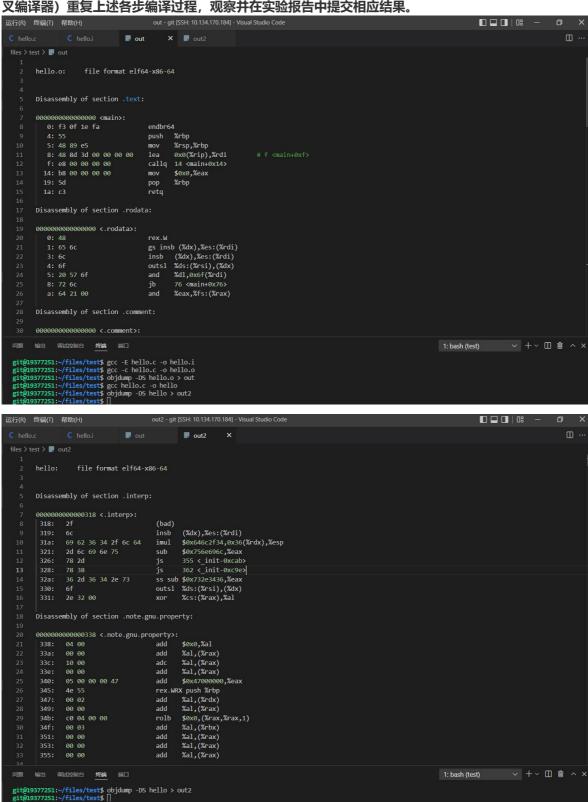
Lab1实验报告

Thinking 1.1

请查阅并给出前述objdump 中使用的参数的含义。使用其它体系结构的编译器(如课程平台的MIPS交叉编译器)重复上述各步编译过程,观察并在实验报告中提交相应结果。



objdump指令详情

--archive-headers

-a

显示档案库的成员信息,类似1s -1将1ib*.a的信息列出。

-b bfdname

--target=bfdname

指定目标码格式。这不是必须的, objdump能自动识别许多格式, 比如:

objdump -b oasys -m vax -h fu.o

显示fu.o的头部摘要信息,明确指出该文件是Vax系统下用Oasys编译器生成的目标文件。objdump -i将给出这里可以指定的目标码格式列表。

-C

--demangle

将底层的符号名解码成用户级名字,除了去掉所开头的下划线之外,还使得C++函数名以可理解的方式显示出来。

--debugging

-g

显示调试信息。企图解析保存在文件中的调试信息并以C语言的语法显示出来。仅仅支持某些类型的调试信息。有些其他的格式被readelf -w支持。

-e

--debugging-tags

类似-g选项,但是生成的信息是和ctags工具相兼容的格式。

--disassemble

-d

从objfile中反汇编那些特定指令机器码的section。

-D

--disassemble-all

与 -d 类似,但反汇编所有section.

-EB

-FI

--endian={big|little}

指定目标文件的小端。这个项将影响反汇编出来的指令。在反汇编的文件没描述小端信息的时候用。例如S-records.

-f

--file-headers

显示objfile中每个文件的整体头部摘要信息。

-h

--section-headers

--headers

显示目标文件各个section的头部摘要信息。

-H

--help

简短的帮助信息。

-i

--info

显示对于 -b 或者 -m 选项可用的架构和目标格式列表。

-j name

--section=name

仅仅显示指定名称为name的section的信息

-1

--line-numbers

用文件名和行号标注相应的目标代码,仅仅和-d、-D或者-r一起使用使用-1d和使用-d的区别不是很大,在源码级调试的时候有用,要求编译时使用了-q之类的调试编译选项。

-m machine

--architecture=machine

指定反汇编目标文件时使用的架构,当待反汇编文件本身没描述架构信息的时候(比如S-records),这个选项很有用。可以用-i选项列出这里能够指定的架构.

--reloc

-r

显示文件的重定位入口。如果和-d或者-D一起使用,重定位部分以反汇编后的格式显示出来。

--dynamic-reloc

-R

显示文件的动态重定位入口,仅仅对于动态目标文件意义,比如某些共享库。

-s

--full-contents

显示指定section的完整内容。默认所有的非空section都会被显示。

-S

--source

尽可能反汇编出源代码,尤其当编译的时候指定了-g这种调试参数时,效果比较明显。隐含了-d参数。

--show-raw-insn

反汇编的时候,显示每条汇编指令对应的机器码,如不指定--prefix-addresses,这将是缺省选项。

--no-show-raw-insn

反汇编时,不显示汇编指令的机器码,如不指定--prefix-addresses,这将是缺省选项。

--start-address=address

从指定地址开始显示数据,该选项影响-d、-r和-s选项的输出。

--stop-address=address

显示数据直到指定地址为止,该项影响-d、-r和-s选项的输出。

-+

--syms

显示文件的符号表入口。类似于nm -s提供的信息

-T

--dynamic-syms

显示文件的动态符号表入口,仅仅对动态目标文件意义,比如某些共享库。它显示的信息类似于 nm -D|--dynamic 显示的信息。

-V

--version

版本信息

--all-headers

->

显示所可用的头信息,包括符号表、重定位入口。-x 等价于-a -f -h -r -t 同时指定。

-7

--disassemble-zeroes

一般反汇编输出将省略大块的零,该选项使得这些零块也被反汇编。

@file 可以将选项集中到一个文件中,然后使用这个@file选项载入。

Thinking 1.2

也许你会发现我们的readelf程序是不能解析之前生成的内核文件(内核文件是可执行文件)的,而我们之后将要介绍的工具readelf则可以解析,这是为什么呢?(提示:尝试使用readelf-h,观察不同)

由于生成的vmlinux内核文件是大端读取,而readelf.c是小段读取,因此无法解析。同时因为这个readelf只能读取32位elf文件,因此也不能用它读取在实验平台上例如gcc生成的elf文件。

Thinking 1.3

在理论课上我们了解到,MIPS 体系结构上电时,启动入口地址为0xBFC00000(其实启动入口地址是根据具体型号而定的,由硬件逻辑确定,也有可能不是这个地址,但一定是一个确定的地址),但实验操作系统的内核入口并没有放在上电启动地址,而是按照内存布局图放置。思考为什么这样放置内核还能保证内核入口被正确跳转到?

(提示:思考实验中启动过程的两阶段分别由谁执行。)

由于boot loader的第二部分stage2中设置了cpu指令寄存器的内容为内核入口函数的地址,这个地址就是内存布局图中内核所在的地址。而且stage2运行在ram中,此时拥有足够的运行环境,这个c环境会帮助找到内核对应的地址。

Thinking 1.4

sg_size 和bin_size 的区别它的开始加载位置并非页对齐,同时bin_size的结束位置(va+i)也并非页对齐,最终整个段加载完毕的sg_size 末尾的位置也并非页对齐,请思考,为了保证页面不冲突(不重复为同一地址申请多个页,以及页上数据尽可能减少冲突),这样一个程序段应该怎样加载内存空间中。

彻底并透彻地理解上图能帮助大家在后续实验中节约很多时间

若需要加载的程序段为p,遍历其他所有进程,对于其中的进程s,如果s.sg_size结束位置s.addr+s.sg size到s.addr+s.sg size+p.sg size之间不存在进程,那么就可以把p加载到这里。

Thinking 1.5

内核入口在什么地方? main 函数在什么地方? 我们是怎么让内核进入到想要的 main 函数的呢? 又是怎么进行跨文件调用函数的呢?

内核的入口在boot/start.S

main函数在 init/main.c

内核进入main的方式:设置sp栈指针,然后跳转到main函数

通过链接器来跨文件调用函数

Thinking 1.6

查阅《See MIPS Run Linux》一书相关章节,解释boot/start.S 中下面几行对CPO 协处理器寄存器进行读写的意义。具体而言,它们分别读/写了哪些寄存器的哪些特定位,从而达到什么目的?

```
/* Disable interrupts */
mtc0 zero, CP0_STATUS
.....
/* disable kernel mode cache */
mfc0 t0, CP0_CONFIG
and t0, ~0x7
ori t0, 0x2
mtc0 t0, CP0_CONFIG
```

第一条指令把cpu通用寄存器s的内容传送到协处理器0寄存器n,数据为32位。

第二条指令把通用寄存器d装入cpu控制寄存器n的值,这是查看控制寄存器的值的唯一方法。后续是变更新控制寄存器内部的单个域。

Exercise 1.1

请修改include.mk文件,使交叉编译器的路径正确。之后执行make指令,如果配置一切正确,则会在gxemul目录下生成vmlinux的内核文件。

直接增加路径即可

Exercise 1.2

阅读./readelf 文件夹中 kerelf.h、readelf.c 以及 main.c 三个文件中的代码,并完成 readelf.c 中缺少的代码,readelf 函数需要输出 ELF 文件的所有 section header 的序号和其中记录的相应section 的地址信息,对每个 section header,输出格式为:"%d:0x%x\n",两个标识符分别代表序号和地址。序号即第几个section header,从0开始。

正确完成 readelf.c 代码之后,在 readelf 文件夹下执行 make 命令,即可生成可执行文件 readelf,它接受文件名作为参数,对 ELF文件进行解析**。**

通过查看elf文件说明来了解32位elf文件头的内容,然后编写c代码即可。

实验难点:理解elf文件头内的内容,比如section头表偏移量,section头表中表项数量以及读取section节内容的方法。

首先从文件头中获得段头表相对文件头地址的偏移量和文件头地址相加获得段头表开始的地址,之后根据文件头中获得的段头表入口大小和数量循环每个段头开始处的偏移量在于段投标开始地址相加得到段头的真实地址。

Exercise 1.3

填写tools/scse0_3.lds中空缺的部分,将内核调整到正确的位置上。**在 lab1 中,只需要填补.text、.data 和.bss 段将内核调整到正确的位置上即可。**

模仿指导书中的写法设置好位置即可,值得注意的是他们在内核中的位置在内存布局图中可以找到。

Exercise 1.4

完成boot/start.S中空缺的部分。设置栈指针,跳转到main函数。 使用/OSLAB/gxemul -E testmips -C R3000 -M 64 elf-file运行(其中elf-file是你编译生成的vmlinux文件的路径)。关于如何退出gxemul,请看教程中"如何退出gxemul"一节。

在内存布局图中找到sp指针指向的位置,设置好然后跳转main函数即可。

Exercise 1.5

阅读相关代码和下面对于函数规格的说明,补全lib/print.c中lp_Print()函数中缺失的部分来实现字符输出。

根据指导书中关于printf的说明完成即可。依次读取前面%[flags][width][.precision][length]specifier的值

实验难点:每次fmt匹配后需要fmt++使其指向下一个字符。代码中很多变量没有初始化,因此在最开始应对他们赋上合适的初始值,例如padc初始值为空格字符。最后应当所有变量都有初始值。

%d中需要判断输出的内容是否为负数,同时要把其变回正数,在调用的函数中使用negflag增加负号。

Makefile

变量

变量名	含义
\$<	第一个依赖文件
\$@	目标
\$^	所有不重复的依赖文件,以空格分开

ELF文件

文件头

定义

```
#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_shentsize;
Elf32_Half e_shentsize;
Elf32_Half e_shentsize;
Elf32_Half e_shentsize;
Elf32_Half e_shrum;
Elf32_Half e_shrum;
Elf32_Half e_shstrndx;
} Elf32_Ehdr;
```

含义

• **e_ident:**文件的前16个字节包含ELF文件的识别标志,而且提供了用于解码和解析文件内容的数据,不依赖操作系统。

索引	名称	含义
0	EI_MAG0	文件标志,固定值0x7f
1	EI_MAG1	文件标志,固定值'E'
2	EI_MAG2	文件标志,固定值'L'
3	EI_MAG3	文件标志,固定值'F'
4	EL_CLASS	文件类型, 64 (2) 或32 (1) 位
5	EL_DATA	编码格式,LSB小头编码(1)MSB大头编码(2)
6	EL_VERSION	文件头版本1
7-15	EL_PAD	暂不使用

- e_type: 2bytes,表明本文本文件时遇哪种类型,如:可重定位文件,可执行文件等。
- e_machine: 2bytes用于指定该文件适用的处理器体系结构
- e_version: 4bytes指明目标文件的版本
- e_entry: 4 (64位 8) bytes程序入口的虚拟地址
- e_phoff: 4 (64位 8) bytes程序头表在文件中的偏移量
- e_shoff: 4 (64位 8) bytes节头表在文件中的偏移量
- e_flags: 4bytes处理器特定的标志位
- e_ehsize: 2bytesELF文件头大小
- **e_phentsize**: **2bytes**程序头表(program header table)中每个表项大小
- e_phnum: 2bytes程序头表中有多少个表项
- e_shentsize: 2bytes节头表(section header table)中每一个表项的大小
- e_shnum: 2bytes节头表中一共有多少个表项
- e_shstrndx: 2bytes节头表中与节名字表对应的表项的索引

- sh_name 本节的名字,是一个索引号,指向字符串表中的一个位置
- sh_type 本节的所有类型
- sh_flags 本节的一些属性
- sh_addr 若本节的内容需要映射到进程空间取,则指定映射的起始位置,否则为0
- sh_offset指明了本节所在的位置,相对于文件开头的偏移量
- sh_size节的大小,单位是字节
- sh_link索引值,指向节头表中本节所对应的位置
- sh_info 有关此节的附加信息
- sh_addralign 至命本节内容如何对齐字节
- **sh_entsize** 一些节的内容是一张表,其中每个表项大小固定,对于这种表,指定表项大小。此值为0说明不是表。