操作系统实验报告

院系: 计算机学院

班级: 计科2班

学号: 19335174

姓名: 施天予

指导老师:凌应标

一、实验题目

独立内核开发

二、实验目的

- 1、熟悉掌握 C 语言和汇编的混合编译
- 2、将实验二的操作系统分离为引导程序和内核,用引导程序加载内核

三、实验要求

- 1、寻找或测试一套匹配的汇编与 c 编译器组合。利用 c 编译器,将一个样板 C 程序进行编译,获得符号列表文档,分析全局变量、局部变量、变量初始化、函数调用、参数传递情况,确定一种匹配的汇编语言工具,在实验报告中描述这些工作。
- 2、写一个汇编程和 c 程序混合编程实例,展示你所用的这套组合环境的使用。 汇编模块中定义一个字符串,调用 C 语言的函数,统计其中某个字符出现的次数(函数返回),汇编模块显示统计结果。执行程序可以在 DOS 中运行。
- 3、重写实验二程序,实验二的的监控程序从引导程序分离独立,生成一个 COM 格式程序的独立内核,在 1.44MB 软盘映像中,保存到特定的几个扇区。利用汇编程和 c 程序混合编程监控程序命令保留原有程序功能,如可以按操作选择,执行一个或几个用户程序、加载用户程序和返回监控程序;执行完一个用户程序后,可以执行下一个。
- 4、利用汇编程和 c 程序混合编程的优势,多用 c 语言扩展监控程序命令处理能力。
- 5、重写引导程序,加载 COM 格式程序的独立内核。
- 6、拓展自己的软件项目管理目录,管理实验项目相关文档。

四、实验方案

【实验环境】

1、实验运行环境: Windows10

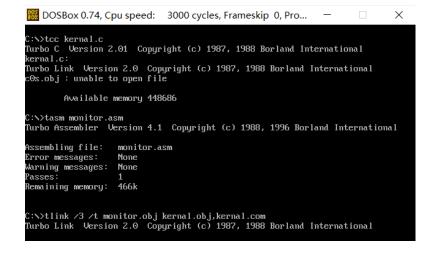
- 2、虚拟机软件: VirtualBox
- 3、NASM编译器
- 4、DOSBox 下 TCC+TASM+TLINK 混合编译

【实验工具】

- 1、汇编语言: NASM, TASM
- 2、文本编辑器: Notepad++
- 3、相关软件: WinHex

【实验思想】

- 1、汇编与 C 语言的关系
 - 1) 汇编语言的目的
 - a. 设置自身运行模式和环境,设置硬件寄存器
 - b. 设置 I/O 端口和实现 I/O 操作
 - c. 初始化中断向量表和实现中断处理
 - d. 实现控制原语
 - 2) C语言的目的
 - a. 适合构造复杂的数据结构和相关数据结构的管理
 - b. 实现复杂的功能或算法
- 2、汇编与 C 语言的混合编译链接
 - 1) TCC 编译 C 语言程序
 - 2) TASM 编译汇编语言程序
 - 3) TLINK 链接 C 和汇编生成的 obj文件,生成 COM 可执行文件



3)编译步骤

将老师的 tcc-tasm 文件夹中的六个文件放入之前编译 nasm 的文件夹下,在 DOSBox 中定位到该文件夹,就可以使用 TCC, TASM, TLINK 了(使用 TLINK 时要把汇编生成的 obj文件写在前面,C生成的 obj文件写在后面)。

五、实验过程

1、样板程序的编译与分析(将字符串大写)

【源代码】

```
1 ;程序源代码 (showstr.asm)
2 extrn upper:near ;声明一个c程序函数upper
3 extrn Message:near ;声明一个外部变量
  .8086
4
5
  TEXT segment byte public 'CODE'
6 DGROUP group _TEXT,_DATA,_BSS
7
         assume cs: TEXT
8 org 100h
9 start:
10
      mov
           ax,
               CS
11
      mov ds, ax
                          ; DS = CS
                          ; ES = CS
12
      mov es, ax
13
      mov ss, ax
                          ; SS = cs
14
      mov sp, 100h
15
      call near ptr _upper
16
      mov bp, offset _Message ; BP=当前串的偏移地址
17
                         ; ES:BP = 串地址
      mov ax, ds
18
                            ; 置ES=DS
      mov es, ax
   mov cx, 18
                           ; CX = 串长 (=18)
19
20
      mov ax, 1301h
21
      ; AH = 13h (功能号)、AL = 01h (光标置于串尾)
22
      mov bx, 0007h
      ; 页号为0(BH = 0) 黑底白字(BL = 07h)
23
      mov dh, 20 ; 行号=10 mov dl, 10 ; 列号=10
24
25
      mov dl, 10
26
      int 10h ; BIOS的10h功能: 显示一行字符
27
      jmp $
28
29
  TEXT ends
30 ; ***********DATA segment*********
31 DATA segment word public 'DATA'
32
   DATA ends
33 ; *************BSS segment*********
34 BSS
          segment word public 'BSS'
  BSS ends
36 ;***********end of file*******
37 end start
```

```
/*程序源代码(upper.c)*/
2
    char Message[18]="OSaaabbbAaBbCcDdEe";
3
   □upper(){
4
        int i=0;
5
        while (Message[i]) {
          if (Message[i]>='a'&&Message[i]<='z')</pre>
6
7
           Message[i]=Message[i]+'A'-'a';
8
           i++;
9
10
```

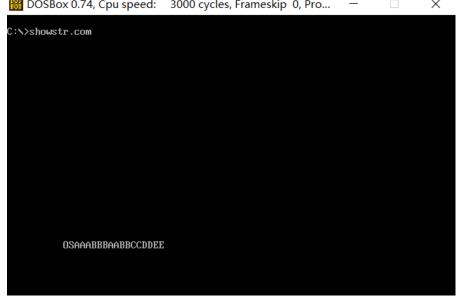
【编译运行】

```
C:\>tcc -mt -c -oupper.obj upper.c
Turbo C Version 2.01 Copyright (c) 1987, 1988 Borland International
upper.c:

Available memory 456414

C:\>tasm showstr.asm showstr.obj
Turbo Assembler Version 4.1 Copyright (c) 1988, 1996 Borland International
Assembling file: showstr.asm
Error messages: None
Warning messages: None
Passes: 1
Remaining memory: 468k

C:\>tlink /3 /t showstr.obj upper.obj,showstr.com
Turbo Link Version 2.0 Copyright (c) 1987, 1988 Borland International
```



【变量与传参分析】

全局变量: Message 是 C 中定义的字符串,在汇编程序中用 extrn _Message: near 声明,C 程序中的变量名在汇编要加下划线。

局部变量: i 用于 upper 函数中的一个循环变量。

变量初始化: Message 在 C 中初始化为"OSaaabbbAaBbCcDdEe"。

函数调用: 汇编调用 C 的函数要先声明 extrn _upper: near, 调用 upper 函数 call near ptr _upper, 调用 C 中的函数也要在函数名前加下划线。

参数传递情况: 参数传递时,用 push word ptr _DATA:_b 就是实参 b 压栈,而 push word ptr _DATA:_a 就是实参 a 压栈,然后调用函数 f(a,b),之后两个 POP 指令是调用后清除栈中的参数。说明调用 C 函数时,参数按后面参数先进栈的顺序压栈,进栈出栈以字为单位。

【列表文件】

在 DOSBox 下运用 tcc 将 upper. c 编译成 upper. asm 文件

```
C:\>tcc -mt -S -oupper.asm upper.c
Turbo C Version 2.01 Copyright (c) 1987, 1988 Borland International
upper.c:
Available memory 458678
```

tcc -mt -S -oupper.asm upper.c

```
DATA
        segment word public 'DATA'
Message
            label
                     byte
   db
        79
       83
   db
       97
   db
       97
   db
   db
       97
   db
       98
   db
       98
   db
       98
       65
   db
   db
       97
   db
       66
   db
       98
   db
       67
   db
       99
   db
       68
   db
       100
   db
       69
       101
   db
DATA
        ends
```

可以看到初始化的字符串变量 Message 被放在数据段,如果未初始化会放在 BSS 段。

2、混合编程实例(统计字符次数)

【程序分析】

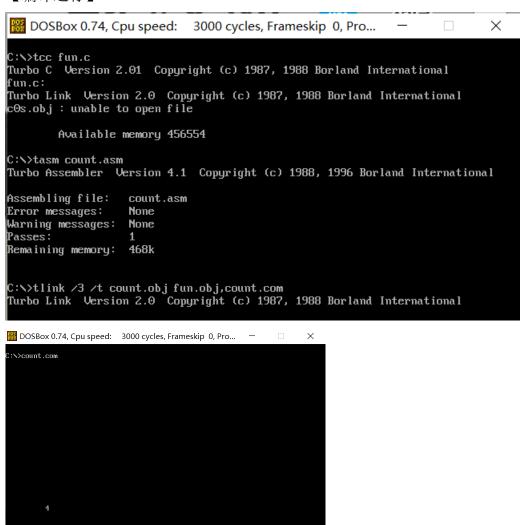
C 中的 fun 函数用于统计字符串 Message 中 a 的个数。在汇编程序中先用 public 声明字符串变量 Message 为全局变量,在 datadef 中定义 Message 并初始化。用 extrn 声明 C 中的 fun 函数,,在 C 中用 extern char Message [] 声明 汇编中定义的 Message 字符串变量。再用 call 调用 fun 函数,利用 BIOS 的 int 10h 输出字符串功能,设置串长 CX=1(统计结果是一个字符),成功实现。

【设计代码】

```
1 ;程序源代码 (count.asm)
2 extrn fun:near
3 public Message
4
5
   TEXT segment byte public 'CODE'
6
   DGROUP group _TEXT,_DATA,_BSS
    assume cs: TEXT
9
  org 100h
10
11
  start:
12
     mov ax, cs
                          ; DS = CS
13
      mov ds, ax
                          ; ES = CS
14
      mov es, ax
15
      mov ss, ax
                          ; SS = cs
16
      mov sp, 100h
17
      call near ptr _fun
18
      mov bp, offset _Message ; BP=当前串的偏移地址
19
      mov ax, ds ; ES:BP = 串地址
20
                        ; 置ES=DS
      mov es, ax
                        ; CX = 串长 (=1)
21
      mov cx, 1
22
      mov ax, 1301h
23
      ; AH = 13h (功能号)、AL = 01h (光标置于串尾)
24
      mov bx, 0007h
25
      ; 页号为0(BH = 0) 黑底白字(BL = 07h)
      mov dh, 20 ; 行号=10 mov dl, 10 ; 列号=10
26
27
28
      int 10h ; BIOS的10h功能:显示一行字符
29
      jmp $
30
   datadef:
31
      Message db "OSaaabbbAaBbCcDdEe", 0
32
33
34
   TEXT ends
35 ; ************DATA segment**********
36
  DATA segment word public 'DATA'
37
   DATA ends
38 ;***********BSS segment********
          segment word public 'BSS'
39
   BSS
40
   BSS ends
41 ;***********end of file*******
42 end start
```

```
/*程序源代码(fun.c)*/
1
2
    extern char Message[];
3
4
   □fun(){
 5
         int i=0, k=0;
 6
         while (Message[i]) {
7
             if (Message[i]=='a') k++;
             i++;
9
         }
10
         Message[0]=k+'0';
11
```

【编译运行】



"OSaaabbbAaBbCcDdEe"中 a 的数量是 4, 结果正确!

注意:_Message 必须定义在数据段_datadef,如果定义在代码段,会显示乱码!

3、生成独立内核

1) 引导程序(bootloader.asm)

上次实验是把引导程序和监控程序合在一起,这次分离后引导程序只要跳转到内核所在的扇区即可。我把引导程序放在首扇区,4个用户程序放在2、

3、4、5扇区,内核放在第6扇区起始,内核的代码一共占了4个扇区

```
;BIOS将把引导扇区加载到0:7C00h处,并开始执行
1 org 7c00h
2 OffsetOfMonitor equ OA100h ;加载操作系统内核
4 ReadOs:
                ;段地址; 存放数据的内存基地址,设置原地址
     mov ax,cs
     mov es, ax ;设置段地址 (不能直接mov es,段地址)
mov bx, OffsetOfMonitor ;存放内核的内存偏移地址OffsetOfMonitor
     mov ah,2
              ;功能号
8
                         ;扇区数,内核占用扇区数 注意: 不止加载了一个扇区
9
     mov al, 4
                        ;驱动器号; 软盘为0, 硬盘和U盘为80H
     mov dl,0
             ; 磁头号; 起始编号为0
; 柱面号; 起始编号为0
     mov dh, 0
12
     mov ch, 0
                   ;存放内核的起始扇区号6
     mov cl,6
13
                      ;调用读磁盘BIOS的13h功能
;控制权移交给内核
14
      int 13H
     jmp 0a00h:100h
15
16
     times 510 - ($ - $$) db 0;将前510字节不是0就填0
18
      db 0x55, 0xaa
```

2) 内核

a) 监控程序(monitor.asm)

监控程序的代码参考了老师的 afile. asm。其中 include kliba. asm 获取汇编库函数,并声明 extern _cmain: near,使用 call near ptr _cmain来调用 kernal.c 中的 cmain 函数进入操作系统主界面。

```
extern macro %1 ;统一用extern导入外部标识符
      extrn %1
   endm
  extern cmain:near
9 .8086
10 _TEXT segment byte public 'CODE'
11 assume cs: TEXT
12 DGROUP group _TEXT,_DATA,_BSS
  org 100h
13
14
15 start:
17
18
19
      mov ax,cs
20
      mov ds, ax; DS = CS
      mov es, ax; ES = CS
22
      mov ss, ax; SS = cs
23
      mov sp, 0FFF0h
24
      mov ah,2
25
      mov bh,0
      mov dx,0
26
      int 10h
27
28
```

```
call near ptr cmain
     jmp $
31
32
    include kliba.asm
33 ;
34 ;***********************
35
36
   TEXT ends
   IBAT CHUS
37
   _DATA segment word public 'DATA'
38
   39
40 DATA ends
41
42 BSS segment word public 'BSS'
43 BSS ends
44
45 end start
46 ;**********************
47
```

b) 汇编库函数(kliba.asm)

这段代码在老师的 kliba. asm 的基础上修改。 实现了以下四个函数:

_cls:清屏

loadUser: 跳转到指定扇区加载运行用户程序

getChar: 从键盘读取一个字符

printChar: 显示一个字符

```
36 ;****************
37 ; void cls()
38 ;****************
39 public cls
40 cls proc
41 ; 清屏
42
             push ax
43
            push bx
44
            push cx
45
             push dx
                 mov ax, 600h ; AH = 6, AL = 0
mov bx, 700h ; 黑底白字(BL = 7)
mov cx, 0 ; 左上角: (0, 0)
46
47
48
                 mov dx, 184fh ; 右下角: (24, 79)
int 10h ; 显示中断
49
50
51
                 mov ah,2
52
                 mov bh, 0
53
                 mov dx, 0
54
                      int 10h
             pop dx
55
56
             pop cx
57
             pop bx
58
             pop ax
59
60 cls endp
```

```
87 ; void _loadUser();
89 OffSetOfUserPrg equ 8100h
90 public _loadUser
91 loadUser proc
92
     push ds
93
     push es
94
     push bp
95
     mov bp,sp
                      ;段地址; 存放数据的内存基地址
96
     mov ax,cs
     mov es, ax ;设置段地址(不能直接mov es,段地址)
mov bx, OffSetOfUserPrg ;偏移地址;存放数据的内存偏移地址
97
98
                       ;功能号
99
     mov ah,2
100
                      ;扇区数
     mov al,1
                      ;驱动器号; 软盘为0, 硬盘和U盘为80H
101
     mov dl, 0
                      ;磁头号;起始编号为0
;柱面号;起始编号为0
102
     mov dh,0
103
     mov ch,0
     mov cl, [bp+8]
                      ;用户程序在2,3,4,5扇区
104
                      调用读磁盘BIOS的13h功能
105
     int 13H ;
     mov bx,OffSetOfUserPrq
106
107
     call bx
108
     pop bp
109
     pop es
     pop ds
     ret
   loadUser endp
61
    62 ; void getChar()
   63
64 public _getChar
65
   getChar proc
66
       mov ah, 0
67
       int 16h
68
       mov byte ptr [_x],al
69
       ret
70 getChar endp
72 ; void printChar(char ch);
74 public printChar
75 _printChar proc
76
      push bp
77
      mov bp, sp
    mov al,[bp+4];char\ip\bp
78
79
     mov bl,0
      mov ah, 0eh
81
      int 10h
82
      mov sp,bp
83
      pop bp
84
      ret
   printChar endp
```

_cls 基于老师给的大致没有改变。_loadUser 利用压栈实现参数传递,参数就是跳转的扇区号。_getChar 函数使用 int 16h 中断指令,_x 是 kernal.c

中定义的字符,用于读入后传到 C 库中。_printChar 函数使用 int 13h 的 0eh 功能输出字符。

c) C库函数(kernel.c)

printf: 输出字符串

在 C 语言程序中, 实现了四个基于汇编库函数的拓展函数:

strcmp: 判断字符串是否相等 InputCom: 从键盘读取命令 RunCom: 运行命令函数

```
⇒void printf(char* str) {
     int i=0;
     while(str[i]) {
         printChar(str[i]);
         i++;
     }
└ }
int strcmp(char* a,char* b) {
    int i=0, j=0;
    while(a[i]!=0&&b[j]!=0) {
        if(a[i]!=b[j])
            return 0;
        i++;
        j++;
    if(a[i]==0&&b[j]==0)
        return 1;
        return 0;
}
```

在 kliba. asm 中实现了一个字符的输入与输出,在 C中只要循环调用就能变成 printf 输出字符串。

```
void InputCom(char* command) {
    int i=0;
    getChar();
    while(x!=13) {
        if(x==8) {
            printChar(x);
            printChar(32);
            i--;
            printChar(x);
            getChar();
            continue;
        printChar(x);
        command[i]=x;
        i++;
        getChar();
    command[i]='\0';
}
```

```
void RunCom(char *command) {
    if (strcmp (command, "help"))
        help();
    else if(strcmp(command, "cls"))
       cls();
    else if(strcmp(command, "ls"))
        file();
    else if(strcmp(command,"u1")) {
        cls();
        loadUser(2);
        cls();
    else if(strcmp(command,"u2")) {
        cls();
        loadUser(3);
        cls();
    else if(strcmp(command, "u3")) {
        cls();
        loadUser(4);
        cls();
    else if(strcmp(command,"u4")) {
        cls();
        loadUser(5);
        cls();
    else if(strcmp(command, "runall")) {
        cls();
        tasks = 4;
        cls();
    else if(strcmp(command, "quit"))
        symbol=0;
        printf("\n\rWrong command!\n\r");
}
```

这是关于命令的读取与执行的两个函数。在 InputCom 函数中我用 ASCII 是 否为 8 判断退格,实现前一个填空白,防止裸机操作系统的退格让光标后移。在 RunCom 函数中用 strcmp 判断不同的命令,从而执行不同的操作,各种命令的含义在后面会讲到,如果命令是无效的,那么会输出"Wrong Command!"

d) 操作系统内核的功能(kernel.c)

Screen(): 在进入操作系统时显示的一个简单界面,用 printf 即可完成

help():帮助菜单,可以告诉用户可以执行哪些命令。Help用于获取帮助菜单,cls用于清屏,ul-u4是分别执行4个用户程序,runall是依次执行一遍用户程序ul-u4,ls用于获取文件列表,quit退出操作系统。

```
void help() {
    printf("\n\rhelp: get the command list\n\r");
    printf("cls: clear the screen\n\r");
    printf("ul:
                    run the user program 1\n\r");
    printf("u2:
                    run the user program 2\n\r");
    printf("u3:
                    run the user program 3\n\r");
    printf("u4:
                   run the user program 4\n\r");
    printf("runall: run all the user programs\n\r");
    printf("ls:
                   get the file list\n\r");
    printf("quit:
                    exit the OS\n\r");
}
```

file(): 一个文件列表。利用 printf 函数显示 4 个用户程序的文件大小。

cmain 主程序: 监控程序会调用 cmain 来实现操作系统的正常运行, cmain 里可以调用各种函数。在 cmain 中先调用 Screen 函数显示个性界面。Symbol 是一个标志变量,当输入命令 quit 时 symbol=0 从而跳出 while 循环,退出操作系统。对于命令 runall,我用一个变量 tasks 使其保留剩余要执行的用户程序个数,在每次 while 循环中如果 tasks 不为 0 就执行用户程序,否则就输入新的命令。因为一共有 4 个用户程序,所以用 loadUser (6-tasks) 可以完美按顺序执行 ul-u4 用户程序(用户程序在 2-5 扇区),用一个 getChar 函数在每次用户程序返回时停顿,否则在执行完第一个用户程序后后面的用户程序只会一闪而过。

```
cmain(){
    cls();
    Screen();
    symbol=1;
    while(symbol) {
        if (tasks!=0) {
            loadUser(6-tasks);
            tasks--;
            getChar();
            continue;
        printf("\n\r>>");
        InputCom(command):
        RunCom (command);
   printf("\n\rThank you for using!\n\rSee you next time!\n\r");
    getChar();
}
```

3) 用户程序

用户程序沿用了实验二的 4 个用户程序。唯一的改进是使用了 call 而不是 jmp 跳入用户程序,在用户程序中也用 ret 进行返回。用户程序在 2-5 扇区。在用户程序界面按空格返回操作系统界面。

```
;输入空格后弹出程序
150
       mov ah, 1
151
       int 16h
       mov bl,20h
152
153
       cmp al,bl
154
       jz Quit
155
        jmp loop1
156
157 Quit:
158
       ret
159 end:
160
                              ;停止画框,无限循环
        jmp $
```

4) NASM 编译与 TCC-TASM-TLINK 混合编译链接

用户程序和引导程序用 NASM 编译

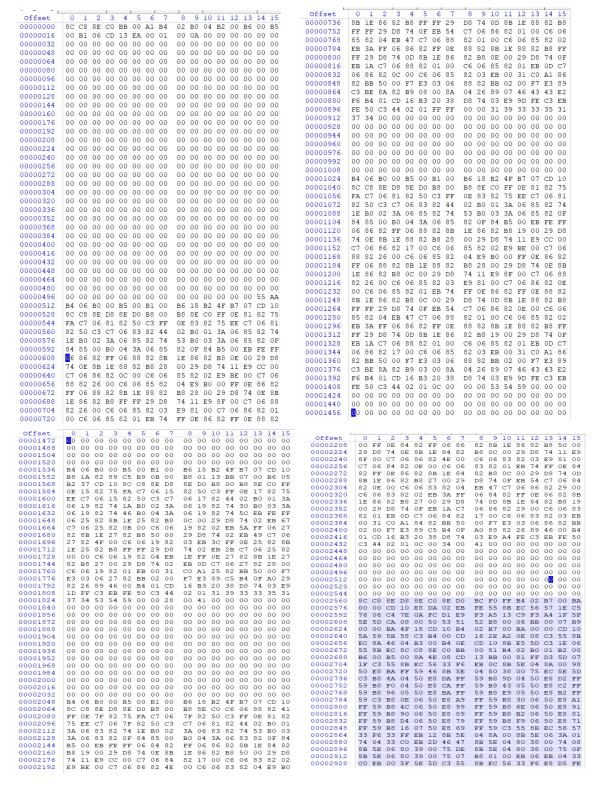
```
Microsoft Windows [版本 10.0.18363.1441]
(c) 2019 Microsoft Corporation。保留所有权利。
E:\nasm>nasm -f bin 1.asm -o 1.bin
E:\nasm>nasm -f bin 2.asm -o 2.bin
E:\nasm>nasm -f bin 3.asm -o 3.bin
E:\nasm>nasm -f bin 4.asm -o 4.bin

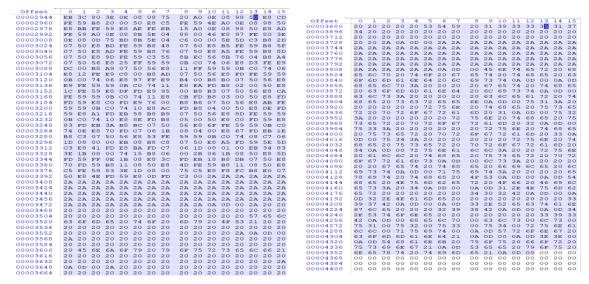
E:\nasm>nasm -f bin bootloader.asm -o bootloader.bin
```

```
内核的三个程序用 TCC-TASM-TLINK 混合编译链接
BOSBox 0.74, Cpu speed: 3000 cycles, Frameskip 0, Pro...
                                                                                 X
C:\>tcc kernal.c
Turbo C Version 2.01 Copyright (c) 1987, 1988 Borland International
kernal.c:
Turbo Link Version 2.0 Copyright (c) 1987, 1988 Borland International
cOs.obj : unable to open file
        Available memory 448686
C:\>tasm monitor.asm
Turbo Assembler Version 4.1 Copyright (c) 1988, 1996 Borland International
Assembling file:
                    monitor.asm
Error messages:
                    None
Warning messages:
                    None
Passes:
Remaining memory:
C:\>tlink /3 /t monitor.obj kernal.obj,kernal.com
Turbo Link Version 2.0 Copyright (c) 1987, 1988 Borland International
```

5) 使用 WinHex 修改软盘

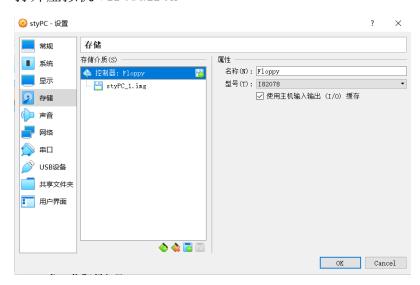
使用 WinHex 打开所有的 COM 文件和 BIN 文件。将引导程序 bootloader. bin 放入首扇区,将 4 个用户程序分别放入 2-5 扇区,将内核 kernal. com 放入第 6 扇区。





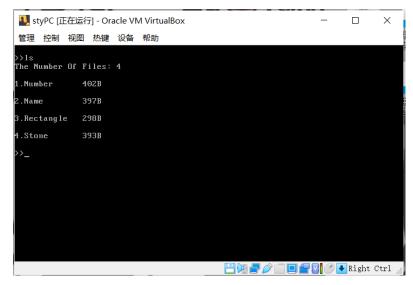
6) 用软盘启动裸机

打开虚拟机 VirtualBox

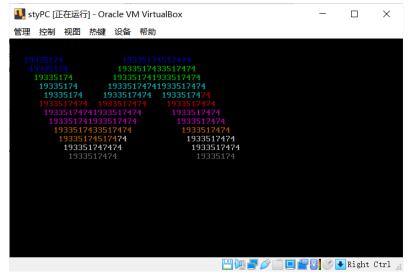


操作系统界面

文件列表



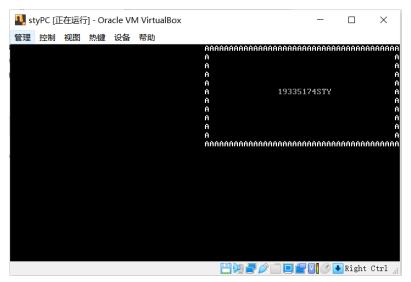
用户程序1



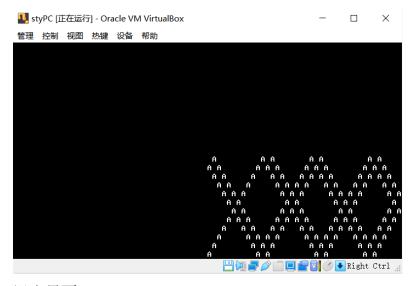
用户程序2



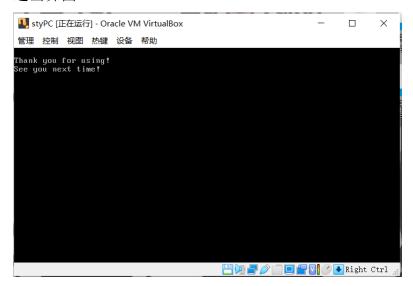
用户程序3



用户程序4



退出界面



Runall 命令是依次执行四个用户程序,截图很难显示动画的效果,我在"运行视频"中有演示。

六、实验总结

这次实验中相比前几次实验难度一下提升了很多,在混合编译链接这块我就花了很多时间才明白。我原本是想用 gcc-nasm 的,但是我总是编译链接出错,于是就换用了 tcc-tasm,但是由于 tcc-tasm 过于古老,网上关于这方面的资料也比较少,所以我还是遇到了不少麻烦。好在经过不懈的努力,我终于还是顺利完成了这次实验。

引导程序的设计相对比较简单,只要跳到内核的扇区即可。但是内核的设计十分艰难,特别是在传递参数这一块。压栈的原理比较复杂,让我一开始疑惑了好久,好在老师有代码参考也让我弄明白了。在实验过程中,我发现一些函数的参数如果用全局变量来设计可能相对简单一些,我也使用了一些全局变量来传递参数。在调用用户程序时我也学到了一个新方法,用 call 代替 jmp,这样在用户程序使用 ret 就能返回。

这次实验我也遇到了一些细节问题。我在用 TLINK 混合编译链接时,一开始把 C 生成的 obj 文件写在了前面,这样就会出错,后来我才发现要把会变生成的 obj 文件写在前面才能正常编译。在使用 tcc 编译时,我一开始用"//"来添加中文注释,但是却报错了,研究了一下发现 tcc 比较古老,只能用"/**/"的方式注释。我在设计 runall 命令时,一开始直接想用连续跳转到四个不同扇区的方法,却发现这样不可能实现——用户程序 1 还能执行,后面的用户程序只会一闪而过。经过调试我发现,在 cmain 中添加一个 tasks 变量判断是否有剩余的任务没完成,并在每次跳转到用户程序后加一个 getChar()停顿,就可以实现用户程序的直接跳转了。

总而言之,这次实验的探索之旅让我学到了很多东西。虽然花费的时间和精力很多,但一切都是值得的。希望在后面的实验中,我也能再接再厉,勇往直前!

七、参考文献

《汇编语言 (第3版)》

《X86 汇编语言: 从实模式到保护模式》

https://blog.csdn.net/wuxch8/article/details/106921308/

https://blog.csdn.net/hua19880705/article/details/8125706

https://blog.csdn.net/longintchar/article/details/79511747

https://blog.csdn.net/mirage1993/article/details/45116881

https://blog.csdn.net/knxw0001/article/details/7248683