

Uboot中start.S源码的指令级的详尽解析

版本：v1.8

Crifan Li

摘要

本文对Uboot中的Start.S的源码的几乎每一行，都进行了详细的解析



本文提供多种格式供：

在线阅读	HTML ¹	HTMLs ²	PDF ³	CHM ⁴	TXT ⁵	RTF ⁶	WEBHELP ⁷
下载（7zip压缩包）	HTML ⁸	HTMLs ⁹	PDF ¹⁰	CHM ¹¹	TXT ¹²	RTF ¹³	WEBHELP ¹⁴

HTML版本的在线地址为：

http://www.crifan.com/files/doc/docbook/uboot_starts_analysis/release/html/uboot_starts_analysis.html

有任何意见，建议，提交bug等，都欢迎去讨论组发帖讨论：

http://www.crifan.com/bbs/categories/uboot_starts_analysis/

修订历史

修订 1.0	2011-04-17	crl
1. 详细解释了uboot的start.s中的每行代码 2. 添加了相关知识点的详细解释		
修订 1.6	2011-05-01	crl
1. 添加汇编学习记录 2. 添加了如何查看C或汇编的源代码所对应的真正的汇编代码 3. 添加Start S的总结		

¹ http://www.crifan.com/files/doc/docbook/uboot_starts_analysis/release/html/uboot_starts_analysis.html

² http://www.crifan.com/files/doc/docbook/uboot_starts_analysis/release/htmls/index.html

³ http://www.crifan.com/files/doc/docbook/uboot_starts_analysis/release/pdf/uboot_starts_analysis.pdf

⁴ http://www.crifan.com/files/doc/docbook/uboot_starts_analysis/release/chm/uboot_starts_analysis.chm

⁵ http://www.crifan.com/files/doc/docbook/uboot_starts_analysis/release/txt/uboot_starts_analysis.txt

⁶ http://www.crifan.com/files/doc/docbook/uboot_starts_analysis/release/rtf/uboot_starts_analysis.rtf

⁷ http://www.crifan.com/files/doc/docbook/uboot_starts_analysis/release/webhelp/index.html

⁸ http://www.crifan.com/files/doc/docbook/uboot_starts_analysis/release/html/uboot_starts_analysis.html.7z

⁹ http://www.crifan.com/files/doc/docbook/uboot_starts_analysis/release/htmls/index.html.7z

¹⁰ http://www.crifan.com/files/doc/docbook/uboot_starts_analysis/release/pdf/uboot_starts_analysis.pdf.7z

¹¹ http://www.crifan.com/files/doc/docbook/uboot_starts_analysis/release/chm/uboot_starts_analysis.chm.7z

¹² http://www.crifan.com/files/doc/docbook/uboot_starts_analysis/release/txt/uboot_starts_analysis.txt.7z

¹³ http://www.crifan.com/files/doc/docbook/uboot_starts_analysis/release/rtf/uboot_starts_analysis.rtf.7z

¹⁴ http://www.crifan.com/files/doc/docbook/uboot_starts_analysis/release/webhelp/uboot_starts_analysis.webhelp.7z

<div><div><div>a. Start.S的各个部分的总结</div><div>b. Uboot中的内存的layout</div></div><div>4. 更加详细地解释了为何ARM9中PC=PC+8</div><div>5. 添加了一些其他的细节的内容</div><div>6. 修正一些拼写错误</div></div>		
修订 1.8	2012-11-05	crl
<div><div>1. 通过Docbook发布</div><div>2. 修正了一些琐碎的笔误，同时增添了些琐碎内容</div><div>3. 修正了0xdeadbeef的解释</div></div>		

Uboot中start.S源码的指令级的详尽解析:

Crifan Li

版本 : **v1.8**

出版日期 2012-11-05

版权 © 2012 Crifan, <http://crifan.com>

本文章遵从 : [署名-非商业性使用 2.5 中国大陆\(CC BY-NC 2.5\)](#)¹⁵

¹⁵ http://www.crifan.com/files/doc/docbook/soft_dev_basic/release/html/soft_dev_basic.html#cc_by_nc

目录

正文之前	ix
1. 本文内容	ix
2. 本文目标	ix
3. 代码来源	ix
4. 阅读此文所要具有的前提知识	ix
5. 声明	ix
1. start.S详解	1
1.1. 设置CPU模式	1
1.1.1. globl	1
1.1.2. _start	2
1.1.3. ldr	3
1.1.4. .word	5
1.1.5. .balign	6
1.1.6. _TEXT_BASE _armboot_start	7
1.1.7. _bss_start _bss_end	7
1.1.8. FREE_RAM_END FREE_RAM_SIZE	8
1.1.9. IRQ_STACK_START FIQ_STACK_START	9
1.1.10. cpsr	10
1.1.11. bic	12
1.1.12. orr	12
1.1.13. msr	13
1.2. 关闭看门狗	14
1.2.1. pWTCON INTMOD INTMSK INTSUBMSK CLKDIVN	14
1.2.2. ldr pWTCON	15
1.2.3. mov	17
1.2.4. str	17
1.3. 关闭中断	19
1.3.1. set INTMSK	19
1.3.2. set INTSUBMSK	20
1.3.3. set CLKDIVN	21
1.3.4. bl	22
1.4. 设置堆栈sp指针	23
1.4.1. stack_setup	23
1.4.2. calc stack	25
1.4.3. clock_init	26
1.4.4. adr	27
1.4.5. clear_bss	29
1.4.6. cal armboot size	29
1.4.7. cal armboot size	31
1.5. 清除bss段	33
1.5.1. clear_bss	33
1.5.2. clear css loop	33
1.5.3. ldr pc	33
1.5.4. cpu_init_crit	34
1.5.5. disable MMU	39
1.5.6. clear bits	40
1.5.7. bl lowlevel_init	43
1.6. 异常中断处理	43
1.6.1. macros stmia	43
1.6.2. cal reg value and store	47
1.6.3. irq_save_user_regs irq_restore_user_regs	48
1.6.4. exception handlers	50
1.6.5. Launch	52
1.6.6. int_return	53
2. start.S的总结	55

2.1. start.S各个部分的总结	55
2.2. Uboot中的内存的Layout	56
3. 相关知识点详解	58
3.1. 如何查看C或汇编的源代码所对应的真正的汇编代码	58
3.2. uboot初始化中，为何要设置CPU为SVC模式而不是设置为其他模式	59
3.3. 什么是watchdog + 为何在要系统初始化的时候关闭watchdog	61
3.3.1. 什么是watchdog	61
3.3.2. 为何在要系统初始化的时候关闭watchdog	61
3.4. 为何ARM7中PC=PC+8	62
3.4.1. 为何ARM9和ARM7一样，也是PC=PC+8	63
3.5. AMR寄存器的别名 + APCS	69
3.5.1. ARM中的寄存器的别名	69
3.5.2. 什么是APCS	71
3.6. 为何C语言（的函数调用）需要堆栈，而汇编语言却不需要堆栈	71
3.6.1. 保存现场/上下文	71
3.6.1.1. 什么叫做上下文context	71
3.6.2. 传递参数	72
3.6.3. 举例分析C语言函数调用是如何使用堆栈的	72
3.7. 关于为何不直接用mov指令，而非要用adr伪指令	73
3.8. mov指令的操作数的取值范围到底是多少	74
3.9. 汇编学习总结记录	77
3.9.1. 汇编中的标号=C中的标号	77
3.9.2. 汇编中的跳转指令=C中的goto	77
3.9.3. 汇编中的.globl=C语言中的extern	78
3.9.4. 汇编中用bl指令和mov pc, lr来实现子函数调用和返回	78
3.9.5. 汇编中的对应位置有存储值的标号 = C语言中的指针变量	79
3.9.6. 汇编中的ldr+标号，来实现C中的函数调用	81
3.9.7. 汇编中设置某个寄存器的值或给某个地址赋值	82
参考书目	85

插图清单

1.1. LDR指令的语法	4
1.2. CPSR/SPSR的位域结构	11
1.3. pWTCON	14
1.4. INTMOD	14
1.5. INTMSK	15
1.6. INTSUBMSK	15
1.7. CLKDIVN	15
1.8. WTCON寄存器的位域	18
1.9. INTMSK寄存器的位域	20
1.10. INTSUBMSK寄存器的位域	21
1.11. INTSUBMSK寄存器的位域	22
1.12. macro的语法	45
1.13. LDM/STM的语法	45
1.14. 条件码的含义	46
2.1. Uboot中的内存的Layout	57
3.1. AMR7三级流水线	62
3.2. ARM7三级流水线状态	62
3.3. ARM7三级流水线示例	63
3.4. ARM7三级流水线 vs ARM9五级流水线	64
3.5. ARM7三级流水线到ARM9五级流水线的映射	64
3.6. ARM9的五级流水线示例	65
3.7. ARM9的五级流水线中为何PC=PC+8	67
3.8. ARM Application Procedure Call Standard (AAPCS)	69
3.9. 数据处理指令的指令格式	75

表格清单

1.1. global的语法	2
1.2. .word的语法	5
1.3. balignl的语法	6
1.4. CPSR Bitfield	11
1.5. CPSR=0xD3的位域及含义	13
1.6. 控制寄存器1的位域含义	40
1.7. 时钟模式	41
1.8. 关于访问控制位在域访问控制寄存器中的含义	42
1.9. 关于访问允许(AP)位的含义	42
3.1. ARM中CPU的模式	60
3.2. ARM寄存器的别名	70
3.3. mov指令0xe3a00453的位域含义解析	76

范例清单

3.1. 汇编中的ldr加标号实现函数调用 示例	81
3.2.	83
3.3.	83

正文之前

1. 本文内容

此文主要内容就是分析start.S这个汇编文件的内容，即ARM上电后的最开始那一段的启动过程。

2. 本文目标

本文的目标是，希望看完此文的读者，可以达到：

1. 微观上，对此start.S的每一行，都有了基本的了解
2. 宏观上，对基于ARM核的S3C24X0的CPU的启动过程，有更加清楚的概念

这样的目的，是为了读者看完本文后，再去看看其他类似的启动相关的源码，能明白需要做什么事情，然后再看别的系统是如何实现相关的内容的，达到一定程度的触类旁通。

总体说就是，要做哪些，为何要这么做，如何实现的，即英语中常说的：

- do what
- why do
- how do

此三方面都清楚理解了，那么也才能算真正懂了。

3. 代码来源

所用代码来自TQ2440官网，天嵌的bbs上下载下来的uboot中的源码：

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\cpu\arm920t\start.S

下载地址为：[2010年6月 最新TQ2440光盘下载（Linux内核，WinCE的eboot，uboot均有更新）](http://bbs.embedsky.net/viewthread.php?tid=859)¹

4. 阅读此文所要具有的前提知识

阅读此文之前，你至少要对TQ2440的板子有个基本的了解，

以及要了解开发板初始化的大概要做的事情，比如设置输入频率，设置堆栈等等。

另外，至少要有一定的C语言的基础，这样更利于理解汇编代码。

5. 声明

由于水平有限，难免有误，欢迎指正：admin (at) crifan.com

欢迎转载，但请注明作者。

¹ <http://bbs.embedsky.net/viewthread.php?tid=859>

第 1 章 start.S详解

下面将详细解释uboot中的start.S中的每一行代码。详细到，每个指令的语法和含义，都进行详细讲解，使得此文读者可以真正搞懂具体的含义，即what。

以及对于一些相关的问题，深入探究为何要这么做，即why。

对于uboot的start.S，主要做的事情就是系统的各个方面的初始化。

从大的方面分，可以分成这几个部分：

- 设置CPU模式
- 关闭看门狗
- 关闭中断
- 设置堆栈sp指针
- 清除bss段
- 异常中断处理

下面来对start.S进行详细分析，看看每一个部分，是如何实现的。

1.1. 设置CPU模式

1.1.1. globl

```
/*
 * armboot - Startup Code for ARM920 CPU-core
 *
 * Copyright (c) 2001 Marius Gröber <mag@sysgo.de>
 * Copyright (c) 2002 Alex Zerkov <azu@sysgo.de>
 * Copyright (c) 2002 Gary Jennejohn <gj@denx.de>
 *
 * See file CREDITS for list of people who contributed to this
 * project.
 *
 * This program is free software; you can redistribute it and/or
 * modify it under the terms of the GNU General Public License as
 * published by the Free Software Foundation; either version 2 of
 * the License, or (at your option) any later version.
 *
 * This program is distributed in the hope that it will be useful,
 * but WITHOUT ANY WARRANTY; without even the implied warranty of
 * MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the
 * GNU General Public License for more details.
 *
 * You should have received a copy of the GNU General Public License
 * along with this program; if not, write to the Free Software
 * Foundation, Inc., 59 Temple Place, Suite 330, Boston,
 * MA 02111-1307 USA
 */
#include <config.h>
#include <version.h>
```

```

/*
*****
*
* Jump vector table as in table 3.1 in [1]
*
*****
*/

.globl① _start

```

- ① `globl`是个关键字，对应含义为：
<http://re-eject.gbadev.org/files/GasARMRef.pdf>

表 1.1. `global`的语法

Directive	Description	Syntax	Example
<code>.global</code>	Makes <code>symbol</code> visible to the linker	<code>.global symbol</code>	<code>.global MyAsmFunc</code>
<code>.globl</code>	Same as <code>.global</code>	<code>.globl symbol</code>	<code>.globl MyOtherAsmFunc</code>

所以，意思很简单，就是相当于C语言中的Extern，声明此变量，并且告诉链接器此变量是全局的，外部可以访问

所以，你可以看到

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\board\EmbedSky\u-boot.lds

中，有用到此变量:

`ENTRY(_start)`

即指定入口为`_start`,而由下面的`_start`的含义可以得知，`_start`就是整个start.S的最开始，即整个uboot的代码的开始。

1.1.2. `_start`

```

_start①: b    reset

```

- ① `_start`后面加上一个冒号‘:’，表示其是一个标号Label，类似于C语言goto后面的标号。

而同时，`_start`的值，也就是这个代码的位置了，此处即为代码的最开始，相对的0的位置。

而此处最开始的位置，在程序开始运行的时候，如果是从NorFlash启动，那么其地址是0，

`_stat=0`

如果是重新relocate代码之后，就是我们定义的值了，即，在

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\board\EmbedSky\config.mk

中的：

```
TEXT_BASE = 0x33D00000
```

表示是代码段的基地址，即

```
_start=TEXT_BASE=0x33D00000
```

关于标号的语法解释:

<http://sourceware.org/binutils/docs-2.20/as/Labels.html#Labels>

A label is written as a symbol immediately followed by a colon `:`. The symbol then represents the current value of the active location counter, and is, for example, a suitable instruction operand. You are warned if you use the same symbol to represent two different locations: the first definition overrides any other definitions.

而_start标号后面的:

```
b    reset
```

就是跳转到对应的标号为reset的位置。

1.1.3. ldr

```
ldr① pc, _undefined_instruction  
ldr pc, _software_interrupt  
ldr pc, _prefetch_abort  
ldr pc, _data_abort  
ldr pc, _not_used  
ldr pc, _irq  
ldr pc, _fiq
```

① ldr命令的语法为：

http://infocenter.arm.com/help/topic/com.arm.doc.dui0206hc/DUI0206HC_rvct_linker_and_utilities_guide.pdf

图 1.1. LDR指令的语法

4.1.20 LDR

31	28	27	26	25	24	23	22	21	20	19	16	15	12	11	0
cond	0	1	I	P	U	0	W	1	Rn	Rd	addr_mode				

The LDR (Load Register) instruction loads a word from the memory address calculated by <addressing_mode> and writes it to register <Rd>. If the address is not word-aligned, the loaded value is rotated right by 8 times the value of bits[1:0] of the address. For a little-endian memory system, this rotation causes the addressed byte to occupy the least significant byte of the register. For a big-endian memory system, it causes the addressed byte to occupy bits[31:24] or bits[15:8] of the register, depending on whether bit[0] of the address is 0 or 1 respectively.

If the PC is specified as register <Rd>, the instruction loads a data word which it treats as an address, then branches to that address. In ARM architecture version 5 and above, bit[0] of the loaded value determines whether execution continues after this branch in ARM state or in Thumb state, as though a BX (loaded_value) instruction had been executed. In earlier versions of the architecture, bits[1:0] of the loaded value are ignored and execution continues in ARM state, as though a MOV PC, (loaded_value) instruction had been executed.

Syntax

LDR{<cond>} <Rd>, <addressing_mode>

where:

<cond> Is the condition under which the instruction is executed. The conditions are defined in *The condition field* on page A3-5. If <cond> is omitted, the AL (always) condition is used.

<Rd> Specifies the destination register for the loaded value.

<addressing_mode>

Is described in *Addressing Mode 2 - Load and Store Word or Unsigned Byte* on page A5-18. It determines the I, P, U, W, Rn and addr_mode bits of the instruction.

The syntax of all forms of <addressing_mode> includes a base register <Rn>. Some forms also specify that the instruction modifies the base register value (this is known as *base register writeback*).

<http://wenku.baidu.com/view/f7cc280102020740be1e9bea.html>

LDR指令的格式为：

LDR{条件} 目的寄存器，<存储器地址>

LDR指令用于从存储器中将一个32位的字数据传送到目的寄存器中。该指令通常用于从存储器中读取32位的字数据到通用寄存器，然后对数据进行处理。当程序计数器PC作为目的寄存器时，指令从存储器中读取的字数据被当作目的地址，从而可以实现程序流程的跳转。该指令在程序设计中比较常用，且寻址方式灵活多样，请读者认真掌握。

指令示例：

LDR R0, [R1] ; 将存储器地址为R1的字数据读入寄存器R0。

LDR R0, [R1, R2] ; 将存储器地址为R1+R2的字数据读入寄存器R0。

LDR R0, [R1, #8] ; 将存储器地址为R1+8的字数据读入寄存器R0。

LDR R0, [R1, R2]! ; 将存储器地址为R1+R2的字数据读入寄存器R0,并将新地址R1 + R2写入R1。

LDR R0, [R1, #8]! ; 将存储器地址为R1+8的字数据读入寄存器R0,并将新地址R1 + 8写入R1。

LDR R0, [R1], R2 ; 将存储器地址为R1的字数据读入寄存器R0,并将新地址R1 + R2写入R1。

LDR R0, [R1, R2, LSL # 2] ! ; 将存储器地址为 $R1 + R2 \times 4$ 的字数据读入寄存器R0, 并将新地址 $R1 + R2 \times 4$ 写入R1。

LDR R0, [R1], R2, LSL # 2 ; 将存储器地址为R1的字数据读入寄存器R0, 并将新地址 $R1 + R2 \times 4$ 写入R1。"

<http://www.pczpg.com/a/2010/0607/11062.html>

ARM是RISC结构, 数据从内存到CPU之间的移动只能通过L/S指令来完成, 也就是ldr/str指令。

比如想把数据从内存中某处读取到寄存器中, 只能使用ldr

比如:

```
ldr r0, 0x12345678
```

就是把0x12345678这个地址中的值存放到r0中。

上面那些ldr的作用, 以第一个_undefined_instruction为例, 就是将地址为_undefined_instruction中的一个word的值, 赋值给pc。

1.1.4. .word

```
_undefined_instruction: .word ① undefined_instruction
_software_interrupt: .word software_interrupt
_prefetch_abort: .word prefetch_abort
_data_abort: .word data_abort
_not_used: .word not_used
_irq: .word irq
_fiq: .word fiq
```

① http://blogold.chinaunix.net/u3/115924/showart_2280163.html

.word .word expr {expr}... 分配一段字内存单元, 并用expr初始化字内存单元(32bit)

<http://re-eject.gbadev.org/files/GasARMRef.pdf>

表 1.2. .word的语法

Directive	Description	Syntax	Example
.word	Define word expr (32bit numbers)	.word expr {, ...}	.word 144511, 0x11223

所以上面的含义, 以_undefined_instruction为例, 就是, 此处分配了一个word=32bit=4字节的地址空间, 里面存放的值是undefined_instruction。

而此处_undefined_instruction也就是该地址空间的地址了。用C语言来表达就是:

```
_undefined_instruction = &undefined_instruction
```

或

```
*_undefined_instruction = undefined_instruction
```

在后面的代码，我们可以看到，`undefined_instruction`也是一个标号，即一个地址值，对应着就是在发生“未定义指令”的时候，系统所要去执行的代码。

（其他几个对应的“软件中断”，“预取指错误”，“数据错误”，“未定义”，“（普通）中断”，“快速中断”，也是同样的做法，跳转到对应的位置执行对应的代码。）

所以：

```
ldr pc, 标号1
.....
标号1: .word 标号2
.....
标号2:
..... (具体要执行的代码)
```

的意思就是，将地址为标号1中内容载入到pc，而地址为标号1中的内容，正好装的是标号2。

用C语言表达其实很简单：

$PC = *(\text{标号1}) = \text{标号2}$

对PC赋值，即是实现代码跳转，所以整个这段汇编代码的意思就是：

跳转到标号2的位置，执行对应的代码。

1.1.5. .balignl

.balignl 16,0xdeadbeef

- ❶ .balignl这个标号的语法及含义：

<http://re-eject.gbadev.org/files/GasARMRef.pdf>

表 1.3. .balignl的语法

Directive	Description	Syntax	Example
.balignl	Word align the following code to alignment byte boundary (default=4). Fill skipped words with fill (default=0 or NOP). If the number of bytes skipped is greater than max, then don't align (default=alignment).	.balignl {alignment} {, fill} {, max}	.balignl

所以意思就是，接下来的代码，都要16字节对齐，不足之处，用0xdeadbeef填充。

其中关于所要填充的内容0xdeadbeef，刚开始没看懂是啥意思，后来终于搞懂了。

经过（<645116977@qq.com>等）多位网友提示和纠正，觉得这样解释会更加合理些：

此处0xdeadbeef本身没有真正的意义，但是很明显，字面上的意思是，（坏）死的牛肉。

虽然其本身没有实际意义，但是其是在十六进制下，能表示出来的，为数不多的，可读的单词之一了。

另外一个相对常见的是：**0xbadc0de**，意思是bad code，坏的代码，注意其中的o是0，因为十六进制中是没有o的。

这些“单词”，相对的作用是，使得读代码的人，以及在查看程序运行结果时，容易看懂，便于引起注意。

而关于自己之前，随意杜撰出来的，希望起到搞笑作用，表示good beef（好的牛肉）的0xgoodbeef，实际上，在十六进制下，会出错的，因为十六进制下没有o和g这两个字母。

1.1.6. _TEXT_BASE _armboot_start

```
/*
*****
*
* Startup Code (reset vector)
*
* do important init only if we don't start from memory!
* relocate armboot to ram
* setup stack
* jump to second stage
*
*****
*/

_TEXT_BASE①:
.word TEXT_BASE

②.globl _armboot_start
_armboot_start:
.word _start
```

- ① 此处和上面的类似，_TEXT_BASE是一个标号地址，此地址中是一个word类型的变量，变量名是TEXT_BASE，此值见名知意，是text的base，即代码的基地址，可以在

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\board\EmbedSky\config.mk

中找到其定义：

```
TEXT_BASE = 0x33D00000
```

- ② 同理，此含义可用C语言表示为：

```
*(_armboot_start) = _start
```

1.1.7. _bss_start _bss_end

```
/*
* These are defined in the board-specific linker script.
*/
.globl _bss_start
_bss_start:
.word __bss_start
```



```
❶
.globl _bss_end
_bss_end:
.word _end
```

- ❶ 关于_bss_start和_bss_end都只是两个标号，对应着此处的地址。

而两个地址里面分别存放的值是__bss_start和_end,这两个的值，根据注释所说，是定义在开发板相关的链接脚本里面的，我们此处的开发板相关的链接脚本是：

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\board\EmbedSky\u-boot.lds

其中可以找到__bss_start和_end的定义：

```
__bss_start = .;
.bss : { *(.bss) }
_end = .;
```

而关于__bss_start和__bss_end定义为.global即全局变量，是因为uboot的其他源码中要用到这两个变量，详情请自己去搜索源码。

1.1.8. FREE_RAM_END FREE_RAM_SIZE

```
.globl FREE_RAM_END
FREE_RAM_END:
.word 0x0badc0de
```

```
❶
.globl FREE_RAM_SIZE
FREE_RAM_SIZE:
.word 0x0badc0de
```

- ❶ 关于FREE_RAM_END和FREE_RAM_SIZE，这里只是两个标号，之所以也是声明为全局变量，是因为uboot的源码中会用到这两个变量。

但是这里有点特别的是，这两个变量，将在本源码start.S中的后面要用到，而在后面用到这两个变量之前，uboot的C源码中，会先去修改这两个值，具体的逻辑是：

本文件start.S中，后面有这两句：

```
ldr pc, _start_armboot
_start_armboot: .word start_armboot
```

意思很明显，就是去调用start_armboot函数。

而start_armboot函数是在：

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\lib_arm\board.c

中：

```

init_fnc_t *init_sequence[] = {
    cpu_init, /* basic cpu dependent setup */
    .....,
    NULL,
};

void start_armboot (void)
{
    init_fnc_t **init_fnc_ptr;
    .....,

    for (init_fnc_ptr = init_sequence; *init_fnc_ptr; ++init_fnc_ptr) {
        if ((*init_fnc_ptr)() != 0) {
            hang ();
        }
    }
    .....,
}

```

即在start_armboot去调用了cpu_init。

cpu_init函数是在：

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\cpu\arm920t\cpu.c

中：

cpu_init源码。

```

int cpu_init (void)
{
    /*
     * setup up stacks if necessary
     */
#ifdef CONFIG_USE_IRQ
    IRQ_STACK_START = _armboot_start - CFG_MALLOC_LEN - CFG_GBL_DATA_SIZE - 4;
    FIQ_STACK_START = IRQ_STACK_START - CONFIG_STACKSIZE_IRQ;
    FREE_RAM_END = FIQ_STACK_START - CONFIG_STACKSIZE_FIQ - CONFIG_STACKSIZE;
    FREE_RAM_SIZE = FREE_RAM_END - PHYS_SDRAM_1;
#else
    FREE_RAM_END = _armboot_start - CFG_MALLOC_LEN - CFG_GBL_DATA_SIZE - 4 -
        CONFIG_STACKSIZE;
    FREE_RAM_SIZE = FREE_RAM_END - PHYS_SDRAM_1;
#endif
    return 0;
}

```

在cpu_init中，根据我们的一些定义，比如堆栈大小等等，去修改了IRQ_STACK_START，FIQ_STACK_START，FREE_RAM_END和FREE_RAM_SIZE的值。

至于为何这么修改，后面遇到的时候会具体再解释。

1.1.9. IRQ_STACK_START FIQ_STACK_START

```

#ifdef CONFIG_USE_IRQ
/* IRQ stack memory (calculated at run-time) */
.globl IRQ_STACK_START
IRQ_STACK_START:
.word 0x0badc0de
❶
/* FIQ stack memory (calculated at run-time) */
.globl FIQ_STACK_START
FIQ_STACK_START:
.word 0x0badc0de
#endif

```

❶ 同上，IRQ_STACK_START和FIQ_STACK_START，也是在cpu_init中用到了。

不过此处，是只有当定义了宏CONFIG_USE_IRQ的时候，才用到这两个变量，其含义也很明显，只有用到了中断IRQ，才会用到中断的堆栈，才有中端堆栈的起始地址。

快速中断FIQ，同理。

1.1.10. cpsr

```

/*
 * the actual reset code
 */

reset:
/*
 * set the cpu to SVC32 mode
 */
mrs❶ r0,cpsr❷

```

❷ CPSR 是当前的程序状态寄存器(Current Program Status Register)，

而 SPSR 是保存的程序状态寄存器(Saved Program Status Register)。

具体细节，可参考：

[ARM7体系结构](http://www.docin.com/p-73665362.html)¹

<http://www.csie.nctu.edu.tw/~wjtsai/EmbeddedSystemDesign/Ch2-bootloader.pdf>

¹ <http://www.docin.com/p-73665362.html>

图 1.2. CPSR/SPSR的位域结构

• Program Status Register (PSR)

- **CPSR**: Current Program Status Register
- **SPSR**: Saved Program Status Register

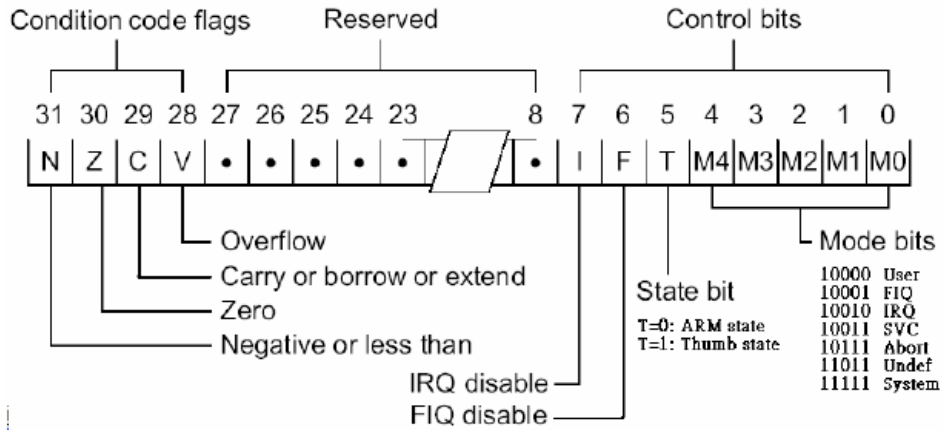


表 1.4. CPSR Bitfield

31	30	29	28	---	7	6	-	4	3	2	1	0	说明
N	Z	C	V		I	F		M4	M3	M2	M1	M0	
								0	0	0	0	0	User26 模式
								0	0	0	0	1	FIQ26 模式
								0	0	0	1	0	IRQ26 模式
								0	0	0	1	1	SVC26 模式
								1	0	0	0	0	User 模式
								1	0	0	0	1	FIQ 模式
								1	0	0	1	0	IRQ 模式
								1	0	0	1	1	SVC 模式
								1	0	1	1	1	ABT 模式
								1	1	0	1	1	UND 模式

❶ MRS - Move From Status Register

MRS指令的语法为：

四、程序状态寄存器访问指令

1、MRS指令

MRS指令的格式为：

MRS{条件} 通用寄存器，程序状态寄存器 (CPSR或SPSR)

MRS指令用于将程序状态寄存器的内容传送到通用寄存器中。该指令一般用在以下两种情况：

I. 当需要改变程序状态寄存器的内容时，可用MRS将程序状态寄存器的内容读入通用寄存器，修改后再写回程序状态寄存器。

II.当在异常处理或进程切换时，需要保存程序状态寄存器的值，可先用该指令读出程序状态寄存器的值，然后保存。

指令示例：

MRS R0, CPSR ; 传送CPSR的内容到R0

MRS R0, SPSR ; 传送SPSR的内容到R0”

所以，上述汇编代码含义为，将CPSR的值赋给R0寄存器。

1.1.11. bic

```
bic① r0,r0,#0x1f
```

① bic指令的语法是：

16、BIC指令

BIC指令的格式为：

BIC{条件}{S} 目的寄存器，操作数1，操作数2

BIC指令用于清除操作数1的某些位，并把结果放置到目的寄存器中。操作数1应是一个寄存器，

操作数2可以是一个寄存器，被移位的寄存器，或一个立即数。操作数2为32位的掩码，如果在掩码中设置了某一位，则清除这一位。未设置的掩码位保持不变。

而0x1f=11111b

所以，此行代码的含义就是，清除r0的bit[4:0]位。

1.1.12. orr

```
orr① r0,r0,#0xd3
```

① orr指令的语法是：

14、ORR指令

ORR指令的格式为：

ORR{条件}{S} 目的寄存器，操作数1，操作数2

ORR指令用于在两个操作数上进行逻辑或运算，并把结果放置到目的寄存器中。操作数1应是一个寄存器，操作数2可以是一个寄存器，被移位的寄存器，或一个立即数。该指令常用于设置操作数1的某些位。

指令示例：

ORR R0, R0, #3 ; 该指令设置R0的0、1位，其余位保持不变。

所以此行汇编代码的含义为：

而0xd3=1101 0111[4:0]位。

将r0与0xd3算数或运算，然后将结果给r0,即把r0的bit[7:6]和bit[4]和bit[2:0]置为1。

1.1.13. msr

msr^❶ cpsr,r0

❶ MSR - Move to Status Register

msr的指令格式是：

四、程序状态寄存器访问指令

.....

2、MSR指令

MSR指令的格式为：

MSR{条件} 程序状态寄存器 (CPSR或SPSR) _<域> , 操作数

MSR指令用于将操作数的内容传送到程序状态寄存器的特定域中。其中，操作数可以为通用寄存器或立即数。<域>用于设置程序状态寄存器中需要操作的位，32位的程序状态寄存器可分为4个域：

位[31：24]为条件标志位域，用f表示；

位[23：16]为状态位域，用s表示；

位[15：8]为扩展位域，用x表示；

位[7：0]为控制位域，用c表示；

该指令通常用于恢复或改变程序状态寄存器的内容，在使用时，一般要在MSR指令中指明将要操作的域。

指令示例：

MSR CPSR , R0 ; 传送R0的内容到CPSR

MSR SPSR , R0 ; 传送R0的内容到SPSR

MSR CPSR_c , R0 ; 传送R0的内容到SPSR，但仅仅修改CPSR中的控制位域

此行汇编代码含义为，将r0的值赋给CPSR。

所以，上面四行汇编代码的含义就很清楚了。

先是把CPSR的值放到r0寄存器中，然后清除bit[4:0]，然后再或上

0xd3=11 0 10111b

表 1.5. CPSR=0xD3的位域及含义

CPSR位域	7	6	5	4	3	2	1	0
位域含义	I	F		M4	M3	M2	M1	M0
0xD3	1	1	0	1	0	0	1	1

对应含义	关闭中断 IRQ	关闭快速 中断FIQ		设置CPU为SVC模式,这和上面代码注释中的“set the cpu to SVC32 mode”，也是一致的。
------	-------------	---------------	--	--

关于为何设置CPU为SVC模式，而不是设置为其他模式，请参见本文档后面的章节：[第 3.2 节 “uboot 初始化中，为何要设置CPU为SVC模式而不是设置为其他模式”](#)

1.2. 关闭看门狗

1.2.1. pWTCON INTMOD INTMSK INTSUBMSK CLKDIVN

```
/* turn off the watchdog */
#ifdef CONFIG_S3C2400
# define pWTCON 0x15300000
# define INTMSK 0x14400008 /* Interrupt-Controller base addresses */
# define CLKDIVN 0x14800014 /* clock divisor register */
#elif defined(CONFIG_S3C2410) || defined(CONFIG_S3C2440)❶
# define pWTCON 0x53000000
# define INTMOD 0x4A000004
# define INTMSK 0x4A000008 /* Interrupt-Controller base addresses */
# define INTSUBMSK 0x4A00001C
# define CLKDIVN 0x4C000014 /* clock divisor register */
#endif
```

❶ 上面几个宏定义所对应的地址，都可以在对应的datasheet中找到对应的定义：

其中，S3C2410和TQ2440开发板所用的CPU S3C2440，两者在这部分的寄存器定义，都是一样的，所以此处，采用CONFIG_S3C2410所对应的定义。

关于S3C2440相关的软硬件资料，这个网站提供的很全面：

<http://just4you.springnote.com/pages/1052612>

其中有S3C2440的CPU的datasheet：

s3c2440a_um_rev014_040712.pdf

其中有对应的寄存器定义：

图 1.3. pWTCON

Watchdog Timer					
WTCON	0x53000000	←	W	R/W	Watchdog Timer Mode
WTDAT	0x53000004				Watchdog Timer Data
WTCNT	0x53000008				Watchdog Timer Count

图 1.4. INTMOD

Register	Address	R/W	Description	Reset Value
INTMOD	0x4A000004	R/W	Interrupt mode register. 0 = IRQ mode 1 = FIQ mode	0x00000000

图 1.5. INTMSK

Register	Address	R/W	Description	Reset Value
INTMSK	0X4A000008	R/W	Determine which interrupt source is masked. The masked interrupt source will not be serviced. 0 = Interrupt service is available. 1 = Interrupt service is masked.	0xFFFFFFFF

图 1.6. INTSUBMSK

Register	Address	R/W	Description	Reset Value
INTSUBMSK	0X4A00001C	R/W	Determine which interrupt source is masked. The masked interrupt source will not be serviced. 0 = Interrupt service is available. 1 = Interrupt service is masked.	0xFFFF

图 1.7. CLKDIVN

Register Name	Address (B. Endian)	Address (L. Endian)	Acc. Unit	Read/Write	Function
Clock & Power Management					
LOCKTIME	0x4C000000	←	W	R/W	PLL Lock Time Counter
MPLLCON	0x4C000004				MPLL Control
UPLLCON	0x4C000008				UPLL Control
CLKCON	0x4C00000C				Clock Generator Control
CLKSLOW	0x4C000010				Slow Clock Control
CLKDIVN	0x4C000014				Clock divider Control
CAMDIVN	0x4C000018				Camera Clock divider Control

而关于每个寄存器的具体含义，见后面的分析。

1.2.2. ldr pWTCON

```
#if defined(CONFIG_S3C2400) || defined(CONFIG_S3C2410) || defined(CONFIG_S3C2440)
ldr① r0, =pWTCON
```

- ❶ 这里的ldr和前面介绍的ldr指令不是一个意思。

这里的ldr是伪指令ldr。



伪指令

伪指令，就是“伪”的指令，是针对“真”的指令而言的。

真的指令就是那些常见的指令，比如上面说的arm的ldr, bic, msr等等指令，是arm体系架构中真正存在的指令，你在arm汇编指令集中找到对应的含义。

而伪指令是写出来给汇编程序看的，汇编程序能看的伪指令具体表示的是啥意思，然后将其翻译成真正的指令或者进行相应的处理。

伪指令ldr语法和含义：

<http://blog.csdn.net/lihaoweiV/archive/2010/11/24/6033003.aspx>

另外还有一个就是ldr伪指令，虽然ldr伪指令和ARM的ldr指令很像，但是作用不太一样。ldr伪指令可以在立即数前加上=，以表示把一个地址写到某寄存器中，比如：

```
ldr r0, =0x12345678
```

这样，就把0x12345678这个地址写到r0中了。所以，ldr伪指令和mov是比较相似的。

只不过mov指令后面的立即数是有限制的，这个立即数，能够必须由一个8位的二进制数，即属于0x00-0xFF内的某个值，经过偶数次右移后得到，这样才是合法数据，而ldr伪指令没有这个限制。

那为何ldr伪指令的操作数没有限制呢，那是因为其是伪指令，写出来的伪指令，最终会被编译器解释成为真正的，合法的指令的，一般都是对应的mov指令。

这样的话，写汇编程序的时候，使用MOV指令是比较麻烦的，因为有些简单的数据比较容易看出来，有些数据即不容易看出来是否是合法数据。所以，对此，ldr伪指令的出现，就是为了解决这个问题的，你只管放心用ldr伪指令，不用关心操作数，而写出的ldr伪指令，编译器会帮你翻译成对应的真正的汇编指令的。

而关于编译器是如何将这些ldr伪指令翻译成为真正的汇编指令的，我的理解是，其会自动会去算出来对应的操作数，是否是合法的mov的操作数，如果是，就将该ldr伪指令翻译成mov指令，否则就用别的方式处理，我所观察到的，其中一种方式就是，单独申请一个4字节的空间用于存放操作数，然后用ldr指令实现。

在uboot中，最后make完毕之后，会生产u-boot，

通过：

```
arm-linux-objdump -d u-boot > dump_u-boot.txt
```

就可以把对应的汇编代码输出到该txt文件了，其中就能找到伪指令：

```
ldr    r0, =0x53000000
```

所对应的，真正的汇编代码：

```
33d00068: e3a00453  mov r0, #1392508928 ; 0x53000000
```

所以被翻译成了mov指令。

而经过我的尝试，故意将0x53000000改为0x53000010，对应的生产的汇编代码为：

```
33d00068: e59f0408  ldr r0, [pc, #1032] ; 33d00478 <fiq+0x58>
.....
33d00478: 53000010  .word 0x53000010
```

其中可以看到，由于0x53000010不是有效的mov的操作数，没法找到合适的0x00-0Xff去通过偶数次循环右移而得到，所以只能换成此处这种方式，即在另外申请一个word的空间用于存放这个值：

```
33d00478: 53000010  .word 0x53000010
```

然后通过计算出相对当前PC的偏移，得到的地址，用ldr指令去除该地址中的值，即0x53000010，送给r0,比起mov指令，要复杂的多，也多消耗了一个word的空间。

对应地，其他的方式，个人理解，好像也可以通过MVN指令来实现，具体细节，有待进一步探索。

而这里的：

```
ldr    r0, =pWTCON
```

意思就很清楚了，就是把宏pWTCON的值赋值给r0寄存器，即

r0=0x53000000

1.2.3. mov

```
mov①   r1, #0x0
```

① mov指令语法：

1、MOV指令

MOV指令的格式为：

MOV{条件}{S} 目的寄存器，源操作数

MOV指令可完成从另一个寄存器、被移位的寄存器或将一个立即数加载到目的寄存器。其中S选项决定指令的操作是否影响CPSR中条件标志位的值，当没有S时指令不更新CPSR中条件标志位的值。

指令示例：

MOV R1, R0 ; 将寄存器R0的值传送到寄存器R1

MOV PC, R14 ; 将寄存器R14的值传送到PC，常用于子程序返回

MOV R1, R0, LSL # 3 ; 将寄存器R0的值左移3位后传送到R1

不过对于MOV指令多说一句，那就是，一般可以用类似于：

```
MOV R0, R0
```

的指令来实现NOP操作。

上面这句mov指令很简单，就是把0x0赋值给r1，即

r1=0x0

1.2.4. str

```
str①   r1, [r0]
```

① str指令语法：

4、STR指令

STR指令的格式为：

STR{条件} 源寄存器，<存储器地址>

STR指令用于从源寄存器中将一个32位的字数据传送到存储器中。该指令在程序设计中比较常用，且寻址方式灵活多样，使用方式可参考指令LDR。

指令示例：

STR R0, [R1], # 8 ; 将R0中的字数据写入以R1为地址的存储器中，并将新地址R1 + 8写入R1。

STR R0, [R1, # 8] ; 将R0中的字数据写入以R1 + 8为地址的存储器中。

所以这句str的作用也很简单，那就是将r1寄存器的值，传送到地址值为r0的（存储器）内存中。

用C语言表示就是：

```
*r0 = r1
```

所以，上面几行代码意思也很清楚：

先用r0寄存器存pWTCN的值，然后r1=0，再将r1中的0写入到pWTCN中，其实就是

```
pWTCN = 0;
```

而pWTCN寄存器的具体含义是什么呢？下面就了解其详细含义：

图 1.8. WTCN寄存器的位域

WTCN	Bit	Description	Initial State
Prescaler value	[15:8]	Prescaler value. The valid range is from 0 to 255(2^8-1).	0x80
Reserved	[7:6]	Reserved. These two bits must be 00 in normal operation.	00
Watchdog timer	[5]	Enable or disable bit of Watchdog timer. 0 = Disable 1 = Enable	1
Clock select	[4:3]	Determine the clock division factor. 00: 16 01 : 32 10: 64 11 : 128	00
Interrupt generation	[2]	Enable or disable bit of the interrupt. 0 = Disable 1 = Enable	0
Reserved	[1]	Reserved. This bit must be 0 in normal operation.	0
Reset enable/disable	[0]	Enable or disable bit of Watchdog timer output for reset signal. 1: Assert reset signal of the S3C2440A at watchdog time-out 0: Disable the reset function of the watchdog timer.	1

注意到bit[0]是Reset Enable/Disable，而设置为0的话，那就是关闭Watchdog的reset了，所以其他位的配置选项，就更不需要看了。

我们只需要了解，在此处禁止了看门狗WatchDog（的复位功能），即可。

关于看门狗的作用，以及为何要在系统初始化的时候关闭看门狗，请参见本文档后面的章节：[第 3.3 节 “什么是watchdog + 为何在要系统初始化的时候关闭watchdog”](#)

1.3. 关闭中断

1.3.1. set INTMSK

```
/*  
 * mask all IRQs by setting all bits in the INTMR - default  
 */  
mov r1, #0xffffffff  
ldr r0, =INTMSK  
str r1, [r0]❶
```

- ❶ 上面这几行代码，和前面的很类似，作用很简单，就是将INTMSK寄存器设置为0xffffffff,即，将所有的中端都mask了。

关于每一位的定义，其实可以不看的，反正此处都已mask了，不过还是贴出来，以备后用：

图 1.9. INTMSK寄存器的位域

INTMSK	Bit	Description	Initial State
INT_ADC	[31]	0 = Service available, 1 = Masked	1
INT_RTC	[30]	0 = Service available, 1 = Masked	1
INT_SPI1	[29]	0 = Service available, 1 = Masked	1
INT_UART0	[28]	0 = Service available, 1 = Masked	1
INT_IIC	[27]	0 = Service available, 1 = Masked	1
INT_USBH	[26]	0 = Service available, 1 = Masked	1
INT_USBD	[25]	0 = Service available, 1 = Masked	1
INT_NFCON	[24]	0 = Service available, 1 = Masked	1
INT_UART1	[23]	0 = Service available, 1 = Masked	1
INT_SPI0	[22]	0 = Service available, 1 = Masked	1
INT_SDI	[21]	0 = Service available, 1 = Masked	1
INT_DMA3	[20]	0 = Service available, 1 = Masked	1
INT_DMA2	[19]	0 = Service available, 1 = Masked	1
INT_DMA1	[18]	0 = Service available, 1 = Masked	1
INT_DMA0	[17]	0 = Service available, 1 = Masked	1
INT_LCD	[16]	0 = Service available, 1 = Masked	1
INT_UART2	[15]	0 = Service available, 1 = Masked	1
INT_TIMER4	[14]	0 = Service available, 1 = Masked	1
INT_TIMER3	[13]	0 = Service available, 1 = Masked	1
INT_TIMER2	[12]	0 = Service available, 1 = Masked	1
INT_TIMER1	[11]	0 = Service available, 1 = Masked	1
INT_TIMER0	[10]	0 = Service available, 1 = Masked	1
INT_WDT_AC97	[9]	0 = Service available, 1 = Masked	1
INT_TICK	[8]	0 = Service available, 1 = Masked	1
nBATT_FLT	[7]	0 = Service available, 1 = Masked	1
INT_CAM	[6]	0 = Service available, 1 = Masked	1
EINT8_23	[5]	0 = Service available, 1 = Masked	1
EINT4_7	[4]	0 = Service available, 1 = Masked	1
EINT3	[3]	0 = Service available, 1 = Masked	1
EINT2	[2]	0 = Service available, 1 = Masked	1
EINT1	[1]	0 = Service available, 1 = Masked	1
EINT0	[0]	0 = Service available, 1 = Masked	1

此处，关于mask这个词，解释一下。

mask这个单词，是面具的意思，而中断被mask了，就是中断被掩盖了，即虽然硬件上中断发生了，但是此处被屏蔽了，所以从效果上来说，就相当于中断被禁止了，硬件上即使发生了中断，CPU也不会去执行对应中断服务程序ISR了。

关于中断的内容的详细解释，推荐看这个，解释的很通俗易懂：[【转】ARM9 2410移植之ARM中断原理, 中断嵌套的误区, 中断号的怎么来的²](http://www.crifan.com/switch_arm9_2410_transplant_arm_interrupt_principle_the_error_interrupt_nesting_how_come_the_interrupt_number/)

1.3.2. set INTSUBMSK

²

http://www.crifan.com/switch_arm9_2410_transplant_arm_interrupt_principle_the_error_interrupt_nesting_how_come_the_interrupt_number/

```
# if defined(CONFIG_S3C2410)
ldr r1, =0x3ff❶
ldr r0, =INTSUBMSK
str r1, [r0]
# elif defined(CONFIG_S3C2440)
ldr r1, =0x7fff❷
ldr r0, =INTSUBMSK
str r1, [r0]
# endif
```

- ❶ 此处是将2410的INTSUBMSK设置为0x3ff。

后经HateMath³的提醒后，去查证，的确此处设置的0x3ff，是不严谨的。

因为，根据2410的datasheet中关于INTSUBMSK的解释，bit[10:0]共11位，虽然默认reset的每一位都是1，但是此处对应的mask值，应该是11位全为1=0x7ff。

即写成0x3ff，虽然是可以正常工作的，但是却不够严谨的。

- ❷ 此处CPU是S3C2440，所以用到0x7fff这段代码。

其意思也很容易看懂，就是将INTSUBMSK寄存器的值设置为0x7fff。

先贴出对应每一位的含义：

图 1.10. INTSUBMSK寄存器的位域

INTSUBMSK	Bit	Description	Initial State
Reserved	[31:15]	Not used	0
INT_AC97	[14]	0 = Service available, 1 = Masked	1
INT_WDT	[13]	0 = Service available, 1 = Masked	1
INT_CAM_P	[12]	0 = Service available, 1 = Masked	1
INT_CAM_C	[11]	0 = Service available, 1 = Masked	1
INT_ADC_S	[10]	0 = Service available, 1 = Masked	1
INT_TC	[9]	0 = Service available, 1 = Masked	1
INT_ERR2	[8]	0 = Service available, 1 = Masked	1
INT_TXD2	[7]	0 = Service available, 1 = Masked	1
INT_RXD2	[6]	0 = Service available, 1 = Masked	1
INT_ERR1	[5]	0 = Service available, 1 = Masked	1
INT_TXD1	[4]	0 = Service available, 1 = Masked	1
INT_RXD1	[3]	0 = Service available, 1 = Masked	1
INT_ERR0	[2]	0 = Service available, 1 = Masked	1
INT_TXD0	[1]	0 = Service available, 1 = Masked	1
INT_RXD0	[0]	0 = Service available, 1 = Masked	1

然后我们再来分析对应的0x7fff是啥含义。

其实也很简单，意思就是：

0x7fff = bit[14:0]全是1 = 上表中的全部中断都被屏蔽 (mask) 。

1.3.3. set CLKDIVN

³ hatemath2005@163.com

```

#if 0
/* FCLK:HCLK:PCLK = 1:2:4 */
/* default FCLK is 120 MHz ! */
ldr r0, =CLKDIVN❶
mov r1, #3
str r1, [r0]
#endif
#endif /* CONFIG_S3C2400 || CONFIG_S3C2410 || CONFIG_S3C2440 */❷

```

❶ 此处，关于CLKDIVN的具体含义，参见下表：

图 1.11. INTSUBMSK寄存器的位域

CLKDIVN	Bit	Description	Initial State
DIVN_UPLL	[3]	UCLK select register(UCLK must be 48MHz for USB) 0: UCLK = UPLL clock 1: UCLK = UPLL clock / 2 Set to 0, when UPLL clock is set as 48Mhz Set to 1. when UPLL clock is set as 96Mhz.	0
HDIVN	[2:1]	00 : HCLK = FCLK/1. 01 : HCLK = FCLK/2. 10 : HCLK = FCLK/4 when CAMDIVN[9] = 0. HCLK= FCLK/8 when CAMDIVN[9] = 1. 11 : HCLK = FCLK/3 when CAMDIVN[8] = 0. HCLK = FCLK/6 when CAMDIVN[8] = 1.	00
PDIVN	[0]	0: PCLK has the clock same as the HCLK/1. 1: PCLK has the clock same as the HCLK/2.	0

而此处代码被#if 0注释掉了。

问：为何要注释掉，难道想要使用其默认的值，即HDIVN和PDIVN上电后，默认值Initial State，都是0，对应的含义为，FCLK:HCLK:PCLK = 1:1:1 ???

答：不是，是因为我们在其他地方会去初始化时钟，去设置对应的CLKDIVN，详情参考后面的代码第 1.4.3 节 “clock_init” 的部分。

❷ 此处是结束上面的#endif

1.3.4. bl

```

/*
 * we do sys-critical inits only at reboot,
 * not when booting from ram!
 */
#ifndef CONFIG_SKIP_LOWLEVEL_INIT
bl❶ cpu_init_crit
#endif

```

❶ 关于bl指令的含义：

b指令，是单纯的跳转指令，即CPU直接跳转到某地址继续执行。

而BL是Branch with Link，带分支的跳转，而Link指的是Link Register，链接寄存器R14，即lr，所以，bl的含义是，除了包含b指令的单纯的跳转功能，在跳转之前，还把r15寄存器=PC=cpu地址，赋值给r14=lr，然后跳转到对应位置，等要做的事情执行完毕之后，再用

```
mov pc, lr
```

使得cpu再跳转回来，所以整个逻辑就是调用子程序的意思。

bl的语法为：

2、BL指令

BL指令的格式为：

BL(条件) 目标地址

BL 是另一个跳转指令，但跳转之前，会在寄存器R14中保存PC的当前内容，因此，可以通过将R14 的内容重新加载到PC中，来返回到跳转指令之后的那个指令处执行。该指令是实现子程序调用的一个基本但常用的手段。以下指令：

BL Label ；当程序无条件跳转到标号Label处执行时，同时将当前的PC值保存到R14中

对于上面的代码来说，意思就很清晰了，就是当没有定义CONFIG_SKIP_LOWLEVEL_INIT的时候，就掉转到cpu_init_crit的位置，而在后面的代码cpu_init_crit中，你可以看到最后一行汇编代码就是

```
mov pc, lr ,
```

又将PC跳转回来，所以整个的含义就是，调用子程序cpu_init_crit,等cpu_init_crit执行完毕，再返回此处继续执行下面的代码。

于此对应地b指令，就只是单纯的掉转到某处继续执行，而不能再通过mov pc, lr跳转回来了。

1.4. 设置堆栈sp指针

1.4.1. stack_setup

```
/* Set up the stack      */
stack_setup:
ldr r0, _TEXT_BASE /* upper 128 KiB: relocated uboot */❶
sub r0, r0, #CFG_MALLOC_LEN /* malloc area */
sub r0, r0, #CFG_GBL_DATA_SIZE /* bdfinfo */
```

- ❶ 此句含义是，把地址为_TEXT_BASE的内存中的内容给r0,即，将所有的中断都mask了。

而查看前面的相关部分的代码，即：

```
_TEXT_BASE:
.word TEXT_BASE
```

得知，地址为_TEXT_BASE的内存中的内容，就是

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\board\EmbedSky\config.mk

中的：

```
TEXT_BASE = 0x33D00000
```


所以，此处即：

```
r0
= TEXT_BASE
= 0x33D00000
```

而关于sub指令：

SUB：减法

(Subtraction)

SUB{条件}{S} <dest>, <op 1>, <op 2>

dest = op_1 - op_2

SUB 用操作数 one 减去操作数 two，把结果放置到目的寄存器中。操作数 1 是一个寄存器，操作数 2 可以是一个寄存器，被移位的寄存器，或一个立即值：

SUB R0, R1, R2 ; R0 = R1 - R2

SUB R0, R1, #256 ; R0 = R1 - 256

SUB R0, R2, R3, LSL#1 ; R0 = R2 - (R3 << 1)

减法可以在有符号和无符号数上进行。

所以对应含义为：

r0 = r0 - #CFG_MALLOC_LEN

r0 = r0 - #CFG_GBL_DATA_SIZE

其中，对应的两个宏的值是：

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\include\configs\EmbedSky.h

中：

```
#define CONFIG_64MB_Nand 0 //添加了对64MB Nand Flash支持

/*
 * Size of malloc() pool
 */
#define CFG_MALLOC_LEN (CFG_ENV_SIZE + 128*1024)
#define CFG_GBL_DATA_SIZE 128 /* size in bytes reserved for initial data */

#if(CONFIG_64MB_Nand == 1)
#define CFG_ENV_SIZE 0xc000 /* Total Size of Environment Sector */
#else
#define CFG_ENV_SIZE 0x20000 /* Total Size of Environment Sector */
#endif
```

所以，从源码中的宏定义中可以看出，

CFG_MALLOC_LEN

```

= (CFG_ENV_SIZE + 128*1024)
= 0x20000 + 128*1024
= 0x40000
= 256*1024
= 256KB

```

```

CFG_GBL_DATA_SIZE
= 128

```

所以，此三行的含义就是算出r0的值：

```

r0
= (r0 - #CFG_MALLOC_LEN) - #CFG_GBL_DATA_SIZE
= r0 - 0x40000 - 128
= r0 - 0x40080
= 33CBFF80

```

1.4.2. calc stack

```

#ifdef CONFIG_USE_IRQ
sub r0, r0, #(CONFIG_STACKSIZE_IRQ+CONFIG_STACKSIZE_FIQ)❶
#endif
sub sp, r0, #12❷ /* leave 3 words for abort-stack */

```

- ❶ 如果定义了CONFIG_USE_IRQ，即如果使用中断的话，那么再把r0的值减去IRQ和FIQ的堆栈的值，

而对应的宏的值也是在

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\include\configs\EmbedSky.h

中：

```

/*-----
 * Stack sizes
 *
 * The stack sizes are set up in start.S using the settings below
 */
#define CONFIG_STACKSIZE (128*1024) /* regular stack */
#ifdef CONFIG_USE_IRQ
#define CONFIG_STACKSIZE_IRQ (4*1024) /* IRQ stack */
#define CONFIG_STACKSIZE_FIQ (4*1024) /* FIQ stack */
#endif

```

所以，此时r0的值就是：

```
#ifdef CONFIG_USE_IRQ
```

```
r0
```

```
= r0 - #(CONFIG_STACKSIZE_IRQ+CONFIG_STACKSIZE_FIQ)
```

```
= r0 - (4*1024 + 4*1024)
```

```
= r0 - 8*1024
```

```
= 33CBFF80 - 8*1024
```

```
= 33CBDF80
```

```
#endif
```

- ❷ 最后，再减去终止异常所用到的堆栈大小，即12个字节。

现在r0的值为：

```
#ifdef CONFIG_USE_IRQ
```

```
r0
```

```
= r0 - 12
```

```
= 33CBDF80 - 12
```

```
= 33CBDF74
```

```
#else
```

```
r0
```

```
= r0 - 12
```

```
= 33CBFF80 - 12
```

```
= 33CBFF74
```

```
#endif
```

然后将r0的值赋值给sp,即堆栈指针。

关于：

sp代表stack pointer，堆栈指针；

和后面要提到的ip寄存器：

ip代表instruction pointer，指令指针。

更多详情参见下面的解释。

关于ARM的寄存器的别名和相关的APCS，参见本文后面的内容：[第 3.5 节 “AMR寄存器的别名 + APCS”](#)

1.4.3. clock_init

```
bl clock_init❶
```

- ❶ 在上面，经过计算，算出了堆栈的地址，然后赋值给了sp，此处，接着才去调用函数clock_init去初始化时钟。

其中此函数是在C文件：

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\board\EmbedSky\boot_init.c

中：

```
void clock_init(void)
{
    ...设置系统时钟clock的相关代码...
}
```

看到这里，让我想起，关于其他人的关于此start.S代码解释中说到的，此处是先去设置好堆栈，即初始化sp指针，然后才去调用C语言的函数clock_init的。

而我们可以看到，前面那行代码：

```
#ifndef CONFIG_SKIP_LOWLEVEL_INIT
    bl cpu_init_crit
#endif
```

就不需要先设置好堆栈，再去进行函数调用。

其中cpu_init_crit对应的代码也在start.S中（详见后面对应部分的代码），是用汇编实现的。

而对于C语言，为何就需要堆栈，而汇编却不需要堆栈的原因，请参见本文后面的内容：[第 3.6 节 “为何C语言（的函数调用）需要堆栈，而汇编语言却不需要堆栈”](#)

1.4.4. adr

```
#ifndef CONFIG_SKIP_RELOCATE_UBOOT
relocate: /* relocate U-Boot to RAM */
    adr❶ r0, _start /* r0 <- current position of code */
```

- ❶ adr指令的语法和含义：

http://blog.mcuol.com/User/cdkfGao/article/8057_1.htm

1、ADR伪指令--- 小范围的地址读取

ADR伪指令将基于PC相对偏移的地址值或基于寄存器相对偏移的地址值读取到寄存器中。

在汇编编译器编译源程序时，ADR伪指令被编译器替换成一条合适的指令。通常，编译器用一条ADD指令或SUB指令来实现该ADR伪指令的功能，若不能用一条指令实现，则产生错误，编译失败。

ADR伪指令格式：**ADR{cond} register, expr**

地址表达式expr的取值范围：

当地址值是字节对齐时，其取值范围为: +255 ~ 255B；

当地址值是字对齐时，其取指范围为: -1020 ~ 1020B；

所以，上述:

```
adr r0, _start
```

的意思其实很简单，就是将_start的地址赋值给r0.但是具体实现的方式就有点复杂了，对于用adr指令实现的话，说明_start这个地址，相对当前PC的偏移，应该是很小的，意思就是向前一段后者向后一段去找，肯定能找到_start这个标号地址的，此处，自己通过看代码也可以看到_start，就是在当前指令的前面，距离很近，编译后，对应汇编代码，也可以猜得出，应该是上面所说的，用sub来实现，即当前PC减去某个值，得到_start的值，

参照前面介绍的内容，去：

```
arm-inux-objdump -d u-boot > dump_u-boot.txt
```

然后打开dump_u-boot.txt,可以找到对应的汇编代码，如下：

```
33d00000 <_start>:
33d00000: ea000014 b 33d00058 <reset>
      ° ° °
33d000a4 <relocate>:
33d000a4: e24f00ac sub r0, pc, #172 ; 0xac
```

可以看到，这个相对当前PC的距离是0xac=172,细心的读者可以看到，那条指令的地址减去0xac，却并不等于_start的值，即

$$33d000a4 - 33d00000 = 0xa4 \neq 0xac$$

而 $0xac - 0xa4 = 8$ ，

那是因为，由于ARM920T的五级流水线的缘故导致指令执行那一时刻的PC的值等于该条指令PC的值加上8，即

sub r0, pc, #172中的PC的值是

sub r0, pc, #172

指令地址：33d000a4,再加上8，即 $33d000a4 + 8 = 33d000ac$ ，

所以， $33d000ac - 0xac$ ，才等于我们看到的33d00000，才是_start的地址。

这个由于流水线导致的PC的值和当前指令地址不同的现象，就是我们常说的，ARM中， $PC = PC + 8$ 。

对于为何是 $PC = PC + 8$ ，请参见后面的内容：[第 3.4 节 “为何ARM7中 \$PC = PC + 8\$ ”](#)

对于此处为何不直接用mov指令，却使用adr指令，请参见后面内容：[第 3.7 节 “关于为何不直接用mov指令，而非要用adr伪指令”](#)

对于mov指令的操作数的取值范围，请参见后面内容：[第 3.8 节 “mov指令的操作数的取值范围到底是多少”](#)

```
adr r0, _start
```

的伪代码，被翻译成实际汇编代码为：

```
33d000a4: e24f00ac sub r0, pc, #172 ; 0xac
```

其含义就是，通过计算PC+8-172 # _start的地址，

而_start的地址，即相对代码段的0地址，是这个地址在运行时刻的值，而当ARM920T加电启动后，，此处是从Nor Flash启动，对应的代码，也是在Nor Flash中，对应的物理地址是0x0,所以，此时_start的值就是0，而不是0x33d00000。

所以，此时：

`r0 = 0x0`

1.4.5. clear_bss

```
ldr r1, _TEXT_BASE① /* test if we run from flash or RAM */
②cmp    r0, r1      /* don't reloc during debug    */
③beq    clear_bss
```

① 这里的_TEXT_BASE的含义，前面已经说过了，那就是：

```
_TEXT_BASE:
.word TEXT_BASE
```

得知，地址为_TEXT_BASE的内存中的内容，就是

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\board\EmbedSky\config.mk

中的：

```
TEXT_BASE = 0x33D00000
```

所以，此处就是

`r1 = 0x33D00000`

② 含义很简单，就是比较r0和r1。而

`r0 = 0x0`

`r1 = 0x33D00000`

所以不相等

③ 因此beq发现两者不相等，就不会去跳转到clear_bss，不会去执行对应的将bss段清零的动作了。

1.4.6. cal armboot size

```
①ldr r2, _armboot_start
ldr r3, _bss_start
②sub r2, r3, r2 /* r2 <- size of armboot    */
```

① 这两行代码意思也很清楚，分别装载_armboot_start和_bss_start地址中的值，赋值给r2和r3
而_armboot_start和_bss_start的值，前面都已经提到过了，就是：

```
.globl _armboot_start
_armboot_start:
```

```
.word _start

.globl _bss_start
_bss_start:
.word __bss_start
```

```
_TEXT_BASE:
.word TEXT_BASE
```

而其中的_start，是我们uboot的代码的最开始的位置，而__bss_start的值，是在u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\board\EmbedSky\u-boot.lds中的：

```
SECTIONS
{
. = 0x00000000;

. = ALIGN(4);
.text :

...
. = ALIGN(4);
.rodata : { *(.rodata) }

. = ALIGN(4);
.data : { *(.data) }

...
. = ALIGN(4);
__bss_start = .;
.bss : { *(.bss) }
_end = .;
}
```

所以，可以看出，__bss_start的位置，是bss的start开始位置，同时也是text+rodata+data的结束位置，即代码段，只读数据和已初始化的可写的的最末尾的位置。

其实我们也可以通过前面的方法，objdump出来，看到对应的值：

```
33d00048 <_bss_start>:
33d00048: 33d339d4 .word 0x33d339d4
```

是0x33d339d4。



注意

【总结】

$r2 = _start = 0x33d00000$

$r3 = __bss_start = 0x33d339d4$

- ② 此处的意思就很清楚了，就是 $r2 = r3 - r2$ ，计算出

text + rodata + data

的大小，即整个需要载入的数据量是多少，用于下面的函数去拷贝这么多的数据到对应的内存的位置。

这里的实际的值是

$r2$

$= r3 - r2$

$= 0x33d339d4 - 0x33d00000$

$= 0x000339d4$



注意

【总结】

到此刻位置,假定是从Nor Flash启动的话：

$r0 = 0x0$ = 我们代码此刻所在的位置

$r1 = 0x33D00000$ = 我们想要把我们的代码放到哪里

$r2 = 0x000339d4$ = 对应的代码的大小 (此处的代码 = text + rodata + data)

1.4.7. cal armboot size

```
#if 1
❶ bl CopyCode2Ram /* r0: source, r1: dest, r2: size */
#else
add r2, r0, r2 /* r2 <- source end address */

copy_loop:
ldmia r0!, {r3-r10} /* copy from source address [r0] */
stmia r1!, {r3-r10} /* copy to target address [r1] */
cmp r0, r2 /* until source end addreee [r2] */
ble copy_loop
#endif
#endif /* CONFIG_SKIP_RELOCATE_UBOOT */
```

- ❶ 此处，代码很简单，只是注释掉了原先的那些代码，而单纯的只是去调用CopyCode2Ram这个函数。

CopyCode2Ram函数，前面也提到过了，是在：

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\board\EmbedSky\boot_init.c

中：

```
int CopyCode2Ram(unsigned long start_addr, unsigned char *buf, int size)
{
    unsigned int *pdwDest;
    unsigned int *pdwSrc;
```



```

int i;

if (bBootFrmNORFlash())
{
    pdwDest = (unsigned int *)buf;
    pdwSrc = (unsigned int *)start_addr;
    /* 从 NOR Flash启动 */
    for (i = 0; i < size / 4; i++)
    {
        pdwDest[i] = pdwSrc[i];
    }
    return 0;
}
else
{
    /* 初始化NAND Flash */
    nand_init_ll();

    /* 从 NAND Flash启动 */
    if (NF_ReadID() == 0x76 )
        nand_read_ll(buf, start_addr, (size + NAND_BLOCK_MASK)&~(NAND_BLOCK_MASK));
    else
        nand_read_ll_lp(buf, start_addr, (size +
        NAND_BLOCK_MASK_LP)&~(NAND_BLOCK_MASK_LP));
    return 0;
}
}

```

可以看到，其有三个参数，start_addr，*buf和size，这三个参数，分别正好对应着我们刚才所总结的r0,r1和r2。

这些寄存器和参数的对应关系，也是APCS中定义的：

实际参数

APCS 没有定义记录、数组、和类似的格局。这样语言可以自由的定义如何进行这些活动。但是，如果你自己的实现实际上不符合 APCS 的精神，那么将不允许来自你的编译器的代码与来自其他编译器的代码连接在一起。典型的，使用 C 语言的惯例。

- 前4个整数实参(或者更少!)被装载到 a1 - a4
- 前 4 个浮点实参(或者更少!)被装载到 f0 - f3
- 其他任何实参(如果有的话)存储在内存中，用进入函数时紧接在 sp 的值上面的字来指向。换句话说，其余的参数被压入栈顶。所以要想简单。最好定义接受 4 个或更少的参数的函数

上面说的a1-a4,就是寄存器r0-r3。

而CopyCode2Ram函数的逻辑也很清晰，就是先去判断是从Nor Flash启动还是从Nand Flash启动，然后决定从哪里拷贝所需要的代码到对应的目标地址中。

1.5. 清除bss段

1.5.1. clear_bss

```
clear_bss:①
ldr r0, _bss_start② /* find start of bss segment      */
ldr r1, _bss_end③ /* stop here                        */
mov r2, #0x00000000 /* clear                      */
```

② 此处的_bss_start是：

```
.globl _bss_start
_bss_start:
.word _bss_start
```

③ 而_bss_end,是：

```
.globl _bss_end
_bss_end:
.word _end
```

① 对应的，__bss_start和_end，都在前面提到过的那个链接脚本里面：
u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\board\EmbedSky\u-boot.lds

中的：

```
__bss_start = .;
.bss : { *(.bss) }
_end = .;
```

即bss段的起始地址和结束地址。

1.5.2. clear css loop

```
① clbss_l: str r2, [r0] /* clear loop...      */
add r0, r0, #4
cmp r0, r1
ble clbss_l
```

① 此段代码含义也很清晰，那就是，

先将r2,即0x0，存到地址为r0的内存中去，然后r0地址加上4，比较r0地址和r1地址，即比较当前地址是否到了bss段的结束位置，如果le, little or equal, 小于或等于，那么就跳到clbss_l,即接着这几个步骤，直到地址超过了bss的_end位置，即实现了将整个bss段，都清零。

1.5.3. ldr pc

```

#if 0❶
/* try doing this stuff after the relocation */
ldr    r0, =pWTCON
mov    r1, #0x0
str    r1, [r0]

/*
 * mask all IRQs by setting all bits in the INTMR - default
 */
mov r1, #0xffffffff
ldr r0, =INTMR
str r1, [r0]

/* FCLK:HCLK:PCLK = 1:2:4 */
/* default FCLK is 120 MHz ! */
ldr r0, =CLKDIVN
mov r1, #3
str r1, [r0]
/* END stuff after relocation */
#endif

ldr pc, _start_armboot
❷
_start_armboot: .word start_armboot

```

- ❶ 此处忽略已经注释掉的代码
- ❷ 最后的那两行，意思也很简单，那就是将地址为_start_armboot中的内容，即

start_armboot，赋值给PC，即调用start_armboot函数。

至此，汇编语言的start.S的整个工作，就完成了。

而start_armboot函数，在C文件中：

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\board\EmbedSky\EmbedSky.c
中：

```

void start_armboot (void)
{
    .....
}

```

这就是传说中的，调用第二层次，即C语言级别的初始化了，去初始化各个设备了。

其中包括了CPU，内存等，以及串口，正常初始化后，就可以从串口看到uboot的打印信息了。

1.5.4. cpu_init_crit

```

/*
*****
*
* CPU_init_critical registers
*
* setup important registers

```

```

* setup memory timing
*
*****
*/
#ifndef CONFIG_SKIP_LOWLEVEL_INIT
cpu_init_crit:
/*
* flush v4 I/D caches
*/
mov r0, #0
mcr❶ p15, 0, r0, c7, c7, 0 /* flush v3/v4❷ cache */
mcr p15, 0, r0, c8, c7, 0 /* flush v4 TLB */❸

```

❶ 关于mcr的来龙去脉：

<http://apps.hi.baidu.com/share/detail/32319228>

ARM 微处理器可支持多达 16 个协处理器，用于各种协处理操作，在程序执行的过程中，每个协处理器只执行针对自身的协处理指令，忽略 ARM 处理器和其他协处理器的指令。ARM 的协处理器指令主要用于 ARM 处理器初始化 ARM 协处理器的数据处理操作，以及在 ARM 处理器的寄存器和协处理器的寄存器之间传送数据，和在 ARM 协处理器的寄存器和存储器之间传送数据。ARM 协处理器指令包括以下 5 条：

1. CDP 协处理器数操作指令
2. LDC 协处理器数据加载指令
3. STC 协处理器数据存储指令
4. MCR ARM 处理器寄存器到协处理器寄存器的数据传送指令
5. MRC 协处理器寄存器到ARM 处理器寄存器的数据传送指令

.....

CP15系统控制协处理器

CP15 —系统控制协处理器（ the system control coprocessor ）他通过协处理器指令MCR和MRC提供具体的寄存器来配置和控制caches、MMU、保护系统、配置时钟模式（在bootloader时钟初始化用到）

CP15的寄存器只能被MRC和MCR（ Move to Coprocessor from ARM Register ）指令访问

一些要说明的内容，见下：：

http://infocenter.arm.com/help/topic/com.arm.doc.ddi0151c/ARM920T_TRM1_S.pdf

you can only access CP15 registers with MRC and MCR instructions in a privileged mode. The assembler for these instructions is:

MCR/MRC{cond} P15,opcode_1,Rd,CRn,CRm,opcode_2

The CRn field of MRC and MCR

instructions specifies the coprocessor register to access. The CRm field and opcode_2 fields specify a particular action when addressing registers. The L bit distinguishes between an MRC (L=1) and an MCR (L=0).

Note:

Attempting to read from a nonreadable register, or to write to a nonwritable register causes unpredictable results.

The opcode_1, opcode_2, and CRm fields should be zero, except when the values specified are used to select the desired operations, in all instructions that access CP15.

Using other values results in unpredictable behavior

CP15有很多个寄存器，分别叫做寄存器0(Register 0)，到寄存器15 (Register 15)，

每个寄存器分别控制不同的功能，而且有的是只读，有的是只写，有的是可读写。

而且这些寄存器的含义，随着版本ARM内核版本变化而不断扩展，详情请参考：[Processor setup via co-processor 15 and about co-processors](http://www.arm.com/processor/setup/via-co-processor-15-and-about-co-processors)⁴

其中，根据我们此处关心的内容，摘录部分内容如下：

ARM 710

- Register 7 - IDC flush (write only)
Any data written to this location will cause the IDC (Instruction/Data cache) to be flushed.

.....

StrongARM SA110

.....

- Register 7 - Cache control (write only)
Any data written to this location will cause the selected cache to be flushed.

The OPC_2 and CRm co-processor fields select which cache

operation should occur:

Function OPC_2 CRm Data

Flush I + D %0000 %0111 -

Flush I %0000 %0101 -

Flush D %0000 %0110 -

Flush D single %0001 %0110 Virtual address

Clean D entry %0001 %1010 Virtual address

Drain write buf. %0100 %1010 -

- Register 8 - TLB operations (write only)
Any data written to this location will cause the selected TLB flush operation.

⁴ <http://www.heyrick.co.uk/assembler/coprocsmnd.html>

The OPC_2 and CRm co-processor fields select which cache

operation should occur:

Function OPC_2 CRm Data

Flush I + D %0000 %0111 -

Flush I %0000 %0101 -

Flush D %0000 %0110 -

Flush D single %0001 %0110 Virtual address"

而MCR的详细的语法为：

MCR指令

MCR指令将ARM处理器的寄存器中的数据传送到协处理器寄存器中。如果协处理器不能成功地执行该操作，将产生未定义的指令异常中断。

指令语法格式

MCR{<cond>} <p> , < opcode_1> , <Rd>,<CRn>,<CRm>{,<opcode_2>}

MCR{<cond>} p15 , 0 , <Rd>,<CRn>,<CRm>{,<opcode_2>}

其中

- <cond>
指令执行的条件码.当<cond>忽略时指令为无条件执行。
- <opcode_1>
协处理器将执行的的操作的操作码。对于CP15协处理器来说，<opcode_1>永远为0b000，当<opcode_1>不为0b000时，该指令操作结果不可预知。
- <Rd>
作为源寄存器的ARM寄存器，其值将被传送到协处理器寄存器中
- <CRn>
作为目标寄存器的协处理器寄存器，其编号可能是C0，C1，...，C15。

<CRm>和<opcode_2>两者组合决定对协处理器寄存器进行所需要的操作，如果没有指定，则将为<CRm>为C0，opcode_2为0

对照上面的那行代码：

```
mcr p15, 0, r0, c7, c7, 0 /* flush v3/v4 cache */
```

可以看出，其中

rd为r0=0

CRn为C7

CRm为C7

对于这行代码的作用，以此按照语法，来一点点解释如下：

首先，mcr做的事情，其实很简单，就是“ARM处理器的寄存器中的数据传送到协处理器寄存器中”，

此处即是，将ARM的寄存器r0中的数据，此时r0=0,所以就是把0这个数据，传送到协处理器CP15中。

而对应就是写入到“<CRn>”这个“目标寄存器的协处理器寄存器”，此处CRn为C7，即将0写入到寄存器7（Register 7）中去。

而上面关于Register 7的含义中也说了，“Any data written to this location will cause the selected cache to be flushed”，即你往这个寄存器7中写入任何数据，都会导致对应的缓存被清空。而到底那个缓存被清空呢，即我们这行指令

```
mcr p15, 0, r0, c7, c7, 0
```

起了什么作用呢

那是由“<CRm>和<opcode_2>两者组合决定”的。

而此处CRm为C7，opcode_2为0，而对于C7和0的组合的作用，参见上面的那个表中[Register 7中的Flash I+D \[36\]](#)那一行，

当opcode_2为0，CRm为0111=7，就是我们要找的，其作用是“Flush I + D”，即清空指令缓存I Cache和数据缓存D Cache。

根据该表，同理，如果是opcode_2=0，而CRm=0101b=5，那么对应的就是去“Flush I”，即只清除指令缓存I Cache了。

而对应的指令也就是

```
mcr p15, 0, r0, c7, c5, 0
```

了。

- ② 此注释说此行代码的作用是，清理v3或v4的缓存

其中v4，我们很好理解，因为我们此处的CPU是ARM920T的核心，是属于ARM V4的，而为何又说，也可以清除v3的cache呢？

那是因为，本身这些寄存器位域的定义，都是向下兼容的，参见上面引用的内容，也写到了：

ARM 710

- Register 7 - IDC flush (write only)
Any data written to this location will cause the IDC (Instruction/Data cache) to be flushed.

即，对于ARM7的话，你写同样的这行代码

```
mcr p15, 0, r0, c7, c7, 0
```

也还是向register 7中写入了数据0，这也同样满足了其所说的“Any data written to this location”，也会产生同样的效果“cause the IDC (Instruction/Data cache) to be flushed”。

- ③ 同理，可以看出此行是去操作寄存器8，而对应的各个参数为：

rd为r0=0

CRn为C8

CRm为C7

opcode_2为0

对照寄存器8的表：

- Register 8 - TLB operations (write only)
Any data written to this location will **cause the selected TLB flush operation.**

The OPC_2 and CRm co-processor fields select which cache operation should occur:

Function OPC_2 CRm Data

Flush I + D %0000 %0111 -

Flush I %0000 %0101 -

Flush D %0000 %0110 -

Flush D single %0001 %0110 Virtual address"

其含义为：

向寄存器8中写入数据，会导致对应的TLB被清空。具体是哪个TLB，由opcode_2和CRm组合决定，

此处opcode_2为0，CRm为7=0111b，所以对应的作用是“Flush I + D”，即清空指令和数据的TLB。



提示

上述两行代码，其实都可以ARM的官方网站上面找到：

<http://infocenter.arm.com/help/index.jsp?topic=/com.arm.doc.ddi0184b/Chdcfejb.html>

Function	Rd	Instruction
Invalidate ICache and DCache	SBZ	MCR p15,0,Rd,c7,c7,0

<http://infocenter.arm.com/help/index.jsp?topic=/com.arm.doc.ddi0184b/Chdifbjc.html>

Function	Rd	Instruction
Invalidate TLB(s)	SBZ	MCR p15,0,Rd,c8,c7,0

1.5.5. disable MMU

```
/*
 * disable MMU stuff and caches
 */
mrc p15, 0, r0, c1, c0, 0①
```

- ① 此处，对应的值为：

rd为r0=0

CRn为C1

CRm为C0

opcode_2为0

即，此行代码是将r0的值，即0，写入到CP15的寄存器1中。

寄存器1的相关的定义为：

<http://www.heyrick.co.uk/assembler/coprocmnd.html>

StrongARM SA110

- Register 1 - Control (read/write)
All values set to 0 at power-up.
- Bit 0 - On-chip MMU turned off (0) or on (1)
- Bit 1 - Address alignment fault disabled (0) or enabled (1)
- Bit 2 - Data cache turned off (0) or on (1)
- Bit 3 - Write buffer turned off (0) or on (1)
- Bit 7 - Little-endian operation if 0, big-endian if 1
- Bit 8 - System bit - controls the MMU permission system
- Bit 9 - ROM bit - controls the MMU permission system
- Bit 12 - Instruction cache turned off (0) or on (1)

所以，对应内容就是，向bit[CRm]中写入opcode_2，即向bit[0]写入0，对应的作用为“On-chip MMU turned off”，即关闭MMU。

1.5.6. clear bits

```
①bic r0, r0, #0x00002300 @ clear bits 13, 9:8 (--V- --RS)②
bic r0, r0, #0x00000087 @ clear bits 7, 2:0 (B--- -CAM)③
orr r0, r0, #0x00000002 @ set bit 2 (A) Align④
orr r0, r0, #0x00001000 @ set bit 12 (I) I-Cache⑤
mcr p15, 0, r0, c1, c0, 0⑥
```

- ① 此处几行代码，注释中写的也很清楚了，就是去清楚对应的位和设置对应的位，具体位域的含义见下：

<http://infocenter.arm.com/help/index.jsp?topic=/com.arm.doc.ddi0184b/Chdifbjc.html>

表 1.6. 控制寄存器1的位域含义

Register bits	Name	Function	Value
31	iA bit	Asynchronous clock select	See Table 2.11 ⁵

⁵ <http://infocenter.arm.com/help/topic/com.arm.doc.ddi0151c/I66051.html#Cegbgbae>

Register bits	Name	Function	Value
30	nF bit	notFastBus select	See Table 2.11 ⁶
29:15	-	Reserved	Read = Unpredictable Write = Should be zero
14	RR bit	Round robin replacement	0 = Random replacement 1 = Round-robin replacement
13	V bit	Base location of exception registers	0 = Low addresses = 0x00000000 1 = High addresses = 0xFFFF0000
12	I bit	ICache enable	0 = ICache disabled 1 = ICache enabled
11:10	-	Reserved	Read = 00 Write = 00
9	R bit	ROM protection	This bit modifies the MMU protection system. See Domain access control ⁷
8	S bit	System protection	This bit modifies the MMU protection system. See Domain access control ⁸
7	B bit	Endianness	0 = Little-endian operation 1 = Big-endian operation
6:3	-	Reserved	Read = 1111 Write = 1111
2	C bit	DCache enable	0 = DCache disabled 1 = DCache enabled
1	A bit	Alignment fault enable	Data address alignment fault checking 0 = Fault checking disabled 1 = Fault checking enabled
0	M bit	MMU enable	0 = MMU disabled 1 = MMU enabled

表 1.7. 时钟模式

Clocking mode	iA	nF
FastBus mode	0	0
Synchronous	0	1
Reserved	1	0
Asynchronous	1	1

<http://infocenter.arm.com/help/index.jsp?topic=/com.arm.doc.ddi0151c/I273867.html>

Domain access control

⁶ <http://infocenter.arm.com/help/topic/com.arm.doc.ddi0151c/I66051.html#Cegbgbae>

⁷ <http://infocenter.arm.com/help/topic/com.arm.doc.ddi0151c/I273867.html>

⁸ <http://infocenter.arm.com/help/topic/com.arm.doc.ddi0151c/I273867.html>

表 1.8. 关于访问控制位在域访问控制寄存器中的含义

Value	Meaning	Description
00	No access	Any access generates a domain fault
01	Client	Accesses are checked against the access permission bits in the section or page descriptor
10	Reserved	Reserved. Currently behaves like the no access mode
11	Manager	Accesses are not checked against the access permission bits so a permission fault cannot be generated

表 1.9 “关于访问允许(AP)位的含义” shows how to interpret the Access Permission (AP) bits and how their interpretation is dependent on the S and R bits (control register bits 8 and 9)

表 1.9. 关于访问允许(AP)位的含义

AP	S	R	Supervisor permissions	User permissions	Description
00	0	0	No access	No access	Any access generates a permission fault
00	1	0	Read-only	No access	Only Supervisor read permitted
00	0	1	Read-only	Read-only	Any write generates a permission fault
00	1	1	Reserved	-	-
01	x	x	Read/write	No access	Access allowed only in Supervisor mode
10	x	x	Read/write	Read-only	Writes in User mode cause permission fault
11	x	x	Read/write	Read/write	All access types permitted in both modes
xx	1	1	Reserved	-	

② 此行作用是：

- 清除bit[13]
Base location of exception register(异常寄存器基地址)

0 = Low address = 0x0000 0000

- 清除bit[9]和bit[8]
此处不是很懂，待后续深入了解。

目前的理解是：

不论是Supervisor还是user，谁都不能访问，否则就出现权限错误 “Any access generates a permission fault”

③ 此行作用是：

- 清除bit[7]
使用little endian

2. 清除bit[2-0]
DCache disabled,关闭Dcache ;

Alignment Fault checking disabled , 关闭地址对齐的错误检查 ;

MMU disabled , 关闭MMU。

④ 此行作用是：

1. 设置bit[1]
“Enable Data address alignment fault checking” 打开数据地址对齐的错误检查，即如果数据地址为非法（奇数？）地址，就报错。

⑤ 此行作用是：

1. 设置bit[12]
开启指令缓存I cache。

⑥ mcr指令，将刚才设置的r0的值，再写入到寄存器1中。

1.5.7. bl lowlevel_init

```
/*
 * before relocating, we have to setup RAM timing
 * because memory timing is board-dependend, you will
 * find a lowlevel_init.S in your board directory.
 */
mov ip, lr
bl lowlevel_init
mov lr, ip①
mov pc, lr②
#endif /* CONFIG_SKIP_LOWLEVEL_INIT */③
```

① 将lr的值给ip，即指令指针r12，此处之所以要保存一下lr是因为此处是在子函数cpu_init_crit中，lr已经保存了待会用于返回主函数的地址，即上次调用时候的pc的值，而此处如果在子函数cpu_init_crit中继续调用其他子函数lowlevel_init,而不保存lr的话，那么调用完lowlevel_init返回来时候，就丢失了cpu_init_crit要返回的位置。

说白了就是，每次你要调用函数之前，你自己要确保是否已经正确保存了lr的值，要保证函数调用完毕后，也能正常返回。当然，如果你此处根本不需要返回，那么就不用去保存lr的值了。

② 典型的子函数调用，通过将lr的值赋值给pc，实现函数调用完成后而返回的。

③ 这里，其是和前面的代码：

```
#ifndef CONFIG_SKIP_LOWLEVEL_INIT
bl cpu_init_crit
#endif
```

是对应的。

1.6. 异常中断处理

1.6.1. macros stmia

```
/*
```

```

*****
*
* Interrupt handling
*
*****
*/

@
@ IRQ stack frame.
@
#define S_FRAME_SIZE 72
❶
#define S_OLD_R0 68
#define S_PSR 64
#define S_PC 60
#define S_LR 56
#define S_SP 52

#define S_IP 48
#define S_FP 44
#define S_R10 40
#define S_R9 36
#define S_R8 32
#define S_R7 28
#define S_R6 24
#define S_R5 20
#define S_R4 16
#define S_R3 12
#define S_R2 8
#define S_R1 4
#define S_R0 0

#define MODE_SVC 0x13
#define I_BIT 0x80

/*
* use bad_save_user_regs for abort/prefetch/undef/swi ...
* use irq_save_user_regs / irq_restore_user_regs for IRQ/FIQ handling
*/
.macro❷ bad_save_user_regs

sub sp, sp, #S_FRAME_SIZE❸
stmia❹ sp, {r0 - r12} @ Calling r0-r12
ldr r2, _armboot_start ❺

```

❶ 此处很简单，只是一些宏定义而已。

后面用到的时候再解释。

❷ .macro和后面的.endm相对应，其语法是：

图 1.12. macro的语法

Macro Directives			
Directive	Description	Syntax	Example
.macro	Define a macro.	.macro name {args, ...}	
	A macro can be defined without arguments,		.macro NoArgsMacro
	and can be called simply by specifying its name.	name {args, ...}	NoArgsMacro
	A macro can also be defined with arguments,		.macro ArgMacro arg, arg2
	and can be called the same way with commas seperating its arguments.		ArgMacro 10, 11
	Arguments can be accessed by their name prefixed with a \.	\arg	mov r0, \arg
	You can define default macro arguments.		.macro ArgMacro arg=1, arg2
	Arguments are omitted by simply placing a comma and no value, or ignoring them all together (trailing only).		ArgMacro , 11
	Arguments can be set in a modified order by referencing them by name.		ArgMacro arg2=11, arg=10
	Macros can be recursive.		
.endm	Mark the end of a macro.	.endm	.endm
.exitm	Exit a macro early.	.exitm	.exitm
\@	Pseudo variable that contains the macro number executed. Can be used for a unique number on every macro definition.	\@	MyLabel\@:
.purgem	Undefine a macro, so that further uses do not evaluate.	.purgem name	.purgem NoArgsMacro

所以，此处就相当于一个无参数的宏bad_save_user_regs，也就相当于一个函数了。

③ 即

sp

= sp- S_FRAME_SIZE

= sp - 72

④ stmia的语法为：

图 1.13. LDM/STM的语法

<LDM|STM>{cond}<FD|ED|FA|EA|IA|IB|DA|DB> Rn{!},<Rlist>{^}

where:

- {cond} is a two-character condition mnemonic, see *Figure 5-2: Condition codes* on page 5-2
- Rn is an expression evaluating to a valid register number
- <Rlist> is a list of registers and register ranges enclosed in {} (e.g. {R0,R2-R7,R10}).
- {!} (if present) requests write-back (W=1), otherwise W=0
- {^} (if present) set S bit to load the CPSR along with the PC, or force transfer of user bank when in privileged mode

其中，条件域的具体含义如下：

图 1.14. 条件码的含义

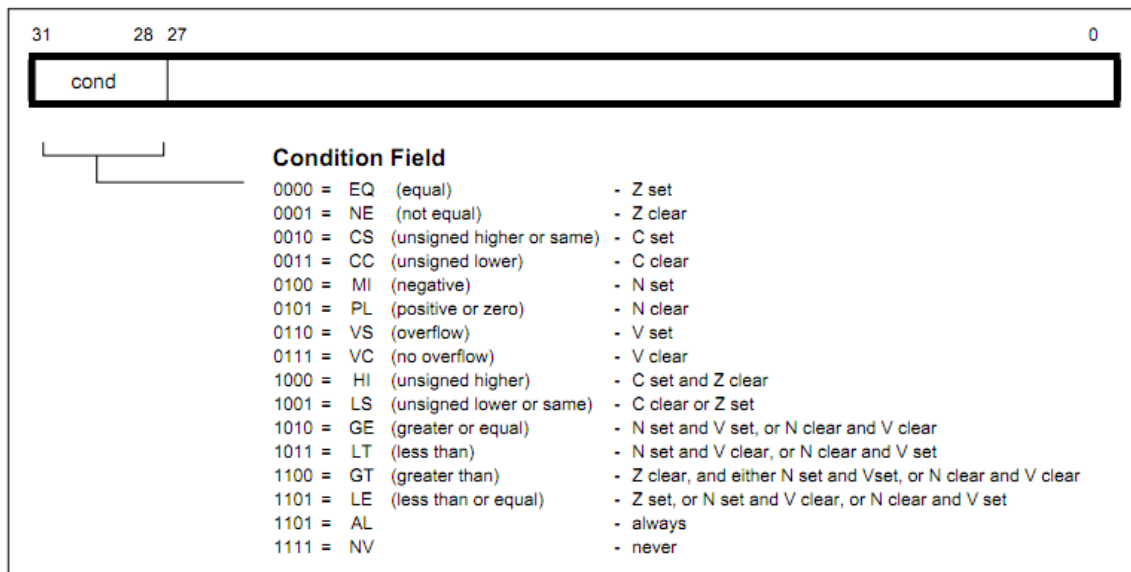


Figure 5-2: Condition codes

更具体的含义：

六、批量数据加载/存储指令ARM微处理器所支持批量数据加载/存储指令可以一次在一片连续的存储器单元和多个寄存器之间传送数据，批量加载指令用于将一片连续的存储器中的数据传送到多个寄存器，批量数据存储指令则完成相反的操作。常用的加载存储指令如下：

LDM (或STM) 指令

LDM (或STM) 指令的格式为：

LDM (或STM) {条件}{类型} 基址寄存器{!}，寄存器列表{^}

LDM (或STM) 指令用于从由基址寄存器所指示的一片连续存储器到寄存器列表所指示的多个寄存器之间传送数据，该指令的常见用途是将多个寄存器的内容入栈或出栈。

其中，{类型}为以下几种情况：

IA 每次传送后地址加1；

IB 每次传送前地址加1；

DA 每次传送后地址减1；

DB 每次传送前地址减1；

FD 满递减堆栈；

ED 空递减堆栈；

FA 满递增堆栈；

EA 空递增堆栈；

{!}为可选后缀，若选用该后缀，则当数据传送完毕之后，将最后的地址写入基址寄存器，否则基址寄存器的内容不改变。

基址寄存器不允许为R15，寄存器列表可以为R0～R15的任意组合。

{^}为可选后缀，当指令为LDM且寄存器列表中包含R15，选用该后缀时表示：除了正常的数据传送之外，还将SPSR复制到CPSR。同时，该后缀还表示传入或传出的是用户模式下的寄存器，而不是当前模式下的寄存器。

指令示例：

STMFD R13!, {R0, R4-R12, LR} ; 将寄存器列表中的寄存器 (R0, R4到R12, LR) 存入堆栈。

LDMFD R13!, {R0, R4-R12, PC} ; 将堆栈内容恢复到寄存器 (R0, R4到R12, LR) 。

所以，此行的含义是，

将r0到r12的值，一个个地传送到对应的地址上，基地址是sp的值，传完一个，sp的值加4，一直到传送完为止。

此处，可见，前面那行代码：

`sp = sp - 72`

就是为此处传送r0到r12，共13个寄存器，地址空间需要13*4=72个字节，

即前面sp减去72，就是为了腾出空间，留此处将r0到r12的值，放到对应的位置的。

❶ 此处的含义就是，将_armboot_start中的值，参考前面内容，即为_start，

而_start的值：

从 Nor Flash启动时：`_stat=0`

relocate代码之后为：`_start=TEXT_BASE=0x33D00000`

此处是已经relocate代码了，所以应该理解为后者，即`_start=0x33D00000`

所以：

`r2=0x33D00000`

1.6.2. cal reg value and store

```
sub r2, r2, #(CONFIG_STACKSIZE+CFG_MALLOC_LEN)❶
sub r2, r2, #(CFG_GBL_DATA_SIZE+8) @ set base 2 words into abort stack❷
ldmia r2, {r2 - r3} @ get pc, cpsr❸
add r0, sp, #S_FRAME_SIZE @ restore sp_SVC❹
add r5, sp, #S_SP❺
mov r1, lr❻
stmia r5, {r0 - r3} @ save sp_SVC, lr_SVC, pc, cpsr❼
mov r0, sp❽
.endm❾
```

❶ 此处：

r2

= r2 - (CONFIG_STACKSIZE+CFG_MALLOC_LEN)

= r2 - (128*1024 + 256*1024)

= 0x33D00000 - 384KB

= 0x33CA0000

② 此处：

r2

= r2 - (CFG_GBL_DATA_SIZE + 8)

= 0x33CA0000 - (128 + 8)

= 0x33C9FF78

③ 分别将地址为r2和r2+4的内容，即地址为0x33C9FF78和0x33C9FF7C中的内容，load载入给r2和r3寄存器。

④ 将sp的值，加上72，送给r0

⑤ 前面的定义是：

```
#define S_SP 52
```

所以此处就是将sp的值，加上52，送给r5

⑥ 将lr给r1

⑦ 然后将r0到r3中的内容，存储到地址为r5-r5+12中的位置去。

⑧ 将sp再赋值给r0

⑨ 结束宏bad_save_user_regs

此处虽然每行代码基本看懂了，但是到底此bad_save_user_regs函数是做什么的，还是不太清楚，有待以后慢慢深入理解。

1.6.3. irq_save_user_regs irq_restore_user_regs

```
.macro irq_save_user_regs
sub sp, sp, #S_FRAME_SIZE
stmia sp, {r0 - r12} @ Calling r0-r12
add r8, sp, #S_PC
stmdb r8, {sp, lr}^ @ Calling SP, LR
str lr, [r8, #0] @ Save calling PC
mrs r6, spsr
str r6, [r8, #4] @ Save CPSR
str r0, [r8, #8] @ Save OLD_R0
mov r0, sp
.endm

.macro irq_restore_user_regs
ldmia sp, {r0 - lr}^ @ Calling r0 - lr
mov r0, r0
ldr lr, [sp, #S_PC] @ Get PC
add sp, sp, #S_FRAME_SIZE
subs pc, lr, #4 @ return & move spsr_svc into cpsr
.endm

① .macro get_bad_stack
ldr r13, _armboot_start @ setup our mode stack
```

```

sub r13, r13, #(CONFIG_STACKSIZE+CFG_MALLOC_LEN)
sub r13, r13, #(CFG_GBL_DATA_SIZE+8) @ reserved a couple spots in abort stack
str lr, [r13] @ save caller lr / spsr
mrs lr, spsr
str lr, [r13, #4]
mov r13, #MODE_SVC @ prepare SVC-Mode
@ msr spsr_c, r13
msr spsr, r13
mov lr, pc
movs pc, lr
.endm
❷
.macro get_irq_stack @ setup IRQ stack
ldr sp, IRQ_STACK_START
.endm
❸
.macro get_fiq_stack @ setup FIQ stack
ldr sp, FIQ_STACK_START
.endm
❹

```

- ❶ 上面两段代码，基本上和前面很类似，虽然每一行都容易懂，但是整个两个函数的意思，除了看其宏的名字irq_save_user_regs和irq_restore_user_regs，分别对应着中断中，保存和恢复用户模式寄存器，之外，其他的，个人目前还是没有太多了解。
- ❷ 此处的get_bad_stack被后面undefined_instruction，software_interrupt等处调用，目前能理解的意思是，在出错的时候，获得对应的堆栈的值。
- ❸ 此处的含义很好理解，就是把地址为IRQ_STACK_START中的值赋值给sp。

即获得IRQ的堆栈的起始地址。

而对于IRQ_STACK_START，是前面就提到过的[cpu_init源码](#)

而此处，就是用到了，前面已经在cpu_init()中重新计算正确的值了。

即算出IRQ堆栈的起始地址，其算法很简单，就是：

```
IRQ_STACK_START = _armboot_start - CFG_MALLOC_LEN - CFG_GBL_DATA_SIZE - 4;
```

即，先减去malloc预留的空间，和global data，即在

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\board\EmbedSky\board.c

中定义的全局变量：

```
DECLARE_GLOBAL_DATA_PTR;
```

而此宏对应的值在：

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\include\asm-arm\global_data.h

中：

```
#define DECLARE_GLOBAL_DATA_PTR register volatile gd_t *gd asm ("r8")
```

即，用一个固定的寄存器r8来存放此结构体的指针。



提示

这也对应着编译uboot的时候，你所看到的编译参数-ffixed-r8

此gd_t的结构体，不同体系结构，用的不一样。

而此处arm的平台中，gd_t的定义在同一文件中：

```
typedef struct global_data {
    bd_t *bd;
    unsigned long flags;
    unsigned long baudrate;
    unsigned long have_console; /* serial_init() was called */
    unsigned long reloc_off; /* Relocation Offset */
    unsigned long env_addr; /* Address of Environment struct */
    unsigned long env_valid; /* Checksum of Environment valid? */
    unsigned long fb_base; /* base address of frame buffer */
#ifdef CONFIG_VFD
    unsigned char vfd_type; /* display type */
#endif
#if 0
    unsigned long cpu_clk; /* CPU clock in Hz! */
    unsigned long bus_clk;
    unsigned long ram_size; /* RAM size */
    unsigned long reset_status; /* reset status register at boot */
#endif
    void **jt; /* jump table */
} gd_t;
```

而此全局变量gd_t *gd会被其他很多文件所引用，详情自己去代码中找。

- ④ 此处和上面类似，把地址为FIQ_STACK_START中的内容，给sp。

其中：

```
FIQ_STACK_START = IRQ_STACK_START - CONFIG_STACKSIZE_IRQ;
```

即FIQ的堆栈起始地址，是IRQ堆栈起始地址减去IRQ堆栈的大小。

1.6.4. exception handlers

```
/*
 * exception handlers
 */
.align 5
undefined_instruction:①
    get_bad_stack
    bad_save_user_regs
    bl do_undefined_instruction
    .align 5

software_interrupt:
    get_bad_stack
    bad_save_user_regs
    bl do_software_interrupt

    .align 5
prefetch_abort:
```

```

get_bad_stack
bad_save_user_regs
bl do_prefetch_abort

.align 5
data_abort:
get_bad_stack
bad_save_user_regs
bl do_data_abort

.align 5
not_used:
get_bad_stack
bad_save_user_regs
bl do_not_used
❷

```

- ❶ 如果发生未定义指令异常，CPU会掉转到start.S开头中对应的位置：

```
ldr pc, _undefined_instruction
```

即把地址为_undefined_instruction中的内容给pc，即跳转到此处执行对应的代码。

其做的事情依次是：

获得出错时候的堆栈

保存用户模式寄存器

跳转到对应的函数：do_undefined_instruction

而do_undefined_instruction函数是在：

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\cpu\arm920t\interrupts.c

中：

```

void bad_mode (void)
{
    panic ("Resetting CPU ...\n");
    reset_cpu (0);
}

void do_undefined_instruction (struct pt_regs *pt_regs)
{
    printf ("undefined instruction\n");
    show_regs (pt_regs);
    bad_mode ();
}

```

可以看到，此处起始啥事没错，只是打印一下出错时候的寄存器的值，然后跳转到bad_mode中取reset CPU，直接重启系统了。

- ❷ 以上几个宏，和前面的do_undefined_instruction是类似的，就不多说了。

1.6.5. Launch

```

@ HJ
.globl Launch
.align 4
Launch:
    mov r7, r0
    @ disable interrupt
    @ disable watch dog timer
    mov r1, #0x53000000
    mov r2, #0x0
    str r2, [r1]

    ldr r1,=INTMSK
    ldr r2,=0xffffffff @ all interrupt disable
    str r2,[r1]

    ldr r1,=INTSUBMSK
    ldr r2,=0x7ff @ all sub interrupt disable
    str r2,[r1]

    ldr r1, = INTMOD
    mov r2, #0x0 @ set all interrupt as IRQ (not FIQ)
    str r2, [r1]

    @
    mov ip, #0
    mcr p15, 0, ip, c13, c0, 0 @ /* zero PID */
    mcr p15, 0, ip, c7, c7, 0 @ /* invalidate I,D caches */
    mcr p15, 0, ip, c7, c10, 4 @ /* drain write buffer */
    mcr p15, 0, ip, c8, c7, 0 @ /* invalidate I,D TLBs */
    mrc p15, 0, ip, c1, c0, 0 @ /* get control register */
    bic ip, ip, #0x0001 @ /* disable MMU */
    mcr p15, 0, ip, c1, c0, 0 @ /* write control register */

    @ MMU_EnableICache
    @mrc p15,0,r1,c1,c0,0
    @orr r1,r1,#(1<<12)
    @mcr p15,0,r1,c1,c0,0

#ifdef CONFIG_SURPORT_WINCE
    bl Wince_Port_Init
#endif

    @ clear SDRAM: the end of free mem(has wince on it now) to the end of SDRAM
    ldr r3, FREE_RAM_END
    ldr r4, =PHYS_SDRAM_1+PHYS_SDRAM_1_SIZE @ must clear all the memory unused to
zero
    mov r5, #0

    ldr r1, _armboot_start
    ldr r2, =On_Steppingstone
    sub r2, r2, r1
    mov pc, r2
On_Steppingstone:
2: stmia r3!, {r5}

```

```

cmp    r3, r4
bne    2b

@ set sp = 0 on sys mode
mov sp, #0

@ add by HJ, switch to SVC mode
msr cpsr_c, #0xdf @ set the I-bit = 1, diable the IRQ interrupt
msr cpsr_c, #0xd3 @ set the I-bit = 1, diable the IRQ interrupt
ldr sp, =0x31ff5800

nop
nop
nop
nop

mov    pc, r7 @ Jump to PhysicalAddress
nop
mov pc, lr

```

- ❶ 此处相当于一个叫做Launch的函数，做了也是类似的系统初始化的动作。
但是没找到此函数在哪里被调用的。具体不太清楚。

1.6.6. int_return

```

❶
#ifdef CONFIG_USE_IRQ
.align 5
irq:
/* add by www.embedsky.net to use IRQ for USB and DMA */
sub lr, lr, #4 @ the return address
ldr sp, IRQ_STACK_START @ the stack for irq
stmdb sp!, { r0-r12,lr } @ save registers

ldr lr, =int_return @ set the return addr
ldr pc, =IRQ_Handle❷ @ call the isr
int_return:
ldmia sp!, { r0-r12,pc }^ @ return from interrupt
.align 5
fiq:❸
get_fiq_stack
/* someone ought to write a more effiction fiq_save_user_regs */
irq_save_user_regs
bl do_fiq❹
irq_restore_user_regs
#else
❺
.align 5
irq:
get_bad_stack
bad_save_user_regs
bl do_irq

.align 5
fiq:

```

```

get_bad_stack
bad_save_user_regs
bl do_fiq

#endif

```

- ❶ 此处，做的事情，很容易看懂，就是中断发生后，掉转到这里，然后保存对应寄存器，然后跳转到对应irq函数IRQ_Handle中去。

但是前面为何sp为何去减去4，原因不太懂。

- ❷ 关于IRQ_Handle，是在：

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\cpu\arm920t\s3c24x0\interrupts.c
中：

```

void IRQ_Handle(void)
{
    unsigned long oft = intregs->INTOFFSET;
    S3C24X0_GPIO * const gpio = S3C24X0_GetBase_GPIO();

    // printk("IRQ_Handle: %d\n", oft);

    //清中断
    if( oft == 4 ) gpio->EINTPEND = 1<<7;
    intregs->SRCPND = 1<<oft;
    intregs->INTPND = intregs->INTPND;

    /* run the isr */
    isr_handle_array[oft]();
}

```

此处细节就不多解释了，大体含义是，找到对应的中断源，然后调用对应的之前已经注册的中断服务函数ISR。

- ❸ 此处也很简单，就是发生了快速中断FIQ的时候，保存IRQ的用户模式寄存器，然后调用函数do_fiq,调用完毕后，再恢复IRQ的用户模式寄存器。
- ❹ do_fiq()是在：

u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\cpu\arm920t\interrupts.c
中：

```

void do_fiq (struct pt_regs *pt_regs)
{
    printf ("fast interrupt request\n");
    show_regs (pt_regs);
    bad_mode ();
}

```

和前面提到过的do_undefined_instruction的一样，就是打印寄存器信息，然后跳转到bad_mode()去重启CPU而已。

- ❺ 此处就是，如果没有定义CONFIG_USE_IRQ，那么就用这段代码，可以看到，都只是直接调用do_irq和do_fiq，也没做什么实际工作。

第 2 章 start.S的总结

2.1. start.S各个部分的总结

其实关于start.S这个汇编文件，主要做的事情就是系统的各个方面的初始化。

关于每个部分，上面具体的代码实现，也都一行行的解释过了，此处不再赘述。

此处，只是简单总结一下，其实现的方式，或者其他需要注意的地方。

1. 设置CPU模式

总的来说，就是将CPU设置为SVC模式。

至于为何设置CPU是SVC模式，请参见后面章节的详细解释。

2. 关闭看门狗

就是去设置对应的寄存器，将看门狗关闭。

至于为何关闭看门狗，请参见后面章节的详细解释。

3. 关闭中断

关闭中断，也是去设置对应的寄存器，即可。

4. 设置堆栈sp指针

所谓的设置堆栈sp指针，这样的句子，之前听到N次了，但是说实话，一直不太理解，到底更深一层的含义是什么。

后来，看了更多的代码，才算有一点点了解。所谓的设置堆栈sp指针，就是设置堆栈，而所谓的设置堆栈，要做的事情，看起来很简单，就只是一个很简单的动作：让sp等于某个地址值，即可。

但是背后的逻辑是：

首先你自己要搞懂当前系统是如何使用堆栈的，堆栈是向上生长的还是向下生长的。

然后知道系统如何使用堆栈之后，给sp赋值之前，你要保证对应的地址空间，是专门分配好了，专门给堆栈用的，保证堆栈的大小相对合适，而不要太小以至于后期函数调用太多，导致堆栈溢出，或者堆栈太大，浪费存储空间，等等。

所有这些背后的逻辑，都是要经过一定的编程经验，才更加容易理解其中的含义的。

此处，也只是简单说说，更多相关的内容，还是要靠每个人自己多实践，慢慢的更加深入的理解。

5. 清除bss段

此处很简单，就是将对应bss段，都设置为0，即清零。

其对应的地址空间，就是那些未初始化的全局变量之类的地址。

6. 异常中断处理

异常中断处理，就是实现对应的常见的那些处理中断的部分内容。

说白了就是实现一个个中断函数。uboot在初始化的时候，主要目的只是为了初始化系统，及引导系统，所以，此处的中断处理部分的代码，往往相对比较简单，不是很复杂。

2.2. Uboot中的内存的Layout

总结了start.S做的事情之后，另外想在此总结一下，uboot中，初始化部分的代码执行后，对应的内存空间，都是如何规划，什么地方放置了什么内容。此部分内容，虽然和start.S没有直接的关系，但是start.S中，堆栈sp的计算等，也和这部分内容有关。

下面这部分的uboot的内存的layout，主要是根据：

1. start.S中关于设置堆栈指针的部分的代码

```
/* Set up the stack      */
stack_setup:
ldr r0, _TEXT_BASE /* upper 128 KiB: relocated uboot */
sub r0, r0, #CFG_MALLOC_LEN /* malloc area          */
sub r0, r0, #CFG_GBL_DATA_SIZE /* bdfinfo          */

#ifdef CONFIG_USE_IRQ
sub r0, r0, #(CONFIG_STACKSIZE_IRQ+CONFIG_STACKSIZE_FIQ)
#endif
sub sp, r0, #12 /* leave 3 words for abort-stack */

bl clock_init
```

2. u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\cpu\arm920t\cpu.c中的代码

```
int cpu_init (void)
{
/*
* setup up stacks if necessary
*/
#ifdef CONFIG_USE_IRQ
IRQ_STACK_START = _armboot_start - CFG_MALLOC_LEN - CFG_GBL_DATA_SIZE - 4;
FIQ_STACK_START = IRQ_STACK_START - CONFIG_STACKSIZE_IRQ;
FREE_RAM_END = FIQ_STACK_START - CONFIG_STACKSIZE_FIQ - CONFIG_STACKSIZE;
FREE_RAM_SIZE = FREE_RAM_END - PHYS_SDRAM_1;
#else
FREE_RAM_END = _armboot_start - CFG_MALLOC_LEN - CFG_GBL_DATA_SIZE - 4 -
CONFIG_STACKSIZE;
FREE_RAM_SIZE = FREE_RAM_END - PHYS_SDRAM_1;
#endif
return 0;
}
```

3. u-boot-1.1.6_20100601\opt\EmbedSky\u-boot-1.1.6\board\EmbedSky\config.mk中的定义

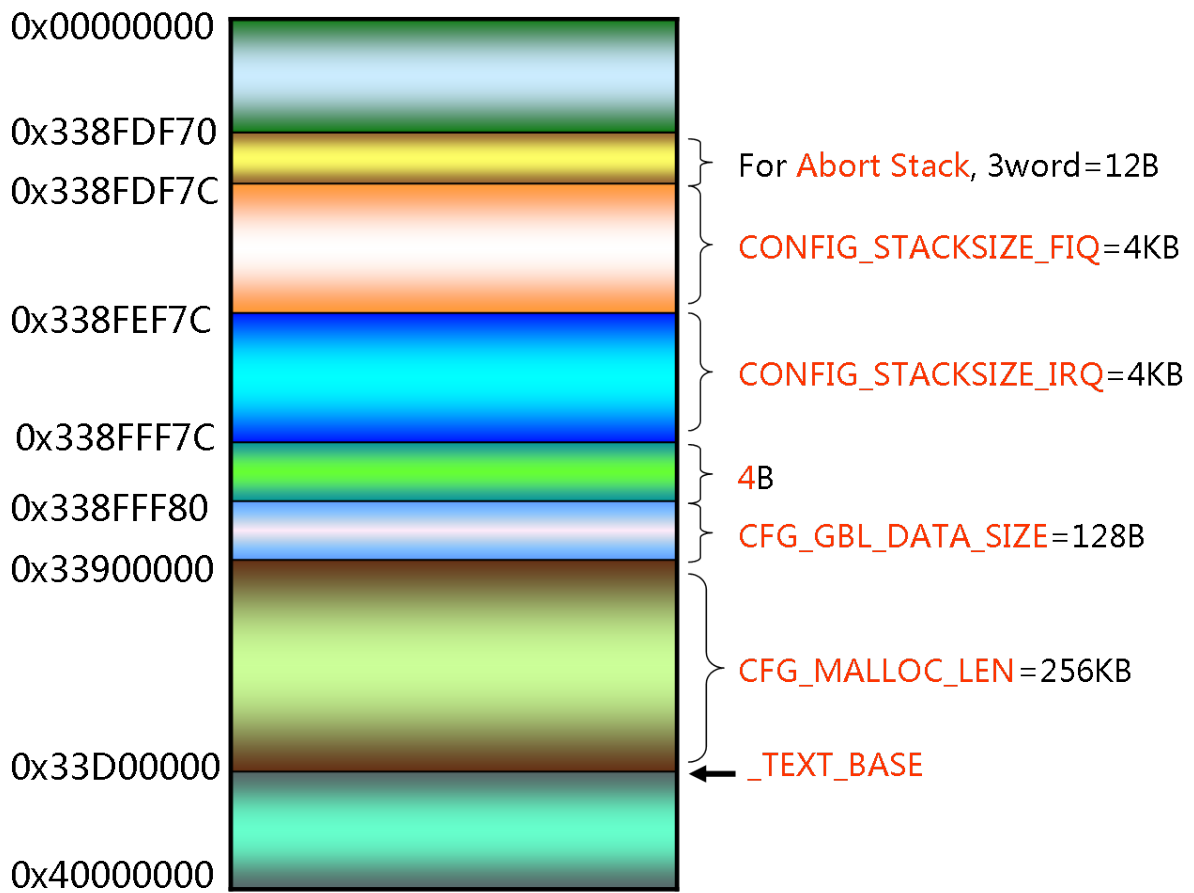
```
TEXT_BASE = 0x33D00000
```

分析而得出的。

uboot的内存的layout，用图表表示就是：

图 2.1. Uboot中的内存的Layout

Uboot Memory Layout



第 3 章 相关知识点详解

3.1. 如何查看C或汇编的源代码所对应的真正的汇编代码

首先解释一下，由于汇编代码中会存在一些伪指令等内容，所以，写出来的汇编代码，并不一定是真正可以执行的代码，这些类似于伪指令的汇编代码，经过汇编器，转换或翻译成真正的可以执行的汇编指令。所以，上面才会有将“汇编源代码”转换为“真正的汇编代码”这一说。

然后，此处对于有些人不是很熟悉的，如何查看源代码真正对应的汇编代码。

此处，对于汇编代码，有两种：

1. 一种是只是进过编译阶段，生成了对应的汇编代码
2. 另外一种，是，编译后的汇编代码，经过链接器链接后，对应的汇编代码。

总的来说，两者区别很小，后者主要是更新了外部函数的地址等等，对于汇编代码本身，至少对于我们一般所去查看源代码所对应的汇编来说，两者可以视为没区别。

在查看源代码所对应的真正的汇编代码之前，先要做一些相关准备工作：

1. 编译uboot

在Linux下，一般编译uboot的方法是：

a. make distclean

去清除之前配置，编译等生成的一些文件。

b. make EmbedSky_config

去配置我们的uboot

c. make

去执行编译

2. 查看源码所对应的汇编代码

对于我们此处的uboot的start.S来说：

- a. 对于编译所生成的汇编的查看方式是
用交叉编译器的dump工具去将汇编代码都导出来：

```
arm-linux-objdump -d cpu/arm920t/start.o > uboot_start.o_dump_result.txt
```

这样就把start.o中的汇编代码导出到uboot_start.o_dump_result.txt中了。

然后查看uboot_start.o_dump_result.txt，即可找到对应的汇编代码。

举例来说，对于start.S中的汇编代码：

```
/* Set up the stack */
stack_setup:
ldr r0, _TEXT_BASE /* upper 128 KiB: relocated uboot */
sub r0, r0, #CFG_MALLOC_LEN /* malloc area */
sub r0, r0, #CFG_GBL_DATA_SIZE /* bdfinfo */
```

```
#ifdef CONFIG_USE_IRQ
sub r0, r0, #(CONFIG_STACKSIZE_IRQ+CONFIG_STACKSIZE_FIQ)
#endif
sub sp, r0, #12 /* leave 3 words for abort-stack */

bl clock_init
```

去uboot_start.o_dump_result.txt中，搜索stack_setup，即可找到对应部分的汇编代码：

```
00000090 <stack_setup>:
90: e51f0058 ldr r0, [pc, #-88]; 40 <_TEXT_BASE>
94: e2400701 sub r0, r0, #262144; 0x400000
98: e2400080 sub r0, r0, #128; 0x80
9c: e240d00c sub sp, r0, #12; 0xc
a0: ebfffffe bl 0 <clock_init>
```

- b. 对于链接所生成的汇编的查看方式是和上面方法一样，即：

```
arm-linux-objdump -d u-boot > whole_uboot_dump_result.txt
```

然后打开该txt，找到stack_setup部分的代码：

```
33d00090 <stack_setup>:
33d00090: e51f0058 ldr r0, [pc, #-88]; 33d00040 <_TEXT_BASE>
33d00094: e2400701 sub r0, r0, #262144; 0x400000
33d00098: e2400080 sub r0, r0, #128; 0x80
33d0009c: e240d00c sub sp, r0, #12; 0xc
33d000a0: eb000242 bl 33d009b0 <clock_init>
```

两者不一样地方在于，我们uboot设置了text_base，即代码段的基地址，上面编译后的汇编代码，经过链接后，更新了对应的基地址，所以看起来，所以代码对应的地址，都变了，但是具体地址中的汇编代码，除了个别调用函数的地址和跳转到某个标号的地址之外，其他都还是一样的。

对于C语言的源码，也是同样的方法，用对应的dump工具，去从该C语言的.o文件中，dump出来汇编代码。



注意

【总结】

不论是C语言还是汇编语言的源文件，想要查看其对应的生成的汇编代码的话，方法很简单，就是用dump工具，从对应的.o目标文件中，导出对应的汇编代码，即可。

3.2. uboot初始化中，为何要设置CPU为SVC模式而不是设置为其他模式

在看Uboot的start.S文件时候，发现其最开始初始化系统，做的第一件事情，就是将CPU设置为SVC模式，但是S3C2440的CPU的core是ARM920T，其有7种模式，为何非要设置为SVC模式，而不是设置为其他模式呢？对此，经过一些求证，得出如下原因：

首先，先要了解ARM的CPU的7种模式是哪些：

<http://www.docin.com/p-73665362.html>

表 3.1. ARM中CPU的模式

处理器模式	说明	备注
用户(usr)	正常程序工作模式	此模式下程序不能够访问一些受操作系统保护的系统资源，应用程序也不能直接进行处理器模式的切换。
系统(sys)	用于支持操作系统的特权任务等	与用户模式类似，但具有可以直接切换到其它模式等特权
快中断(fiq)	支持高速数据传输及通道处理	FIQ异常响应时进入此模式
中断(irq)	用于通用中断处理	IRQ异常响应时进入此模式
管理(svc)	操作系统保护代码	系统复位和软件中断响应时进入此模式
中止(abt)	用于支持虚拟内存和/或存储器保护	在ARM7TDMI没有大用处
未定义(und)	支持硬件协处理器的软件仿真	未定义指令异常响应时进入此模式

另外，7种模式中，除用户usr模式外，其它模式均为特权模式。

对于为何此处是svc模式，而不是其他某种格式，其原因，可以从两方面来看：

1. 我们先简单的来分析一下那7种模式：

- a. 中止abt和未定义und模式
首先可以排除的是，中止abt和未定义und模式，那都是不太正常的模式，此处程序是正常运行的，所以不应该设置CPU为其中任何一种模式，所以可以排除。
- b. 快中断fiq和中断irq模式
其次，对于快中断fiq和中断irq来说，此处uboot初始化的时候，也还没啥中断要处理和能够处理，而且即使是注册了终端服务程序后，能够处理中断，那么这两种模式，也是自动切换过去的，所以，此处也不应该设置为其中任何一种模式。
- c. 用户usr模式
虽然从理论上来说，可以设置CPU为用户usr模式，但是由于此模式无法直接访问很多的硬件资源，而uboot初始化，就必须要去访问这类资源，所以此处可以排除，不能设置为用户usr模式。
- d. 系统sys模式 vs 管理svc模式
首先，sys模式和usr模式相比，所用的寄存器组，都是一样的，但是增加了一些访问一些在usr模式下不能访问的资源。

而svc模式本身就属于特权模式，本身就可以访问那些受控资源，而且，比sys模式还多了些自己模式下的影子寄存器，所以，相对sys模式来说，可以访问资源的能力相同，但是拥有更多的硬件资源。

所以，从理论上来说，虽然可以设置为sys和svc模式的任一种，但是从uboot方面考虑，其要做的事情是初始化系统相关硬件资源，需要获取尽量多的权限，以方便操作硬件，初始化硬件。

从uboot的目的是初始化硬件的角度来说，设置为svc模式，更有利于其工作。

因此，此处将CPU设置为SVC模式。

2. uboot作为一个bootloader来说，最终目的是为了启动Linux的kernel，在做好准备工作（即初始化硬件，准备好kernel和rootfs等）跳转到kernel之前，本身就要满足一些条件，其中一个条件，就是要求CPU处于SVC模式的。

所以，uboot在最初的初始化阶段，就将CPU设置为SVC模式，也是最合适的。



提示

关于满足哪些条件，详情请参考

[ARM Linux Kernel Boot Requirements](#)¹

或者Linux内核文档:

kernel_source_root\documentation\arm\booting

中也是同样的解释：

The CPU must be in SVC mode

所以，uboot在最初的初始化阶段，就将CPU设置为SVC模式，也是最合适的。

综上所述，uboot在初始化阶段，就应该将CPU设置为SVC模式。

3.3. 什么是watchdog + 为何在要系统初始化的时候关闭watchdog

关于Uboot初始化阶段，在start.S中，为何要去关闭watchdog，下面解释具体的原因：

3.3.1. 什么是watchdog

参考[嵌入式系统之WATCHDOG\(看门狗\)概述](#)²

简要摘录如下：

watchdog一般是一个硬件模块，其作用是，在嵌入式操作系统中，很多应用情况是系统长期运行且无人看守，所以难免或者怕万一出现系统死机，那就杯具了，这时，watchdog就会自动帮你重启系统。

那么其是如何实现此功能的呢？那么就要简单解释一下其实现原理了。

watchdog硬件的逻辑就是，其硬件上有个记录超时功能，然后要求用户需要每隔一段时间（此时间可以根据自己需求而配置）去对其进行一定操作，比如往里面写一些固定的值，俗称“喂狗”，那么我发现超时了，即过了这么长时间你还不给偶喂食，那么偶就认为你系统是死机了，出问题了，偶就帮你重启系统。

说白了就是弄个看家狗dog，你要定期给其喂食，如果超时不喂食，那么狗就认为你，他的主人，你的系统，死机了，就帮你reset重启系统。

3.3.2. 为何在要系统初始化的时候关闭watchdog

了解了watchdog的原理后，此问题就很容易理解了。

如果不禁用watchdog，那么就要单独写程序去定期“喂狗”，那多麻烦，多无聊啊。

毕竟咱此处只是去用uboot初始化必要的硬件资源和系统资源而已，完全用不到这个watchdog的机制。需要用到，那也是你linux内核跑起来了，是你系统关心的事情，和我uboot没啥关系的，所以肯定此处要去关闭watchdog（的reset功能）了。

¹ <http://www.arm.linux.org.uk/developer/booting.php>

² <http://wenku.baidu.com/view/e5cd52ff04a1b0717fd5dd27.html>

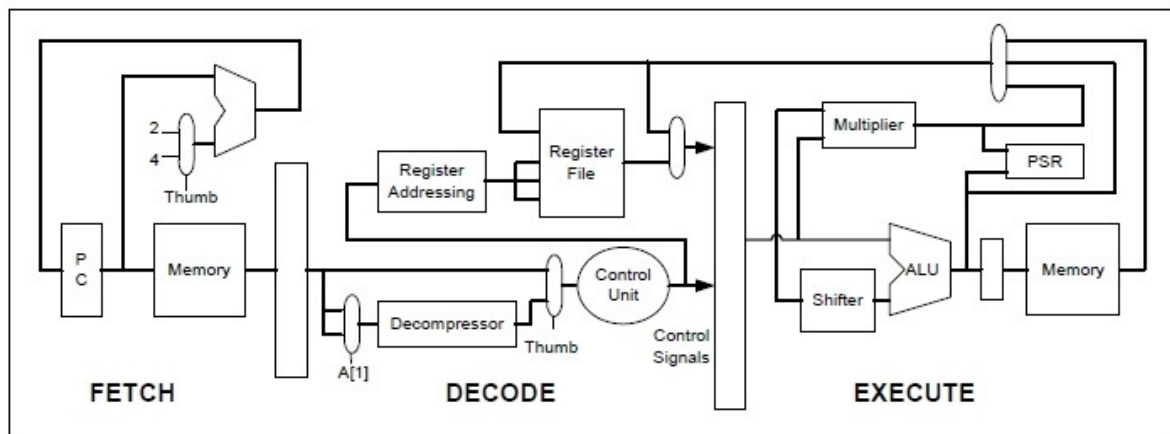
3.4. 为何ARM7中PC=PC+8

此处解释为何ARM7中，CPU地址，即PC，为何有PC=PC+8这一说法：

众所周知，AMR7，是三级流水线，其细节见图：

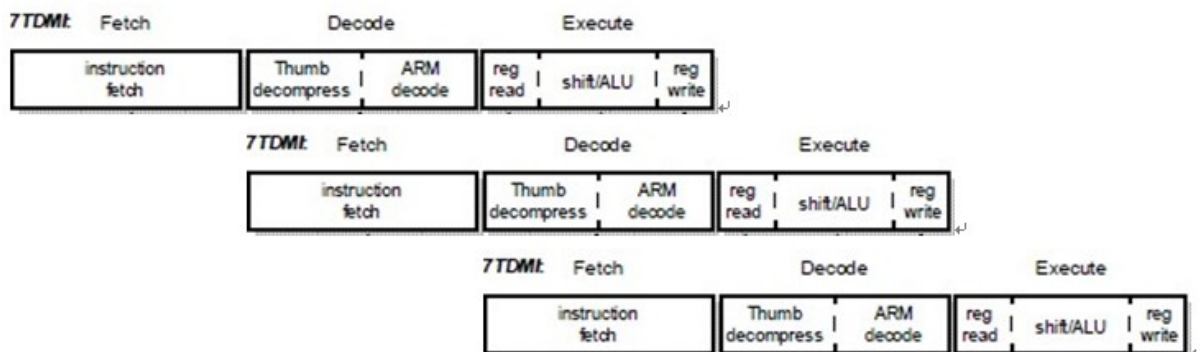
图 3.1. AMR7三级流水线

ARM7TM Core and Pipeline



首先，对于ARM7对应的流水线的执行情况，如下面这个图所示：

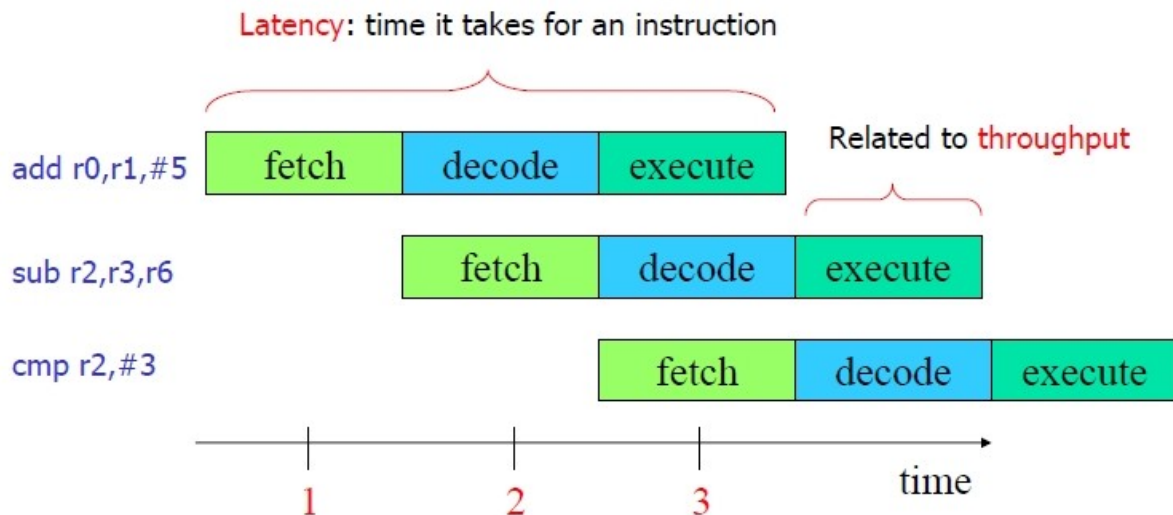
图 3.2. ARM7三级流水线状态



然后对于三级流水线举例如下：

图 3.3. ARM7三级流水线示例

ARM7 Single-Cycle Instruction 3-Stage Pipeline



从上图，其实很容易看出，第一条指令：

`add r0, r1,$5`

执行的时候，此时PC已经指向第三条指令：

`cmp r2,#3`

的地址了，所以，是 $PC=PC+8$ 。

3.4.1. 为何ARM9和ARM7一样，也是 $PC=PC+8$

ARM7的三条流水线， $PC=PC+8$ ，很好理解，但是ARM9中，是五级流水线，为何还是 $PC=PC+8$ ，而不是

PC

$=PC+(5-1)*4$

$=PC + 16$ ，

呢？

下面就需要好好解释一番了。

具体解释之前，先贴上ARM7和ARM9的流水线的区别和联系：

图 3.4. ARM7三级流水线 vs ARM9五级流水线

Figure 1 : The ARM7TDMI core and ARM7TDMI-S core pipeline

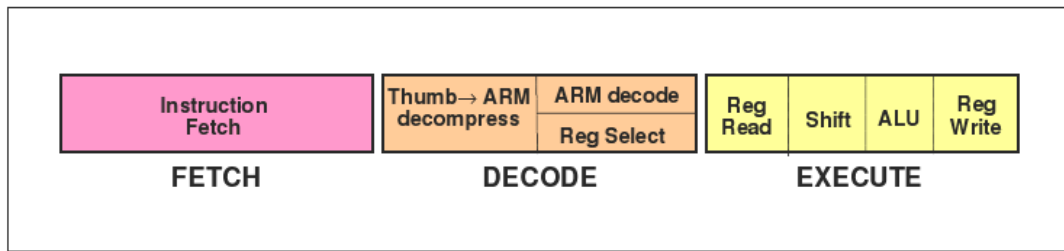


Figure 2 : The ARM9TDMI core pipeline

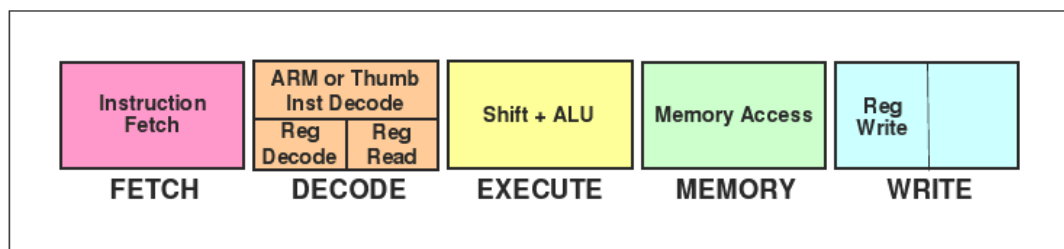
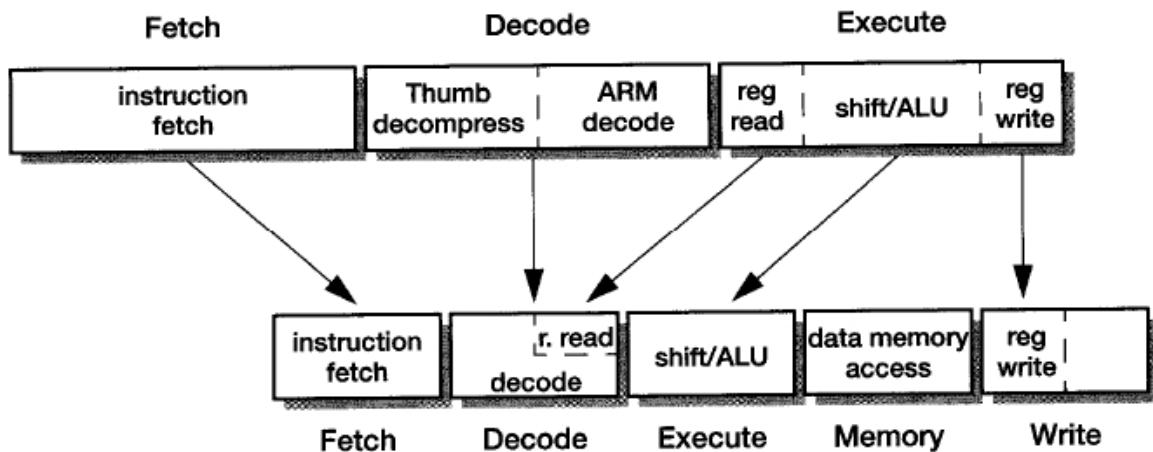


图 3.5. ARM7三级流水线到ARM9五级流水线的映射

ARM7TDMI 3-Stage Pipeline



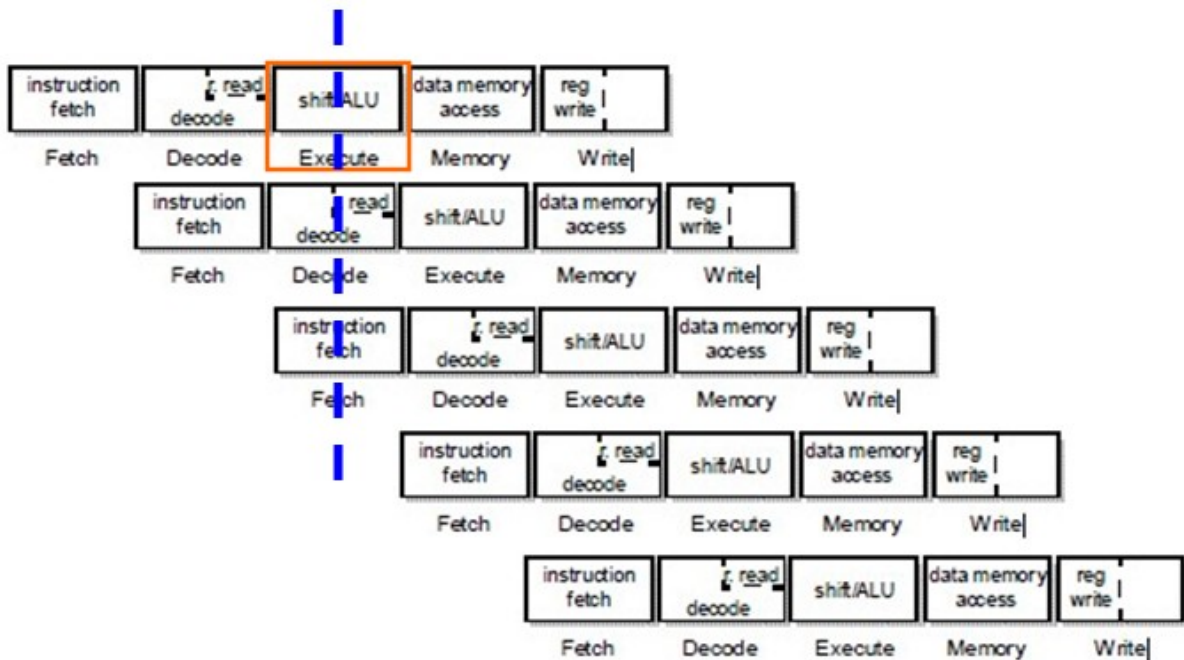
ARM9TDMI 5-Stage Pipeline

下面开始对为何ARM9也是PC=PC+8进行解释。

先列出ARM9的五级流水线的示例：

图 3.6. ARM9的五级流水线示例

ARM9的五级流水线示例



举例分析为何 $PC = PC + 8$

然后我们以下面uboot中的start.S的最开始的汇编代码为例来进行解释：

```
00000000 <_start>:
 0: ea000014 b 58 <reset>
 4: e59ff014 ldr pc, [pc, #20] ; 20 <_undefined_instruction>
 8: e59ff014 ldr pc, [pc, #20] ; 24 <_software_interrupt>
 c: e59ff014 ldr pc, [pc, #20] ; 28 <_prefetch_abort>
10: e59ff014 ldr pc, [pc, #20] ; 2c <_data_abort>
14: e59ff014 ldr pc, [pc, #20] ; 30 <_not_used>
18: e59ff014 ldr pc, [pc, #20] ; 34 <_irq>
1c: e59ff014 ldr pc, [pc, #20] ; 38 <_fiq>

00000020 <_undefined_instruction>:
20: 00000120 .word 0x00000120
```

下面对每一个指令周期，CPU做了哪些事情，分别详细进行阐述：

在看下面具体解释之前，有一句话要牢记，那就是：

PC不是指向你正在运行的指令，而是

PC始终指向你要取的指令的地址

认识清楚了这个前提，后面的举例讲解，就容易懂了。

1. 指令周期Cycle1

a. 取指

PC总是指向将要读取的指令的地址（即我们常说的，指向下一条指令的地址），而当前PC=4，

所以去取物理地址为4对应的指令

```
ldr pc, [pc, #20]
```

其对应二进制代码为e59ff014。

此处取指完之后，自动更新PC的值，即 $PC = PC + 4$ （单个指令占4字节，所以加4） $= 4 + 4 = 8$

2. 指令周期Cycle2

a. 译指

翻译指令e59ff014

b. 同时再去取指

PC总是指向将要读取的指令的地址（即我们常说的，指向下一条指令的地址），而当前PC=8，

所以去物理地址为8所对应的指令“ldr pc, [pc, #20]”其对应二进制代码为e59ff014。

此处取指完之后，自动更新PC的值，即 $PC = PC + 4 = 8 + 4 = 12 = 0xc$

3. 指令周期Cycle3

a. 执行（指令）

执行“e59ff014”，即

```
ldr pc, [pc, #20]
```

所对表达的含义，即PC

$= PC + 20$

$= 12 + 20$

$= 32$

$= 0x20$

此处，只是计算出待会要赋值给PC的值是0x20，这个0x20还只是放在执行单元中内部的缓冲中。

b. 译指

翻译e59ff014

c. 取指

此步骤由于是和上面（1）中的执行同步做的，所以，未受到影响，继续取指，而取指的那一时刻，PC为上一Cycle更新后的值，即 $PC = 0xc$ ，所以是去取物理地址为0xc所对应的指令

```
ldr pc, [pc, #20]
```

对应二进制为e59ff014

在Cycle3的时候，PC的值，刚好已经在Cycle1和Cycle2，分别加了4，所以Cycle3的时候， $PC = PC + 8$ ，而同样道理，对于任何一条指令的，都是在Cycle3，指令的Execute执行阶段，如果用到PC的值，那么PC那一时刻，就是 $PC = PC + 8$ 。

所以，此处虽然是五级流水线，但是却不是 $PC = PC + 16$ ，而是 $PC = PC + 8$ 。

进一步地，我们发现，其实 $PC = PC + N$ 的N，是和指令的执行阶段所处于流水线的深度有关，即此处指令的执行Execute阶段，是五级流水线中的第三个，而这个第三阶段的Execute和指令的第一个阶段的Fetch取指，相差的值是 $3 - 1 = 2$ ，即两个CPU的Cycle，而每个Cycle都会导致 $PC = PC + 4$ ，所以，指令到了Execute阶段，才会发现，此时PC已经变成 $PC = PC + 8$ 了。

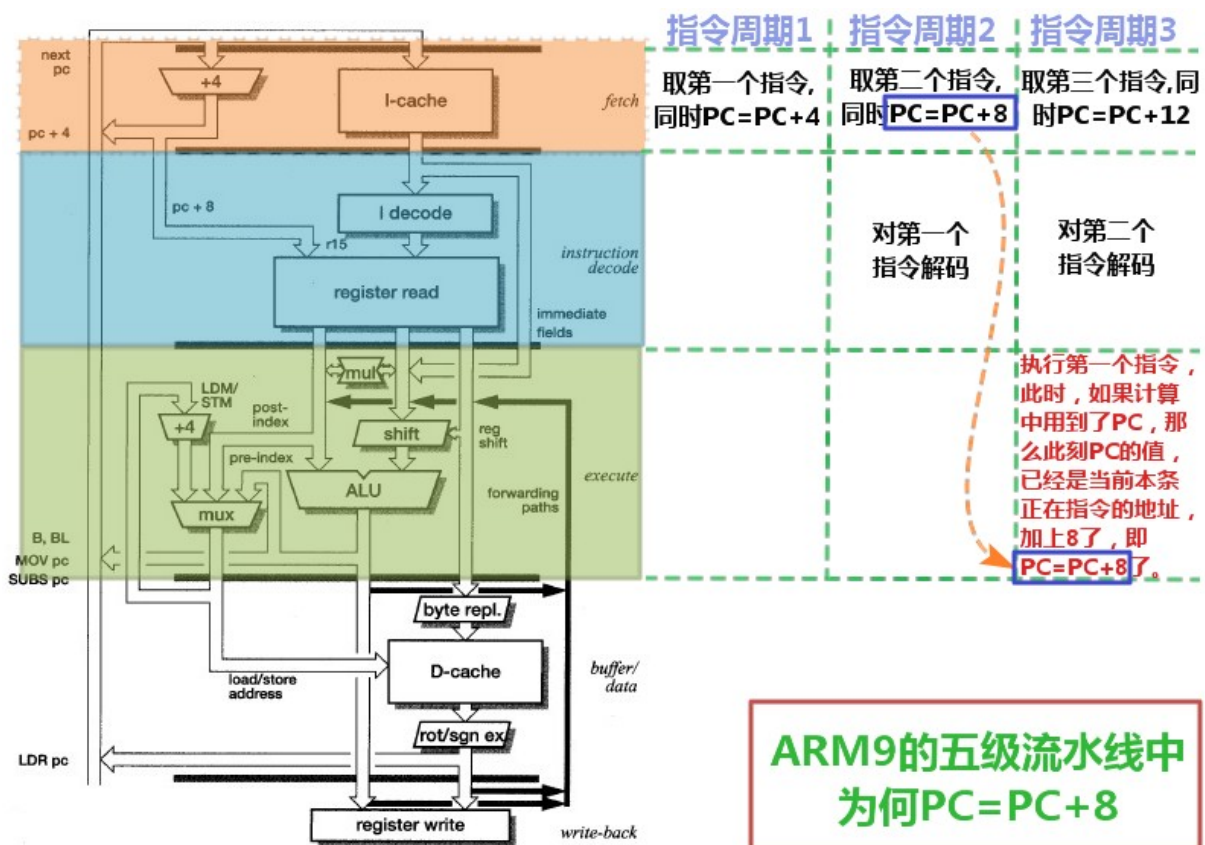
回过头来反观ARM7的三级流水线，也是同样的道理，指令的Execute执行阶段，是处于指令的第三个阶段，同理，在指令计算数据的时候，如果用到PC，就会发现此时 $PC = PC + 8$ 。

同理，假如ARM9的五级流水线，把指令的Execute执行阶段，设计在了第四个阶段，那么就是 $PC = PC + (第4阶段 - 1) * 4个字节 = PC = PC + 12$ 了。

用图来说明 $PC = PC + 8$ 个过程

对于上面的文字的分析过程，可能看起来不是太容易理解，所以，下面这里通过图表来表示具体的流程，就更容易看懂了。其中，下图，是以ARM9的五级流水线的内部架构图为基础，而编辑的出来用于说明为何ARM9的五级流水线，也是 $PC = PC + 8$ ：

图 3.7. ARM9的五级流水线中为何 $PC = PC + 8$



对于上图中的，第一个指令在执行的时候，是使用到了PC的值，其实，我们可以看到，

对于指令在执行中，不论是否用到PC的值，PC都会按照既定逻辑，没一个cycle，自动增加4的，套用《非诚勿扰2》中的经典对白，即为：

你（指令执行的时候）用，

或者不用，

PC就在那里，

自动增4

所以，经过两个cycle的增4，就到了指令执行的时候，此时PC已经增加了8了，即使你指令执行的时候，没有用到PC的值，其也还是已经加了8了。而一般来说，大多数的指令，肯定也都是没有用到PC的，但是其实任何指令执行的那一时刻，也已经是 $PC=PC+8$ ，而多数指令没有用到，所以很多人没有注意到这点罢了。



PC (execute) = PC (fetch) + 8

对于 $PC=PC+8$ 中的两个PC，其实含义不完全一样。其更准确的表达，应该是这样：

$$PC (execute) = PC (fetch) + 8$$

其中：

PC (fetch)：当前正在执行的指令，就是之前取该指令时候的PC的值

PC (execute)：当前指令执行的计算中，如果用到PC，则此时PC的值。



不同阶段的PC值的关系

对应地，在ARM7的三级流水线（取指，译指，执行）和ARM9的五级流水线（取指，译指，执行，存储，写回）中，可以这么说：

PC，总是指向当前正在被取指的指令的地址，

PC-4，总是指向当前正在被译指的指令的地址，

PC-8，总是指向当前的那条指令，即我们一般说的，正在被执行的指令的地址。

【总结】

ARM7的三级流水线， $PC=PC+8$ ，

ARM9的五级流水线，也是 $PC=PC+8$ ，

根本的原因是，两者的流水线设计中，指令的Execute执行阶段，都是处于流水线的第三级。

所以使得 $PC=PC+8$ 。

类似地，可以推导出：

假设，Execute阶段处于流水线中的第E阶段，每条指令是T个字节，那么

PC

$$= PC + N * T$$

$$= PC + (E - 1) * T$$

此处ARM7和ARM9：

Execute阶段都是第3阶段 # E=3

每条指令是4个字节 # T=4

所以：

PC

$$= PC + N * T$$

$$= PC + (3 - 1) * 4$$

$$= PC + 8$$



关于直接改变PC的值，会导致流水线清空的解释

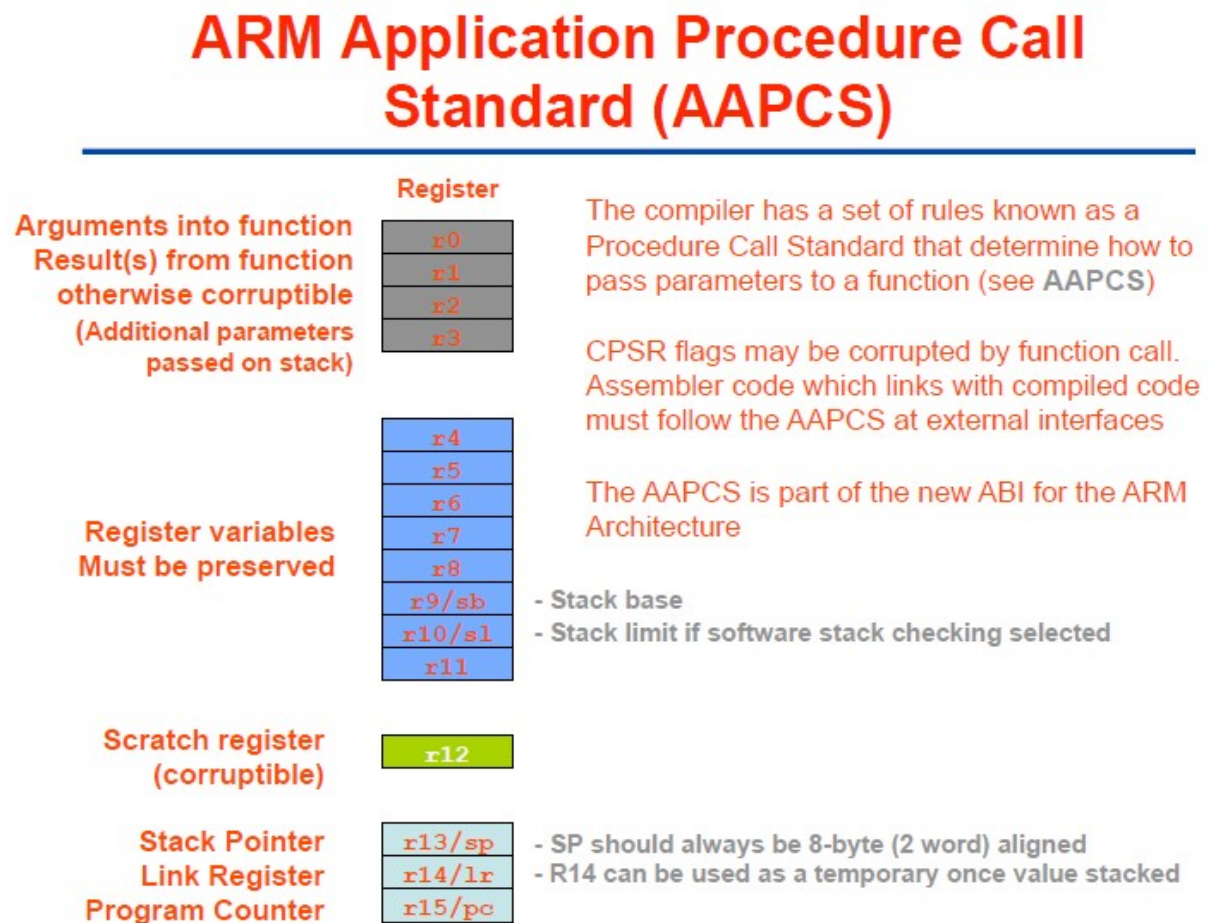
把PC的值直接赋值为0x20。而PC值更改，直接导致流水线的清空，即导致下一个cycle中的，对应的流水线中的其他几个步骤，包括接下来的同一个Cycle中的取指的工作被取消。在PC跳转到0x20的位置之后，流水线重新计算，重新一步步地按照流水线的逻辑，去一点执行。当然要保证当前指令的执行完成，即执行之后，还有两个cycle，分别做的Memory和Write，会继续执行完成。

3.5. AMR寄存器的别名 + APCS

此处简单介绍一下，ARM寄存器的别名，以及什么是APCS。

用文字解释之前，先看这个版本的解释，显得很直观，很好理解：

图 3.8. ARM Application Procedure Call Standard (AAPCS)



3.5.1. ARM中的寄存器的别名

默认的情况下，这些寄存器只是叫做r0,r1,...,r14等，而APCS 对其起了不同的别名。

使用汇编器预处理器的功能，你可以定义 R0 等名字，但在你修改其他人写的代码的时候，最好还是学习使用 APCS 名字。

一般编程过程中，最好按照其约定，使用对应的名字，这样使得程序可读性更好。

关于不同寄存器所对应的名字，见下表：

表 3.2. ARM寄存器的别名

寄存器名字		
Reg#	APCS	意义
R0	a1	工作寄存器
R1	a2	"
R2	a3	"
R3	a4	"
R4	v1	必须保护
R5	v2	"
R6	v3	"
R7	v4	"
R8	v5	"
R9	v6	"
R10	sl	栈限制
R11	fp	帧指针
R12	ip	内部过程调用寄存器
R13	sp	栈指针
R14	lr	连接寄存器
R15	pc	程序计数器

更加详细一点，见下：

[Predeclared register names](#)³

The following register names are predeclared:

1. **r0-r15** and **R0-R15**
2. **a1-a4** (argument, result, or scratch registers, synonyms for r0 to r3)
3. **v1-v8** (variable registers, r4 to r11)
4. **sb** and **SB** (static base, r9)
5. **ip** and **IP** (intra-procedure-call scratch register, r12)
6. **sp** and **SP** (stack pointer, r13)
7. **lr** and **LR** (link register, r14)
8. **pc** and **PC** (program counter, r15).

[Predeclared extension register names](#)⁴

The following extension register names are predeclared:

1. **d0-d31** and **D0-D31**(VFP double-precision registers)
2. **s0-s31** and **S0-S31**(VFP single-precision registers)

³ http://www.keil.com/support/man/docs/armasm/armasm_ch03s03s01.htm

⁴ http://www.keil.com/support/man/docs/armasm/armasm_ch03s03s02.htm

[Predeclared coprocessor names](#)⁵

The following coprocessor names and coprocessor register names are predeclared:

1. p0-p15 (coprocessors 0-15)
2. c0-c15 (coprocessor registers 0-15).

3.5.2. 什么是APCS

APCS，ARM 过程调用标准(ARM Procedure Call Standard)，提供了紧凑的编写例程的一种机制，定义的例程可以与其他例程交织在一起。最显著的一点是对这些例程来自哪里没有明确的限制。它们可以编译自 C、Pascal、也可以是用汇编语言写成的。

APCS 定义了：

- 对寄存器使用的限制。
- 使用栈的惯例。
- 在函数调用之间传递/返回参数。
- 可以被“回溯”的基于栈的结构格式，用来提供从失败点到程序入口的函数(和给予的参数)的列表。

3.6. 为何C语言（的函数调用）需要堆栈，而汇编语言却不需要堆栈

之前看了很多关于uboot的分析，其中就有说要为C语言的运行，准备好堆栈。

而自己在Uboot的start.S汇编代码中，关于系统初始化，也看到有堆栈指针初始化这个动作。但是，从来只是看到有人说系统初始化要初始化堆栈，即正确给堆栈指针sp赋值，但是却从来没有看到有人解释，为何要初始化堆栈。所以，接下来的内容，就是经过一定的探究，试图来解释一下，为何要初始化堆栈，即：

为何C语言的函数调用要用到堆栈，而汇编却不需要初始化堆栈。

要明白这个问题，首先要了解堆栈的作用。

关于堆栈的作用，要详细讲解的话，要很长的篇幅，所以此处只是做简略介绍。

总的来说，堆栈的作用就是：保存现场/上下文，传递参数。

3.6.1. 保存现场/上下文

现场，意思就相当于案发现场，总有一些现场的情况，要记录下来的，否则被别人破坏掉之后，你就无法恢复现场了。而此处说的现场，就是指CPU运行的时候，用到了一些寄存器，比如r0,r1等等，对于这些寄存器的值，如果你不保存而直接跳转到子函数中去执行，那么很可能就被其破坏了，因为其函数执行也要用到这些寄存器。

因此，在函数调用之前，应该将这些寄存器等现场，暂时保持起来，等调用函数执行完毕返回后，再恢复现场。这样CPU就可以正确的继续执行了。

在计算机中，你常可以看到上下文这个词，对应的英文是context。那么：

3.6.1.1. 什么叫做上下文context

保存现场，也叫保存上下文。

⁵ http://www.keil.com/support/man/docs/armasm/armasm_ch03s03s03.htm

上下文，英文叫做context，就是上面的文章，和下面的文章，即与你此刻，当前CPU运行有关系的内容，即那些你用到寄存器。所以，和上面的现场，是一个意思。

保存寄存器的值，一般用的是push指令，将对应的某些寄存器的值，一个个放到堆栈中，把对应的值压入到堆栈里面，即所谓的**压栈**。

然后待被调用的子函数执行完毕的时候，再调用pop，把堆栈中的一个一个的值，赋值给对应的那些你刚开始压栈时用到的寄存器，把对应的值从堆栈中弹出去，即所谓的**出栈**。

其中保存的寄存器中，也包括lr的值（因为用bl指令进行跳转的话，那么之前的pc的值是存在lr中的），然后在子程序执行完毕的时候，再把堆栈中的lr的值pop出来，赋值给pc，这样就实现了子函数的正确的返回。

3.6.2. 传递参数

C语言进行函数调用的时候，常常会传递给被调用的函数一些参数，对于这些C语言级别的参数，被编译器翻译成汇编语言的时候，就要找个地方存放一下，并且让被调用的函数能够访问，否则就没法实现传递参数了。对于找个地方放一下，分两种情况。

一种情况是，本身传递的参数就很少，就可以通过寄存器传送参数。

因为在前面的保存现场的动作中，已经保存好了对应的寄存器的值，那么此时，这些寄存器就是空闲的，可以供我们使用的了，那就可以放参数，而参数少的情况下，就足够存放参数了，比如参数有2个，那么就用r0和r1存放即可。（关于参数1和参数2，具体哪个放在r0，哪个放在r1，就是和APCS中的“在函数调用之间传递/返回参数”相关了，APCS中会有详细的约定。感兴趣的自己去研究。）

但是如果参数太多，寄存器不够用，那么就得把多余的参数堆栈中了。

即，可以用堆栈来传递所有的或寄存器放不下的那些多余的参数。

3.6.3. 举例分析C语言函数调用是如何使用堆栈的

对于上面的解释的堆栈的作用显得有些抽象，此处再用例子来简单说明一下，就容易明白了：

用：

```
arm-inux-objdump -d u-boot > dump_u-boot.txt
```

可以得到dump_u-boot.txt文件。该文件就是中，包含了u-boot中的程序的可执行的汇编代码，其中我们可以看到C语言的函数的源代码，到底对应着那些汇编代码。

下面贴出两个函数的汇编代码，

一个是clock_init，另一个是与clock_init在同一C源文件中的，另外一个函数CopyCode2Ram

```
33d0091c <CopyCode2Ram>:
33d0091c: e92d4070 push {r4, r5, r6, lr}❶
33d00920: e1a06000 mov r6, r0
33d00924: e1a05001 mov r5, r1
33d00928: e1a04002 mov r4, r2
33d0092c: ebffffef bl 33d008f0 <bBootFrmNORFlash>❷
... ..
33d00984: ebffff14 bl 33d005dc <nand_read_ll>
... ..
33d009a8: e3a00000 mov r0, #0 ; 0x0❸
33d009ac: e8bd8070 pop {r4, r5, r6, pc}❹

33d009b0 <clock_init>:
33d009b0: e3a02313 mov r2, #1275068416 ; 0x4c000000❺
```

```
33d009b4: e3a03005  mov r3, #5 ; 0x5
33d009b8: e5823014  str r3, [r2, #20]
... ..
33d009f8: e1a0f00e  mov pc, lr⑥
```

❶ 此处就是我们所期望的，用push指令，保存了r4,r5,r以及lr。

用push去保存r4,r5,r6，那是因为所谓的保存现场，以后后续函数返回时候再恢复现场，

❷ 上述用push去保存lr，那是因为函数CopyCode2Ram里面在此处调用了bBootFrmNORFlash以及也调用了nand_read_ll：

```
33d00984: ebffff14  bl 33d005dc <nand_read_ll>
```

也用到了bl指令，会改变我们最开始进入clock_init时候的lr的值，所以我们要用push也暂时保存起来。

❸ 把0赋值给r0寄存器，这个就是我们所谓返回值的传递，是通过r0寄存器的。

此处的返回值是0，也对应着C语言的源码中的“return 0”。

❹ 把之前push的值，给pop出来，还给对应的寄存器，其中最后一个是将开始push的lr的值，pop出来给赋给PC，因为实现了函数的返回。

❺ 可以看到此处是该函数第一行

其中没有我们所期望的push指令，没有去将一些寄存器的值放到堆栈中。这是因为，我们clock_init这部分的内容，所用到的r2,r3等寄存器，和前面调用clock_init之前所用到的寄存器r0，没有冲突，所以此处可以不用push去保存这类寄存器的值，不过有个寄存器要注意，那就是r14，即lr，其是在前面调用clock_init的时候，用的是bl指令，所以会自动把跳转时候的pc的值赋值给lr，所以也不需要push指令去将PC的值保存到堆栈中。

❻ 而此处是clock_init的代码的最后一行

就是我们常见的mov pc, lr，把lr的值，即之前保存的函数调用时候的PC值，赋值给现在的PC，这样就实现了函数的正确的返回，即返回到了函数调用时候下一个指令的位置。

这样CPU就可以继续执行原先函数内剩下那部分的代码了。



对于使用哪个寄存器来传递返回值

当然你也可以用其他暂时空闲没有用到的寄存器来传递返回值，但是这些处理方式，本身是根据ARM的APCS的寄存器的使用的约定而设计的，你最好不要随便改变使用方式，最好还是按照其约定的来处理，这样程序更加符合规范。

3.7. 关于为何不直接用mov指令，而非要用adr伪指令

在分析uboot的start.S中，看到一些指令，比如：

```
adr r0, _start
```

觉得好像可以直接用mov指令实现即可，为啥还要这么麻烦地，去用ldr去实现？

关于此处的代码，为何要用adr指令：

```
adr r0, _start
```



注意

其被编译器编译后，会被翻译成：

```
sub r0, pc, #172
```

而不直接用mov指令直接将_start的值赋值给r0，类似于这样：

```
mov r0, _start
```

呢？

其原因主要是，

```
sub r0, pc, #172
```

这样的代码，所处理的值，都是相对于PC的偏移量来说的，这样的代码中，没有绝对的物理地址值，都是相对的值，利用产生位置无关代码。因为如果用mov指令：

```
mov r0, _start
```

那么就会被编译成这样的代码：

```
mov r0, 0x33d00000
```

如果用了上面这样的代码：

```
mov r0, 0x33d00000
```

那么，如果整个代码，即要执行的程序的指令，被移动到其他位置，那么

```
mov r0, 0x33d00000
```

这行指令，执行的功能，就是跳转到绝对的物理地址，而不是跳转到相对的_start的位置了，就不能实现我们想要的功能了，这样包含了绝对物理地址的代码，也就不是位置无关的代码了。

与此相对，这行指令：

```
sub r0, pc, #172
```

即使程序被移动到其他位置，那么该行指令还是可以跳转到相对PC往前172字节的地方，也还是我们想要的_start的位置，这样包含的都是相对的偏移位置的代码，就叫做位置无关代码。其优点就是不用担心你的代码被移动，即使程序的基地址变了，所有的代码的相对位置还是固定的，程序还是可以正常运行的。

关于，之所以不用上面的：

```
mov r0, 0x33d00000
```

类似的代码，除了上面说的，不是位置无关的代码之外，其还有个潜在的问题，那就是，关于mov指令的源操作数，此处即为0x33d00000，不一定是合法的mov指令所允许的值，这也正是下面要详细解释的内容[第 3.8 节 “mov指令的操作数的取值范围到底是多少”](#)

【总结】

之所以用adr而不用mov，主要是为了生成地址无关代码，以及由于不方便判断一个数，是否是有效的mov的操作数。

3.8. mov指令的操作数的取值范围到底是多少

关于mov指令操作数的取值范围，网上看到一些人说是0x00-0xFF,也有人说是其他的值的，但是经过一番求证，发现这些说法都不对。下面就是来详细解释，mov指令的操作数的取值范围，到底是多少。

在看了我说的，关于这行代码：

```
mov r0, 0x33d00000
```

的源操作数0x33d0000，可能是mov指令所不允许的，这句话后，可能有人会说，我知道，那是因为mov的操作数的值，不允许大于255,至少网上很多人的资料介绍中，都是这么说的。

对此，要说的是，你的回答是错误的。

关于mov操作数的真正的允许的取值范围，还真的不是那么容易就能搞懂的，下面就来详细解释解释。

总的来说，我是从[ARM 汇编的mov操作立即数的疑问](http://blog.chinaunix.net/space.php?uid=20799298&do=blog&cuid=2055392)⁶

里面，才算清楚mov的取值范围，以及找了相应的datasheet，才最终看懂整个事情的来龙去脉的。

首先，mov的指令，是属于ARM指令集中，数据处理（Data Process）分类中的其中一个指令，

而数据处理指令的具体格式是：[ARM Processor Instruction Set](http://netwinder.osuosl.org/pub/netwinder/docs/arm/ARM7500FvB_3.pdf)⁷

图 3.9. 数据处理指令的指令格式

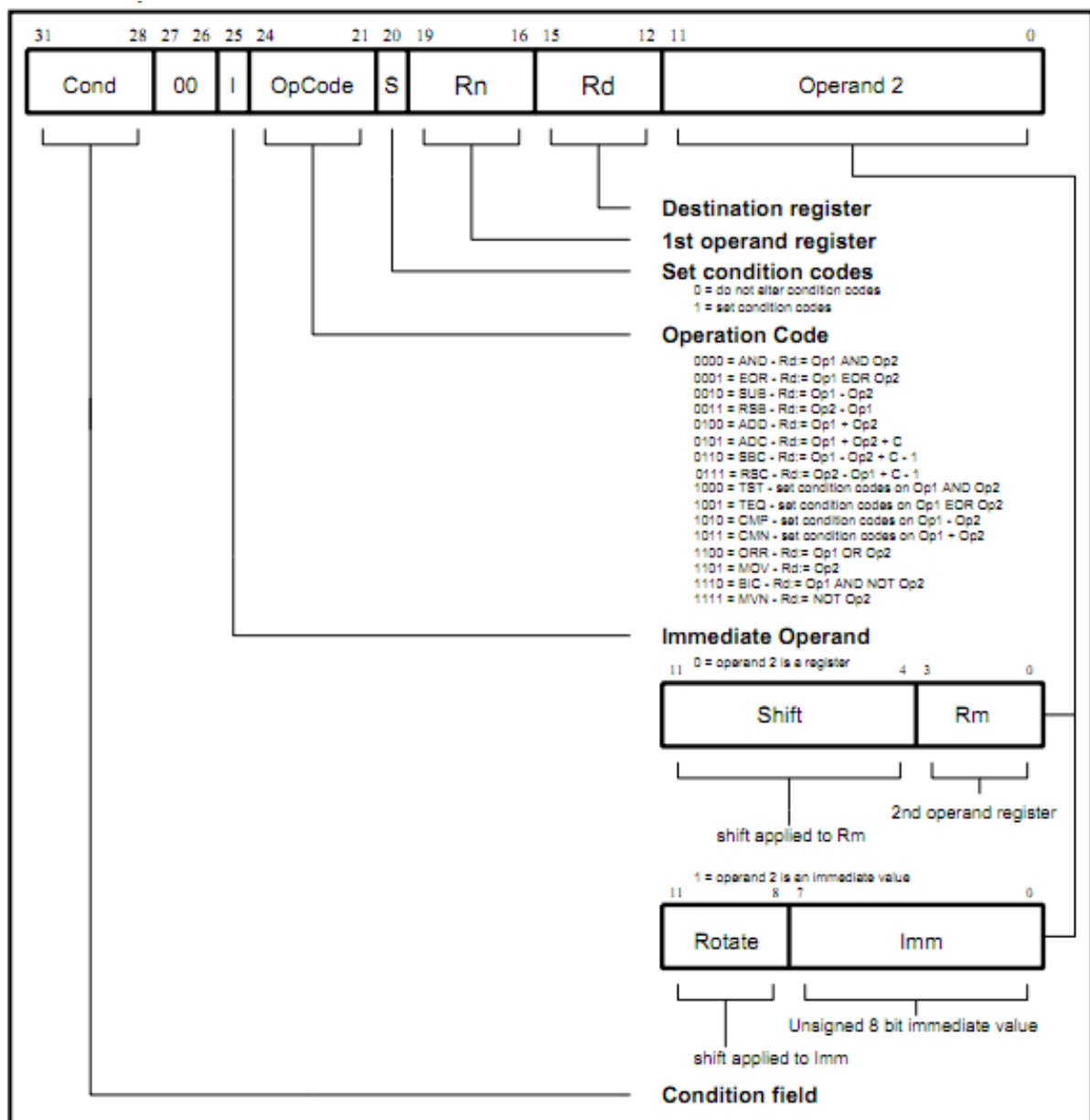


Figure 5-4: Data processing instructions

⁶ <http://blog.chinaunix.net/space.php?uid=20799298&do=blog&cuid=2055392>

⁷ http://netwinder.osuosl.org/pub/netwinder/docs/arm/ARM7500FvB_3.pdf

对于此格式，我们可以拿：

```
arm-linux-objdump -d u-boot > dump_u-boot.txt
```

中得到的汇编代码中关于：

```
ldr    r0, =0x53000000
```

所对应的，真正的汇编代码：

```
33d00068: e3a00453  mov r0, #1392508928 ; 0x53000000
```

来分析，就容易看懂了：

```
mov r0, #1392508928
```

```
= mov r0, #0x53000000
```

的作用就是，把0x53000000移动到r0中去。

其对应的二进制指令是上面的：

```
0xe3a00453 = 1110 0011 1010 0000 0000 0100 0101 0011 b
```

下面对照mov指令的格式，来分析这些位所对应的含义：

表 3.3. mov指令0xe3a00453的位域含义解析

31-28	27-26	25	24-21	20	19-16	15-12	11-0
Condition Field	Op2	I(Immediate Operand)	OpCode(Condition Code)	Op2	Rn(1st Operand Register)	Rd(Destination Register)	on Operand 2
							1 = operand 2 is an immediate value
							11-8: Rotate
1110	00	1	1101	0	0000	0000	0100 0101 0011
		表明是立即数	1101对应的是MOV指令		MOV指令做的事情是：Rd:=Op2,和Rn无关，所以忽略这个Rn	表示0000号寄存器，即r0	0100=4, 0x53含义参见注释1



注意

上述datasheet中写到：

5.4.3 Immediate operand rotates

The immediate operand rotate field is a 4 bit unsigned integer which specifies a shift operation on the 8 bit immediate value. This value is zero extended to 32 bits, and then subject to a **rotate right** by **twice** the value in the rotate field. This enables many common constants to be generated, for example all powers of 2

意思是，对于bit[11:8]的值，是个4位，无符号的整型，其指定了bit[7:0]的8bit立即数值的位移操作。具体如何指定呢，那就是将bit[7:0]的值，循环右移2x bit[11:8]位。

对于我们的例子，就是，将bit[7:0]的值0x53,循环右移 2xbit[11:8]= 2 x 4 = 8位，

而0x53循环右移8位，就得到了0x53000000，就是我们要mov值，mov到目的寄存器rd，此处为r0中。

而上面英文最后一句说的是，通过将bit[7:0]的值,循环右移 2xbit[11:8]的方式，就可以产生出很多个数值了，即mov的操作数中，其中符合可以通过0x00-0xFF循环右移偶数位而产生的数值，都是合法的mov的操作数，而这样的数，其实是很多的。



【总结】

所以，**mov指令的操作数的真正的取指范围**，即不是0-0xFF（0-255），也不是只有2的倍数，而是：

只要该数，可以通过0x00-0xFF中某个数，循环右移偶数位而产生，就是合法的mov的操作数，否则就是非法的mov的操作数。

3.9. 汇编学习总结记录

对于我们之前分析的start.S中，涉及到很多的汇编的语句，其中，可以看出，很多包含了很多种不同的语法，使用惯例等，下面，就对此进行一些总结，借以实现一定的举一反三或者说触类旁通，这样，可以起到一定的借鉴功能，方便以后看其他类似汇编代码，容易看懂汇编代码所要表达的含义。

3.9.1. 汇编中的标号=C中的标号

像前面汇编代码中，有很多的，以点开头，加上一个名字的形式标号，比如：

```
reset:
/*
 * set the cpu to SVC32 mode
 */
mrs r0,cpsr
```

中的reset，就是汇编中的标号，相对来说，比较容易理解，就相当于C语言的标号。

比如，C语言中定义一个标号ERR_NODEV：

```
ERR_NODEV: /* no device error */
... /* c code here */
```

然后对应在别处，使用goto去跳转到这个标号ERR_NODEV：

```
if (something)
goto ERR_NODEV;
```



【总结】

汇编中的标号 = C语言中的标号Label

3.9.2. 汇编中的跳转指令=C中的goto

对应地，和上面的例子中的C语言中的编号和掉转到标号的goto类似，汇编中，对于定义了标号，那么也会有对应的指令，去跳转到对应的汇编中的标号。

这些跳转的指令，就是b指令，b是branch的缩写。

b指令的格式是：

```
b{cond} label
```

简单说就是跳转到label处。

用和上面的例子相关的代码来举例：

```
.globl _start
_start: b    reset
```

就是用b指令跳转到上面那个reset的标号。



【总结】

汇编中的b跳转指令 = C语言中的goto

3.9.3. 汇编中的.global=C语言中的extern

对于上面例子中：

```
.globl _start
```

中的.global，就是声明_start为全局变量/标号，可以供其他源文件所访问。

即汇编器，在编译此汇编代码的时候，会将此变量记下来，知道其是个全局变量，遇到其他文件是用到此变量的时候，知道是访问这个全局变量的。

因此，从功能上来说，就相当于C语言用extern去声明一个变量，以实现本文件外部访问此变量。



【总结】

汇编中的.global或.global = C语言中的extern

3.9.4. 汇编中用bl指令和mov pc, lr来实现子函数调用和返回

和b指令类似的，另外还有一个bl指令，语法是：

```
BL{cond} label
```

其作用是，除了b指令跳转到label之外，在跳转之前，先把下一条指令地址存到lr寄存器中，以方便跳转到那边执行完毕后，将lr再赋值给pc，以实现函数返回，继续执行下面的指令的效果。

用下面这个start.S中的例子来说明：

```
bl cpu_init_crit
.....
cpu_init_crit:
.....
mov pc, lr
```


其中，就是先调用bl掉转到对应的标号cpu_init_crit，其实就是相当于一个函数了，

然后在cpu_init_crit部分，执行完毕后，最后调用 mov pc, lr，将lr中的值，赋给pc，即实现函数的返回原先 bl cpu_init_crit下面那条代码，继续执行函数。

上面的整个过程，用C语言表示的话，就相当于

```
.....
cpu_init_crit();
.....

void cpu_init_crit(void)
{
.....
}
```

而关于C语言中，**函数的跳转前后所要做的事情**，都是C语言编译器帮我们实现好了，会将此C语言中的函数调用，转化为对应的汇编代码的。

其中，此处所说的，函数掉转前后所要做的事情，就是：

- **函数跳转前**

要将当前指令的下一条指令的地址，保存到lr寄存器中

- **函数调用完毕后**

将之前保存的lr的值给pc，实现函数跳转回来。继续执行下一条指令。

而如果你本身自己写汇编语言的话，那么这些函数跳转前后要做的事情，都是你程序员自己要关心，要实现的事情。



【总结】

汇编中bl + mov pc, lr = C语言中的子函数调用和返回

3.9.5. 汇编中的对应位置有存储值的标号 = C语言中的指针变量

像前文所解析的代码中类似于这样的：

```
LABEL1 : .word Value2
```

比如：

```
_TEXT_BASE:
.word TEXT_BASE
```

所对应的含义是，有一个标号_TEXT_BASE

而该标号中对应的位置，所存放的是一个word的值，具体的数值是TEXT_BASE，此处的TEXT_BASE是在别处定义的一个宏，值是0x33D00000。

所以，即为：

有一个标号_TEXT_BASE，其对应的位置中，所存放的是一个word的值，值为

```
TEXT_BASE=0x33D00000
```

总的来说，此种用法的含义，如果用C语言来表示，其实更加容易理解：

```
int *_TEXT_BASE = TEXT_BASE = 0x33D00000
```

即：

```
int *_TEXT_BASE = 0x33D00000
```



C语言中如何引用汇编中的标号

不过，对于这样的类似于C语言中的指针的汇编中的标号，在C语言中调用到的话，却是这样引用的：

```
/* for the following variables, see start.S */
extern ulong _armboot_start; /* code start */
extern ulong _bss_start; /* code + data end == BSS start */
.....
IRQ_STACK_START = _armboot_start - CFG_MALLOC_LEN - CFG_GBL_DATA_SIZE -
4;
.....
```

而不是我原以为的，直接当做指针来引用该变量的方式：

```
*IRQ_STACK_START = *_armboot_start - CFG_MALLOC_LEN - CFG_GBL_DATA_SIZE
- 4;
```

其中，对应的汇编中的代码为：

```
.globl _armboot_start
_armboot_start:
.word _start
```

所以，针对这点，还是需要注意一下的。至少以后如果自己写代码的时候，在C语言中引用汇编中的global的标号的时候，知道是如何引用该变量的。



【总结】

汇编中类似这样的代码：

```
label1: .word value2
```

就相当于C语言中的：

```
int *label1 = value2
```

但是在C语言中引用该标号/变量的时候，却是直接拿来用的，就像这样：

```
label1 = other_value
```

其中label1就是个int型的变量。

3.9.6. 汇编中的ldr+标号，来实现C中的函数调用

接着上面的内容，继续解释，对于汇编中这样的代码：

第一种：

```
ldr pc, 标号1
.....
标号1: .word 标号2
.....
标号2:
..... ( 具体要执行的代码 )
```

或者是，

第二种：

```
ldr pc, 标号1
.....
标号1: .word XXX ( C语言中某个函数的函数名 )
```

的意思就是，将地址为标号1中内容载入到pc中。

而地址为标号1中的内容，就是标号2。

```
TEXT_BASE=0x33D00000
```

所以上面第一种的意思:

就很容易看出来，就是把标号2这个地址值，给pc，即实现了跳转到标号2的位置执行代码，

就相当于调用一个函数，该函数名为标号2。

第二种的意思，和上面类似，是将C语言中某个函数的函数名，即某个地址值，给pc，实现调用C中对应的那个函数。

两种做法，其含义用C语言表达，其实很简单：

$PC = * (\text{标号1}) = \text{标号2}$

例 3.1. 汇编中的ldr加标号实现函数调用 示例

举个例子就是：

第一种：

```
.....
ldr pc, _software_interrupt
.....
_software_interrupt: .word software_interrupt
.....
software_interrupt:
get_bad_stack
bad_save_user_regs
```

```
bl do_software_interrupt
```

就是实现了将标号1, _software_interrupt, 对应的位置中的值, 标号2, software_interrupt, 给pc, 即实现了将pc掉转到software_interrupt的位置, 即实现了调用函数software_interrupt的效果。

第二种：

```
ldr pc, _start_armboot
_start_armboot: .word start_armboot
```

含义就是, 将标号1, _start_armboot, 所对应的位置中的值, start_armboot给pc, 即实现了调用函数start_armboot的目的。

其中, start_armboot是C语言文件中某个C语言的函数。



【总结】

汇编中, 实现函数调用的效果, 有如下两种方法：

1. 方法1

```
ldr pc, 标号1
.....
标号1: .word 标号2
.....
标号2:
..... ( 具体要执行的代码 )
```

2. 方法2

```
ldr pc, 标号1
.....
标号1: .word XXX ( C语言中某个函数的函数名 )
```

3.9.7. 汇编中设置某个寄存器的值或给某个地址赋值

在汇编代码start.S中, 看到不止一处, 类似于这样的代码：

形式1：

```
# define pWTCON 0x53000000
.....
ldr r0, =pWTCON
mov r1, #0x0
str r1, [r0]
```

或者是,

形式2：

```
# define INTSUBMSK 0x4A00001C
.....
ldr r1, =0x7fff
ldr r0, =INTSUBMSK
str r1, [r0]
```

其含义，都是将某个值，赋给某个地址，此处的地址，是用宏定义来定义的，对应着某个寄存器的地址。

其中，形式1是直接通过mov指令来将0这个值赋给r1寄存器，和形式2中的通过ldr伪指令来将0x3fff赋给r1寄存器，两者区别是，前者是因为已经确定所要赋的值0x0是mov的有效操作数，而后者对于0x3fff不确定是否是mov的有效操作数



警告

如果不是，则该指令无效，编译的时候，也无法通过编译，会出现类似于这样的错误：

```
start.S: Assembler messages:
start.S:149: Error: invalid constant -- 'mov r1,#0xFFEFDFFF'
make[1]: *** [start.o] 错误 1
make: *** [cpu/arm920t/start.o] 错误 2
```

所以才用ldr伪指令，让编译器来帮你自动判断：

1. 如果该操作数是mov的有效操作数，那么ldr伪指令就会被翻译成对应的mov指令

例 3.2.

举例说明：

汇编代码：

```
# define pWTCON 0x53000000
.....
ldr r0, =pWTCON
```

被翻译后的真正的汇编代码：

```
33d00068: e3a00453 mov r0, #1392508928 ; 0x53000000
```

2. 如果该操作数不是mov的有效操作数，那么ldr伪指令就会被翻译成ldr指令

例 3.3.

举例说明：

汇编代码：

```
ldr r1, =0x7fff
```

被翻译后的真正的汇编代码：

```
33d00080: e59f13f8 ldr r1, [pc, #1016] ; 33d00480 <fiq+0x60>
.....
33d00480: 00007fff .word 0x00007fff
```

即把ldr伪指令翻译成真正的ldr指令，并且另外分配了一个word的地址空间用于存放该数值，然后用ldr指令将对应地址中的值载入，赋值给r1寄存器。



【总结】

汇编中，一个常用的，用来给某个地址赋值的方法，类似如下形式：

```
#define 宏的名字 寄存器地址
.....
ldr r1, =要赋的值
ldr r0, =宏的名字
str r1, [r0]
```

参考书目

[1] [2010年6月 最新TQ2440光盘下载 \(Linux内核, WinCE的eboot, uboot均有更新\)](#)¹

[2] [.globl,.word,.balignl的语法](#)²

[3] [label的解释](#)³

[4] [ldr的语法](#)⁴

[5] [ldr指令](#)⁵

[6] [ldr指令](#)⁶

[7] [.word的语法](#)⁷

[8] [ARM7体系结构](#)⁸

[9] [bootloader](#)⁹

[10] [S3C2440相关的软硬件资料](#)¹⁰

[11] [S3C2440的CPU的datasheet : s3c2440a_um_rev014_040712.pdf](#)¹¹

[12] [伪指令ldr语法和含义](#)¹²

[13] [ARM9 2410移植之ARM中断原理, 中断嵌套的误区, 中断号的怎么来的](#)¹³

[14] [adr指令的语法和含义](#)¹⁴

[15] [ARM协处理器](#)¹⁵

[16] [ARM920T](#)¹⁶

[17] [CP15的各个寄存器的含义解释](#)¹⁷

[18] [Invalidate ICache and DCache](#)¹⁸

[19] [Invalidate TLB\(s\)](#)¹⁹

[20] [Control register](#)²⁰

¹ <http://bbs.embedsky.net/viewthread.php?tid=859>

² <http://re-eject.gbadev.org/files/GasARMRef.pdf>

³ <http://sourceware.org/binutils/docs-2.20/as/Labels.html#Labels>

⁴ http://infocenter.arm.com/help/topic/com.arm.doc.dui0206hc/DUI0206HC_rvct_linker_and_utilities_guide.pdf

⁵ <http://wenku.baidu.com/view/f7cc280102020740be1e9bea.html>

⁶ <http://www.pczpg.com/a/2010/0607/11062.html>

⁷ http://blogold.chinaunix.net/u3/115924/showart_2280163.html

⁸ <http://www.docin.com/p-73665362.html>

⁹ <http://www.csie.nctu.edu.tw/~wjtsai/EmbeddedSystemDesign/Ch2-bootloader.pdf>

¹⁰ <http://just4you.springnote.com/pages/1052612>

¹¹ <http://just4you.springnote.com/pages/1052612/attachments/803220>

¹² <http://blog.csdn.net/lihaoweiv/archive/2010/11/24/6033003.aspx>

¹³ http://www.crifan.com/switch_arm9_2410_transplant_arm_interrupt_principle_the_error_interrupt_nesting_how_come_the_interrupt_number/

¹⁴ http://blog.mcuol.com/User/cdkfGao/article/8057_1.htm

¹⁵ <http://apps.hi.baidu.com/share/detail/32319228>

¹⁶ http://infocenter.arm.com/help/topic/com.arm.doc.ddi0151c/ARM920T_TRM1_S.pdf

¹⁷ <http://www.heyrick.co.uk/assembler/coprocmnd.html>

¹⁸ <http://infocenter.arm.com/help/index.jsp?topic=/com.arm.doc.ddi0184b/Chdcfejb.html>

¹⁹ <http://infocenter.arm.com/help/index.jsp?topic=/com.arm.doc.ddi0184b/Chdifbjc.html>

²⁰ <http://infocenter.arm.com/help/index.jsp?topic=/com.arm.doc.ddi0184b/Chdifbjc.html>

- [21] [Domain access control](#)²¹
- [22] [ARM920T的CPU的7种模式](#)²²
- [23] [ARM Linux Kernel Boot Requirements](#)²³
- [24] [嵌入式系统之WATCHDOG\(看门狗\)概述](#)²⁴
- [25] [ARM流水线和program counter\(PC\)的增量](#)²⁵
- [26] [Predeclared register names](#)²⁶
- [27] [Predeclared extension register names](#)²⁷
- [28] [Predeclared coprocessor names](#)²⁸
- [29] [mov的操作数的取指范围](#)²⁹
- [30] [ARM Processor Instruction Set](#)³⁰
- [31] [ARM9流水线PC=PC+8](#)³¹
- [32] [Strange behaviour of ldr \[pc, #value\]](#)³²

²¹ <http://infocenter.arm.com/help/index.jsp?topic=/com.arm.doc.ddi0151c/I273867.html>
²² <http://www.docin.com/p-73665362.html>
²³ <http://www.arm.linux.org.uk/developer/booting.php>
²⁴ <http://wenku.baidu.com/view/e5cd52ff04a1b0717fd5dd27.html>
²⁵ <http://hi.baidu.com/istry/blog/item/f823e1438de0a71972f05d0f.html>
²⁶ http://www.keil.com/support/man/docs/armasm/armasm_ch03s03s01.htm
²⁷ http://www.keil.com/support/man/docs/armasm/armasm_ch03s03s02.htm
²⁸ http://www.keil.com/support/man/docs/armasm/armasm_ch03s03s03.htm
²⁹ <http://blog.chinaunix.net/space.php?uid=20799298&do=blog&cuid=2055392>
³⁰ http://netwinder.osuosl.org/pub/netwinder/docs/arm/ARM7500FEvB_3.pdf
³¹ <http://blog.csdn.net/hamilton1/archive/2011/02/18/6192722.aspx>
³² <http://stackoverflow.com/questions/2102921/strange-behaviour-of-ldr-pc-value>