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 модулей

Студентка гр. 0381	Сарычева А.А
Преподаватель	Ефремов М.А.

Санкт-Петербург

2022

Цель работы.

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

Задание.

Напишите текст исходного .СОМ модуля, который определяет тип РС и версию системы. Ассемблерная программа должна читать содержимое предпоследнего байта ROM BIOS, по таблице, сравнивая коды, определять тип РС и выводить строку с названием модели. Если код не совпадает ни с одним значением, то двоичный код переводиться в символьную строку, содержащую шестнадцатеричного числа запись И выводиться на экран виде соответствующего сообщения. Затем версия определяется 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.

Таблица 1 – Необходимые данные: соотношение типа IBM PS с кодом

таолица 1— необходимые данные, соотношение типа тым г 5 с кодом		
Тип IBM PS	Код	
PS	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	

Выполнение работы.

В файле lb1_com.asm написан код исходного .COM модуля. В начале данного модуля прописаны строки для вывода запрашиваемой информации, для

вывода которых была создана процедура OUTPUT. Кроме того, были использованы процедуры TETR_TO_HEX, BYTE_TO_HEX, WRD_TO_HEX, BYTE_TO_DEC из предоставленного шаблона.

Написана процедура ТҮРЕ_РС, которая определяет тип РС. В регистр AL сохраняется значение байта, который хранит в себе код типа РС. Данный код используется для сравнения с кодами, указанными в табл.1, при наличии совпадения с одним из указанных кодов происходит переход по соответствующей метке, где происходит запись в регистр DX смещение необходимой строки сообщения, после чего происходит вызов процедуры печати сообщения — OUTPUT.

Далее написана процедура SYSTEM_VERSION, которая определяет версию MS-DOS, серийный номер ОЕМ и серийный номер пользователя. С помощью функции 30H прерывания 21H определяются необходимые данные, которые выводятся в соответствующем формате с помощью процедуры OUTPUT.

Командой MASM lb1_com.asm был получен объектный файл lb1_com.obj, из которого командой LINK lb1_com.obj собирается «плохой» .EXE-модуль. При его запуске выводятся строки, представленные на рис.1.



Рисунок 1 – вывод модуля lb1 com.exe

Командой EXE2BIN lb1_com.exe lb1_com.com был получен «хороший» .COM-модуль. При его запуске выводятся строки, представленные на рис.2.

```
F:\>lb1_com.com
Тип PC: AT
Версия MS-DOS: 5.0
Серийный номер ОЕМ: 00Н
Серийный номер пользователя: 000000Н
```

Рисунок 2 – вывод модуля lb1 com.com

Далее в файле lb1_exe.asm был написан код «хорошего» .EXE-модуля, на основе кода из файла lb1_com.asm, в который были добавлены определения сегментов данных и стека, необходимые данные были перенесены в соответсвующие им сегмент, кроме того, была создана далекая головная процедура MAIN, в которо инициализируется регистр DS адресом начала сегмента данных и вызываются процедуры TYPE_PC и SYSTEM_VERSION.

Командой MASM lb1_exe.asm был получен объектный файл lb1_exe.obj, из которого командой LINK lb1_exe.obj собирается «хороший» .EXE-модуль. При его запуске выводятся строки, представленные на рис.3.

```
F:\>lb1_exe.exe
Тип РС: AT
Версия MS-DOS: 5.0
Серийный номер ОЕМ: 00Н
Серийный номер пользователя: 000000Н
```

Рисунок 3 – вывод модуля lb1_exe.com

Выводы.

В ходе лабораторной работы были исследованы различия в структурах исходных текстов модулей типов .COM и .EXE, структуры файлов загрузочных модулей и способы их загрузки в основную память.

Ответы на контрольные вопросы.

Отличия исходных текстов СОМ и ЕХЕ программ

1) Сколько сегментов должна содержать СОМ-программа?

СОМ-программа должна содержать 1 сегмент, являющийся сегментом кода. В нем же определяются данные, а стек генерируется автоматически, в следствие чего он опускается в СОМ-программе.

2) Сколько сегментов должна содержать ЕХЕ-программа?

EXE-программа должна содержать сегмент кода, сегмент данных, сегмент стека. Сегмент стека может быть опущен, в таком случае будет использоваться стек DOS. В общем случае, EXE-программа должна содержать не менее одного сегмента.

3) Какие директивы должны обязательно быть в тексте COM-программы? В тексте COM-программы должны обязательно присутствовать:

-директива ASSUME, которая связывает сегментные регистры с именем единственного сегмента;

-директива ORG 100h, которая обеспечивает смещение на 256 байт от нулевого адреса для устранения попадания в PSP.

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

В СОМ-программе не поддерживаются команды с указанием сегментов в виде операндов, т.к. в СОМ-программах отсутствует таблица настройки, которая используется для определения абсолютных адресов сегментов.

Отличия форматов файлов СОМ и ЕХЕ модулей

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

Файл состоит из одного сегмента, в котором находятся код и данные. Код располагается с адреса 0, так как в СОМ-файлы не содержат заголовка и таблицы настройки адресов. Размер файла не может превышать 65280 байт.

```
Address 0 1 2 3 4 5 6 7 8 9 a b c d e f Dump
00000000 e9 28 02 92 a8 af 20 50 43 3a 20 50 43 0d 0a 24 й(.'ËÏ PC: PC..$
00000010 92 a8 af 20 50 43 3a 20 50 43 2f 58 54 0d 0a 24 'ËÏ PC: PC/XT..$
00000020 92 a8 af 20 50 43 3a 20 41 54 0d 0a 24 92 a8 af 'ËÏ PC: AT..$'ËÏ
00000030 20 50 43 3a 20 50 53 32 20 ac ae a4 a5 ab ec 20 PC: PS2 ¨¤Ґ«м
00000040 33 30 0d 0a 24 92 a8 af 20 50 43 3a 20 50 53 32 30..$'EÏ PC: PS2
00000050 20 ac ae a4 a5 ab ec 20 35 30 20 a8 ab a8 20 36 →®¤Ґ«м 50 Ё«Ё 6
00000060 30 0d 0a 24 92 a8 af 20 50 43 3a 20 50 53 32 20 0..$'ËÏ PC: PS2
00000070 ac ae a4 a5 ab ec 20 38 30 0d 0a 24 92 a8 af 20 ¨¤Ґ«м 80..$°ËΪ
00000080 50 43 3a 20 50 91 6a 72 0d 0a 24 92 a8 af 20 50 PC: P'jr..$'ËÏ P
00000090 43 3a 20 50 43 20 43 6f 6e 76 65 72 74 69 62 6c C: PC Convertibl
0000000a0 65 0d 0a 24 92 a8 af 20 50 43 3a 20 ad a5 ae af e..$'ËÏ PC: -Ґ⊗Ï
000000b0 e0 a5 a4 a5 ab a5 ad 2c 20 aa ae a4 3a 20 20 0d aT'¤T'«T'-, €®¤:
000000c0 0a 24 82 a5 e0 e1 a8 ef 20 4d 53 2d 44 4f 53 3a .$, ҐабЁп MS-DOS:
000000d0 20 20 2e 20 20 0d 0a 24 91 a5 e0 a8 a9 ad eb a9 . ..$'ҐaË©-π©
0000000e0 20 ad ae ac a5 e0 20 4f 45 4d 3a 20 20 20 48 0d -8-1a OEM: H.
000000f0 0a 24 91 a5 e0 a8 a9 ad eb a9 20 ad ae ac a5 e0 .$'ľaË©-π© -®-Ґa
00000100 20 af ae ab ec a7 ae a2 a0 e2 a5 ab ef 3a 20 20 Ї®«м§®ў.⥫п:
00000110 20 20 20 20 20 48 24 24 0f 3c 09 76 02 04 07 04
00000120 30 c3 51 8a e0 e8 ef ff 86 c4 b1 04 d2 e8 e8 e6 0ГОЉаиля†П±.Тииж
00000130 ff 59 c3 53 8a fc e8 e9 ff 88 25 4f 88 05 4f 8a яYГSЉьийя€%О€.ОЉ
00000140 c7 e8 de ff 88 25 4f 88 05 5b c3 51 52 32 e4 33 SuMOs€%O€.[PQR2x3
00000150 d2 b9 0a 00 f7 f1 80 ca 30 88 14 4e 33 d2 3d 0a T№..чсЪКО€.N3T=.
00000160 00 73 fl 3c 00 74 04 0c 30 88 04 5a 59 c3 b4 09 .sc<.t..0€.ZYFr.
00000170 cd 21 c3 b8 00 f0 8e c0 26 a0 fe ff 3c ff 74 le H!Pë.p%A&.mg<st.
00000180 3c fe 74 20 3c fb 74 1c 3c fc 74 1e 3c fa 74 20 < mt < mt < bt < ***
00000190 3c f8 74 22 3c fd 74 24 3c f9 74 26 74 2a ba 03 <mt"<>t$<mt&t*e.
000001a0 01 eb 38 90 ba 10 01 eb 32 90 ba 20 01 eb 2c 90 .л8.е..л2.е .л,.
000001b0 ba 2d 01 eb 26 90 ba 64 01 eb 20 90 ba 7c 01 eb e-.л&.ed.л .e|.л
000001c0 la 90 ba 8b 01 eb 14 90 bf a4 01 83 c7 19 8a c7 ..e<.л..їж.́тЗ.ЉЗ
000001d0 e8 4f ff 89 05 ba a4 01 eb 01 90 e8 90 ff c3 b4 иОяж.сж.л..и.яГг
000001e0 30 cd 21 50 be c2 01 83 c6 0f e8 5e ff 58 8a c4 0H!PsB.ŕЖ.и^яХЪД
000001f0 83 c6 03 e8 55 ff ba c2 01 e8 72 ff bf d8 01 83 fЖ.иUяєВ.игяїШ.́́р
00000200 c7 14 8a c7 e8 lb ff 89 05 ba d8 01 e8 5f ff bf 3.ЉЗи.я%.eШ.и яї
00000210 f2 01 83 c7 22 8b cl e8 19 ff 8a c3 e8 03 ff 83 т.ŕS"<Би.яЉГи.я́г
00000220 ef 02 89 05 ba f2 01 e8 44 ff c3 e8 45 ff e8 ae п.ж.ет.иDяГиЕяи®
00000230 ff 32 c0 b4 4c cd 21
                                                         g2ArLH!
```

Рисунок 4 – содержимое lb1 сот.com

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

«Плохой» ЕХЕ файл содержит в себе заголовок, таблицу настройки адресов, 1 сегмент, в котором находятся код и данные. Код располагается с адреса 300h, а с адреса 0 располагается заголовок ЕХЕ файла, в котором содержится сигнатура ЕХЕ файла, размер файла, размер заголовка, смещение таблицы настроек адресов и др. технические параметры.

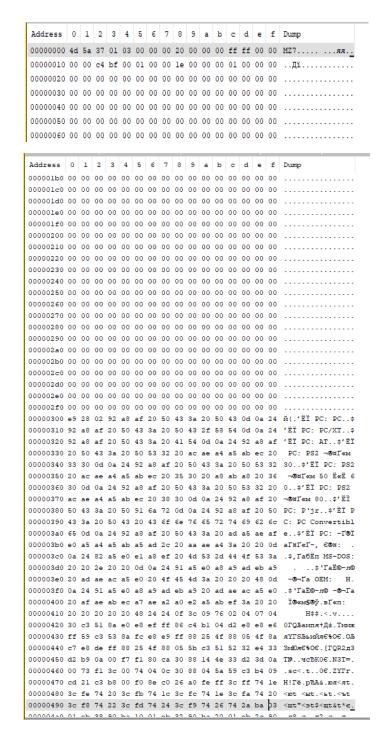


Рисунок 5 – содержимое lb1 сот.ехе

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

«Хороший» ЕХЕ файл содержит заголовок, таблицу настройки адресов, затем сегмент стека, сегмент данных и сегмент кода. «Хороший» ЕХЕ файл от «плохого» отличается наличием деления на сегменты (данные находятся в своем личном сегменте).

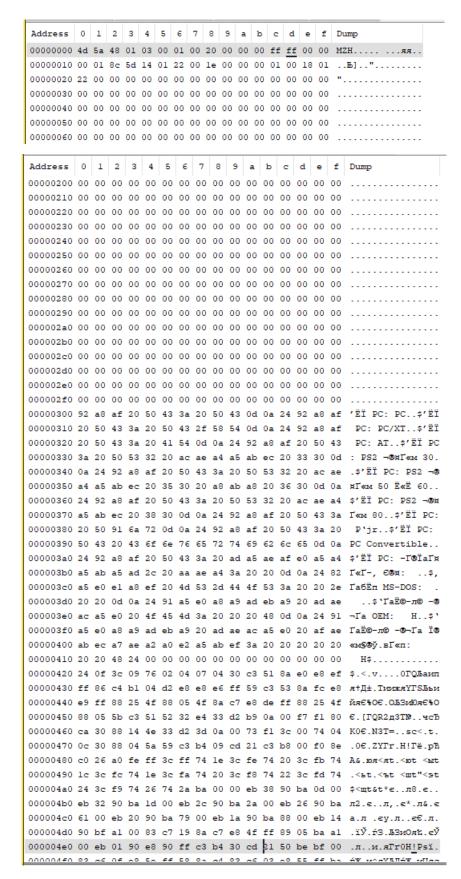


Рисунок 6 – содержимое lb1 exe.exe

Загрузка СОМ модуля в основную память

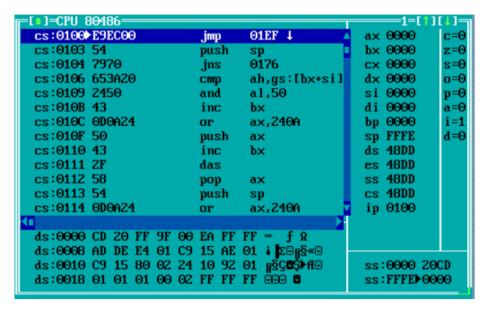


Рисунок 7 – отладчик TD.EXE с открытым COM-файлом

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

Выделяется свободный сегмент памяти достаточного размера. Затем в первые 256 байт записывается PSP, после которого следует содержание СОМфайла. В стек записывается адрес начала PSP, сегментные регистры также указывают на начало PSP. Указатель стека SP принимает значение конца сегмента. IP на начало программы принимает значение 0100h. Код располагается с адреса 0100h.

- 2) Что располагается с адреса 0?
- С адреса 0 располагается PSP (префикс программного сегмента).
- 3) Какие значения имеют сегментные регистры? На какие области памяти они указывают?

Сегментные регистры CS, DS, SS, ES указывают на начало PSP и имеют значение 48BD.

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

Стек в СОМ программе генерируется автоматически и находится в сегменте кода. В него автоматически записывается значение 0000. Указатель стека на начало программа принимает значение конца сегмента — FFFEh. Таким образом, по мере заполнения стека значение указателя будет уменьшаться и так таковой конечной границей является адрес 0000.

Загрузка «хорошего» EXE модуля в основную память

cs:0105 50 push ax bx 0000 z=0 cs:0106 B8FD48 mov ax,48FD cx 0000 s=0 cs:0109 BED8 mov ds,ax dx 0000 o=0 cs:010B E84EFF call 005C si 0000 p=0 cs:010E E8A8FF call 0089 di 0000 a=0 cs:0111 32C0 xor al,al bp 0000 i=1	[□[□]=CPU 80486======		1=[†][↓]¬
cs:0106 B8FD48 mov ax,48FD cx 0000 s=0 cs:0109 8ED8 mov ds,ax dx 0000 s=0 cs:010B E84EFF call 005C si 0000 p=0 cs:010E E8A8FF call 00B9 di 0000 a=0 cs:0111 32C0 xor al,al sp 0100 d=0 cs:0115 CD21 int 21 ds 48DD cs:0117 0000 add [bx+si],al ss 48ED cs:0119 0000 add [bx+si],al cs 490D cs:011B 0000 add [bx+si],al cs 490D cs:011D 0000 add [bx+si],al cs 490D ip 0103	cs:0103>2BC0	sub ax,ax	ax 0000 c=0
cs:0109 8ED8 mov ds,ax dx 0000 p=0 cs:010B E84EFF call 005C si 0000 p=0 cs:010E E8A8FF call 00B9 di 0000 a=0 cs:0111 32C0 xor al,al sp 0100 d=0 cs:0113 B44C mov ah,4C sp 0100 d=0 cs:0115 CD21 int 21 ds 48DD cs:0117 0000 add [bx+si],al cs 48DD cs:0119 0000 add [bx+si],al cs 48ED cs:011B 0000 add [bx+si],al cs 490D cs:011B 0000 add [bx+si],al cs 490D cs:011D 0000 add [bx+si],al ip 0103	cs:0105 50	push ax	bx 0000 z=0
cs:010B E84EFF call 005C si 0000 p=0 cs:010E E8A8FF call 00B9 di 0000 a=0 cs:0111 32C0 xor al,al bp 0000 i=1 cs:0113 B44C mov ah,4C sp 0100 d=0 cs:0115 CD21 int 21 ds 48DD cs:0117 0000 add [bx+si],al cs 48DD cs:0119 0000 add [bx+si],al cs 48ED cs:011B 0000 add [bx+si],al cs 490D cs:011D 0000 add [bx+si],al cs 490D ip 0103	cs:0106 B8FD48	mo∨ a×,48FD	cx 0000 s=0
cs:010E E8A8FF call 00B9 di 0000 a=0 cs:0111 32C0 xor al,al bp 0000 i=1 cs:0113 B44C mov ah,4C sp 0100 d=0 cs:0115 CD21 int 21 ds 48DD cs:0117 0000 add [bx+si],al es 48DD cs:0119 0000 add [bx+si],al ss 48ED cs:011B 0000 cs:011D 0000 add [bx+si],al ip 0103	cs:0109 8ED8	mov ds,ax	d× 0000 o=0
cs:0111 32C0 xor al,al bp 0000 i=1 cs:0113 B44C mov ah,4C sp 0100 d=0 cs:0115 CD21 int 21 ds 48DD cs:0117 0000 add Ibx+sil,al es 48DD cs:0119 0000 add Ibx+sil,al cs 49CD cs:011B 0000 add Ibx+sil,al cs 490D cs:011D 0000 add Ibx+sil,al ip 0103	cs:010B E84EFF	call 005C	si 0000 p=0
cs:0113 B44C mov ah,4C sp 0100 d=0 cs:0115 CD21 int 21 cs:0117 0000 add [bx+si],al cs:0119 0000 add [bx+si],al cs:0119 0000 add [bx+si],al cs:011B 0000 add [bx+si],al cs:011D 0000 add [bx+si],al ip 0103 cs:011D 0000 ip 0103	cs:010E E8A8FF	call 00B9	di 0000 a=0
cs:0115 CD21 int 21 ds 48DD cs:0117 0000 add [bx+si],al es 48DD cs:0119 0000 add [bx+si],al ss 48ED cs:011B 0000 add [bx+si],al cs 490D cs:011D 0000 add [bx+si],al ip 0103	cs:0111 3200	xor al,al	bp 0000 i=1
cs:0117 0000 add [bx+si],al es 48DD cs:0119 0000 add [bx+si],al ss 48ED cs:011B 0000 add [bx+si],al cs 490D cs:011D 0000 add [bx+si],al ip 0103	cs:0113 B44C	mov ah,4C	sp 0100 d=0
cs:0119 0000 add [bx+si],al ss 48ED cs:011B 0000 add [bx+si],al cs 490D cs:011D 0000 add [bx+si],al ip 0103	cs:0115 CD21	int 21	ds 48DD
cs:011B 0000 add [bx+si],al cs 490D cs:011D 0 <mark>0</mark> 00 add [bx+si],al ip 0103	cs:0117 0000	add [bx+si],al	es 48DD
cs:011D 0 <mark>0</mark> 00 add [bx+si],al ip 0103	cs:0119 0000	add [bx+si],al	ss 48ED
(□)		add [bx+si],al	cs 490D
ds:0000 CD 20 FF 9F 00 EA FF FF = f Ω	cs:011D 0 <mark>0</mark> 00	add [b×+si],al	ip 0103
ds:0000 CD 20 FF 9F 00 EA FF FF = f Ω	4 □) -	
ds:0008 AD DE E4 01 C9 15 AE 01 ↓ E□ S«□	ds:0008 AD DE E4 01 C9	15 AE 01 ↓ Σ□ _[§≪□	
ds:0010 C9 15 80 02 24 10 92 01 [§Ç @ \$▶ff□ ss:0102 6570			
ds:0018 01 01 01 00 02 FF FF FF □□□ © ss:0100▶7954	ds:0018 01 01 01 00 02	FF FF FF ©©© ©	ss:0100⊁7954

Рисунок 8 – отладчик TD.EXE с открытым EXE-файлом

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

Определяется сегментный адрес свободного участка памяти, затем создается два блока: блок памяти для переменных среды и блок памяти для PSP и программы. В блок памяти переменных среды помещается путь к файлу программы, затем записывается PSP, с помощью которого определяется сегментный адрес для загрузки программы. Далее после обработки таблицы настройки адресов, определяется абсолютные адреса сегментов. К началу выполнения программы DS и ES устанавливаются на начало сегмента PSP(48DDh), SS – на начало сегмента стека(48EDh), CS – на начало сегмента команд(490Dh). Управление передается по адресу точки входа, указанного в заголовке файла(0103h).

2) На что указывают регистры DS и ES?

В начале выполнения программы регистры DS и ES указывают на начало сегмента PSP.

3) Как определяется стек?

Стек определяется либо упрощенной директивой .STACK, которой обозначается начало сегмента стека, либо стандартной директивой SEGMENT.

Формат: ИмяСегмента SEGMENT STACK

. . .

ИмяСегмента ENDS

4) Как определяется точка входа?

Точка входа определяется в сегменте кода директивой END (start_label).