Стек
Стек продолжение
Числа с плавающей точкой
Числа с плавающей точкой, векторные инструкции
Динамическая память
Строковые операции
Макросы
If Else

стек

вычитаем - увеличиваем память прибавляем к еsp уменьшаем

Задачи

- 1. Сохранить что то \ достать
- 2. Механизм передачи параметров функции

ИНСТРУКЦИЯ PUSH и POP

- 1. Принимает word и double word аргументы
- 2. синтаксис: push(pop) op1 $op1 \in m16/m32/r16/r32/imm$ уточнение при push imm (константа) всегда выделяется 4 байта
- 3. Равносильно добавить

 $\mathrm{sub}\ \mathrm{esp},\ 4$

mov [ss:esp], op1

4. забрать из стека

sub op1, [ss:esp]

add esp, 4

ещё ИНСТРУКЦИИ

pusha/popa 16 бит

pushad/popad 32 бит

Берёт все регистры общего назначения и добавляет их на вершину стека

pushf/popf 16 бит

pushfd/popfd 32 бит

Берёт регистры флагов eflags и добавляет их на вершину стека (но не все но нам это неважно)

в 32 битной системе все 16 битные инстуркции работаю также как и 32 битные

1. call метка равносильно:

```
sub esp, 4 mov ss:esp, eip jmp метка или можно передать регистр eax \to call [eax+eps:eax]
```

2. ret op1 равносильно

```
тень, [esp]; тень - теневой регистр скрытый от нас add esp, 4+{\rm op}1 јтр тень
```

Объеснение работы: Программа доходит до call, она ложит в вершину стека еір (точка возврата) выполнили функцию по метке и дошли до геt который берёт из стека еір и возвращается

call и ret обязательно добавляют и забирают память стека.

3. call foo

foo:

mov eax, [esp]; в eax будет сохранено значение eip это единственный способ его достать

ret

стековый кадр

используем регистр ebp для сохранения изначального значения esp для дальнейши и call
выполнили call на некоторую метку (функцию) имеющую вид:
push ebp
mov ebp, esp
...; расширили стек на сколько надо для локальных переменных
mov [ebp-8], eax
mov [ebp-8+4], edx
локальные переменные ближе к esp внешние параметры ниже по положительеым смещениям
от ebp
обращение к переменным что вызывали
mov eax, [ebp]
mov [eax-8], ebx; сохранили в локальные переменные вышестоящей функции значение ebx

соглашения о передаче переменных

call conventions

1. зависят от языка программмирования

мы можем рекурсивно забираться дальше пока $\mathrm{ebx} \neq 0$

- 2. описывает способ передачи параметров
- 3. описывает способ возвращения результатов
- 4. кто чистит стек от параметров
- 5. какие регистры не меняются функцией

Соглашения:

- 1. systemV ABI Unix системы
- 2. cdecl windows С программы
- 3. stdcall winAPI, Freepascal
- 4. syscall система вызова windows
- 5. fastcall linux версия(3 параметра регистр) и windows версия(2 параметра регистр) вроде
- 6. pascal turbopascal
- 7. this call C++

среди всего этого множества мы изучаем stdcall способы передачи параметров

- 1. через регистры
- 2. через стек $\begin{cases} в & \text{прямом порядке первый элемент ближе к началу памяти} \\ & \text{в обратном порядке первый элемент ближе к концу памяти} \end{cases}$

параметры в обратном порядке через стек cdecl stdcall

al - $8бит <math>\mathbb{Z}$

ax - 16бит $\mathbb Z$

 eax - 32бит $\mathbb Z$

edx:eax - 64бит \mathbb{Z}

 $\mathrm{st}\emptyset$ - с плавающей точкой $\mathbb R$

cdecl, systemV ABC: CTEK чистит тот кто вызывает функцию

stdcall, pascal: стек чистит сама функция

cdecl, stdcall: можно менять eax, ecx edx, eflags

fastcall: ecx, edx

Продолжение со стеком

имя_функции proc public (stdcall)

```
имя_функции endp
      на паскале:
1
            function sum_numbs(a, b: integer): longintege;
2
                sum_numbs := a+b
   на ассемблере:
1
            sum_numbs proc
2
                push ebp
3
                mov ebp, esp
                mov ax, [ebp+8]; взали а
4
5
                mov dx, [ebp+12]; взали b
6
                movsx eax, ax
7
                movsx edx, dx
                add eax, edx
8
9
                pop ebp
                ret 8; 2 параметра
10
                sum_numbs endp
11
12
       код Функции:
13
        .data
            a dw 3
14
            b dw 4
15
16
        .code
17
18
       mov ax, a
19
       mov bx, b
20
       push ebx
21
       push eax
22
       call sum_numbs
1
            function sum_abc(a, b: integer; var c:longint): boolean;
2
                c := a+b
```

на ассемблере:

```
1
            sum_abc proc
2
                push ebp
3
                mov ebp, esp
                mov ax, [ebp+8]; взали а
4
                mov dx, [ebp+12]; взали b
5
6
                add ax, dx
7
                mov ecx, 0
8
                jno @F; прыгнуть на ближа'шую анонимную метку ниже
9
                mov ecx, 1
                00: ; анонимнаа метка преобразуетса в ?? 0001
10
                mov edx, [ebp+16]; взали адрес С
11
12
                cwde; расширили ах
13
                adc eax, 0; дла корректности в случае переполнениа
14
                mov [edx], еах; записали по ссылке
15
                mov eax, ecx
16
                pop ebp
                ret 12; 3*4
17
18
                sum_abc endp
19
       ;код Среды:
20
       .data
21
            a dw 3
22
           b dw 4
23
           c dd?
24
       .code
25
       . . .
26
            push offset c; передадим ссылку
           mov ax, b
27
28
            push eax
29
           mov ax, a
30
            push eax
31
            call sum_abc
```

код с лекции:

print number знаковое значение 32-битное number - выводим в десятичной системе значение

Иллюстрация стека памяти для задачи 2

esp	data
esp+4	ebp
esp+8	eip
esp+10	a
esp+12	garbage
esp+14	b
esp+16	garbage
	addr c

Числа с плавающей точкой

- 1. FPU floating point unit 8087 сопроцессор (чуть инфы 1 потока)
 - позволяет вычислять 10 байтные Числа
 - ! в FPU существуеют регистры $st(\emptyset), st(1), st(2), \ldots, st(7)$, где $st(\emptyset)$ вершина стека
 - ! Каждый регистр имеет 80 бит (10 байт) размер, Freepascal реализует использование 80 битных чисел в типе extended (на C long double) на порядок выделяется 15 бит
 - * fld dword ptr mem; single precision (одинарная точность) float
 - * fld qword ptr mem; double precision (двойная точность) double
 - * fld tword ptr mem; long double|extended
 - ! помещает в вершину стека st0 байты из памяти
 - * fstp dword|qword|tword ptr mem инструкция кладёт в указнный адрес в памяти значение ${
 m st0}$
 - * По соглашению stdcall возвращаем значение в st0
 - ! Вся остальная инфа в лекциях первого потока, мы не этим пользуемся
- 2. Хронология перехода набора инструкций 3DNOW \to MMX \to SSE \to SSE2 \to SSE3 (pentium pro) \to SSE4 но с появлением AMD64 \to AVX \to AVX2 \to AVX512 Мы работает с SSE2, SSE3
- 3. SSE2
 - * есть 7 регистров $xmm0, xmm1, \ldots, xmm7$ каждый по 128 бит
 - ! Различные способы интерпретации регистров

- * Обращение может
 - (a) Скалярное $\rightarrow |_{32}|$ к блоку в мнемонике scalar

(b) Векторное обращение к всем 128 битам в мнемонике **p**acked в продолжении sse3 число регистров выросло до 16 в AVX 32 регистра уmm в AVX 512 регистры zmm 512 битности

Скалярные инстуркции

- (a) **movss** op1, op2 ; 32 битное обращение в память регистр op1 память \to xmm[0...7] или наоборот как укажешь
- (b) **movsd** op1, op2 ; 64 битное обращение в память регистр op1 память \rightarrow xmm[0...7] или наоборот как укажешь

addss	addsd
subss	subsd
mulss	mulsd
divss	divsd

ss работа в в скалярном виде с числом с одинарной точностью sd работа в скалярном виде с числом 2 точностью

- (c) comicss
- (d) comicsd

! выставляют ZF, SF, PF (PF в случае ошибки) операнд это регистр

4. Векторные инстуркции

$$movups(movupd)$$
 op1, op2; берут все 128 бит op1, op2 - (память $|xmm[0...7]$)

movaps(movapd); где a-aligned требует выравненнного стека по 16 бит работает быстрее

Пример (мнемоника инструкций по аналогии)

addps xmm0, xmm1; сложит векторно каждый блок сложит векторно

сложит 4 числа одновременно

xorps(d)

xorps xmm0, xmm0; занулит xmm0

5. Конвертация

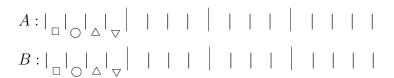
- * $\operatorname{cvt}[p,s][s,d,i]$ 2 [p,s][s,d,i] op1, op2; op1 куда, op2 что
- ! Пример cvtss2sd xmm0, xmm1;
 - ! первая буква [ps] векторно или скалярно (весь регистр или лишь один блок)
 - ! вторая буква [sdi] одинарная точность 32бита, 2 точность 64 бита, в целые числа знаковые

Продолжение чисел с плавающей точкой

```
cvtsd2si xmm, reg32\64
1
2
           cvtsd2ss xmm|mem, xmm
           cvtsi2sd mem32|64(reg32|64), xmm
3
           cvtsi2ss mem32|64(reg32|64), xmm
4
           cvtss2sd xmm|mem32, xmm
5
6
           ; замечаниа по переводу в целочисленны і:
           cvtss2si xmm|mem64, xmm; если число вылезло за диапозон то
7
           ; положит в памать 100000...
8
9
           ; после запато'і часть округлаетса к ближа'шему чётному
10
           ; см 2 бита последние
           cvttsd2si xmm|mem64, reg32|64; t-урезание
11
           ; если число вышло за диапозон допустим порадок 2^2^10
12
           ; то заполнит единицами (как максимальное)
13
14
           ; часть после запато івыкинет сразу
           cvttss2si xmm|mem32, reg32|64
15
```

команда $\operatorname{cvt}[t]$ s - скалярная $\operatorname{cvt}[t]$ p - векторное кладёт в 128 битное место очевидно в AVX инструкциях: мнемоника инструкций начинается с V - VCVTSD2SI к примеру регистры уmm

вычислить скалярное произведение



А и В выравнены в памяти по 16 то есть адрес начала внутри числа одинарной точности 32 бита мы положим в xmm0 командой **movaps** произведём умножение **mulps**

сделаем также для обоих массивов и положим в xmm0 произведение с посощью векторных инструкций

Сложить все произведения

Способ 0 - просто выгрузить в память и сложить по сути скалярно потребует много взаимодействий с памятью

Способ 1 инстурукция **haddps** - горизаонтальное сложение появилось в SSE3

$$\begin{vmatrix} \begin{vmatrix} & & \\ \Box_1 & \bigcirc_1 & \Delta_1 & \nabla_1 \end{vmatrix} & xmm0$$

$$\begin{vmatrix} & & \\ &$$

haddps xmm0, xmm1, xmm0;

$$\left| \begin{array}{c} | \\ \square_1 + \bigcirc_1 & \triangle_1 + \bigtriangledown_1 & \square_2 + \bigcirc_2 & \triangle_2 + \bigtriangledown_2 \end{array} \right| \qquad xmm1$$

Способ 2 с помоощью инстуркции $\mathbf{shuffps}$ xmm0, xmm1, imm8(байт с маской); появилось \mathbf{s} SSE2

$$\left| \begin{array}{c|c} \square & \bigcirc & \triangle & \bigtriangledown \\ 3(11) & 2(10) & 1(01) & 0(00) \end{array} \right| \qquad xmm0$$

- снизу подписаны адреса соответсвенно 11_10_01_00

второй регистр адрес назначения

инстуркция сделает перестановку по байту (поделим на 4) смотрим двойкам

для примерера 00 11 10 01

значит $11 \to 00,10 \to 11,01 \to 10,00 \to 01$ будто сдвиг по циклу влево если указать несколько позиций конечных одинаковых то поведение непредсказуемо shuffps xmm0, xmm1 00111001b

$$|\bigcirc_{10\to11}|\bigcirc_{01\to10}|\bigcirc_{00\to01}|_{01\to00}|$$
 $xmm0$

по итогу мы последовательно циклом пройдёмся сложим и не будем общаться через шину памяти много раз

* перед просмотром примера можно рассмотреть obj файл через objdump -d main.obj; что то на nm (Linux) для того чтобы разобраться с линковкой нескольких объектных файлов с модификатором Public T, нет t, также есть u - undefined

! важен порядок линковки сначала main.obj a потом multiply.obj

Работа с динамической память

код с комментариями в $\mathrm{T}\Gamma!!$ в итоговой версии конспекта будут ссылки

Строковые команды

- 1. флаг **DF** будет отвечать за сторону обхода строки(массива) (a) **cld** - сбросить DF (b) std - выставить DF Будем считать что \mathbf{DF} изначально не определён (a) **movsb** ; байты 8бит (b) **movsw** ; слова 16бит (c) **movsd** ; 2 слова 32бит 2. операции передачи строки \triangle копирует массив ds:esi в es:edi, ds, es - сегментные регистры, для **rep** (чуть ниже объяснение) в есх длину массива О операндов нет, как и все последующие инструкции работают без операндов \square использует есх, esi, edi, ds, es, eflags при условии что есх $\neq 0$ 1 mov [edi], [esi]; псевдокод 2 dec ecx 3 DF = 1: ; то есть если DF выставлен то копирует на -- адресов 4 dec esi 5 dec edi 6 DF = 0: ; то есть если DF не выставлен то копирует на ++ адресов 7 inc esi inc edi
- 3. чтобы многократно вызвать movsb|movsw|movsd нужно использовать $rep\ movsb|movsw|movsd$ пока $ecx \neq 0$ Тогда скопируется массив ECX символов согласно DF из [ESI] в [EDI]
- 4. операции сравнения строк:
 - (a) cmpsb
 - (b) cmpsw
 - (c) cmpsd

	\triangle по аналогии с movsb зависит от DF будет сравнивать попарно байты/слова/двойные слова
	\bigcirc вызов с repe повторять пока есх $\neq 0$ и пока вхождение символов равны так что в [edi] и [esi] останутся байты разные или строки идентичны
	\square вызов с repne повторять пока есх $\neq 0$ и пока вхождение символов не равны
	если строки идентичны то ${f ZF}=0$
5.	Строковые операции уступают в производительности векторным*
6.	Операции сравнения с определённым набором байтов
	(a) scasb ; образец в al
	(b) scasw ; образец в ах
	(c) scasd ; образец в еах
	○ DF аналогично, при 1 адреса на уменьшение, при 0 увеличение
	$\hfill\Box$ Взаимодействует с массивом по адресу es:edi, длина массива всё также в есх
	\triangle вызов с repe повторять пока есх $\neq 0$ и пока вхождение символа с образцом равны
	∇ вызов с repne повторять пока есх $\neq 0$ и пока вхождение символа с образцом не равны
7.	операция сохранения в регистр
	(a) lodsb
	(b) \mathbf{lodsw}
	(c) lodsd
	сохраняет содержимое [edi] в регистре al ax eax
8.	операция заполнения
	(a) stosb
	(b) stosw
	(c) stosd
	О DF логика та же

	∆ берёт из al ax eax и кладёт в [es:edi]
	□ работает с rep, заполнит длиной есх массив образцами из еах
СМОТ	реть пример в ТГ!!

Макросы

- 1. макроподстановка → 2. компиляция → 3. объектный файл
- 1. Ассесмблер
- 2. Сюда подставить код код от макроса
- 3. текст

Ассемблер прочитаем ASM файл и вставит текст из макроса причём рекурсивно то есть он и в текст дополнит макросами сначала он соберёт код полностью с учётом вставок а только потом

```
1 ml ... prog.asm
2 поавитса фа'й:
3 prog.lst
```

В этом файле хранится информация после макроподстановки Директивы

- 1. .listmacro;
- 2. .nolistmacro;

в макросах console.inc включени .notlistmacro, что значит что после подстановки в cprog>.lst

```
1 има еqu выражение; има = вырежение
2 ARR_PARM_OFF = 8+4+4; выражение = (equ) лучше пользоватьса = если это ма кропеременнаа
3 ARR_PARM_OFF[ebp]; подставит 8+4+4
4 %ARR_PARM_OFF[ebp]; подставит 16
```

Макрос синтаксис

```
1 < Mma> macro < cписок_параметров>;
2 ; < cписок_параметров> := < параметр> {, < параметр>}
```

Виды параметров:

```
1. просто <параметр> := <Имя>
```

- 2. обязательный <параметр> := <**Имя>: req**
- 3. параметр со значением по умолчанию := $\langle \mathbf{И} \mathbf{м} \mathbf{x} \rangle$:= текст

как передовать параметры:

```
1 .ВІАКА а, , 12 ; макрос с 3 параметрами 2 параметр - пустое слово|значе ние по умолчанию|Выдаст ошибку если параметр req
2 !, - экранирует запатую
3 .ВІАКА а, <большо'ітекст>, b; конструкциа в скобочках воспринимаетса ед инственным параметром - строко'і
```

Подстановка в текст

```
1 YEAR = 2025
2 "Сегодна & YEAR & год"
```

подставит 2025 как число

```
1 

1 

<
```

Конструкция досрочного выхода: exitm

метка, конструкция

```
1 :after_loop_in_macros ;;метка начинаетса с :
2 goto after_loop_in_macros
```

Метки Ассемблера внутри макроса:

```
1 local список_меток_дла_макроса;
2 ;; пример
3 local mark_err
```

```
4 mark_err:
```

название метки будет удалено сохранится лишь число ??000017 конструкция .err, дойдя до блока будет ошибка

```
    . err текст
    . . . .
    exitm
    ;; выведет текст компилатором в случае ошибки
```

конструкция есно текст

и %есhо выражение соответственно выведет значение или текст во время компиляции СМ пример в тг

Макросы продолжение

меня не было

Многомодульное программирование

a.asm - головной файл

b.asm

```
1 ...
2    add eax, var_b
3 ...
4    end
```

c.asm

- $1 \text{ var_b} \rightarrow \text{имя предоставляется}$
- 2 var_b >- имя испрользуемой в модуле

ИМЯ = символ для линковки

- 1. strong сильные имена, имя должно быть указано лишь однажды
- 2. weak слабые имена, имя можно будет переопределить

Конструкция для предоставляения:

```
1 public има переменно'і функции метки;
```

Конструкция для получения:

```
1 extern список имён
```

для переменных: имя переменной: [byte|word|...]

для функций: имя_переменной: [near|proc] - первый смотрит на соглашение

типичный пролог asm файла:

```
1 .686
2 .model flat cdecl|stdcall|...
```

при написании stdcall, cdecl ассемблер переименует названия под стандарт для паскаля ничего не пишем что имена не менялись

 $.asm \rightarrow .obj$

Секции:

- 1. .data
- 2. .data?
- 3. .code

при компиляции изначально не понять смещение offset var_a, так как всё будет слеплено в один файл после потому существуют заголовки:

имя размер где встречается ... ничерта не понял ... неебические иллюстрации не для моего уровня теха

- 1. Архитектура
- 2. Размер Секций их позиций в файле
- 3. Смещение до точки входа
- 4. Список динамических подгружаемых библиотек
- 5. Размер стека
- 6. С какого адреса размещать программу

Динамически подгружаемые данные.

таблицы этих DLL лок: GOT - Global offset table как правило переменных

PLT - f123: ...

ну тут пиздец

взаимодействие с Pascal

способы взаимодействия:

1. Ассемблерные вставки 2. Интринтики 3. Слинковать с обј Ассемблера

```
1 asm
2 ...
3 ... набор инстукци'і
4 ...
5 end ['eax', 'ecx']
```

в конце список регистров изменённых вставкой для линковки с obj файлом: {L путь_до_объектного_файла}

Макроархитектура

Теневые реистры

в процессоре существует множество регистров, скрытые от пользователя
Пример когда мы обращаемся к еах нат самом деле мы обращаемся к теневому
регистру

Блок команд

Все инструкции попадают в назначенный блок команд

Допустим представим команду так как она хранится в блоке

add byte ptr[edx], 45

преобразуется в

mov тень_i, 45

load тень $_i$, edx; тень $_i \leftarrow [edx]$

add тень $_i$, тень $_i$, тень $_i$

store edx, тень $_k$

команды load и store - самые долгие, тк работают с памятью

*работа с регистрами быстрее чем с константами

Конвеер * иллюстрация с сборкой самолёта коннвер с сцеплением

груду мусора собрать

собрали элементы

собрали частично

собрали

выкатили

дверки для дополнительных если двигатель в самолёт надо что то стороннее вставить

проблемы - данные на вход неполны ждём, зависимости - ждём кучу пузырей чем дальше переход тем хуже

разгон конвеера - время прохождения от начала и до конца

Доступ в память

load store

короче общается через иерархию кэшей

кэш комманд и кэш данных отличаются

Предсказание переходов

запихать, но если мы не угадали, то окажется что то что мы вычислили окажется никому не нужным и придётся пустить в конвеер по метке Существует 2 способа ускорения

- 1. Таблица переходов
- 2. спекулитивное выполнение

Спекулитивное выполнение возможно при нескольких АЛУ

выполнение происходит одновремнно на двух ветках (то что после jne и после метки) после сравнения подставит нужную ветку, ветка просчитывается в рамках буфера

Глубина предсказаний - число различных исходов

Таблица переходов

Откуда	куда	счётчик	счётчик неуспеха

В тот момент когда процессор доходит до нового перехода от вычисляет по некоторой функции что вытащить из таблички для нового перехода по числу срабатываний и неуспеха

Благодаря табличке мы будем чаще предсказывать что действительно произойдёт по отношению числу срабатываний и неуспехов.

Прерывания

существуют 2 вида прерывания:

- 1. Внешние прерывания вызвано внешним устройством
- 2. Внутренние вызвано схемой процессора (пример обращение к несуществующему адресу)

У каждого прерывания есть номер прерывания - всего их 256, чего очень мало так как внешних устройств много допустима ситуация, когда от 10 подключенных жётских дисков прерывания будут сливаться в один номер Прерывания и Исключения - различия

- 1. Прерывание предразумевает что ошибочная ситуация не произошла
- 2. Исключение содержит помимо номера прерывания сведения об ошибки когда случилось прерывание процессор естественный ход прерывает и выполняет обработку этого прерывания

Регистр прерываний (вектор прерываний) - 256 битный регистр, при происхождение прерывания, он заносится в вектор после обработки соответсвующий бит обнуляется обратно

изначалі	ьно

	шина	контроллер
interrupt controller ранее внешние устр	ISA	\mathbf{IC}
Pragramm I	PCI	PIC
Advanced создан с учётом работы между ядрами и про	PCI	APIC
устройство отправялет некоторое сообщение	PCI PCI express	MSI
, TD 0	TT3 /T	-

регистр флагов - IF (Interrupt Flags) IF = 0 - на прерывение не реагирует

IF = 1 - на прерывание регирует

cli sti - ставит 0 и 1 соответсвенно

Idtr Idt - таблицы

Idtr - interrupt description table register - содержит адрес таблицы и размер (регистр)

idt - interrupt description table

поля IDT:

- 1 используется или нет
- 2 адрес обработчика прерываний (виртуальный)
- 3 тип обработчика прерываний
- 4' номер сегмента кода (где лежит обработчик прерываний) в GDT
- 4" Homep TSS (task save segment) сегмента
- 5 DPL desired privilege level
- 6 размер адреса

при возникновении прерывания:

- 1. на старом стеке
- 2. на новом стеке

В зависимости от уровни привелегий процессор сам может сбрасывать IF или нет

на старом стеке: программа выполнялась дошла до N иструкции произошло прерывание

- 1. старый CS, eflags и EIP кладём на стек в случае ошибки информация о ней.
- 2. сменили сегмент на тот что в табличке (CS)
- 3. сменили EIP на тот что в табличке
- 4. CLI аппаратно

*иллюстрация

на новом стеке - всё то же самое но сохраняется SS и ESP сохраним: SS ESP CS eflags EIP *сообщение об ошибке*

также используется iret новый стек возьмём из TSS

в прерывании необходимо сохранить все регистры в случае использования TSS не надо.