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

Студент гр. 0381	 Ефимов Н.Д
Преподаватель	 Губкин А.Ф.

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.

Требуется написать 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 запись числа И выводиться на экран 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 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 .СОМ, далее его построить, отладить и сравнить исходные тексты для .СОМ и .ЕХЕ модулей

Основные теоретические положения.

Изначально были даны несколько процедур, выполняющих некоторые операции:

TETR_TO_HEX - Перевод десятичной цифры в код символа BYTE_TO_HEX - Перевод байта в 16-ной с/с в символьный код WRD_TO_HEX - Перевод слова в 16-ной с/с в символьный код

BYTE_TO_DEC - Перевод байта в 16-ной с/с в символьный код в 10-ной с/с Выполнение работы.

WRITESTRING – Вывод строки на экран

Выполнение работы.

Для выполнения задания нужно написать две процедуры: первая – для определения типа PC, а вторая – для определения характеристик OS. Эти процедуры называются PCTYPE и OSVERSION.

1 шаг – написание файла для .COM модуля. Название файла: lb1 com.asm. Взяв за основу код из методических указаний, был написан остов программы. Дальше следовало добвать ряд переменных, которые будут хранить в себе нужную информацию для работы кода. Эти переменные приведены в разделе «Данные». Дальше были написаны нужные процедуры: процедура РСТҮРЕ работает с информацией по адресу 0F000:0FFFEh. После взятия данных происходи их распределение согласно таблице, приложенной в руководстве к выполнению лабораторной работы. Вторя процедура, OSVERSION, работает с функцией 30h прерывания 21h, а затем, получив нужные данные, помещает их в отведенные переменные в начале кода. Дальше идет запуск программы в DOSBOX. После выполнения программы выводится следующее:

```
F:\>lb1_com.com
Type: AT
Version MS-DOS: 5.0
Serial number OEM: 0
User serial number: 000000H
F:\>_
```

Если попробовать запустить код, но не переводить его из EXE, то получится следующее:

Это и будет «плохой» .EXE модуль.

2 шаг — написание программы для «хорошего» модуля .EXE. Для этого нужно немного исправить файл lb1_com.asm. В этом файле не прописывается стек, поэтому и код, и данные после работы компилятора будут лежать в одной ячейке. Для этого нужно разделить сегмент кода и данных. Добавлением разделения сегмента стека, сегмента данных, сегмента кода получается файл lb1 exe.asm. После сборки в DOSBOX получится следующее:

```
F:\>lb1_exe.exe
Type: AT
Version MS-DOS: 5.0
Serial number OEM: 0
User serial number: 000000H
F:\>
```

3 шаг – сравнение двух исходных тектов модулей .COM и .EXE. В прошлом пункте были описаны отличия, а ответы на вопросы к лабораторной работе будут даны в A.

4,5,6 шаг – ответы на вопросы будут даны в Приложении А.

Выводы.

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

ПРИЛОЖЕНИЕ А ОТВЕТЫ НА ВОПРОСЫ

Отличия исходных текстов СОМ и ЕХЕ программ

1. Сколько сегментов должна содержать СОМ-программа?

Ответ: СОМ-программа должна содержать только один сегмент. Код и данные находятся в одном сегменте, а стек генерируется автоматически.

2. ЕХЕ-программа?

Такой тип программы должен содержать не менее одного сегмента. Сегменты данных, стека и кода описываются отдельно. Если сегмент стека не описан, то будет использоваться стек DOS

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

Должна быть обязательна директива ORG 100h, так как при загрузке модуля все сегментные регистры содержат адрес префикса программного сегмента (PSP), который является 256-байтовым(100H) блоком, поэтому адресация имеет смещение в 256 байт от нулевого адреса. Также необходима процедура ASSUME для того, чтобы сегмент данных и сегмент кода указывали на один общий сегмент. (ASSUME CS:LABONE, DS:LABONE, ES:NOTHING, SS:NOTHING)

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

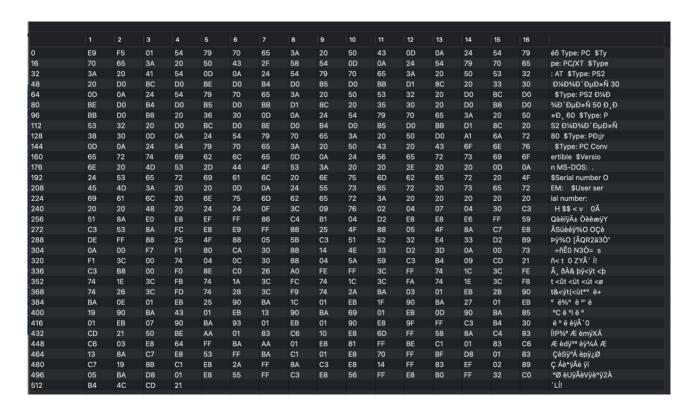
Не все форматы поддерживаются. Нельзя использовать команды вида mov <pегистр>, seg <имя сегмента>, так как в .com-программе отсутствует таблица настроек (содержит описание адресов, которые зависят от размещения загрузочного модуля в ОП).

Отличия форматов файлов СОМ и ЕХЕ модулей

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

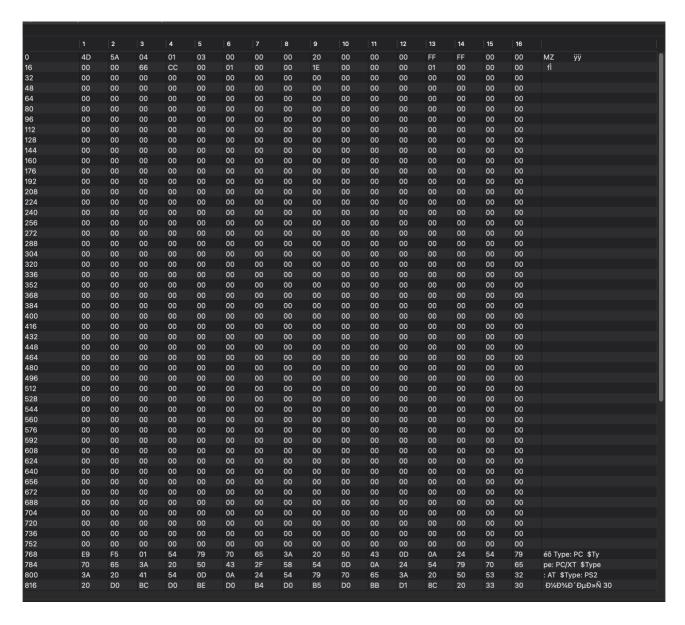
СОМ-файл состоит из одного сегмента, состоящего из сегмент кода и сегмент данных, сегмент стека генерируется автоматически при создании СОМ-программы. СОМ-файл ограничен размером одного сегмента и не превышает 64 Кб

Код начинается с адреса 0h, но при загрузке модуля устанавливается смещение в 100h.



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

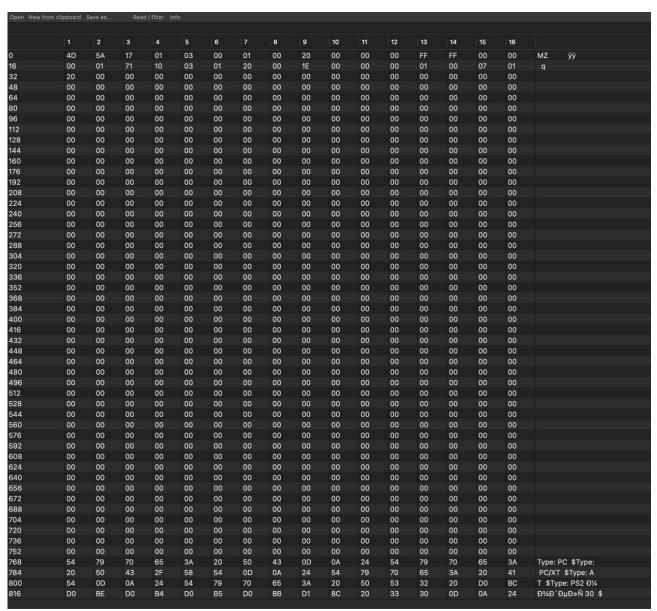
В «плохом» ЕХЕ данные и код располагаются в одном сегменте. Это неправильно, так как код и данные должны быть разделены на отдельные сегменты. Код располагается с адреса 300h, а с адреса 0h идёт таблица настроек.



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

В ЕХЕ-программе код, данные и стек — отдельные сегменты. ЕХЕ-файл имеет заголовок, который используется при его загрузке. Заголовок состоит из форматированной части, содержащей сигнатуру и данные, необходимые для загрузки ЕХЕ-файла, и таблицы для настройки адресов. В отличии от «плохого» ЕХЕ в «хорошем» ЕХЕ присутствуют три сегмента: сегмент кода, сегмент данных и сегмент стека, а «плохой» ЕХЕ содержит один сегмент, совмещающий код и данные. Также в «плохом» ЕХЕ адресация кода начинается с 300h, так как он получается из .СОМ файла, в котором изначально сегмент кода смещён на 100h, а при создании «плохого» ЕХЕ к этому смещению добавляется размер PSP

модуля(200h). А в «хорошем» ЕХЕ присутствует только смещение для PSP модуля, поэтому код начинается с 200h. В данной случае смещение кода 400h так как выделяется память под стек (200h), память под стек находится между PSP и кодом.

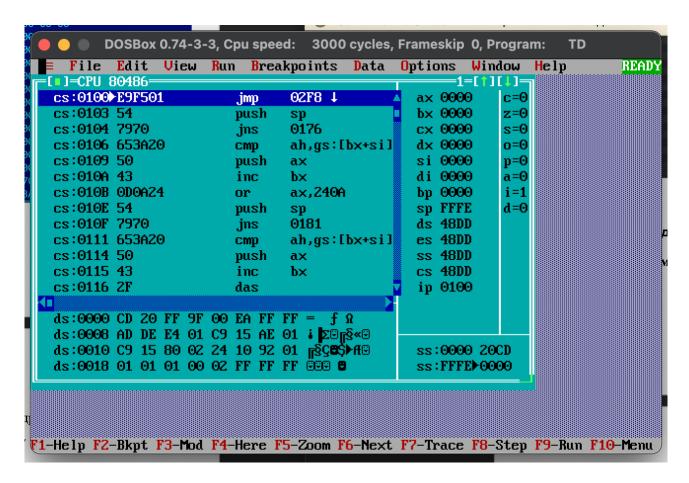


Загрузка СОМ модуля в основную память

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

Определяется сегментный адрес участка ОП, у которого достаточно места для загрузки программы, образ СОМ-файла считывается с диска и помещается в

память, начиная с PSP:0100h. После загрузки двоичного образа COM-программы сегментные регистры CS, DS, ES и SS указывают на PSP(в данном случае сегментные регистры указывают на 48DD), SP указывает на конец сегмента PSP(обычно FFFE), слово 00H помещено в стек, IP содержит 100H в результате команды JMP PSP:100H.



2. Что располагается с адреса 0?

Программный сегмент PSP, размером 256 байт (100h), зарезервированный операционной системой.

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

Ссегментные регистры CS, DS, ES и SS указывают на PSP и имеют значения 48DD.

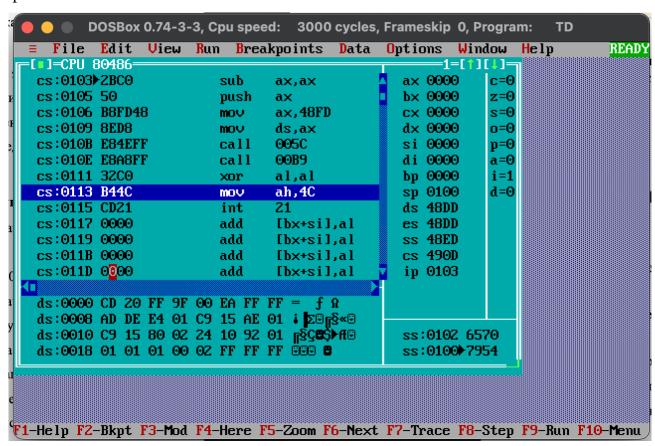
4. Как определяется стек? Какую область памяти он занимает? Какие адреса?

Стек генерируется автоматически при создании COM-программы. SS – на начало (0h), регистр SP указывает на конец стека (FFFEh), Адреса стека расположены в диапазоне 0h – FFFEh (FFFEh, – последний адрес, кратный двум).

Загрузка «хорошего» ЕХЕ модуля в основную память

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

ЕХЕ-файл загружается с адреса PSP:0100h. В процессе загрузки считывается информация заголовка (PSP) 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 EXEфайле.



2. На что указывают регистры DS и ES?

Они указывают на начало сегмента PSP

3. Как определяется стек?

Стек определяется директивой stack, после которой задается размер стека. При исполнении регистр SS указывает на начало этого стека, а SP - на конец стека

4. Как определяется точка входа?

При помощи директивы END.