

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

ОТЧЕТ
по лабораторной работе №1
по дисциплине «Операционные системы»
Тема: Исследование структур загрузочных модулей

Студент гр. 7381

Аженилок В.А.

Преподаватель

Ефремов М.А.

Санкт-Петербург

2018

Цель работы.

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

Необходимые сведения для составления программы.

Тип IBM PC хранится в байте по адресу 0F000:0FFFE, в предпоследнем байте ROM BIOS. Соответствие кода и типа в таблице:

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

Для определения версии MS DOS следует воспользоваться функцией 30H прерывания 21H. Входным параметром является номер функции в AH:

MOV AH,30h

INT 21h

Выходными параметрами являются:

AL – номер основной версии. Если 0, то <2.0;

AH – номер модификации;

BH – серийный номер OEM (Original Equipment Manufacturer);

BL:CH – 24-битовый серийный номер пользователя.

Постановка задачи.

Требуется реализовать текст исходного .COM модуля, который определяет тип PC и версию системы. 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 xx.yy, где xx - номер основной версии, а yy - номер модификации в десятичной системе счисления, формировать строки с серийным номером OEM (Original Equipment Manufacturer) и серийным номером пользователя. Полученные строки выводятся на экран.

Далее необходимо отладить полученный исходный модуль и получить «хороший» .COM модуль, а также необходимо построить «плохой» .EXE, полученный из исходного текста для .COM модуля.

Затем нужно написать текст «хорошего» .EXE модуля, который выполняет те же функции, что и модуль .COM, далее его построить, отладить и сравнить исходные тексты для .COM и .EXE модулей.

Процедуры используемые в программе.

TETR_TO_HEX – Используется для перевода половины байта в шестнадцатеричную систему счисления.

BYTE_TO_HEX – Используется для перевода байта регистра AL в шестнадцатеричную систему счисления, помещая результат в AX.

WRD_TO_HEX – Используется для перевода двух байт регистра AX в шестнадцатеричную систему счисления, помещая результат в регистр DI.

BYTE_TO_DEC – Используется для перевода байта регистра AL в десятичную систему счисления, помещая результат в SI.

TYPE_IBM_PC – Определяет тип IBM PC.

VERS_DOS – Определяет версию MS DOS.

OEM_DOS – Определяет серийный номер OEM.

USER_DOS - Определяет серийный номер пользователя.

PRINT – Вывод на экран.

Структуры данных.

Таблица 1 – Структуры данных

Название поля данных	Тип	Назначение
_type	db	Тип IBM PC
_PC	db	PC
_PC_XT	db	PC/XT
_AT	db	AT
_PS2_30	db	PS2 модель 30
_PS2_50_60	db	PS2 модель 50 или 60
_PS2_80	db	PS2 модель 80
_PCjr	db	PCjr
_PC_Conv	db	PC Convertible
_ver	db	Номер версии MS DOS
_oem	db	Серийный номер OEM
_user	db	Серийный номер пользователя

Ход работы.

Шаг 1. Запуск «хорошего» .COM модуля.

```
C:\>lab1_1.com
Тип IBM PC PC/XT
Номер версии MS DOS: 5.0
Серийный номер OEM: 255
Серийный номер пользователя: 000000
```

Рисунок 1 – «Хороший» .COM модуль

Запуск «плохого» .EXE модуля.

```
C:\>lab1_1.exe

щТѢТип IMB PC

255      000000      5 0
щТѢТип IMB PC      щТѢТип IMB PC
00      255      000000
щТѢТип IMB PC      щТѢТип IMB PC
00      0000
C:\>_
```

Рисунок 2 – «Плохой» .EXE модуль

Шаг 2. Запуск «хорошего» .EXE модуля.

```
C:\>lab1_2.exe
Тип IMB PC PC/XT
Номер версии MS DOS: 5.0
Серийный номер OEM: 255
Серийный номер пользователя: 000000
```

Рисунок 3 – «Хороший» .EXE модуль

Шаг 3. Ответы на контрольные вопросы. Отличия исходных текстов COM и EXE программ.

1) Сколько сегментов должна содержать COM-программа?

Один сегмент.

2) EXE программа?

EXE программа может содержать больше одного сегмента.

3) Какие директивы должны обязательно быть в тексте COM программы?

Директива ORG 100h (смещение 100h), так как при загрузке COM-файла в память DOS занимает первые 256 байт (100h) блоком данных PSP и располагает код программы только после этого блока. Директива ASSUME, ставящая в соответствие начало программы сегментам кода и данных.

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

Нет, не все, так как в отличие от EXE-программы, в которой существует таблица настроек (таблица разметки), называемая Relocation

Table, COM-программа ею не располагает. Адреса сегментов определяются загрузчиком в момент запуска программы на основе информации о местоположении полей адресов в файле из Relocation Table. Следовательно, в связи с отсутствием этой таблицы в COM-программах, команды вида `mov [регистр], seg [сегмент]` недопустимы.

Шаг 4. .COM модуль в шестнадцатеричном виде.

C:\MASM\LAB1_1.COM																	
0000000000:	E9	80	01	92	A8	AF	20	49	4D	42	20	50	43	20	24	50	щАТип IMB PC \$P
0000000010:	43	0D	0A	24	50	43	2F	58	54	0D	0A	24	41	54	0D	0A	С \$PC/XT \$AT \$
0000000020:	24	50	53	32	20	AC	AE	A4	A5	AB	EC	20	33	30	0D	0A	\$PS2 модель 30 \$
0000000030:	24	50	53	32	20	AC	AE	A4	A5	AB	EC	20	35	30	20	A8	\$PS2 модель 50 и
0000000040:	AB	A8	20	36	30	0D	0A	24	50	53	32	20	AC	AE	A4	A5	ли 60 \$PS2 моде
0000000050:	AB	EC	20	38	30	0D	0A	24	50	43	6A	72	0D	0A	24	50	ль 80 \$PCjr \$P
0000000060:	43	20	43	6F	6E	76	65	72	74	69	62	6C	65	0D	0A	24	С Convertible \$
0000000070:	8D	AE	AC	A5	E0	20	A2	A5	E0	E1	A8	A8	20	4D	53	20	Номер версии MS
0000000080:	44	4F	53	3A	20	20	2E	20	20	20	20	0D	0A	24	91	A5	DOS: . \$Ce
0000000090:	E0	A8	A9	AD	EB	A9	20	AD	AE	AC	A5	E0	20	4F	45	4D	рийный номер OEM
00000000A0:	3A	20	20	20	20	0D	0A	24	91	A5	E0	A8	A9	AD	EB	A9	: \$Серийный
00000000B0:	20	AD	AE	AC	A5	E0	20	AF	AE	AB	EC	A7	AE	A2	A0	E2	номер пользоват
00000000C0:	A5	AB	EF	3A	20	20	20	20	20	20	0D	0A	24	24	0F	3C	еля: \$ \$<
00000000D0:	09	76	02	04	07	04	30	C3	51	8A	E0	E8	EF	FF	86	C4	в\$Qкря Ж-
00000000E0:	B1	04	D2	E8	E8	E6	FF	59	C3	53	8A	FC	E8	E9	FF	88	тшщ Y SKNщ И
00000000F0:	25	4F	88	05	4F	8A	C7	E8	DE	FF	88	25	4F	88	05	5B	%OIOK ш И%OIO[
0000000100:	C3	51	52	32	E4	33	D2	B9	0A	00	F7	F1	80	CA	30	88	QR2фЗт \$ üäA-ои
0000000110:	14	4E	33	D2	3D	0A	00	73	F1	3C	00	74	04	0C	30	88	IN3т= \$ sě< t\$Oи
0000000120:	04	5A	59	C3	1E	B8	00	F0	8E	D8	2B	DB	B7	FE	1F	C3	\$ZY \$ ěO+ \$ \$
0000000130:	50	56	BE	70	01	83	C6	15	E8	C6	FF	BE	70	01	83	C6	PV p\$Г \$ п \$ п\$Г
0000000140:	17	8A	C4	E8	BB	FF	5E	58	C3	50	53	56	BE	8E	01	83	\$K-ш ^X-PSV O\$Г
0000000150:	C6	16	8A	C7	E8	AA	FF	5E	5B	58	C3	53	51	57	50	BF	\$K шк ^X-SQWP
0000000160:	A8	01	83	C7	22	8B	C1	E8	7F	FF	8A	C3	E8	69	FF	BF	и\$Г "л-ш \$ K wi 7
0000000170:	A8	01	83	C7	1D	89	05	58	5F	59	5B	C3	50	B4	09	CD	и\$Г \$Й\$X_Y [-P- \$=
0000000180:	21	58	C3	E8	9E	FF	BA	03	01	E8	F0	FF	BA	0F	01	80	!X ш0 \$ ш ě \$ \$A
0000000190:	FF	FF	74	47	BA	14	01	80	FF	FE	74	3F	BA	14	01	80	tG \$A ■t? \$A
00000001A0:	FF	FB	74	37	BA	1C	01	80	FF	FC	74	2F	BA	21	01	80	√t7 \$A №t/ !A
00000001B0:	FF	FA	74	27	BA	31	01	80	FF	FC	74	1F	BA	48	01	80	·t' 1A №t \$ H\$A
00000001C0:	FF	F8	74	17	BA	58	01	80	FF	FD	74	0F	BA	5F	01	80	°t \$ X\$A \$t \$ _A
00000001D0:	FF	F9	74	07	8A	C7	E8	FF	FE	8B	D0	E8	9E	FF	B4	30	·t \$ K ш ■л ш0 -0
00000001E0:	CD	21	E8	4B	FF	E8	61	FF	E8	70	FF	BA	70	01	E8	8B	=!шK ша шр p\$шл
00000001F0:	FF	BA	8E	01	E8	85	FF	BA	A8	01	E8	7F	FF	32	C0	B4	O\$шE и\$ш \$ 2 4
0000000200:	4C	CD	21														L=!

Рисунок 4 - .COM модуль в шестнадцатеричном виде

«Плохой» .EXE модуль в шестнадцатеричном виде.

Рисунок 5 - «Плохой» .EXE модуль в шестнадцатеричном виде
«Хороший» .EXE модуль в шестнадцатеричном виде.

C:\MASM\LAB1_2.EXE																									
0000000000: 4D 5A 0A 00 03 00 01 00	20 00 21 00 FF FF 21 00	MZ	00 00 00 00	!	ÿÿ!																				
0000000010: 00 02 54 D7 00 00 00 00	1E 00 00 00 01 00 BA 00	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00																				
0000000020: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000030: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000040: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000050: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000060: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000070: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000080: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000090: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
00000000A0: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
00000000B0: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
00000000C0: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
00000000D0: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
00000000E0: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
00000000F0: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000100: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000110: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000120: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000130: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000140: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000150: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000160: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000170: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000180: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000190: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
00000001A0: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
00000001B0: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
00000001C0: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
00000001D0: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
00000001E0: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
00000001F0: 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00																								
0000000200: E9 B6 00 24 0F 3C 09 76	02 04 07 04 30 C3 51 8A	éŋ	\$	0	0	0	0	0	0	0	0	0	0	0	0										
0000000210: E0 E8 EF FF 86 C4 B1 04	D2 E8 E8 E6 FF 59 C3 53	àèiÿt	À±	0	0	0	0	0	0	0	0	0	0	0	0										
0000000220: 8A FC E8 E9 FF 88 25 4F	88 05 4F 8A C7 E8 DE FF	Süèëÿ	~%	0	0	0	0	0	0	0	0	0	0	0	0										
0000000230: 88 25 4F 88 05 5B C3 51	52 32 E4 33 D2 B9 0A 00	~%0	0	0	0	0	0	0	0	0	0	0	0	0	0										
0000000240: F7 F1 80 CA 30 88 14 4E	33 D2 3D 0A 00 73 F1 3C	÷ñ€Ê	0	0	0	0	0	0	0	0	0	0	0	0	0										
0000000250: 00 74 04 0C 30 88 04 5A	59 C3 1E B8 00 F0 8E D8	t	0	0	0	0	0	0	0	0	0	0	0	0	0										
0000000260: 2B DB B7 FE 1F C3 50 56	BE 6D 00 83 C6 15 E8 C6	+0	·	p	0	0	0	0	0	0	0	0	0	0	0										
0000000270: FF BE 6D 00 83 C6 17 8A	C4 E8 BB FF 5E 58 C3 50	ÿ% _m	f	À	0	0	0	0	0	0	0	0	0	0	0										
0000000280: 53 56 BE 8B 00 83 C6 16	8A C7 E8 AA FF 5E 5B 58	SV%<	f	À	0	0	0	0	0	0	0	0	0	0	0										
0000000290: C3 53 51 57 50 BF A5 00	83 C7 22 8B C1 E8 7F FF	ÀSQWP	¿	ÿ	f	Ç	0	0	0	0	0	0	0	0	0										
00000002A0: 8A C3 E8 69 FF BF A5 00	83 C7 1D 89 05 58 5F 59	Š	À	0	0	0	0	0	0	0	0	0	0	0	0										
00000002B0: 5B C3 50 B4 09 CD 21 58	C3 B8 14 00 8E D8 8C DB	[À	0	0	0	0	0	0	0	0	0	0	0	0										
00000002C0: E8 97 FF BA 00 00 E8 E9	FF BA 0C 00 80 FF FF 74	è-ÿ	0	0	0	0	0	0	0	0	0	0	0	0	0										
00000002D0: 47 BA 11 00 80 FF FE 74	3F BA 11 00 80 FF FB 74	G	0	0	0	0	0	0	0	0	0	0	0	0	0										
00000002E0: 37 BA 19 00 80 FF FC 74	2F BA 1E 00 80 FF FA 74	7	0	0	0	0	0	0	0	0	0	0	0	0	0										
00000002F0: 27 BA 2E 00 80 FF FC 74	1F BA 45 00 80 FF F8 74	'	0	0	0	0	0	0	0	0	0	0	0	0	0										
0000000300: 17 BA 55 00 80 FF FD 74	0F BA 5C 00 80 FF F9 74	0	0	0	0	0	0	0	0	0	0	0	0	0	0										
0000000310: 07 8A C7 E8 F8 FE 8B D0	E8 97 FF B4 30 CD 21 E8	0	0	0	0	0	0	0	0	0	0	0	0	0	0										
0000000320: 44 FF E8 5A FF E8 69 FF	BA 6D 00 E8 84 FF BA 8B	DÿèZÿèiÿ	0	0	0	0	0	0	0	0	0	0	0	0	0										
0000000330: 00 E8 7E FF BA A5 00 E8	78 FF 32 C0 B4 4C CD 21	è~ÿ	0	0	0	0	0	0	0	0	0	0	0	0	0										
0000000340: 92 A8 AF 20 49 4D 42 20	50 43 20 24 50 43 0D 0A	'	0	0	0	0	0	0	0	0	0	0	0	0	0										
0000000350: 24 50 43 2F 58 54 0D 0A	24 41 54 0D 0A 24 50 43	\$PC/XT	0	0	0	0	0	0	0	0	0	0	0	0	0										
0000000360: 32 20 AC AE A4 A5 AB EC	20 33 30 0D 0A 24 50 43	2	~	0	0	0	0	0	0	0	0	0	0	0	0										
0000000370: 32 20 AC AE A4 A5 AB EC	20 35 30 20 A8 AB A8 20	2	~	0	0	0	0	0	0	0	0	0	0	0	0										
0000000380: 36 30 0D 0A 24 50 43 32	20 AC AE A4 A5 AB EC 20	60	0	0	0	0	0	0	0	0	0	0	0	0	0										
0000000390: 38 30 0D 0A 24 50 43 6A	72 0D 0A 24 50 43 20 43	80	0	0	0	0	0	0	0	0	0	0	0	0	0										
00000003A0: 6F 6E 76 65 72 74 69 62	6C 65 0D 0A 24 8D AE AC	onvertible	0	0	0	0	0	0	0	0	0	0	0	0	0										
00000003B0: A5 E0 20 A2 A5 E0 E1 A8	A8 20 4D 53 20 44 4F 53	ÿà	ø	ÿ	à	á	0	0	0	0	0	0	0	0	0										
00000003C0: 3A 20 20 2E 20 20 20 20	0D 0A 24 91 A5 E0 A8 A9	:	0	0	0	0	0	0	0	0	0	0	0	0	0										
00000003D0: AD EB A9 20 AD AE AC A5	E0 20 4F 45 4D 3A 20 20	-è0	0	0	0	0	0	0	0	0	0	0	0	0	0										
00000003E0: 20 20 0D 0A 24 91 A5 E0	A8 A9 AD EB A9 20 AD AE	0	0	0	0	0	0	0	0	0	0	0	0	0	0										
00000003F0: AC A5 E0 20 AF AE AB EC	A7 AE A2 A0 E2 A5 AB EF	-ÿà	0	0	0	0	0	0	0	0	0	0	0	0	0										
0000000400: 3A 20 20 20 20 20 20 0D	0A 24	:	0	0	0	0	0	0	0	0	0	0	0	0	0										

Рисунок 6 - «Хороший» .EXE модуль в шестнадцатеричном виде

Ответы на контрольные вопросы. Отличия форматов файлов COM и EXE программ.

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

COM файл состоит из одного сегмента и содержит данные и машинные команды. Код начинается с адреса 0h, но при загрузке модуля устанавливается смещение в 100h.

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

В «плохом» EXE файле данные и код содержатся в одном сегменте. Код располагается с адреса 300h. С адреса 0h располагается Relocation Table (таблица разметки).

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

В «хорошем» файле EXE содержится информация для загрузчика, сегмент стека, сегмент данных и сегмент кода (3 сегмента вместо одного в «плохом» .EXE). Код располагается с адреса 200h в отличии от 300h в «плохом» .EXE файле.

Шаг 5. Загрузка COM модуля в основную память.

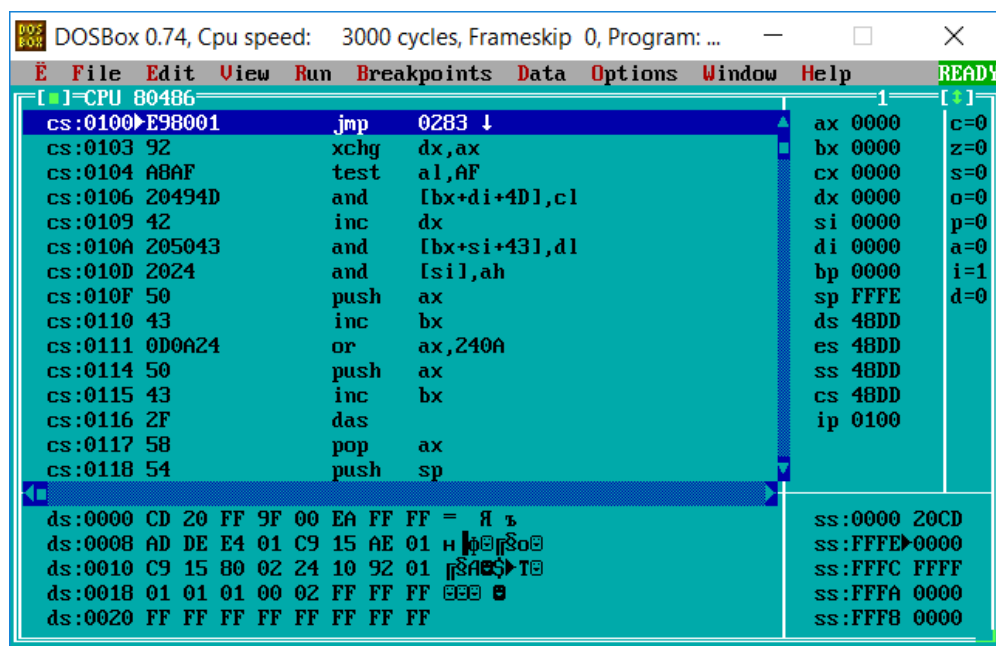


Рисунок 7 – Загрузка COM модуля в основную память

Ответы на контрольные вопросы. Загрузка COM модуля в основную память.

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

После загрузки COM-программы в память сегментные регистры указывают на начало PSP. Код располагается с адреса 100h (ip = 0100h).

2) **Что располагается с 0 адреса?**

Адрес начала PSP.

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

48DDh. Они указывают на начало PSP.

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

Стек определяется автоматически, указатель стека устанавливается на конец сегмента. Если для программы размер сегмента в 64КБ является достаточным, то DOS устанавливает в регистре SP адрес конца сегмента – FFFEH. Адреса расположены в диапазоне 0000h-FFFEh.

Шаг 6. Загрузка «хорошего» EXE модуля в память.

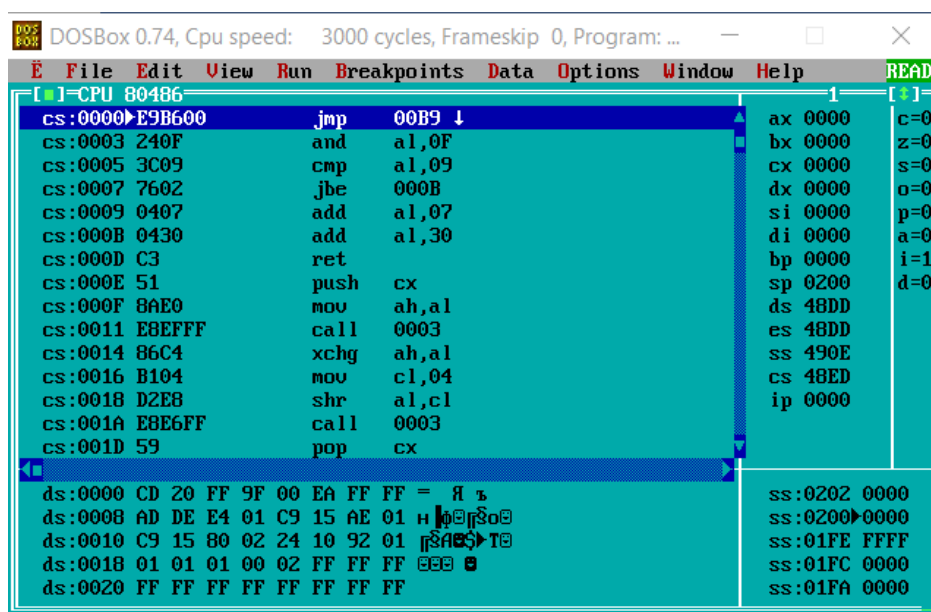


Рисунок 8 – Загрузка «хорошего» EXE модуля в память

Ответы на контрольные вопросы. Загрузка «хорошего» EXE модуля в память.

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

В области памяти строится PSP, стандартная часть заголовка считывается в память, определяется длина тела загрузочного модуля, определяется начальный сегмент, загрузочный модуль считывается в начальный сегмент, таблица настройки считывается в рабочую память, определяются значения сегментных регистров. DS и ES устанавливаются на начало PSP, SS - на начало стека, CS - на начало сегмента кода.

2) На что указывают регистры DS и ES?

DS и ES указывают на начало PSP. После выполнения команд `mov ax, @data` и `mov ds, ax` регистре DS содержит адрес начала сегмента данных.

3) Как определяется стек?

В исходном коде модуля стек определяется при помощи директивы `STACK`, а при исполнении в регистры SS и SP записываются адрес начала сегмента стека и его вершины соответственно.

4) Как определяется точка входа?

При помощи команды `END`.

Вывод.

В ходе работы было проведено исследование различий в структурах исходных текстов модулей `.COM` и `.EXE`, структур файлов загрузочных модулей и способов их загрузки в основную память.