第2章 基础知识

UNIX运行在由数字设备公司 (DEC)制造的PDP11系列计算机的较大型号机上。本章对这些计算机，特别是PDP11/40机的某些重要特征提供概括性的摘要。

如果读者以前并不熟悉 PDP11系列计算机，那么应当先读 DEC出版的《PDP11处理机手册》。

一台PDP11计算机由下列几部分组成：一个处理器 (也称为CPU)、与之相连接的一个或多个内存部件以及若干外设控制器，将这三部分连接起来的是一根称之为“ Unibus ”(单总线)的双向平行通信线。

2.1 处理机

处理机是按16位字长的指令、数据和程序地址设计的，它包含一组高速寄存器。

2.2 处理机状态字

这是一个16位寄存器，它分成若干位段，它们的意义说明如下：

位	说 明
14，15	当前状态(00＝核心态)
12，13	前状态(11＝用户态)
5，6，7	处理机优先级(范围0－7)
4	陷入位
3	N位，若上次结果为负则设置此位
2	Z位，若上次结果为零则设置此位
1	V位，若上次操作产生溢出则设置此位
0	C位，若上次操作产生进位则设置此位

处理机可以在两种不同模式下操作：核心态和用户态。与用户态相比，核心态具有较高优先权，它由操作系统保留供自身使用。模式选择决定了下列事项：

- 使用哪一组存储管理段寄存器，这些寄存器的作用是将程序虚地址翻译成物理地址。
- 使用哪一个寄存器作为r6，r6是“ 栈指针 ”。
- 某些指令，例如halt(停机)，是否可以执行。

2.3 通用寄存器

处理机包含了一组 16位寄存器，其中有 8个是任何时候都可存访的“ 通用寄存器 ”。这些寄存器被称之为：

r0、r1、r2、r3、r4、r5、r6以及r7。

前6个通用寄存器可用作为累加器、地址指针或变址寄存器。 UNIX使用这些寄存器的惯例是：

- r0、r1在表达式求值时用作临时累加器；在过程返回时存放返回值；在过程调用的某些情况下可用来传递实参。
- r2、r3和r4在过程执行时可用作局部变量。在过程调用入口处存储这些寄存器的值，在退出过程时则恢复这些寄存器值。
- r5用作过程激活记录动态链的链首指针，该动态链存放在当前栈上。r5被称为“环境指针”。

最后两个通用寄存器具有专门的意义和作用：

- r6(也称为sp)用作栈指针。PDP11/40处理机有两个寄存器，分别在核心态和用户态下用作sp。而其他通用寄存器则没有这种双份使用方式。
- r7(也称为pc)用作程序计数器，亦即指令地址寄存器。

2.4 指令集

PDP11指令集包括双、单和零操作数指令。指令长度通常是一个字，某些指令则扩充为两个或叁个字以包括附加的寻址信息。

对于单操作数指令，其操作数通常称为“目的操作数”；在双操作数指令中，两个操作数分别被称为“源操作数”和“目的操作数”。后面将说明各种寻址方式。

文件m40.s是针对11/40处理机的汇编语言支持例程文件。下列指令在该文件中得到应用。注意，N、Z、V、C是处理机状态字(ps)中的条件码位。很多指令以副作用设置这些位，而“bit”、“cmp”和“tst”指令的主要功能则是设置这些条件码位。

adc 将C位的内容加至目的操作数。
add 将源操作数加至目的操作数。
ash 按移位计数将指定寄存器的内容左移相应次数(负
ashc 除涉及两个寄存器外，与ash指令相同。
asl 所有位向左移1位。第0位装入0，而第15位的原值
asr 所有位向右移1位。第15位保持原先值，而第0位
beq 若等于，也就是若Z=1则转移。
bge 若大于或等于，也就是若N=V则转移。
bhi 若高于，也就是若C＝0以及Z＝0则转移。
bhis 若高于或相同，也就是若C＝0则转移。
bic 若源操作数中的位为非零值，则将目的操作数中的
bis 对源和目的操作数执行或操作，将结果存入目的操
bit 对源和目的操作数执行逻辑与操作，设置条件码。
ble 若大于或等于，也就是若Z=1或N=V则转移。
blo 若低于(0)，也就是若C＝1则转移。
bne 若不等于(0)，也就是若Z=0则转移。

br 转移到.－128和.＋127之间的一个单元，其中，·表示当前单元的地址。

clc 清除C。

clr 将目的操作数清为0。

cmp 比较源和目的操作数，设置条件码位。若源操作数的值小于目的操作数值，则设置N位。

dec 将目的操作数的值减1。

div 存在rn和r(n+1)(其中，n是偶数)中的32位二进制补码整数被除以源操作数。商存在 rn中，而余数则存在r(n+1)中。

inc 将目的操作数的值加1。

jmp 跳转至目的地址。

jsr 跳转至子程序。寄存器值则按下列方式从左向右移动：

pc、 rn、－(sp)=dest.、pc、rn

mfpi 将前地址空间中指定字的值压入当前栈。

mov 将源操作数值复制至目的操作数。

mtpi 从当前栈弹出一个字，将其值存入前地址空间中的指定字中。

mul 将rn的内容和源操作数相乘。若n是偶数，则将积存放在rn和r(n+1)中。

reset 将Unibus中的INIT线设置成10ms。其作用是重新启动所有设备控制器。

ror 将目的操作数中各位循环右移1位。第0位的原先值装入至C，而C的原先值装入至第15位。

rts 从子程序中返回。从rn中重装pc，从栈中重装rn。

弹出重装 pc和ps。

结果非0，则转移回“位移”(offset)字。

攵。

低字节。

件码， N和Z。

us，直至发生一硬件中断。

下列字节版指令也用于 m40.s文件中：

艮多来自于它所提供的多种寻址方式，这些寻址方式可以

址方式。

2.5.1 寄存器方式

操作数驻留在一个通用寄存器中，例如：

```
clr             r0
mov             r1,r0
add             r4,r2
```

在下列方式中，指定的寄存器包含一地址值，用其定位操作数。

2.5.2 寄存器延迟方式

寄存器包含操作数的地址，例如：

```
inc             (r1)
asr             (sp)
add             (r2),r1
```

2.5.3 自动增1方式

寄存器包含操作数地址。作为一副作用，在操作后，寄存器值增 1，例如：

```
clr             (r1)+
mfpi            (r0)+
mov             (r1)+,r0
mov             r2,(r0)+
cmp             (sp)+,(sp)+
```

2.5.4 自动减1方式

寄存器值先减1，然后用来定位操作数，例如：

```
inc             －(r0)
mov             －(r1),r2
mov             (r0)+,－(sp)
clr             －(sp)
```

2.5.5 变址方式

操作数地址由两部分组成，一部分是寄存器包含的值，
位字值，两者相加得到操作数地址，例如：

```
clr             2(r0)
movb            6(sp),(sp)
movb            －reloc(r0),r0
```

```
mov                 -10(r2),(r1)
```

在这种寻址方式中，寄存器可被视为变址寄存器或基地址寄存器。 m40.s采用后一种看法。上面的第三个例子是使用寄存器作为变址寄存器的少数几个实例中的一个。 (注意，_reloc是一个可接受的变量名。)

有两种寻址方式，它们只用于下列两个例子中：

```
jsr                 pc,*(r0)+
jmp                 *0f(r0)
```

其中，第一个例子涉及“自动增量延迟”方式的使用。 (这用于例程call中的0785和0799行。)企图执行的子程序的地址在 r0指向的地址字中找到，也就是说涉及到两级间接寻址。在此种使用中r0增1的副作用并无任何实际影响。

第二个例子出现在 1055和1066行，它是“变址延迟”方式的一个例子。“jump”(“跳转”)的目的地址是一个字的内容，该字的地址值是 0f加上r0的值(一个小的正整数)。这是实现多路选择(multi-way switch)的标准方法。

下列两种方式用程序计数器作为指定的寄存器，以达到某种特定的作用。

2.5.6 立即方式

这是pc自动增1方式。操作数从程序串中取得，也就是这是一个立即操作数，例如：

```
add                 $2,r0
add                 $2,(r1)
                        0
                        0,r0
                        6,(r1)+
```

程序计数器值的地址从程序串中取得，并与 pc值相加得到

```
PS
R0
```

```
KISA6
```

变址延迟”、“立即”以及“相对”方式都将指令延长了一

减1”方式，再加上 r6的特殊属性使得我们可以方便地将出)列表中，在存储器中它向下生长。从此可以得到很多子位置独立的代码。

2.6 UNIX汇编程序

UNIX汇编程序是一个两遍汇编程序，它不具有宏功能。《UNIX汇编程序参考手册》对其作了完整的说明。

下面的一些简要注解应对读者有所帮助：

1) 一数字字符串可定义一数值常数。若该数字字符串以“ . ”终止，则被解释为十进制数，否则为八进制数。

2) 字符“ / ”用于表示该行的后继部分是注释。

3) 若在同一行中包含有多条语句，则相互间应当用分号分隔。

4) 字符“ . ”用于表示当前位置。

5) 在DEC汇编用“ # ”和“ @ ”之处，UNIX汇编分别改用了“ $ ”和“ * ”。

6) 标识符由若干字母数字字符组成 (包括下划线)。只有前8个字符是有意义的，而且第一个字符不能是数字字符。

7) 在“ C ”程序中出现的全局变量各在汇编程序中加了一个前缀，它由单个下划线组成。例如，在汇编语言文件“ m40.s ”中第1 025行上出现变量“ _regloc ”，它引用“ trap.c ”文件中2 677行的变量“ regloc ”。

8) 有两类语句标号：名字标号和数值标号。后者由一个数字后面跟一个冒号组成，这种标号无需是唯一的。如若引用 nf,其中“ n ”是一个数字，则表示要引用的是向前搜索遇到的第一个标号“ n：”。

如若引用“ nb ”，则表示要引用的是向后搜索遇到的第一个标号“ n ”。

9) 下列形式赋值语句

标识符＝表达式

将值和标识符的类型相结合。例如：

```
.＝60^.
```

其中，操作符“ ^ ”传递第一个操作数的值以及第二个
元 ”。

10) 字符串引用符号是“ ＜ ”和“ ＞ ”。

11) 下列形式的语句

```
.global x,y,z
```

使名字“ x ”、“ y ”和“ z ”为外部名。

12) 名字“ _edata ”和“ _end ”是装入程序的伪变量，
加上bss段的长度。

2.7 存储管理

在PDP11上运行的程序，其直接寻址最大可为 64K字节
致。考虑到经济因素，较大型号的 PDP11机可以配置较多的

字节)。PDP11机又包括了存管部件(memory management unit)，它将16位虚地址(或称程序地址)变换成18位或更多位的物理地址。PDP11/40的存管部件较PDP11/45或PDP11/70的简单。

PDP11/40的存管部件由两组寄存器组成，它们将虚地址映照为物理地址。这些寄存器被称之为“活动页寄存器”或者“段寄存器”。一组寄存器在处理机处于用户态时使用，另一组则在核心态时使用。更改这些寄存器的内容也就更改了虚、实地址的具体映照。进行这种更改的能力作为一种特权由操作系统保留自用。

2.8 段寄存器

每组段寄存器包含8对寄存器，其中每一对的组成是：“页地址寄存器”(PAR)和“页说明寄存器”(PDR)。

每一对寄存器控制一页的映照，而页长为8K字节(4K字)，程序虚地址空间包含8页。

每一页又分成128块(blocks)，每块长度为64字节(32字)。块长是存储映照的基本单位，也是存储分配的基本单位。

任一虚地址不是属于这一页，就是属于那一页。相应物理地址的产生方式是：将页内相对地址与相应PAR的内容相加，这构成了一个位数扩展的地址 (PDP11/40和11/45是18位；PDP11/70是22位)。于是每个页地址寄存器对相应页起重定位寄存器的作用。

每一页可被分成两部分，一部分被称为上部 (高地址部分)，另一部分则被称为下部 (低地址部分)，它们以32字的整数倍数作为边界而分隔，因此每一部分的长度都是 32字的整数倍。特别地，其中一个部分可以为空，在此情况下，另一部分就可占用整个页。两个部分中只有一部分包含有效虚地址。在另一部分中的地址则被说明为无效的。如果企图引用一

这种方案的优点是只需为页的有效部分分配物理存储空

存储的是32字块数－1)；
分是有效部分时该位被设置为1；
存取”、“只读”或“读/写”。
存取方式字段设置为“不存取”。

存储器的某一区间，但是硬件并不规定此种存储区间的分界上开始和结束外)。这些区间可以以任何顺序进行分配，

当有规律的，我们将在第 7章中看到这一点。相关页对应连续分配的，这样就使得一道程序的各段最多映照到物理

存储器中的两个区间。

2.11 状态寄存器

除段寄存器外，PDP11/40还有两个存储管理状态寄存器：

SR0 包含异常终止出错标志和其他有关操作系统的重要信息。特别地，当 SR0的第0位设置时，存储管理起作用。

SR2 在每条指令存取操作的开始处装入16位虚地址。

2.12 “i”和“d”空间

PDP11/45和11/70系统增加了段寄存器组。pc寄存器(r7)引用的地址被称为属于“ i ”空间，它由一组段寄存器翻译为物理地址；其余地址被称为属于“ d ”空间，它由另一组段寄存器翻译。

这种安排的优点是“ i ”和“ d ”空间可分别为32K字，于是可分配给程序的最大地址空间较PDP11/40扩大了1倍。

2.13 启动条件

当系统第一次启动时，存管部件被禁止操作，处理机处于核心态。

在这种环境下，变化范围在 0～56K字节的虚地址映照为相同的物理地址值。但是，虚地址空间的最高页映照为物理地址空间的最高页，在 PDP11/40或11/45中，地址区间

```
0160000－0177777
```

被映照为下列物理地址区间

```
0760000－0777777
```

2.14 专用设备寄存器

物理存储器的最后一页保留用于处理机及各外部设备
存储器空间减少了1页，但是却不再需要提供特殊的指令类
为寄存器在此页中分配地址的方法是一种创新，其价