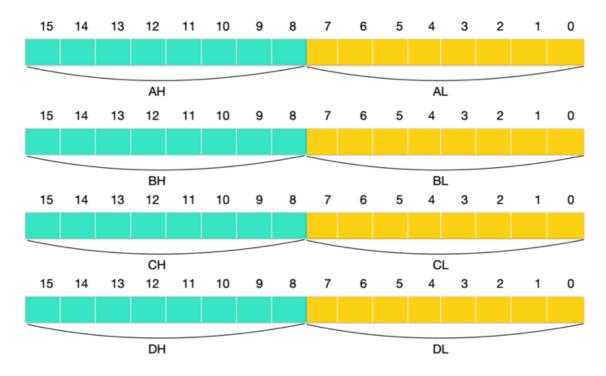
王爽汇编第二章,寄存器

- 1. 寄存器知识
- 2. 通用寄存器
- 2.1. AX寄存器(累加寄存器)
- 2.2. BX寄存器(寻址寄存器)
- 2.3. CX寄存器(计数寄存器)
- 2.4. DX寄存器
- 3. 段寄存器
- 4. 偏移(索引)寄存器
- 4.1. BP基础指针寄存器
- 4.2. SP栈指针寄存器
- 4.3. SI变址寄存器
- 4.4. DI目标变址寄存器
- 5. IP寄存器
- 6. 标志寄存器

1. 寄存器知识

一个典型的CPU由运算器、控制器、寄存器 (CPU工作原理) 等器件构成,这些器件靠内部总线相连。

- 运算器进行信息处理(运算单元)
- 寄存器进行信息存储(存储单元)
- 控制器负责控制各种器件工作(控制单元)



几乎所有的冯·诺伊曼型计算机的 CPU,其工作都可以分为5个阶段: **取指令、指令译码、执行指令、访存取数、结果写回。**

1. 取指令: 取出内存中OPCode字节码。

2. 指令译码:将OPCode字节码翻译成汇编指令。

3. 指令执行: 执行汇编指令。

4. 访存取数: 根据汇编指令需要,可能会访问内存数据,根据地址码得到内存位置并读取该操作数。

5. 结果写回: 运行结果数据写回到 CPU 的内部寄存器, 方便后续指令快速存取。

8086CPU中一共有14 个寄存器,所有寄存器都是16位宽,存2个字节,因为是完全的16位微处理器。

AX, BX, CX, DX, SP, BP, SI, DI, IP, FLAG, CS, DS, SS, ES

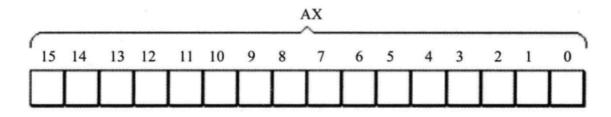


图 2.1 16 位寄存器的逻辑结构

这 14 个寄存器有可能进行具体的划分,按照功能可以分为五种:

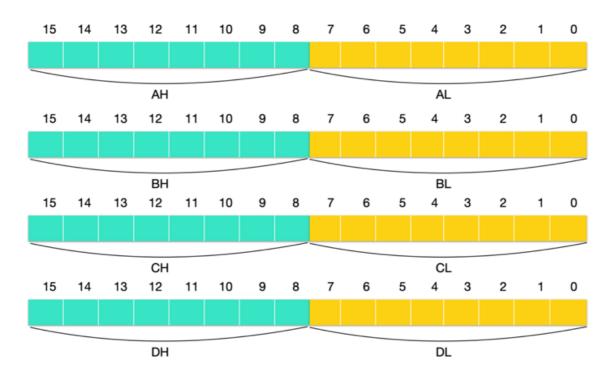
- 通用寄存器
- 段寄存器
- 偏移寄存器
- IP寄存器
- 标志寄存器

2. 通用寄存器

通用寄存器有4个: AX, BX, CX, DX, 一般用来存放数据, 也被称为数据寄存器。

它们可分为两个可独立使用的8位寄存器

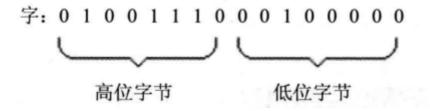
如下图所示。



8086CPU可以一次性处理以下两种尺寸的数据。

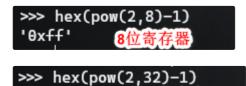
• 字节: 记为byte, 一个字节由8个bit组成, 可以存在8位寄存器中。

• 字:记为word,一个字由两个字节组成,可以存在一个16位寄存器中(16位CPU)

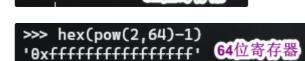


8086采用小端模式: 高地址存放高位字节, 低地址存放低位字节。

一个8位寄存器所能存储的数据范围是: 0~255 (2的8次方-1)



'0xffffffff'



2.1. AX寄存器(累加寄存器)

AX也叫做累加寄存器,主要用于输入/输出大规模的指令运算。

以下例子是我在Windows下编译的32位汇编,其中AX扩展成了EAX,我们可以看见乘法(累加)后的结果他其实就是保存到了EAX中,然后再将EAX赋值给整型变量c,所以<mark>乘法(imul)</mark>和 除法(div) 运算都会用到AX寄存器,除法会将商保存到ax中,余数保存到dx中。

编译命令: cl -FAS .\1.cpp

```
= -8
a$ = -4
main
       PROC
       push
       mov
               ebp, esp
                                                       ; 000000cH
       sub
               esp, 12
                                        ①
                                        变量a
               DWORD PTR _a$[ebp], 2
               DWORD PTR _b$[ebp],
       mov
                                        (3)
              eax, DWORD PTR _as[ebp]
       mov
       imul eax, DWORD PTR _b$[ebp]
               DWORD PTR _c$[ebp], eax
       mov
            return Θ;
                            eax作为累加结果 传给变量C
               eax, eax
               esp, ebp
       mov
               ebp
       pop
       ret
_main
       ENDP
TEXT
       ENDS
END
```

可以利用文章 https://www.cnblogs.com/VxerLee/p/15264290.html 中的方法将C语言转换成DOS汇编代码,然后对比源码和汇编代码发现乘法结果确实会放入到ax中。

乘法:

```
HELLOWOR.C (C:\tools\TC2.0 for dosbox\tc) - GVIM
     _main
        push bp
                                  文件(F) 编辑(E) 工具(T) 语法(S) 缓冲区(B) 窗口(W) 帮助
        mov bp,sp
                                 스 🖺 😭 🕰 (월 lai 🗈 X la) 영 l 🖺 🕰 lai (
        sub sp,2
                                   1 Winclude (stdio.h)
                                   3 int main()
29
                                   4 {
        mov si,2
                                             int a,b,c;
                                   6
                                             a = 2;
    mov di,3
                                             b = 3;
                                             c = a * b;
                                             printf("Hello,World!\r\n");
        mul di
                                  10
                                             return 0;
        mov word ptr [bp-2],ax
                                  11 }^2
        push
        mov ax, offset DGROUP:s@
                far ptr _printf
        call
```

除法(取余数):

```
b@ label
16 v b@w label
                word
                                               DIV.C (C:\tools\TC2.0 for dosbox\tc) - GVIM
        ?debug C E9CE592E53056469762E63
                                              文件(F) 编辑(E) 工具(T) 语法(S) 缓冲区(B) 窗口(W) 帮助(H)
         ?debug C E92A6C2D53112E2F696E636C75646
                                              | 🖰 🖫 🕞 🖺 | 9 @ | X 🗈 🛍 | 🗞 🕰 | 🚣 ;
        ?debug C E900501D1110696E636C7564652F7
     _BSS ends
                                                1 #include <stdio.h>
                segment byte public 'CODE'
    DIV_TEXT
                                               2 int main()
                                               3 {
  ∨_main prod
push be
                   far
                                               4
                                                          int a;
                                               5
                                                          int b;
                                                          int c;
                                               6
                                               7
                                                          a = 12;
                                                         b = 4;
                                               8
                                                         c = a 🎖 b;
                                               9
                                                          printf("%d\n",c);
                                               10
                                               11 }
        mov si,12
        mov word ptr [bp-2],dx
```

ax还常常被用作函数的返回值。

```
DOSBox 0.72, Cpu Cycles: 3000, Framesk
         jmp short @1
                                                     File Edit Run Compile
    01:
                                                         Line 1
                                                                    Col 1
                                                                            Insert In
                                                   #include <stdio.h>
         pop si
        pop bp
                                                 int div(int a,int b)
    _div endp
                                                       int c:
                                                       c = a \times b;
    _main proc
                                                      return c;
      push bp
mov bp,sp
                                                   }
                                                   int main()
       push si
push di
?debug L 15
                                                       int a:
                                                       int b;
                                                       int c:
                                                       a = 12;
                                                       b = 4;
                                                       c = div(a,b);
        mov di,4
                                                       printf("xd\n",c);
                                                  F1-Help F5-Zoom F6-Switch F7-Tr
56
       call→ near ptr _div
       рор сх
                                    函数返回值
        mov word ptr [bp-2],ax
       push word ptr [bp-2]
        mov ax, offset DGROUP:s@
                far ptr _printf
```

2.2. BX寄存器(寻址寄存器)

bx叫做数据寄存器,用来暂存一般数据,不过用的最多的还是用来寻址,bx存放偏移地址,然后根据基地址+偏移进行物理内存地址的定位。

暂存一般数据:

```
1 mov bx,2
2 add bx,bx
3 mov ax,bx
```

用bx寻址:

```
_DATA
BX.C (C:\tools\TC2.0 for dosbox\tc) - GVIM
                                                                           文件(F) 编辑(E) 工具(T) 语法(S) 缓冲区(B) 窗口(
                                                                           스 및 및 를 을 © | X 🖹 📵 🖁 🕏 🕏
                                                                            1 #include <stdio.h>
                                                                            2 int main()
                                                                            3 {
BSS
BX_TEXT segment byte public 'CODE'
                                                                                      int arry[18];
                                                                                      int a,b,c,i
_main prod
push b
                                                                                      for(i=1;i<=10;i++)
             far
                                                                                      {
    mov bp,sp
                                                                            8
                                                                                              arry[i] = i;
    sub sp,24
                                                                           11
                                                                                      b = 2;
                                                                                      c = 6 / 2;
                                                                           12
                                                                                      arry[<mark>5</mark>] = c;
                                                                           13
    mov si,1 ;s
jmp short @5
                                                                                      printf("%d\r\n",arry[5]);
                                                                           14
                                                                           15 }^Z
    lea ax, word ptr [bp-24]; 取临时变量地址给ax add bx,ax;原来如此,这里bx用来做数组下标,因为在8086中最大16位,那么xxx+2表示下一位值
   mov word ptr ss:[bx],si;array[i] = i
                                            用bx表示下
@5:
   cmp si,10 ;i<=10
```

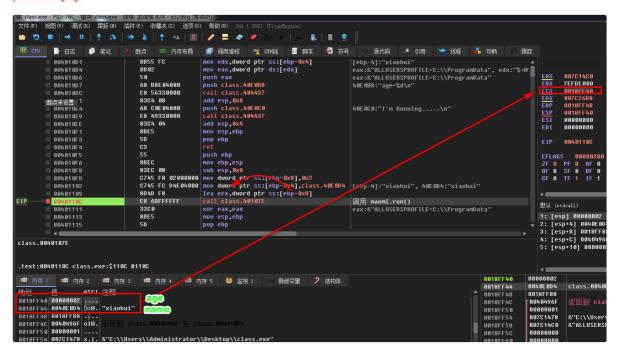
2.3. CX寄存器(计数寄存器)

cx也是叫做数据寄存器,也能暂存一般数据,不过cx还有个专门的用途,根据字面意思c - > "Count",用来在Loop(循环)时候,用cx寄存器来进行计数指定循环的次数。

```
1 mov cx,10 ;循环10次
2 xor ax,ax
3 xor bx,bx ;存放sum
5 int ax ;
6 add bx,ax;/bx = 1+2+3+4+5+6+7+8+9+10
7 loop s
```

除此之外在Win32汇编中,ecx经常会作为类指针进行传入。

进行反汇编后可以发现代码调用 maomi.run() 的时候会把 this指针 传入 ecx 寄存器。 lea ecx,dword ptr ss:[x]

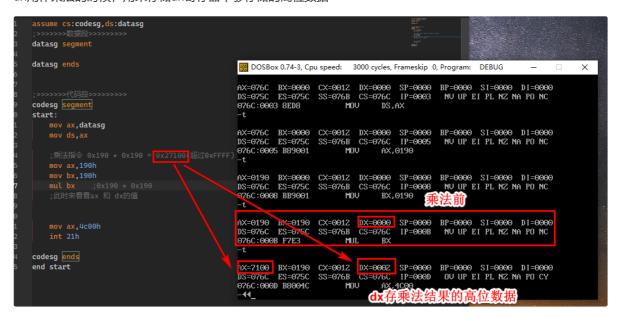


```
this$ = -4
     ?run@Cat@@QAEXXZ PROC NEAR
                                              ; Cat::run
359
         push
          mov ebp, esp
                                            ecx 给 this指针
         push
          mov DWORD PTR _this$[ebp], ecx
364
                  cout << "I'm a Cat." << endl;</pre>
365
         push
                OFFSET FLAT:?endl@std@@YAAAV?$basic_ostream@DU?$char_traits@D@std@@@1@AAV21@@Z ; std::endl
         push
367
                 OFFSET FLAT:$SG7428
                 OFFSET FLAT:?cout@std@@3V?$basic_ostream@DU?$char_traits@D@std@@@1@A ; std::cout
         push
         call
                 ??6std@@YAAAV?$basic_ostream@DU?$char_traits@D@std@@@@AAV10@PBD@Z ; std::operator<<</p>
         add esp, 8
         mov ecx, eax
                 ??6?$basic_ostream@DU?$char_traits@D@std@@@std@@QAEAAV01@P6AAAV01@AAV01@@Z@Z ; std::basic_ostrea
         call.
     : 17 :
                 cout << "My Name is " << this->name << endl;
374
375
                 OFFSET_FLAT:?endl@std@@YAAAV?$basic_ostream@DU?$char_traits@D@std@@@1@AAV21@@Z ; std::endl
        mov eax, DWORD PTR _this$[ebp]
mov ecx, DWORD PTR [eax+4]
                                            this+4 = this->name
         push
                 OFFSET FLAT: $SG7521
                 {\tt OFFSET\ FLAT:?} cout@std@@3V?\$basic\_ostream@DU?\$char\_traits@D@std@@@1@A \textit{; } std::cout@std@partialseam. } \\
         push
                 ??6std@@YAAAV?$basic_ostream@DU?$char_traits@D@std@@@@AAV10@PBD@Z ; std::operator<</pre>
          call.
         add esp, 8
                eax
??6std@@YAAAV?$basic_ostream@DU?$char_traits@D@std@@@0@AAV10@PBD@Z ; std::operator<<
         push
         call
         add esp, 8
         mov ecx, eax
                 ??6?$basic_ostream@DU?$char_traits@D@std@@@AEAAV01@P6AAAV01@AAV01@@Z@Z ; std::basic_ostrea
388
         call
     ; 18 : printf("age=%d\n",this->age);
390
         mov edx, DWORD PTR _this$[ebp]
                                           this+0 = this->age
         mov eax, DWORD PTR [edx]
                 OFFSET FLAT:$SG7522
         push
         call _printf
add esp, 8
396
397
                 printf("I'm Running....\n");
399
400
                OFFSET FLAT:$SG7523
401
         push
         call _printf
add esp, 4
402
403
      ; 20 : }
```

2.4. DX寄存器

dx寄存器和之前的寄存器一样,都是能暂存一般数据,在之前的汇编代码中可以发现编译器会将(通用寄存器)和si、di都拿来暂存一般的数据,除此之外之前还介绍过乘法和除法的值也会写入dx中。

dx用作乘法的时候,用来存储ax寄存器不够存储的高位数据



3. 段寄存器

CPU中包含了四个段寄存器,用作程序指令,数据或栈的基础位置。ps:(不过在Windows中好像这些段寄存器没什么用,因为Windows用了平坦模式,每个段的地址都一样都是0,直接用偏移来定位)

段寄存器主要的功能如下:

CS(Code Segment) : 代码寄存器,程序代码的基础位置
 DS(Data Segment) : 数据寄存器,变量的基本位置
 SS(Stack Segment) : 栈寄存器,栈的基础位置

• ES(Extra Segment): 其他寄存器,内存中变量的其他基本位置。

```
1  mov cs:[xxx],ax
2  mov ds:[xxx],ax
3  mov ss:[xxx],ax
4  mov es:[xxx],ax
```

4. 偏移(索引)寄存器

偏移寄存器或者叫索引寄存器,主要是包含段地址的偏移量,用来进行内存地址的定位。

• BP(Base Pointer):基础指针,它是栈寄存器上的偏移量,用来定位栈上变量

• SP(Stack Pointer): 栈指针,它是栈寄存器上的偏移量,用来定位栈顶

• SI(Source Index): 变址寄存器,用来拷贝源字符串

• DI(Destination Index): 目标变址寄存器,用来复制到目标字符串

4.1. BP基础指针寄存器

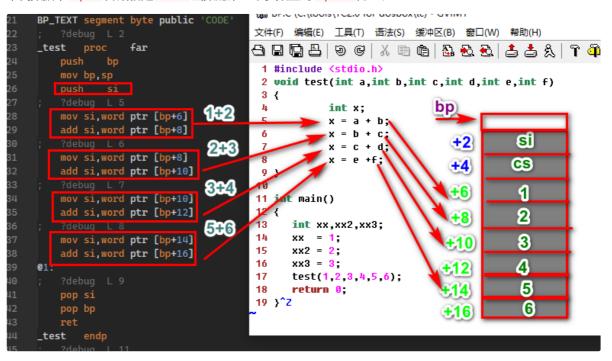
用 [bp-xx] 表示变量

```
9 }
                                   10
         mov bp,sp
                                   11 int main()
         sub sp,6
                                   12 {
                                   13
                                         int xx,xx2,xx3;
         mov~word ptr [bp-6],1
51
                                   14
                                         xx = 1;
                                   15
                                         xx2 = 2;
         mov~word ptr [bp-4],2
                                   16
                                         xx3 = 3;
                                   17
                                         test(1,2,3,4,5,6);
                                   18
         mov-word ptr [bp-2],3
                                          return 0:
                                   19 }^Z
```

用 [bp+xx] 表示函数参数

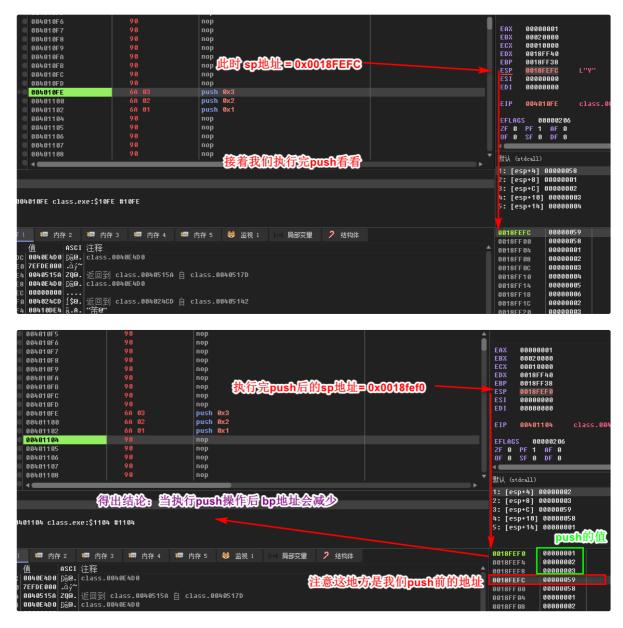
```
mov ax,6
push ax
mov ax,5
push ax
mov ax,4
push ax
mov ax,3
push ax
mov ax,2
push ax
mov ax,1
push 6,5,4,3,2,1,cs
push ax
push cs
call near ptr _test
```

堆栈数据中 bp+6 开始就是 test 函数的第一个参数直到 bp+16 为止。

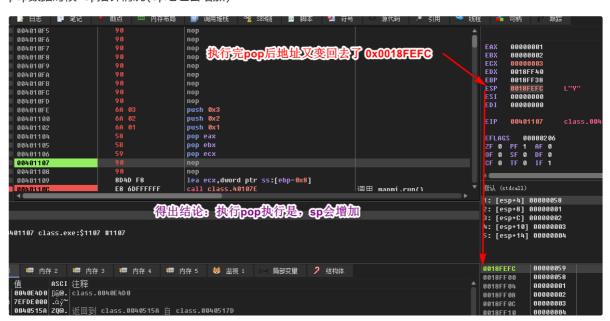


4.2. SP栈指针寄存器

这个其实主要是栈顶指针,无论何时sp指针都执行栈的顶部,[栈是一个从高地址向下生长的内存]。 push数据时候 sp指针情况(sp地址会减少)



pop数据时候 sp指针情况(sp地址会增加)



4.3. SI变址寄存器

si寄存器是变址寄存器,可以用来存放寻址用的偏移,此外si还被用作隐含的源串地址,默认在DS段中。 [无论是si还是di这两个寄存器总感觉和字符串都有点关系,c++的string?]

4.4. DI目标变址寄存器

di寄存器是目标变址寄存器,和si一样可以用来存放寻址用的偏移,此外di还被用作隐含的目的串地址,默认在ES段中。

先来看看STOS指令的介绍:

```
字符串存储指令 STOS

A 格式: STOS OPRD

D 功能: 把AL(字节)或AX(字)中的数据存储到DI为目的串地址指针所寻址的存储器单元中去.指针DI将根据DF的值进行自动调整.
```

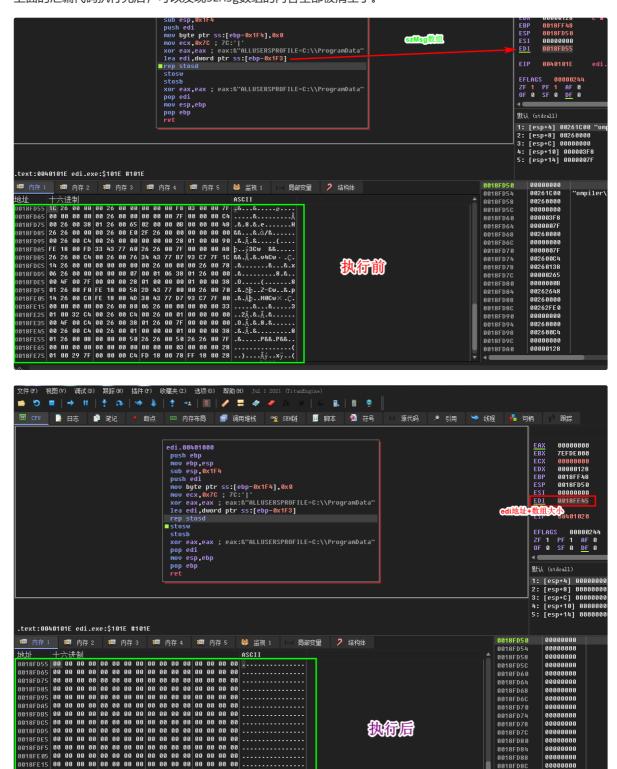
接下来我们来看看,在汇编语言中是如何实现高级语言中清空数组的操作,以下是一个初始化字符数组的高级代码。

```
1 | char szMsg[500]={0};
```

翻译后的汇编代码。

```
edi.00401000
push ebp
 mov ebp,esp
                       ecx循环0x7c*4次
 sub esp,0x1F4
 push edi
 mov byte ptr ss:[eop-0x1F4],0x0
 mov ecx.0x7C - 7C:'|'
xor eax.eax ; eax:&"ALLUSERSPROFILE=C:\\ProgramData
rep stosd
            rep 循环操作
         将eax内容复制到edi内存
 stosb stos
 xor eax,eax ; eax:&"ALLUSERSPROFILE=C:\\ProgramData"
 pop edi
 mov esp,ebp
 pop ebp
 ret
```

上面的汇编代码执行完后,可以发现szMsg数组的内容全部被清空了。



IP寄存器 **5.**

IP(Instruction Pointer): 指令指针寄存器,它是从 Code Segment 代码寄存器处的偏移来存储执行的 下一条指令。

00000000 00000000

00000000 00000000 00000000

ពពពពពពព

00000000 00000000

00000000 00000000

0018FD74 0018FD78 0018FD7C 0018FD80

0018FD84 0018FD88 0018FD8C 0018FD90

0018FD94 0018FD98 0018FD9C 0018FDA0

6. 标志寄存器

就剩下两种寄存器还没聊了,这两种寄存器是指令指针寄存器和标志寄存器:

- FLAG : Flag 寄存器用于存储当前进程的状态,这些状态有
 - 。 位置 (Direction): 用于数据块的传输方向, 是向上传输还是向下传输
 - 。 中断标志位 (Interrupt) : 1 允许; 0 禁止
 - 。 陷入位 (Trap): 确定每条指令执行完成后, CPU 是否应该停止。1 开启, 0 关闭
 - 。 进位 (Carry) : 设置最后一个无符号算术运算是否带有进位
 - 。 溢出 (Overflow) : 设置最后一个有符号运算是否溢出
 - 。符号 (Sign): 如果最后一次算术运算为负,则设置 1 =负,0 =正
 - 。 零位 (Zero): 如果最后一次算术运算结果为零, 1 = 零
 - 辅助进位 (Aux Carry) : 用于第三位到第四位的进位
 - 。 奇偶校验 (Parity) : 用于奇偶校验

参考文献:

https://blog.csdn.net/qq_39654127/article/details/88698911 《王爽汇编笔记》

https://segmentfault.com/a/1190000037478310《十一假期淦了八天寄存器的相关知识》

https://docs.microsoft.com/zh-cn/cpp/build/reference/fa-fa-listing-file?view=msvc-16 0 《/FA/Fa使用》

https://blog.csdn.net/bagboy_taobao_com/article/details/6203705 《了解寄存器:ESI EDI 变址寄存器》

http://c.biancheng.net/view/3679.html 《汇编语言字符串基本指令简介》

https://blog.csdn.net/weixin_43216249/article/details/110728729《汇编语言学习笔记--串操作篇(c++的string???)》