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

ОТЧЕТ

по лабораторной работе №1

по дисциплине «Операционные системы»

Тема: Исследование структур загрузочных модулей

| Студент гр. 9382 | Субботин М.О. |
|------------------|-------------------|
| Преподаватель | Ефремов М.А. |

Санкт-Петербург

2021

Цель работы.

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

Основные теоретические положения.

«Истина познается в сравнении», как говорили древние. К счастью, у нас есть возможность исследовать в одной системе два различных формата загрузочных модулей, сравнить их и лучше понять как система программирования и управляющая программа обращаются с ними. Система программирования включает компилятор с языка ассемблер (часто называется, просто, ассемблер), который изготавливает объектные модули. Компоновщик (Linker) по совокупности объектных модулей, изготавливает загрузочный модуль, а также, функция ядра – загрузчик, которая помещает программу в основную память и запускает на выполнение. Все эти компоненты согласованно работают для изготовления и выполнения загрузочных модулей разного типа. Для выполнения лабораторной работы сначала нужно изготовить загрузочные модули. Ассемблерная программа должна читать содержимое предпоследнего байта ROM BIOS, по таблице, сравнивая коды, определять тип PC и выводить строку с названием модели. Если код не совпадает ни с одним значением, то двоичный код переводиться в символьную строку, содержащую запись шестнадцатеричного числа и выводиться на экран в виде соответствующего сообщения.

Затем определяется версия системы. Ассемблерная программа должна по значениям регистров AL и AH формировать текстовую строку в формате хх.уу, где хх - номер основной версии, а уу - номер модификации в десятичной системе счисления, формировать строки с серийным номером ОЕМ и серийным номером пользователя. Полученные строки выводятся на экран.

Ход выполнения:

Были созданы методы TYPE_IBM_PC и OS_VERSION. Первый метод выводит на экран информацию о типе ПК, вторая о системе. В зависимости от предпоследнего байта ROM BIOS, первый метод выводит соответствующий байту тип модели. На рисунке ниже представлены соответствия тип – байт.

| PC | FF |
|----------------------|--------|
| PC/XT | FE, FB |
| AT | FC |
| PS2 модель 30 | FA |
| PS2 модель 50 или 60 | FC |
| PS2 модель 80 | F8 |
| PCjr | FD |
| PC Convertible | F9 |

Рисунок 1. Соответствие предпоследнего байта ROM BIOS и типа ПК. Метод OS_VERSION выводит следующую информацию: номер основной версии, номер модификации в десятичной системе счисления, серийный номер ОЕМ, серийный номер пользователя.

Были получены следующие результаты (рисунки 2-4):

```
D:\LABOS\LAB1>lab1.com
Type: AT
MS DOS Version: 5.0
OEM: 0
Serial number: 000000
```

Рисунок 2. Результат выполнения .СОМ файла.



Рисунок 3. Результат выполнения "плохого". ЕХЕ файла.

D:\LABOS\LAB1>lab1_exe.exe

Type: AT MS DOS Version: 5.0

Serial number: 000000

Рисунок 4. Результат выполнения "хорошего" .ЕХЕ файла.

Выводы.

Были исследованы различия в структурах исходных текстов модулей типов .СОМ и .ЕХЕ, структур файлов загрузочных модулей и способов их загрузки в основную память.

ПРИЛОЖЕНИЕ А ОТВЕТЫ НА КОНТРОЛЬНЫЕ ВОПРОСЫ

Отличия исходных текстов СОМ и ЕХЕ программ

1. Сколько сегментов должна содержать СОМ-программа?

СОМ-программа должна содержать один сегмент, в этом одном сегменте хранятся код и данные, стек же генерируется автоматически.

2. ЕХЕ-программа?

EXE-программа может содержать ≥1 сегментов. EXE-программа может содержать произвольный объем основной памяти, поэтому кол-во сегментов ограничено размером основной памяти.

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

ASSUME CS:TESTPC, DS:TESTPC, ES:NOTHING, SS:NOTHING и

ORG 100H. Первая директива указывает на то, что сегменты данных и кода находятся в одном сегменте. Вторая директива нужна для смещения адресов на 256 байт, т.к. первые 256 байт указывают на PSP.

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

Нельзя использовать команды типа: "MOV <register>, SEG <segment name>", т.к. в СОМ-программе нет таблицы настроек.

Отличия форматов файлов СОМ и ЕХЕ модулей

1. Какова структура файла СОМ? С какого адреса располагается код?

СОМ-программа состоит из одного сегмента (сегмент кода и сегмент данных), сегмент стека получается автоматически. СОМ-программа не превышает 64 Кб.

Как будет видно, код начинается с адреса 0h, но из-за директивы org 100h происходит смещение на 100h.

```
79 70 65 3A 20 50 43 0D 0A 24 54 79
              B4 01 54
                                                               ⊕ Type: PC..$Ty
00000010
                                                               pe: PC/XT..$Type
              65
                    20 50
                          43
                                 58
                                   54
                                       0D 0A
                                             24
                                                    79
                                                       70 65
                              2F
                                                 54
00000020
           3A 20 41
                    54 0D 0A
                              24 54 79
                                       70 65 3A
                                                   50 53 32
                                                               : AT...$Type: PS2
00000030
                                       0D 0A 24 54 79 70 65
                                                                model 30..$Type
              6D 6F
                    64 65 6C
                             20 33 30
00000040
           3A 20
                 50 53 32 20
                             6D 6F 64
                                       65 6C 20 38 30 0D 0A
                                                               : PS2 model 80..
00000050
           24 54
                    70 65
                          ЗА
                              20 50
                                       6A
                                          72
                                             0D
                                                 0A
                                                    24
                                                       54
                                                          79
                                                               $Type: PSjr..$Ty
00000060
                    20 50 43
                              20 43 6F
                                       6E
                                             65
                                                       69 62
                                                               pe: PC Convertib
00000070
           6C 65
                    0A 24 4D 53 20
                                       4F
                                              20
                                                       72 73
                                                               le..$MS DOS Vers
                                    44
                                          53
                                                   65
00000080
                    3A 20
                          20
                             20 2E 20
                                       20 0D 0A
                                                                          ..$0EM
00000090
           3A 20
                    20 OD
                          0A
                              24 53 65
                                       72
                                          69 61
                                                 6C
                                                    20
                                                       6E 75
                                                                   ..$Serial nu
000000A0
                           20
                             20 20 20
                                       20 20
                                                 20 20
                                                       20 20
                    72 3A
                                             20
                                                               mber:
000000B0
           20 20 20 20 20 20 20 20 20
                                       20 20 20 20 24
                                                       24 0F
00000C0
           3C 09
                 76 02 04
                          07 04 30 C3 51 8A E0 E8
                                                   EF
                                                       FF 86
                                                               <.v...0 | QèαΦ∩ å
00000D0
           C4 B1 04 D2 E8 E8 E6 FF 59 C3 53 8A FC E8 E9 FF
                                                               —∭.πΦΦμ Υ├SèʰΦΘ
000000E0
           88 25 4F
                    88 05 4F
                             8A C7 E8 DE FF
                                             88 25 4F 88 05
                                                               00000F0
           5B C3 51 52 32 E4 33 D2 B9 0A 00 F7
                                                 F1 80 CA 30
                                                               [ -QR2∑3π4 . . ≈±Ç-0
00000100
           88 14 4E 33 D2 3D
                             0A 00 73 F1 3C 00
                                                74
                                                    04 0C 30
                                                               \hat{e}.N3\pi = ...s \pm < .t..0
00000110
           88 04 5A 59 C3 B8 00 F0 8E
                                                               ê.ZY├q.≡Ä└&á• <
                                       C0 26 A0 FE
                                                       3C FF
00000120
           74 1C 3C FE 74 1E
                             3C FB 74
                                       1A 3C FC
                                                74 1C 3C FA
                                                               t.<•t.<√t.<nt.<.
00000130
           74 1E 3C F8 74 20 3C FD 74 22 3C F9 74 24 BA 03
                                                               t.<°t <²t"<·t$∥.
00000140
           01 EB 22 90 BA 0E
                             01 EB 1C
                                       90
                                          BA 1C 01 EB 16 90
                                                               .δ"έ∥..δ.έ∥..δ.έ
                    EB 10 90 BA 3C 01 EB 0A 90 BA 51 01 EB
00000150
           BA 27 01
                                                               ∥'.δ.É∥<.δ.É∥Q.δ
00000160
           04 90 BA 5E 01 B4 09 CD 21 C3 B4 30 CD 21 50 BE
                                                               .É∥^.┤.=! ├ 0=! Þ╛
00000170
           75 01 83 C6 11 E8 7A FF 58 8A C4 83 C6 03 E8 71
                                                               u.â ⊨.Фz Xè—â ⊨.Фq
00000180
           FF BA 75 01 B4 09 CD 21 BE 8D 01 83 C6 05 8A C7
                                                                ∥u.-|.=!∃ì.â⊧.è∥
                                             97 01 83 C7 14
                                                                  ||ì.-|.=!¬ù.â|-.
00000190
                       8D 01 B4 09 CD 21
                                          BF
                    BA
000001A0
           8B C1 E8
                    35 FF 83 EF 02 8A C3 E8 1C FF 89 05 BA
                                                               ï<sup>⊥</sup>Φ5 â∩.è Φ. ë.
000001B0
           97 01 B4 09 CD 21 C3 E8 5B FF E8 AD FF 32 C0 B4
                                                               ù.┥.=! ┡Ф[ Ф; 2 ┗
000001C0
           4C CD 21 +
                                                               L=!
```

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

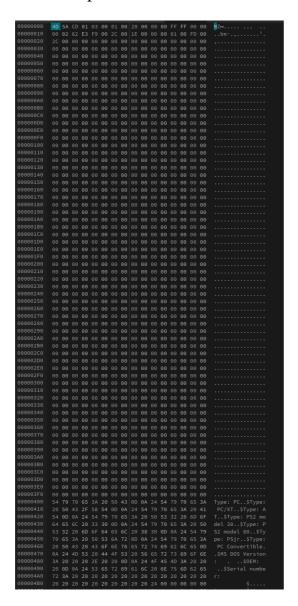
В "плохом" EXE-файле сегменты данных и кода расположены в одном сегменте, что не очень хорошо, т.к. в EXE-файле сегменты кода и данных должны быть разделены. Код начинается с адреса 300h(PSP -100h, заголовок— 100h, org 100h), с адреса 0h располагается заголовок, а конкретно сигнатура MZ формата 2 байта (4D 5A).

```
: PS2 model 80.
```

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

В "хорошем" ЕХЕ-файле сегменты кода, данных и стека должны быть раздельными. ЕХЕ-файл имеет заголовок, в нем находится сигнатура и данные, необходимые для загрузки ЕХЕ-файла, а также там находится таблица настройки адресов. ЕХЕ-файл может иметь любой размер.

В "плохом" EXE адресация начинается с 300h. В "хорошем" EXE 400h, т.к. теперь нет смещения в 100h (org 100h), а вместо этого отведено место под стек равное 200h.

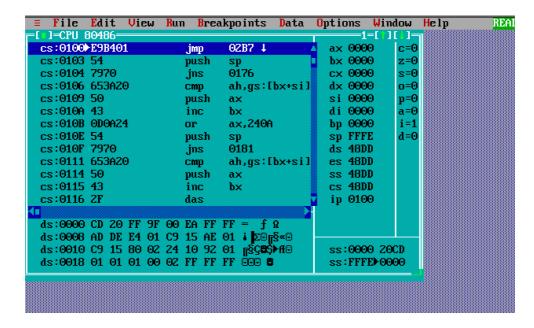


Загрузка СОМ модуля в основную память

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

Сначала определяется место в ОП, в которое можно загрузить программу, СОМ-файл считается и помещается в память. Код начинается со смещением в 100h.

Как видно в IP лежит 0100h.



2. Что располагается с адреса 0?

С адреса 0h по 100h находится PSP.

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

Сегментные регистры CS,ES,SS,CS равны 48DD(указывают на PSP), а SP равен FFFE и указывает на конец сегмента PSP.

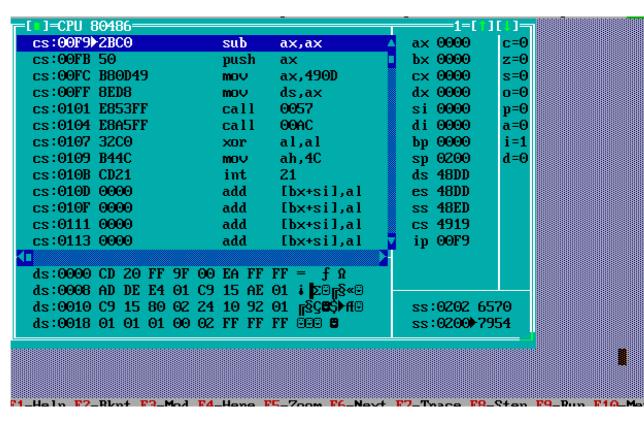
4. Как определяется стек? Какую область памяти он занимает? Какие адреса?

Стек получается автоматически. SS указывает на начало 0h, SP – на конец FFFEh, т.е. стек расположен с 0h по FFFEh.

Загрузка "хорошего" ЕХЕ модуля в основную память

1. Как загружается "хороший" ЕХЕ? Какие значения имеют сегментные регистры?

EXE-файл загружается начиная с адреса 0100h, считывается информация заголовка PSP и выполняется перемещение адресов сегментов. DS, ES равны 48DD (указывают на начало PSP), SS равен 48ED (указывает на начало сегмента стека), CS равен 4919 (указывает на начало сегмента команд).



2. На что указывают регистры DS и ES?

DS и ES указывают на начало PSP.

3. Как определяется стек?

Стек определяется следующим образом:

AStack SEGMENT STACK

DW 256 DUP(?)

AStack ENDS

Где 256 — размер стека. Регистр SS указывает на начало сегмента стека, SP — на конец.

4. Как определяется точка входа?

Точка входа определяется при помощи директивы END. Эта директива обозначает конец программы, но также прописав END <entry point> можно указать точку входа в программу.

приложение б

КОД ПРОГРАММ

Lab1.asm:

TESTPC SEGMENT ASSUME CS:TESTPC, DS:TESTPC, ES:NOTHING, SS:NOTHING 100H **ORG START: JMP BEGIN** ;ДАННЫЕ PC STR db 'Type: PC',0DH,0AH,'\$' 'Type: PC/XT',0DH,0AH,'\$' PC XT STR db 'Type: AT',0DH,0AH,'\$' AT STR db 'Type: PS2 model 30',0DH,0AH,'\$' PS2 30 STR db 'Type: PS2 model 80',0DH,0AH,'\$' PS2 80 STR db PSjr_STR db 'Type: PSir',0DH,0AH,'\$' PC Conv STR db 'Type: PC Convertible',0DH,0AH,'\$' 'MS DOS Version: . ',0DH,0AH,'\$' MS DOS VERSION db **OEM** 'OEM: ',0DH,0AH,'\$' db \$' **NUMBER** db 'Serial number: ;ПРОЦЕДУРЫ TETR TO HEX PROC near and AL,0Fh AL,09 cmp **NEXT** ibe add AL,07 NEXT: add AL,30h ret TETR TO HEX **ENDP** <u>:-----</u> BYTE TO HEX **PROC** near ;байт в AL переводится в два символа шестн. числа в AX push CX mov AH,AL TETR TO HEX call xchg AL,AH mov CL,4 shr AL,CL TETR TO HEX; в AL старшая цифра call

; в АН младшая

CX

pop

```
ret
BYTE TO HEX
                     ENDP
<u>:-----</u>
WRD TO HEX
                     PROC
                                    near
;перевод в 16 сс 16-ти разрядного числа
;в АХ - число, DI - адрес последнего символа
               push BX
               mov BH,AH
                     BYTE TO HEX
               call
               mov
                          [DI],AH
               dec
                          DI
                          [DI],AL
               mov
               dec
                          DI
                          AL,BH
               mov
                     BYTE TO HEX
               call
                          [DI],AH
               mov
               dec
                          DI
                          [DI],AL
               mov
                          BX
               pop
               ret
WRD TO HEX ENDP
                     PROC
BYTE TO DEC
;перевод в 10cc, SI - адрес поля младшей цифры
               push CX
               push DX
               xor
                          AH,AH
                          DX,DX
               xor
                          CX,10
               mov
loop_bd:
                     CX
          div
                    DL,30h
               or
               mov
                          [SI],DL
                     SI
               dec
                     DX,DX
               xor
                    AX,10
               cmp
                          loop bd
               jae
               cmp
                          AL,00h
               je
                          end 1
                          AL,30h
               or
               mov
                          [SI],AL
```

DX

ENDP

CX

end 1:

BYTE_TO_DEC

pop

pop ret

| ; | | |
|-------------|------|----------------------|
| TYPE_IBM_PC | PROC | near near |
| | mov | ax,0F000h |
| | mov | • |
| | mov | al,es:[0FFFEh] |
| | cmp | al,0FFh |
| | je | PC |
| | | 1 0001 |
| | cmp | |
| | je | PC_XT |
| | cmp | |
| | je | PC_XT |
| | cmp | al,0FCh |
| | je | AT |
| | cmn | al,0FAh |
| | je | PS2 30 |
| | Je | 1 5 2 _50 |
| | cmp | al,0F8h |
| | je | PS2_80 |
| | cmp | al,0FDh |
| | je | PSjr |
| | | • |
| | | al,0F9h |
| | je | PC_Conv |
| PC: | | |
| | mov | dx,offset PC_STR |
| | jmp | write_type |
| PC_XT: | | 1 CC 4 DC VT CTD |
| | | dx,offset PC_XT_STR |
| AT: | Jmp | write_type |
| 111. | mov | dx,offset AT STR |
| | | write type |
| PS2_30: | J 1 | _ 71 |
| | mov | dx,offset PS2_30_STR |
| | jmp | write_type |
| PS2_80: | | 4 00 |
| | | dx,offset PS2_80_STR |
| | jmp | write_type |
| | | |

PSjr: mov dx,offset PSjr STR jmp write type PC Conv: mov dx,offset PC Conv STR write type: mov AH,09h 21h int ret TYPE IBM PC **ENDP** OS VERSION **PROC** near ah,30h mov 21h int push ax mov si, offset MS DOS VERSION si,17 ;подвинемся на место первой цифры версии в выводе BYTE TO DEC; переводим число al в десятичное и call сохраняем в память куда указывает si pop ax mov al, ah si, 3; чуть подвинем указатель add BYTE TO DEC call mov dx, offset MS DOS VERSION mov AH,09h int 21h mov si, offset OEM add si, 5 mov al, bh BYTE TO DEC call mov dx, offset OEM mov AH,09h int 21h mov di, offset NUMBER

add

di, 20

mov ax, cx

call WRD TO HEX; уже занесли число в память.

sub di,2

mov al, bl

call BYTE_TO_HEX

mov [di], ax

mov dx, offset NUMBER

mov AH,09h int 21h

ret

OS VERSION ENDP

; КОД BEGIN:

call TYPE_IBM_PC call OS VERSION

;Выход в DOS

xor AL,AL

mov AH,4Ch

int 21H

TESTPC ENDS

END START ;конец модуля, START - точка входа

Lab1 exe.asm:

AStack SEGMENT STACK

DW 256 DUP(?)

AStack ENDS

DATA SEGMENT

PC STR db 'Type: PC',0DH,0AH,'\$'

PC XT STR db 'Type: PC/XT',0DH,0AH,'\$'

AT STR db 'Type: AT',0DH,0AH,'\$'

PS2_30_STR db 'Type: PS2 model 30',0DH,0AH,'\$'

PS2_80_STR db 'Type: PS2 model 80',0DH,0AH,'\$'

PSjr STR db 'Type: PSjr',0DH,0AH,'\$'

PC Conv STR db 'Type: PC Convertible',0DH,0AH,'\$'

MS_DOS_VERSION db 'MS DOS Version: . ',0DH,0AH,'\$'

OEM db 'OEM: ',0DH,0AH,'\$'

NUMBER db 'Serial number: \$'

DATA ENDS

CODE SEGMENT

ASSUME CS:CODE,DS:DATA,SS:AStack

;ПРОЦЕДУРЫ

;-----

TETR_TO_HEX PROC near

and AL,0Fh

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 шестн. числа в AX

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 сс 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 10сс, 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 1

or AL,30h

mov [SI],AL

end_l: pop DX

pop CX

ret

BYTE_TO_DEC ENDP

;-----

TYPE_IBM_PC PROC near

mov ax,0F000h

mov es,ax

mov al,es:[0FFFEh]

cmp al,0FFh

je PC

cmp al,0FEh

je PC_XT

cmp al,0FBh

je PC_XT

cmp al,0FCh

je AT

cmp al,0FAh

je PS2_30

cmp al,0F8h

je PS2_80

cmp al,0FDh

je PSjr

cmp al,0F9h

je PC_Conv

PC:

mov dx,offset PC_STR

jmp write_type

PC_XT:

mov dx,offset PC_XT_STR

jmp write_type

AT:

mov dx,offset AT_STR

jmp write_type

PS2_30:

mov dx,offset PS2_30_STR

jmp write_type

PS2_80:

mov dx,offset PS2_80_STR

jmp write type PSjr: mov dx,offset PSjr STR jmp write type PC_Conv: mov dx,offset PC_Conv_STR write type: mov AH,09h 21h int ret TYPE IBM PC **ENDP** OS VERSION PROC near ah,30h mov 21h int push ax mov si, offset MS_DOS_VERSION si,17 ;подвинемся на место первой цифры версии в add выводе BYTE_TO_DEC; переводим число al в десятичное и call сохраняем в память куда указывает si pop ax

mov al, ah

add si, 3; чуть подвинем указатель

call BYTE_TO_DEC

mov dx, offset MS_DOS_VERSION

mov AH,09h

int 21h

mov si, offset OEM

add si, 5

mov al, bh

call BYTE_TO_DEC

mov dx, offset OEM

mov AH,09h

int 21h

mov di, offset NUMBER

add di, 20

mov ax, cx

call WRD_TO_HEX; уже занесли число в память.

sub di,2

mov al, bl

call BYTE_TO_HEX

mov [di], ax

mov dx, offset NUMBER

mov AH,09h

int 21h

ret

OS_VERSION ENDP

; КОД

Main PROC FAR

sub AX,AX

push AX

mov AX,DATA

mov DS,AX

call TYPE_IBM_PC

call OS_VERSION

;Выход в DOS

xor AL,AL

mov AH,4Ch

int 21H

Main ENDP

CODE ENDS

END Main