МИНОБРНАУКИ РОССИИ САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ «ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА) Кафедра МО ЭВМ

отчет

по лабораторной работе №1 по дисциплине «Операционные системы»

ТЕМА: Исследование структур загрузочных моделей

Студент гр.0382	Злобин А.С.
Преподаватель	Ефремов М.А

2022

Санкт-Петербург

Цель работы.

Исследование различий в структурах исходных текстов модулей типов .COM и .EXE, структур файлов загрузочных моделей и способов их загрузки в основную память.

Задание.

- 1. Написать текст исходного .COM модуля, который определяет тип PC и версию системы. Результатом будет "хороший" .COM модуль, а также "плохой" .EXE, полученный из исходного текста для .COM модуля.
- 2. Написать текст исходного .EXE модуля, который выполняет те же функции, что и модуль в Шаге 1. Результатом будет "хороший" .EXE.
- 3. Сравнить исходные тексты для .COM и .EXE модулей. Ответить на контрольные вопросы "Отличия исходных текстов COM и EXE программ".
- 4. Запустить FAR и открыть файл загрузочного модуля .COM и файл "плохого" .EXE в шестнадцатеричном виде. Затем открыть файл загрузочного модуля "хорошего" .EXE и сравнить его с предыдущими файлами. Ответить на контрольные вопросы "Отличия форматов файлов COM и EXE модулей".
- 5. Открыть отладчик TD.EXE и загрузить .COM. Ответить на контрольные вопросы "Загрузка COM модуля в основную память". Представить в отчете план загрузки .COM модуля в основную память.
- 6. Открыть отладчик TD.EXE и загрузить "хороший" .EXE. Ответить на контрольные вопросы "Загрузка "хорошего" EXE модуля в основную память".
- 7. Оформить отчет в соответствии с требованиями. В отчете необходимо привести скриншоты. Для файлов их вид в шестнадцатеричном виде, для загрузочных модулей в отладчике.

Выполнение работы:

1. При написании .COM модуля был использован шаблон из методический указаний, в котором реализованы процедуры преобразования двоичных кодов в символы шестнадцатеричных и десятичных чисел. Для определения типа РС и версии системы были написаны процедуры: PC_TYPE, OS_VER.

PC_TYPE — процедура для вывода типа ПК, получаемого считыванием предпоследнего байта ROM BIOS (по адресу 0F000:0FFFEh). Далее процедура соотносит численное значение типа известным названиям, а если такого нет, то конвертирует число в строку. Наконец, строковое значение типа выводится с помощью процедуры PRINT.

OS_VER — процедура для вывода версии MS DOS, а также серийных номеров OEM и пользователя. Вся необходимая информация получается при вызове функции 30h прерывания 21h. После вызова функции в AL содержится номер основной версии (если версия раньше 2.0 то AL=0), а в AH — номер модификации. В ВН хранится серийный номер ОЕМ, однако по неизвестным причинам это значение напрямую зависит от значения в AL до вызова функции (если AL = 0, то значение серийного номера будет 255 (FF), если 1 — то 16, иначе 0). В ВL и СХ хранится 3-хбайтовый серийный номер 3 пользователя. Все полученные значения конвертируются в строки и выводятся на экран.

В результате шага имеем "хороший" .COM модуль и "плохой" .EXE модуль. Выводы, полученные при их запуске, представлены на рисунке 1 и рисунке 2 соответственно.

C:\>lb1_com IBM type: AT MS DOS version: 5.0 OEM number:255 User_s number: 000000h

Рисунок 1 – результат запуска модуля lb1_com.com



Рисунок 2 – результат запуска модуля lb1_com.exe

2. Для написания "хорошего" .EXE модуля разобьем программу на сегменты кода, данных и стека, также добавим главную процедуру. В .COM модуле имелась директива org 100h, что нужна, поскольку при загрузке COM модуля в память DOS первые 256 байт блоком данных занимает PSP, код программы располагается лишь после этого блока. В .EXE модуле же мы в этом не нуждаемся, поскольку блок PSP расположен вне сегмента кода. На рисунке 3 представлены результаты запуска данного модуля.



Рисунок 3 – результат запуска модуля lb1_exe.exe

Вопросы.

Отличия исходных текстов СОМ и ЕХЕ программ

- 1) Сколько сегментов должна содержать СОМ-программа? Один сегмент (содержит и код, и данные; стек же генерируется сам).
 - 2) EXE-программа?

Обязательно один — сегмент кода. Также можно добавить сегмент данных и стека, они описываются отдельно друг от друга.

3) Какие директивы должны обязательно быть в тексте СОМ-программы?

Обязательно нужна директива ORG 100h, поскольку первые 256 байт занимает блок данных PSP. Поэтому нужно, чтобы CS:IP имел смещение в 256 байт от нулевого адреса. Так же необходима директива ASSUME для указания сегмента кода и данных на один сегмент.

4) Все ли форматы команд можно использовать в СОМ-программе?

Нет, так как в отличие от EXE, адреса сегментных регистров определяются при запуске программы и отсутствует таблица relocation table, соответственно мы не можем использовать команды с указанием сегментов

Отличия форматов файлов СОМ и ЕХЕ модулей

1) Какова структура файла COM? С какого адреса располагается код?

.COM модуль состоит из одного сегмента, в котором содержится и код, и данные. Код располагается с нулевого адреса, но устанавливается смещение 100h на конец PSP



Рисунок 4 – .СОМ в 16-м виде

2) Какова структура файла "плохого" EXE? С какого адреса располагается код? Что располагается с адреса 0?

Состоит из одного сегмента. Код, стек и данные располагается с адреса 300h (200h занимает заголовок и relocation table, 100h - смещение).

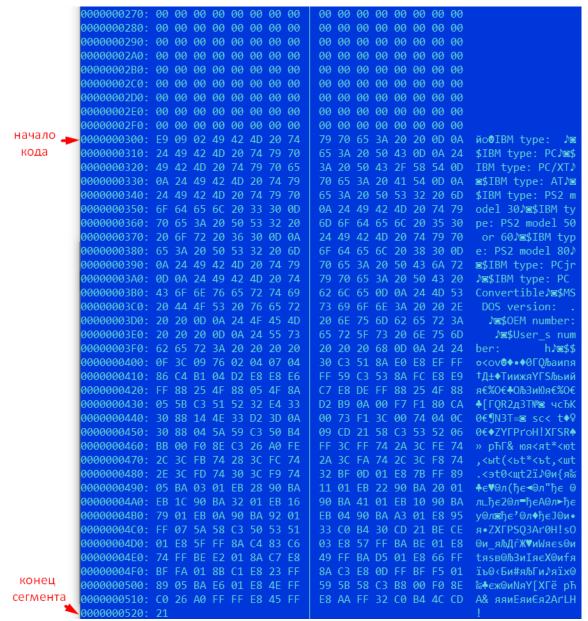


Рисунок 5 – "плохой" .ЕХЕ в 16-м виде

3) Какова структура файла "хорошего" EXE? Чем он отличается от файла "плохого" EXE?

В начале модуля располагается заголовок и relocation table (200h байт), затем сегменты в порядке их определения в коде, т.е. сначала сегмент стека, потом сегмент данных, а затем – сегмент кода.

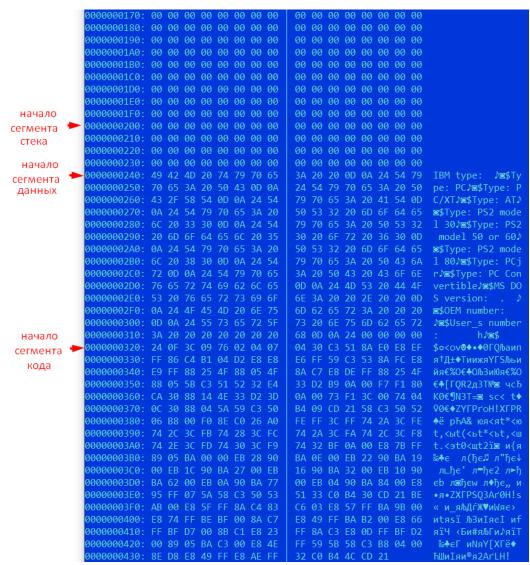


Рисунок 6 – "хороший" .ЕХЕ в 16-м виде

Загрузка СОМ модуля в основную память".

Результат загрузки программы в отладчик представлены на рисунке 7.

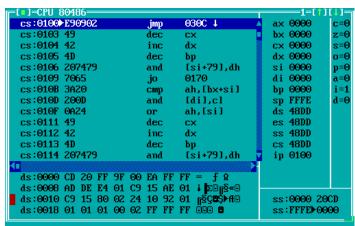


Рисунок 7 – .СОМ модуль

1) Какой формат загрузки модуля СОМ? С какого адреса располагается код?

Сначала определяется сегментный адрес участка ОП, в который помещается модуль. Затем создается блок памяти для PSP и программы. Сегментные регистры устанавливаются на начало PSP, а SP — на конец модуля. Начиная с PSP:0100h помещается считанный образ СОМ-файла. Код же располагается с адреса CS:0100h.

2) Что располагается с адреса 0?

Сегмент PSP.

3) Какие значения имеют сегментные регистры? На какие области памяти они указывают?

Сегментные регистры указывают на PSP.

4) Как определяется стек? Какую область памяти он занимает? Какие адреса?

Стек находится в одном сегменте со всем остальным, и поэтому область памяти под него составляет 64Кб. Адреса от 0h до FFFEh. SP (FFFEh) указывает на конец стека, SS (0h) – на начало.

Загрузка "хорошего" ЕХЕ модуля в основную память.

Результат загрузки программы в отладчик представлены на рисунке 8.

[-[■]=CPU 80486			1=[†]	[]=
cs:010D>B8F148	MOV	ax,48F1 🔺	ax 0000	c=0
cs:0110 8ED8	MOV	ds,ax	b× 0000	z=0
cs:0112 E849FF	call	005E	cx 0000	s=0
cs:0115 E8AEFF	call	0006	d× 0000	0=0
cs:0118 3200	xor	al,al	si 0000	p=0
cs:011A B44C	MOV	ah,4C	di 0000	a=0
cs:011C CD21	int	21	bp 0000	i=1
cs:011E 0000	add	[bx+si],al	sp 0040	d=0
cs:0120 0000	add	[bx+si],al	ds 48DD	
cs:0122 0000	add	[bx+si],al	es 48DD	
cs:0124 0000	add	[bx+si],al	ss 48ED	
cs:0126 0000	add	[bx+si],al	cs 48FF	
cs:0128 0000	add	[bx+si],al 🔽	ip 010D	
4		>		
ds:0000 CD 20 FF 9F 00	EA FF	$\mathbf{F}\mathbf{F} = \mathbf{f} \mathbf{\Omega}$		
ds:0008 AD DE E4 01 C9				
ds:0010 C9 15 80 02 24			ss:0042 204D	
ds:0018 01 01 01 00 02	FF FF	FF 999 8	ss:0040>4249	

Рисунок 8 – загруженный в отладчик "хороший" .СОМ модуль

1) Как загружается "хороший" EXE? Какие значения имеют сегментные регистры?

ЕХЕ-файл загружается, начиная с адреса PSP:0100h. В процессе загрузки считывается информация заголовка (PSP) ЕХЕ в начале файла и выполняется перемещение адресов сегментов, то есть DS и ES устанавливаются на начало сегмента PSP(DS=ES=48DD), SS(SS=48ED) — на начало сегмента стека, CS(CS=490D) — на начало сегмента команд. В IP загружается смещение точки входа в программу, которая берётся из метки после директивы END. Причём дополнительный программный сегмент (PSP) присутствует в каждом EXE-файле.

2) На что указывают регистры DS и ES?

Эти регистры указывают на начало PSP.

3) Как определяется стек?

Стек определяется с помощью директивы .stack, после которой задаётся размер стека. При исполнение регистр SS указывает на начало сегмента стека, а SP на конца стека(его смещение)

4) Как определяется точка входа?

Точка входа определяется директивой END. Данная метка определяет адрес, с которого начинается выполнение программы, т.е. — точка входа в программу.

Исходный код программы см. в приложении А.

Выводы.

В ходе работы были исследованы различия в структурах исходных текстов модулей типов .COM и .EXE, в структурах файлов загрузочных моделей и способов их загрузки в основную память; была написана программа, выводящую информацию о типе PC и версии системы.

ПРИЛОЖЕНИЕ А

ИСХОДНЫЙ КОД ПРОГРАММЫ

Название файла: lab1_com.asm

```
TESTPC SEGMENT
  ASSUME CS:TESTPC, DS:TESTPC, ES:NOTHING, SS:NOTHING
  ORG 100H
START: JMP BEGIN
; Данные
TYPE PC db 'Type: PC', ODH, OAH, '$'
TYPE PC XT db 'Type: PC/XT', ODH, OAH, '$'
TYPE AT db 'Type: AT', ODH, OAH, '$'
TYPE PS2 M30 db 'Type: PS2 модель 30',0DH,0AH,'$'
TYPE PS2 M50 60 db 'Type: PS2 модель 50 или 60', 0DH, 0AH, '$'
TYPE PS2 M80 db 'Type: PS2 модель 80', 0DH, 0AH, '$'
TYPE PC JR db 'Type: PCjr', ODH, OAH, '$'
TYPE PC CONV db 'Type: PC Convertible', ODH, OAH, '$'
VERSIONS db 'Version MS-DOS: . ', ODH, OAH, '$'
SERIAL NUMBER db 'Serial number OEM: ',ODH,OAH,'$'
USER NUMBER db 'User serial number:
; Процедуры
;-----
TETR TO HEX PROC near
  and AL, OFh
  cmp AL,09
  jbe next
  add AL,07
next:
  add AL, 30h
  ret
TETR TO HEX ENDP
;-----
BYTE TO HEX PROC near
;байт в AL переводится в два символа шест. числа в АХ
  push CX
  mov AH, AL
  call TETR TO HEX
  xchg AL, AH
  mov CL, 4
  shr AL, CL
  call TETR TO HEX ;в AL старшая цифра
  рор СХ ; в АН младшая
  ret
BYTE TO HEX ENDP
;-----
WRD TO HEX PROC near
;перевод в 16 с/с 16-ти разрядного числа
; в АХ - число, DI - адрес последнего символа
  push BX
```

```
mov BH, AH
  call BYTE TO HEX
  mov [DI], AH
  dec DI
  mov [DI], AL
  dec DI
  mov AL, BH
  call BYTE TO HEX
  mov [DI], AH
  dec DI
  mov [DI], AL
  pop BX
  ret
WRD TO HEX ENDP
;------
BYTE TO DEC PROC near
; перевод в 10c/c, SI - адрес поля младшей цифры
  push CX
  push DX
  xor AH, AH
  xor DX, DX
  mov CX, 10
loop bd:
  div CX
  or DL, 30h
  mov [SI], DL
  dec SI
  xor DX, DX
  cmp AX, 10
  jae loop bd
  cmp AL,00h
  je end l
  or AL, 30h
  mov [SI], AL
end 1:
  pop DX
  pop CX
  ret
BYTE TO DEC ENDP
;-----
WRITESTRING PROC near
  mov AH,09h
  int 21h
  ret
WRITESTRING ENDP
PC TYPE PROC near
  mov ax, 0f000h; получаем номер модели
  mov es, ax
  mov al, es:[Offfeh]
  cmp al, Offh; начинаем стравнивать
  je pc
```

```
cmp al, Ofeh
   je pc xt
   cmp al, 0fbh
   je pc xt
   cmp al, Ofch
   je pc at
   cmp al, Ofah
   je pc ps2 m30
   cmp al, 0f8h
   je pc ps2 m80
   cmp al, 0fdh
   je pc jr
   cmp al, 0f9h
   je pc conv
pc:
      mov dx, offset TYPE PC
      jmp writetype
pc xt:
      mov dx, offset TYPE PC XT
      jmp writetype
pc at:
      mov dx, offset TYPE AT
      jmp writetype
pc ps2 m30:
      mov dx, offset TYPE PS2 M30
      jmp writetype
pc ps2 m50 60:
      mov dx, offset TYPE PS2 M50 60
      jmp writetype
pc_ps2 m80:
      mov dx, offset TYPE_PS2_M80
      jmp writetype
pc jr:
      mov dx, offset TYPE PC JR
      jmp writetype
pc conv:
      mov dx, offset TYPE PC CONV
      jmp writetype
writetype:
      call WRITESTRING
   ret
PC TYPE ENDP
OS VER PROC near
   mov ah, 30h
   int 21h
   push ax
   mov si, offset VERSIONS
   add si, 16
   call BYTE TO DEC
   pop ax
   mov al, ah
```

```
mov si, offset SERIAL NUMBER
        add si, 19
        mov al, bh
        call BYTE TO DEC
        mov dx, offset SERIAL NUMBER
        call WRITESTRING
        mov di, offset USER NUMBER
        add di, 25
        mov ax, cx
        call WRD TO HEX
        mov al, bl
        call BYTE TO HEX
        sub di, 2
        mov [di], ax
        mov dx, offset USER NUMBER
        call WRITESTRING
        ret
     OS VER ENDP
     ; Код
     BEGIN:
        call PC TYPE
        call OS VER
        xor AL, AL
        mov AH, 4Ch
        int 21H
     TESTPC ENDS
     END START
     Название файла: lab1_exe.asm
          SEGMENT STACK
AStack
          DW 128 DUP(?)
AStack
         ENDS
DATA SEGMENT
   TYPE PC db 'Type: PC', ODH, OAH, '$'
   TYPE PC XT db 'Type: PC/XT', ODH, OAH, '$'
   TYPE AT db 'Type: AT', ODH, OAH, '$'
   TYPE PS2 M30 db 'Type: PS2 модель 30', ODH, OAH, '$'
   TYPE PS2 M50 60 db 'Type: PS2 модель 50 или 60', ODH, OAH, '$'
   TYPE PS2 M80 db 'Type: PS2 модель 80',0DH,0AH,'$'
   TYPE PC JR db 'Type: PCjr', ODH, OAH, '$'
   TYPE PC CONV db 'Type: PC Convertible', ODH, OAH, '$'
   VERSIONS db 'Version MS-DOS: . ', ODH, OAH, '$'
```

add si, 3

call BYTE TO DEC

call WRITESTRING

mov dx, offset VERSIONS

```
SERIAL NUMBER db 'Serial number OEM: ',ODH,OAH,'$'
  USER NUMBER db 'User serial number:
                                       н $'
DATA ENDS
CODE SEGMENT
  ASSUME CS:CODE, DS:DATA, SS:AStack
  ; Процедуры
;-----
TETR TO HEX PROC near
  and AL, OFh
  cmp AL,09
  jbe next
  add AL,07
next:
  add AL, 30h
  ret
TETR TO HEX ENDP
;-----
BYTE TO HEX PROC near
;байт в AL переводится в два символа шест. числа в АХ
  push CX
  mov AH, AL
  call TETR TO HEX
  xchg AL, AH
  mov CL, 4
  shr AL, CL
  call TETR TO HEX ;в AL старшая цифра
  рор СХ ;в АН младшая
  ret
BYTE TO HEX ENDP
;-----
WRD TO HEX PROC near
;перевод в 16 с/с 16-ти разрядного числа
; в АХ - число, DI - адрес последнего символа
  push BX
  mov BH, AH
  call BYTE TO HEX
  mov [DI], AH
  dec DI
  mov [DI], AL
  dec DI
  mov AL, BH
  call BYTE TO HEX
  mov [DI], AH
  dec DI
  mov [DI], AL
  pop BX
  ret
WRD TO HEX ENDP
;-----
BYTE TO DEC PROC near
; перевод в 10c/c, SI - адрес поля младшей цифры
  push CX
```

```
push DX
  xor AH, AH
  xor DX, DX
  mov CX, 10
loop bd:
  div CX
   or DL, 30h
  mov [SI], DL
  dec SI
  xor DX, DX
   cmp AX, 10
   jae loop bd
   cmp AL,00h
   je end l
   or AL, 30h
  mov [SI], AL
end 1:
  pop DX
  pop CX
  ret
BYTE TO DEC ENDP
;-----
WRITESTRING PROC near
  mov AH,09h
   int 21h
   ret
WRITESTRING ENDP
PC TYPE PROC near
  mov ax, 0f000h; получаем номер модели
  mov es, ax
  mov al, es:[0fffeh]
   cmp al, Offh; начинаем стравнивать
   je pc
   cmp al, Ofeh
   je pc xt
   cmp al, 0fbh
   je pc xt
   cmp al, Ofch
   je pc at
   cmp al, Ofah
   je pc ps2 m30
   cmp al, 0f8h
   je pc ps2 m80
   cmp al, 0fdh
   je pc_jr
   cmp al, 0f9h
   je pc conv
pc:
   mov dx, offset TYPE PC
   jmp writetype
pc xt:
```

```
mov dx, offset TYPE PC XT
   jmp writetype
pc at:
   mov dx, offset TYPE AT
   jmp writetype
pc ps2 m30:
   mov dx, offset TYPE PS2 M30
   jmp writetype
pc ps2 m50 60:
   mov dx, offset TYPE PS2 M50 60
   jmp writetype
pc ps2 m80:
   mov dx, offset TYPE PS2 M80
   jmp writetype
pc jr:
   mov dx, offset TYPE PC JR
   jmp writetype
pc conv:
   mov dx, offset TYPE PC CONV
   jmp writetype
writetype:
   call WRITESTRING
   ret
PC TYPE ENDP
OS VER PROC near
   mov ah, 30h
   int 21h
   push ax
   mov si, offset VERSIONS
   add si, 16
   call BYTE TO DEC
   pop ax
   mov al, ah
   add si, 3
   call BYTE TO DEC
   mov dx, offset VERSIONS
   call WRITESTRING
   mov si, offset SERIAL NUMBER
   add si, 19
   mov al, bh
   call BYTE TO DEC
   mov dx, offset SERIAL NUMBER
   call WRITESTRING
   mov di, offset USER NUMBER
   add di, 25
   mov ax, cx
   call WRD TO HEX
   mov al, bl
```

```
call BYTE TO HEX
   sub di, 2
   mov [di], ax
   mov dx, offset USER NUMBER
   call WRITESTRING
   ret
OS_VER ENDP
Main PROC FAR
   sub AX, AX
  push AX
  mov AX, DATA
mov DS, AX
  call PC_TYPE
  call OS VER
  xor AL, AL
  mov AH, 4Ch
   int 21H
   ;ret
Main ENDP
CODE ENDS
   END Main
```