**МАТЕМАТИЧКА ГИМНАЗИЈА**

**У БЕОГРАДУ**

**МАТУРСКИ РАД**

**из предмета**

**Рачунарства и информатике**

на тему:

Основе оперативних система

Ментор

Никола Тасић

Ученик

Борис Ћеранић

**Београд, Мај 2023.**

# **Увод**

Технолошка развића донеле су многе промене у живот просечног човека, само дигитални уређаји трансформисали су начине на које комуницирамо, баратамо информацијама, складиштимо драгоцена сећања у виду слика и бивамо продуктивни у друштву. Још од почетка друге половине 20-ог века тече непрекидан развој рачунара, та категорија се касније шири на мобилне телефоне, конзоле за игрице, паметне фрижидере... Данас је опус разноврсности дигиталних уређаја готово непојмљив једном човеку, они се увлаче у сваки аспект људског живота. Свим тим уређајима је потребно нешто да их управља, неки софтвер, преферабилно што стандардизованији могући. У супротном би завладао хаос, за сваки уређај од њих бесконачно би било потребно писати различиту верзију софтвера. Тај проблем решавају, или барем ублажавају оперативни системи.

# Оперативни систем

Оперативни систем је софтвер или програм који контролише операције компјутерског система и његових ресурса. Осим тога, постоји једна веома важна критерија која је заједничка за све оперативни системе, а то је, да је оперативни систем способан да учитава и извршава програме независно од хардвера (уређаја) на коме се налазе док им омогућава стандардизован интерфејс за улаз и излаз. Неке од главних функција које опепративни систем може имати су:

* Управљање меморијом, уређајима за складиштење и другим системским ресурсима
* Одржавање реда и спровођење сигурносних полиса другим програмима
* Извршавање више програма истовремено, одређивање редоследа и важности извршавања,

давајући им одговарајућу предност – multitasking

* Динамичко покретање и гашење других програма
* Обезбеђење основног корисничког интерфејса као и *API*-ова

Не подржава сваки оперативни систем све ове функције. *MS-DOS*, на пример, није могао да да извршава више програма од једном, био је *single-tasking*. Важно је такође напоменути шта оперативни системи нису:

* Комјутерски хардвер
* Нека специфична апликација као што су игрице, процесори текста и *Web browser*-и.
* Колекција алатки попут *GNU*
* Окружење за развој софтвера, додуше поједији, попут *UCSD Pascal*, поседују интегрисана окужења
* Графички корисники интерфејс (GUI) без обзира на то што већина оперативних система долази са једним

# Историјат

# Историјат оперативних система може се трагати до раних деценија 1950-их година, када су рачунари почели да се развијају. У том периоду, рачунари су били гломазни, скупи и сложени системи којима је било потребно управљати на ефикасан начин.

# Један од првих оперативних система био је General Motors Research Laboratories Operating System (GM-NAA I/O) који је развила компанија IBM за General Motors у 1956. години. Он је имао основне функционалности за управљање улазно-излазним операцијама и алокацију ресурса.

# Касније, у 1960. години, IBM је представио оперативни систем OS/360 за своје рачунаре. Овај систем је био значајан јер је пружао комплетну инфраструктуру за рад са рачунарима, укључујући систем за прекиде, улазно-излазне операције и меморијске управљачке функције.

# У наредним деценијама, оперативни системи су се развијали и усавршавали. У 1970-им годинама, Бел Лабс је развио оперативни систем UNIX који се и данас користи. UNIX је био отворени систем, што је омогућило његово широко распрострањење и употребу на различитим рачунарским платформама.

# Следећа значајна промена у оперативним системима дошла је са појавом персоналних рачунара у 1980-им годинама. Микрософт је представио MS-DOS, оперативни систем заснован на командној линији, који је постао широко коришћен на PC рачунарима.

# У 1990-им годинама, Микрософт је лансирао Windows оперативни систем који је имао графички кориснички интерфејс и омогућавао једноставније и интуитивније коришћење рачунара. Ово је довело до експлозије примене рачунара у различитим сферама, укључујући и домаће кориснике.

# Касније, оперативни системи су се развијали и пристосавали потребама нових технологија. Данас, оперативни системи као што су Windows, macOS и Linux пружају бројне функционалности, укључујући мрежне везе, мултимедијалне могућности и подршку за различите апликације.

# Оперативни системи су проследили пут развоја рачунара и играли кључну улогу у олакшавању коришћења и управљања рачунарима. Кроз иновације и напредак у технологији, оперативни системи настављају да се развијају и у будућности ће наставити да обликују начин на који користимо рачунаре.

# Методологија и опсег

У опсег овог рада спадају, поводом ограниченог времена, само неки аспекти оперативних система. Ради дубљег и лакшег разумевања рад ће бити обрађен из перспективе прављења једног врло елементарног оперативног система. То потенцијано укључује коришћење стандарда и пракси које су се користиле у прошлости, 90-их година прошлог века, а сада само заузимају место на старим интернет форумима или чак у музејима. Новији стандарди и праксе су много компликованији, али раде на сличном или истом принципу, чиме се тематика овог рада не нарушава.

## Boris-toolchain

Сходно

## SmallerC

Smaller C је једноставан и мали single-pass C компајлер, тренутно подржава већину фукција C језика између *C89* и *C99* стандарда. *Self-host*-ован је на оперативним системима DOS, Windows и Linux, тиме може и да се покрене на њима и да компајлује програме за њих. Погодан је за писање оперативних система у различитим модовима процесора попут *Real Mode*, *Virtual 8086 Mode*, *Unreal Mode* и *32-bit Protected mode.*

Компајлер је праћен Preprocessor-ом и Linker-ом који може да произведе програмске фајл формате попут *COM* и *MZ* који користи *DOS, PE* који користи *Windows* и *ELF* који користи *Linux* оперативни систем. Подржава и стандардну C библиотеку написану за ове оперативне системе. За правилно функционисање му је потребан екстерни асемблер. Направљен је од стране *alexfru*-а и може се видети на GitHub-у путем [линка](https://github.com/alexfru/SmallerC).

Овај компајлер је поприлично згодан за писање једноставних оперативних система тиме што подржава *Unreal Mode* и поприлично је, по својој сврси, лаган и ненабуџен. Мане су додуше то што је врло слабо документиран као и недостатак људи који га користе чиме је страна подршка и помоћ страшно лимитирана.

## *NASM*

The Netwide Assembler- асемблер и дисасемблер намењен за Intel x86 архитектуру рачунара. Користи се за писање *16-Bit*, *32-Bit* и *64-Bit* програма. Оригинално написан од стране Симона Татхама. Подржава више излазних формата као *ELF* и .bin који се често користи за писање оперативних система.

Ми ћемо га већински користити индиректно преко *SmallerC*-а, а и помало директно где нам је потребан већи степен контроле поготово при првобитном покретању оперативног система. Директно написан assembly код биће примарно 16-Bit-ни.

## QEMU

QEMU је open-source емулатор машина и виртуализатор. Може се користити више сврха али се најчешће користи за емулацију система где омогућава виртуелно окружење као читав рачунар (процесор, меморија, периферни уређаји) за покретање оперативног система. Користи се помоћу обезбеђених команди за терминал. Подржава различите архитектуре емулације као ARM, MIPS, RISC-V и наравно x86.

Ми ћемо користити x86 емулацију, примарно због приступачности и брзине откривања грешака у односу да покретање оперативног система на правој машини.

## *Make*

Make је GNU алатка која је намењена за контролу генерације извршних фајлова и осталих пропратних фајлова приликом компилације и прављења програма. Може да се користи и за покретање направљених програма. Користи такозване *Makefile*-ове да би одредио редослед баратања фајлова као и који фајлови су потребни у ком тренутку.

Један *Makefile* садржи обично више правила које изгледају овако:

target: dependencies ...

commands

...

Где су *dependencies* target-и потребни да би се извршило правило, а commands команде које ће правило покренути.

(дати пример стрвари)

# x86 Архитектура – кратак потсетник

x86 је фамилија архитектура развијена од стране Интел-а базирана на 8086 микропроцесору. 8086 је настао као *16-Bit*-на екстензија *8080* *8-Bit*-ног микропроцесора. Касније, 1985. настаје *80386* са 32-bit-ном величином регистра.

У периоду од 1999. до 2003. *AMD* развија архитектуру са величином регистра од *64-Bit*-а, она се често води под називом x86\_64 или amd64.

Ми ћемо се фокусирати на *32-Bit*-ну x86 архитектуру. Прецизније на *i386*, oва верзија базирана је на Интел-овом *80386* микропроцесору где број 3 означава трећу генерацију архитектуре.

## Регистри процесора

Процесор садржи више типова регистара. Неки регистри који су мањи од 32-Bit-а могу се такође адресирати. Они су остатак претходних верзија архитектуре, упркос томе могу често могу бити корисни.

Регистри опште намене

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 8Bit | al | bl | cl | dl | sil | dil | spl | bpl |
| 16Bit | ax | bx | cx | dx | si | di | sp | Bp |
| 32Bit | eax | ebx | ecx | edx | esi | edi | esp | ebp |
| 64Bit | rax | rbx | rcx | rdx | rsi | rdi | rsp | rbp |

EFLAGS регистар

Сегментни регистри:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| cs | Ds | ss | es | fs | gs |

Контролни регистри:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| cr0 | cr2 | cr3 | cr4 | cr8 |

Системски регистри (показивачи на табеле):

|  |  |  |
| --- | --- | --- |
| gdtr | ldtr | idtr |

Почев од 16-Bit-не x86 архитектуре, могу се адресирати и горње половине (од 9. до 16. Bit-а) регистара ax, bx, cx и dx редом под именом ah, bh, ch и dh. У наредној табели је приказана однос поменутих регистара као и које битове они адресирају:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 63-56 | 55-48 | 47-40 | 39-32 | 31-24 | 23-16 | 15-8 | 7-0 |
|  |  |  |  |  |  | ah | al |
|  |  |  |  |  |  | ax | |
|  |  |  |  | eax | | | |
| rax | | | | | | | |

У адресирању делова регистара могу помоћи такође *bitwise* логичкеоперације, на пример *ah*, *ax&0xFF00,* *eax&0x0000FF00* и *rax&0x000000000000FF00* се сви односе на исто.

### Регистри опште намене

Регистри опште намене имају улогу да чувају операнде и показиваче:

* Операнде за логичке и аритметичке операције
* Операције калкулацију адресирања
* Показиваче на локацију у меморији

Регистри опште намене се, као што назив су сугерише, могу користити поизвољно према потреби. Међутим, дизајнери хардвера су увидели могућност даљих оптимизација у томе да сваком регистри додадтно доделе неку специфичну улогу.

* *eax* – акумулатор за операнде и резултате неких операција
* *ebx* – показивач на податке у *ds* сегменту
* *ecx* – бројач петље и операције над стринговима
* *edx* – показивач на улаз и излаз
* *esi* – показивач на податке у сегменту ds регистра, изворни показивач за операције над стринговима
* *edi* – показивач на податке у сегменту es регистра, крајњи показивач за операције над стринговима
* *esp* – показивач на врх стека
* *ebp* – показивач на део стека, често одређује тренутни stack-frame

### EFLAGS регистар

EFLAGS регистар

Садржи заставице о тренутном стању процесора као и резултате неких операција

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Bit | 21 | 20 | 19 | 18 | 17 | 16 | 15 | 14 | 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |
| Име | ID | VIP | VIF | AC | VM | RF | R | NT | IOPL | | OF | DF | IF | TF | SF | ZF | R | AF | R | CF | R | CF |

* CF – Carry flag
* PF – Parity flag
* AF – Auxiliary flag
* ZF – Zero flag
* SF – Sign flag
* TF – Trap flag
* IF – Interrupt enable flag
* DF – Direction flag
* OF – Overflow flag
* IOPL – I/O privilege level flag
* NT – Nested task flag
* RF – Resume flag
* VM – Virtual 8086 mode flag
* AC – Alignment check
* VIF – Virtual interrupt flag
* VIP – Virtual interrupt pending flag
* ID – CPUID instruction flag

Битови који нису приказани или они који су обележени са R су резервисани. Наведене су само неке од неких функција застава, различите инструкције их различито користе.

### Сегментни регистри

Сегментни регистри су *16-Bit*-ни показивачи на сегмент у меморији. Да би смо радили на неком делу меморије, у сегментним ретистрима морамо да имамо одговарајући сегментни селектор(који показује на тај део меморије.

* *cs* – *code segment* садржи селектор сегмента који показује на сегмент у коме је код тј. инструкције које се извршавају
* *ds* – *data segment* садржи селектор сегмента података, сличну функцију врше и *es*, *fs* као и *gs*
* *ss* – *stack segment*, показује на сегмент у коме је тренутни стек, будући се овај регистар може мењати подржано је да програми баратају са више стекова и да мењају између њих.

## Оперативни режими процесора

Неки од оперативних режима процесора у x86 архитектуруи су:

* Real mode
* Protected mode
* Virtual 8086 mode
* Unreal mode

### Real Mode

Реални режим је једноставни 16-Bit-ни режим подржан од стране свих x86 процесора. Сваки x86 процесор се покреће у овом режиму ради компатибилности. И ако је режим у својој основи 16-Bit-ни коришћење 32-Bit-них регистара је дозвољено уз додатак одговарајућег префикса на инструкцију (0x66). Неке од предности овог режима су:

* BIOS преузима одговорност драјвера за контролу осталих уређаја и контролу прекида
* BIOS функције обезбеђују напредну колекцију *low-level API функција*
* Приступ меморији је бржи и једноставнији поводом недостатка *descriptor* таблица
* Нема сигурносних ограничења процесора на меморију и хардверски I/O

Овај режим, из перспективе конструкције оперативног система, има и мане:

* Доступно је само 1 MB меморије
* Адресирање више од 64kB меморије од једном је праћено честим променама селектора сегмента.

Овај режим ћемо користити само за покретање оперативног система, део BOOT, чиме ћемо што пре прећи у Unreal mode.

### Сегментација меморије

Сегментација меморије је решење на проблем када имамо више од 64kB меморије и 16-Bit-не регистре да њима адресирамо меморију. То се ради увођењем поменутих сегментних регистара где се у рачунању физичке адресе они множе са 16. Да би добили адресу коју којој желимо да приступимо на то треба да додамо одговарајући регистар за адресирање. То нам даје ефективно адресирање меморије од 20 Bit-ова тј. имамо приступ 2^20 = (nesto) B. На пример адресирање [ds:bx] би у облику физичке адресеизгледало као

(стави овде формулу)

Сегментацију користи реални режим и дозвољава следеће начине адресирања:

* [bx + val]
* [si + val]
* [di + val]
* [bp + val]
* [bx + si + val]
* [bx + di + val]
* [bp + si + val]
* [bp + di + val]
* [val]

Различите инструкције ће уз ове начине адресирања имплицирати различите селекторе сегмента:

mov [si], ax – ће подразумевати ds селектор сегмента тј. *ds:si*

mov es:[si], ax – екплицитно користи es селектор сегмента

cmpsb – инструкција ће поредити бајтове на *ds:si и* es:di и складиштити одговоарајући резултат у EFLAGS регистар

### Protected Mode

Заштићени режим је главни оперативи режим модерних Интел процесора још од 80286 (16-Bit-ног процесора). Максимална доступна меморија је 2^32 B тј. 4GB.

Мане из перспективе конструкције једноставног оперативног система овог режима су:

* Потреба за писање драјвера за сваки уређај који се користи
* Строга контрола меморије и I/O уређаја од стране процесора као и дељење програма по привилегијама на прстење

### Virtual 8086 Mode

Виртуелни 8086 режим је под-режим заштићеног режима где процесор емулира реални режим док је у заштићеном режиму. Сходно тиме се користи за покретање програма који подржавају само реални режим. Да би процесор био у овом режиму 17. Bit EFLAGS регистра мора да буде сетован(једнак 1).

### Unreal Mode

Нереални режим је варијанта реалног режима у коме су један или више *descpripor-*асегмента учитани са нестандардним верностима, попут 32-Bit-ни лимити који дозвољавају приступ целој меморији. Користи се у 80286 и каснијим x86 процесорима.

Компајлер који ми користимо подржава нереални режим, када SmallerC компајлује за нереални режим доста ствари попут far поинтера и far call-ова су олакшани. Из наше перспективе овај режим спаја предности реалног и заштићеног режима:

* BIOS функције стоје на располагању тако да је писање драјвера за већину једноставних ствари непотребно
* На располагању нам стоји цела меморија која може линеарно да се адресира за податке, а и код са употребом пар трикова
* Једноставније је

У нашем случају мењаћемо лимите само ds, es, fs и gs сегмената тако да буду максималне могуће веиличине тј. 4GB(2^32), док cs i ss, сегменти за код и стек, остају 64kB.

## Прекиди…

# **Пример основа оперативног система**

Оперативни систем је софтвер који решава један или више проблема, ти проблеми и потребе су већ претходно поменуте у раду. Консеквентно се тај проблем може решити на практично бесконачно много начина, опет логичким следом већина оперативних система има заједничке особине. У дањем раду биће изложен један начин решавања тог проблема, као и сличности тог решења са постојећим познатијим оперативним системима. Преглед функционисања BorisOS-а се налази на следећем графику. Корисно је поменути да је разумевање овог графика ”на прву руку” није обавезно и охрабрено је враћање на њега.

# MBR – boot

Када се рачунар покрене између осталог покреће се и BIOS. Шта BIOS тачно ради спада ван опсега овог рада, али за нас је важно да он учитава први сектор, један сектор је 512B, хард диска, USB stick-а или неког другог уређаја за складиштење. Тај сектор обично називамо *MBR***-** *Master Boot**Record*, он уобичајено садржи информације о партицијама уређаја за складиштење. Ради једноставности ми ћемо прескочити могућност више партиција на место *MBR*-a ставити само *boot sector.*

*Boot sector* се по правилу учитава на меморијску адресу 0x7C00 и предаје му се контрола.

*Boot sector* иначе само треба да садржи инструкције за дање покретање система, али по спецификацији *FAT16 file system* заузима део овог сектора, о њему ће бити више речено касније, за сада нам је само битно да по спецификацији првих пар бајтова садрже инструкцију за скакање на boot.

## Boot

По горе написаном контрола је предата boot-у. Он има следеће функције:

* Обезбеђење повољног окружења за Kernel
* Учитавање Kernel-а или његовог дела (и свега њему потребног) у меморију
* Снабдевање Kernel-а са свим информацијама потребним
* Трансфер контроле Kernel-у или делу њега

Из спецификације x86 архитектуре може се закључити да boot ради у реалном режиму процесора, тиме има приступ BIOS-овим функцијама и ресурсима што знатно олакшава његову израду.

### Обезбеђење повољног окружења за Kernel

У нашем случају прво ћемо подесити стек и одмах ући у нереалан режим. Стек ћемо ради избегавања преклапања са другим деловима оперативног система и програма ставити у други сегмент тј. у интервал [0x10000,0xFFFF]. Пошто стек у x86 архитектури расте ”на доле” почетна вредности sp и ss су 0xFFF0 (стандардна је пракса да стек буде усклађен по 4 или 16 бајтова) и 0x1000. Овиме је стек има своју максималну могућу дужину да режим коме ћемо радити, нереални.

Даље да би се пребацили у нереални режим морамо прво да се пребацимо у заштићени. Пре тога морамо учитати *Global Descriptor Table* чији опис спада ван опуса овог рада и да привремено онемогућимо прекиде. Када у *Global Descriptor Table-у* променимо лимите враћамо се у ”реални режим” који је заправо сада нереални. Горе поменуто извршава следећи код:

(стави код за ово у текстбокс)

### Учитавање Kernel-a

За правилно објашњење овог дела потребно нам да знамо начин *функционисања FAT16 file system*-а. За сада је само битно да boot тражи фајл под називом ”*LOADER.SYS*” и да га учитава на адресу 0x500.

Важно је напоменути да је 512 бајтова врло мало простора па форсира нека нестандардна, чак и занимљива решења. Пример тога био би провера да ли се тренутни фајл који проверавамо зове ”*LOADER.SYS*”:

(убаци део комичног кода овде у текстбокс )

### Снабдевање Kernel-а са свим информацијама потребним

Због ограниченог времена и једноставности оперативног система ова функција није имплементирана, али би иначе требало да садржи податке о подизању и о диску на коме се оперативни систем налази.

### Трансфер контроле Kernel-у

Ово се у нашем случају ради far call-ом са фиксним параметрима

(текстбокс кал инструкције). Где је 20 подразумевани померај од почетка фајла за извршавање MZ фајлова када је *header* празан.

Лепо је овде напоменути да су обично boot и loader спојени у bootloader који учитава цео Kernel. У нашем случају boot учитава loader који има више функција од учитавања остатка Kernel-а, што ће бити описано у даљем раду.

# Kernel

*Kernel* је централни део оперативног система, природна је последица системове потребе да управља ресурсима, код једноставнијих оперативних система овај део посебно добија на значењу. Неке од важнијих функција *Kernel-*амогу бити*:*

* Менаџмент меморије и осталих периферних уређаја
* Покреће апликације и обезбеђујући им одговарајуће окружење
* Заштита програма и приступа
* Multitasking

*Kernel*-исе отприлике могу поделити у неколико категорија:

## Монолитски *Kernel*

Монолитски *Kernel*-ифункционишу са целим *Kernel-*оми драјверима у једном меморијском адресном простору тј. *kernel-space-у.* У принципу овај приступ добија поприлично на ефикасности, јер је мењање контекста поприлично рачунски захтевна операција на x86 архитектури.

## Микро *Kernel*

Микро *Kernel* тежи да што више функција и сервиса покреће у User-space-у. Тиме се доста добија на пољу флексибилности јер *Kernel* није само један масиван програм којим је тешко баратати у смислу покретања и одржавања као и у смислу писања и debug-овања. Као директну последицу добијамо и повећану стабилност и сигурност. Мана је што у доста случајева цена мењања контекста није вредна поменутих бенефиција.

## Хибридни *Kernel*

За *Kernel* којим семи бавимо је тешко рећи ког је типа, јер је превише мали и примитиван*.* Испољава неке карактеристике Микро *Kernel-*а тиме што раније и брже гура своје функције у User-space. Са друге стране подела User-space Kernel-space је овде врло апстрактна и виртуелна и цео оперативни систем се понаша као један програм чиме га можемо назвати и монолитским. Тачно због овога постоји термин Хибридни *Kernel*

## Loader

Често су термини boot и loader спојени у bootloader као једну целину чији је задатак да, слично горе наведеном, учита цео *Kernel* претходно обезбедивши му повољно окружење. Овде ћемо узимати те појмове за одвојене, штавише као део овог решења на проблем оперативног система узећемо значење *Loader*-а у ширем смислу. Погодна је у овом тренутку *Kernel* поделити на две етапе рани *Loader* и касни *Loader* при томе треба пазити да је прелаз између ових етапа врло природан тј. не стриктно зацртан. Сходно са овим су функције раног *Loader*-а следеће:

* Системски позиви
* Учитавање остатка *Kernel*-a и оперативног система
* Учитавање системских позива као и њихово складиштење

Више о другој и трећој тачки биће објашњено у секцији о Системски позиви и Дељене библиотеке.

Функције касног *Loader*-а се примарно ослањају на функције раног, утолико да су оне практично само унапређене верзије њихових претходника:

* Покретање User-space програма
* Динамично учитавање библиотека
* Складиштење меморијских места учитаних библиотека
* Давање потребних података User-space програмима

За потпуно разумевање наведених функција неопходно је знати поглавља Системски позиви и дељене библиотеке и BEX формат.

# Системски позиви и дељене библиотеке

## Системски позиви

Системски позиви служе да се позове неки сервис *Kernel-*а из *User-space*-а, на пример да се алоцира меморија или да се учита програм. Могуће методе да се имплеметирају могу бити:

* Прекиди
* Sysenter/Sysexit (Интел)
* Софтверски, far показивачима

Најчешћи начин да се имплеметирају системски позиви су прекиди, Linux користи прекид 0x80 у те сврхе. Мана овог приступа је што често, да би се писали хендлери, делови кода који обрађују прекид, морамо да користимо асемблер. То чини процес писања знатно тежим и мање флексибилнијим, додуше бржим.

Интелови процесори од Pentium II подржавају специјалне инструкције за мењање контекста из извршавање системских позива, **sysenter** и **sysexit**. Овај метод имплеметације је доста комплексан и спада ван опсега овог рада.

Једноставан начин да се имплементирају системски позиви јесте као обичне C функције које неко може да позове ***far*** показивачем. Овај начин је додатно олакшан тиме што имамо глобални стек и тиме што немамо реалну поделу на User-space и Kernel-space.

## Дељене библиотеке

Концепт дељених библиотека тј. динамичког link-овања је поприлично једноставан и биће овде описан. Већини програма су потребне неке библиотеке, готово написан код да обавља неку функцију. Често се у пракси деси да је различитим програмима потребна једна те иста библиотека. Најједноставније решење је да се за сваки програм учита засебна инстанца те библиотеке чиме је она **статички link-ована** заједно са остаком програма.

Мало компликованије, а и аугументативно занимљивије, решење јесте да приметимо да програми могу да деле библиотеке тј. да више програма може да користи једну инстанцу једне библиотеке које је само једном учитана у меморију тек када је била потребна неком програму. Овакво решење се зове **динамичко link-овање**. Сликовити приказ меморије може овде бити од помоћи:

Статично link-овање

…

Програм 1

**…**

Програм 2

Библ.

Библ.

**…**

**…**

Динамичко link-овање

…

Програм 1

**…**

Програм 2

Библ.

**…**

Овај приступ постаје компликованији чињеницом да немамо попуну контролу над *linker*-ом пошто користимо онај који је дошао уз компајлер, а прављење скроз новог *linker-*а спада ван опсега овог рада.

## Тренутна имплементација

Тренутна имплементација ова два сервиса је суштински спојена у једно. Оба су обичне C функције које се могу позивати преко *far* показивача. То је урађено на следећи начин:

(убаци график о овоме 3 етапе који су скраћени наслови доле)

### Учитавање дељених библиотека и функција

*Loader* у овој фази, као што назив сугерише, учитава потребне податке о програму да би се он користио као библиотека. Он ће, након учитавања програма наћи мапу програма која по спецификацији мора имати исти назив као програм осим што има екстензију *.MAP* и мора се налазити у истом фолдеру као и програм. Више о фолдерима биће у секцији *File system*. Битно је да се офсети који се налазе у мапи приликом учитавања додају на апсолутну адресу где је учитана библиотека, тиме добијамо апсолутну адресу функција тј. *far* показиваче.

### Складиштење дељених библиотека и функција

Складиштење локација библиотека ради се уз помоћ два низа. Један садржи структуре које су састава:

* Име *.MAP* фајла
* Показивач на почетак одговарајућих функција
* Број функција на који се показује

Други низ само садржи показиваче на функције. Битно је напоменути да су ово низови са фиксном величином, тиме смо у датом тренутку лимитирани да пратимо **64 библиотека** и **256 функција.**

**(**убаци график овде о овоме, 2 вертикална стуба са стрелицом између**)**

### Прослеђивање програму

По BEX спецификацији која је већ описана *Loader* ће ишчитати библиотеке потребне програму. У зависности да ли је програм у *Kernel-space-у* или у *User-space*-удешава се следеће:

1. У *Kernel-space-у Loader* има приступ само сопственим примитивним функцијама учитавања фајлова, тиме може да учитава само из полазног директоријума. Ако не нађе потребне библиотеке односно програме и мапе, он ће пријавити грешку и неће учитати жељени *Kernel-space* програм
2. У *User-space*-у *Loader* већ има приступ напреднијим функцијама учитавања из *File manager*-а, сходно тиме он ће сваку поражену библиотеку потражити у **PATH** директоријуму, где је за сада **PATH** само сачувану коду *Loader-*а. Ако је нађе он ће пробати да је учита и покрене, покретање програма успева само ако су успешна учитавања свих библиотека. **Важно је напоменути да превенција застоја и кружног учитавања није имплементирана, при томе се мора водити рачуна о потребним библиотекама.**

Поводом већ поменуте лимитације коришћења готовог *linker*-а имплементација је захтевала креативна решења. Програму се прослеђује низ показивача који је састављен од показивача на низ функција из прве табеле. Приметимо да не морамо да прослеђујемо величину низа програму јер је она иста као и број тражених библиотека, који програм очигледно зна. Ово је такође низ ограничене величине чиме смо ограничени на **128 библиотека које можемо да проследимо,** то је тренутно веће од броја библиотека које тренутно можемо да пратимо али је тако због последица имплементације.

Код за формирање низа код Kernel-space програма

void\*\*\* lib\_store = &temp\_sector; //where in the temp sector to put the funcs

for(char \*lib\_search= &(mz->lib\_name); \*lib\_search; lib\_search+=16){//per BEX spec

struct shared\_lib\* sh = get\_shared\_lib(lib\_search);

if(sh==NULL){

return 2; // LIB NOT FOUND

}

\*lib\_store = sh->funs\_ptr;

lib\_store++;

}

Програм добија тако конструисан низ на стеку и статички *link*-ован код попут следећег би га дочекао:

Иницијализатор link-овања, пример

extern void start\_program();

extern void init\_MEMMNG(void\*\* funcs);

void \_\_start\_\_(void\*\*\* libs){

init\_MEMMNG(libs[0]);

start\_program();

}

Као што видимо он прослеђује header фајлу који је статички *link*-ован заједно са програмом попут овог:

Пример header фајла библиотеке

#define func\_num 3

typedef void\* (\*malloc\_ptr)(size\_t size);

malloc\_ptr malloc;

typedef void (\*dalloc\_ptr)(uint32\_t begin, size\_t size);

dalloc\_ptr dalloc;

typedef void (\*memcpy\_ptr)(void\* src, void \*dest, size\_t n);

memcpy\_ptr memcpy;

typedef void (\*print\_mem\_ptr)();

print\_mem\_ptr print\_mem;

void init\_MEMMNG(void\*\* funcs){

malloc = (malloc\_ptr)(funcs[0]);

dalloc = (dalloc\_ptr)(funcs[1]);

memcpy = (memcpy\_ptr)(funcs[2]);

print\_mem = (print\_mem\_ptr)(funcs[3]);

}

Овим приступом смо заобишли поменуте лимитације и одржали једноставност употребе библиотека приликом писања кода.

## BEX формат

BEX je формат извршних фајлова, скраћеница је за Boris EXectutable, састоји се од врло примитивне екстензије MZ извршног формата. Битно је имати у глави да ово поглавље не може бити скроз јасно док се не прочита одељак Системски позиви и дељене библиотеке.

### MZ извршни формат

МЗ овог назива представља потпис Микрософтовог инжењера Марка Збиковског. Дизајнирана је као преместиви извршни формат који би радио под реалним режимом процесора. Оригинално је направљен за MS-DOS и имао .EXE фајл екстензију, разни формати попут NE, LE и PE који користи Windows имају уграђен у себи овај формат. Спецификација структуре фајла дата је табелом:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Офсет** | | **Поље** | **Величина** | **Опис** |
| 0 | 0x00 | Signature | реч | 0x5A4D (ASCII за слова 'M' и 'Z') |
| 2 | 0x02 | Extra bytes | реч | Број бајтова у последњем сектору. |
| 4 | 0x04 | Pages | реч | Број целих сектора |
| 6 | 0x06 | Relocation items | реч | Број уноса у таблици премештања |
| 8 | 0x08 | Header size | реч | Количина параграфа (дужина 16 бајтова) који заузима header |
| 10 | 0x0A | Minimum allocation | реч | The number of paragraphs **required** by the program, excluding the PSP and program image. If no free block is big enough, the loading stops.  Број параграфа који су неопходни да се алоцирају да би програм могао да се покрене. |
| 12 | 0x0C | Maximum allocation | реч | Број параграфа које пгограм тражи да се алоцирају |
| 14 | 0x0E | Initial SS | реч | Преместива вредност ss регистра, иницјални стек програма |
| 16 | 0x10 | Initial SP | реч | Преместива вредност sp регистра, иницјални стек програма |
| 18 | 0x12 | Checksum | реч | Када се ова реч дода на све остале речи фајла, резултат треба бити нула |
| 20 | 0x14 | Initial IP | реч | Иницијална вредност ip регистра |
| 22 | 0x16 | Initial CS | реч | Иницијална вредност cs регистра |
| 24 | 0x18 | Relocation table | реч | Апсолутни офсет таблице премештања |
| 26 | 0x1A | Overlay | реч | Вредност која се користи за overlay management |
| 28 | 0x1C | Overlay information | N/A | Фајлови понекад садрже екстра информације potrebne за overlay managment |

Ради једноставности ми ћемо употребљавати само део горе наведених података, примарно за одређивање почека извршавања програма. Сходно тиме офсет почетка извршавања програма се може добити формулом: 16\*Header size + 16 \* Initial CS + Initial CS. То додато на апсолутну адресу на којој је програм учитан даје апсолутну адресу почетка извршавања.

Претходно је поменуто да је ово преместиви извршни формат. То значи да адреса на којој се програм учитава у меморију није фиксна, већ се може динамички мењати по потреби оперативног система или корисника.

У x86 архитектури да би програм могао да функционише свака инструкција која мења ток извршавања кода, попут jump, call и ret инструкција, мора да има апсолутну адресу на коју ”скаче”. То није тако једноставно када програм има могућност да се налази на различитим местима у меморији.

Решење на претходно поменути проблем су таблице премештаја( *relocation table*), оне нам показују где се све налазе инструкције чије адресе на које ”скачу” морамо да мењамо. То је обично приказано у релативном офсету у односу на почетак програма. Уобичајено је да део оперативног система који учитава програм, врло вероватно *Kernel,* пре извршавања али после учитавања програма у меморију погледа таблицу премештања и промени адресе, тј. дода апсолутну адресу почетка програма.

Ми на жалост нећемо користити таблицу премештања по MZ спецификацији, објашњење и замена дата је у следећем параграфу.

### Лимитације *linker*-а

Због одабраних алатки, посебно *linker*-а који долази уз компајлер који користимо морамо занемарити таблицу премештања коју нам нуди MZ спецификација. Уместо ње, морамо да користимо таблицу премештања везану за сам компајлер. Она је читљива само пре фазе link-овања програма, због тога ,као део BEX екстензије, link-ујемо мали део кода који чита ту таблицу и по њој, после, за време извршавања програма промени све потребне адресе.

### Формат екстензије

Екстензија садржи следеће:

* Библиотеке и системске позиве на које се програм ослања
* Код за учитавање библиотека
* Код за премештање адреса

Почетак екстензије су специфицира програму потребне библиотеке или системске позиве низом која се састоји од неког броја параграфа од 16 бајтова да би било у сагласности са MZ спецификацијом. Низ је null-terminated што значи да се чита док се на наиђе на параграф који је састављен искључиво од нула. Од тих 16 бајтова ми користимо само првих 11, они складиште стандаризовано име мапе, о којој ће касније бити речи, програма који нам служи као библиотека или део оперативног система са системским позивима.

Код за учитавање библиотека узима са стека, који је глобалан, показивач на низ локација потребних библиотека, дужина низа је позната јер је она иста као и број тражених библиотека(у супротном се програм не учитава). Чланови тог низа прослеђују различитим header фајловима одговарајућих библиотека, тиме је омогућено да су библиотеке интегрисане тако да се приликом писања C кода ни не примети да нису директно link-оване. На крају учитавања пребацује контролу почетку програма тј. функцији типа *void* означеном *start\_program*

Код за премештање адреса обавља задатак описан у делу Лимитације Linker-а. И ако је налази у меморије после кода за учитавање библиотека, извршава се први. То је логично јер за правилно функционисање било ког дела програма је потребно да све адресе за ”скакање” буду на месту. Овај код се извршава не дирајући стек, након завршетка преправки адреса пребацује контролу коду из претходног параграфа.

Потенцијално корисно је да се формат приказе кроз пример:

|  |  |
| --- | --- |
| Офсет | Садржај |
| 0x00 | MEMMNG MAP |
| 0x10 | FILEMNG MAP |
| 0x20 | Null |
| 0x30 | asm("db '</MAP NAME/>'");  asm("times 5 db 0");  asm("times 16 db 0"); //null terminated  extern void \_start\_program();  //extern void init\_</FILE NAME WITHOUT EXT/>(void\* funcs);  void \_\_start\_\_(void\*\* libs){  //call externs above  \_start\_program();  }  (горњи асемблер код је остављен примера ради али је у суштини оно што приказују поља горе) |
| … | …  .relo\_data\_loop:  cmp esi, edx  jae .relo\_data\_done  lea edi, [ebx + esi] ; edi = physical address of a relocation table element  ror edi, 4  mov ds, di  shr edi, 28  …  (ово је само део кода, читав код не би могао практично стати) |

### MAP пропратни фајл

Овај тип фајла носи екстензију .MAP и стоје уз библиотеку или било који фајл чије функције требају другом фајлу. Као што име сугерише овај фајл садржи мапу програма уз који долази, а са њим дели и име тј. њихово име се разликује само у екстензији.

Мапа програма је врло једноставна, састоји се прво од једне речи која нам означава број N мапираних функција, нису све функције мапиране, а потом наредних N речи означава офсете мапираних функција у односу на почетак програма. Редослед по коме су функције је унапред одређен од онога који конструише мапу и не мора одговарати редоследу по коме су функције написане у C фајлу.

Тренутно се овај фајл мора ручно конструисати, кратак програм који на C конструише мапу је у плану.

Формат фајла се може приказати табелом:

|  |  |
| --- | --- |
| Офсет | Садржај, пример |
| 0x00 | N |
| 0x02 | 0x1337 |
| 0x04 | 0xDEAD |
| … | … |
| N-1 | 0xBEEF |

## Менаџер меморије

Менаџмент меморије је, као што је већ поменуто, једна од најважнијих задатака *Kernel*-а. Имплементација алоцирана меморије може да се уради на разне начине:

* Bitmap-ом
* Buddy Allocation System који користи Linux
* Равним низом

Оптимална имплементација зависи од ситуације и њено одређивање захтева барем:

* Интимно познавање хардвера, брзине писања и читања меморије и уређаја за складиштење
* Статистички анализу и математичко моделовање правилности у приступању меморији
* Дубоко разумевања процеса кеширања

Што спада ван опсега овог рада, па ћемо ми остати на најједноставнијој имплементацији. Такође сложенији оперативни системи деле алоцирање меморије на више нивоа, на пример на оперативног система који даје меморију програмима и на ниво програма који раздељује меморију својим деловима која му је дата од претходног нивоа

То је низ структура које описују почетак и дужину алоцираног сегмента меморије (**сегменти поменути овде нису исти као и сегменти из одељка о сегментацији меморије**).

Тражење слободне меморије неке одређене величине се врши функцијом malloc која тражи прву слободну рупу између алоцираних сегмената које је исте или веће величине него затражена.

Функција malloc

/// @brief Allocates memory, the user is responsible for deallocating it

/// @param size the size of memory to be allocated

/// @return pointer to the location of available memory

void\* malloc(size\_t size){

for(int i = 0;memory[i].len && i < max\_memory\_sectors;i++){

size\_t cur\_available = memory[i+1].begin-memory[i].begin-memory[i].len;

if(cur\_available>size){ //we have found the available space

void\* res =(void\*) (memory[i].begin+memory[i].len);

memory[i].len+=size;

return res;

}

else if(cur\_available==size && memory[i+1].len){ //we have found the available space but we need to merge the segments

void\* res =(void\*) (memory[i].begin+memory[i].len);

memory[i].len+=size;

memory[i].len+=memory[i+1].len;

for(int j = i+1;memory[j].len & j < max\_memory\_sectors;j++){//we shift the rest of the array one to the left

memory[j] = memory[j+1];

}

return res;

}

}

return NULL; //no memory slot found

}

Деалоцирање меморије врши се функцијом dalloc (free у другим оперативним системима) која само брише или измешта алоциране сегменте

## IPC – међипроцеска комуникација

# User-space

## File system

## FAT16 file system

### Header у boot sector-у

### File Allocation Table

### Формат Directory-а

### Организација у фајлове и фолдере

## Shell

# Мотивација и закључак