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. 7383 _____ Лосев М.Л.

Преподаватель _____ Ефремов М.А.

Санкт-Петербург

2018

Постановка задачи.

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

Ход работы

Для выполнения работы была написана программа на языке ассемблера. Она транслировалась с помощью компилятора TASM 5.0 и собиралась с помощью TLINK.

Сведения об используемых функциях и структурах данных.

BYTE_TO_HEX – переводит значение регистра AL в его запись в шестнадцатеричной с/с, помещает ее в AX

WRD_TO_HEX – переводит значение регистра AX в его запись в шестнадцатеричной с/с, помещает ее в память так, что DI указывает на младшую цифру.

BYTE_TO_DEC – переводит значение регистра AL в его запись в десятичной ичной с/с, помещает результат в память так, что SI указывает на младшую цифру.

OUTPUT_PROC – вызывает прерывание DOS вывода строки.

Последовательность действий, выполняемых утилитой.

Утилита получает байт кода типа компьютера и сопоставляет его с некоторыми кодами. В случае совпадения с одним из них она выводит соответствующее сообщение, если же совпадения нет, то код типа компьютера переводится в символьную запись в 16 с/с и выводится на экран. С помощью функции 30h прерывания 21h определяется версия DOS, OEM, номер пользователя. Версия DOS переводится в символьное представление в 10 с/с и выводится. OEM и номер пользователя переводятся в символьное представление в 16 с/с и выводятся на экран. Происходит вызод в DOS.

Результаты исследования.

Отличия исходных текстов COM и EXE программ:

1) *Сколько сегментов должна содержать COM-программа?*

COM-программа должна содержать один сегмент.

2) *EXE-программа?*

EXE-программа может содержать один или несколько сегментов. Обычно она содержит сегменты данных, стека и кода. Может содержать дополнительный сегмент. Может содержать более 4 сегментов.

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

ORG 100h - директива ORG служит для резервирования 256 байт от начального адреса под PSP, устанавливает счетчик адреса в нужное абсолютное значение.

SEGMENT — определяет начало и конец сегмента памяти.

ASSUME — задает значения сегментных регистров.

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

Нет. Нельзя использовать команды, содержащие адреса сегментов, например, CODE и DATA.

Отличия форматов файлов COM и EXE модулей:

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

COM-файл содержит данные программы и машинный код. Код располагается с самого начала файла (см. рис. 1).

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

Первые два байта заняты сигнатурой EXE-файла, далее до 1Bh включительно остальная часть стандартного заголовка (заголовок у «плохого» и «хорошего» EXE-файлов одинаковый, см. рис. 3), далее следует таблица настройки адресов (в случае «плохого» EXE-файла нули). Таблица состоит из элементов, число которых записано в байтах 06-07 заголовка. Элемент таблицы настройки состоит из двух полей: 2-байтного смещения и 2-байтного сегмента, и указывает слова в загрузочном модуле, содержащее адрес, который должен быть настроен на место памяти, в которое загружается задача. Для каждого элемента таблицы настройки к полю сегмента прибавляется сегментный адрес начального

сегмента. В результате элемент таблицы указывает на слово в памяти, к которому прибавляется сегментный адрес начального сегмента. В COM такой таблицы нет, потому что все находится в одном сегменте, и не используются команды, содержащие адреса сегментов, например, CODE и DATA. Код начинается с 300h по счету байта (см. рис. 2).

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

Он похож на «плохой», но в нем данные начинаются с 220h байта (см. рис. 4).

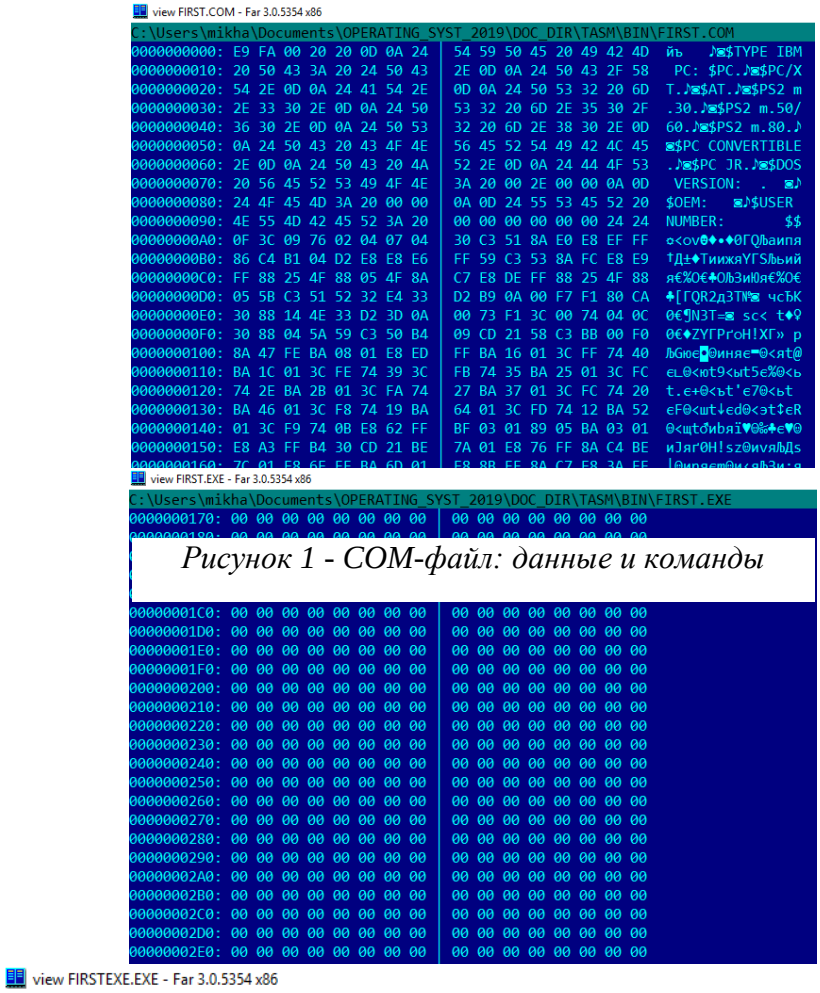

Рисунок 1 - COM-файл: данные и команды

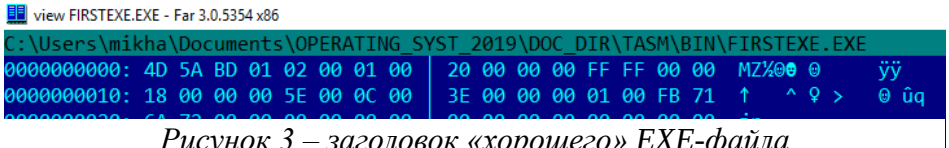

Рисунок 3 – заголовок «хорошего» EXE-файла

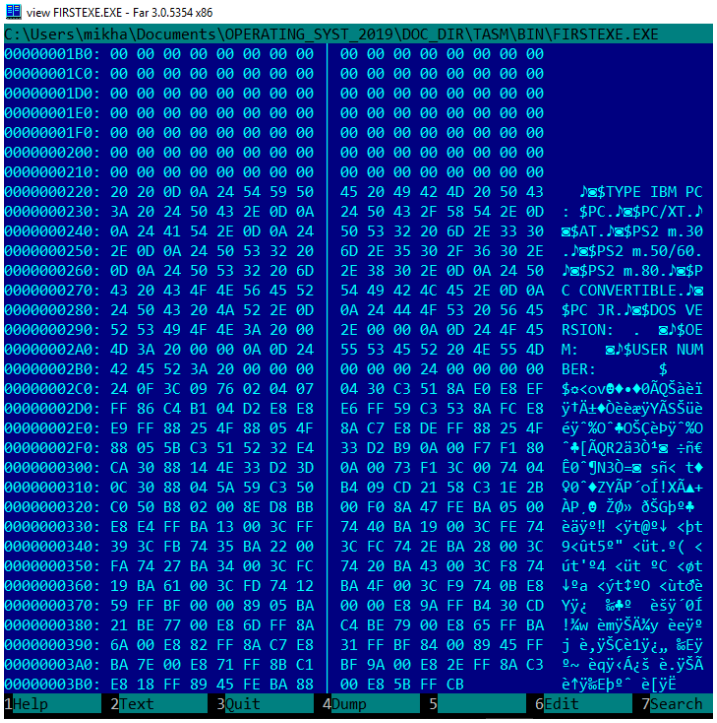

Рисунок 4 – данные и команды «хорошего» EXE-файла

Загрузка COM-модуля в основную память:

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

Система выделяет свободный сегмент памяти и заносит его адрес во все сегментные регистры (CS, DS, ES и SS). В первые 256 байт этого сегмента записывается PSP. Непосредственно за ним загружается содержимое COM-файла без изменений. Указатель стека (регистр SP) устанавливается на конец сегмента. В стек записывается 0000h (адрес возврата для команды ret). Управление передаётся по адресу CS:0100h, где находится первый байт исполняемого файла.

2) *Что располагается с адреса 0?*

Префикс программного сегмента.

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

Все они имеют одно и то же значение 869h (см. рис. 5). Они указывают на начало сегмента программы, то есть на начало PSP. IP имеет значение 0100h(см.

рис. 5), то есть указывает на начало программы (потому что она начинается после PSP, который занимает 256 байт).

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

Стек располагается в конце сегмента программы. Его вершина имеет адрес 0FFFEh (в начале работы программы такое значение имеет SP(см. рис. 5)). Стек растет в сторону уменьшения адресов. Адреса теоретически могут быть от 00h до 0FFFEh, но стек не должен занимать память, в которой находятся данные и код, чтобы не переписать их.

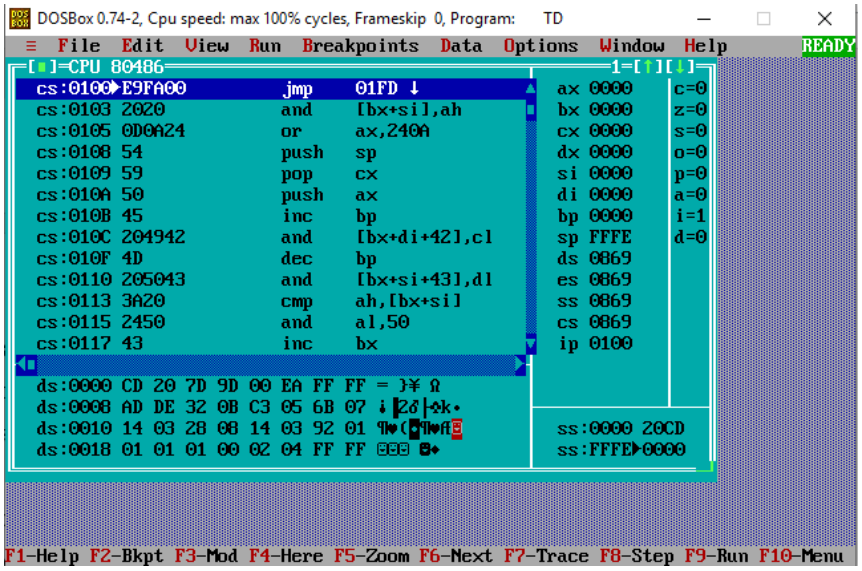

Рисунок 5 – COM-программа в отладчике

Загрузка «хорошего» EXE модуля в основную память:

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

DS и ES устанавливаются на начало PSP, SS – на начало сегмента стека, CS – на начало сегмента команд. В IP загружается смещение точки входа в программу.

2) *На что указывают регистры DS и ES?*

На начало PSP.

3) *Как определяется стек?*

В исходном коде модуля стек определяется при помощи директивы STACK. При запуске программы в SS заносится адрес сегмента стека, а в SP – адрес его вершины (см. рис. 6).

4) *Как определяется точка входа?*

Точка входа определяется по метке в ассемблерном коде, для которой указана директива END. 14-ый и 15-ый байты стандартного заголовка EXE-файла указывают на эту метку.

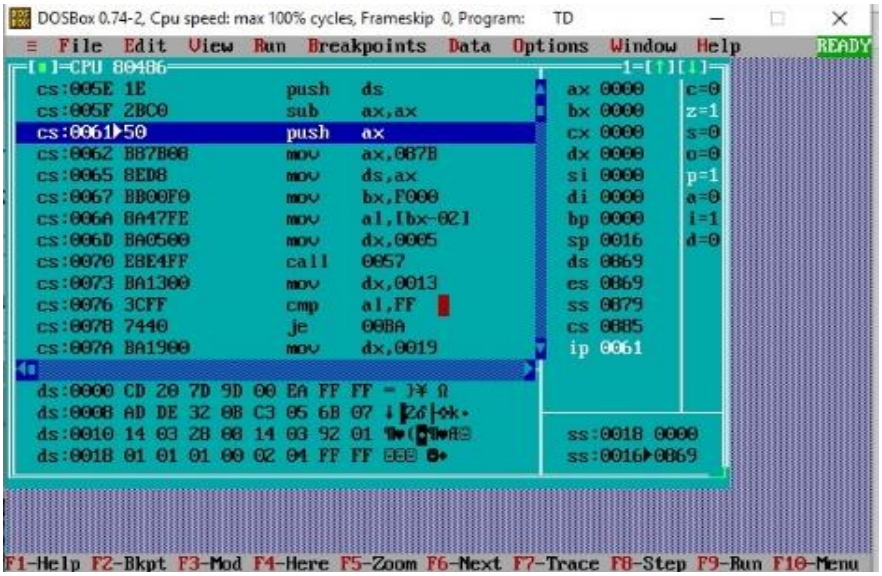

Рисунок 6 – EXE-программа в отладчике

Заключение.

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