# 13. Введение в программирование на языке ассемблера. Часть 2

#### Разделы:

- Система команд х86 с примерами использования
- Конвенции вызовов подпрограмм и рекурсии
- «Плавающая» арифметика
- Системные вызовы
- Работа со структурами и файлами

- Команда *loop*
- Синтаксис:
- 100р метка
- Принцип работы:
  - -уменьшить значение регистра %есх на 1;
  - -если %есх == 0, передать управление следующей за loop команде;
  - -если %есх <> 0, передать управление на метку.

```
.data
printf format:
      .string "%d\n"
.text
.globl main
main:
 > movl $0, %eax /* в %eax будет результат, поэтому в
                             начале его нужно обнулить
      movl $100, %ecx /* 100 шагов цикла
                                                           * /
sum:
      */
      loop
           sum
      /* %eax = 5050, %ecx = 0 */
 * следующий код выводит число в %еах на экран и завершает программу
      pushl %eax
      pushl $printf format
      call printf
      addl $8, %esp
      movl $0, %eax
      ret
```

```
.data
printf format:
        .string "%d\n"
array:
        .long 100, 15, 148, 12, 151, 3, 72
array end:
.text
.globl main
main:
       movl $0, %eax
                                /* в %еах будет храниться результат;
                                   в начале наибольшее значение - 0
        movl $array, %ebx
                                /* в %ebx находится адрес текущего
                                                                      * /
                                   элемента массива
```

```
loop start:
                                /* начало цикла
                                /* сравнить текущий элемент массива с
        cmpl %eax, (%ebx)
                                   текущим наибольшим значением из %еах
        jbe
                                /* если текущий элемент массива меньше
             less
                                  или равен наибольшему, пропустить
                                  следующий код
       movl (%ebx), %eax
                                /* а вот если элемент массива
                                  превосходит наибольший, значит, его
                                  значение и есть новый максимум
less:
        addl $4, %ebx
                               /* увеличить %ebx на размер одного
                                  элемента массива, 4 байта
        cmpl $array end, %ebx /* сравнить адрес текущего элемента и
                                                                     * /
                                  адрес конца массива
       jе
             loop end
                               /* если они равны, выйти из цикла
        jmp
             loop start
                               /* иначе повторить цикл снова
loop end:
  следующий код выводит число из %еах на экран и завершает программу
       pushl %eax
       pushl $printf format
       call printf
       addl $8, %esp
       movl $0, %eax
```

ret

```
#include <stdio.h>
int main()
  int array[] = \{ 100, 15, 148, 12, 151, 3, 72 \};
  int* array end = &array[sizeof(array) / sizeof(int)];
  int max = 0;
  int *p = array;
  do
   if(*p > max)
     max = *p;
  } while(++p != array end);
  printf("%d\n", max);
  return 0;
```

```
.data
printf format:
        .string "%d\n"
array:
       .long 100, 15, 148, 12, 151, 3, 72
array end:
array size:
        .long (array end - array)/4 /* количество элементов массива */
.text
.globl main
main:
       movl $0, %eax /* в %eax будет храниться результат;
                                  в начале наибольшее значение - 0 */
       movl $0, %ecx
                              /* начать просмотр с нулевого элемента
                               /* начало цикла
loop start:
                                                                    * /
       cmpl %eax, array(,%ecx,4) /* сравнить текущий элемент
                                  массива с текущим наибольшим
                                  значением из %еах
       jbe
                               /* если текущий элемент массива меньше
            less
                                  или равен наибольшему, пропустить
                                  следующий код
       movl array(,%ecx,4), %eax /* а вот если элемент массива
                                  превосходит наибольший, значит, его
                                  значение и есть новый максимум
```

```
less:
       incl %ecx
                               /* увеличить на 1 номер текущего
                                                                   * /
                                  элемента
       cmpl array size, %ecx /* сравнить номер текущего элемента с
                                  общим числом элементов
       je loop end
                           /* если они равны, выйти из цикла
                                                                   * /
                                                                   * /
            loop start
                              /* иначе повторить цикл снова
       jmp
loop end:
 * следующий код выводит число в %еах на экран и завершает программу
       pushl %eax
       pushl $printf format
       call printf
       addl $8, %esp
       movl $0, %eax
       ret
```

```
loop_start: /* начало цикла */

/* вот тут находится тело цикла */

cmpl ... /* что-то с чем-то сравнить для
принятия решения о выходе из цикла */

је loop_end /* подобрать соответствующую команду
условного перехода для выхода из
цикла */

jmp loop_start /* иначе повторить цикл снова */

loop end:
```

Цикл:

```
for(init; cond; incr)
{
  body;
}
```

• эквивалентен:

```
init;
while(cond)
{
  body;
  incr;
}
```

- *and* источник, приёмник
- *or* источник, приёмник
- *хог* источник, приёмник
- not операнд
- test операнд\_1, операнд\_2
- Сдвиги см. далее

```
testb $0b00001000, %al /* установлен ли 3-й (с нуля) бит? */
    je not_set
    /* нужные биты установлены */
not_set:
    /* биты не установлены */

testl %eax, %eax
    je is_zero
    /* %eax != 0 */
is_zero:
    /* %eax == 0 */
```

```
/* Shift Arithmetic Left/SHift logical Left */
sal/shl количество_сдвигов, назначение

/* SHift logical Right */
shr количество_сдвигов, назначение

/* Shift Arithmetic Right */
sar количество_сдвигов, назначение
```

```
До сдвига:
         | ? | | 1000100010001000100010001011
        Флаг CF Операнд
                     Сдвиг влево на 1 бит:
           1 | <-- | 0001000100010001000100010110 | <-- 0
        Флаг CF Операнд
                    Сдвиг влево на 3 бита:
        | 0 | <-- | 0100010001000100010001011000 | <-- 000
Улетели Флаг CF Операнд
в никуда
```

```
До сдвига:
         10001000100010001000100010001011 |
       Операнд
                                                Флаг CF
         Логический сдвиг вправо на 1 бит:
   --> | 0100010001000100010001000101 | -->
       Операнд
                                               Флаг CF
         Логический сдвиг вправо на 3 бита:
000 --> | 0001000100010001000100010001 | --> |
       Операнд
                                               Флаг CF Улетели
                                                       в никуда
```

```
До сдвига:
         10001000100010001000100010001011 |
       Операнд
                                                 Флаг CF
        старший бит равен 1 ==>
          ==> значение отрицательное ==>
          ==> "вдвинуть" бит 1 ---+
         Арифметический сдвиг вправо на 1 бит:
        | 1100010001000100010001000101 | --> | 1
       Операнд
                                                Флаг CF
         Арифметический сдвиг вправо на 3 бита:
111 --> | 1111000100010001000100010001 | --> | 0 |
       Операнд
                                                Флаг CF Улетели
                                                        в никуда
```

```
/* ROtate Right */
ror количество_сдвигов, назначение
/* ROtate Left */
rol количество_сдвигов, назначение
```

```
До сдвига:
                 | 10001000100010001000100010011
   Флаг CF
                 Операнд
                   Циклический сдвиг влево на 1 бит:
       | <--+-- | 0001000100010001000100010111 |
   Флаг CF V Операнд
                   Циклический сдвиг влево на 3 бита:
      0 | <--+-- | 0100010001000100010001011100 |
   Флаг СБ
               Операнд
             100
```

```
До сдвига:
         10001000100010001000100010001011
       Операнд
                                                   Флаг CF
         Циклический сдвиг вправо на 1 бит:
        11000100010001000100010001000101
       Операнд
                                                  Флаг CF
         Циклический сдвиг вправо на 3 бита:
         01110001000100010001000100010001 | ---+-->
       Операнд
                                                   Флаг CF
                                            011
```

```
/* Rotate through Carry Right */
rcr количество_сдвигов, назначение
/* Rotate through Carry Left */
rcl количество_сдвигов, назначение
```

```
До сдвига:
        | X | | 1000100010001000100010001011
       Флаг CF Операнд
                   Циклический сдвиг влево через CF на 1 бит:
   +-<- | 1 | <--- | 000100010001000100010001011X | ---+
     Флаг CF Операнд
                   Циклический сдвиг влево через CF на 3 бита:
   +-<- | 0 | <--- | 0100010001000100010001011X10 | ---+
     Флаг CF Операнд
```

```
До сдвига:
        10001000100010001000100010001011
      Операнд
                                          Флаг СЕ
        Циклический сдвиг вправо через CF на 1 бит:
      | X100010001000100010001000101 | ---> | 1 | ->-+
                                          Флаг CF V
      Операнд
             ______
        Циклический сдвиг вправо через СF на 3 бита:
       11X1000100010001000100010001 | ---> | 0 | ->-+
      Операнд
                                          Флаг CF V
```

```
int main()
{
  int a = 0x11223344;
  int shift_count = 8;

a = (a << shift_count) | (a >> (32 - shift_count));
  printf("%x\n", a);
  return 0;
}
```

### Основные конвенции вызовов подпрограмм

```
pushl $p3
pushl $p2
pushl $p1
call proc
addl $12, %esp
```

#### Рекурсии

```
.data
printf format:
        .string "%d\n"
 .text
/* int factorial(int) */
factorial:
       pushl %ebp
       movl %esp, %ebp
       /* извлечь аргумент в %еах */
       movl 8(%ebp), %eax
        /* факториал 0 равен 1 */
        cmpl $0, %eax
        jne not zero
        movl $1, %eax
        jmp return
not zero:
        /* следующие 4 строки вычисляют выражение
          %eax = factorial(%eax - 1) */
        decl %eax
        pushl %eax
        call factorial
        addl $4, %esp
       /* извлечь в %ebx аргумент и вычислить %eax = %eax * %ebx */
       movl 8(%ebp), %ebx
       mull %ebx
```

#### Рекурсии

```
результат в паре %edx:%eax, но старшие 32 бита нужно
          отбросить, так как они не помещаются в int */
return:
       movl %ebp, %esp
       popl %ebp
        ret
.globl main
main:
       pushl %ebp
       movl %esp, %ebp
       pushl $5
       call factorial
       pushl %eax
       pushl $printf format
       call printf
        /* стек можно не выравнивать, это будет сделано
          во время выполнения эпилога */
       movl $0, %eax
                                      /* завершить программу */
       movl %ebp, %esp
       popl %ebp
       ret
```

#### Рекурсии

```
/* здесь нет рекурсии! */
factorial:
       movl 4(%esp), %ecx
       cmpl $0, %ecx
       jne not zero
       movl $1, %eax
       ret
not zero:
       movl $1, %eax
loop_start:
       mull %ecx
       loop loop start
       ret
```

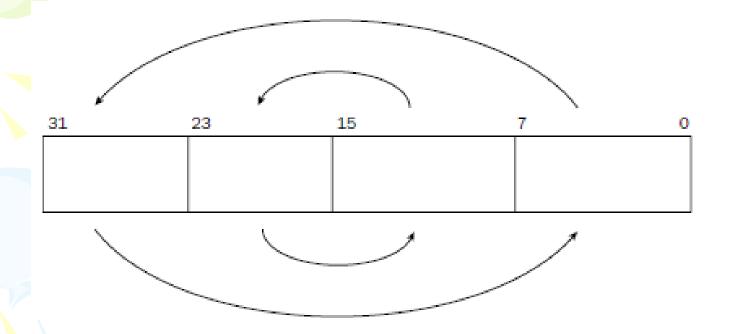
 Команда xchg производит обмен значениями двух операндов и имеет очень простой синтаксис

```
xchg операнд 1, операнд 2
```

- Один из операндов или оба должны быть регистрами
- В первом случае для доступа к другому операнду используется непосредственная адресация

```
.data
printf_format:
        .string "%d %d\n"
val1:
        .long 100
val2:
         .long 200
array:
        .rept 80
        .byte 0
         .endr
endarray:
```

```
.text
.globl main
main:
       pushl val1
       pushl val2
       pushl $printf format
        call printf
        addl $12, %esp
       movl vall, %eax
        xchq %eax, val2
        movl %eax, val1
       pushl val1
        pushl val2
       pushl $printf format
        call printf
        addl $12, %esp
       movl $0, %eax
        ret
```



 Команда bswap позволяет изменять порядок байтов в 32-битном длинном слове

movl \$0x12345678, %ebx bswap %ebx

 В ассемблере для последовательной обработки цепочек есть специализированные команды

lods

#### stos

• Все цепочечные команды подразумевают, что в регистре %esi находится указатель на следующий необработанный элемент цепочки-источника, а в регистре %edi – указатель на следующий элемент цепочки-приёмника

```
.data
printf format:
       .string "%s\n"
str in:
       .string "abc123()!@!777"
        .set str in length, .-str in
.bss
str out:
       .space str in length
.text
.globl main
main:
       pushl %ebp
       movl %esp, %ebp
       movl $str in, %esi /* цепочка-источник
       movl $str out, %edi /* цепочка-приёмник
       movl $str in length - 1, %ecx /* длина строки без нулевого
                                  байта (нулевой байт не обрабатываем)
```

lodsb /\* загрузить байт из источника в %al \*/ /\* произвести какую-то операцию с %al incb %al stosb /\* сохранить %al в приёмнике loop 1b /\* копировать нулевой байт movsb /\* важно: сейчас %edi указывает на конец цепочки-приёмника \*/ pushl \$str out pushl \$printf format call printf /\* вывести на печать movl \$0, %eax movl %ebp, %esp popl %ebp ret.

- Команда movs выполняет копирование одного элемента из цепочки-источника в цепочку-приёмник.
- Команда стрз выполняет сравнение элемента из цепочки-источника и цепочки-приёмника
- Команда scas предназначена для поиска определённого элемента в цепочке

- Нужно организовать что-то вроде цикла для обработки всей цепочки
- Для этих целей существуют префиксы команд:
  - rep
  - repe/repz
  - repne/repnz
- Префиксы ставятся перед командой, например: *repe scas*

- Команды для управления флагом df:
  - cld (CLear Direction flag) сбрасывает флаг df
  - -std (SeT Direction flag) устанавливает флаг df

```
.data
printf format:
        .string "%s\n"
str in:
        .string "abc123()!@!777"
        .set str in length, .-str in
.bss
str out:
        .space str in length
.text
/* void *my memcpy(void *dest, const void *src, size t n); */
my memcpy:
       pushl %ebp
       movl %esp, %ebp
       pushl %esi
       pushl %edi
       movl 8(%ebp), %edi /* цепочка-назначение
                                                                    * /
       movl 12(%ebp), %esi /* цепочка-источник
                                                                    */
       movl 16(%ebp), %ecx
                             /* длина
       rep movsb
       movl 8(%ebp), %eax /* вернуть dest
                                                                    * /
       popl %edi
       popl %esi
       movl %ebp, %esp
       popl %ebp
```

ret

```
.glob1 main
main:

push1 %ebp

mov1 %esp, %ebp

push1 $str_in_length

push1 $str_out

call my_memcpy

push1 $str_out

push1 $printf_format

call printf

mov1 $0, %eax

mov1 %ebp, %esp

pop1 %ebp

ret
```

```
.data
printf format:
       .string "%u\n"
str in:
       .string "abc123()!@!777"
.text
/* size t my strlen(const char *s); */
my strlen:
       pushl %ebp
       movl %esp, %ebp
       pushl %edi
       movl 8(%ebp), %edi
                                     /* цепочка */
       movl $0xffffffff, %ecx
                        /* %eax = 0 */
       xorl %eax, %eax
       repne scasb
       notl %ecx
       decl %ecx
       movl %ecx, %eax
       popl %edi
       movl %ebp, %esp
       popl %ebp
       ret
```

```
.globl main
main:
       pushl %ebp
       movl %esp, %ebp
       pushl $str in
        call my strlen
       pushl %eax
       pushl $printf format
        call printf
       movl $0, %eax
             %ebp, %esp
       movl
             %ebp
       popl
        ret
```

FPU Register	Description
Data registers	Eight 80-bit registers for floating-point data
Status register	16-bit register to report the status of the FPU
Control register	16-bit register to control the precision of the FPU
Tag register	16-bit register to describe the contents of the eight data registers
FIP register	48-bit FPU instruction pointer (FIP) points to the next FPU instruction
FDP register	48-bit FPU data pointer (FDP) points to the data in memory
Opcode register	11-bit register to hold the last instruction processed by the FPU

Data Type	Length	Significand Bits	Exponent Bits	Range
Single precision	32	24	8	1.18 x 10^-38 to 3.40 x 10^38
Double precision	64	53	11	2.23 x 10^-308 to 1.79 x 10^308
Double extended	80	64	15	3.37 x 10^-4932 to 1.18 x 10^4932

```
fld/fst source /* операнд */
.section .data
value1:
.float 12.34
value2:
   .double 2353.631
.section .bss
  .lcomm data, 8
.section .text
.globl start
start:
  nop
  flds value1
   fldl value2
   fstl data
  movl $1, %eax
   movl $0, %ebx
   int $0x80
```

Instruction	Description
FLD1	Push +1.0 into the FPU stack
FLDL2T	Push log(base 2) 10 onto the FPU stack
FLDL2E	Push log(base 2) e onto the FPU stack
FLDPI	Push the value of pi onto the FPU stack
FLDLG2	Push log(base 10) 2 onto the FPU stack
FLDLN2	Push log(base e) 2 onto the FPU stack
FLDZ	Push +0.0 onto the FPU stack

FPU Register Stack

R7	ST(0)
R6	ST(1)
R5	ST(2)
R4	ST(3)
R3	ST(4)
R2	ST(5)
R1	ST(6)
RO	ST(7)

Instruction Descrip		Description
	FADD	Floating-point addition
	FDIV	Floating-point division
	FDIVR	Reverse floating-point division
	FMUL	Floating-point multiplication
	FSUB	Floating-point subtraction
	FSUBR	Reverse floating-point subtraction

- ☐ FADD source: Add a 32- or 64-bit value from memory to the STO register
- FADD %st(x), %st(0): Add st(x) to st(0) and store the result in st(0)
- □ FADD %st(0), %st(x): Add st(0) to st(x) and store the result in st(x).
- FADDP %st(0), %st(x): Add st(0) to st(x), store the result in st(x), and pop st(0)
- □ FADDP: Add st(0) to st(1), store the result in st(1), and pop st(0)
- ☐ FIADD source: Add a 16- or 32-bit integer value to st(0) and store the result in st(0)

```
fadds data1 # add the 32-bit value at data1 to the STO register fmull data1 # multiply the 64-bit value at data1 with the STO register fidiv data1 # divide STO by the 32-bit integer value at data1 fsub %st, %st(1) # subtract the value in STO from ST1, and store in ST1 fsub %st(0), %st(1) # subtract the value in STO from ST1, and store in ST1 fsub %st(1), %st(0) #subtract the value in ST1 from STO, and store in STO
```

((43.65 / 22) + (76.34 \* 3.1)) / ((12.43 \* 6) - (140.2 / 94.21))

43.65

1.98409

3 76.34 1.98409

4 3.1 76.34 1.98409 5 236.654 76.34 1.98409

238.63809

7 12.43 238.63809

8 74.58 238.63809 9 140.2 74.58 238.63809 10 94.21 140.2 74.58 238.63809 11 1.48816 74.58 238.63809 12 13 73.09184 3.264907 74.58 238.63809

```
.section .data
value1:
  .float 43.65
value2:
   int 22
value3:
   .float 76.34
value4:
   .float 3.1
value5:
 .float 12.43
value6:
   .int 6
value7:
   .float 140.2
value8:
   .float 94.21
output:
   .asciz "The result is %f\n"
```

```
.section .text
.globl start
start:
  finit
  flds value1
  fidiv value2
  flds value3
  flds value4
  fmul %st(1), %st(0)
  fadd %st(2), %st(0)
  flds value5
  fimul value6
  flds value7
  flds value8
  fdivrp
  fsubr %st(1), %st(0)
  fdivr %st(2), %st(0)
  subl $8, %esp
  fstpl (%esp)
  pushl $output
  call printf
  add $12, %esp
  pushl $0
  call exit
```

_			
1	Instruction	Description	
	F2XM1	Computes 2 to the power of the value in ST0, minus 1	
	FABS	Computes the absolute value of the value in ST0	
	FCHS	Changes the sign of the value in ST0	
	FCOS	Computes the cosine of the value in ST0	
	FPATAN	Computes the partial arctangent of the value in ST0	
	FPREM	Computes the partial remainders from dividing the value in $ST0by$ the value in $ST1$	
	FPREM1	Computes the IEEE partial remainders from dividing the value in ST0 by the value in ST1	
	FPTAN	Computes the partial tangent of the value in ST0	
	FRNDINT	Rounds the value in ST0 to the nearest integer	
	FSCALE	Computes ST0 to the ST1st power	
	FSIN	Computes the sine of the value in ST0	
	FSINCOS	Computes both the sine and cosine of the value in ST0	
	FSQRT	Computes the square root of the value in ST0	
	FYL2X	Computes the value ST1 * log ST0 (base 2 log)	
	FYL2XP1	Computes the value ST1 * $\log$ (ST0 + 1) (base 2 $\log$ )	

Instruction	Description	
FCOM	Compare the ST0 register with the ST1 register.	
FCOM ST(x)	Compare the ST0 register with another FPU register.	
FCOM source	Compare the ST0 register with a 32- or 64-bit memory value.	
FCOMP	Compare the ST0 register with the ST1 register value and pop the stack.	
FCOMP ST(x)	Compare the ST0 register with another FPU register value and pop the stack.	
FCOMP source	Compare the ST0 register with a 32 or 64-bit memory value and pop the stack.	
FCOMPP	Compare the ST0 register with the ST1 register and pop the stack twice.	
FTST	Compare the ST0 register with the value 0.0.	

Condition	C3	C2	C0	
ST0 > source	0	0	0	
ST0 < source	0	0	1	
ST0 = source	1	0	0	:

#### Системные вызовы

- Перечислены в файле
  /usr/include/asm/unistd\_32.h
- Номер задается в регистре %еах
- Шесть параметров, передаваемых в следующем порядке через регистры:

```
– первый – в %ebx;
```

- второй в %*есх*;
- третий в %*edx*;
- четвертый в %*esi*;
- пятый в %*edi*;
- шестой в %*ebp*.

- В наших программах мы работаем с файлами следующими способами:
  - Сообщаем системе имя открываемого файла и то, в каком режиме он должен быть открыт
     читаем, пишем, читаем-пишем, создаем, если он не существует, и т.д.
  - Номер файлового дескриптора будет возвращен системным вызовом через регистр %еах
  - Далее можно оперировать с файлом, выполняя чтение и/или запись, всякий раз указывая ОС номер дескриптора
  - По окончании работы с файлами, мы должны сообщить ОС об их закрытии

.section .bss
.lcomm my\_buffer, 500

movl \$my\_buffer, %ecx
movl 500, %edx
movl 3, %eax
int \$0x80

- Как известно, при старте Linuxпрограммы открыты три файла с дескрипторами:
  - STDIN стандартный ввод
  - STDOUT стандартный вывод
  - STDERR стандартный вывод ошибок

- Программа работает с двумя файлами, читает из первого, переводит все строчные буквы в прописные и записывает результат в другой файл
- Код организован следующим образом:
  - Функция, которая принимает блок памяти и преобразует его в «верхний регистр»
  - Секция кода, которая циклически считывает данные из входного файла, размещая их в буфере, вызывает нашу функцию, и затем измененный буфер записывает в выходной файл
  - Начало функционирования программы с открытием необходимых файлов

- В нашей программе мы с помощью .equ свяжем имена с номерами
- Например, мы напишем
- .equ LINUX SYSCALL, 0x80
- Тогда всякий раз при использовании LINUX\_SYSCALL, ассемблер будет подставлять вместо него 0x80
- Можно будет написать
- int \$LINUX SYSCALL
- Это удобнее для чтения и запоминания

```
section data
/** CONSTANTS */
/* system call numbers */
.equ SYS OPEN, 5
.equ SYS WRITE, 4
.equ SYS READ, 3
.equ SYS CLOSE, 6
.equ SYS EXIT, 1
/* options for open */
.equ O RDONLY, 0
.equ O CREAT WRONLY TRUNC, 03101
/* standard file descriptors */
.equ STDIN, 0
.equ STDOUT, 1
.equ STDERR, 2
/* system call interrupt */
.equ LINUX SYSCALL, 0x80
.equ END OF FILE, 0 /* This is the return value */
                    /* of read which means we've */
                    /* hit the end of the file */
.equ NUMBER ARGUMENTS, 2
```

```
.section .bss
/** BUFFERS
* Buffer - this is where the data is loaded into
* from the data file and written from
* into the output file. This should
* never exceed 16,000 for various
* reasons.
* /
.equ BUFFER SIZE, 500
.lcomm BUFFER DATA, BUFFER SIZE
.section .text
/* STACK POSITIONS */
.equ ST SIZE RESERVE, 8
.equ ST FD IN, -4
.equ ST FD OUT, -8
.equ ST ARGC, 0 /* Number of arguments */
.equ ST ARGV 0, 4 /* Name of program */
.equ ST ARGV 1, 8 /* Input file name */
.equ ST ARGV 2, 12 /* Output file name */
```

```
.globl start
 start:
  /* save the stack pointer */
 movl %esp, %ebp
  /** Allocate space for our file descriptors
   * on the stack
   * /
  subl $ST SIZE RESERVE, %esp
open files:
open fd in:
  /* open syscall */
  movl $SYS OPEN, %eax
  /* input filename into %ebx */
  movl ST ARGV 1 (%ebp), %ebx
  /* read-only flag */
  movl $0 RDONLY, %ecx
  /* this doesn't really matter for reading */
  movl $0666, %edx
  /* call Linux */
  int $LINUX SYSCALL
```

```
store fd in:
  /* save the given file descriptor */
 movl %eax, ST FD IN(%ebp)
open fd out:
  /* open the file */
 movl $SYS OPEN, %eax
  /* output filename into %ebx */
  movl ST ARGV 2(%ebp), %ebx
  /* flags for writing to the file */
  mov1 $0 CREAT WRONLY TRUNC, %ecx
  /* mode for new file (if it's created) */
 movl $0666, %edx
 /* call Linux */
  int $LINUX SYSCALL
store fd out:
  /* store the file descriptor here */
  movl %eax, ST FD OUT(%ebp)
```

```
/** BEGIN MAIN LOOP */
read loop begin:
 /* READ IN A BLOCK FROM THE INPUT FILE */
 movl $SYS READ, %eax
  /* get the input file descriptor */
 movl ST FD IN(%ebp), %ebx
 /* the location to read into */
 movl $BUFFER DATA, %ecx
  /* the size of the buffer */
 movl $BUFFER SIZE, %edx
  /* Size of buffer read is returned in %eax */
  int $LINUX SYSCALL
  /* EXIT IF WE'VE REACHED THE END */
  /* check for end of file marker */
  cmpl $END OF FILE, %eax
  /* if found or on error, go to the end */
  jle end loop
```

```
continue read loop:
  /** CONVERT THE BLOCK TO UPPER CASE */
  pushl $BUFFER DATA /* location of buffer */
  pushl %eax /* size of the buffer */
  call convert to upper
  popl %eax /* get the size back */
  addl $4, %esp /* restore %esp */
  /** WRITE THE BLOCK OUT TO THE OUTPUT FILE */
  /* size of the buffer */
  movl %eax, %edx
  mov1 $SYS WRITE, %eax
  /* file to use */
  movl ST FD OUT(%ebp), %ebx
  /*location of the buffer */
  movl $BUFFER DATA, %ecx
  int $LINUX SYSCALL
  /** CONTINUE THE LOOP */
  jmp read loop begin
```

```
end loop:
  /** CLOSE THE FILES */
  movl $SYS CLOSE, %eax
  movl ST FD OUT(%ebp), %ebx
  int $LINUX SYSCALL
 movl $SYS CLOSE, %eax
 movl ST FD IN(%ebp), %ebx
  int $LINUX SYSCALL
  /** EXIT */
  movl $SYS EXIT, %eax
  movl $0, %ebx
  int $LINUX SYSCALL
/** This function actually does the
 * conversion to upper case for a block
 * VARIABLES:
 * %eax - beginning of buffer
* %ebx - length of buffer
 * %edi - current buffer offset
 * %cl - current byte being examined
 * (first part of %ecx)
 * CONSTANTS:
* The lower boundary of our search
.equ LOWERCASE A, 'a'
/* The upper boundary of our search */
.equ LOWERCASE Z, 'z'
```

```
/* Conversion between upper and lower case */
.equ UPPER CONVERSION, 'A' - 'a'
/* STACK STUFF */
.equ ST BUFFER LEN, 8 /* Length of buffer */
.equ ST BUFFER, 12 /* actual buffer */
convert to upper:
 pushl %ebp
 movl %esp, %ebp
 /* SET UP VARIABLES */
 movl ST BUFFER (%ebp), %eax
 movl ST BUFFER LEN(%ebp), %ebx
 movl $0, %edi
 /* if a buffer with zero length was given
  * to us, just leave
  * /
 cmpl $0, %ebx
  je end convert loop
```

```
convert loop:
  /* get the current byte */
  movb (%eax, %edi, 1), %cl
  /* go to the next byte unless it is between */
  /* 'a' and 'z' */
  cmpb $LOWERCASE A, %cl
  jl next byte
  cmpb $LOWERCASE Z, %cl
  jg next byte
  /* otherwise convert the byte to uppercase */
  addb $UPPER CONVERSION, %cl
 /* and store it back */
  movb %cl, (%eax, %edi, 1)
next byte:
 incl %edi #next byte
  cmpl %edi, %ebx #continue unless
 /* we've reached the end */
  jne convert loop
end convert loop:
  /* no return value, just leave 8/
  movl %ebp, %esp
  popl %ebp
  ret
```

# Работа со структурами данных

- Мы попробуем написать программу для манипуляции простыми структурами фиксированной длины
- Предположим, мы хотим сохранять информацию о знакомых нам людях
- Мы могли бы представить следующий пример структуры фиксированной длины:
  - Имя 40 байт
  - Фамилия 40 байт
  - Адрес 240 байт
  - Возраст 4 байта

# Работа со структурами данных

```
/* record-def.s */
.equ RECORD_FIRSTNAME, 0
.equ RECORD_LASTNAME, 40
.equ RECORD_ADDRESS, 80
.equ RECORD_AGE, 320
.equ RECORD_SIZE, 324
```

# Работа со структурами данных

```
/* linux.s */
#Common Linux Definitions
#System Call Numbers
equ SYS EXIT, 1
.equ SYS READ, 3
.equ SYS WRITE, 4
.equ SYS OPEN, 5
.equ SYS CLOSE, 6
.equ SYS BRK, 45
#System Call Interrupt Number
equ LINUX SYSCALL, 0x80
#Standard File Descriptors
.equ STDIN, 0
.equ STDOUT, 1
.equ STDERR, 2
#Common Status Codes
.equ END OF FILE, 0
```

```
/* read-record.s */
.include "record-def.s"
.include "linux.s"
#PURPOSE: This function reads a record from the file
# descriptor
#INPUT: The file descriptor and a buffer
#
#OUTPUT: This function writes the data to the buffer
# and returns a status code.
#STACK LOCAL VARIABLES
.equ ST READ BUFFER, 8
.equ ST FILEDES, 12
```

```
.section .text
globl read record
.type read record, @function
read record:
pushl %ebp
mov1 %esp, %ebp
pushl %ebx
movl ST FILEDES (%ebp), %ebx
mov1 ST READ BUFFER (%ebp), %ecx
movl $RECORD SIZE, %edx
movl $SYS READ, %eax
int $LINUX SYSCALL
#NOTE - %eax has the return value, which we will
# give back to our calling program
popl %ebx
movl %ebp, %esp
popl %ebp
ret
```

```
/* write-record.s */
.include "linux.s"
.include "record-def.s"
#PURPOSE: This function writes a record to
# the given file descriptor
#INPUT: The file descriptor and a buffer
#OUTPUT: This function produces a status code
#STACK LOCAL VARIABLES
.equ ST WRITE BUFFER, 8
.equ ST FILEDES, 12
.section .text
.globl write record
.type write record, @function
```

```
write record:
pushl %ebp
movl %esp, %ebp
pushl %ebx
movl $SYS WRITE, %eax
movl ST FILEDES (%ebp), %ebx
movl ST WRITE BUFFER(%ebp), %ecx
movl $RECORD SIZE, %edx
int $LINUX SYSCALL
#NOTE - %eax has the return value, which we will
# give back to our calling program
popl %ebx
movl %ebp, %esp
popl %ebp
ret
```

```
/* write-records.s */
.include "linux.s"
.include "record-def.s"
.section .data
#Constant data of the records we want to write
#Each text data item is padded to the proper
#length with null (i.e. 0) bytes.
#.rept is used to pad each item. .rept tells
#the assembler to repeat the section between
#.rept and .endr the number of times specified.
#This is used in this program to add extra null
#characters at the end of each field to fill
#it up
record1:
.ascii "Fredrick\0"
.rept 31 #Padding to 40 bytes
.byte 0
.endr
.ascii "Bartlett\0"
.rept 31 #Padding to 40 bytes
.bvte 0
.endr
.ascii "4242 S Prairie\nTulsa, OK 55555\0"
.rept 209 #Padding to 240 bytes
.byte 0
.endr
.long 45
```

```
record2:
.ascii "Marilyn\0"
.rept 32 #Padding to 40 bytes
.byte 0
.endr
.ascii "Taylor\0"
.rept 33 #Padding to 40 bytes
.bvte 0
.endr
.ascii "2224 S Johannan St\nChicago, IL 12345\0"
.rept 203 #Padding to 240 bytes
.byte 0
.endr
.long 29
record3:
.ascii "Derrick\0"
.rept 32 #Padding to 40 bytes
.byte 0
.endr
.ascii "McIntire\0"
.rept 31 #Padding to 40 bytes
.byte 0
.endr
```

```
.ascii "500 W Oakland\nSan Diego, CA 54321\0"
.rept 206 #Padding to 240 bytes
.byte 0
.endr
.long 36
#This is the name of the file we will write to
file name:
.ascii "test.dat\0"
.equ ST FILE DESCRIPTOR, -4
.globl start
start:
#Copy the stack pointer to %ebp
movl %esp, %ebp
#Allocate space to hold the file descriptor
subl $4, %esp
#Open the file
movl $SYS OPEN, %eax
movl $file name, %ebx
movl $0101, %ecx #This says to create if it
#doesn't exist, and open for
#writing
movl $0666, %edx
int $LINUX SYSCALL
```

```
#Store the file descriptor away
movl %eax, ST FILE DESCRIPTOR(%ebp)
#Write the first record
pushl ST FILE DESCRIPTOR(%ebp)
pushl $record1
call write record
addl $8, %esp
#Write the second record
pushl ST FILE DESCRIPTOR(%ebp)
pushl $record2
call write record
addl $8, %esp
#Write the third record
pushl ST FILE DESCRIPTOR(%ebp)
pushl $record3
call write record
addl $8, %esp
#Close the file descriptor
movl $SYS CLOSE, %eax
movl ST FILE DESCRIPTOR (%ebp), %ebx
int $LINUX SYSCALL
#Exit the program
movl $SYS EXIT, %eax
movl $0, %ebx
int $LINUX SYSCALL
```

```
/* count-chars.s */
#PURPOSE: Count the characters until a null byte is reached.
#INPUT: The address of the character string
#OUTPUT: Returns the count in %eax
#PROCESS:
# Registers used:
# %ecx - character count
# %al - current character
# %edx - current character address
.type count chars, @function
.globl count chars
#This is where our one parameter is on the stack
.equ ST STRING START ADDRESS, 8
count chars:
pushl %ebp
movl %esp, %ebp
#Counter starts at zero
movl $0, %ecx
#Starting address of data
movl ST STRING START ADDRESS (%ebp), %edx
```

```
count loop begin:
#Grab the current character
movb (%edx), %al
#Is it null?
cmpb $0, %al
#If yes, we're done
je count loop end
#Otherwise, increment the counter and the pointer
incl %ecx
incl %edx
#Go back to the beginning of the loop
jmp count loop begin
count loop end:
#We're done. Move the count into %eax
#and return.
movl %ecx, %eax
popl %ebp
ret
```

```
/* write-newline.s */
.include "linux.s"
.globl write newline
.type write newline, @function
.section .data
newline:
.ascii "\n"
.section .text
.equ ST FILEDES, 8
write newline:
pushl %ebp
movl %esp, %ebp
movl $SYS WRITE, %eax
movl ST FILEDES (%ebp), %ebx
movl $newline, %ecx
movl $1, %edx
int $LINUX SYSCALL
movl %ebp, %esp
popl %ebp
ret
```

```
/* read-records.s */
.include "linux.s"
.include "record-def.s"
.section .data
file name:
.ascii "test.dat\0"
.section .bss
.lcomm record buffer, RECORD SIZE
.section .text
#Main program
.globl start
start:
#These are the locations on the stack where
#we will store the input and output descriptors
#(FYI - we could have used memory addresses in
#a .data section instead)
.equ ST INPUT DESCRIPTOR, -4
.equ ST OUTPUT DESCRIPTOR, -8
#Copy the stack pointer to %ebp
movl %esp, %ebp
#Allocate space to hold the file descriptors
subl $8, %esp
```

```
#Open the file
movl $SYS OPEN, %eax
movl $file name, %ebx
movl $0, %ecx #This says to open read-only
movl $0666, %edx
int $LINUX SYSCALL
#Save file descriptor
movl %eax, ST INPUT DESCRIPTOR(%ebp)
#Even though it's a constant, we are
#saving the output file descriptor in
#a local variable so that if we later
#decide that it isn't always going to
#be STDOUT, we can change it easily.
movl $STDOUT, ST OUTPUT DESCRIPTOR(%ebp)
record read loop:
pushl ST INPUT DESCRIPTOR(%ebp)
pushl $record buffer
call read record
addl $8, %esp
```

```
#Returns the number of bytes read.
#If it isn't the same number we
#requested, then it's either an
#end-of-file, or an error, so we're
#quitting
cmpl $RECORD SIZE, %eax
ine finished reading
#Otherwise, print out the first name
#but first, we must know it's size
pushl $RECORD FIRSTNAME + record buffer
call count chars
addl $4, %esp
movl %eax, %edx
movl ST OUTPUT DESCRIPTOR (%ebp), %ebx
movl $SYS WRITE, %eax
movl $RECORD FIRSTNAME + record buffer, %ecx
int $LINUX SYSCALL
pushl ST OUTPUT DESCRIPTOR(%ebp)
call write newline
addl $4, %esp
jmp record read loop
finished reading:
movl $SYS EXIT, %eax
movl $0, %ebx
int $LINUX SYSCALL
```

```
/* add-year.s */
.include "linux.s"
.include "record-def.s"
.section .data
input file name:
.ascii "test.dat\0"
output file name:
.ascii "testout.dat\0"
.section .bss
.lcomm record buffer, RECORD SIZE
#Stack offsets of local variables
.equ ST INPUT DESCRIPTOR, -4
.equ ST OUTPUT DESCRIPTOR, -8
.section .text
.globl start
start:
#Copy stack pointer and make room for local variables
movl %esp, %ebp
subl $8, %esp
```

```
#Open file for reading
movl $SYS OPEN, %eax
movl $input file name, %ebx
movl $0, %ecx
movl $0666, %edx
int $LINUX SYSCALL
movl %eax, ST INPUT DESCRIPTOR(%ebp)
#Open file for writing
movl $SYS OPEN, %eax
movl $output file name, %ebx
movl $0101, %ecx
movl $0666, %edx
int $LINUX SYSCALL
movl %eax, ST OUTPUT DESCRIPTOR(%ebp)
loop begin:
pushl ST INPUT DESCRIPTOR(%ebp)
pushl $record buffer
call read record
addl $8, %esp
```

```
#Returns the number of bytes read.
#If it isn't the same number we
#requested, then it's either an
#end-of-file, or an error, so we're
#quitting
cmpl $RECORD SIZE, %eax
jne loop end
#Increment the age
incl record buffer + RECORD AGE
#Write the record out.
pushl ST OUTPUT DESCRIPTOR(%ebp)
pushl $record buffer
call write record
addl $8, %esp
jmp loop begin
loop end:
movl $SYS EXIT, %eax
movl $0, %ebx
int $LINUX SYSCALL
```

#### See also

- Магда, Ю.С. Ассемблер для процессоров Intel Pentium/ Ю.С. Магда. – СПб.: Питер, 2006. – 416 с.
- Робачевский, А. Операционная система Unix, 2 изд./
   А.Робачевский, С.Немнюгин, О.Стесик. СПб.: БХВ-Петербург, 2010. 656 с.
- Столяров, А.В. Программирование на языке ассемблера NASM для ОС UNIX: учеб.пособие. М.: Макс, 2011. 188 с. Доступ: http://www.stolyarov.info/books/asm\_unix
- flat assembler http://www.flatassembler.net/
- The Netwide Assembler: NASM http://nasm.us/
- Intel® 64 and IA-32 Architectures Optimization Reference Manual, 3.5.1.3 Using LEA.
- Intel® 64 and IA-32 Architectures Software Developer's Manual, 4.1 Instructions (N-Z), PUSH
- Ассемблер в Linux для программистов С –
   http://ru.wikibooks.org/wiki/Ассемблер\_в\_Linux\_для\_програ