目　　录

“黑色经典”系列之《嵌入式 Linux 应用程序开发详解》

第 1 章 Linux 快速入门

本章目标

嵌入式 Linux 是以 Linux 为基础的操作系统，只有对 Linux 系统有了较为熟练的使用之后，才能在嵌入式 Linux 开发领域得心应手。通过本章的学习，读者能够掌握如下内容。

- 能够独立安装 Linux 操作系统
- 能够熟练使用 Linux 系统的基本命令
- 认识 Linux 系统启动过程
- 能够独立在 Linux 系统中安装软件
- 能够独立设置 Linux 环境变量
- 能够独立定制 Linux 服务

1.1 嵌入式 Linux 基础

随着摩托罗拉手机 A760、IBM 智能型手表 WatchPad、夏普 PDA Zaurus 等一款款高性能“智能数码产品”的出现，以及 Motolola、三星、MontaVista、飞利浦、Nokia、IBM、SUN 等众多国际顶级巨头的加入，嵌入式 Linux 的队伍越来越庞大了。在通信、信息、数字家庭、工业控制等领域，随处都能见到嵌入式 Linux 的身影。根据美国 VDC（Venture Development Co.）统计数据显示，嵌入式 Linux 的市场规模从 2001 年的 5520 万美元，到 2006 将会增长至 3.46 亿美元，在未来两年将占嵌入式操作系统市场份额的 50%。

究竟是什么原因让嵌入式 Linux 发展如此迅速呢？又究竟是什么原因让它能与强劲的 Vxworks、Window CE 相抗衡呢？这一切还是要归根于它的父亲——Linux 的功劳。可以说，嵌入式 Linux 正是继承和发展了 Linux 的诱人之处才走到今天的，而 Linux 也正是有了嵌入式 Linux 的广泛应用才使其更加引人瞩目。以下就从 Linux 开始，一层层揭开嵌入式 Linux 的面纱。

1.1.1 Linux 发展概述

简单地说，Linux 是指一套免费使用和自由传播的类 UNIX 操作系统。人们通常所说的 Linux 是指 Linus Torvalds 所写的 Linux 操作系统内核。

当时的 Linus 还是芬兰赫尔辛基大学的一名学生，他主修的课程中有一门课是操作系统，而且这门课是专门研究程序的设计和执行。最后这门课程提供了一种称为 Minix 的初期 UNIX 系统。Minix 是一款仅为教学而设计的操作系统，而且功能有限。因此，和 Minix 的众多使用者一样，Linus 也希望能给它添加一些功能。

在之后的几个月里，Linus 根据实际的需要，编写了磁盘驱动程序以便下载访问新闻组的文件，又写了个文件系统以便能够阅读 Minix 文件系统中的文件。这样，“当你有了任务切换，有了文件系统和设备驱动程序后，这就是 UNIX，或者至少是其内核。”于是，0.0.1 版本的 Linux 就诞生了。

Linus 从一开始就决定自由传播 Linux，他把源代码发布在网上，于是，众多的爱好者和程序员也都通过互联网加入到 Linux 的内核开发工作中。这个思想与 FSF（Free Software Foundation）资助发起的 GNU（GNU’s Not UNIX）的自由软件精神不谋而合。

GNU 是为了推广自由软件的精神以实现一个自由的操作系统，然后从应用程序开始，实现其内核。而当时 Linux 的优良性能备受 GNU 的赏识，于是 GNU 就决定采用 Linus 及其开发者的内核。在他们的共同努力下，Linux 这个完整的操作系统诞生了。其中的程序开发共同遵守 General Public License（GPL）协议，这是最开放也是最严格的许可协议方式，这个协议规定了源码必须可以无偿的获取并且修改。因此，从严格意义上说，Linux 应该叫做 GNU/Linux，其中许多重要的工具如 gcc、gdb、make、Emacs 等都是 GNU 贡献。

这个“婴儿版”的操作系统以平均两星期更新一次的速度迅速成长，如今的 Linux 已经有超过 250 种发行版本，且可以支持所有体系结构的处理器，如 X86、PowerPC、ARM、XSCALE 等，也可以支持带 MMU 或不带 MMU 的处理器。到目前为止，它的内核版本也已经从原先的 0.0.1 发展到现在的 2.6.xx。

自由软件（free software）中的 free 并不是指免费，而是指自由。它赋予使用者四种自由。

- 自由之一：有使用软件的自由。
- 自由之二：有研究该软件如何运作的自由，并且得以改写该软件来符合使用者自身的需求。取得该软件的源码是达成此目的前提。
- 自由之三：有重新散布该软件的自由，所以每个人都可以藉由散布自由软件来敦亲睦另。
- 自由之四：有改善再利用该软件的自由，并且可以发表改写版供公众使用，如此一来，整个社群都可以受惠。如前项，取得该软件的源码是达成此目的前提。

小知识

GPL：GPL 协议是 GNU 组织、维护的一种版权协议，遵守这个协议的软件可以自由地获取、查看、使用其源代码。GPL 协议是整个开源世界的精神基础。

Linux 的内核版本号：

Linux 内核版本号格式是 x.y.zz-www，数字 x 代表版本类型，数字 y 为偶数时是稳定版本，为奇数时是开发版本，如 2.0.40 为稳定版本，2.3.42 为开发版本，测试版本为 3 个数字加上测试号，如 2.4.12-rc1。最新的 Linux 内核版本可从 http://www.kernel.org 上获得。

1.1.2 Linux 作为嵌入式操作系统的优势

从 Linux 系统的发展过程可以看出，Linux 从最开始就是一个开放的系统，并且它始终遵循着源代码开放的原则，它是一个成熟而稳定的网络操作系统，作为嵌入式操作系统有如下优势。

1．低成本开发系统

Linux 的源码开放性允许任何人可以获取并修改 Linux 的源码。这样一方面大大降低了开发的成本，另一方面又可以提高开发产品的效率。并且还可以在 Linux 社区中获得支持，用户只需向邮件列表发一封邮件，即可获得作者的支持。

2．可应用于多种硬件平台

Linux 可支持 X86、PowerPC、ARM、XSCALE、MIPS、SH、68K、Alpha、SPARC 等多种体系结构，并且已经被移植到多种硬件平台。这对于经费、时间受限制的研究与开发项目是很有吸引力的。Linux 采用一个统一的框架对硬件进行管理，同时从一个硬件平台到另一个硬件平台的改动与上层应用无关。

3．可定制的内核

Linux 具有独特的内核模块机制，它可以根据用户的需要，实时地将某些模块插入到内核中或者从内核中移走，并能根据嵌入式设备的个性需要量体裁衣。经裁减的 Linux 内核最小可达到 150KB 以下，尤其适合嵌入式领域中资源受限的实际情况。当前的 2.6 内核加入了许多嵌入式友好特性，如构建用于不需要用户界面的设备的小占板面积内核选项。

4．性能优异

Linux 系统内核精简、高效和稳定，能够充分发挥硬件的功能，因此它比其他操作系统的运行效率更高。在个人计算机上使用 Linux，可以将它作为工作站。它也非常适合在嵌入式领域中应用，对比其他操作系统，它占用的资源更少，运行更稳定，速度更快。

5．良好的网络支持

Linux 是首先实现 TCP/IP 协议栈的操作系统，它的内核结构在网络方面是非常完整的，并提供了对包括十兆位、百兆位及千兆位的以太网，还有无线网络、Token ring（令牌环）和光纤甚至卫星的支持，这对现在依赖于网络的嵌入式设备来说无疑是很好的选择。

1.1.3 Linux 发行版本

由于 Linux 属于 GNU 系统，而这个系统采用的 GPL 协议，并保证了源代码的公开。于是众多组织或公司在 Linux 内核源代码的基础上进行了一些必要的修改加工，然后再开发一些配套的软件，并把它整合成一个自己的发布版 Linux。除去非商业组织 Debian 开发的 Debian GNU/Linux 外，美国的 Red Hat 公司发行了 Red Hat Linux，法国的 Mandrake 公司发行了 Mandrake Linux，德国的 SUSE 公司发行了 SUSE Linux，国内众多公司也发行了中文版的 Linux，如著名的红旗 Linux。Linux 目前已经有超过 250 个发行版本。

下面仅对 Red Hat、Debian、Mandrake 等有代表性的 Linux 发行版本进行介绍。

1．Red Hat

国内，乃至是全世界的 Linux 用户最熟悉的发行版想必就是 Red Hat 了。Red Hat 最早是由 Bob Young 和 Marc Ewing 在 1995 年创建的。目前 Red Hat 分为两个系列：由 Red Hat 公司提供收费技术支持和更新的 Red Hat Enterprise Linux（RHEL，Red Hat 的企业版），以及由社区开发的免费的桌面版 Fedora Core。

Red Hat 企业版有三个版本——AS、ES 和 WS。AS 是其中功能最为强大和完善的版本。而正统的桌面版 Red Hat 版本更新早已停止，最后一版是 Red Hat 9.0。本书就以稳定性高的 RHEL AS 作为安装实例进行讲解。

官方主页：http://www.redhat.com/。

2．Debian

之所以把 Debian 单独列出，是因为 Debian GNU/Linux 是一个非常特殊的版本。在 1993 年，伊恩·默多克（Ian Murdock）发起 Debian 计划，它的开发模式和 Linux 及其他开放性源代码操作系统的精神一样，都是由超过 800 位志愿者通过互联网合作开发而成的。一直以来，Debian GNU/Linux 被认为是最正宗的 Linux 发行版本，而且它是一个完全免费的、高质量的且与 UNIX 兼容的操作系统。

Debian 系统分为三个版本，分别为稳定版（Stable），测试版（Testing）和不稳定版（Unstable）。并且每次发布的版本都是稳定版，而测试版在经过一段时间的测试证明没有问题后会成为新的稳定版。Debian 拥有超过 8710 种不同的软件，而且每一种软件都是自由的，而且有非常方便的升级安装指令，基本囊括了用户需要。Debian 也是最受欢迎的嵌入式 Linux

之一。

官方主页：http://www.debian.org/。

3．国内的发行版本及其他

目前国内的红旗、新华等都发行了自己的 Linux 版本。

除了前面所提到的这些版本外，业界还存在着诸如 gentoo、LFS 等适合专业人士使用的版本。在此不做介绍，有兴趣的读者可以自行查找相关的资料做进一步的了解。

1.1.4 如何学习 Linux

正如人们常说的“实践出真知”，学习 Linux 的过程也一样。只有通过大量的动手实践才能真正地领会 Linux 的精髓，才能迅速掌握在 Linux 上的应用开发，相信有编程语言经验的读者一定会认同这一点。因此，在本书中笔者安排了大量的实验环节和课后实践环节，希望读者尽可能多参与。

另外要指出的是，互联网也是一个很好的学习工具，一定要充分地加以利用。正如编程语言一样，实践的过程中总会出现多种多样的问题，笔者在写作的过程当中会尽可能地考虑可能出现的问题，但限于篇幅和读者的实际情况，不可能考虑到所有可能出现的问题，所以希望读者能充分利用互联网这一共享的天空，在其中寻找答案。以下列出了国内的一些 Linux 论坛：

http://www.linuxfans.org
http://www.linuxforum.net/
http://www.linuxeden.com/forum/
http://www.newsmth.net

1.2 Linux 安装

有了一个初步的了解后，读者是否想亲自试一下？其实安装 Linux 是一件很容易的事情，不过在开始安装之前，还需要了解一下在 Linux 安装过程中可能遇到的一些基本知识以及它与 Windows 的区别。

1.2.1 基础概念

1．文件系统、分区和挂载

文件系统是指操作系统中与管理文件有关的软件和数据。Linux 的文件系统和 Windows 中的文件系统有很大的区别，Windows 文件系统是以驱动器的盘符为基础的，而且每一个目录是与相应的分区对应，例如“E:\workplace”是指此文件在 E 盘这个分区下。而 Linux 恰好相反，文件系统是一个文件树，且它的所有文件和外部设备（如硬盘、光驱等）都是以文件的形式挂结在这个文件树上，例如“\usr\local”。对于 Windows 而言，就是指所有分区都是在一些目录下。总之，在 Windows 下，目录结构属于分区；Linux 下，分区属于目录结构。其

关系如下图 1.1 和 1.2 所示。

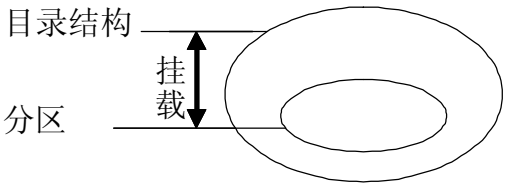


图 1.1 Linux 下目录与分区关系

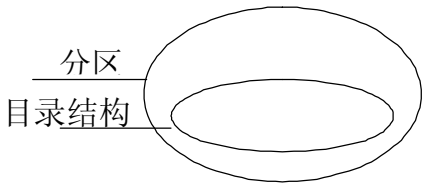


图 1.2 Windows 下目录与分区关系图

因此，在 Linux 中把每一个分区和某一个目录对应，以后在对这个目录的操作就是对这个分区的操作，这样就实现了硬件管理手段和软件目录管理手段的统一。这个把分区和目录对应的过程叫做**挂载**（Mount），而这个挂载在文件树中的位置就是**挂载点**。这种对应关系可以由用户随时中断和改变。

想一想 Linux 文件系统的挂载特性给用户能带来怎样的好处呢？

2．主分区、扩展分区和逻辑分区

硬盘分区是针对一个硬盘进行操作的，它可以分为：主分区、扩展分区、逻辑分区。其中主分区就是包含操作系统启动所必需的文件和数据的硬盘分区，要在硬盘上安装操作系统，则该硬盘必须要有一个主分区，而且其主分区的数量可以是 1～3 个；扩展分区也就是除主分区外的分区，但它不能直接使用，必须再将它划分为若干个逻辑分区才可使用，其数量可以有 0 或 1 个；而逻辑分区则在数量上没有什么限制。它们的关系如图 1.3 所示。

一般而言，对于先装了 Windows 的用户，则 Windows 的 C 盘是装在主分区上的，可以把 Linux 安装在另一个主分区或者扩展分区上。通常为了安装方便安全起见，一般采用把 Linux 装在多余的逻辑分区上。如图 1.4 所示。

图 1.3 Linux 下主分区、扩展分区、逻辑分区示意图

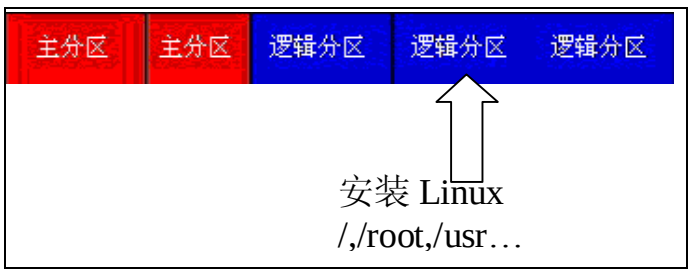


图 1.4 Linux 安装的分区示意图

小知识

通常在 Windows 下的盘符和 Linux 设备文件的对应关系如下:
C 盘——/dev/hda1（主分区）
D 盘——/dev/hda5（逻辑分区）
E 盘——/dev/hda6（逻辑分区）

3．SWAP 交换分区

在硬件条件有限的情况下，为了运行大型的程序，Linux 在硬盘上划出一个区域来当作临时的内存，而 Windows 操作系统把这个区域叫做虚拟内存，Linux 把它叫做交换分区 swap。在安装 Linux 建立交换分区时，一般将其设为内存大小的 2 倍，当然也可以设为更大。

4．分区格式

不同的操作系统选择了不同的格式，同一种操作系统也可能支持多种格式。微软公司的 Windows 就选择了 FAT32、NTFS 两种格式，但是 Windows 不支持 Linux 上常见的分区格式。Linux 是一个开放的操作系统，它最初使用 EXT2 格式，后来使用 EXT3 格式，但是它同时支持非常多的分区格式，包括很多大型机上 UNIX 使用的 XFS 格式，也包括微软公司的 FAT 以及 NTFS 格式。

5．GRUB

GRUB 是一种引导装入器（类似在嵌入式中非常重要的 bootloader）——它负责装入内核并引导 Linux 系统，位于硬盘的起始部分。由于 GRUB 多方面的优越性，如今的 Linux 一般都默认采用 GRUB 来引导 Linux 操作系统。但事实上它还可以引导 Windows 等多种操作系统。

小知识

在安装了 Windows 和 Linux 双系统后，系统是以 Linux 的 GRUB 作为引导装入器来选择启动 Windows 或 Linux 的，因此，若此时直接在 Windows 下把 Linux 的分区删除，会导致系统因没有引导装入器而无法启动 Windows，这点要格外小心。

6．root 权限

Linux 也是一个多用户的系统（在这一点上类似 Windows XP），不同的用户和用户组会有不同的权限，其中把具有超级权限的用户称为 root 用户。root 的默认主目录在“/root”下，而其他普通用户的目录则在“/home”下。root 的权限极高，它甚至可以修改 Linux 的内核，因此建议初学者要慎用 root 权限，不然一个小小参数的设置错误很有可能导致系统的严重问题。

1.2.2 硬件需求

Linux 对硬件的需求非常低。如果要是只想在字符方式下运行，那么一台 386 的计算机已经可以用来安装 Linux 了；如果想运行 X-Windows，那也只需要一台 16MB 内存，600MB 硬盘的 486 计算机即可。这听起来比那些需要 256MB 内存，2.0GBHz 的操作系统要好得多，事实上也正是如此。

现在软件和硬件行业的趋势是让用户购买更快的计算机，不断扩充内存和硬盘，而 Linux 却不受这个趋势的影响。随着 Linux 的发展，由于在其上运行的软件越来越多，因此它所需要的配置越来越高，但是用户可以有选择地安装软件，从而节省资源。既可以运行在最新的 Pentium 4 处理器上，也可以运行在 400MHz 的 Pentium II 上，甚至如果用户需要，也可以在只有文本界面的更低配置的机器上运行。由此可见 Linux 非常适合需求各异的嵌入式硬件平台。而且 Linux 可以很好地支持标准配件。如果用户的计算机是采用了标准配件，那么运行 Linux 应该没有任何问题。

1.2.3 安装准备

在开始安装之前，首先需要了解一下机器的硬件配置，包括以下几个问题。

（1）有几个硬盘，每个硬盘的大小，如果有两个以上的硬盘哪个是主盘。

（2）内存有多大。

（3）显卡的厂家和型号，有多大的显存。

（4）显示器的厂家和型号。

（5）鼠标的类型。

如果用户的计算机需要联网，那么还需要注意以下问题。

（1）计算机的 IP 地址，子网掩码，网关，DNS 的地址，主机名。

（2）或许有的时候还需要搞清楚网卡的型号和厂商。

如果不确定系统对硬件的兼容性，或者想了解 Linux 是否支持一些比较新或不常见的硬件，用户可以到 http://hardware.redhat.com 和 http://xfree86.org 进行查询。

其次，用户可以选择从网络安装（如果带宽够大，笔者推荐从商家手中购买 Linux 的安装盘，一般会获得相应的产品手册，售后服务和众多附赠的商业软件），也可以从他人那里复制，放心，这是合法的，因为 Linux 是免费的。如果用户需要获得最新的，或需要一个不易于购买到的版本，那么用户可以从 http://www.Linuxiso.org 下载一个需要的 Linux 版本。

最后，应在安装前确认磁盘上是否有足够的空间，一般的发行版本全部安装需要 3GB 左右，最小安装可以到数十兆字节，当然还需要给未来的使用留下足够的空间。如果用户拥有的是一个已经分区的空闲空间，那么可以选择在安装前在 Windows 下删除相应分区，也可以选择在安装时删除。

1.2.4 安装过程

Red Hat Enterprise 4 AS 于 2005 年 2 月发布的，是基于 2.6.9 版本的 Linux 内核。它可以选择的安装模式有光盘安装、硬盘安装和网络安装。由于 Red Hat Enterprise 4 AS 易于上手，较为稳定，因此笔者向初学者推荐此版本。

Red Hat Enterprise 4 AS 的安装盘共有 4 张，安装模式有图形安装模式和 Linux text 安装模式两种，对于初学者，推荐图形安装模式。

1．开机启动界面

将第一张 CD 光盘插入后会有图 1.5 所示选项，直接按 Enter 键就可以进入图形安装

模式。

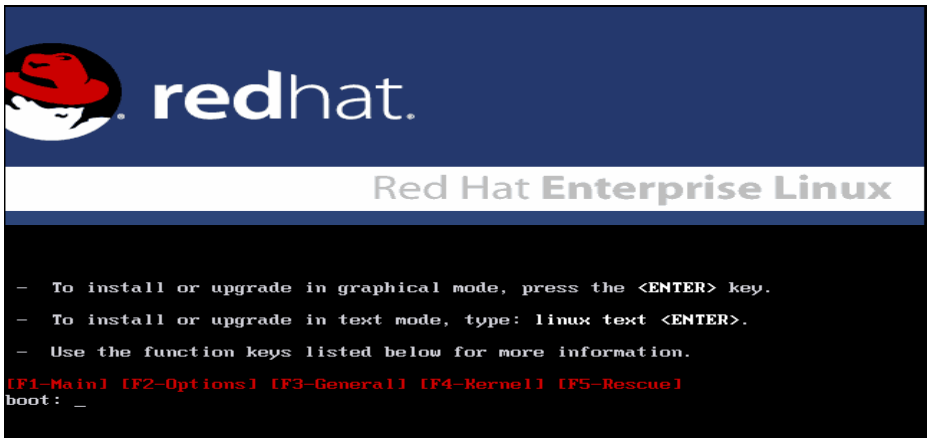


图 1.5　Linux 安装模式选择

小知识　如果想修复已经安装好的系统，请在提示符 boot: 后输入“Linux rescue”命令。

2．检测安装盘

图 1.6 出现后，如果是一张完整的安装盘，则可以直接单击“Skip”按钮跳过，否则单击“OK”按钮检测安装盘的完整性，不过要等很长时间。

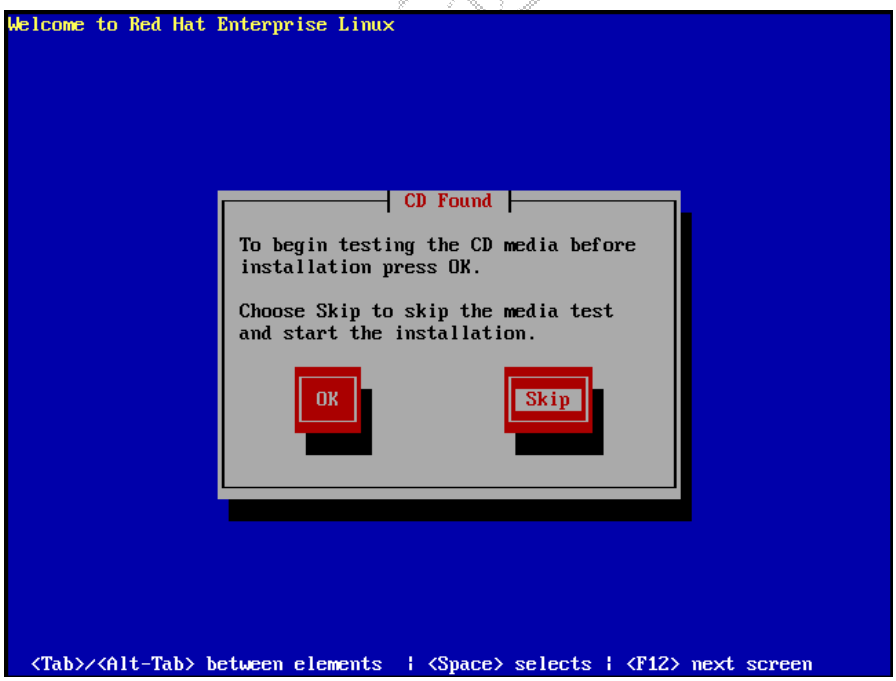


图 1.6　Linux 光盘检测

3．安装过程中的语言、键盘的选择

下面两步是安装过程中的语言和键盘的选择，要想在安装过程中可以看到中文的提示，就要选“简体中文”。接下来的键盘布局类型选择中直接选定默认的“美国英语式”即可。

4．磁盘分区

按照提示往下进行，会提示选择安装方式是“个人桌面”、“工作站”、“服务器”还是“定制”，其中若选择“个人桌面”、“工作站”或是“服务器”，则之后系统会自动选择它所要需要的软件，为了增加灵活性，可选择“定制”安装。此后将出现图 1.7 所示的磁盘分区设置对话框，这是众多初学者为之犯错和迷惑的地方。

注意 如果用户的计算机没有安装任何操作系统，那么可以考虑选择自动分区。否则要选择用 Disk Druid 手工分区。如果选用自动分区。自动分区会把用户的磁盘全部清空，然后进行自动分区。

图 1.7　Linux 安装分区选择

单击“下一步”，将会看到目前磁盘的分区情况，如图 1.8 所示。可以通过双击空闲磁盘空间，或者单击“新建”为 Linux 来创建一个新的磁盘分区。

图 1.8　Linux 分区情况

注意　笔者为截图方便，采用的是虚拟机安装，故在设备驱动器中显示的是“/dev/sda”，若正常硬盘安装应显示“/dev/hda1”等，读者需按实际情况进行选择。

单击“新建”按钮，出现图 1.9 所示对话框，挂载点和文件系统的问题已在 1.2.1 节中已叙述过，首先在挂载点的下拉列表中选择“/”，然后在文件系统的下拉列表中选择“ext3”，大小视用户计算机硬盘情况而定。设置完毕后单击“确定”按钮。

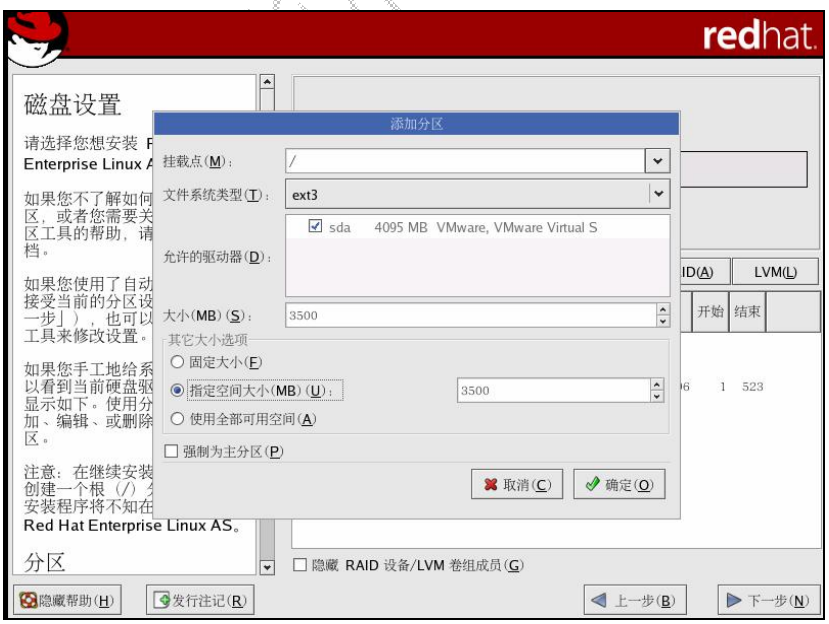


图 1.9　选择挂载根文件系统

动手

如果还想添加其他文件系统该如何操作呢？看看最后效果又是如何？

试试

单击“下一步”，进入图 1.10 所示对话框，交换分区选择中，文件系统选择为“swap”，如果用户的内存小于等于 256MB，那么笔者推荐把交换分区设为内存的两倍甚至更大（注意，最好大小为 2 的 N 次幂的数字），如果用户的内存大于等于 512MB，那么设为 512MB 即可。

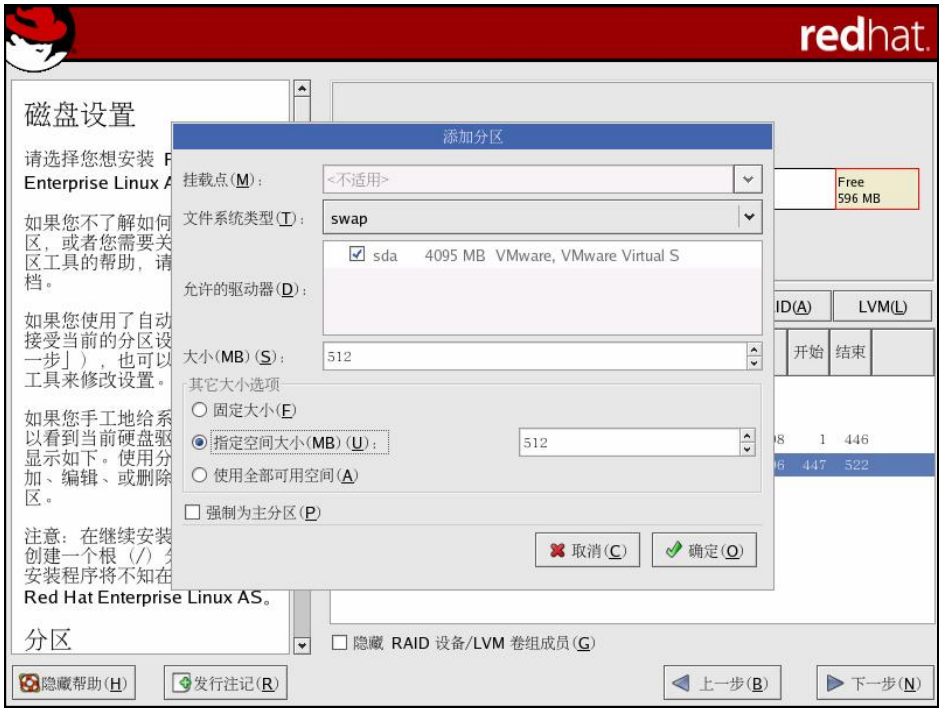


图 1.10 挂载交换分区

5．引导程序选择

磁盘分区后面的步骤比较简单。

由于现在往往存在着多系统，因此 Linux 为用户提供了 GRUB 系统引导管理器，如图 1.11 所示，直接依次单击“确定”和“下一步”按钮即可。

图 1.11　选择 GRUB 方式引导

接下来的网络设置要按网络情况而定。为了便于安装以后的驱动，要把防火墙关掉，SELinux 也关掉。

之后的时区选择请选择上海（没有北京时间可选，不过两者等效），进行 root 用户（Linux 系统中的权限最高者）的密码设置。

如果是初学者，为了能够更好的学习 Linux，也为了避免以后的麻烦，请在软件包组安装选择中选择“全部”。接下来的语言支持选定中文。

大约 60～80 分钟，Linux 的安装就会初步完成，重启进入 Linux 后还需要进行一些设置，包括时间、刷新率等。

1.3　Linux 文件及文件系统

在安装完 Linux 之后，下面先对 Linux 中一些非常重要的概念做一些介绍，以便进一步学习使用 Linux。

1.3.1　文件类型及文件属性

1．文件类型

Linux 中的文件类型与 Windows 有显著的区别，其中最显著的区别在于 Linux 对目录和设备都当作文件来进行处理，这样就简化了对各种不同类型设备的处理，提高了效率。Linux

中主要的文件类型分为 4 种：普通文件、目录文件、链接文件和设备文件。

（1）普通文件

普通文件如同 Windows 中的文件一样，是用户日常使用最多的文件。它包括文本文件、shell 脚本（shell 的概念在第 2 章会进行讲解）、二进制的可执行程序和各种类型的数据。

（2）目录文件

在 Linux 中，目录也是文件，它们包含文件名和子目录名以及指向那些文件和子目录的指针。目录文件是 Linux 中存储文件名的惟一地方，当把文件和目录相对应起来时，也就是用指针将其链接起来之后，就构成了目录文件。因此，在对目录文件进行操作时，一般不涉及对文件内容的操作，而只是对目录名和文件名的对应关系进行了操作。

另外，在 Linux 系统中的每个文件都被赋予一个惟一的数值，而这个数值被称做索引节点。索引节点存储在一个称作索引节点表（Inode Table）中，该表在磁盘格式化时被分配。每个实际的磁盘或分区都有其自己的索引节点表。一个索引节点包含文件的所有信息，包括磁盘上数据的地址和文件类型。

Linux 文件系统把索引节点号 1 赋于根目录，这也就是 Linux 的根目录文件在磁盘上的地址。根目录文件包括文件名、目录名及它们各自的索引节点号的列表，Linux 可以通过查找从根目录开始的一个目录链来找到系统中的任何文件。

Linux 通过上下链接目录文件系统来实现对整个文件系统的操作。比如把文件从一个磁盘目录移到另一实际磁盘的目录时（实际上是通过读取索引节点表来检测这种行动的），这时，原先文件的磁盘索引号删除，而且在新磁盘上建立相应的索引节点。它们之间的相应关系如图 1.12 所示。

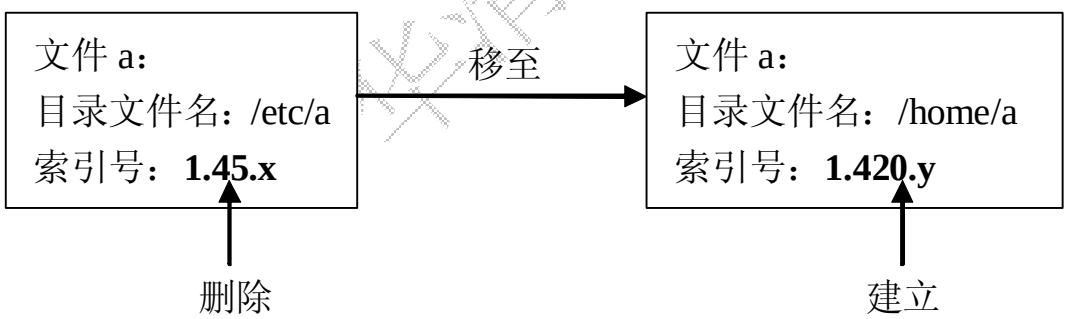


图 1.12　目录文件与索引节点关系

（3）链接文件

链接文件有些类似于 Windows 中的“快捷方式”，但是它的功能更为强大。它可以实现对不同的目录、文件系统甚至是不同的机器上的文件直接访问，并且不必重新占用磁盘空间。

（4）设备文件

Linux 把设备都当作文件一样来进行操作，这样就大大方便了用户的使用（在后面的 Linux 编程中可以更为明显地看出）。在 Linux 下与设备相关的文件一般都在/dev 目录下，它包括两种，一种是块设备文件，另一种是字符设备文件。

Ø 块设备文件是指数据的读写，它们是以块（如由柱面和扇区编址的块）为单位的设备，最简单的如硬盘（/dev/hda1）等。

Ø 字符设备主要是指串行端口的接口设备。

2．文件属性

Linux 中的文件属性如图 1.13 如示。

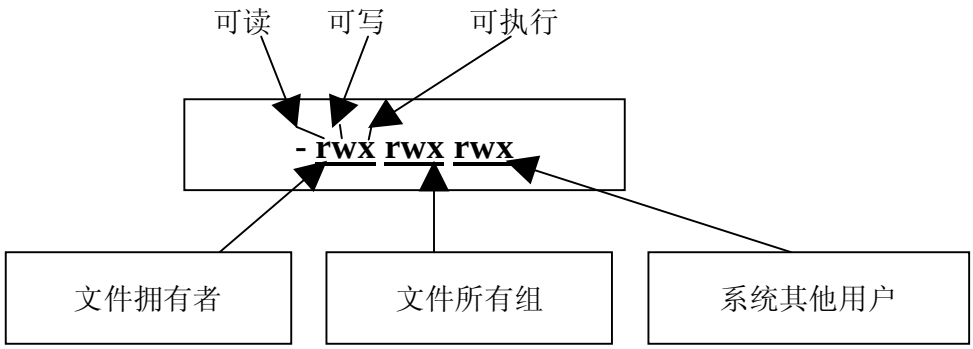


图 1.13　Linux 文件属性表示方法

首先，Linux 中文件的拥有者可以把文件的访问属性设成 3 种不同的访问权限：可读（r）、可写（w）和可执行（x）。文件又有 3 个不同的用户级别：文件拥有者（u）、所属的用户组（g）和系统里的其他用户（o）。

第一个字符显示文件的类型：

Ø “-”表示普通文件;
Ø “d”表示目录文件;
Ø “l”表示链接文件;
Ø “c”表示字符设备;
Ø “b”表示块设备;
Ø “p”表示命名管道比如 FIFO 文件（First In First Out，先进先出）;
Ø “f”表示堆栈文件比如 LIFO 文件（Last In First Out，后进先出）。

第一个字符之后有 3 个三位字符组：

Ø 第一个三位字符组表示对于文件拥有者（u）对该文件的权限;
Ø 第二个三位字符组表示文件用户组（g）对该文件的权限;
Ø 第三个三位字符组表示系统其他用户（o）对该文件的权限;
Ø 若该用户组对此没有权限，一般显示“-”字符。

小知识　目录权限和文件权限有一定的区别。对于目录而言，r 代表允许列出该目录下的文件和子目录，w 代表允许生成和删除该目录下的文件，x 代表允许访问该目录。

1.3.2　文件系统类型介绍

1．ext2 和 ext3

ext3 是现在 Linux（包括 Red Hat，Mandrake 下）常见的默认的文件系统，它是 ext2 的

升级版本。正如 Red Hat 公司的首席核心的开发人员 Michael K.Johnson 所说，从 ext2 转换到 ext3 主要有以下 4 个理由：可用性、数据完整性、速度以及易于转化。ext3 中采用了日志式的管理机制，它使文件系统具有很强的快速恢复能力，并且由于从 ext2 转换到 ext3 无须进行格式化，因此，更加推进了 ext3 文件系统的大大推广。

2．swap 文件系统

该文件系统是 Linux 中作为交换分区使用的。在安装 Linux 的时候，交换分区是必须建立的，并且它所采用的文件系统类型必须是 swap 而没有其他选择。

3．vfat 文件系统

Linux 中把 DOS 中采用的 FAT 文件系统（包括 FAT12，FAT16 和 FAT32）都称为 vfat 文件系统。

4．NFS 文件系统

NFS 文件系统是指网络文件系统，这种文件系统也是 Linux 的独到之处。它可以很方便地在局域网内实现文件共享，并且使多台主机共享同一主机上的文件系统。而且 NFS 文件系统访问速度快、稳定性高，已经得到了广泛的应用，尤其在嵌入式领域，使用 NFS 文件系统可以很方便地实现文件本地修改，而免去了一次次读写 flash 的忧虑。

5．ISO9660 文件系统

这是光盘所使用的文件系统，在 Linux 中对光盘已有了很好的支持，它不仅可以提供对光盘的读写，还可以实现对光盘的刻录。

1.3.3 Linux 目录结构

Linux 的目录结构如图 1.14 所示。

下面以 Red Hat Enterprise 4 AS 为例，详细列出了 Linux 文件系统中各主要目录的存放内容，如表 1.1 所示。

表 1.1 Linux 文件系统目录结构

目　　录	目 录 内 容
/bin	bin 就是二进制（binary）英文缩写。在这里存放前面 Linux 常用操作命令的执行文件，如 mv、ls、mkdir 等。有时，这个目录的内容和/usr/bin 里面的内容一样，它们都是放置一般用户使用的执行文件
/boot	这个目录下存放操作系统启动时所要用到的程序。如启动 grub 就会用到其下的/boot/grub 子目录
/dev	该目录中包含了所有 Linux 系统中使用的外部设备。要注意的是，这里并不是存放的外部设备的驱动程序，它实际上是一个访问这些外部设备的端口。由于在 Linux 中，所有的设

	备都当作文件一样进行操作，比如：/dev/cdrom 代表光驱，用户可以非常方便地像访问文件、目录一样对其进行访问
/etc	该目录下存放了系统管理时要用到的各种配置文件和子目录。如网络配置文件、文件系统、x 系统配置文件、设备配置信息设置用户信息等都在这个目录下。系统在启动过程中需要读取其参数进行相应的配置
/etc/rc.d	该目录主要存放 Linux 启动和关闭时要用到的脚本文件，在后面的启动详解中还会进一步地讲解

续表

目　　录	目 录 内 容
/etc/rc.d/init	该目录存放所有 Linux 服务默认的启动脚本（在新版本的 Linux 中还用到的是/etc/xinetd.d 目录下的内容）
/home	该目录是 Linux 系统中默认的用户工作根目录。如前面在 1.3.1 节中所述，执行 adduser 命令后系统会在/home 目录下为对应账号建立一个名为同名的主目录
/lib	该目录是用来存放系统动态链接共享库的。几乎所有的应用程序都会用到这个目录下的共享库。因此，千万不要轻易对这个目录进行什么操作
/lost+found	该目录在大多数情况下都是空的。只有当系统产生异常时，会将一些遗失的片段放在此目录下
/media	该目录下是光驱和软驱的挂载点，Fedora Core 4 已经可以自动挂载光驱和软驱
/misc	该目录下存放从 DOS 下进行安装的实用工具，一般为空
/mnt	该目录是软驱、光驱、硬盘的挂载点，也可以临时将别的文件系统挂载到此目录下
/proc	该目录是用于放置系统核心与执行程序所需的一些信息。而这些信息是在内存中由系统产生的，故不占用硬盘空间
/root	该目录是超级用户登录时的主目录
/sbin	该目录是用来存放系统管理员的常用的系统管理程序
/tmp	该目录用来存放不同程序执行时产生的临时文件。一般 Linux 安装软件的默认安装路径就是这里
/usr	这是一个非常重要的目录，用户的很多应用程序和文件都存放在这个目录下，类似与 Windows 下的 Program Files 的目录
/usr/bin	系统用户使用的应用程序
/usr/sbin	超级用户使用的比较高级的管理程序和系统守护程序
/usr/src	内核源代码默认的放置目录
/srv	该目录存放一些服务启动之后需要提取的数据

/sys	这是 Linux 2.6 内核的一个很大的变化。该目录下安装了 2.6 内核中新出现的一个文件系统 sysfs sysfs 文件系统集成了下面 3 种文件系统的信息：针对进程信息的 proc 文件系统、针对设备的 devfs 文件系统以及针对伪终端的 devpts 文件系统。该文件系统是内核设备树的一个直观反映。当一个内核对象被创建的时候，对应的文件和目录也在内核对象子系统中被创建
/var	这也是一个非常重要的目录，很多服务的日志信息都存放在这里

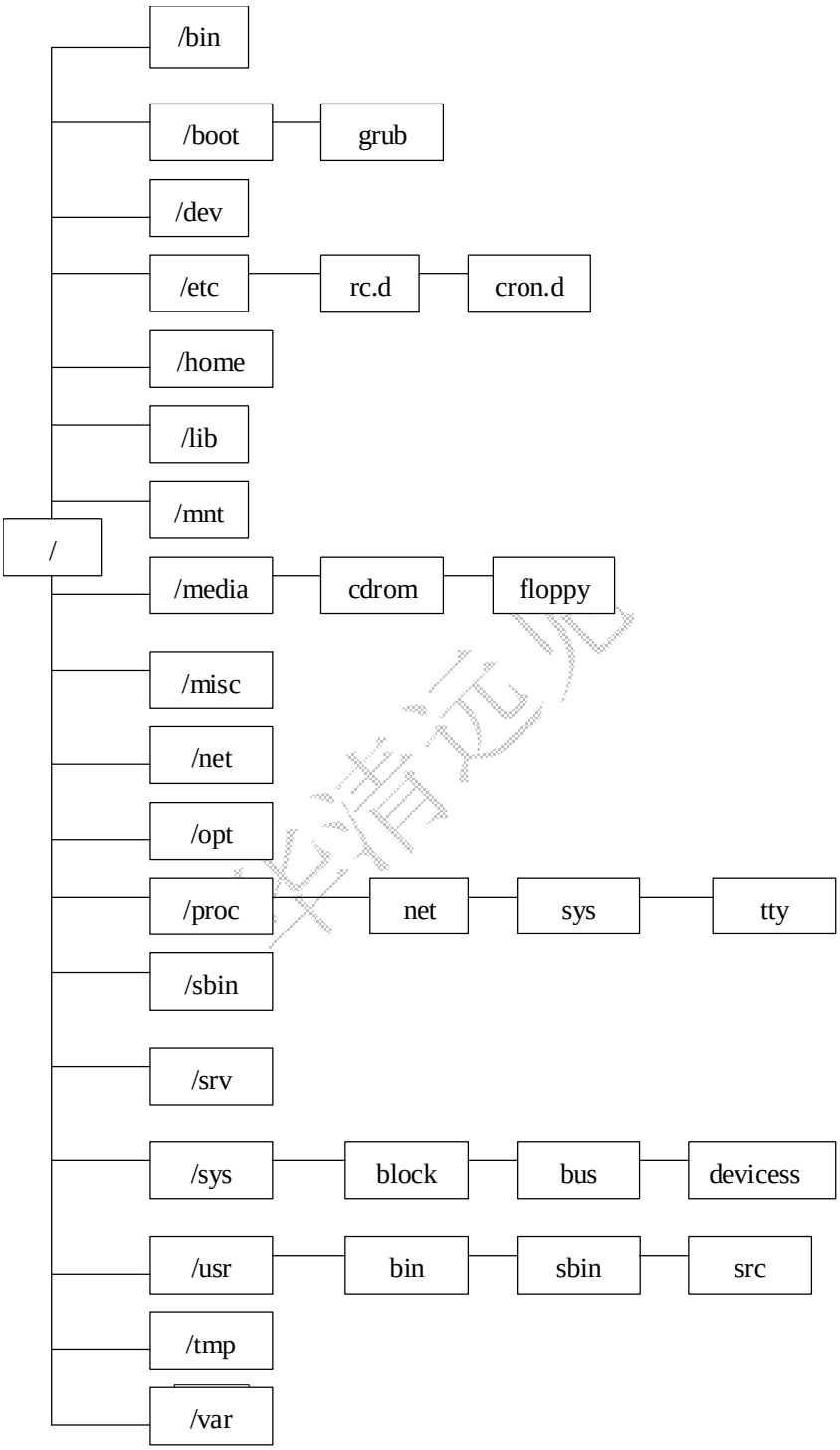


图 1.14　Linux 目录结构

1.4 实验内容——安装 Linux 操作系统

1．实验目的

通过读者亲自动手实践安装 Linux 操作系统，已经对 Linux 有个初步的认识，并且加深了对 Linux 中的基本概念的理解，并熟悉 Linux 文件系统目录结构。

2．实验内容

安装 Linux（Red Hat Enterprise 4 AS 版本）操作系统，查看 Linux 的目录结构。

3．实验步骤

（1）磁盘规划。
在这步中，需要规划出最好有 5GB 的空间来安装 Linux。
（2）下载 Linux 版本。
可以从 Linux 的映像网站上下载各版本的 Linux。
（3）搜集主机硬件信息。
查看相应版本的 Linux 是否已有了对相应各硬件的驱动支持。较新版本的 Linux 一般对硬件的支持都比教好。
（4）确认用户网络信息。
包括 IP、子网掩码、DNS 地址等。
（5）按照本书 1.2 小节讲述的步骤安装 Linux，对关键的步骤：如配置文件系统及硬盘分区要倍加小心。
（6）选择安装套件，建议新手可以使用全部安装来减少以后学习的难度。
（7）配置用户信息、网络信息等。
（8）安装完成，用普通用户登录到 Linux 下。
（9）使用文件浏览器熟悉文件的目录结构。

4．实验结果

能够成功安装上 Linux 操作系统，并且对 Linux 文件系统的目录结构能有一个总体的掌握。

本 章 小 结

本章首先介绍了 Linux 的历史、嵌入式 Linux 操作系统的优势、Linux 不同发行版本的区别以及如何学习 Linux。在这里要着重掌握的是 Linux 内核与 GNU 的关系，了解 Linux 版本号的规律，同时还要了解 Linux 多硬件平台支持、低开发成本等优越性。对于 Linux 的不同

发行版本，读者可以到各自主页上了解相关信息。

本章接着介绍了如何安装 Linux，这里最关键的一步是分区。希望读者能很好地掌握主分区、扩展分区的概念。Linux 文件系统与 Windows 文件系统的区别以及 Linux 中“挂载”与“挂载点”的含义，这几个都是 Linux 中的重要概念，希望读者能够切实理解其含义。

在安装完 Linux 之后，本章讲解了 Linux 中文件和文件系统的概念。这些是 Linux 中最基础最常见的概念，只有真正理解之后才能为进一步学习 Linux 打下很好的基础。读者要着重掌握 Linux 的文件分类、文件属性的表示方法，并且能够通过实际查看 Linux 目录结构来熟悉 Linux 中重要目录的作用。

最后本章还设计了本书中的第一个实验——安装 Linux，这也是读者必须要完成的最基础的实验，读者在安装前要做好充分的准备工作，包括备份重要资料、复习本章相关内容，相信经过认真学习的读者都能安装成功。

思考与练习

1．请查找资料，看看 GNU 所规定的自由软件的具体协议是什么？
2．请问 Linux 下的文件系统和 Windows 下的文件系统有什么区别？
3．试指出读者系统中的磁盘划分情况（如主分区、扩展分区的对应情况）。
4．如何安装 Linux？
5．Linux 中的文件有哪些类？这样分类有什么好处？
6．若有一个文件，其属性为“-rwxr—rw-”，说出这代表的什么？
7．请说出下列目录放置的是什么数据：
/etc/：
/etc/rc.d/init.d/：
/usr/bin：
/bin：
/usr/sbin：
/sbin：
/var/log：

“黑色经典”系列之《嵌入式 Linux 应用程序开发详解》

第 2 章 Linux 基础命令

本章目标

Linux 是个高可靠、高性能的系统，而所有这些优越性只有在直接使用 Linux 命令行（Shell 环境）才能充分地体现出来。在本章将帮助读者学会如下内容。

- 掌握 Shell 基本概念
- 熟练使用 Linux 中用户管理命令
- 熟练使用 Linux 中系统相关命令
- 熟练使用 Linux 中文件目录相关命令
- 熟练使用 Linux 中打包压缩相关命令
- 熟练使用 Linux 中文件比较合并相关命令
- 熟练使用 Linux 中网络相关命令
- 了解 Linux 的启动过程
- 深入了解 INIT 进程及其配置文件
- 能够独立完成在 Linux 中解压软件
- 学会添加环境变量
- 能够独立定制 Linux 中系统服务

2.1 Linux 常用操作命令

在安装完 Linux 再次启动之后，就可以进入到与 Windows 类似的图形化界面了。这个界面就是 Linux 图形化界面 X 窗口系统（简称 X）的一部分。要注意的是，X 窗口系统仅仅是 Linux 上面的一个软件（或者也可称为服务），它不是 Linux 自身的一部分。虽然现在的 X 窗口系统已经与 Linux 整合地相当好了，但毕竟还不能保证绝对的可靠性。另外，X 窗口系统是一个相当耗费系统资源的软件，它会大大地降低 Linux 的系统性能。因此，若是希望更好地享受 Linux 所带来的高效及高稳定性，建议读者尽可能地使用 Linux 的命令行界面，也就是 Shell 环境。

当用户在命令行下工作时，不是直接同操作系统内核交互信息的，而是由命令解释器接受命令，分析后再传给相关的程序。Shell 是一种 Linux 中的命令行解释程序，就如同 Command.com 是 DOS 下的命令解释程序一样，为用户提供使用操作系统的接口。它们之间的关系如图 2.1 所示。用户在提示符下输入的命令都由 Shell 先解释然后传给 Linux 内核。

小知识

- Shell 是命令语言、命令解释程序及程序设计语言的统称。它不仅拥有自己内建的 Shell 命令集，同时也能被系统中其他应用程序所调用。
- Shell 的另一个重要特性是它自身就是一个解释型的程序设计语言，Shell 程序设计语言支持绝大多数在高级语言中能见到的程序元素，如函数、变量、数组和程序控制结构。Shell 编程语言简单易学，任何在提示符中能键入的命令都能放到一个可执行的 Shell 程序中。关于 Shell 编程的详细讲解，感兴趣的读者可以参见其他相关书籍。

Linux 中运行 Shell 的环境是“系统工具”下的“终端”，读者可以单击“终端”以启动 Shell 环境。这时屏幕上显示类似“**[sunq@www home]$**”的信息，其中，sunq 是指系统用户，而 home 是指当前所在的目录。

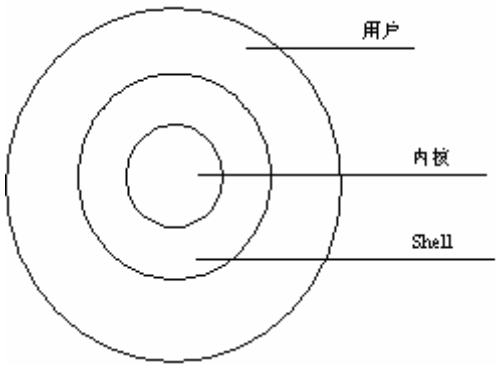


图 2.1 内核、Shell 和用户的关系

由于 Linux 中的命令非常多，要全部介绍几乎不可能。因此，在本书中按照命令的用途进行分类讲解，并且对每一类中最常用的命令详细讲解，同时列出同一类中的其他命令。由于同一类的命令都有很大的相似性，因此，读者通过学习本书中所列命令，可以很快地掌握其他命令。

命令格式说明。

- 格式中带[]的表明为可选项，其他为必选项。
- 选项可以多个连带写入。
- 本章后面选项参数列表中加粗的含义是：该选项是非常常用的选项。

2.1.1 用户系统相关命令

Linux 是一个多用户的操作系统，每个用户又可以属于不同的用户组，下面，首先来熟悉一下 Linux 中的用户切换和用户管理的相关命令。

1．用户切换（su）

（1）作用

变更为其它使用者的身份，主要用于将普通用户身份转变为超级用户，而且需输入相应用户密码。

（2）格式

su [选项] [使用者]

其中的使用者为要变更的对应使用者。

（3）常见参数

主要选项参数见表 2.1 所示。

表 2.1　su 命令常见参数列表

选　　项	参 数 含 义
-，-l，--login	为该使用者重新登录，大部分环境变量（如 HOME、SHELL 和 USER 等）和工作目录都是以该使用者（USER）为主。若没有指定 USER，缺省情况是 root
-m，-p	执行 su 时不改变环境变量
-c，--command	变更账号为 USER 的使用者，并执行指令（command）后再变回原来使用者

（4）使用示例

```
[sunq@www sunq]$ su - root
Password:
[root@www root]#
```

示例通过 su 命令将普通用户变更为 root 用户，并使用选项“-”携带 root 环境变量。

（5）使用说明

- 在将普通用户变更为 root 用户时建议使用”-”选项，这样可以将 root 的环境变量和工作目录同时带入，否则在以后的使用中可能会由于环境变量的原因而出错。
- 在转变为 root 权限后，提示符变为#。

小知识

环境变量实际上就是用户运行环境的参数集合。Linux 是一个多用户的操作系统。而且在每个用户登录系统后，都会有一个专有的运行环境。通常每个用户默认的环境都是相同的，而这个默认环境实际上就是一组环境变量的定义。用户可以对自己的运行环境进行定制，其方法就是修改相应的系统环境变量。

常见的环境变量如下。

☆PATH 是系统路径。

☆HOME 是系统根目录。

☆HISTSIZE 是指保存历史命令记录的条数。

☆LOGNAME 是指当前用户的登录名。

☆HOSTNAME 是指主机的名称，若应用程序要用到主机名的话，通常是从这个环境变量中来取得的。

☆SHELL 是指当前用户用的是哪种 Shell。

☆LANG/LANGUGE 是和语言相关的环境变量，使用多种语言的用户可以修改此环境变量。

☆MAIL 是指当前用户的邮件存放目录。

设置环境变量方法如下。

- ü 通过 etho 显示字符串（指定环境变量）。
- ü 通过 export 设置新的环境变量。
- ü 通过 env 显示所有环境变量。
- ü 通过 set 命令显示所有本地定义的 Shell 变量。
- ü 通过 unset 命令来清除环境变量。

读者可以试着用“env”命令查看“su - root”和“su root”的区别。

2．用户管理（useradd 和 passwd）

Linux 中常见用户管理命令如表 2.2 所示，本书仅以 useradd 和 passwd 为例进行详细讲解，其他命令类似，请读者自行学习使用。

表 2.2 Linux 常见用户管理命令

命　令	命 令 含 义	格　式
useradd	添加用户账号	useradd [选项] 用户名
usermod	设置用户账号属性	usermod [选项] 属性值
userdel	删除对应用户账号	userdel [选项] 用户名
groupadd	添加组账号	groupadd [选项] 组账号
groupmod	设置组账号属性	groupmod [选项] 属性值
groupdel	删除对应组账号	groupdel [选项] 组账号
passwd	设置账号密码	passwd [对应账号]
id	显示用户 ID、组 ID 和用户所属的组列表	id [用户名]
groups	显示用户所属的组	groups [组账号]
who	显示登录到系统的所有用户	who

（1）作用

① useradd：添加用户账号。

② passwd：更改对应用户账号密码。

（2）格式

① useradd：useradd [选项] 用户名。

② passwd：passwd [选项] [用户名]。

其中的用户名为修改账号密码的用户，若不带用户名，缺省为更改当前使用者账号密码。

（3）常用参数

① useradd 主要选项参数见表 2.3 所示。

表 2.3 useradd 命令常见参数列表

选　　项	参 数 含 义
-g	指定用户所属的群组
-m	自动建立用户的登入目录
-n	取消建立以用户名称为名的群组

② passwd：一般很少使用选项参数。

（4）使用实例

```
[root@www root]# useradd yul
[root@www root]# passwd yul
New password:
Retype new password:
passwd: all authentication tokens updated successfully
[root@www root]# su - yul
[yul@www yul]$
[yul@www yul]$ pwd（查看当前目录）
/home/yul
```

实例中先添加了用户名为 yul 的用户，接着又为该用户设置了账号密码。并从 su 的命令可以看出，该用户添加成功，其工作目录为”/home/yul”。

（5）使用说明

- 在使用添加用户时，这两个命令是一起使用的，其中，useradd 必须用 root 的权限。而且 useradd 指令所建立的账号，实际上是保存在“/etc/passwd”文本文件中，文件中每一行包含一个账号信息。
- 在缺省情况下，useradd 所做的初始化操作包括在“/home”目录下为对应账号建立一个名为同名的主目录，并且还为该用户单独建立一个与用户名同名的组。
- adduser 只是 useradd 的符号链接（关于符号链接的概念在本节后面会有介绍），两者是相同的。
- passwd 还可用于普通用户修改账号密码，Linux 并不采用类似 windows 的密码回显（显示为*号），所以输入的这些字符用户是看不见的。密码最好包括字母、数字和特殊符号，并且设成 6 位以上。

3．系统管理命令（ps 和 kill）

Linux 中常见的系统管理命令如表 2.4 所示，本书以 ps 和 kill 为例进行讲解。

表 2.4　　Linux 常见系统管理命令

命　　令	命 令 含 义	格　　式
ps	显示当前系统中由该用户运行的进程列表	ps [选项]
top	动态显示系统中运行的程序（一般为每隔 5s）	top
kill	输出特定的信号给指定 PID（进程号）的进程	kill [选项] 进程号（PID）

uname	显示系统的信息（可加选项-a）	uname [选项]
setup	系统图形化界面配置	setup
crontab	循环执行例行性命令	crontab [选项]
shutdown	关闭或重启 Linux 系统	shutdown [选项] [时间]
uptime	显示系统已经运行了多长时间	uptime
clear	清除屏幕上的信息	clear

（1）作用

① ps：显示当前系统中由该用户运行的进程列表。

② kill：输出特定的信号给指定 PID（进程号）的进程，并根据该信号而完成指定的行为。其中可能的信号有进程挂起、进程等待、进程终止等。

（2）格式

① ps：ps [选项]。

② kill：kill [选项] 进程号（PID）。

kill 命令中的进程号为信号输出的指定进程的进程号，当选项是缺省时为输出终止信号给该进程。

（3）常见参数

① ps 主要选项参数见表 2.5 所示。

表 2.5 ps 命令常见参数列表

选　项	参 数 含 义
-ef	查看所有进程及其 PID（进程号）、系统时间、命令详细目录、执行者等
-aux	除可显示-ef 所有内容外，还可显示 CPU 及内存占用率、进程状态
-w	显示加宽并且可以显示较多的信息

② kill 主要选项参数见表 2.6 所示。

表 2.6 kill 命令常见参数列表

选　项	参 数 含 义
-s	根据指定信号发送给进程
-p	打印出进程号（PID），但并不送出信号
-l	列出所有可用的信号名称

（4）使用实例

```
[root@www root]# ps –ef
UID        PID  PPID  C STIME TTY          TIME CMD
root         1     0  0  2005 ?        00:00:05 init
root         2     1  0  2005 ?        00:00:00 [keventd]
root         3     0  0  2005 ?        00:00:00 [ksoftirqd_CPU0]
```

```
root       4    0  0  2005 ?       00:00:00 [ksoftirqd_CPU1]
root    7421    1  0  2005 ?       00:00:00 /usr/local/bin/ntpd -c /etc/ntp.
root    21787 21739  0 17:16 pts/1     00:00:00 grep ntp
[root@www root]# kill 7421
[root@www root]# ps -ef|grep ntp
root    21789 21739  0 17:16 pts/1    00:00:00 grep ntp
```

该实例中首先查看所有进程，并终止进程号为 7421 的 ntp 进程，之后再次查看时已经没有该进程号的进程。

（5）使用说明

- ps 在使用中通常可以与其他一些命令结合起来使用，主要作用是提高效率。
- ps 选项中的参数 w 可以写多次，通常最多写 3 次，它的含义表示加宽 3 次，这足以显示很长的命令行了。例如：ps –auxwww。

小知识 管道是 Linux 中信息通信的重要方式。它是把一个程序的输出直接连接到另一个程序的输入，而不经过任何中间文件。管道线是指连接二个或更多程序管道的通路。在 shell 中字符”|”表示管道线。如前例子中的 ps –ef|grep ntp 所示，ps –ef 的结果直接输入到 grep ntp 的程序中（关于 grep 命令在后面会有详细的介绍）。grep、pr、sort 和 wc 都可以在上述管道线上工作。读者可以灵活地运用管道机制提高工作效率。

4．磁盘相关命令（fdisk）

Linux 中与磁盘相关的命令如表 2.7 所示，本书仅以 fdisk 为例进行讲解。

表 2.7　　Linux 常见系统管理命令

选　项	参 数 含 义	格　　式
free	查看当前系统内存的使用情况	free [选项]
df	查看文件系统的磁盘空间占用情况	df [选项]
du	统计目录（或文件）所占磁盘空间的大小	du [选项]
fdisk	查看硬盘分区情况及对硬盘进行分区管理	fdisk [-l]

（1）作用

fdisk 可以查看硬盘分区情况，并可对硬盘进行分区管理，这里主要向读者介绍查看硬盘分区情况，另外，fdisk 也是一个非常好的硬盘分区工具，感兴趣的读者可以另外查找资料学习使用 fdisk 进行硬盘分区。

（2）格式

fdisk [-l]

（3）使用实例

```
[root@sunq ~]# fdisk -l
Disk /dev/hda: 40.0 GB, 40007761920 bytes
240 heads, 63 sectors/track, 5168 cylinders
Units = cylinders of 15120 * 512 = 7741440 bytes
```

```
   Device Boot      Start         End      Blocks   Id  System
/dev/hda1   *           1        1084     8195008+   c  W95 FAT32 (LBA)
/dev/hda2            1085        5167    30867480    f  W95 Ext'd (LBA)
/dev/hda5            1085        2439    10243768+   b  W95 FAT32
/dev/hda6            2440        4064    12284968+   b  W95 FAT32
/dev/hda7            4065        5096     7799526   83  Linux
/dev/hda8            5096        5165      522081   82  Linux swap
```

可以看出，使用”fdisk –l”列出了文件系统的分区情况。

（4）使用说明

- 使用 fdisk 必须拥有 root 权限。
- IDE 硬盘对应的设备名称分别为 hda、hdb、hdc 和 hdd，SCSI 硬盘对应的设备名称则为 sda、sdb、…此外，hda1 代表 hda 的第一个硬盘分区，hda2 代表 hda 的第二个分区，依此类推。
- 通过查看/var/log/messages 文件，可以找到 Linux 系统已辨认出来的设备代号。

5．磁盘挂载命令（mount）

（1）作用

挂载文件系统，它的使用权限是超级用户或/etc/fstab 中允许的使用者。正如 1.2.1 节中所述，挂载是指把分区和目录对应的过程，而挂载点是指挂载在文件树中的位置。mount 命令就可以把文件系统挂载到相应的目录下，并且由于 Linux 中把设备都当作文件一样使用，因此，mount 命令也可以挂载不同的设备。

通常，在 Linux 下“/mnt”目录是专门用于挂载不同的文件系统的，它可以在该目录下新建不同的子目录来挂载不同的设备文件系统。

（2）格式

mount [选项] [类型] 设备文件名 挂载点目录

其中的类型是指设备文件的类型。

（3）常见参数

mount 常见参数如表 2.8 所示。

表 2.8 mount 命令选项常见参数列表

选　项	参 数 含 义
-a	依照/etc/fstab 的内容装载所有相关的硬盘
-l	列出当前已挂载的设备、文件系统名称和挂载点
-t 类型	将后面的设备以指定类型的文件格式装载到挂载点上。常见的类型有前面介绍过的几种：vfat、ext3、ext2、iso9660、nfs 等
-f	通常用于除错。它会使 mount 不执行实际挂上的动作，而是模拟整个挂上的过程，通常会和-v 一起使用

（4）使用实例

使用 mount 命令主要通过以下几个步骤。

① 确认是否为 Linux 可以识别的文件系统，Linux 可识别的文件系统只要是以下几种。

- Windows95/98 常用的 FAT32 文件系统：vfat。
- WinNT/2000 的文件系统：ntfs。
- OS/2 用的文件系统：hpfs。
- Linux 用的文件系统：ext2、ext3、nfs。
- CD-ROM 光盘用的文件系统：iso9660。

② 确定设备的名称，确定设备名称可通过使用命令“fdisk -l”查看。

③ 查找挂接点。

必须确定挂载点已经存在，也就是在“/mnt”下的相应子目录已经存在，一般建议在“/mnt”下新建几个如“/mnt/windows”，“/mnt/usb”的子目录，现在有些新版本的 Linux（如红旗 Linux、中软 Linux、MandrakeLinux）都可自动挂载文件系统，Red Hat 仅可自动挂载光驱。

④ 挂载文件系统如下所示。

```
[root@sunq mnt]# mount -t vfat /dev/hda1 /mnt/c
[root@sunq mnt]# cd /mnt/c
24.s03e01.pdtv.xvid-sfm.rmvb  Documents and Settings  Program Files
24.s03e02.pdtv.xvid-sfm.rmvb  Downloads                Recycled
...
```

C 盘是原先笔者 Windows 系统的启动盘。可见，在挂载了 C 盘之后，可直接访问 Windows 下的 C 盘的内容。

⑤ 在使用完该设备文件后可使用命令 umount 将其卸载。

```
[root@sunq mnt]# umount /mnt/c
[root@sunq mnt]# cd /mnt/c
[root@sunq c]# ls
```

可见，此时目录“/mnt/c”下为空。Windows 下的 C 盘成功卸载。

小知识

- 在 Linux 下如何使用 U 盘呢？

一般 U 盘为 SCSI 格式的硬盘，其格式为 vfat 格式，其设备号可通过“fdisk –l”进行查看，假若设备名为“/dev/sda1”，则可用如下命令就可将其挂载：

mount -t vfat /dev/sda1 /mnt/u

- 若想设置在开机时自动挂载，可在文件“/etc/fstab”中加入相应的设置行即可。

2.1.2 文件目录相关命令

由于 Linux 中有关文件目录的操作非常重要，也非常常用，因此在本节中，作者将基本所有的文件操作命令都进行了讲解。

1．cd

（1）作用

改变工作目录。

（2）格式

cd [路径]

其中的路径为要改变的工作目录，可为相对路径或绝对路径。

（3）使用实例

```
[root@www uclinux]# cd /home/sunq/
[root@www sunq]# pwd
[root@www sunq]# /home/sunq/
```

该实例中变更工作目录为“/home/sunq/”，在后面的 pwd（显示当前目录）的结果中可以看出。

（4）使用说明

- 该命令将当前目录改变至指定路径的目录。若没有指定路径，则回到用户的主目录。为了改变到指定目录，用户必须拥有对指定目录的执行和读权限。
- 该命令可以使用通配符。
- 可使用“cd –”可以回到前次工作目录。
- “./”代表当前目录，“../”代表上级目录。

2．ls

（1）作用

列出目录的内容。

（2）格式：ls [选项] [文件]

其中文件选项为指定查看指定文件的相关内容，若未指定文件，默认查看当前目录下的所有文件。

（3）常见参数

ls 主要选项参数见表 2.9 所示

表 2.9　　ls 命令常见参数列表

选　　项	参 数 含 义
-1，--format=single-column	一行输出一个文件（单列输出）
-a，-all	列出目录中所有文件，包括以“.”开头的文件
-d	将目录名和其他文件一样列出，而不是列出目录的内容
-l,--format=long, --format=verbose	除每个文件名外，增加显示文件类型、权限、硬链接数、所有者名、组名、大小（Byte）及时间信息（如未指明是其他时间即指修改时间）
-f	不排序目录内容，按它们在磁盘上存储的顺序列出

（4）使用实例

```
[yuling@www /]$ ls -l
total 220
drwxr-xr-x   2 root    root        4096 Mar 31  2005 bin
drwxr-xr-x   3 root    root        4096 Apr  3  2005 boot
-rw-r--r--   1 root    root           0 Apr 24  2002 test.run
…
```

该实例查看当前目录下的所有文件，并通过选项“-l”显示出详细信息。

显示格式说明如下。

文件类型与权限 链接数 文件属主 文件属组 文件大小 修改的时间 名字

（5）使用说明

- 在 ls 的常见参数中，-l（长文件名显示格式）的选项是最为常见的。可以详细显示出各种信息。
- 若想显示出所有“.”开头的文件，可以使用-a，这在嵌入式的开发中很常用。

注意 Linux 中的可执行文件不是与 Windows 一样通过文件扩展名来标识的，而是通过设置文件相应的可执行属性来实现的。

3．mkdir

（1）作用

创建一个目录。

（2）格式

mkdir [选项] 路径

（3）常见参数

mkdir 主要选项参数如表 2.10 所示

表 2.10 mkdir 命令常见参数列表

选　项	参 数 含 义
-m	对新建目录设置存取权限，也可以用 chmod 命令（在本节后会有详细说明）设置
-p	可以是一个路径名称。此时若此路径中的某些目录尚不存在，在加上此选项后，系统将自动建立好那些尚不存在的目录，即一次可以建立多个目录

（4）使用实例

```
[root@www sunq]# mkdir -p ./hello/my
[root@www my]# pwd（查看当前目录命令）
/home/sunq/hello/my
```

该实例使用选项“-p”一次创建了./hello/my 多级目录。

```
[root@www my]# mkdir -m 777 ./why
[root@www my]# ls -l
total 4
drwxrwxrwx    2 root    root        4096 Jan 14 09:24 why
```

该实例使用改选项“-m”创建了相应权限的目录。对于“777”的权限在本节后面会有详细的说明。

（5）使用说明

该命令要求创建目录的用户在创建路径的上级目录中具有写权限，并且路径名不能是当前目录中已有的目录或文件名称。

4．cat

（1）作用

连接并显示指定的一个和多个文件的有关信息。

（2）格式

cat[选项]文件 1 文件 2…

其中的文件 1、文件 2 为要显示的多个文件。

（3）常见参数

cat 命令的常见参数如表 2.11 所示。

表 2.11 cat 命令常见参数列表

选　项	参 数 含 义
-n	由第一行开始对所有输出的行数编号
-b	和-n 相似，只不过对于空白行不编号

（4）使用实例

```
[yul@www yul]$ cat -n hello1.c hello2.c
     1  #include <stdio.h>
     2  void main()
     3  {
     4          printf("Hello!This is my home!\n");
     5  }
     6  #include <stdio.h>
     7  void main()
     8  {
     9          printf("Hello!This is your home!\n");
    10  }
```

在该实例中，指定对 hello1.c 和 hello2.c 进行输出，并指定行号。

5．cp、mv 和 rm

（1）作用

① cp：将给出的文件或目录复制到另一文件或目录中。

② mv：为文件或目录改名或将文件由一个目录移入另一个目录中。

③ rm：删除一个目录中的一个或多个文件或目录。

（2）格式

① cp：cp [选项] 源文件或目录 目标文件或目录。

② mv：mv [选项] 源文件或目录 目标文件或目录。

③ rm：rm [选项] 文件或目录。

（3）常见参数

① cp 主要选项参数见表 2.12 所示。

表 2.12 cp 命令常见参数列表

选　项	参 数 含 义
-a	保留链接、文件属性，并复制其子目录，其作用等于 dpr 选项的组合

-d	拷贝时保留链接
-f	删除已经存在的目标文件而不提示
-i	在覆盖目标文件之前将给出提示要求用户确认。回答 y 时目标文件将被覆盖，而且是交互式拷贝
-p	此时 cp 除复制源文件的内容外，还将把其修改时间和访问权限也复制到新文件中
-r	若给出的源文件是一目录文件，此时 cp 将递归复制该目录下所有的子目录和文件。此时目标文件必须为一个目录名

② mv 主要选项参数如表 2.13 所示。

表 2.13　　mv 命令常见参数列表

选　项	参 数 含 义
-i	若 mv 操作将导致对已存在的目标文件的覆盖，此时系统询问是否重写，并要求用户回答 y 或 n，这样可以避免误覆盖文件
-f	禁止交互操作。在 mv 操作要覆盖某已有的目标文件时不给任何指示，在指定此选项后，i 选项将不再起作用

③ rm 主要选项参数如表 2.14 所示。

表 2.14　　rm 命令常见参数列表

选　项	参 数 含 义
-i	进行交互式删除
-f	忽略不存在的文件，但从不给出提示
-r	指示 rm 将参数中列出的全部目录和子目录均递归地删除

（4）使用实例

① cp

```
[root@www hello]# cp -a ./my/why/ ./
[root@www hello]# ls
my  why
```

该实例使用-a 选项将“/my/why”目录下的所有文件复制到当前目录下。而此时在原先目录下还有原有的文件。

② mv

```
[root@www hello]# mv -i ./my/why/ ./
[root@www hello]# ls
my  why
```

该实例中把“/my/why”目录下的所有文件移至当前目录，则原目录下文件被自动删除。

③ rm

```
[root@www hello]# rm –r -i ./why
rm: descend into directory './why'? y
rm: remove './why/my.c'? y
```

```
rm: remove directory './why'? y
```

该实例使用“-r”选项删除“./why”目录下所有内容，系统会进行确认是否删除。

（5）使用说明

① cp：该命令把指定的源文件复制到目标文件或把多个源文件复制到目标目录中。

② mv：

- 该命令根据命令中第二个参数类型的不同（是目标文件还是目标目录）来判断是重命名还是移动文件，当第二个参数类型是文件时，mv 命令完成文件重命名，此时，它将所给的源文件或目录重命名为给定的目标文件名；
- 当第二个参数是已存在的目录名称时，mv 命令将各参数指定的源文件均移至目标目录中；
- 在跨文件系统移动文件时，mv 先复制，再将原有文件删除，而链至该文件的链接也将丢失。

③ rm：

- 如果没有使用- r 选项，则 rm 不会删除目录；
- 使用该命令时一旦文件被删除，它是不能被恢复的，所以最好使用-i 参数。

6．chown 和 chgrp

（1）作用

① chown：修改文件所有者和组别。

② chgrp：改变文件的组所有权。

（2）格式

① chown：chown [选项]...文件所有者[所有者组名] 文件

其中的文件所有者为修改后的文件所有者。

② chgrp：chgrp [选项]... 文件所有组 文件

其中的文件所有组为改变后的文件组拥有者。

（3）常见参数

chown 和 chgrp 的常见参数意义相同，其主要选项参数如表 2.15 所示。

表 2.15 chown 和 chgrp 命令常见参数列表

选　项	参 数 含 义
-c，-changes	详尽地描述每个 file 实际改变了哪些所有权
-f，--silent,--quiet	不打印文件所有权就不能修改的报错信息

（4）使用实例

在笔者的系统中一个文件的所有者原先是这样的。

```
[root@www sunq]# ls -l
-rwxr-xr-x   15 apectel  sunq        4096  6月  4  2005 uClinux-dist.tar
```

可以看出，这是一个文件，它的文件拥有者是 apectel，具有可读写和执行的权限，它所属的用户组是 sunq，具有可读和执行的权限，但没有可写的全权，同样，系统其他用户对其也只有可读和执行的权限。

首先使用 chown 将文件所有者改为 root。

```
[root@www sunq]# chown root uClinux-dist.tar
[root@www sunq]# ls –l
-rwxr-xr-x   15 root    sunq        4096  6月  4  2005 uClinux-dist.tar
```

可以看出，此时，该文件拥有者变为了 root，它所属文件用户组不变。

接着使用 chgrp 将文件用户组变为 root。

```
[root@www sunq]# chgrp root uClinux-dist.tar
[root@www sunq]# ls –l
-rwxr-xr-x   15 root    root        4096  6月  4  2005 uClinux-dist.tar
```

（5）使用说明

- 使用 chown 和 chgrp 必须拥有 root 权限。

在进行有关文件的操作时，若想避免输入冗长的文件，在文件名没有重复的情况下可以使用输入文件前几个字母 + <Tab>键的方式，即：cd:/uC<tab> = cd:/uClinux-list

7．chmod

（1）作用

改变文件的访问权限。

（2）格式

chmod 可使用符号标记进行更改和八进制数指定更改两种方式，因此它的格式也有两种不同的形式。

① 符号标记：chmod [选项]…符号权限[符号权限]…文件

其中的符号权限可以指定为多个，也就是说，可以指定多个用户级别的权限，但它们中间要用逗号分开表示，若没有显示指出则表示不作更改。

② 八进制数：chmod [选项] …八进制权限 文件…

其中的八进制权限是指要更改后的文件权限。

（3）选项参数

chmod 主要选项参数如表 2.16 所示。

表 2.16　　chmod 命令常见参数列表

选　项	参 数 含 义
-c	若该文件权限确实已经更改，才显示其更改动作
-f	若该文件权限无法被更改也不要显示错误信息
-v	显示权限变更的详细资料

（4）使用实例

chmod 涉及文件的访问权限，在此对相关的概念进行简单的回顾。

在 1.3.1 节中已经提到，文件的访问权限可表示成：**- rwx rwx rwx**。在此设有三种不同的访问权限：读（r）、写（w）和运行（x）。三个不同的用户级别：文件拥有者（u）、所属

的用户组（g）和系统里的其他用户（o）。在此，可增加一个用户级别 a（all）来表示所有这三个不同的用户级别。

① 对于第一种符号连接方式的 chmod 命令中，用加号“+”代表增加权限，用减号“–”删除权限，等于号“=”设置权限。

例如原先笔者系统中有文件 uClinux20031103.tgz，其权限如下所示。

```
[root@www sunq]# ls –l
-rw-r--r--    1 root    root    79708616 Mar 24  2005 uClinux20031103.tgz
[root@www sunq]# chmod a+rx,u+w uClinux20031103.tgz
[root@www sunq]# ls –l
-rwxr-xr-x    1 root    root    79708616 Mar 24  2005 uClinux20031103.tgz
```

可见，在执行了 chmod 之后，文件拥有者除拥有所有用户都有的可读和执行的权限外，还有可写的权限。

② 对于第二种八进制数指定的方式，将文件权限字符代表的有效位设为“1”，即“rw-”、“rw-”和“r--”的八进制表示为“110”、“110”、“100”，把这个 2 进制串转换成对应的 8 进制数就是 6、6、4，也就是说该文件的权限为 664（三位八进制数）。这样对于转化后 8 进制数、2 进制及对应权限的关系如表 2.17 所示。

表 2.17　　转化后 8 进制数、2 进制及对应权限的关系

转换后 8 进制数	2 进 制	对 应 权 限	转换后 8 进制数	2 进 制	对 应 权 限
0	000	没有任何权限	1	001	只能执行
2	010	只写	3	011	只写和执行
4	100	只读	5	101	只读和执行
6	110	读和写	7	111	读，写和执行

同上例，原先笔者系统中有文件 genromfs-0.5.1.tar.gz，其权限如下所示。

```
[root@www sunq]# ls –l
-rw-rw-r--    1 sunq    sunq       20543 Dec 29  2004 genromfs-0.5.1.tar.gz
[root@www sunq]# chmod 765 genromfs-0.5.1.tar.gz
[root@www sunq]# ls –l
-rwxrw-r-x    1 sunq    sunq       20543 Dec 29  2004 genromfs-0.5.1.tar.gz
```

可见，在执行了 chmod 765 之后，该文件的拥有者权限、文件组权限和其他用户权限都恰当地对应了。

（5）使用说明

- 使用 chmod 必须具有 root 权限。

想一想　chmod o+x uClinux20031103.tgz 是什么意思？它所对应的 8 进制数指定更改应如何表示？

8．grep

（1）作用

在指定文件中搜索特定的内容，并将含有这些内容的行标准输出。

（2）格式

grep [选项] 格式 [文件及路径]

其中的格式是指要搜索的内容格式，若缺省“文件及路径”则默认表示在当前目录下搜索。

（3）常见参数

grep 主要选项参数如表 2.18 所示。

表 2.18　　grep 命令常见参数列表

选　　项	参 数 含 义
-c	只输出匹配行的计数
-I	不区分大小写（只适用于单字符）
-h	查询多文件时不显示文件名
-l	查询多文件时只输出包含匹配字符的文件名
-n	显示匹配行及行号
-s	不显示不存在或无匹配文本的错误信息
-v	显示不包含匹配文本的所有行

（4）使用实例

```
    [root@www sunq]# grep "hello" / -r
    Binary file ./iscit2005/备份/iscit2004.sql matches
    ./ARM_TOOLS/uClinux-Samsung/linux-2.4.x/Documentation/s390/Debugging390.t
xt:hello world$2 = 0
    ...
```

该本例中，“hello”是要搜索的内容，“/ -r”是指定文件，表示搜索根目录下的所有文件。

（5）使用说明

- 在缺省情况下，“grep”只搜索当前目录。如果此目录下有许多子目录，“grep”会以如下形式列出：“grep:sound:Is a directory”这会使“grep”的输出难于阅读。但有两种解决的方法：

① 明确要求搜索子目录：grep –r（正如上例中所示）；

② 忽略子目录：grep -d skip。

- 当预料到有许多输出，可以通过管道将其转到“less”（分页器）上阅读：如 grep "h" ./ -r |less 分页阅读。
- grep 特殊用法：

grep pattern1|pattern2 files：显示匹配 pattern1 或 pattern2 的行；

grep pattern1 files|grep pattern2：显示既匹配 pattern1 又匹配 pattern2 的行；

小知识　在文件命令中经常会使用 pattern 正则表达式，它是可以描述一类字符串的模式（Pattern），如果一个字符串可以用某个正则表达式来描述，就称这个字符和该正则表达式匹配。这和 DOS

中用户可以使用通配符“*”代表任意字符类似。在 Linux 系统上，正则表达式通常被用来查找文本的模式，以及对文本执行“搜索－替换”操作等。

正则表达式的主要参数有：

- \：忽略正则表达式中特殊字符的原有含义；
- ^：匹配正则表达式的开始行；
- $：匹配正则表达式的结束行；
- <：从匹配正则表达式的行开始；
- >：到匹配正则表达式的行结束；
- [ ]：单个字符，如[A]即 A 符合要求；
- [-]：范围，如[A-Z]，即 A、B、C 一直到 Z 都符合要求；
- 。：所有的单个字符；
- *：所有字符，长度可以为 0。

9．find

（1）作用

在指定目录中搜索文件，它的使用权限是所有用户。

（2）格式

find [路径][选项][描述]

其中的路径为文件搜索路径，系统开始沿着此目录树向下查找文件。它是一个路径列表，相互用空格分离。若缺省路径，那么默认为当前目录。

其中的描述是匹配表达式，是 find 命令接受的表达式。

（3）常见参数

[选项]主要参数如表 2.19 所示。

表 2.19 find 选项常见参数列表

选　　项	参 数 含 义
-depth	使用深度级别的查找过程方式，在某层指定目录中优先查找文件内容
-mount	不在其他文件系统（如 Msdos、Vfat 等）的目录和文件中查找

[描述]主要参数如表 2.20 所示。

表 2.20 find 描述常见参数列表

选　　项	参 数 含 义
-name	支持通配符*和?
-user	用户名：搜索文件属主为用户名（ID 或名称）的文件
-print	输出搜索结果，并且打印

（4）使用实例

```
[root@www sunq]# find ./ -name qiong*.c
./qiong1.c
./iscit2005/qiong.c
```

在该实例中使用了-name 的选项支持通配符。

（5）使用说明

- 若使用目录路径为“/”，通常需要查找较多的时间，可以指定更为确切的路径以减少查找时间。
- find 命令可以使用混合查找的方法，例如，想在/etc 目录中查找大于 500000 字节，并且在 24 小时内修改的某个文件，则可以使用-and（与）把两个查找参数链接起来组合成一个混合的查找方式，如“find /etc -size +500000c -and -mtime +1”。

10．locate

（1）作用

用于查找文件。其方法是先建立一个包括系统内所有文件名称及路径的数据库，之后当寻找时就只需查询这个数据库，而不必实际深入档案系统之中了。因此其速度比 find 快很多。

（2）格式

locate [选项]

（3）locate 主要选项参数如表 2.21 所示。

表 2.21　　locate 命令常见参数列表

选　　项	参 数 含 义
-u	从根目录开始建立数据库
-U	指定开始的位置建立数据库
-f	将特定的文件系统排除在数据库外，例如 proc 文件系统中的文件
-r	使用正则运算式做寻找的条件
-o	指定数据库存的名称

（4）使用实例

```
[root@www sunq]# locate issue -U ./
[root@www sunq]# updatedb
[root@www sunq]# locate -r issue*
./ARM_TOOLS/uClinux-Samsung/lib/libpam/doc/modules/pam_issue.sgml
./ARM_TOOLS/uClinux-Samsung/lib/libpam/modules/pam_issue
./ARM_TOOLS/uClinux-Samsung/lib/libpam/modules/pam_issue/Makefile
./ARM_TOOLS/uClinux-Samsung/lib/libpam/modules/pam_issue/pam_issue.c
...
```

示例中首先在当前目录下建立了一个数据库，并且在更新了数据库之后进行正则匹配查找。通过运行可以发现 locate 的运行速度非常快。

（5）使用说明

locate 命令所查询的数据库由 updatedb 程序来更新的，而 updatedb 是由 cron daemon 周期性建立的，但若所找到的档案是最近才建立或刚更名的，可能会找不到，因为 updatedb 默认每天运行一次，用户可以由修改 crontab（etc/crontab）来更新周期值。

11. ln

（1）作用

为某一个文件在另外一个位置建立一个符号链接。当需要在不同的目录用到相同的文件时，Linux 允许用户不用在每一个需要的目录下都存放一个相同的文件，而只需将其他目录下文件用 ln 命令链接即可，这样就不必重复地占用磁盘空间。

（2）格式

ln[选项] 目标 目录

（3）常见参数

Ø –s 建立符号链接（这也是通常惟一使用的参数）。

（4）使用实例

```
[root@www uclinux]# ln -s ../genromfs-0.5.1.tar.gz ./hello
[root@www uclinux]# ls -l
total 77948
lrwxrwxrwx   1 root   root    24 Jan 14 00:25 hello -> ../genromfs-0.5.1.tar.gz
```

该实例建立了当前目录的 hello 文件与上级目录之间的符号连接，可以看见，在 hello 的 ls –l 中的第一位为“l”，表示符号链接，同时还显示了链接的源文件。

（5）使用说明

- ln 命令会保持每一处链接文件的同步性，也就是说，不论改动了哪一处，其他的文件都会发生相同的变化。
- ln 的链接又软链接和硬链接两种：

软链接就是上面所说的 ln -s ** **，它只会在用户选定的位置上生成一个文件的镜像，不会重复占用磁盘空间，平时使用较多的都是软链接；

硬链接是不带参数的 ln ** **，它会在用户选定的位置上生成一个和源文件大小相同的文件，无论是软链接还是硬链接，文件都保持同步变化。

2.1.3 压缩打包相关命令

Linux 中打包压缩的如表 2.22 所示，本书以 gzip 和 tar 为例进行讲解。

表 2.22 Linux 常见系统管理命令

命 令	命 令 含 义	格 式
bzip2	.bz2 文件的压缩（或解压）程序	bzip2[选项] 压缩（解压缩）的文件名
bunzip2	.bz2 文件的解压缩程序	bunzip2[选项] .bz2 压缩文件
bzip2recover	用来修复损坏的.bz2 文件	bzip2recover .bz2 压缩文件
gzip	.gz 文件的压缩程序	gzip [选项] 压缩（解压缩）的文件名
gunzip	解压被 gzip 压缩过的文件	gunzip [选项] .gz 文件名
unzip	解压 winzip 压缩的.zip 文件	unzip [选项] .zip 压缩文件
compress	早期的压缩或解压程序（压缩后文件名为.Z）	compress [选项] 文件
tar	对文件目录进行打包或解包	tar [选项] [打包后文件名]文件目录列表

1．gzip

（1）作用

对文件进行压缩和解压缩，而且 gzip 根据文件类型可自动识别压缩或解压。

（2）格式

gzip [选项] 压缩（解压缩）的文件名。

（3）常见参数

gzip 主要选项参数如表 2.23 所示。

表 2.23　　gzip 命令常见参数列表

选　　项	参 数 含 义
-c	将输出信息写到标准输出上，并保留原有文件
-d	将压缩文件解压
-l	对每个压缩文件，显示下列字段：压缩文件的大小、未压缩文件的大小、压缩比、未压缩文件的名字
-r	查找指定目录并压缩或解压缩其中的所有文件
-t	测试，检查压缩文件是否完整
-v	对每一个压缩和解压的文件，显示文件名和压缩比

（4）使用实例

```
[root@www my]# gzip hello.c
[root@www my]# ls
hello.c.gz
[root@www my]# gzip -l hello.c
         compressed            uncompressed  ratio uncompressed_name
61                   39.3% hello.c
```

该实例将目录下的“hello.c”文件进行压缩，选项“-l”列出了压缩比。

（5）使用说明

- 使用 gzip 压缩只能压缩单个文件，而不能压缩目录，其选项“-d”是将该目录下的所有文件逐个进行压缩，而不是压缩成一个文件。

2．tar

（1）作用

对文件目录进行打包或解包。

在此需要对**打包**和**压缩**这两个概念进行区分。**打包**是指将一些文件或目录变成一个总的文件，而**压缩**则是将一个大的文件通过一些压缩算法变成一个小文件。为什么要区分这两个概念呢？这是由于在 Linux 中的很多压缩程序（如前面介绍的 gzip）只能针对一个文件进行压缩，这样当想要压缩较多文件时，就要借助它的工具将这些堆文件先打成一个包，然后再用原来的压缩程序进行压缩。

（2）格式

tar [选项] [打包后文件名]文件目录列表。

tar 可自动根据文件名识别打包或解包动作，其中打包后文件名为用户自定义的打包后文件名称，文件目录列表可以是要进行打包备份的文件目录列表，也可以是进行解包的文件目录列表。

（3）主要参数

tar 主要选项参数如表 2.24 所示。

表 2.24 tar 命令常见参数列表

选　项	参 数 含 义
-c	建立新的打包文件
-r	向打包文件末尾追加文件
-x	从打包文件中解出文件
-o	将文件解开到标准输出
-v	处理过程中输出相关信息
-f	对普通文件操作
-z	调用 gzip 来压缩打包文件，与-x 联用时调用 gzip 完成解压缩
-j	调用 bzip2 来压缩打包文件，与-x 联用时调用 bzip2 完成解压缩
-Z	调用 compress 来压缩打包文件，与-x 联用时调用 compress 完成解压缩

（4）使用实例

```
[root@www home]# tar -cvf yul.tar ./yul
./yul/
./yul/.bash_logout
./yul/.bash_profile
./yul/.bashrc
./yul/.bash_history
./yul/my/
./yul/my/1.c.gz
./yul/my/my.c.gz
./yul/my/hello.c.gz
./yul/my/why.c.gz
[root@www home]# ls -l yul.tar
-rw-r--r--    1 root    root       10240 Jan 14 15:01 yul.tar
```

该实例将“./yul”目录下的文件加以打包，其中选项“-v”在屏幕上输出了打包的具体过程。

```
[root@www sunq]# tar -zxvf linux-2.6.11.tar.gz
linux-2.6.11/
linux-2.6.11/drivers/
linux-2.6.11/drivers/video/
linux-2.6.11/drivers/video/aty/
```

...

该实例用选项“-z”调用 gzip，并-x 联用时完成解压缩。

（5）使用说明

tar 命令除了用于常规的打包之外，使用更为频繁的是用选项“-z”或“-j”调用 gzip 或 bzip2（Linux 中另一种解压工具）完成对各种不同文件的解压。

表 2.25 对 Linux 中常见类型的文件解压命令做一总结。

表 2.25　　Linux 常见类型的文件解压命令一览表

文 件 后 缀	解 压 命 令	示　　例
.a	tar xv	tar xv hello.a
.z	Uncompress	uncompress hello.Z
.gz	Gunzip	gunzip hello.gz
.tar.Z	tar xvZf	tar xvZf hello.tar.Z
.tar.gz/.tgz	tar xvzf	tar xvzf hello.tar.gz
tar.bz2	tar jxvf	tar jxvf hello.tar.bz2
.rpm	安装：rpm –i	安装：rpm -i hello.rpm
	解压：rpm2cpio	解压：rpm2cpio hello.rpm
.deb（Debain 中的文件格式）	安装：dpkg –i	安装：dpkg -i hello.deb
	解压：dpkg-deb --fsys-tarfile	解压：dpkg-deb --fsys-tarhello hello.deb
.zip	Unzip	unzip hello.zip

2.1.4　比较合并文件相关命令

1．diff

（1）作用

比较两个不同的文件或不同目录下的两个同名文件功能，并生成补丁文件。

（2）格式

diff[选项] 文件 1 文件 2

diff 比较文件 1 和文件 2 的不同之处，并按照选项所指定的格式加以输出。diff 的格式分为命令格式和上下文格式，其中上下文格式又包括了旧版上下文格式和新版上下文格式，命令格式分为标准命令格式、简单命令格式及混合命令格式，它们之间的区别会在使用实例中进行详细地讲解。当选项缺省时，diff 默认使用混合命令格式。

（3）主要参数

diff 主要选项参数如表 2.26 所示。

表 2.26　　diff 命令常见参数列表

选　　项	参 数 含 义
-r	对目录进行递归处理
-q	只报告文件是否有不同，不输出结果
-e，-ed	命令格式
-f	RCS（修订控制系统）命令简单格式

-c，--context	旧版上下文格式
-u，--unified	新版上下文格式
-Z	调用 compress 来压缩归档文件，与-x 联用时调用 compress 完成解压缩

（4）使用实例

以下有两个文件 hello1.c 和 hello2.c。

```
//hello1.c
#include <stdio.h>
void main()
{
        printf("Hello!This is my home!\n");
}
//hello2.c
#include <stdio.h>
void main()
{
        printf("Hello!This is your home!\n");
}
```

以下实例主要讲解了各种不同格式的比较和补丁文件的创建方法。

① 主要格式比较

首先使用**旧版上下文格式**进行比较。

```
[root@www yul]# diff -c hello1.c hello2.c
*** hello1.c   Sat Jan 14 16:24:51 2006
--- hello2.c   Sat Jan 14 16:54:41 2006
***************
*** 1,5 ****
  #include <stdio.h>
  void main()
  {
!       printf("Hello!This is my home!\n");
  }
--- 1,5 ----
  #include <stdio.h>
  void main()
  {
!       printf("Hello!This is your home!\n");
  }
```

可以看出，用旧版上下文格式进行输出时，在显示每个有差别行的同时还显示该行的上下三行，区别的地方用“!”加以标出，由于示例程序较短，上下三行已经包含了全部代码。

接着使用**新版的上下文格式**进行比较。

```
[root@www yul]# diff -u hello1.c hello2.c
--- hello1.c    Sat Jan 14 16:24:51 2006
+++ hello2.c    Sat Jan 14 16:54:41 2006
@@ -1,5 +1,5 @@
 #include <stdio.h>
 void main()
 {
-       printf("Hello!This is my home!\n");
+       printf("Hello!This is your home!\n");
 }
```

可以看出，在新版上下文格式输出时，仅把两个文件的不同之处分别列出，而相同之处没有重复列出，这样大大方便了用户的阅读。

接下来使用**命令格式**进行比较。

```
[root@www yul]# diff -e hello1.c hello2.c
4c
       printf("Hello!This is your home!\n");
```

可以看出，命令符格式输出时仅输出了不同的行，其中命令符“4c”中的数字表示行数，字母的含义为 **a——添加，b——删除，c——更改**。因此，-e 选项的命令符表示：若要把 hello1.c 变为 hello2.c，就需要把 hello1.c 的第四行改为显示出的“printf(“Hello!This is your home!\n”);”即可。

选项“-f”和选项“-e”显示的内容基本相同，就是数字和字母的顺序相交换了，从以下的输出结果可以看出。

```
[root@www yul]# diff -f hello1.c hello2.c
c4
       printf("Hello!This is your home!\n");
```

在 diff 选项缺省的情况下，输出结果如下所示。

```
[root@www yul]# diff hello1.c hello2.c
4c4
<      printf("Hello!This is my home!\n");
---
>      printf("Hello!This is your home!\n");
```

可以看出，diff 缺省情况下的输出格式充分显示了如何将 hello1.c 转化为 hello2.c 的方法，即通过“4c4”实现。

② 创建补丁文件（也就是差异文件）是 diff 的功能之一，不同的选项格式可以生成与之相对应的补丁文件。见下例。

```
[root@www yul]# diff hello1.c hello2.c >hello.patch
[root@www yul]# vi hello.patch
4c4
<       printf("Hello!This is my home!\n");
---
>       printf("Hello!This is your home!\n");
```

可以看出，使用缺省选项创建补丁文件的内容和前面使用缺省选项的输出内容是一样的。

小知识

上例中所使用的”>“是输出重定向。通常在 Linux 上执行一个 shell 命令行时，会自动打开三个标准文件：标准输入文件（stdin），即通常对应终端的键盘；标准输出文件（stdout）和标准错误输出文件（stderr），前两个文件都对应终端的屏幕。进程将从标准输入文件中得到输入数据，并且将正常输出数据输出到标准输出文件，而将错误信息送到标准错误文件中。这就是通常使用的**标准输入/输出方式**。

直接使用标准输入/输出文件存在以下问题：首先，用户输入的数据只能使用一次。当下次希望再次使用这些数据时就不得不重新输入。同样，用户对输出信息不能做更多的处理，只能等待程序的结束。

为了解决上述问题，Linux 系统为输入、输出的信息传送引入了两种方式：**输入/输出重定向机制和管道**（在 1.3.1 的小知识中已有介绍）。其中，输入重定向是指把命令（或可执行程序）的标准输入重定向到指定的文件中。也就是说，输入可以不来自键盘，而来自一个指定的文件。同样，输出重定向是指把命令（或可执行程序）的标准输出或标准错误输出重新定向到指定文件中。这样，该命令的输出就可以不显示在屏幕上，而是写入到指定文件中。就如上述例子中所用到的把“diff hello1.c hello2.c”的结果重定向到 hello.patch 文件中。这就大大增加了输入/输出的灵活性。

2．patch

（1）作用

命令跟 diff 配合使用，把生成的补丁文件应用到现有代码上。

（2）格式

patch [选项] [待 patch 的文件[patch 文件]]。

常用的格式为：patch -pnum [patch 文件]，其中的-pnum 是选项参数，在后面会详细介绍。

（3）常见参数

patch 主要选项参数如表 2.27 所示。

表 2.27 patch 命令常见参数列表

选　项	参 数 含 义
-b	生成备份文件
-d	把 dir 设置为解释补丁文件名的当前目录

-e	把输入的补丁文件看作是 ed 脚本
-pnum	剥离文件名中的前 NUM 个目录成分
-t	在执行过程中不要求任何输入
-v	显示 patch 的版本号

以下对-punm 选项进行说明。

首先查看以下示例（对分别位于 xc.orig/config/cf/Makefile 和 xc.bsd/config/cf/Makefile 的文件使用 patch 命令）。

```
diff -ruNa xc.orig/config/cf/Makefile xc.bsd/config/cf/Makefile
```

以下是 patch 文件的头标记。

```
--- xc.orig/config/cf/Imake.cf Fri Jul 30 12:45:47 1999
+++ xc.new/config/cf/Imake.cf Fri Jan 21 13:48:44 2000
```

这个 patch 如果直接应用，那么它会去找“xc.orig/config/cf”目录下的 Makefile 文件，假如用户源码树的根目录是缺省的 xc 而不是 xc.orig，则除了可以把 xc.orig 移到 xc 处之外，还有什么简单的方法应用此 patch 吗？NUM 就是为此而设的：patch 会把目标路径名剥去 NUM 个“/”，也就是说，在此例中，-p1 的结果是 config/cf/Makefile，-p2 的结果是 cf/Makefile。因此，在此例中就可以用命令 cd xc；patch _p1 < /pathname/xxx.patch 完成操作。

（4）使用实例

```
[root@www yul]# diff hello1.c hello2.c >hello1.patch
[root@www yul]# patch ./hello1.c < hello1.patch
patching file ./hello1.c
[root@www yul] ]# vi hello1.c
#include <stdio.h>
void main()
{
       printf("Hello!This is your home!\n");
}
```

在该实例中，由于 patch 文件和源文件在同一目录下，因此直接给出了目标文件的目录，在应用了 patch 之后，hello1.c 的内容变为了 hello2.c 的内容。

（5）使用说明

- 如果 patch 失败，patch 命令会把成功的 patch 行补上其差异，同时（无条件）生成备份文件和一个.rej 文件。.rej 文件里是没有成功提交的 patch 行，需要手工打上补丁。这种情况在原码升级的时候有可能会发生。
- 在多数情况下，patch 程序可以确定补丁文件的格式，当它不能识别时，可以使用-c、-e、-n 或者-u 选项来指定输入的补丁文件的格式。由于只有 GNU patch 可以创建和读取新版上下文格式的 patch 文件，因此，除非能够确定补丁所面向的只是那些使用 GNU 工具的用户，否则应该使用旧版上下文格式来生成补丁文件。

- 为了使 patch 程序能够正常工作，需要上下文的行数至少是 2 行（即至少是有一处差别的文件）。

2.1.5 网络相关命令

Linux 下网络相关的常见命令如下表 2.28 所示，本书仅以 ifconfig 和 ftp 为例进行说明。

表 2.28 Linux 下网络相关命令

选　　项	参 数 含 义	常见选项格式
netstat	显示网络连接、路由表和网络接口信息	netstat [-an]
nslookup	查询一台机器的 IP 地址和其对应的域名	Nslookup [IP 地址/域名]
finger	查询用户的信息	finger [选项] [使用者] [用户@主机]
ping	用于查看网络上的主机是否在工作	ping [选项] 主机名/IP 地址
ifconfig	查看和配置网络接口的参数	ifconfig [选项] [网络接口]
ftp	利用 ftp 协议上传和下载文件	在本节中会详细讲述
telnet	利用 telnet 协议浏览信息	telent [选项] [IP 地址/域名]
ssh	利用 ssh 登录对方主机	ssh [选项] [IP 地址]

1．ifconfig

（1）作用

用于查看和配置网络接口的地址和参数，包括 IP 地址、网络掩码、广播地址，它的使用权限是超级用户。

（2）格式

ifconfig 有两种使用格式，分别用于查看和更改网络接口。

① ifconfig [选项] [网络接口]：用来查看当前系统的网络配置情况。

② ifconfig 网络接口 [选项] 地址：用来配置指定接口（如 eth0，eth1）的 IP 地址、网络掩码、广播地址等。

（3）常见参数

ifconfig 第二种格式常见选项参数如表 2.29 所示。

表 2.29 ftp 命令选项常见参数列表

选　　项	参 数 含 义
-interface	指定的网络接口名，如 eth0 和 eth1
up	激活指定的网络接口卡
down	关闭指定的网络接口
broadcast address	设置接口的广播地址
poin to point	启用点对点方式
address	设置指定接口设备的 IP 地址
netmask address	设置接口的子网掩码

（4）使用实例

首先，在本例中使用 ifconfig 的第一种格式来查看网口配置情况。

```
[root@sunq workplace]# ifconfig
eth0      Link encap:Ethernet  HWaddr 00:08:02:E0:C1:8A
          inet addr:59.64.205.70  Bcast:59.64.207.255  Mask:255.255.252.0
          inet6 addr: fe80::208:2ff:fee0:c18a/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:26931 errors:0 dropped:0 overruns:0 frame:0
          TX packets:3209 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:6669382 (6.3 MiB)  TX bytes:321302 (313.7 KiB)
          Interrupt:11

lo        Link encap:Local Loopback
          inet addr:127.0.0.1  Mask:255.0.0.0
          inet6 addr: ::1/128 Scope:Host
          UP LOOPBACK RUNNING  MTU:16436  Metric:1
          RX packets:2537 errors:0 dropped:0 overruns:0 frame:0
          TX packets:2537 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:2093403 (1.9 MiB)  TX bytes:2093403 (1.9 MiB)
```

可以看出，使用 ifconfig 的显示结果中详细列出了所有活跃接口的 IP 地址、硬件地址、广播地址、子网掩码、回环地址等。

```
[root@sunq workplace]# ifconfig eth0
eth0      Link encap:Ethernet  HWaddr 00:08:02:E0:C1:8A
          inet addr:59.64.205.70  Bcast:59.64.207.255  Mask:255.255.252.0
          inet6 addr: fe80::208:2ff:fee0:c18a/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:27269 errors:0 dropped:0 overruns:0 frame:0
          TX packets:3212 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:6698832 (6.3 MiB)  TX bytes:322488 (314.9 KiB)
          Interrupt:11
```

在此例中，通过指定接口显示出对应接口的详细信息。另外，用户还可以通过指定参数“-a”来查看所有接口（包括非活跃接口）的信息。

接下来的示例指出了如何使用 ifconfig 的第二种格式来改变指定接口的网络参数配置。

```
[root@sunq ~]# ifconfig eth0 down
[root@sunq ~]# ifconfig
```

```
lo        Link encap:Local Loopback
          inet addr:127.0.0.1  Mask:255.0.0.0
          inet6 addr: ::1/128 Scope:Host
          UP LOOPBACK RUNNING  MTU:16436  Metric:1
          RX packets:1931 errors:0 dropped:0 overruns:0 frame:0
          TX packets:1931 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:2517080 (2.4 MiB)  TX bytes:2517080 (2.4 MiB)
```

在此例中，通过将指定接口的状态设置为 DOWN，暂时暂停该接口的工作。

```
[root@sunq workplace]# ifconfig eth0 210.25.132.142 netmask 255.255.255.0
[root@sunq workplace]# ifconfig
eth0      Link encap:Ethernet  HWaddr 00:08:02:E0:C1:8A
          inet addr:210.25.132.142  Bcast:210.25.132.255  Mask:255.255.255.0
          inet6 addr: fe80::208:2ff:fee0:c18a/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:1722 errors:0 dropped:0 overruns:0 frame:0
          TX packets:5 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:147382 (143.9 KiB)  TX bytes:398 (398.0 b)
          Interrupt:11
...
```

从上例可以看出，ifconfig 改变了接口 eth0 的 IP 地址、子网掩码等，在之后的 ifconfig 查看中可以看出确实发生了变化。

（5）使用说明

用 ifconfig 命令配置的网络设备参数不需重启就可生效，但在机器重新启动以后将会失效。

2．ftp

（1）作用

该命令允许用户利用 ftp 协议上传和下载文件。

（2）格式

ftp [选项] [主机名/IP]。

ftp 相关命令包括使用命令和内部命令，其中使用命令的格式如上所列，主要用于登录到 ftp 服务器的过程中使用的。内部命令是指成功登录后进行的一系列操作，下面会详细列出。若用户缺省“主机名/IP”，则可在转入到 ftp 内部命令后继续选择登录。

（3）常见参数

ftp 常见选项参数如表 2.30 所示。

表 2.30　　ftp 命令选项常见参数列表

选　　项	参 数 含 义
-v	显示远程服务器的所有响应信息
-n	限制 ftp 的自动登录
-d	使用调试方式
-g	取消全局文件名

ftp 常见内部命令如表 2.31 所示。

表 2.31　　ftp 命令常见内部命令

命　　令	命 令 含 义
account[password]	提供登录远程系统成功后访问系统资源所需的补充口令
Ascii	使用 ascii 类型传输方式，为缺省传输模式
bin/ type binary	使用二进制文件传输方式（嵌入式开发中常见方式）
Bye	退出 ftp 会话过程
cd remote-dir	进入远程主机目录
Cdup	进入远程主机目录的父目录
chmod mode file-name	将远程主机文件 file-name 的存取方式设置为 mode
Close	中断与远程服务器的 ftp 会话（与 open 对应）
delete remote-file	删除远程主机文件
debug[debug-value]	设置调试方式，显示发送至远程主机的每条命令
dir/ls[remote-dir][local-file]	显示远程主机目录，并将结果存入本地文件 local-file
Disconnection	同 Close
get remote-file[local-file]	将远程主机的文件 remote-file 传至本地硬盘的 local-file
lcd[dir]	将本地工作目录切换至 dir
mdelete[remote-file]	删除远程主机文件
mget remote-files	传输多个远程文件

续表

命　　令	命 令 含 义
mkdir dir-name	在远程主机中建一目录
mput local-file	将多个文件传输至远程主机
open host[port]	建立指定 ftp 服务器连接，可指定连接端口
Passive	进入被动传输方式（在这种模式下，数据连接是由客户程序发起的）
put local-file[remote-file]	将本地文件 local-file 传送至远程主机
reget remote-file[local-file]	类似于 get，但若 local-file 存在，则从上次传输中断处续传
size file-name	显示远程主机文件大小
System	显示远程主机的操作系统类型

（4）使用实例

首先，在本例中使用 ftp 命令访问“ftp://study.byr.edu.cn”站点。

```
[root@sunq ~]# ftp study.byr.edu.cn
Connected to study.byr.edu.cn.
220 Microsoft FTP Service
500 'AUTH GSSAPI': command not understood
500 'AUTH KERBEROS_V4': command not understood
KERBEROS_V4 rejected as an authentication type
Name (study.byr.edu.cn:root): anonymous
331 Anonymous access allowed, send identity (e-mail name) as password.
Password:
230 Anonymous user logged in.
Remote system type is Windows_NT.
```

注意 由于该站点可以匿名访问，因此，在用户名处输入 anonymous，在 Password 处输入任意一个 e-mail 地址即可登录成功。

```
ftp> dir
227 Entering Passive Mode (211,68,71,83,11,94).
125 Data connection already open; Transfer starting.
11-20-05  05:00PM       <DIR>          Audio
12-04-05  09:41PM       <DIR>          BUPT_NET_Material
01-07-06  01:38PM       <DIR>          Document
11-22-05  03:47PM       <DIR>          Incoming
01-04-06  11:09AM       <DIR>          Material
226 Transfer complete.
```

以上使用 ftp 内部命令 dir 列出了在改目录下文件及目录的信息。

```
ftp> cd /Document/Wrox/Wrox.Beginning.SQL.Feb.2005.eBook-DDU
250 CWD command successful.
ftp> pwd
257 "/Document/Wrox/Wrox.Beginning.SQL.Feb.2005.eBook-DDU" is current directory.
```

以上实例通过 cd 命令进入相应的目录，可通过 pwd 命令进行验证。

```
ftp> lcd /root/workplace
Local directory now /root/workplace
ftp> get d-wbsq01.zip
local: d-wbsq01.zip remote: d-wbsq01.zip
200 PORT command successful.
```

```
150 Opening ASCII mode data connection for d-wbsq01.zip(1466768 bytes).
WARNING! 5350 bare linefeeds received in ASCII mode
File may not have transferred correctly.
226 Transfer complete.
1466768 bytes received in 1.7 seconds (8.6e+02 Kbytes/s)
```

接下来通过 lcd 命令首先改变用户的本地工作目录，也就是希望下载或上传的工作目录，就着通过 get 命令进行下载文件。由于 ftp 默认使用 ASCII 模式，因此，若希望改为其他模式如“bin”，直接输入 bin 即可，如下所示：

```
ftp> bin
200 Type set to I.
ftp> bye
221
```

最后用 bye 命令退出 ftp 程序。

（5）使用说明

- 若是需要匿名登录，则在“Name (**.**.**.**):”处键入 anonymous，在“Password:”处键入自己的 E-mail 地址即可。
- 若要传送二进制文件，务必要把模式改为 bin。

2.2 Linux 启动过程详解

在了解了 Linux 的常见命令之后，下面来详细了解一下 Linux 的启动过程。Linux 的启动过程包含了 Linux 工作原理的精髓，而且在嵌入式的开发过程也非常需要这方面知识的积累。为了降低阅读的难度，在这部分作者尽量避免直接对大段的汇编代码进行分析，而主要阐述其原理及 INIT 进程。希望读者能认真耐心阅读，相信在真正掌握这一节的内容之后对 Linux 的认识会有一个质的飞跃。

本书假定读者对微机原理、操作系统、汇编和 C 语言已有了一定的基础，因此对以上知识的基础概念不再做详细的说明。

2.2.1 概述

用户开机启动 Linux 过程总体上是这样的：

首先当用户打开 PC 的电源时，CPU 将自动进入实模式，并从地址 0xFFFF0 开始自动执行程序代码，这个地址通常是 ROM-BIOS 中的地址。这时 BIOS 进行开机自检，并按 BIOS 中设置的启动设备（通常是硬盘）进行启动，接着启动设备上安装的引导程序 lilo 或 grub 开始引导 Linux（也就是启动设备的第一个扇区），这时，Linux 才获得了启动权。

接下来的第二阶段，Linux 首先进行内核的引导，主要完成磁盘引导、读取机器系统数据、实模式和保护模式的切换、加载数据段寄存器以及重置中断描述符表等。

第三阶段执行 init 程序（也就是系统初始化工作），init 程序调用了 rc.sysinit 和 rc 等程序，而 rc.sysinit 和 rc 在完成系统初始化和运行服务的任务后，返回 init。

之后的第四阶段，init 启动 mingetty，打开终端供用户登录系统，用户登录成功后进入了 Shell，这样就完成了从开机到登录的整个启动过程。

Linux 启动总体流程图如图 2.2 所示，其中的 4 个阶段分别由同步棒隔开。由于第一阶段不涉及 Linux 自身的启动过程，因此，下面分别对第二和第三阶段进行详细讲解。

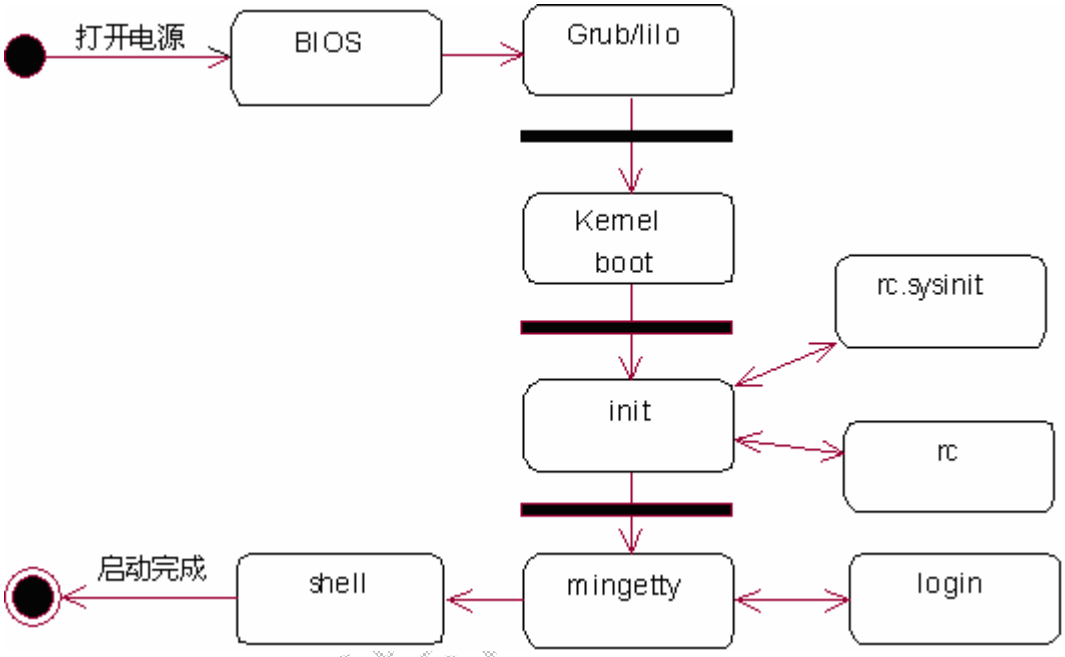


图 2.2 Linux 启动总体流程图

2.2.2 内核引导阶段

在 grub 或 lilo 等引导程序成功完成引导 Linux 系统的任务后，Linux 就从它们手中接管了 CPU 的控制权。用户可以从 www.kernel.org 上下载最新版本的源码进行阅读，其目录为：linux-2.6.*.*\arch\i386\boot。在这过程中主要用到该目录下的这几个文件：bootsect.S、setup.S 以及 compressed 目录下的 head.S 等。

首先要介绍一下，Linux 的内核通常是压缩过后的，包括如上述提到的那几个重要的汇编程序，它们都是在压缩内核 vmlinuz 中的。因为 Linux 中提供的内核包含了众多驱动和功能，因而比较大，所以在采用压缩内核可以节省大量的空间。

（1）bootsect 阶段

当 grub 读入 vmlinuz 后，会根据 bootsect（正好 512bytes）把它自身和 setup 程序段读到不大于 0x90000 开始的的内存里（注意：在以往的引导协议里是放在 0x90000，但现在有所变化），然后 grub 会跳过 bootsect 那 512bytes 的程序段，直接运行 setup 里的第一跳指令。就是说 bzImage 里 bootsect 的程序没有再被执行了，而 bootsect.S 在完成了指令搬移以后就退出了。之后执行权就转到了 setup.S 的程序中。

（2）setup 阶段

setup.S 的主要功能就是利用 ROM BIOS 中断读取机器系统数据，并将系统参数（包括内存、磁盘等）保存到 0x90000～0x901FF 开始的内存中位置。

此外，setup.S 还将 video.S 中的代码包含进来，检测和设置显示器和显示模式。

最后，它还会设置 CPU 的控制寄存器 CR0（也称机器状态字），从而进入 32 位保护模式运行，并跳转到绝对地址为 0x100000（虚拟地址 0xC0000000+0x100000）处。当 CPU 跳到 0x100000 时，将执行“arch/i386/kernel/head.S”中的 startup_32。

（3）head.S 阶段

当运行到 head.S 时，系统已经运行在保护模式，而 head.S 完成的一个重要任务就是将内核解压。就如本节前面提到的，内核是通过压缩的方式放在内存中的，head.S 通过调用 misc.c 中定义的 decompress_kernel()函数，将内核 vmlinuz 解压到 0x100000 的。

接下来 head.S 程序完成完成寄存器、分页表的初始化工作，但要注意的是，这个 head.S 程序与完成解压缩工作的 head.S 程序是不同的，它在源代码中的位置是 arch/i386/kernel/head.S。

在完成了初始化之后，head.S 就跳转到 start_kernel()函数中去了。

（4）main.c 阶段

start_kernel()是“init/main.c”中的定义的函数，start_kernel()调用了一系列初始化函数，进行内核的初始化工作。要注意的是，在初始化之前系统中断仍然是被屏蔽的，另外内核也处于被锁定状态，以保证只有一个 CPU 用于 Linux 系统的启动。

在 start_kernel()的最后，调用了 init()函数，也就是下面要讲述的 INIT 阶段。

2.2.3 init 阶段

在加载了内核之后，由内核执行引导的第一个进程就是 INIT 进程，该进程号始终是“1”。INIT 进程根据其配置文件“/etc/inittab”主要完成系统的一系列初始化的任务。由于该配置文件是 INIT 进程执行的惟一依据，因此先对它的格式进行统一讲解。

inittab 文件中除了注释行外，每一行都有如下格式：

```
id:runlevels:action:process
```

（1）id

id 是配置记录标识符，由 1～4 个字符组成，对于 getty 或 mingetty 等其他 login 程序项，要求 id 与 tty 的编号相同，否则 getty 程序将不能正常工作。

（2）runlevels

runlevels 是运行级别记录符，一般使用 0～6 以及 S 和 s。其中，0、1、6 运行级别为系统保留：0 作为 shutdown 动作，1 作为重启至单用户模式，6 为重启；S 和 s 意义相同，表示单用户模式，且无需 inittab 文件，因此也不在 inittab 中出现。7～9 级别也是可以使用的，传统的 UNIX 系统没有定义这几个级别。

runlevel 可以是并列的多个值，对大多数 action 来说，仅当 runlevel 与当前运行级别匹配成功才会执行。

（3）action

action 字段用于描述系统执行的特定操作，它的常见设置有：initdefault、sysinit、boot、

bootwait、respawn 等。

initdefault 用于标识系统缺省的启动级别。当 init 由内核激活以后，它将读取 inittab 中的 initdefault 项，取得其中的 runlevel，并作为当前的运行级别。如果没有 inittab 文件，或者其中没有 initdefault 项，init 将在控制台上请求输入 runlevel。

sysinit、boot、bootwait 等 action 将在系统启动时无条件运行，忽略其中的 runlevel。

respawn 字段表示该类进程在结束后会重新启动运行。

（4）process

process 字段设置启动进程所执行的命令。

以下结合笔者系统中的 inittab 配置文件详细讲解该配置文件完成的功能：

1．确定用户登录模式

在“/etc/inittab”中列出了如下所示的登录模式，主要有单人维护模式、多用户无网络模式、文字界面多用户模式、X-Windows 多用户模式等。其中的单人维护模式（run level 为 1）是类似于 Windows 中的“安全模式”，在这种情况下，系统不加载复杂的模式从而使系统能够正常启动。在这些模式中最为常见的是 3 或 5，其中本系统中默认的为 5，也就是 X-Windows 多用户模式。

```
# Default runlevel. The runlevels used by RHS are:
#   0 - halt (Do NOT set initdefault to this)
#   1 - Single user mode
#   2 - Multiuser, without NFS (The same as 3, if you do not have networking)
#   3 - Full multiuser mode
#   4 - unused
#   5 - X11
#   6 - reboot (Do NOT set initdefault to this)
#
id:5:initdefault:
```

2．执行内容/etc/rc.d/rc.sysinit

在确定了登录模式之后，就要开始将 Linux 的主机信息读入 Linux 系统，其内容就是文件“/etc/rc.d/rc.sysinit”中的。查看此文件可以看出，在这里确定了默认路径、主机名称、“/etc/sysconfig/network”中所记录的网络信息等。

```
# System initialization.
si::sysinit:/etc/rc.d/rc.sysinit
```

3．启动内核的外挂模块及各运行级的脚本

在此，主要是选择模块的型态以进行驱动程序的加载。接下来会根据不同的运行级（run level）加载不同的模块，启动系统服务。

```
l0:0:wait:/etc/rc.d/rc 0
l1:1:wait:/etc/rc.d/rc 1
```

```
l2:2:wait:/etc/rc.d/rc 2
l3:3:wait:/etc/rc.d/rc 3
l4:4:wait:/etc/rc.d/rc 4
l5:5:wait:/etc/rc.d/rc 5
l6:6:wait:/etc/rc.d/rc 6

# Trap CTRL-ALT-DELETE
ca::ctrlaltdel:/sbin/shutdown -t3 -r now

# When our UPS tells us power has failed, assume we have a few minutes
# of power left.  Schedule a shutdown for 2 minutes from now.
# This does, of course, assume you have powerd installed and your
# UPS connected and working correctly.
pf::powerfail:/sbin/shutdown -f -h +2 "Power Failure; System Shutting Down"

# If power was restored before the shutdown kicked in, cancel it.
pr:12345:powerokwait:/sbin/shutdown -c "Power Restored; Shutdown Cancelled"
# Run gettys in standard runlevels
1:2345:respawn:/sbin/mingetty tty1
2:2345:respawn:/sbin/mingetty tty2
3:2345:respawn:/sbin/mingetty tty3
4:2345:respawn:/sbin/mingetty tty4
5:2345:respawn:/sbin/mingetty tty5
6:2345:respawn:/sbin/mingetty tty6
# Run xdm in runlevel 5
x:5:respawn:/etc/X11/prefdm -nodaemon
```

2.3 Linux 系统服务

INIT 进程的一个重要作用就是启动 Linux 系统服务（也就是运行在后台的守护进程）。Linux 的系统服务包括两种，第一种是独立运行的系统服务，它们常驻内存中，自开机后一直启动着（如 httpd），具有很快的响应速度；第二种是由 xinet 设定的服务。xinet 能够同时监听多个指定的端口，在接受用户请求时，它能够根据用户请求的端口不同，启动不同的网络服务进程来处理这些用户请求。因此，可以把 xinetd 看作一个启动服务的管理服务器，它决定把一个客户请求交给那个程序处理，然后启动相应的守护进程。以下来分别介绍这两种系统服务。

2.3.1 独立运行的服务

独立运行的系统服务的启动脚本都放在目录“/etc/rc.d/init.d/”中。如笔者系统中的系统服务的启动脚本有（关于 Linux 中服务的具体含义见附录）：

```
[root@sunq init.d]# ls /etc/rc.d/init.d
acpid dc_client iptables named pand rpcsvcgssd tux
anacron dc_server irda netdump pcmcia saslauthd vncserver
apmd diskdump irqbalance netfs portmap sendmail vsftpd
arptables_jf dovecot isdn netplugd psacct single watchquagga
atd dund killall network rawdevices smartd winbind
autofs firstboot kudzu NetworkManager readahead smb xfs
…
```

为了指定特定运行级别服务的开启或关闭，系统的各个不同运行级别都有不同的脚本文件，其目录为“/etc/rc.d/rcN.d”，其中的 N 分别对应不用的运行级别。读者可以进入到各个不同的运行级别目录里查看相应服务的开启或关闭状态，如进入“/rc3.d”目录中的文件如下所示：

```
[root@sunq rc3.d]# ls /etc/rc.d/rc3.d
K02NetworkManager K35winbind K89netplugd S10networ S28autofs S95anacron
K05saslauthd K36lisa K90bluetooth S12syslog S40smartd S95atd
K10dc_server K45named K94diskdump S13irqbalance S44acpid S97messagebus
K10psacct K50netdump K99microcode_ctl S13portmap S55cups S97rhnsd
…
```

可以看到，每个对应的服务都以“K”或“S”开头，其中的 K 代表关闭（kill），其中的 S 代表启动（start），用户可以使用命令“+start|stop|status|restart”来对相应的服务进行操作。

在执行完相应的 rcN.d 目录下的脚本文件后，INIT 最后会执行 rc.local 来启动本地服务，因此，用户若想把某些非系统服务设置为自启动，可以编辑 rc.local 脚本文件，加上相应的执行语句即可。

另外，读者还可以使用命令“service+系统服务+操作”来方便地实现相应服务的操作，如下所示：

```
[root@sunq xinetd.d]# service xinetd restart
停止 xinetd:                                               [  确定  ]
开启 xinetd:                                               [  确定  ]
```

2.3.2 xinetd 设定的服务

xinetd 管理系统中不经常使用的服务，这些服务程序只有在有请求时才由 xinetd 服务负责启动，一旦运行完毕服务自动结束。xinetd 的配置文件为“/etc/xinetd.conf”，它对 xinet 的

默认参数进行了配置：

```
#
# Simple configuration file for xinetd
#
# Some defaults, and include /etc/xinetd.d/
defaults
{
      instances                 = 60
      log_type                  = SYSLOG authpriv
      log_on_success        = HOST PID
      log_on_failure        = HOST
      cps               = 25 30
}
includedir /etc/xinetd.d
```

从该配置文件的最后一行可以看出，xinetd 启动“/etc/xinetd.d”为其配置文件目录。再在对应的配置文件目录中可以看到每一个服务的基本配置，如 tftp 服务的配置脚本文件为：

```
service tftp
{
      socket_type   = dgram//数据包格式
      protocol      = udp//使用 UDP 传输
      wait          = yes
      user          = root
      server        = /usr/sbin/in.tftpd
      server_args   = -s /tftpboot
      disable       = yes//不启动
      per_source    = 11
      cps           = 100 2
      flags         = IPv4
}
```

2.3.3 设定服务命令常用方法

设定系统服务除了在本节中提到的使用 service 之外，chkconfig 也是一个很好的工具，它能够为不同的系统级别设置不同的服务。

常用格式

（1）chkconfig –list（注意在 list 前有两个小连线）：查看系统服务设定。

示例：

```
[root@sunq xinetd.d]# chkconfig --list
```

```
sendmail          0:关闭  1:关闭  2:打开  3:打开  4:打开  5:打开  6:关闭
snmptrapd         0:关闭  1:关闭  2:关闭  3:关闭  4:关闭  5:关闭  6:关闭
gpm               0:关闭  1:关闭  2:打开  3:打开  4:打开  5:打开  6:关闭
syslog           0:关闭  1:关闭  2:打开  3:打开  4:打开  5:打开  6:关闭
…
```

（2）chkconfig--level N [服务名称] 指定状态：对指定级别指定系统服务。

```
[root@sunq xinetd.d]# chkconfig --list|grep ntpd
ntpd              0:关闭  1:关闭  2关闭  3:关闭  4:关闭  5:关闭  6:关闭
[root@sunq ~]# chkconfig --level 3 ntpd on
[root@sunq ~]# chkconfig --list|grep ntpd
ntpd              0:关闭  1:关闭  2:关闭  3:打开  4:关闭  5:关闭  6:关闭
```

另外，在 2.1.1 节系统命令列表中指出的 setup 程序中也就可以设定，而且是图形界面，操作较为方便，读者可以自行尝试。

2.4 实验内容

2.4.1 在 Linux 下解压常见软件

1．实验目的

通过在 Linux 下安装一个完整的软件（嵌入式 Linux 的必备工具——交叉编译工具），掌握 Linux 常见命令，学会设置环境变量，并同时搭建起了嵌入式 Linux 的交叉编译环境（关于交叉编译的具体概念在本书后面会详细讲解），为今后的实验打下良好的基础。

2．实验内容

在 Linux 中解压 cross-3.3.2.tar.bz2，并添加到系统环境变量中去。

3．实验步骤

（1）在光盘中的 cross-3.3.2.tar.bz2 拷贝到 Windows 下的任意盘中。

（2）重启机器转到 Linux 下，并用普通用户身份登录。

（3）打开“终端”，并切换到超级用户模式下。

命令为：su - root

（4）查看 cross-3.3.2.tar.bz2 所在的 Windows 下对应分区的格式，并记下其文件设备名称，如“/dev/hda1”等。

命令为：fdisk -l

（5）使用 mkdir 命令在“/mnt”新建子目录作为挂载点。

命令为：mkdir /mnt/windows

（6）挂载 Windows 相应分区。

若是 vfat 格式，则命令为：mount –t vfat /dev/had* /mnt/windows

注意 由于 ntfs 格式在 Linux 下是不安全的，只能读，不能写，因此最好把文件放到 fat32 格式的文件系统中。

（7）进入挂载目录下，查看是否确实挂载上。

命令为：cd /mnt/windows；ls

（8）在/usr/local 下建一名为 arm 的目录。

命令为：mkdir /usr/local/arm

（9）将 cross-3.3.2.tar.bz2 复制到刚刚新建的目录中。

命令为：cp /mnt/windows/cross-3.3.2.tar.bz2 /usr/local/arm

注意 若 cross-3.3.2.tar.bz2 在当前目录中，则可将命令简写为：cp ./cross-3.3.2.tar.bz2 /usr/local/arm

（10）将当前工作目录转到“/usr/local/arm”下。

命令为：cd /usr/local/arm

想一想 为什么要将此目录创建在“/usr/local”下？

（11）解压缩该软件包。

命令为：tar –jxvf cross-3.3.2.tar.bz2

（12）将此目录下的/bin 目录添加到环境变量中去。

命令为：export PATH=/usr/local/arm/3.3.2/bin :$PATH

注意 用此方法添加的环境变量在掉电后会丢失，因此，可以使用在“/etc/bashrc”的最后一行添加以上命令。

（13）查看该路径是否已添加到环境变量中。

命令为：echo $PATH

4．实验结果

成功搭建了嵌入式 Linux 的交叉编译环境，熟悉 Linux 下常用命令，如“su”、“mkdir”、“mount”、“cp”、“tar”等，并学会添加环境变量，同时也对 Linux 的目录结构有了更深一步的理解。

2.4.2 定制 Linux 系统服务

1．实验目的

通过定制 Linux 系统服务，进一步理解 Linux 的守护进程，能够更加熟练运用 Linux 操作基本命令，同时也加深对 INIT 进程的了解和掌握。

2．实验内容

查看 Linux 系统服务，并定制其系统服务。

3．实验步骤

（1）查看系统的默认运行级别。

命令为：cat /etc/inittab（设其结果为 N）

（2）进入到相应级别的服务脚本目录，查看哪些服务是系统启动的独立运行的服务，并做下记录。

命令为：cd /etc/rc.d/rcN.d

（3）利用命令查看系统开机自启动服务，并与上次查看结果进行比较，找出其中的区别，并思考其中的原因。

命令为：chkconfig –list

（4）记录 chkconfig –list 命令中由 xinet 管理的服务，并将其中启动的服务做下记录。

（5）进入 xinet 配置管理的相应目录，查看是否于 chkconfig –list 所得结果相吻合并查看相应脚本文件。

命令为：cd /etc/xinetd.d

（6）将 sshd 服务停止。

命令为：service sshd stop

（7）将 sshd 服务设置为开机不启动。

命令为：chkconfig –level N sshd stop

（8）查看该设置是否生效。

命令为：chkconfig –list

（9）查看系统中所有服务及其端口号列表。

命令为：cat /etc/services

（10）将 sshd 服务端口改为 4022。

命令为：vi /etc/services；转到插入模式并修改其端口号

（11）重启 sshd 服务，验证所改的端口号是否生效。

命令为：service sshd start

（12）重启 Linux 系统，验证所改的服务开机启动是否生效。

4．实验结果分析

首先，该实验通过验证 Linux 系统服务的启动状态，进一步明确了 Linux 系统服务启动的流程，更深一步地理解了 Linux 系统操作。

另外，实验还通过定制 Linux 系统服务 sshd 的开机启动状态和端口号，熟悉了 Linux 的系统定制步骤。

本 章 小 结

本章首先讲解了 Linux 操作的基本命令，这些命令是使用 Linux 的基础。Linux 基础命令包括用户系统相关命令、文件目录相关命令、压缩打包相关命令、比较合并相关命令以及网络相关命令。在本书中，笔者着重介绍了每一类命令中有代表性的重要命令，并给出了具体实例加以讲解，对其他命令列出了其使用方法。希望读者能举一反三，灵活应用。

接下来，本章讲解了 Linux 启动过程，这部分的内容比较难，但对深入理解 Linux 是非常有帮助的，希望读者能反复阅读。

最后，本章还讲解了 Linux 系统服务，包括独立运行的服务和 xinetd 设定的服务，并且讲解了 Linux 中设定服务的常用方法。

本章安排了两个实验，实验一通过一个完整的操作使读者能够熟练使用 Linux，实验二讲解了如何定制 Linux 系统服务，希望读者能够认真动手实践。

思考与练习

1．更改目录的名称，如把/home/sunq 变为/home/kang。
2．若有一文件属性为-rwxr-xrw-，指出其代表什么意思？
3．如何将文件属性变为-rwxrw-r--？
4．下载最新 Linux 源码，并解开至/usr/src 目录下。
5．修改 TELNET、FTP 服务的端口号。

“黑色经典”系列之《嵌入式 Linux 应用程序开发详解》

第 3 章 Linux 下的 C 编程基础

本章目标

在熟悉了 Linux 常见命令，能够在 Linux 中熟练操作之后，本章将带领读者学习在 Linux 中进行 C 语言编程的基本技能。学习了本章后，读者能够掌握如下内容。

- 熟悉 Linux 系统下的开发环境
- 熟悉 Vi 的基本操作
- 熟练 Emacs 的基本操作
- 熟悉 Gcc 编译器的基本原理
- 熟练使用 Gcc 编译器的常用选项
- 熟练使用 Gdb 调试技术
- 熟悉 Makefile 基本原理及语法规范
- 熟练使用 autoconf 和 automake 来生成 Makefile

3.1 Linux 下 C 语言编程概述

3.1.1 C 语言简单回顾

C 语言最早是由贝尔实验室的 Dennis Ritchie 为了 UNIX 的辅助开发而编写的，它是在 B 语言的基础上开发出来的。尽管 C 语言不是专门针对 UNIX 操作系统或机器编写的，但它与 UNIX 系统的关系十分紧密。由于它的硬件无关性和可移植性，使 C 语言逐渐成为世界上使用最广泛计算机语言。

为了进一步规范 C 语言的硬件无关性，1987 年，美国国家标准协会（ANSI）根据 C 语言问世以来各种版本对 C 语言的发展和扩充，制定了新的标准，称为 ANSI C。ANSI C 语言比原来的标准 C 语言有了很大的发展。目前流行的 C 语言编译系统都是以它为基础的。

C 语言的成功并不是偶然的，它强大的功能和它的可移植性让它能在各种硬件平台上游刃自如。总体而言，C 语言有如下特点。

- C 语言是“中级语言”。它把高级语言的基本结构和语句与低级语言的实用性结合起来。C 语言可以像汇编语言一样对位、字节和地址进行操作，而这三者是计算机最基本的工作单元。
- C 语言是结构化的语言。C 语言采用代码及数据分隔，使程序的各个部分除了必要的信息交流外彼此独立。这种结构化方式可使程序层次清晰，便于使用、维护以及调试。C 语言是以函数形式提供给用户的，这些函数可方便地调用，并具有多种循环、条件语句控制程序流向，从而使程序完全结构化。
- C 语言功能齐全。C 语言具有各种各样的数据类型，并引入了指针概念，可使程序效率更高。另外，C 语言也具有强大的图形功能，支持多种显示器和驱动器，而且计算功能、逻辑判断功能也比较强大，可以实现决策目的。
- C 语言可移植性强。C 语言适合多种操作系统，如 DOS、Windows、Linux，也适合多种体系结构，因此尤其适合在嵌入式领域的开发。

3.1.2 Linux 下 C 语言编程环境概述

Linux 下的 C 语言程序设计与在其他环境中的 C 程序设计一样，主要涉及到编辑器、编译链接器、调试器及项目管理工具。现在我们先对这 4 种工具进行简单介绍，后面会对其一一进行讲解。

（1）编辑器

Linux 下的编辑器就如 Windows 下的 word、记事本等一样，完成对所录入文字的编辑功能。Linux 中最常用的编辑器有 Vi（Vim）和 Emacs，

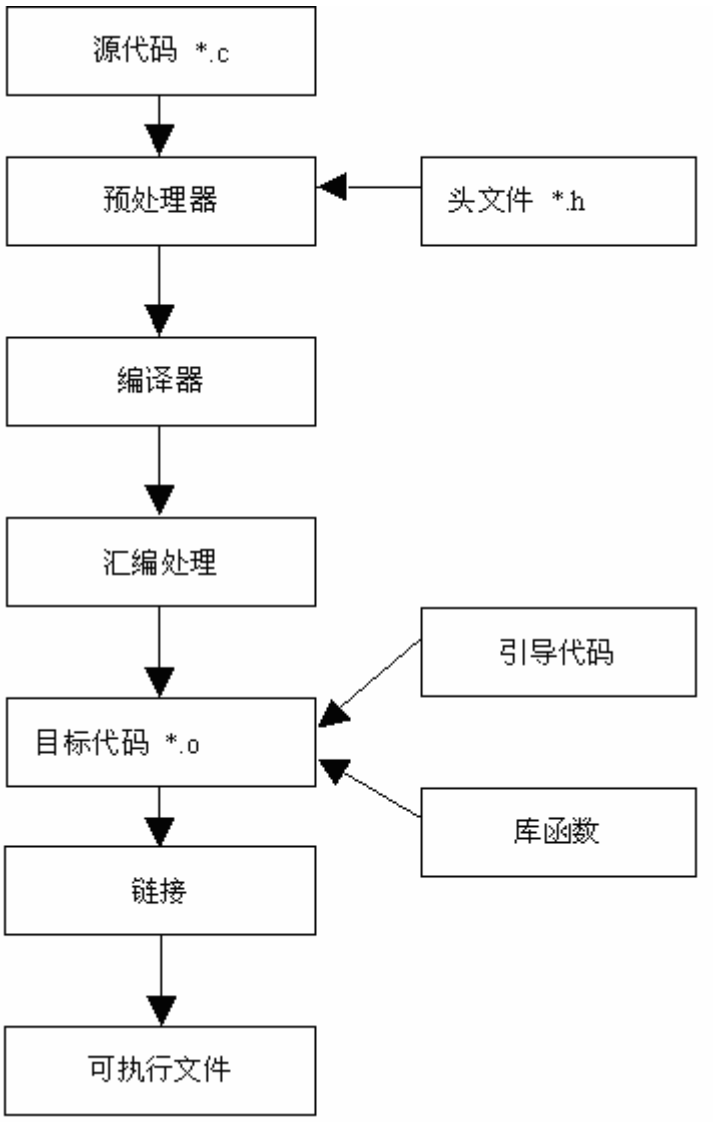


图 3.1　编译过程

它们功能强大，使用方便，广受编程爱好者的喜爱。在本书中，着重介绍 Vi 和 Emacs。

（2）编译链接器

编译是指源代码转化生成可执行代码的过程，它所完成工作主要如图 3.1 所示。

可见，在编译过程是非常复杂的，它包括词法、语法和语义的分析、中间代码的生成和优化、符号表的管理和出错处理等。在 Linux 中，最常用的编译器是 Gcc 编译器。它是 GNU 推出的功能强大、性能优越的多平台编译器，其执行效率与一般的编译器相比平均效率要高 20%～30%，堪称为 GNU 的代表作品之一。

（3）调试器

调试器并不是代码执行的必备工具，而是专为程序员方便调试程序而用的。有编程经验的读者都知道，在编程的过程当中，往往调试所消耗的时间远远大于编写代码的时间。因此，

有一个功能强大、使用方便的调试器是必不可少的。Gdb 是绝大多数 Linux 开发人员所使用的调试器，它可以方便地设置断点、单步跟踪等，足以满足开发人员的需要。

（4）项目管理器

Linux 中的项目管理器“make”有些类似于 Windows 中 Visual C++里的“工程”，它是一种控制编译或者重复编译软件的工具，另外，它还能自动管理软件编译的内容、方式和时机，使程序员能够把精力集中在代码的编写上而不是在源代码的组织上。

3.2 进入 Vi

Linux 系统提供了一个完整的编辑器家族系列，如 Ed、Ex、Vi 和 Emacs 等。按功能它们可以分为两大类：行编辑器（Ed、Ex）和全屏幕编辑器（Vi、Emacs）。行编辑器每次只能对一行进行操作，使用起来很不方便。而全屏幕编辑器可以对整个屏幕进行编辑，用户编辑的文件直接显示在屏幕上，从而克服了行编辑的那种不直观的操作方式，便于用户学习和使用，具有强大的功能。

Vi 是 Linux 系统的第一个全屏幕交互式编辑程序，它从诞生至今一直得到广大用户的青睐，历经数十年仍然是人们主要使用的文本编辑工具，足以见其生命力之强，而强大的生命力是其强大的功能带来的。由于大多数读者在此之前都已经用惯了 Windows 的 word 等编辑器，因此，在刚刚接触时总会或多或少不适应，但只要习惯之后，就能感受到它的方便与快捷。

3.2.1 Vi 的模式

Vi 有 3 种模式，分别为命令行模式、插入模式及命令行模式各模式的功能，下面具体进行介绍。

（1）命令行模式

用户在用 Vi 编辑文件时，最初进入的为一般模式。在该模式中可以通过上下移动光标进行“删除字符”或“整行删除”等操作，也可以进行“复制”、“粘贴”等操作，但无法编辑文字。

（2）插入模式

只有在该模式下，用户才能进行文字编辑输入，用户课按[ESC]键回到命令行模式。

（3）底行模式

在该模式下，光标位于屏幕的底行。用户可以进行文件保存或退出操作，也可以设置编辑环境，如寻找字符串、列出行号等。

3.2.2 Vi 的基本流程

（1）进入 Vi，即在命令行下键入 Vi hello（文件名）。此时进入的是命令行模式，光标位于屏幕的上方，如图 3.2 所示。

（2）在命令行模式下键入 i 进入到插入模式，如图 3.3 所示。可以看出，在屏幕底部显示有“插入”表示插入模式，在该模式下可以输入文字信息。

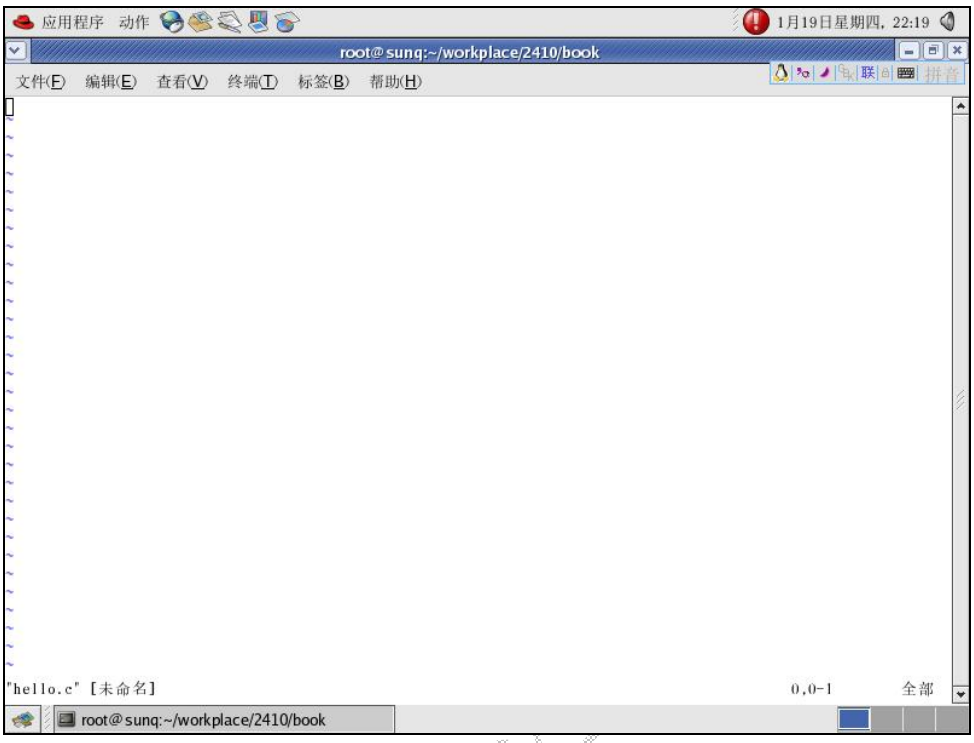


图 3.2　进入 Vi 命令行模式

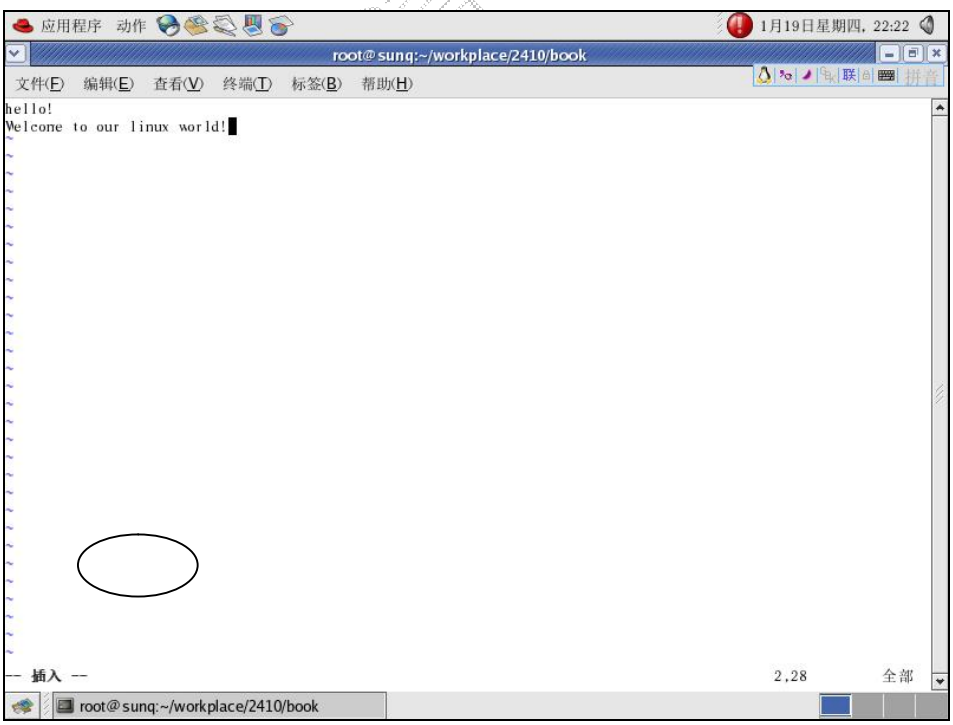


图 3.3　进入 Vi 插入模式

（3）最后，在插入模式中，输入“Esc”，则当前模式转入命令行模式，并在底行行中输入“:wq”（存盘退出）进入底行模式，如图 3.4 所示。

这样，就完成了一个简单的 Vi 操作流程：命令行模式→插入模式→底行模式。由于 Vi 在不同的模式下有不同的操作功能，因此，读者一定要时刻注意屏幕最下方的提示，分清所在的模式。

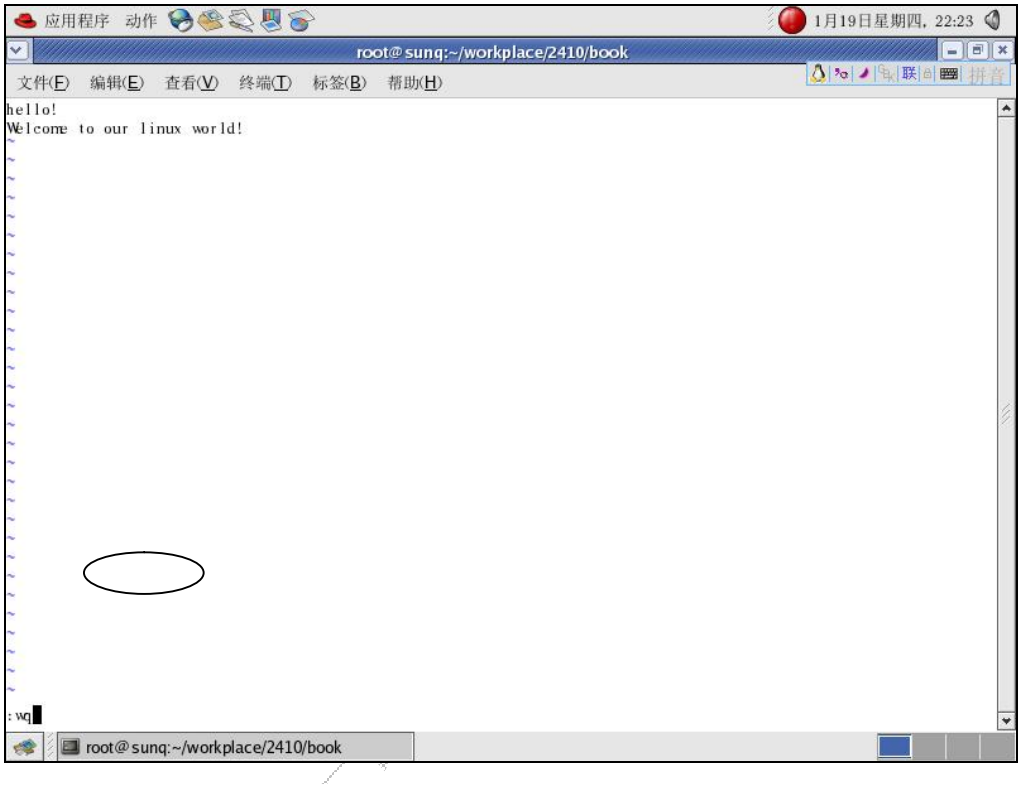


图 3.4　进入 Vi 底行模式

3.2.3　Vi 的各模式功能键

（1）命令行模式常见功能键如表 3.1 所示。

表 3.1　　Vi 命令行模式功能键

目　　录	目 录 内 容
I	切换到插入模式，此时光标当于开始输入文件处
A	切换到插入模式，并从目前光标所在位置的下一个位置开始输入文字
O	切换到插入模式，且从行首开始插入新的一行
[ctrl]+[b]	屏幕往“后”翻动一页
[ctrl]+[f]	屏幕往“前”翻动一页
[ctrl]+[u]	屏幕往“后”翻动半页
[ctrl]+[d]	屏幕往“前”翻动半页

0（数字 0）	光标移到本行的开头
G	光标移动到文章的最后
nG	光标移动到第 n 行
$	移动到光标所在行的“行尾”
n<Enter>	光标向下移动 n 行
/name	在光标之后查找一个名为 name 的字符串
?name	在光标之前查找一个名为 name 的字符串
X	删除光标所在位置的“后面”一个字符

续表

目　　录	目 录 内 容
X	删除光标所在位置的“前面”一个字符
dd	删除光标所在行
ndd	从光标所在行开始向下删除 n 行
yy	复制光标所在行
nyy	复制光标所在行开始的向下 n 行
p	将缓冲区内的字符粘贴到光标所在位置（与 yy 搭配）
U	恢复前一个动作

（2）插入模式的功能键只有一个，也就是 Esc 退出到命令行模式。

（3）底行模式常见功能键如表 3.2 所示。

表 3.2　　Vi 底行模式功能键

目　　录	目 录 内 容
:w	将编辑的文件保存到磁盘中
:q	退出 Vi（系统对做过修改的文件会给出提示）
:q!	强制退出 Vi（对修改过的文件不作保存）
:wq	存盘后退出
:w [filename]	另存一个命为 filename 的文件
:set nu	显示行号，设定之后，会在每一行的前面显示对应行号
:set nonu	取消行号显示

注意　Vi 的升级版 Vim 已经问世了，功能相当强大，且保持与 Vi 的 90%相兼容，因此，感兴趣的读者可以查看相关资料进行学习。

3.3 初探 Emacs

正如前面所述，Vi 是一款功能非常强大的编辑器，它能够方便、快捷、高效地完成用户的任务，那么，在此再次向读者介绍另一款编辑器是否多此一举呢？答案是否定的。因为 Emacs 不仅仅是一款功能强大的编译器，而且是一款融合编辑、编译、调试于一体的开发环境。虽然，它没有 Visual Sdiao 一样绚丽的界面，但是它可以在没有图形显示的终端环境下出色的工作，相信追求强大功能和工作效率的任务并不会介意它朴素的界面的。

Emacs 的使用和 Vi 截然不同。在 Emacs 里，没有类似于 Vi 的 3 种“模式”。Emacs 只有一种模式，也就是编辑模式，而且它的命令全靠功能键完成。因此，功能键也就相当重要了。

但 Emacs 却还使用一个不同 Vi 的“模式”，它的“模式”是指各种辅助环境。比如，当编辑普通文本时，使用的是“文本模式（Txt Mode）”，而当他们写程序时，使用的则是如“c 模式”、“Shell 模式”等。

下面，首先来介绍一下 Emacs 中作为编辑器的使用方法，以帮助读者熟悉 Emacs 的环境。

注释 Emacs 缩写注释：
C-<chr>表示按住 Ctrl 键的同时键入字符<chr>。因此，C-f 就表示按住 Ctrl 键同时键入 f。
M-<chr>表示当键入字符<chr>时同时按住 Meta 或 Edit 或 Alt 键（通常为 Alt 键）。

3.3.1 Emacs 的基本操作

1．Emacs 安装

现在较新版本的 Linux（如本书中所用的 Red Hat Enterprise 4 AS）的安装光盘中一般都自带有 Emacs 的安装包，用户可以通过安装光盘进行安装（一般在第 2 张光盘中）。

2．启动 Emacs

安装完 Emacs 之后，只需在命令行键入“emacs [文件名]”（若缺省文件名，也可在 emacs 编辑文件后另存时指定），也可从“编程”→“emacs”打开，3.5 图中所示的就是从“编程”→“emacs”打开的 Emacs 欢迎界面。

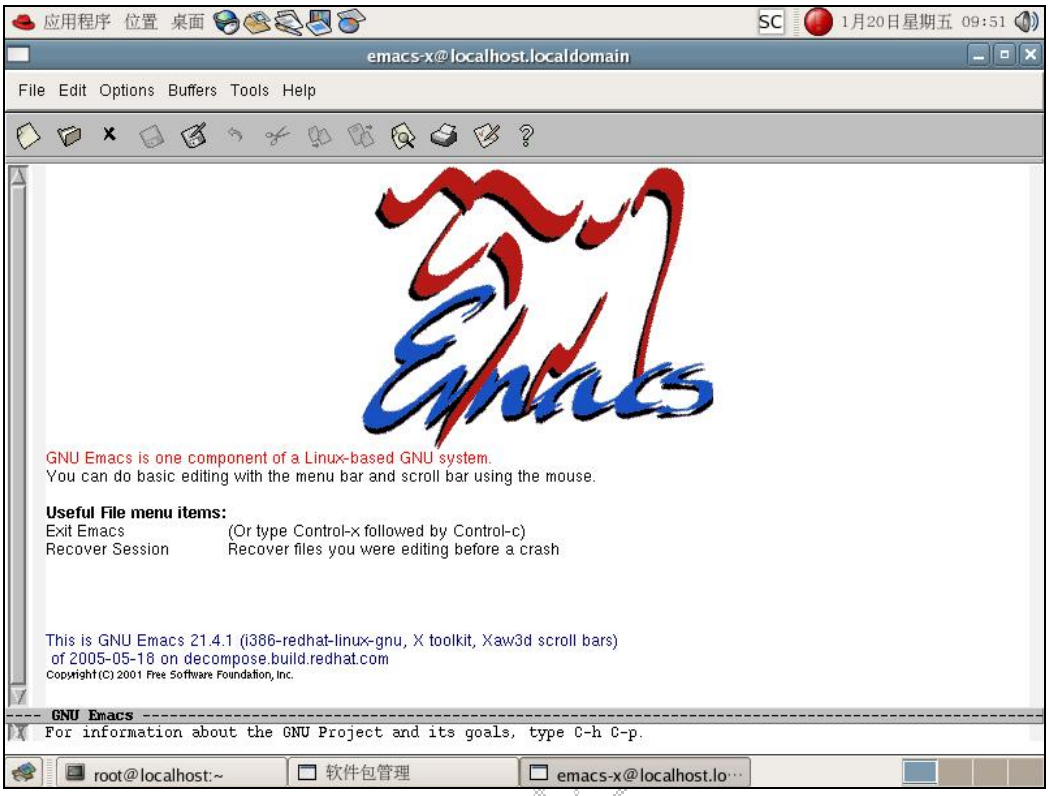


图 3.5　Emacs 欢迎界面

接着可单击任意键进入 Emacs 的工作窗口，如图 3.6 所示。

从图中可见，Emacs 的工作窗口分为上下两个部分，上部为编辑窗口，底部为命令显示窗口，用户执行功能键的功能都会在底部有相应的显示，有时也需要用户在底部窗口输入相应的命令，如查找字符串等。

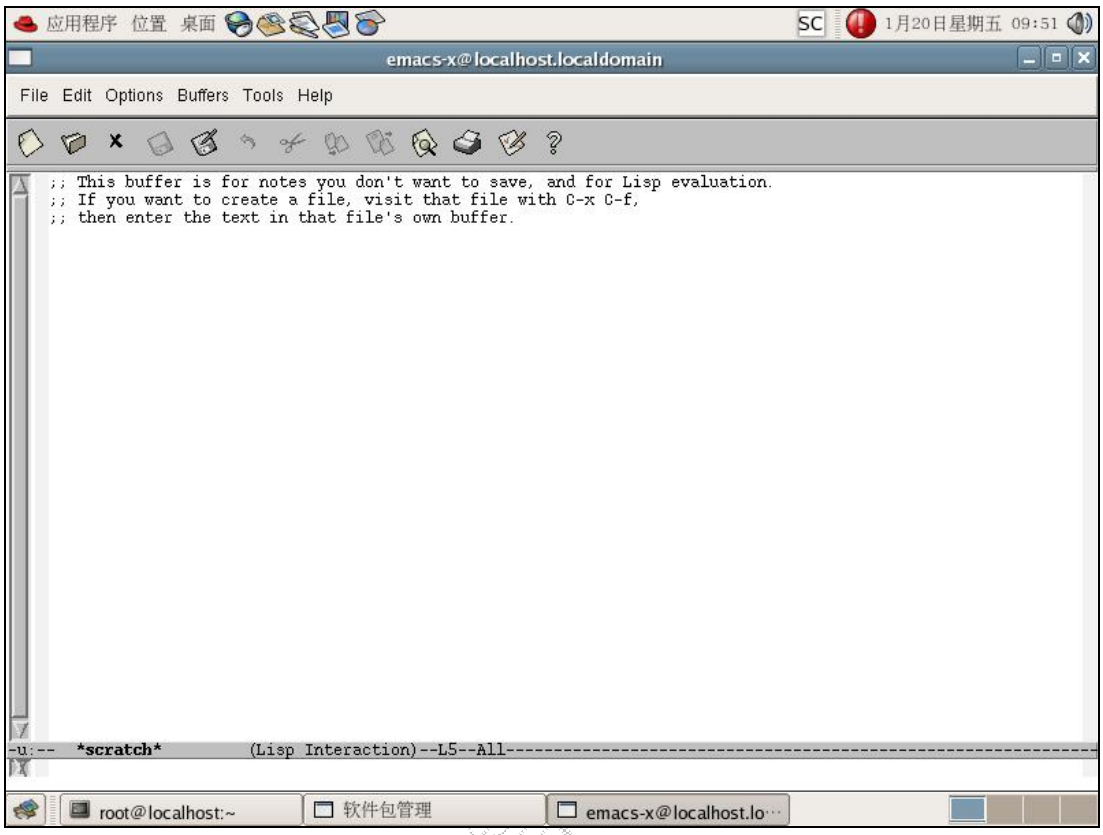


图 3.6 Emacs 的工作窗口

3．进入 Emacs

在进入 Emacs 后，即可进行文件的编辑。由于 Emacs 只有一种编辑模式，因此用户无需进行模式间的切换。下面介绍 Emacs 中基本编辑功能键。

（1）移动光标

虽然在 Emacs 中可以使用“上”、“下”、“左”、“右”方向键来移动单个字符，但笔者还是建议读者学习其对应功能键，因为它们不仅能在所有类型的终端上工作，而且读者将会发现在熟练使用之后，输入这些 Ctrl 加字符会比按方向键快很多。下表 3.3 列举了 Emacs 中光标移动的常见功能键。

表 3.3 Emacs 光标移动功能键

目　　录	目 录 内 容	目　　录	目 录 内 容
C-f	向前移动一个字符	M-b	向后移动一个单词
C-b	向后移动一个字符	C-a	移动到行首
C-p	移动到上一行	C-e	移动到行尾
C-n	移动到下一行	M-<（M 加“小于号”）	移动光标到整个文本的开头

M-f	向前移动一个单词	M->（M 加“大于号”）	移动光标到整个文本的末尾

（2）剪切和粘贴

在 Emacs 中可以使用“Delete”和“BackSpace”删除光标前后的字符，这和用户之前的习惯一致，在此就不赘述。以词和行为单位的剪切和粘贴功能键如表 3.4 所示。

表 3.4　　Emacs 剪切和粘贴

目　录	目 录 内 容	目　录	目 录 内 容
M-Delete	剪切光标前面的单词	M-k	剪切从光标位置到句尾的内容
M-d	剪切光标前面的单词	C-y	将缓冲区中的内容粘贴到光标所在的位置
C-k	剪切从光标位置到行尾的内容	C-x u	撤销操作（先操作 C-x，接着再单击 u）

注意　在 Emacs 中对单个字符的操作是“删除”，而对词和句的操作是“剪切”，即保存在缓冲区中，以备后面的“粘贴”所用。

（3）复制文本

在 Emacs 中的复制文本包括两步：选择复制区域和粘贴文本。

选择复制区域的方法是：首先在复制起始点（A）按下“C-Spase”或“C-@(C-Shift-2)”使它成为一个表示点，再将光标移至复制结束电（B），再按下“M-w”，就可将 A 与 B 之间的文本复制到系统的缓冲区中。在使用功能键 C-y 将其粘贴到指定位置。

（4）查找文本

查找文本的功能键如表 3.5 所示。

表 3.5　　Emacs 查找文本功能键

目　录	目 录 内 容
C-s	查找光标以后的内容，并在对话框的“I-search:”后输入查找字符串
C-r	查找光标以前的内容，并在对话框的“I-search backward:”后输入查找字符串

（5）保存文档

在 Emacs 中保存文档的功能键为“C-x C-s”（即先操作 C-x，接着再操作 C-s），这时，屏幕底下的对话框会出现如“Wrote /root/workplace/editor/why”字样，如图 3.7 所示。

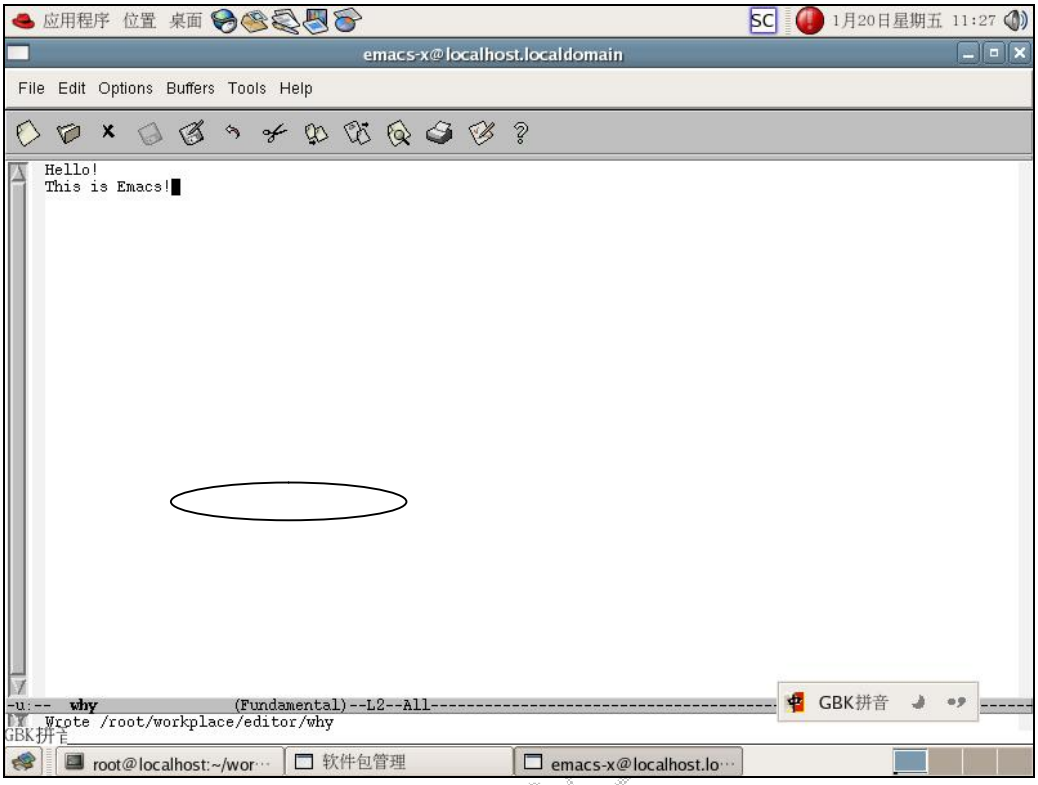


图 3.7 Emacs 中保存文档

另外，Emacs 在编辑时会为每个文件提供“自动保存（auto save）”的机制，而且自动保存的文件的文件名前后都有一个“#”，例如，编辑名为“hello.c”的文件，其自动保存的文件的文件名就叫“#hello.c#”。当用户正常的保存了文件后，Emacs 就会删除这个自动保存的文件。这个机制当系统发生异常时非常有用。

（6）退出文档

在 Emacs 中退出文档的功能键为“C-x C-c”。

3.3.2 Emacs 的编译概述

正如本节前面所提到的，Emacs 不仅仅是个强大的编译器，它还是一个集编译、调试等于一体的工作环境。在这里，读者将会了解到 Emacs 作为编译器的最基本的概念，感兴趣的读者可以参考《Learning GNU Emacs，Second Edition》一书进一步深入学习 Emacs。

1．Emacs 中的模式

正如本节前面提到的，在 Emacs 中并没有像 Vi 中那样的“命令行”、“编辑”模式，只有一种编辑模式。这里所说的“模式”，是指 Emacs 里的各种辅助环境。下面就着重了解一下 C 模式。

当我们启动某一文件时，Emacs 会判断文件的类型，从而自动选择相应的模式。当然，用户也可以手动启动各种模式，用功能键“M-x”，然后再输入模式的名称，如图所示 3.8 所示就启动了“C 模式”。

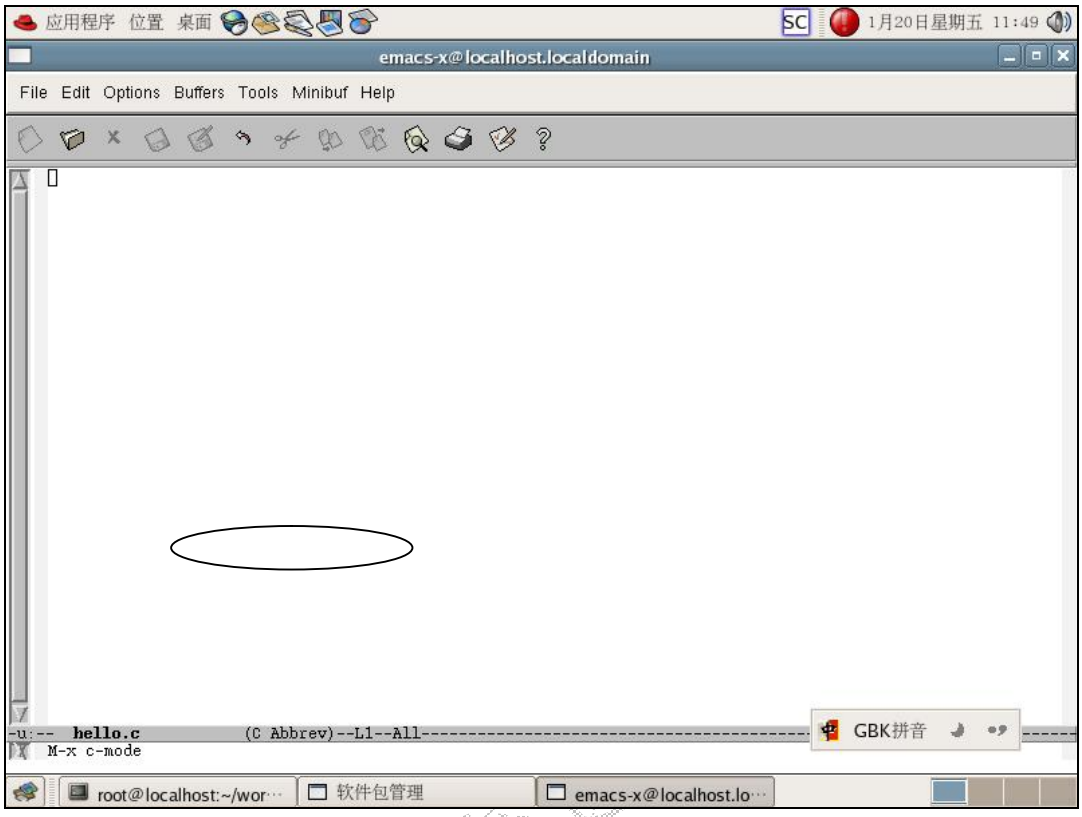


图 3.8　Emacs 中选择模式

在强大的 C 模式下，用户拥有“自动缩进”、“注释”、“预处理扩展”、“自动状态”等强大功能。在“C 模式”下编辑代码时，可以用“Tab”键自动的将当前行的代码产生适当的缩进，使代码结构清晰、美观，它也可以指定缩进的规则。

源代码要有良好可读性，必须要有良好的注释。在 Emacs 中，用“M-”可以产生一条右缩进的注释。C 模式下是“/* comments */”形式的注释，C++模式下是“// comments”形式的注释。当用户高亮选定某段文本，然后操作“C-c C-c”，就可以注释该段文字。

Emacs 还可以使用 C 预处理其运行代码的一部分，以便让程序员检测宏、条件编译以及 include 语句的效果。

2．Emacs 编译调试程序

Emacs 可以让程序员在 Emacs 环境里编译自己的软件。此时，编辑器把编译器的输出和程序代码连接起来。程序员可以像在 Windows 的其他开发工具一样，将出错位置和代码定位联系起来。

Emacs 默认的编辑命令是对一个 make（在本章 3.6 节中会详细介绍）的调用。用户可以打开“tool”下的“Compile”进行查看。Emacs 可以支持大量的工程项目，以方便程序员的开发。

另外，Emacs 为 Gdb 调试器提供了一个功能齐全的接口。在 Emacs 中使用 Gdb 的时候，程序员不仅能够获得 Gdb 用其他任何方式运行时所具有的全部标准特性，还可以通过接口增强而获得的其他性能。

3.4 Gcc 编译器

GNU CC（简称为 Gcc）是 GNU 项目中符合 ANSI C 标准的编译系统，能够编译用 C、C++和 Object C 等语言编写的程序。Gcc 不仅功能强大，而且可以编译如 C、C++、Object C、Java、Fortran、Pascal、Modula-3 和 Ada 等多种语言，而且 Gcc 又是一个交叉平台编译器，它能够在当前 CPU 平台上为多种不同体系结构的硬件平台开发软件，因此尤其适合在嵌入式领域的开发编译。本章中的示例，除非特别注明，否则均采用 Gcc 版本为 4.0.0。

下表 3.6 是 Gcc 支持编译源文件的后缀及其解释。

表 3.6 Gcc 所支持后缀名解释

后 缀 名	所对应的语言	后 缀 名	所对应的语言
.c	C 原始程序	.s/.S	汇编语言原始程序
.C/.cc/.cxx	C++原始程序	.h	预处理文件（头文件）
.m	Objective-C 原始程序	.o	目标文件
.i	已经过预处理的 C 原始程序	.a/.so	编译后的库文件
.ii	已经过预处理的 C++原始程序		

3.4.1 Gcc 编译流程解析

如本章开头提到的，Gcc 的编译流程分为了 4 个步骤，分别为：

- 预处理（Pre-Processing）；
- 编译（Compiling）；
- 汇编（Assembling）；
- 链接（Linking）。

下面就具体来查看一下 Gcc 是如何完成 4 个步骤的。

首先，有以下 hello.c 源代码：

```
#include<stdio.h>
int main()
{
     printf("Hello! This is our embedded world!\n");
     return 0;
}
```

（1）预处理阶段

在该阶段，编译器将上述代码中的 stdio.h 编译进来，并且用户可以使用 Gcc 的选项“-E”进行查看，该选项的作用是让 Gcc 在预处理结束后停止编译过程。

注意 Gcc 指令的一般格式为：Gcc [选项] 要编译的文件 [选项][目标文件]
其中，目标文件可缺省，Gcc 默认生成可执行的文件，命为：编译文件.out

```
[root@localhost Gcc]# Gcc –E hello.c –o hello.i
```

在此处，选项“-o”是指目标文件，由表 3.6 可知，“.i”文件为已经过预处理的 C 原始程序。以下列出了 hello.i 文件的部分内容：

```
typedef int (*__gconv_trans_fct) (struct __gconv_step *,
        struct __gconv_step_data *, void *,
        __const unsigned char *,
        __const unsigned char **,
        __const unsigned char *, unsigned char **,
        size_t *);
…
# 2 "hello.c" 2
int main()
{
 printf("Hello! This is our embedded world!\n");
 return 0;
}
```

由此可见，Gcc 确实进行了预处理，它把“stdio.h”的内容插入到 hello.i 文件中。

（2）编译阶段

接下来进行的是编译阶段，在这个阶段中，Gcc 首先要检查代码的规范性、是否有语法错误等，以确定代码的实际要做的工作，在检查无误后，Gcc 把代码翻译成汇编语言。用户可以使用“-S”选项来进行查看，该选项只进行编译而不进行汇编，生成汇编代码。

```
[root@localhost Gcc]# Gcc –S hello.i –o hello.s
```

以下列出了 hello.s 的内容，可见 Gcc 已经将其转化为汇编了，感兴趣的读者可以分析一下这一行简单的 C 语言小程序是如何用汇编代码实现的。

```
    .file   "hello.c"
    .section    .rodata
    .align 4
.LC0:
    .string    "Hello! This is our embedded world!"
    .text
.globl main
    .type main, @function
main:
    pushl %ebp
    movl %esp, %ebp
    subl $8, %esp
```

```
        andl $-16, %esp
        movl $0, %eax
        addl $15, %eax
        addl $15, %eax
        shrl $4, %eax
        sall $4, %eax
        subl %eax, %esp
        subl $12, %esp
        pushl $.LC0
        call puts
        addl $16, %esp
        movl $0, %eax
        leave
        ret
        .size   main, .-main
        .ident  "GCC: (GNU) 4.0.0 20050519 (Red Hat 4.0.0-8)"
        .section    .note.GNU-stack,"",@progbits
```

（3）汇编阶段

汇编阶段是把编译阶段生成的“.s”文件转成目标文件，读者在此可使用选项“-c”就可看到汇编代码已转化为“.o”的二进制目标代码了。如下所示：

```
[root@localhost Gcc]# Gcc -c hello.s -o hello.o
```

（4）链接阶段

在成功编译之后，就进入了链接阶段。在这里涉及到一个重要的概念：函数库。

读者可以重新查看这个小程序，在这个程序中并没有定义“printf”的函数实现，且在预编译中包含进的“stdio.h”中也只有该函数的声明，而没有定义函数的实现，那么，是在哪里实现“printf”函数的呢？最后的答案是：系统把这些函数实现都被做到名为 libc.so.6 的库文件中去了，在没有特别指定时，Gcc 会到系统默认的搜索路径“/usr/lib”下进行查找，也就是链接到 libc.so.6 库函数中去，这样就能实现函数“printf”了，而这也就是链接的作用。

函数库一般分为静态库和动态库两种。静态库是指编译链接时，把库文件的代码全部加入到可执行文件中，因此生成的文件比较大，但在运行时也就不再需要库文件了。其后缀名一般为“.a”。动态库与之相反，在编译链接时并没有把库文件的代码加入到可执行文件中，而是在程序执行时由运行时链接文件加载库，这样可以节省系统的开销。动态库一般后缀名为“.so”，如前面所述的 libc.so.6 就是动态库。Gcc 在编译时默认使用动态库。

完成了链接之后，Gcc 就可以生成可执行文件，如下所示。

```
[root@localhost Gcc]# Gcc hello.o -o hello
```

运行该可执行文件，出现正确的结果如下。

```
[root@localhost Gcc]# ./hello
Hello! This is our embedded world!
```

3.4.2 Gcc 编译选项分析

Gcc 有超过 100 个的可用选项，主要包括总体选项、告警和出错选项、优化选项和体系结构相关选项。以下对每一类中最常用的选项进行讲解。

（1）总体选项

Gcc 的总结选项如表 3.7 所示，很多在前面的示例中已经有所涉及。

表 3.7　　Gcc 总体选项列表

后　缀　名	所对应的语言
-c	只是编译不链接，生成目标文件“.o”
-S	只是编译不汇编，生成汇编代码
-E	只进行预编译，不做其他处理
-g	在可执行程序中包含标准调试信息
-o file	把输出文件输出到 file 里
-v	打印出编译器内部编译各过程的命令行信息和编译器的版本
-I dir	在头文件的搜索路径列表中添加 dir 目录
-L dir	在库文件的搜索路径列表中添加 dir 目录
-static	链接静态库
-llibrary	连接名为 library 的库文件

对于“-c”、“-E”、“-o”、“-S”选项在前一小节中已经讲解了其使用方法，在此主要讲解另外两个非常常用的库依赖选项“-I dir”和“-L dir”。

- “-I dir”

正如上表中所述，“-I dir”选项可以在头文件的搜索路径列表中添加 dir 目录。由于 Linux 中头文件都默认放到了“/usr/include/”目录下，因此，当用户希望添加放置在其他位置的头文件时，就可以通过“-I dir”选项来指定，这样，Gcc 就会到相应的位置查找对应的目录。

比如在“/root/workplace/Gcc”下有两个文件：

```
/*hello1.c*/
#include<my.h>
int main()
{
    printf("Hello!!\n");
    return 0;
}
/*my.h*/
#include<stdio.h>
```

这样，就可在 Gcc 命令行中加入“-I”选项：

```
[root@localhost Gcc] Gcc hello1.c –I /root/workplace/Gcc/ -o hello1
```

这样，Gcc 就能够执行出正确结果。

小知识 在 include 语句中，“<>”表示在标准路径中搜索头文件，““”表示在本目录中搜索。故在上例中，可把 hello1.c 的“#include<my.h>”改为“#include “my.h””，就不需要加上“-I”选项了。

- “-L dir”

选项“-L dir”的功能与“-I dir”类似，能够在库文件的搜索路径列表中添加 dir 目录。例如有程序 hello_sq.c 需要用到目录“/root/workplace/Gcc/lib”下的一个动态库 libsunq.so，则只需键入如下命令即可：

```
[root@localhost Gcc] Gcc hello_sq.c –L /root/workplace/Gcc/lib –lsunq –o hello_sq
```

需要注意的是，“-I dir”和“-L dir”都只是指定了路径，而没有指定文件，因此不能在路径中包含文件名。

另外值得详细解释一下的是“-l”选项，它指示 Gcc 去连接库文件 libsunq.so。由于在 Linux 下的库文件命名时有一个规定：必须以 l、i、b 3 个字母开头。因此在用-l 选项指定链接的库文件名时可以省去 l、i、b 3 个字母。也就是说 Gcc 在对“-lsunq”进行处理时，会自动去链接名为 libsunq.so 的文件。

（2）告警和出错选项

Gcc 的告警和出错选项如表 3.8 所示。

表 3.8 Gcc 总体选项列表

选　项	含　义
-ansi	支持符合 ANSI 标准的 C 程序
-pedantic	允许发出 ANSI C 标准所列的全部警告信息

续表

选　项	含　义
-pedantic-error	允许发出 ANSI C 标准所列的全部错误信息
-w	关闭所有告警
-Wall	允许发出 Gcc 提供的所有有用的报警信息
-werror	把所有的告警信息转化为错误信息，并在告警发生时终止编译过程

下面结合实例对这几个告警和出错选项进行简单的讲解。

如有以下程序段：

```
#include<stdio.h>

void main()
```

```
{
    long long tmp = 1;
    printf("This is a bad code!\n");
    return 0;
}
```

这是一个很糟糕的程序，读者可以考虑一下有哪些问题？

- “-ansi”

该选项强制 Gcc 生成标准语法所要求的告警信息，尽管这还并不能保证所有没有警告的程序都是符合 ANSI C 标准的。运行结果如下所示：

```
[root@localhost Gcc]# Gcc –ansi warning.c –o warning
warning.c: 在函数“main”中:
warning.c:7 警告: 在无返回值的函数中，“return”带返回值
warning.c:4 警告: “main”的返回类型不是“int”
```

可以看出，该选项并没有发现“long long”这个无效数据类型的错误。

- “-pedantic”

允许发出 ANSI C 标准所列的全部警告信息，同样也保证所有没有警告的程序都是符合 ANSI C 标准的。其运行结果如下所示：

```
[root@localhost Gcc]# Gcc –pedantic warning.c –o warning
warning.c: 在函数“main”中:
warning.c:5 警告: ISO C90 不支持“long long”
warning.c:7 警告: 在无返回值的函数中，“return”带返回值
warning.c:4 警告: “main”的返回类型不是“int”
```

可以看出，使用该选项查看出了“long long”这个无效数据类型的错误。

- “-Wall”

允许发出 Gcc 能够提供的所有有用的报警信息。该选项的运行结果如下所示：

```
[root@localhost Gcc]# Gcc –Wall warning.c –o warning
warning.c:4 警告: “main”的返回类型不是“int”
warning.c: 在函数“main”中:
warning.c:7 警告: 在无返回值的函数中，“return”带返回值
warning.c:5 警告: 未使用的变量“tmp”
```

使用“-Wall”选项找出了未使用的变量 tmp，但它并没有找出无效数据类型的错误。

另外，Gcc 还可以利用选项对单独的常见错误分别指定警告，有关具体选项的含义感兴趣的读者可以查看 Gcc 手册进行学习。

（3）优化选项

Gcc 可以对代码进行优化，它通过编译选项“-On”来控制优化代码的生成，其中 n 是一个代表优化级别的整数。对于不同版本的 Gcc 来讲，n 的取值范围及其对应的优化效果可能

并不完全相同，比较典型的范围是从 0 变化到 2 或 3。

不同的优化级别对应不同的优化处理工作。如使用优化选项“-O”主要进行线程跳转（Thread Jump）和延迟退栈（Deferred Stack Pops）两种优化。使用优化选项“-O2”除了完成所有“-O1”级别的优化之外，同时还要进行一些额外的调整工作，如处理器指令调度等。选项“-O3”则还包括循环展开和其他一些与处理器特性相关的优化工作。

虽然优化选项可以加速代码的运行速度，但对于调试而言将是一个很大的挑战。因为代码在经过优化之后，原先在源程序中声明和使用的变量很可能不再使用，控制流也可能会突然跳转到意外的地方，循环语句也有可能因为循环展开而变得到处都有，所有这些对调试来讲都将是一场噩梦。所以笔者建议在调试的时候最好不使用任何优化选项，只有当程序在最终发行的时候才考虑对其进行优化。

（4）体系结构相关选项

Gcc 的体系结构相关选项如表 3.9 所示。

表 3.9　　Gcc 体系结构相关选项列表

选　项	含　义
-mcpu=type	针对不同的 CPU 使用相应的 CPU 指令。可选择的 type 有 i386、i486、pentium 及 i686 等
-mieee-fp	使用 IEEE 标准进行浮点数的比较
-mno-ieee-fp	不使用 IEEE 标准进行浮点数的比较
-msoft-float	输出包含浮点库调用的目标代码
-mshort	把 int 类型作为 16 位处理，相当于 short int
-mrtd	强行将函数参数个数固定的函数用 ret NUM 返回，节省调用函数的一条指令

这些体系结构相关选项在嵌入式的设计中会有较多的应用，读者需根据不同体系结构将对应的选项进行组合处理。在本书后面涉及到具体实例会有针对性的讲解。

3.5　Gdb 调试器

调试是所有程序员都会面临的问题。如何提高程序员的调试效率，更好更快地定位程序中的问题从而加快程序开发的进度，是大家共同面对的。就如读者熟知的 Windows 下的一些调试工具，如 VC 自带的如设置断点、单步跟踪等，都受到了广大用户的赞赏。那么，在 Linux 下有什么很好的调试工具呢？

本文所介绍的 Gdb 调试器是一款 GNU 开发组织并发布的 UNIX/Linux 下的程序调试工具。虽然，它没有图形化的友好界面，但是它强大的功能也足以与微软的 VC 工具等媲美。下面就请跟随笔者一步步学习 Gdb 调试器。

3.5.1　Gdb 使用流程

这里给出了一个短小的程序，由此带领读者熟悉一下 Gdb 的使用流程。建议读者能够实际动手操作。

首先，打开 Linux 下的编辑器 Vi 或者 Emacs，编辑如下代码（由于为了更好地熟悉 Gdb

的操作，笔者在此使用 Vi 编辑，希望读者能够参见 3.3 节中对 Vi 的介绍，并熟练使用 Vi）。

```
/*test.c*/
#include <stdio.h>
int sum(int m);
int main()
{
     int i,n=0;
     sum(50);
     for(i=1; i<=50; i++)
      {
        n += i;
      }
     printf("The sum of 1-50 is %d \n", n );

}
int sum(int m)
{
       int i,n=0;
       for(i=1; i<=m;i++)
         n += i;
       printf("The sum of 1-m is %d\n", n);
}
```

在保存退出后首先使用 Gcc 对 test.c 进行编译，注意一定要加上选项“-g”，这样编译出的可执行代码中才包含调试信息，否则之后 Gdb 无法载入该可执行文件。

```
[root@localhost Gdb]# gcc -g test.c -o test
```

虽然这段程序没有错误，但调试完全正确的程序可以更加了解 Gdb 的使用流程。接下来就启动 Gdb 进行调试。注意，Gdb 进行调试的是可执行文件，而不是如“.c”的源代码，因此，需要先通过 Gcc 编译生成可执行文件才能用 Gdb 进行调试。

```
[root@localhost Gdb]# gdb test
GNU Gdb Red Hat Linux (6.3.0.0-1.21rh)
Copyright 2004 Free Software Foundation, Inc.
GDB is free software, covered by the GNU General Public License, and you are
welcome to change it and/or distribute copies of it under certain conditions.
Type "show copying" to see the conditions.
There is absolutely no warranty for GDB.  Type "show warranty" for details.
This GDB was configured as "i386-redhat-linux-gnu"...Using host libthread_db
library "/lib/libthread_db.so.1".
```

```
(gdb)
```

可以看出，在 Gdb 的启动画面中指出了 Gdb 的版本号、使用的库文件等信息，接下来就进入了由“(gdb)”开头的命令行界面了。

（1）查看文件

在 Gdb 中键入“l”（list）就可以查看所载入的文件，如下所示：

注意 在 Gdb 的命令中都可使用缩略形式的命令，如“l”代便“list”，“b”代表“breakpoint”，“p”代表“print”等，读者也可使用“help”命令查看帮助信息。

```
(Gdb) l
1       #include <stdio.h>
2       int sum(int m);
3       int main()
4       {
5            int i,n=0;
6            sum(50);
7            for(i=1; i<=50; i++)
8            {
9              n += i;
10           }
(Gdb) l
11           printf("The sum of 1~50 is %d \n", n );
12
13      }
14      int sum(int m)
15      {
16               int i,n=0;
17               for(i=1; i<=m;i++)
18                    n += i;
19               printf("The sum of 1~m is = %d\n", n);
20      }
```

可以看出，Gdb 列出的源代码中明确地给出了对应的行号，这样就可以大大地方便代码的定位。

（2）设置断点

设置断点是调试程序中是一个非常重要的手段，它可以使程序到一定位置暂停它的运行。因此，程序员在该位置处可以方便地查看变量的值、堆栈情况等，从而找出代码的症结所在。

在 Gdb 中设置断点非常简单，只需在“b”后加入对应的行号即可（这是最常用的方式，

另外还有其他方式设置断点）。如下所示：

```
(Gdb) b 6
Breakpoint 1 at 0x804846d: file test.c, line 6.
```

要注意的是，在 Gdb 中利用行号设置断点是指代码运行到对应行之前将其停止，如上例中，代码运行到第 5 行之前暂停（并没有运行第 5 行）。

（3）查看断点情况

在设置完断点之后，用户可以键入“info b”来查看设置断点情况，在 Gdb 中可以设置多个断点。

```
(Gdb) info b
Num Type          Disp Enb Address    What
1   breakpoint    keep y   0x0804846d in main at test.c:6
```

（4）运行代码

接下来就可运行代码了，Gdb 默认从首行开始运行代码，可键入“r”（run）即可（若想从程序中指定行开始运行，可在 r 后面加上行号）。

```
(Gdb) r
Starting program: /root/workplace/Gdb/test
Reading symbols from shared object read from target memory...done.
Loaded system supplied DSO at 0x5fb000

Breakpoint 1, main () at test.c:6
6               sum(50);
```

可以看到，程序运行到断点处就停止了。

（5）查看变量值

在程序停止运行之后，程序员所要做的工作是查看断点处的相关变量值。在 Gdb 中只需键入“p”＋变量值即可，如下所示：

```
(Gdb) p n
$1 = 0
(Gdb) p i
$2 = 134518440
```

在此处，为什么变量“i”的值为如此奇怪的一个数字呢？原因就在于程序是在断点设置的对应行之前停止的，那么在此时，并没有把“i”的数值赋为零，而只是一个随机的数字。但变量“n”是在第四行赋值的，故在此时已经为零。


小技巧 Gdb 在显示变量值时都会在对应值之前加上“$N”标记，它是当前变量值的引用标记，所以以后若想再次引用此变量就可以直接写作“$N”，而无需写冗长的变量名。

（6）单步运行

单步运行可以使用命令“n”（next）或“s”（step），它们之间的区别在于：若有函数调用的时候，“s”会进入该函数而“n”不会进入该函数。因此，“s”就类似于 VC 等工具中的“step in”，“n”类似与 VC 等工具中的“step over”。它们的使用如下所示：

```
(Gdb) n
The sum of 1-m is 1275
7          for(i=1; i<=50; i++)
(Gdb) s
sum (m=50) at test.c:16
16            int i,n=0;
```

可见，使用“n”后，程序显示函数 sum 的运行结果并向下执行，而使用“s”后则进入到 sum 函数之中单步运行。

（7）恢复程序运行

在查看完所需变量及堆栈情况后，就可以使用命令“c”（continue）恢复程序的正常运行了。这时，它会把剩余还未执行的程序执行完，并显示剩余程序中的执行结果。以下是之前使用“n”命令恢复后的执行结果：

```
(Gdb) c
Continuing.
The sum of 1-50 is :1275

Program exited with code 031.
```

可以看出，程序在运行完后退出，之后程序处于“停止状态”。

小知识 在 Gdb 中，程序的运行状态有“运行”、“暂停”和“停止”3 种，其中“暂停”状态为程序遇到了断点或观察点之类的，程序暂时停止运行，而此时函数的地址、函数参数、函数内的局部变量都会被压入“栈”（Stack）中。故在这种状态下可以查看函数的变量值等各种属性。但在函数处于“停止”状态之后，“栈”就会自动撤销，它也就无法查看各种信息了。

3.5.2 Gdb 基本命令

Gdb 的命令可以通过查看 help 进行查找，由于 Gdb 的命令很多，因此 Gdb 的 help 将其分成了很多种类（class），用户可以通过进一步查看相关 class 找到相应命令。如下所示：

```
(gdb) help
List of classes of commands:

aliases -- Aliases of other commands
breakpoints -- Making program stop at certain points
data -- Examining data
files -- Specifying and examining files
```

```
internals -- Maintenance commands
…
Type "help" followed by a class name for a list of commands in that class.
Type "help" followed by command name for full documentation.
Command name abbreViations are allowed if unambiguous.
```

上述列出了 Gdb 各个分类的命令，注意底部的加粗部分说明其为分类命令。接下来可以具体查找各分类种的命令。如下所示：

```
(gdb) help data
Examining data.

List of commands:

call -- Call a function in the program
delete display -- Cancel some expressions to be displayed when program stops
delete mem -- Delete memory region
disable display -- Disable some expressions to be displayed when program stops
…
Type "help" followed by command name for full documentation.
Command name abbreViations are allowed if unambiguous.
```

至此，若用户想要查找 call 命令，就可键入“help call”。

```
(gdb) help call
Call a function in the program.
The argument is the function name and arguments, in the notation of the
current working language.  The result is printed and saved in the value
history, if it is not void.
```

当然，若用户已知命令名，直接键入“help [command]”也是可以的。

Gdb 中的命令主要分为以下几类：工作环境相关命令、设置断点与恢复命令、源代码查看命令、查看运行数据相关命令及修改运行参数命令。以下就分别对这几类的命令进行讲解。

1．工作环境相关命令

Gdb 中不仅可以调试所运行的程序，而且还可以对程序相关的工作环境进行相应的设定，甚至还可以使用 shell 中的命令进行相关的操作，其功能极其强大。表 3.10 所示为 Gdb 常见工作环境相关命令。

表 3.10 Gdb 工作环境相关命令

命 令 格 式	含 义
set args 运行时的参数	指定运行时参数，如 set args 2
show args	查看设置好的运行参数
path dir	设定程序的运行路径
show paths	查看程序的运行路径
set enVironment var [=value]	设置环境变量
show enVironment [var]	查看环境变量
cd dir	进入到 dir 目录，相当于 shell 中的 cd 命令
pwd	显示当前工作目录
shell command	运行 shell 的 command 命令

2．设置断点与恢复命令

Gdb 中设置断点与恢复的常见命令如表 3.11 所示。

表 3.11 Gdb 设置断点与恢复相关命令

命 令 格 式	含 义
bnfo b	查看所设断点
break 行号或函数名 <条件表达式>	设置断点
tbreak 行号或函数名 <条件表达式>	设置临时断点，到达后被自动删除
delete [断点号]	删除指定断点，其断点号为“info b”中的第一栏。若缺省断点号则删除所有断点
disable [断点号]]	停止指定断点，使用“info b”仍能查看此断点。同 delete 一样，省断点号则停止所有断点
enable [断点号]	激活指定断点，即激活被 disable 停止的断点
condition [断点号] <条件表达式>	修改对应断点的条件
ignore [断点号]<num>	在程序执行中，忽略对应断点 num 次
step	单步恢复程序运行，且进入函数调用
next	单步恢复程序运行，但不进入函数调用
finish	运行程序，直到当前函数完成返回
c	继续执行函数，直到函数结束或遇到新的断点

由于设置断点在 Gdb 的调试中非常重要，所以在此再着重讲解一下 Gdb 中设置断点的方法。

Gdb 中设置断点有多种方式：其一是按行设置断点，设置方法在 3.5.1 节已经指出，在此就不重复了。另外还可以设置函数断点和条件断点，在此结合上一小节的代码，具体介绍后两种设置断点的方法。

① 函数断点

Gdb 中按函数设置断点只需把函数名列在命令“b”之后，如下所示：

```
(gdb) b sum
Breakpoint 1 at 0x80484ba: file test.c, line 16.
(gdb) info b
Num Type         Disp Enb Address    What
1   breakpoint    keep y   0x080484ba in sum at test.c:16
```

要注意的是，此时的断点实际是在函数的定义处，也就是在 16 行处（注意第 16 行还未执行）。

② 条件断点

Gdb 中设置条件断点的格式为：b 行数或函数名 if 表达式。具体实例如下所示：

```
(gdb) b 8 if i==10
Breakpoint 1 at 0x804848c: file test.c, line 8.
(gdb) info b
Num Type         Disp Enb Address    What
1   breakpoint    keep y   0x0804848c in main at test.c:8
       stop only if i == 10
(gdb) r
Starting program: /home/yul/test
The sum of 1-m is 1275


Breakpoint 1, main () at test.c:9
9            n += i;
(gdb) p i
$1 = 10
```

可以看到，该例中在第 8 行（也就是运行完第 7 行的 for 循环）设置了一个“i==0”的条件断点，在程序运行之后可以看出，程序确实在 i 为 10 时暂停运行。

3．Gdb 中源码查看相关命令

在 Gdb 中可以查看源码以方便其他操作，它的常见相关命令如表 3.12 所示。

表 3.12 Gdb 源码查看相关相关命令

命 令 格 式	含 义
list <行号>|<函数名>	查看指定位置代码
file [文件名]	加载指定文件
forward-search 正则表达式	源代码前向搜索
reverse-search 正则表达式	源代码后向搜索
dir dir	停止路径名
show directories	显示定义了的源文件搜索路径
info line	显示加载到 Gdb 内存中的代码

4．Gdb 中查看运行数据相关命令

Gdb 中查看运行数据是指当程序处于“运行”或“暂停”状态时，可以查看的变量及表达式的信息，其常见命令如表 3.13 所示：

表 3.13 Gdb 查看运行数据相关命令

命 令 格 式	含 义
print 表达式|变量	查看程序运行时对应表达式和变量的值
x <n/f/u>	查看内存变量内容。其中 n 为整数表示显示内存的长度，f 表示显示的格式，u 表示从当前地址往后请求显示的字节数
display 表达式	设定在单步运行或其他情况中，自动显示的对应表达式的内容

5．Gdb 中修改运行参数相关命令

Gdb 还可以修改运行时的参数，并使该变量按照用户当前输入的值继续运行。它的设置方法为：在单步执行的过程中，键入命令“set 变量＝设定值”。这样，在此之后，程序就会按照该设定的值运行了。下面，笔者结合上一节的代码将 n 的初始值设为 4，其代码如下所示：

```
(Gdb) b 7
Breakpoint 5 at 0x804847a: file test.c, line 7.
(Gdb) r
Starting program: /home/yul/test
The sum of 1-m is 1275

Breakpoint 5, main () at test.c:7
7               for(i=1; i<=50; i++)
(Gdb) set n=4
(Gdb) c
Continuing.
```

```
The sum of 1-50 is 1279


Program exited with code 031.
```

可以看到，最后的运行结果确实比之前的值大了 4。

Gdb 的使用切记点：

- 在 Gcc 编译选项中一定要加入“-g”。
- 只有在代码处于“运行”或“暂停”状态时才能查看变量值。
- 设置断点后程序在指定行之前停止。

3.6 Make 工程管理器

到此为止，读者已经了解了如何在 Linux 下使用编辑器编写代码，如何使用 Gcc 把代码编译成可执行文件，还学习了如何使用 Gdb 来调试程序，那么，所有的工作看似已经完成了，为什么还需要 Make 这个工程管理器呢？

所谓工程管理器，顾名思义，是指管理较多的文件的。读者可以试想一下，有一个上百个文件的代码构成的项目，如果其中只有一个或少数几个文件进行了修改，按照之前所学的 Gcc 编译工具，就不得不把这所有的文件重新编译一遍，因为编译器并不知道哪些文件是最近更新的，而只知道需要包含这些文件才能把源代码编译成可执行文件，于是，程序员就不能不再重新输入数目如此庞大的文件名以完成最后的编译工作。

但是，请读者仔细回想一下本书在 3.1.2 节中所阐述的编译过程，编译过程是分为编译、汇编、链接不同阶段的，其中编译阶段仅检查语法错误以及函数与变量的声明是否正确声明了，在链接阶段则主要完成是函数链接和全局变量的链接。因此，那些没有改动的源代码根本不需要重新编译，而只要把它们重新链接进去就可以了。所以，人们就希望有一个工程管理器能够自动识别更新了的文件代码，同时又不需要重复输入冗长的命令行，这样，Make 工程管理器也就应运而生了。

实际上，Make 工程管理器也就是个“自动编译管理器”，这里的“自动”是指它能够根据文件时间戳自动发现更新过的文件而减少编译的工作量，同时，它通过读入 Makefile 文件的内容来执行大量的编译工作。用户只需编写一次简单的编译语句就可以了。它大大提高了实际项目的工作效率，而且几乎所有 Linux 下的项目编程均会涉及它，希望读者能够认真学习本节内容。

3.6.1 Makefile 基本结构

Makefile 是 Make 读入的惟一配置文件，因此本节的内容实际就是讲述 Makefile 的编写规则。在一个 Makefile 中通常包含如下内容：

- 需要由 make 工具创建的目标体（target），通常是目标文件或可执行文件；
- 要创建的目标体所依赖的文件（dependency_file）；
- 创建每个目标体时需要运行的命令（command）。

它的格式为：

```
target: dependency_files
      command
```

例如，有两个文件分别为 hello.c 和 hello.h，创建的目标体为 hello.o，执行的命令为 gcc 编译指令：gcc –c hello.c，那么，对应的 Makefile 就可以写为：

```
#The simplest example
hello.o: hello.c hello.h
      gcc –c hello.c –o hello.o
```

接着就可以使用 make 了。使用 make 的格式为：make target，这样 make 就会自动读入 Makefile（也可以是首字母小写 makefile）并执行对应 target 的 command 语句，并会找到相应的依赖文件。如下所示：

```
[root@localhost makefile]# make hello.o
gcc –c hello.c –o hello.o
[root@localhost makefile]# ls
hello.c  hello.h  hello.o  Makefile
```

可以看到，Makefile 执行了“hello.o”对应的命令语句，并生成了“hello.o”目标体。

注意 在 Makefile 中的每一个 command 前必须有“Tab”符，否则在运行 make 命令时会出错。

3.6.2 Makefile 变量

上面示例的 Makefile 在实际中是几乎不存在的，因为它过于简单，仅包含两个文件和一个命令，在这种情况下完全不必要编写 Makefile 而只需在 Shell 中直接输入即可，在实际中使用的 Makefile 往往是包含很多的文件和命令的，这也是 Makefile 产生的原因。下面就可给出稍微复杂一些的 Makefile 进行讲解：

```
sunq:kang.o yul.o
Gcc kang.o bar.o -o myprog
kang.o : kang.c kang.h head.h
Gcc –Wall –O -g –c kang.c -o kang.o
yul.o : bar.c head.h
Gcc - Wall –O -g –c yul.c -o yul.o
```

在这个 Makefile 中有 3 个目标体（target），分别为 sunq、kang.o 和 yul.o，其中第一个目标体的依赖文件就是后两个目标体。如果用户使用命令“make sunq”，则 make 管理器就是找到 sunq 目标体开始执行。

这时，make 会自动检查相关文件的时间戳。首先，在检查“kang.o”、“yul.o”和“sunq”3 个文件的时间戳之前，它会向下查找那些把“kang.o”或“yul.o”作为目标文件的时间戳。

比如，“kang.o”的依赖文件为“kang.c”、“kang.h”、“head.h”。如果这些文件中任何一个的时间戳比“kang.o”新，则命令“gcc –Wall –O -g –c kang.c -o kang.o”将会执行，从而更新文件“kang.o”。在更新完“kang.o”或“yul.o”之后，make 会检查最初的“kang.o”、“yul.o”和“sunq”3 个文件，只要文件“kang.o”或“yul.o”中的任比文件时间戳比“sunq”新，则第二行命令就会被执行。这样，make 就完成了自动检查时间戳的工作，开始执行编译工作。这也就是 Make 工作的基本流程。

接下来，为了进一步简化编辑和维护 Makefile，make 允许在 Makefile 中创建和使用变量。变量是在 Makefile 中定义的名字，用来代替一个文本字符串，该文本字符串称为该变量的值。在具体要求下，这些值可以代替目标体、依赖文件、命令以及 makefile 文件中其他部分。在 Makefile 中的变量定义有两种方式：一种是递归展开方式，另一种是简单方式。

递归展开方式定义的变量是在引用在该变量时进行替换的，即如果该变量包含了对其他变量的应用，则在引用该变量时一次性将内嵌的变量全部展开，虽然这种类型的变量能够很好地完成用户的指令，但是它也有严重的缺点，如不能在变量后追加内容（因为语句：CFLAGS = $(CFLAGS) -O 在变量扩展过程中可能导致无穷循环）。

为了避免上述问题，简单扩展型变量的值在定义处展开，并且只展开一次，因此它不包含任何对其他变量的引用，从而消除变量的嵌套引用。

递归展开方式的定义格式为：VAR=var。

简单扩展方式的定义格式为：VAR：=var。

Make 中的变量使用均使用格式为：$(VAR)。

变量名是不包括“:”、“#”、“=”结尾空格的任何字符串。同时，变量名中包含字母、数字以及下划线以外的情况应尽量避免，因为它们可能在将来被赋予特别的含义。

注意 变量名是大小写敏感的，例如变量名“foo”、“FOO”、和“Foo”代表不同的变量。

推荐在 makefile 内部使用小写字母作为变量名，预留大写字母作为控制隐含规则参数或用户重载命令选项参数的变量名。

下面给出了上例中用变量替换修改后的 Makefile，这里用 OBJS 代替 kang.o 和 yul.o，用 CC 代替 Gcc，用 CFLAGS 代替“-Wall -O –g”。这样在以后修改时，就可以只修改变量定义，而不需要修改下面的定义实体，从而大大简化了 Makefile 维护的工作量。

经变量替换后的 Makefile 如下所示：

```
OBJS = kang.o yul.o
CC = Gcc
CFLAGS = -Wall -O -g
sunq : $(OBJS)
      $(CC) $(OBJS) -o sunq
kang.o : kang.c kang.h
      $(CC) $(CFLAGS) -c kang.c -o kang.o
yul.o : yul.c yul.h
      $(CC) $(CFLAGS) -c yul.c -o yul.o
```

可以看到，此处变量是以递归展开方式定义的。

Makefile 中的变量分为用户自定义变量、预定义变量、自动变量及环境变量。如上例中的 OBJS 就是用户自定义变量，自定义变量的值由用户自行设定，而预定义变量和自动变量为通常在 Makefile 都会出现的变量，其中部分有默认值，也就是常见的设定值，当然用户可以对其进行修改。

预定义变量包含了常见编译器、汇编器的名称及其编译选项。表 3.14 列出了 Makefile 中常见预定义变量及其部分默认值。

表 3.14 Makefile 中常见预定义变量

命 令 格 式	含 义
AR	库文件维护程序的名称，默认值为 ar
AS	汇编程序的名称，默认值为 as
CC	C 编译器的名称，默认值为 cc
CPP	C 预编译器的名称，默认值为$(CC) –E
CXX	C++编译器的名称，默认值为 g++
FC	FORTRAN 编译器的名称，默认值为 f77
RM	文件删除程序的名称，默认值为 rm –f
ARFLAGS	库文件维护程序的选项，无默认值
ASFLAGS	汇编程序的选项，无默认值
CFLAGS	C 编译器的选项，无默认值
CPPFLAGS	C 预编译的选项，无默认值
CXXFLAGS	C++编译器的选项，无默认值
FFLAGS	FORTRAN 编译器的选项，无默认值

可以看出，上例中的 CC 和 CFLAGS 是预定义变量，其中由于 CC 没有采用默认值，因此，需要把“CC=Gcc”明确列出来。

由于常见的 Gcc 编译语句中通常包含了目标文件和依赖文件，而这些文件在 Makefile 文件中目标体的一行已经有所体现，因此，为了进一步简化 Makefile 的编写，就引入了自动变量。自动变量通常可以代表编译语句中出现目标文件和依赖文件等，并且具有本地含义（即下一语句中出现的相同变量代表的是下一语句的目标文件和依赖文件）。表 3.15 列出了 Makefile 中常见自动变量。

表 3.15 Makefile 中常见自动变量

命 令 格 式	含 义
$*	不包含扩展名的目标文件名称
$+	所有的依赖文件，以空格分开，并以出现的先后为序，可能包含重复的依赖文件
$<	第一个依赖文件的名称

$?	所有时间戳比目标文件晚的依赖文件，并以空格分开

续表

命 令 格 式	含　义
$@	目标文件的完整名称
$^	所有不重复的依赖文件，以空格分开
$%	如果目标是归档成员，则该变量表示目标的归档成员名称

自动变量的书写比较难记，但是在熟练了之后会非常的方便，请读者结合下例中的自动变量改写的 Makefile 进行记忆。

```
OBJS = kang.o yul.o
CC = Gcc
CFLAGS = -Wall -O -g
sunq : $(OBJS)
     $(CC) $^ -o $@
kang.o : kang.c kang.h
     $(CC) $(CFLAGS) -c $< -o $@
yul.o : yul.c yul.h
     $(CC) $(CFLAGS) -c $< -o $@
```

另外，在 Makefile 中还可以使用环境变量。使用环境变量的方法相对比较简单，make 在启动时会自动读取系统当前已经定义了的环境变量，并且会创建与之具有相同名称和数值的变量。但是，如果用户在 Makefile 中定义了相同名称的变量，那么用户自定义变量将会覆盖同名的环境变量。

3.6.3　Makefile 规则

Makefile 的规则是 Make 进行处理的依据，它包括了目标体、依赖文件及其之间的命令语句。一般的，Makefile 中的一条语句就是一个规则。在上面的例子中，都显示地指出了 Makefile 中的规则关系，如“$(CC) $(CFLAGS) -c $< -o $@”，但为了简化 Makefile 的编写，make 还定义了隐式规则和模式规则，下面就分别对其进行讲解。

1．隐式规则

隐含规则能够告诉 make 怎样使用传统的技术完成任务，这样，当用户使用它们时就不必详细指定编译的具体细节，而只需把目标文件列出即可。Make 会自动搜索隐式规则目录来确定如何生成目标文件。如上例就可以写成：

```
OBJS = kang.o yul.o
CC = Gcc
CFLAGS = -Wall -O -g
sunq : $(OBJS)
     $(CC) $^ -o $@
```

为什么可以省略后两句呢？因为 Make 的隐式规则指出：所有“.o”文件都可自动由“.c”文件使用命令“$(CC) $(CPPFLAGS) $(CFLAGS) -c file.c –o file.o”生成。这样“kang.o”和“yul.o”就会分别调用“$(CC) $(CFLAGS) -c kang.c -o kang.o”和“$(CC) $(CFLAGS) -c yul.c -o yul.o”生成。

注意 在隐式规则只能查找到相同文件名的不同后缀名文件，如“kang.o”文件必须由“kang.c”文件生成。

表 3.16 给出了常见的隐式规则目录：

表 3.16 Makefile 中常见隐式规则目录

对应语言后缀名	规　　则
C 编译：**.c** 变为**.o**	$(CC) –c $(CPPFLAGS) $(CFLAGS)
C++编译：**.cc 或.C** 变为**.o**	$(CXX) -c $(CPPFLAGS) $(CXXFLAGS)
Pascal 编译：**.p** 变为**.o**	$(PC) -c $(PFLAGS)
Fortran 编译：**.r** 变为**-o**	$(FC) -c $(FFLAGS)

2．模式规则

模式规则是用来定义相同处理规则的多个文件的。它不同于隐式规则，隐式规则仅仅能够用 make 默认的变量来进行操作，而模式规则还能引入用户自定义变量，为多个文件建立相同的规则，从而简化 Makefile 的编写。

模式规则的格式类似于普通规则，这个规则中的相关文件前必须用“%”标明。使用模式规则修改后的 Makefile 的编写如下：

```
OBJS = kang.o yul.o
CC = Gcc
CFLAGS = -Wall -O -g
sunq : $(OBJS)
     $(CC) $^ -o $@
%.o : %.c
     $(CC) $(CFLAGS) -c $< -o $@
```

3.6.4 Make 管理器的使用

使用 Make 管理器非常简单，只需在 make 命令的后面键入目标名即可建立指定的目标，如果直接运行 make，则建立 Makefile 中的第一个目标。

此外 make 还有丰富的命令行选项，可以完成各种不同的功能。下表 3.17 列出了常用的 make 命令行选项。

表 3.17 make 的命令行选项

命 令 格 式	含　　义
-C dir	读入指定目录下的 Makefile

-f file	读入当前目录下的 file 文件作为 Makefile

续表

命 令 格 式	含　义
-i	忽略所有的命令执行错误
-I dir	指定被包含的 Makefile 所在目录
-n	只打印要执行的命令，但不执行这些命令
-p	显示 make 变量数据库和隐含规则
-s	在执行命令时不显示命令
-w	如果 make 在执行过程中改变目录，则打印当前目录名

3.7　使用 autotools

在上一小节，读者已经了解到了 make 项目管理器的强大功能。的确，Makefile 可以帮助 make 完成它的使命，但要承认的是，编写 Makefile 确实不是一件轻松的事，尤其对于一个较大的项目而言更是如此。那么，有没有一种轻松的手段生成 Makefile 而同时又能让用户享受 make 的优越性呢？本节要讲的 autotools 系列工具正是为此而设的，它只需用户输入简单的目标文件、依赖文件、文件目录等就可以轻松地生成 Makefile 了，这无疑是广大用户的所希望的。另外，这些工具还可以完成系统配置信息的收集，从而可以方便地处理各种移植性的问题。也正是基于此，现在 Linux 上的软件开发一般都用 autotools 来制作 Makefile，读者在后面的讲述中就会了解到。

3.7.1　autotools 使用流程

正如前面所言，autotools 是系列工具，读者首先要确认系统是否装了以下工具（可以用 which 命令进行查看）。

- aclocal
- autoscan
- autoconf
- autoheader
- automake

使用 autotools 主要就是利用各个工具的脚本文件以生成最后的 Makefile。其总体流程是这样的。

- 使用 aclocal 生成一个“aclocal.m4”文件，该文件主要处理本地的宏定义；
- 改写“configure.scan”文件，并将其重命名为“configure.in”，并使用 autoconf 文件生成 configure 文件。

接下来，笔者将通过一个简单的 hello.c 例子带领读者熟悉 autotools 生成 makefile 的过程，由于在这过程中有涉及较多的脚本文件，为了更清楚地了解相互之间的关系，强烈建议读者

实际动手操作以体会其整个过程。

1．autoscan

它会在给定目录及其子目录树中检查源文件，若没有给出目录，就在当前目录及其子目录树中进行检查。它会搜索源文件以寻找一般的移植性问题并创建一个文件“configure.scan”，该文件就是接下来 autoconf 要用到的“configure.in”原型。如下所示：

```
[root@localhost automake]# autoscan
autom4te: configure.ac: no such file or directory
autoscan: /usr/bin/autom4te failed with exit status: 1
[root@localhost automake]# ls
autoscan.log  configure.scan  hello.c
```

由上述代码可知 autoscan 首先会尝试去读入“configure.ac”（同 configure.in 的配置文件）文件，此时还没有创建该配置文件，于是它会自动生成一个“configure.in”的原型文件“configure.scan”。

2．autoconf

configure.in 是 autoconf 的脚本配置文件，它的原型文件“configure.scan”如下所示：

```
#                                               -*- Autoconf -*-
# Process this file with autoconf to produce a configure script.
AC_PREREQ(2.59)
#The next one is modified by sunq
#AC_INIT(FULL-PACKAGE-NAME,VERSION,BUG-REPORT-ADDRESS)
AC_INIT(hello,1.0)
# The next one is added by sunq
AM_INIT_AUTOMAKE(hello,1.0)
AC_CONFIG_SRCDIR([hello.c])
AC_CONFIG_HEADER([config.h])
# Checks for programs.
AC_PROG_CC
# Checks for libraries.
# Checks for header files.
# Checks for typedefs, structures, and compiler characteristics.
# Checks for library functions.
AC_CONFIG_FILES([Makefile])
AC_OUTPUT
```

下面对这个脚本文件进行解释。

- 以“#”号开始的行为注释。

- AC_PREREQ 宏声明本文件要求的 autoconf 版本，如本例使用的版本 2.59。
- AC_INIT 宏用来定义软件的名称和版本等信息，在本例中省略了 BUG-REPORT-ADDRESS，一般为作者的 E-mail。
- AM_INIT_AUTOMAKE 是笔者另加的，它是 automake 所必备的宏，也同前面一样，PACKAGE 是所要产生软件套件的名称，VERSION 是版本编号。
- AC_CONFIG_SRCDIR 宏用来侦测所指定的源码文件是否存在，来确定源码目录的有效性。在此处为当前目录下的 hello.c。
- AC_CONFIG_HEADER 宏用于生成 config.h 文件，以便 autoheader 使用。
- AC_CONFIG_FILES 宏用于生成相应的 Makefile 文件。
- 中间的注释间可以添加分别用户测试程序、测试函数库、测试头文件等宏定义。

接下来首先运行 aclocal，生成一个“aclocal.m4”文件，该文件主要处理本地的宏定义。如下所示：

```
[root@localhost automake]# aclocal
```

再接着运行 autoconf，生成“configure”可执行文件。如下所示：

```
[root@localhost automake]# autoconf
[root@localhost automake]# ls
aclocal.m4  autom4te.cache  autoscan.log  configure  configure.in  hello.c
```

3．autoheader

接着使用 autoheader 命令，它负责生成 config.h.in 文件。该工具通常会从“acconfig.h”文件中复制用户附加的符号定义，因此此处没有附加符号定义，所以不需要创建“acconfig.h”文件。如下所示：

```
[root@localhost automake]# autoheader
```

4．automake

这一步是创建 Makefile 很重要的一步，automake 要用的脚本配置文件是 Makefile.am，用户需要自己创建相应的文件。之后，automake 工具转换成 Makefile.in。在该例中，笔者创建的文件为 Makefile.am 如下所示：

```
AUTOMAKE_OPTIONS=foreign
bin_PROGRAMS= hello
hello_SOURCES= hello.c
```

下面对该脚本文件的对应项进行解释。

- 其中的 AUTOMAKE_OPTIONS 为设置 automake 的选项。由于 GNU（在第 1 章中已经有所介绍）对自己发布的软件有严格的规范，比如必须附带许可证声明文件 COPYING 等，否则 automake 执行时会报错。automake 提供了 3 种软件等级：foreign、gnu 和 gnits，让用户选择采用，默认等级为 gnu。在本例使用 foreign 等级，它只检测必须的文件。

- bin_PROGRAMS 定义要产生的执行文件名。如果要产生多个执行文件，每个文件名用空格隔开。
- hello_SOURCES 定义“hello”这个执行程序所需要的原始文件。如果“hello”这个程序是由多个原始文件所产生的，则必须把它所用到的所有原始文件都列出来，并用空格隔开。例如：若目标体“hello”需要“hello.c”、“sunq.c”、“hello.h”三个依赖文件，则定义 hello_SOURCES=hello.c sunq.c hello.h。要注意的是，如果要定义多个执行文件，则对每个执行程序都要定义相应的 file_SOURCES。

接下来可以使用 automake 对其生成“configure.in”文件，在这里使用选项“—adding-missing”可以让 automake 自动添加有一些必需的脚本文件。如下所示：

```
[root@localhost automake]# automake --add-missing
configure.in: installing './install-sh'
configure.in: installing './missing'
Makefile.am: installing 'depcomp'
[root@localhost automake]# ls
aclocal.m4     autoscan.log  configure.in  hello.c    Makefile.am  missing
autom4te.cache  configure    depcomp    install-sh  Makefile.in  config.h.in
```

可以看到，在 automake 之后就可以生成 configure.in 文件。

5．运行 configure

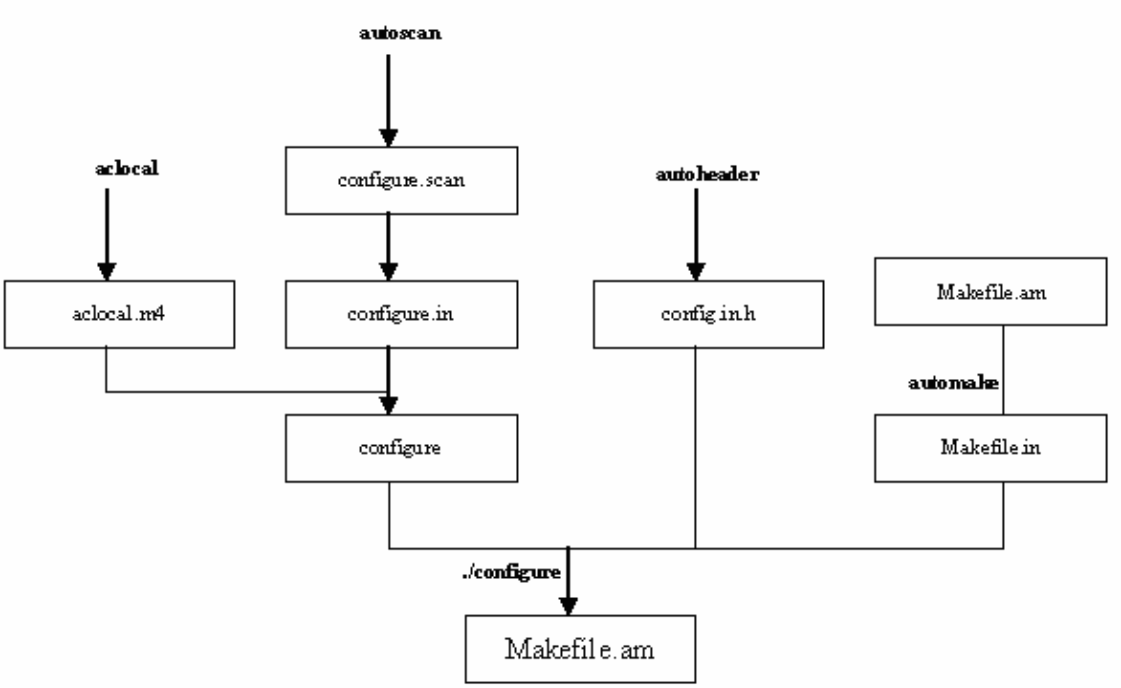
在这一步中，通过运行自动配置设置文件 configure，把 Makefile.in 变成了最终的 Makefile。如下所示：

```
[root@localhost automake]# ./configure
checking for a BSD-compatible install... /usr/bin/install -c
checking whether build enVironment is sane... yes
checking for gawk... gawk
checking whether make sets $(MAKE)... yes
checking for Gcc... Gcc
checking for C compiler default output file name... a.out
checking whether the C compiler works... yes
checking whether we are cross compiling... no
checking for suffix of executables...
checking for suffix of object files... o
checking whether we are using the GNU C compiler... yes
checking whether Gcc accepts -g... yes
checking for Gcc option to accept ANSI C... none needed
checking for style of include used by make... GNU
checking dependency style of Gcc... Gcc3
configure: creating ./config.status
```

```
config.status: creating Makefile
config.status: executing depfiles commands
```

可以看到，在运行 configure 时收集了系统的信息，用户可以在 configure 命令中对其进行方便地配置。在./configure 的自定义参数有两种，一种是开关式（--enable-XXX 或 --disable-XXX），另一种是开放式，即后面要填入一串字符（--with-XXX=yyyy）参数。读者可以自行尝试其使用方法。另外，读者可以查看同一目录下的“config.log”文件，以方便调试之用。

到此为止，makefile 就可以自动生成了。回忆整个步骤，用户不再需要定制不同的规则，而只需要输入简单的文件及目录名即可，这样就大大方便了用户的使用。autotools 生成 Makefile 流程图如图 3.9 所示。



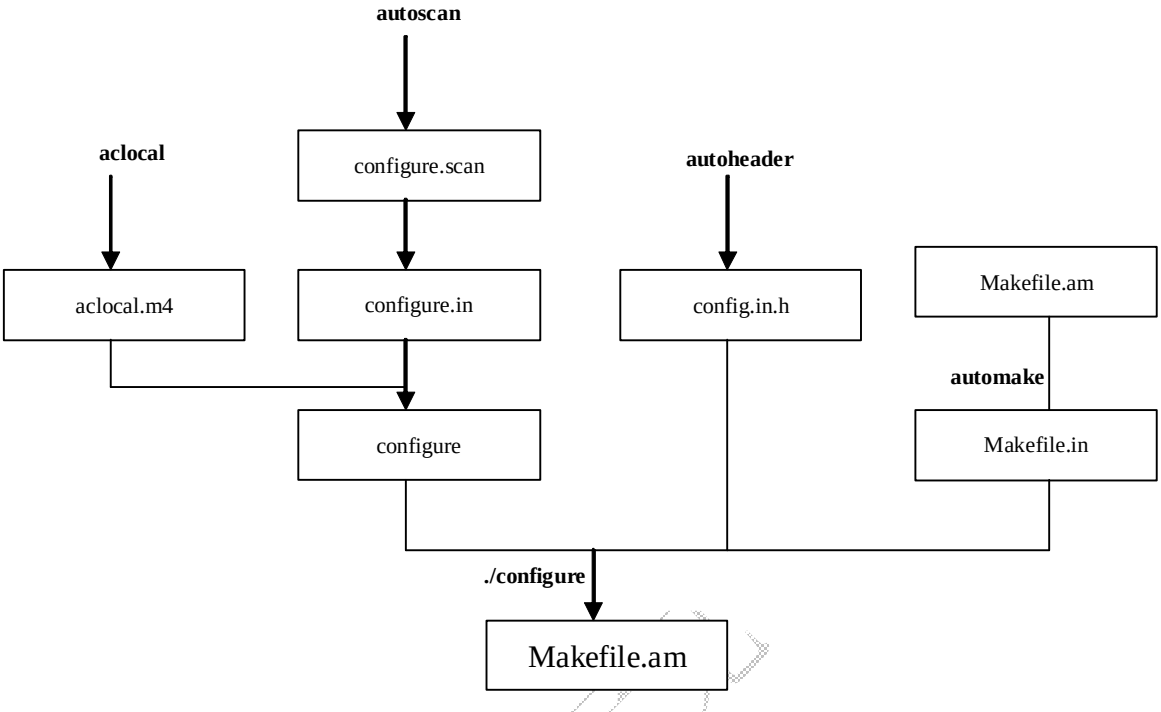


图 3.9　autotools 生成 Makefile 流程图

3.7.2　使用 autotools 所生成的 Makefile

autotools 生成的 Makefile 除具有普通的编译功能外，还具有以下主要功能（感兴趣的读者可以查看这个简单的 hello.c 程序的 makefile）。

1．make

键入 make 默认执行“make all”命令，即目标体为 all，其执行情况如下所示：

```
    [root@localhost automake]# make
    if  Gcc  -DPACKAGE_NAME=\"\"  -DPACKAGE_TARNAME=\"\"  -DPACKAGE_VERSION=\"\"
-DPACKAGE_STRING=\"\" -DPACKAGE_BUGREPORT=\"\" -DPACKAGE=\"hello\" -DVERSION=\"1.0\"
-I. -I.    -g -O2 -MT hello.o -MD -MP -MF ".deps/hello.Tpo" -c -o hello.o hello.c; \
    then mv -f ".deps/hello.Tpo" ".deps/hello.Po"; else rm -f ".deps/hello.Tpo";
exit 1; fi
    Gcc  -g -O2  -o hello  hello.o
```

此时在本目录下就生成了可执行文件“hello”，运行“./hello”能出现正常结果，如下所示：

```
    [root@localhost automake]# ./hello
    Hello!Autoconf!
```

2．make install

此时，会把该程序安装到系统目录中去，如下所示：

```
    [root@localhost automake]# make install
    if  Gcc  -DPACKAGE_NAME=\"\"  -DPACKAGE_TARNAME=\"\"  -DPACKAGE_VERSION=\"\"
-DPACKAGE_STRING=\"\" -DPACKAGE_BUGREPORT=\"\" -DPACKAGE=\"hello\" -DVERSION=\"1.0\"
-I. -I.    -g -O2 -MT hello.o -MD -MP -MF ".deps/hello.Tpo" -c -o hello.o hello.c; \
    then mv -f ".deps/hello.Tpo" ".deps/hello.Po"; else rm -f ".deps/hello.Tpo";
exit 1; fi
    Gcc  -g -O2  -o hello  hello.o
    make[1]: Entering directory '/root/workplace/automake'
    test -z "/usr/local/bin" || mkdir -p -- "/usr/local/bin"
      /usr/bin/install -c 'hello' '/usr/local/bin/hello'
    make[1]: Nothing to be done for 'install-data-am'.
    make[1]: LeaVing directory '/root/workplace/automake'
```

此时，若直接运行 hello，也能出现正确结果，如下所示：

```
    [root@localhost automake]# hello
    Hello!Autoconf!
```

3．make clean

此时，make 会清除之前所编译的可执行文件及目标文件（object file, *.o），如下所示：

```
    [root@localhost automake]# make clean
    test -z "hello" || rm -f hello
    rm -f *.o
```

4．make dist

此时，make 将程序和相关的文档打包为一个压缩文档以供发布，如下所示：

```
    [root@localhost automake]# make dist
    [root@localhost automake]# ls hello-1.0-tar.gz
    hello-1.0-tar.gz
```

可见该命令生成了一个 hello-1.0-tar.gz 的压缩文件。

由上面的讲述读者不难看出，autotools 确实是软件维护与发布的必备工具，鉴于此，如今 GUN 的软件一般都是由 automake 来制作的。

想一想 对于 automake 制作的这类软件，应如何安装呢？

3.8 实验内容

3.8.1 Vi 使用练习

1．实验目的

通过指定指令的 Vi 操作练习，使读者能够熟练使用 Vi 中的常见操作，并且熟悉 Vi 的 3 种模式，如果读者能够熟练掌握实验内容中所要求的内容，则表明对 Vi 的操作已经很熟练了。

2．实验内容

（1）在“/root”目录下建一个名为“/Vi”的目录。
（2）进入“/Vi”目录。
（3）将文件“/etc/inittab”复制到“/Vi”目录下。
（4）使用 Vi 打开“/Vi”目录下的 inittab。
（5）设定行号，指出设定 initdefault（类似于“id:5:initdefault”）的所在行号。
（6）将光标移到该行。
（7）复制该行内容。
（8）将光标移到最后一行行首。
（9）粘贴复制行的内容。
（10）撤销第 9 步的动作。
（11）将光标移动到最后一行的行尾。
（12）粘贴复制行的内容。
（13）光标移到“si::sysinit:/etc/rc.d/rc.sysinit”。
（14）删除该行。
（15）存盘但不退出。
（16）将光标移到首行。
（17）插入模式下输入“Hello,this is Vi world!”。
（18）返回命令行模式。
（19）向下查找字符串“0:wait”。
（20）再向上查找字符串“halt”。
（21）强制退出 Vi，不存盘。
分别指出每个命令处于何种模式下？

3．实验步骤

（1）mkdir /root/Vi
（2）cd /root/Vi
（3）cp /etc/inittab ./
（4）Vi ./inittab

（5）:set nu（底行模式）
（6）17<enter>（命令行模式）
（7）yy
（8）G
（9）p
（10）u
（11）$
（12）p
（13）21G
（14）dd
（15）:w（底行模式）
（16）1G
（17）i 并输入“Hello,this is Vi world!”（插入模式）
（18）Esc
（19）/0:wait（命令行模式）
（20）?halt
（21）:q!（底行模式）

4．实验结果

该实验最后的结果只对“/root/inittab”增加了一行复制的内容：“id:5:initdefault”。

3.8.2 用 Gdb 调试有问题的程序

1．实验目的

通过调试一个有问题的程序，使读者进一步熟练使用 Vi 操作，而且熟练掌握 Gcc 编译命令及 Gdb 的调试命令，通过对有问题程序的跟踪调试，进一步提高发现问题和解决问题的能力。这是一个很小的程序，只有 35 行，希望读者认真调试。

2．实验内容

（1）使用 Vi 编辑器，将以下代码输入到名为 greet.c 的文件中。此代码的原意为输出倒序 main 函数中定义的字符串，但结果显示没有输出。代码如下所示：

```
#include <stdio.h>
int display1(char *string);
int display2(char *string);

int main ()
{
      char string[] = "Embedded Linux";
      display1 (string);
```

```
        display2 (string);
}
int display1 (char *string)
{
        printf ("The original string is %s \n", string);
}
int display2 (char *string1)
{
        char *string2;
        int size,i;
        size = strlen (string1);
        string2 = (char *) malloc (size + 1);
        for (i = 0; i < size; i++)
                string2[size - i] = string1[i];
        string2[size+1] = ' ';
        printf("The string afterward is %s\n",string2);
}
```

（2）使用 Gcc 编译这段代码，注意要加上“-g”选项以方便之后的调试。
（3）运行生成的可执行文件，观察运行结果。
（4）使用 Gdb 调试程序，通过设置断点、单步跟踪，一步步找出错误所在。
（5）纠正错误，更改源程序并得到正确的结果。

3．实验步骤

（1）在工作目录上新建文件 greet.c，并用 Vi 启动：vi greet.c。
（2）在 Vi 中输入以上代码。
（3）在 Vi 中保存并退出：wq。
（4）用 Gcc 编译：gcc -g greet.c -o greet。
（5）运行 greet：./greet，输出为：

```
The original string is Embedded Linux
The string afterward is
```

可见，该程序没有能够倒序输出。
（6）启动 Gdb 调试：gdb greet。
（7）查看源代码，使用命令“l”。
（8）在 30 行（for 循环处）设置断点，使用命令“b 30”。
（9）在 33 行（printf 函数处）设置断点，使用命令“b 33”。
（10）查看断点设置情况，使用命令“info b”。
（11）运行代码，使用命令“r”。
（12）单步运行代码，使用命令“n”。

（13）查看暂停点变量值，使用命令“p string2[size - i]”。

（14）继续单步运行代码数次，并使用命令查看，发现 string2[size-1]的值正确。

（15）继续程序的运行，使用命令“c”。

（16）程序在 printf 前停止运行，此时依次查看 string2[0]、string2[1]…，发现 string[0]没有被正确赋值，而后面的复制都是正确的，这时，定位程序第 31 行，发现程序运行结果错误的原因在于“size-1”。由于 i 只能增到“size-1”，这样 string2[0]就永远不能被赋值而保持 NULL，故输不出任何结果。

（17）退出 Gdb，使用命令 q。

（18）重新编辑 greet.c，把其中的“string2[size - i] = string1[i]”改为“string2[size – i - 1] = string1[i];”即可。

（19）使用 Gcc 重新编译：gcc -g greet.c -o greet。

（20）查看运行结果：./greet

```
The original string is Embedded Linux
The string afterward is xuniL deddedbmE
```

这时，输入结果正确。

4．实验结果

将原来有错的程序经过 Gdb 调试，找出问题所在，并修改源代码，输出正确的倒序显示字符串的结果。

3.8.3 编写包含多文件的 Makefile

1．实验目的

通过对包含多文件的 Makefile 的编写，熟悉各种形式的 Makefile，并且进一步加深对 Makefile 中用户自定义变量、自动变量及预定义变量的理解。

2．实验过程

（1）用 Vi 在同一目录下编辑两个简单的 Hello 程序，如下所示：

```
#hello.c
#include "hello.h"
int main()
{
      printf("Hello everyone!\n");
}
#hello.h
#include <stdio.h>
```

（2）仍在同一目录下用 Vi 编辑 Makefile，且不使用变量替换，用一个目标体实现（即直

接将 hello.c 和 hello.h 编译成 hello 目标体）。然后用 make 验证所编写的 Makefile 是否正确。

（3）将上述 Makefile 使用变量替换实现。同样用 make 验证所编写的 Makefile 是否正确。

（4）用编辑另一 Makefile，取名为 Makefile1，不使用变量替换，但用两个目标体实现（也就是首先将 hello.c 和 hello.h 编译为 hello.o，再将 hello.o 编译为 hello），再用 make 的“-f”选项验证这个 Makefile1 的正确性。

（5）将上述 Makefile1 使用变量替换实现。

3．实验步骤

（1）用 Vi 打开上述两个代码文件“hello.c”和“hello.h”。

（2）在 shell 命令行中用 Gcc 尝试编译，使用命令：“Gcc hello.c –o hello”，并运行 hello 可执行文件查看结果。

（3）删除此次编译的可执行文件：rm hello。

（4）用 Vi 编辑 Makefile，如下所示：

```
hello:hello.c hello.h
	Gcc hello.c -o hello
```

（5）退出保存，在 shell 中键入：make，查看结果。

（6）再次用 Vi 打开 Makefile，用变量进行替换，如下所示：

```
OBJS :=hello.o
CC :=Gcc
hello:$(OBJS)
	$(CC) $^ -o $@
```

（7）退出保存，在 shell 中键入 make，查看结果。

（8）用 Vi 编辑 Makefile1，如下所示：

```
hello:hello.o
	Gcc hello.o -o hello
hello.o:hello.c hello.h
	Gcc -c hello.c -o hello.o
```

（9）退出保存，在 shell 中键入：make -f Makefile1，查看结果。

（10）再次用 Vi 编辑 Makefile1，如下所示：

```
OBJS1 :=hello.o
OBJS2 :=hello.c hello.h
CC :=Gcc
hello:$(OBJS1)
	$(CC) $^ -o $@
$(OBJS1):$(OBJS2)
	$(CC) -c $< -o $@
```

在这里请注意区别“$^”和“$<”。

（11）退出保存，在 shell 中键入 make -f Makefile1，查看结果。

4．实验结果

各种不同形式的 makefile 都能完成其正确的功能。

3.8.4 使用 autotools 生成包含多文件的 Makefile

1．实验目的

通过使用 autotools 生成包含多文件的 Makefile，进一步掌握 autotools 的正确使用方法。同时，掌握 Linux 下安装软件的常用方法。

2．实验过程

（1）在原目录下新建文件夹 auto。
（2）利用上例的两个代码文件“hello.c”和“hello.h”，并将它们复制到该目录下。
（3）使用 autoscan 生成 configure.scan。
（4）编辑 configure.scan，修改相关内容，并将其重命名为 configure.in。
（5）使用 aclocal 生成 aclocal.m4。
（6）使用 autoconf 生成 configure。
（7）使用 autoheader 生成 config.in.h。
（8）编辑 Makefile.am。
（9）使用 automake 生成 Makefile.in。
（10）使用 configure 生成 Makefile。
（11）使用 make 生成 hello 可执行文件，并在当前目录下运行 hello 查看结果。
（12）使用 make install 将 hello 安装到系统目录下，并运行，查看结果。
（13）使用 make dist 生成 hello 压缩包。
（14）解压 hello 压缩包。
（15）进入解压目录。
（16）在该目录下安装 hello 软件。

3．实验步骤

（1）mkdir ./auto。
（2）cp hello.* ./auto（假定原先在“hello.c”文件目录下）。
（3）命令：autoscan。
（4）使用 Vi 编辑 configure.scan 为：

```
#                                               -*- Autoconf -*-
# Process this file with autoconf to produce a configure script.

AC_PREREQ(2.59)
```

```
AC_INIT(hello, 1.0)
AM_INIT_AUTOMAKE(hello,1.0)
AC_CONFIG_SRCDIR([hello.h])
AC_CONFIG_HEADER([config.h])
# Checks for programs.
AC_PROG_CC
# Checks for libraries.
# Checks for header files.
# Checks for typedefs, structures, and compiler characteristics.
# Checks for library functions.
AC_OUTPUT(Makefile)
```

（5）保存退出，并重命名为 configure.in。
（6）运行：aclocal。
（7）运行：autoconf，并用 ls 查看是否生成了 configure 可执行文件。
（8）运行：autoheader。
（9）用 Vi 编辑 Makefile.am 文件为：

```
AUTOMAKE_OPTIONS=foreign
bin_PROGRAMS=hello
hello_SOURCES=hello.c hello.h
```

（10）运行：automake。
（11）运行：./configure。
（12）运行：make。
（13）运行：./hello，查看结果是否正确。
（14）运行：make install。
（15）运行：hello，查看结果是否正确。
（16）运行：make dist。
（17）在当前目录下解压 hello-1.0.tar.gz：tar –zxvf hello-1.0.tar.gz。
（18）进入解压目录：cd ./hello-1.0。
（19）下面开始 Linux 下常见的安装软件步骤：./configure。
（20）运行：make。
（21）运行：./hello（在正常安装时这一步可省略）。
（22）运行：make install。
（23）运行：hello，查看结果是否正确。

4．实验结果

能够正确使用 autotools 生成 Makefile，并且能够安装成功短小的 Hello 软件。

本 章 小 结

本章是 Linux 中进行 C 语言编程的基础，首先讲解了 C 语言编程的关键点，这里关键要了解编辑器、编译链接器、调试器及项目管理工具等关系。

接下来，本章介绍了两个 Linux 中常见的编辑器——Vi 和 Emacs，并且主要按照它们的使用流程进行讲解。

再接下来，本章介绍了 Gcc 编译器的使用和 Gdb 调试器的使用，并结合了具体的实例进行讲解。虽然它们的选项比较多，但是常用的并不多，读者着重掌握笔者例子中使用的一些选项即可。

之后，本章又介绍了 Make 工程管理器的使用，这里包括 Makefile 的基本结构、Makefile 的变量定义及其规则和 Make 的使用。

最后介绍的是 autotools 的使用，这是非常有用的工具，希望读者能够掌握。

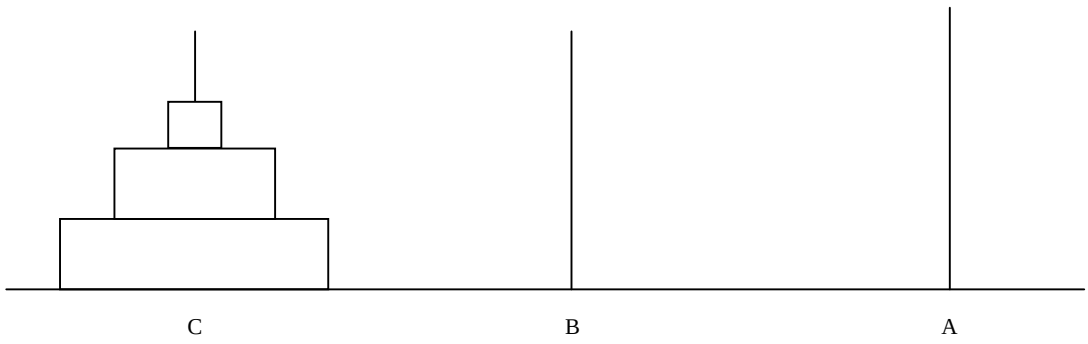
本章的实验安排比较多，包括了 Vi、Gdb、Makefile 和 autotool 的使用，由于这些都是 Linux 中的常用软件，因此希望读者确实掌握。

思考与练习

在 Linux 下使用 Gcc 编译器和 Gdb 调试器编写汉诺塔游戏程序。

汉诺塔游戏介绍如下。

约 19 世纪末，在欧洲的商店中出售一种智力玩具，在一块铜板上有三根杆，如图 3.10 所示。其中，最左边的杆上自上而下、由小到大顺序串着由 64 个圆盘构成的塔。目的是将最左边杆上的盘全部移到右边的杆上，条件是一次只能移动一个盘，且不允许大盘放在小盘的上面。



汉诺塔

图 3.10　汉诺塔游戏示意图

“黑色经典”系列之《嵌入式 Linux 应用程序开发详解》

第 4 章 嵌入式系统基础

本章目标

从本章开始，读者开始真正进入嵌入式领域学习。本章讲解嵌入式系统的基础知识及基本服务的配置，学习完本章读者将掌握如下内容。

- 了解嵌入式系统的含义及其发展情况
- 了解嵌入式系统的体系结构
- 了解 ARM 处理器及 ARM9 的相关知识
- 熟悉三星处理器 S3C2410
- 了解嵌入式系统的基本调试手段

4.1 嵌入式系统概述

4.1.1 嵌入式系统简介

尼葛洛庞帝 2001 年访华时的预言“4～5 年后，嵌入式智能电脑将是继 PC 和 Internet 后的最伟大发明！”如今，嵌入式系统已成功当今最为热门的领域之一，它迅猛的发展势头引起了社会各方面人士的关注。如家用电器、手持通信设备、信息终端、仪器仪表、汽车、航天航空、军事装备、制造工业、过程控制等。今天，嵌入式系统带来的工业年产值已超过 1 万亿美元。用市场观点来看，PC 已经从高速增长进入到平稳发展时期，其年增长率由 20 世纪 90 年代中期的 35%逐年下降，使单纯由 PC 机带领电子产业蒸蒸日上的时代成为历史。根据 PC 时代的概念，美国 *Business Week* 杂志提出了“后 PC 时代”概念，即计算机、通信和消费产品的技术将结合起来，以 3C 产品的形式通过 Internet 进入家庭。这必将培育出一个庞大的嵌入式应用市场。那么究竟什么是嵌入式系统呢？

按照电器工程协会的定义，嵌入式系统是用来控制或者监视机器、装置、工厂等大规模系统的设备。这个定义主要是从嵌入式系统的用途方面来进行定义的。

那么，下面再来看一个在多数书籍资料中的关于嵌入式系统的定义：嵌入式系统是指以应用为中心，以计算机技术为基础，软件硬件可剪裁，适应应用系统对功能、可靠性、成本、体积、功耗严格要求的专用计算机系统。它主要由嵌入式微处理器、外围硬件设备、嵌入式操作系统以及用户应用软件等部分组成。它具有“嵌入性”、“专用性”与“计算机系统”的三个基本要素。

从这个定义可以看出，人们平常所广泛使用的手机、PDA、MP3、机顶盒都属于嵌入式系统设备；而车载 GPS 系统、机器人也是属于嵌入式系统。图 4.1 展出了人们日常生活中形形色色的嵌入式产品。的确，嵌入式系统已经进入了人们生活的方方面面。

图 4.1　生活中的嵌入式设备

4.1.2　嵌入式系统发展历史

嵌入式系统经过 30 年的发展历程，主要经历了 4 个阶段。

第 1 阶段是以单芯片为核心的可编程控制器形式的系统。这类系统大部分应用于一些专业性强的工业控制系统中，一般没有操作系统的支持，通过汇编语言编程对系统进行直接控制。这一阶段系统的主要特点是：系统结构和功能相对单一，处理效率较低，存储容量较小，几乎没有用户接口。由于这种嵌入式系统使用简单、价格低，因此以前在国内工业领域应用较为普遍，但是现在已经远不能适应高效的、需要大容量存储的现代工业控制和新兴信息家电等领域的需求。

第 2 阶段是以嵌入式 CPU 为基础、以简单操作系统为核心的嵌入式系统。其主要特点是：CPU 种类繁多，通用性比较弱；系统开销小，效率高；操作系统达到一定的兼容性和扩展性；应用软件较专业化，用户界面不够友好。

第 3 阶段是以嵌入式操作系统为标志的嵌入式系统。其主要特点是：嵌入式操作系统能运行于各种不同类型的微处理器上，兼容性好；操作系统内核小、效率高，并且具有高度的模块化和扩展性；具备文件和目录管理、支持多任务、支持网络应用、具备图形窗口和用户界面；具有大量的应用程序接口 API，开发应用程序较简单；嵌入式应用软件丰富。

第 4 阶段是以 Internet 为标志的嵌入式系统。这是一个正在迅速发展的阶段。目前大多数嵌入式系统还孤立于 Internet 之外，但随着 Internet 的发展以及 Internet 技术与信息家电、

工业控制技术结合日益密切，嵌入式设备与 Internet 的结合将代表嵌入式系统的未来。

4.1.3 嵌入式系统的特点

（1）面向特定应用的特点。从前面图 1.1 中也可以看出，嵌入式系统与通用型系统的最大区别就在于嵌入式系统大多工作在为特定用户群设计的系统中，因此它通常都具有低功耗、体积小、集成度高等特点，并且可以满足不用应用的特定需求。

（2）嵌入式系统的硬件和软件都必须进行高效地设计，量体裁衣、去除冗余，力争在同样的硅片面积上实现更高的性能，这样才能在具体应用中对处理器的选择更具有竞争力。

（3）嵌入式系统是将先进的计算机技术、半导体技术和电子技术与各个行业的具体应用相结合后的产物。这一点就决定了它必然是一个技术密集、资金密集、高度分散、不断创新的知识集成系统，从事嵌入式系统开发的人才也必须是复合型人才。

（4）为了提高执行速度和系统可靠性，嵌入式系统中的软件一般都固化在存储器芯片或单片机本身中，而不是存储于磁盘中。

（5）嵌入式开发的软件代码尤其要求高质量、高可靠性，由于嵌入式设备所处的环境往往是无人职守或条件恶劣的情况下，因此，其代码必须有更高的要求。

（6）嵌入式系统本身不具备二次开发能力，即设计完成后用户通常不能对其中的程序功能进行修改，必须有一套开发工具和环境才能进行再次开发。

4.1.4 嵌入式系统的体系结构

嵌入式系统作为一类特殊的计算机系统，一般包括以下 3 个方面：硬件设备、嵌入式操作系统和应用软件。它们之间的关系如图 4.2 所示。

硬件设备包括嵌入式处理器和外围设备。其中的嵌入式处理器（CPU）是嵌入式系统的核心部分，它与通用处理器最大的区别在于，嵌入式处理器大多工作在为特定用户群所专门设计的系统中，它将通用处理器中许多由板卡完成的任务集成到芯片内部，从而有利于嵌入式系统在设计时趋于小型化，同时还具有很高的效率和可靠性。如今，全世界嵌入式处理器已经超过 1000 多种，流行的体系结构有 30 多个系列，其中以 ARM、PowerPC、MC 68000、MIPS 等使用得最为广泛。

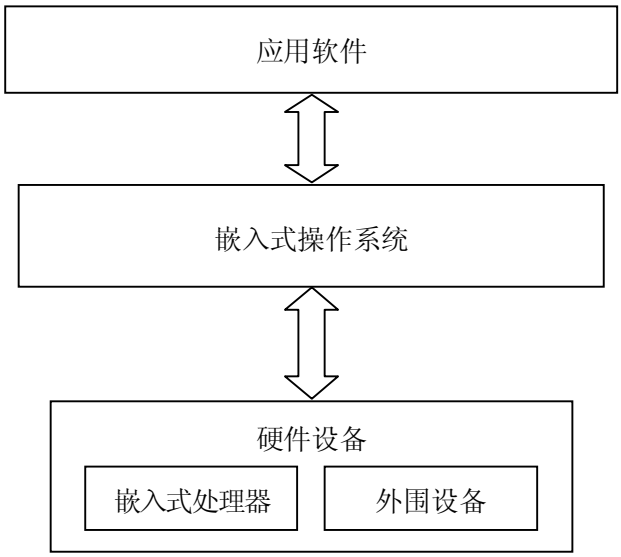


图 4.2　嵌入式体系结构图

外围设备是嵌入式系统中用于完成存储、通信、调试、显示等辅助功能的其他部件。目前常用的嵌入式外围设备按功能可以分为存储设备（如 RAM、SRAM、Flash 等）、通信设备（如 RS-232 接口、SPI 接口、以太网接口等）和显示设备（如显示屏等）3 类。

常见存储器概念辨析：RAM、SRAM、SDRAM、ROM、EPROM、EEPROM、Flash

存储器可以分为很多种类，其中根据掉电数据是否丢失可以分为 RAM（随机存取存储器）和 ROM（只读存储器），其中 RAM 的访问速度比较快，但掉电后数据会丢失，而 ROM 掉电后数据不会丢失。人们通常所说的内存即指系统中的 RAM。

RAM 又可分为 SRAM（静态存储器）和 DRAM（动态存储器）。SRAM 是利用双稳态触发器来保存信息的，只要不掉电，信息是不会丢失的。DRAM 是利用 MOS（金属氧化物半导体）电容存储电荷来储存信息，因此必须通过不停的给电容充电来维持信息，所以 DRAM 的成本、集成度、功耗等明显优于 SRAM。

而通常人们所说的 SDRAM 是 DRAM 的一种，它是同步动态存储器，利用一个单一的系统时钟同步所有的地址数据和控制信号。使用 SDRAM 不但能提高系统表现，还能简化设计、提供高速的数据传输。在嵌入式系统中经常使用。

EPROM、EEPROM 都是 ROM 的一种，分别为可擦除可编程 ROM 和电可擦除 ROM，但使用不是很方便。

Flash 也是一种非易失性存储器（掉电不会丢失），它擦写方便，访问速度快，已大大取代了传统的 EPROM 的地位。由于它具有和 ROM 一样掉电不会丢失的特性，因此很多人称其为 Flash ROM。

嵌入式操作系统从嵌入式发展的第 3 阶段起开始引入。嵌入式操作系统不仅具有通用操作系统的一般功能，如向上提供对用户的接口（如图形界面、库函数 API 等），向下提供与硬件设备交互的接口（硬件驱动程序等），管理复杂的系统资源，同时，它还在系统实时性、

硬件依赖性、软件固化性以及应用专用性等方面，具有更加鲜明的特点。

应用软件是针对特定应用领域，基于某一固定的硬件平台，用来达到用户预期目标的计算机软件。由于嵌入式系统自身的特点，决定了嵌入式应用软件不仅要求做到准确性、安全性和稳定性等方面需要，而且还要尽可能地进行代码优化，以减少对系统资源的消耗，降低硬件成本。

4.1.5 几种主流嵌入式操作系统分析

1．嵌入式 Linux

嵌入式 Linux（Embedded Linux）是指对标准 Linux 经过小型化裁剪处理之后，能够固化在容量只有几 KB 或者几 MB 字节的存储器芯片或者单片机中，是适合于特定嵌入式应用场合的专用 Linux 操作系统。在目前已经开发成功的嵌入式系统中，大约有一半使用的是 Linux。这与它自身的优良特性是分不开的。

嵌入式 Linux 同 Linux 一样，具有低成本、多种硬件平台支持、优异的性能和良好的网络支持等优点。另外，为了更好地适应嵌入式领域的开发，嵌入式 Linux 还在 Linux 基础上做了部分改进，如下所示。

- 改善的内核结构

Linux 内核采用的是整体式结构（Monolithic），整个内核是一个单独的、非常大的程序，这样虽然能够使系统的各个部分直接沟通，提高系统响应速度，但与嵌入式系统存储容量小、资源有限的特点不相符合。因此，在嵌入式系统经常采用的是另一种称为微内核（Microkernel）的体系结构，即内核本身只提供一些最基本的操作系统功能，如任务调度、内存管理、中断处理等，而类似于文件系统和网络协议等附加功能则运行在用户空间中，并且可以根据实际需要进行取舍。这样就大大减小了内核的体积，便于维护和移植。

- 提高的系统实时性

由于现有的 Linux 是一个通用的操作系统，虽然它也采用了许多技术来加快系统的运行和响应速度，但从本质上来说并不是一个嵌入式实时操作系统。因此，利用 Linux 作为底层操作系统，在其上进行实时化改造，从而构建出一个具有实时处理能力的嵌入式系统，如 RT-Linux 已经成功地应用于航天飞机的空间数据采集、科学仪器测控和电影特技图像处理等各种领域。

嵌入式 Linux 同 Linux 一样，也有众多的版本，其中不同的版本分别针对不同的需要在内核等方面加入了特定的机制。嵌入式 Linux 的主要版本如表 4.1 所示。

表 4.1 嵌入式 Linux 主要版本

版　本	简 单 介 绍
μCLinux	开放源码的嵌入式 Linux 的典范之作。它主要是针对目标处理器没有存储管理单元 MMU，其运行稳定，具有良好的移植性和优秀的网络功能，对各种文件系统有完备的支持，并提供标准丰富的 API
RT-Linux	由美国墨西哥理工学院开发的嵌入式 Linux 硬实时操作系统。它已有广泛的应用
Embedix	根据嵌入式应用系统的特点重新设计的 Linux 发行版本。它提供了超过 25 种的 Linux

	系统服务，包括 Web 服务器等。此外还推出了 Embedix 的开发调试工具包、基于图形界面的浏览器等。可以说，Embedix 是一种完整的嵌入式 Linux 解决方案
XLinux	采用了“超字元集”专利技术，使 Linux 内核不仅能与标准字符集相容，还涵盖了 12 个国家和地区的字符集。因此，XLinux 在推广 Linux 的国际应用方面有独特的优势
PoketLinux	它可以提供跨操作系统并且构造统一的、标准化的和开放的信息通信基础结构，在此结构上实现端到端方案的完整平台
红旗嵌入式 Linux	由北京中科院红旗软件公司推出的嵌入式 Linux，它是国内做得较好的一款嵌入式操作系统。目前，中科院计算机研究所自行开发的开放源码的嵌入式操作系统——Easy Embedded OS（EEOS）也已经开始进入实用阶段了

为了不失一般性，本书说所用的嵌入式 Linux 是标准内核裁减的 Linux，而不是上表中的任何一种。

2．VxWorks

VxWorks 操作系统是美国 WindRiver 公司于 1983 年设计开发的一种嵌入式实时操作系统（RTOS），它是在当前市场占有率最高的嵌入式操作系统。VxWorks 的实时性做得非常好，其系统本身的开销很小，进程调度、进程间通信、中断处理等系统公用程序精练而有效，使得它们造成的延迟很短。另外 VxWorks 提供的多任务机制，对任务的控制采用了优先级抢占（Linux 2.6 内核也采用了优先级抢占的机制）和轮转调度机制，这充分保证了可靠的实时性，并使同样的硬件配置能满足更强的实时性要求。另外 VxWorks 具有高度的可靠性，从而保证了用户工作环境的稳定。同时，VxWorks 还有很完备强大的集成开发环境，这也大大方便了用户的使用。

但是，由于 VxWorks 的开发和使用都需要交高额的专利费，因此大大增加了用户的开发成本。同时，由于 VxWorks 的源码不公开，造成它部分功能的更新（如网络功能模块）滞后。

3．QNX

QNX 是业界公认的 X86 平台上最好的嵌入式实时操作系统之一，它具有独一无二的微内核实时平台，是建立在微内核和完全地址空间保护基础之上的，它同样具有实时性强、稳定可靠的优点。

4．Windows CE

WINDOWS CE 是微软开发的一个开放的、可升级的 32 位嵌入式操作系统，是基于掌上型电脑类的电子设备操作系统。它是精简的 Windows 95。Windows CE 的图形用户界面相当出色。Windows CE 具有模块化、结构化和基于 Win32 应用程序接口以及与处理器无关等特点。它不仅继承了传统的 Windows 图形界面，并且用户在 Windows CE 平台上可以使用 Windows 95/98 上的编程工具（如 Visual Basic、Visual++等）、也可以使用同样的函数、使用同样的界面风格，使绝大多数 Windows 上的应用软件只需简单的修改和移植就可以在 WindowsCE 平台上继续使用。但与 VxWorks 相同，WindowsCE 也是比较昂

贵的。

5．Palm OS

Paml OS 在 PDA 和掌上电脑有着很大的用户群。Plam OS 最明显的特点在精简，它的内核只有几千个字节，同时用户也可以方便地开发定制，具有较强的可操作性。

4.2 ARM 处理器硬件开发平台

4.2.1 ARM 处理器简介

ARM 是一类嵌入式微处理器，同时也是一个公司的名字。ARM 公司于 1990 年 11 月成立于英国剑桥，它是一家专门从事 16/32 位 RISC 微处理器知识产权设计的供应商。ARM 公司本身不直接从事芯片生产，而只是授权 ARM 内核，再给生产和销售半导体的合作伙伴，同时也提供基于 ARM 架构的开发设计技术。世界各大半导体生产商从 ARM 公司处购买其设计的 ARM 微处理器核，根据各自不同的应用领域，加入适当的外围电路，从而形成自己的 ARM 微处理器芯片进入市场。

ARM 公司从成立至今，在短短几十年的时间就占据了 75%的市场份额，如今，ARM 微处理器及技术的应用几乎已经深入到各个领域。采用 ARM 技术的微处理器现在已经遍及各类电子产品，汽车、消费娱乐、影像、工业控制、海量存储、网络、安保和无线等市场。到 2001 年就几乎已经垄断了全球 RISC 芯片市场，成为业界实际的 RISC 芯片标准。图 4.3 列举了使用 ARM 微处理器的公司名称。

图 4.3 采用 ARM 微处理器的公司

ARM 的成功，一方面得益于它独特的公司运作模式，另一方面，当然来自于 ARM 处理器自身的优良性能。ARM 处理器有如下特点。

- 体积小、低功耗、低成本、高性能。
- 支持 Thumb（16 位）/ARM（32 位）双指令集，能很好的兼容 8 位/16 位器件。
- 大量使用寄存器，指令执行速度更快。
- 大多数数据操作都在寄存器中完成。
- 寻址方式灵活简单，执行效率高。
- 指令长度固定。

小知识

常见的 CPU 指令集分为 CISC 和 RISC 两种。

CISC（Complex Instruction Set Computer）是“复杂指令集”。自 PC 机诞生以来，32 位以前的处理器都采用 CISC 指令集方式。由于这种指令系统的指令不等长，因此指令的数目非常多，编程和设计处理器时都较为麻烦。但由于基于 CISC 指令架构系统设计的软件已经非常普遍了，所以包括 Intel、AMD 等众多厂商至今使用的仍为 CISC。

RISC（Reduced Instruction Set Computing）是“精简指令集”。研究人员在对 CISC 指令集进行测试时发现，各种指令的使用频度相当悬殊，其中最常使用的是一些比较简单的指令，它们仅占指令总数的 20%，但在程序中出现的频度却占 80%。RISC 正是基于这种思想提出的。采用 RISC 指令集的微处理器处理能力强，并且还通过采用超标量和超流水线结构，大大增强并行处理能力。

4.2.2 ARM 体系结构简介

1．ARM 微处理器工作状态

ARM 微处理器的工作状态一般有两种，并可在两种状态之间切换。

- 第一种为 ARM 状态，此时处理器执行 32 位的字对齐的 ARM 指令。
- 第二种为 Thumb 状态，此时处理器执行 16 位的、半字对齐的 Thumb 指令。

2．ARM 体系结构的存储格式

- 大端格式：在这种格式中，字数据的高字节存储在低地址中，而字数据的低字节则存放在高地址中。
- 小端格式：与大端存储格式相反，在小端存储格式中，低地址中存放的是字数据的低字节，高地址存放的是字数据的高字节。

3．ARM 处理器模式

ARM 微处理器支持 7 种运行模式，分别如下。

- 用户模式（usr）：ARM 处理器正常的程序执行状态。
- 快速中断模式（fiq）：用于高速数据传输或通道处理。
- 外部中断模式（irq）：用于通用的中断处理。
- 管理模式（svc）：操作系统使用的保护模式。

- 数据访问终止模式（abt）：当数据或指令预取终止时进入该模式，可用于虚拟存储及存储保护。
- 系统模式（sys）：运行具有特权的操作系统任务。

4.2.3 ARM9 体系结构

1．ARM 微处理器系列简介

ARM 微处理器系列主要特点及其主要应用领域如表 4.2 所示。

表 4.2 ARM 微处理器系列

型　　号	主 要 特 点
ARM7	低功耗的 32 位 RISC 处理器，最适合用于对价位和功耗要求较高的消费类应用。 主要应用领域为：工业控制、Internet 设备、网络和调制解调器设备、移动电话等多种多媒体和嵌入式应用
ARM9	在高性能和低功耗特性方面提供最佳的性能。 主要应用领域为：无线设备、仪器仪表、安全系统、机顶盒、高端打印机、数字照相机和数字摄像机等
ARM9E	综合处理器，使用单一的处理器内核，提供了微控制器、DSP、Java 应用系统的解决方案，极大的减少了芯片的面积和系统的复杂程度。

续表

型　　号	主 要 特 点
ARM9E	主要应用领域为：下一代无线设备、数字消费品、成像设备、工业控制、存储设备和网络设备等
ARM10E	高性能、低功耗。由于采用了新的体系结构，与同等的 ARM9 器件相比较，在同样的时钟频率下，性能提高近 50%，同时，ARM10E 系列微处理器采用了两种先进的节能方式，使其功耗极低。 主要应用领域为：下一代无线设备、数字消费品、成像设备、工业控制、存储设备和网络设备等
SecurCore	专为安全需要而设计，带有灵活的保护单元，以确保操作系统和应用数据的安全。 主要应用领域为：对安全性要求较高的应用产品及应用系统，如电子商务、电子政务、电子银行业务、网络和认证系统等领域
StrongARM	融合了 Intel 公司的设计和处理技术以及 ARM 体系结构的电源效率，采用在软件上兼容 ARMv4 体系结构、同时采用具有 Intel 技术优点的体系结构。 主要应用领域为：便携式通讯产品和消费类电子产品

2．ARM9 主要特点

ARM 处理器凭借它的低功耗、高性能等特点，被广泛应用于个人通信等嵌入式领域，而 ARM7 也曾在中低端手持设备中占据了一席之地。然而，ARM7 的处理性能逐渐无法满足人们日益增长的高性能功能需求的处理，它开始退出主流应用领域，取而代之的是性能更加强大的 ARM9 系列处理器。

新一代的 ARM9 处理器，通过全新的设计，能够达到两倍以上于 ARM7 处理器的处理能力。它的主要特点如下所述。

（1）5 级流水线

ARM7 处理器采用的 3 级流水线设计，而 ARM9 则采用 5 级流水线设计，如图 4.4 所示。

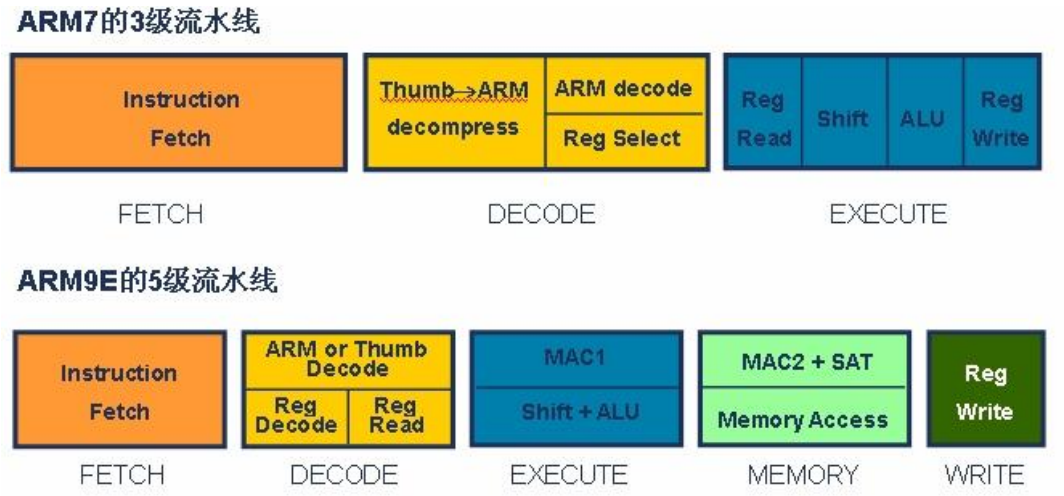


图 4.4　ARM7 与 ARM9 流水线比较

通过使用 5 级流水线机制，在每一个时钟周期内可以同时执行 5 条指令。这样就大大提高了处理性能。在同样的加工工艺下，ARM9 处理器的时钟频率是 ARM7 的 1.8～2.2 倍。

（2）采用哈佛结构

首先读者需要了解什么叫哈佛结构？在计算机中，根据计算机的存储器结构及其总线连接形式，计算机系统可以被分为冯・诺依曼结构和哈佛结构，其中冯・诺依曼结构共用数据存储空间和程序存储空间，它们共享存储器总线，这也是以往设计时常用的方式；而哈佛结构则具有分离的数据和程序空间及分离的访问总线。所以哈佛结构在指令执行时，取址和取数可以并行，因此具有更高的执行效率。ARM9 采用的就是哈佛结构，而 ARM7 采用的则是冯・诺依曼结构。如图 4.5 和图 4.6 分别体现了冯・诺依曼结构和哈佛结构的数据存储方式。

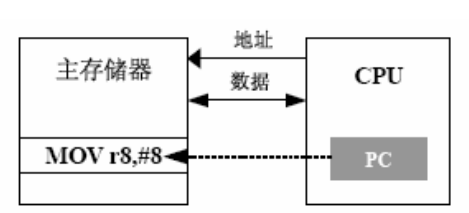


图 4.5　冯・诺依曼结构

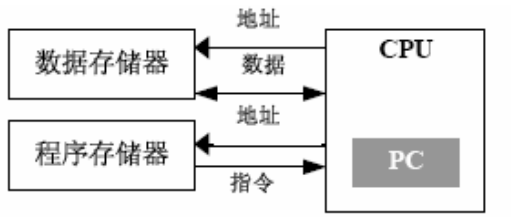


图 4.6　哈佛结构

由于在 RISC 架构的处理器中，程序中大约有 30%的指令是 Load-Store 指令，而采用哈佛结构大大提升了这两个指令的执行速度，因此对提高系统效率的贡献是非常明显的。

（3）高速缓存和写缓存的引入

由于在处理器中，一般处理器速度远远高于存储器访问速度，那么，如果存储器访问成

为系统性能的瓶颈，则处理器再快也都毫无作用。在这种情况下，高速缓存（Cache）和写缓存（Write Buffer）可以很好地解决这个问题，它们存储了最近常用的代码和数据，以供 CPU 快速存储，如图 4.7 所示：

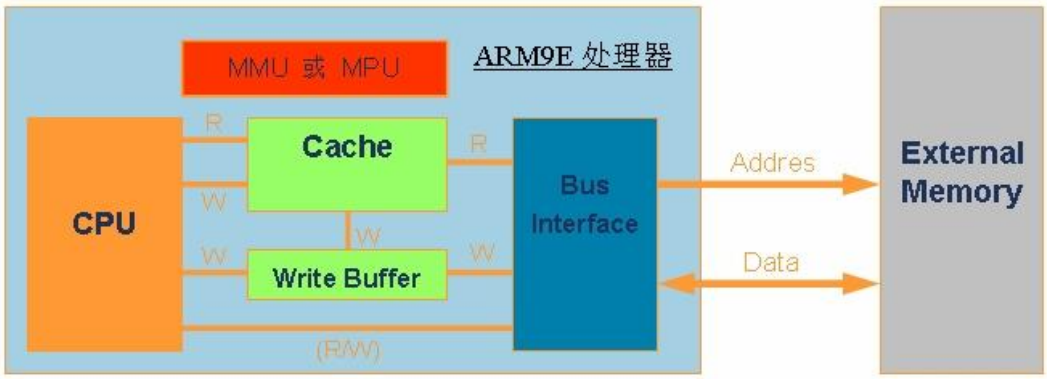


图 4.7　ARM9 的高速缓存和读缓存

（4）支持 MMU

MMU 是内存管理单元，它把内存以“页（page）”为单位来进行处理。一页内存是指一个具有一定大小的连续的内存块，通常为 4096B 或 8192B。操作系统为每个正在运行的程序建立并维护一张被称为进程内存映射（Process Memory Map）的表，表中记录了程序可以存取的所有内存页以及它们的实际位置。

每当程序存取一块内存时，它会把相应的虚拟地址（virtual address）传送给 MMU，而 MMU 会在 PMM 中查找这块内存的实际位置，也就是物理地址（physical address），物理地址可以在内存中或磁盘上的任何位置。如果程序要存取的位置在磁盘上，就必须把包含该地址的页从磁盘上读到内存中，并且必须更新 PMM 以反映这个变化（这被称为 pagefault，即页错）。MMU 的实现过程如图 4.8 所示。

只有拥有了 MMU 才能真正实现内存保护。例如当 A 进程的程序试图直接访问属于 B 进程的虚拟地址中的数据，那么 MMU 会产生一个异常（Exception）来阻止 A 的越界操作。这样，通过内存保护，一个进程的失败并不会影响其他进程的运行，从而增强了系统的稳定性，如图 4.9 所示。ARM9 也正是因此拥有了 MMU，比 ARM7 有了更强的稳定性和可靠性。

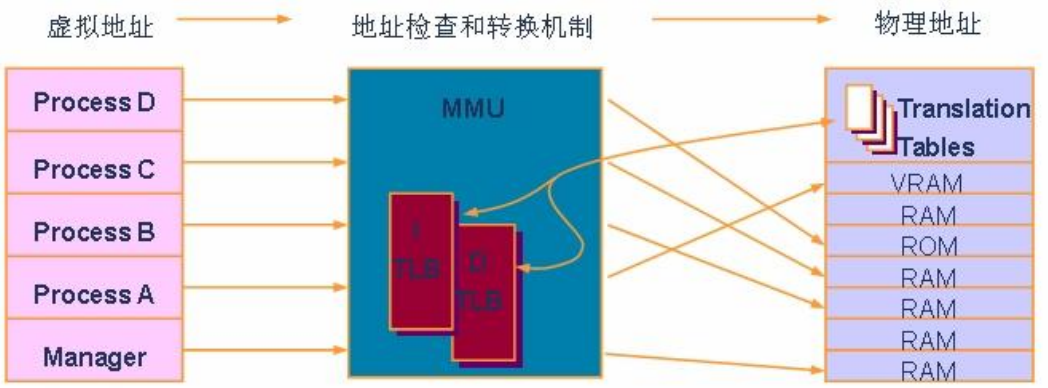


图 4.8　MMU 的实现过程

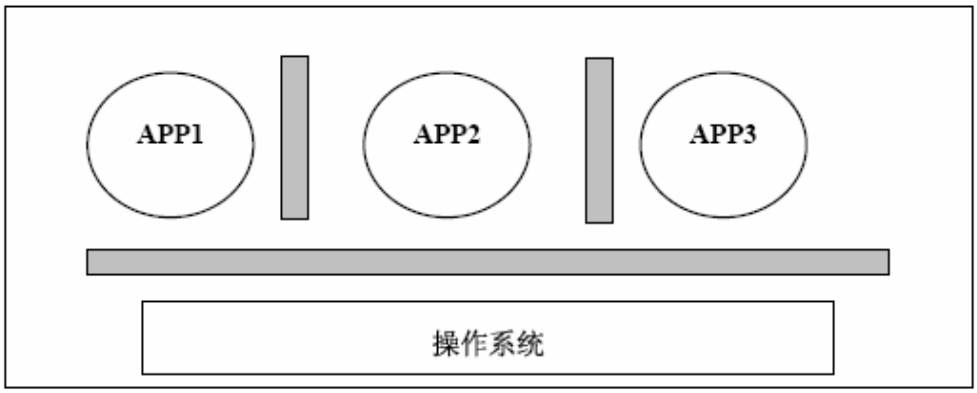


图 4.9　内存保护示意图

4.2.4　S3C2410 处理器详解

本书所采用的硬件平台是深圳优龙科技有限公司的开发板 FS2410（如图 4.10 所示），它的中央处理器是三星公司的 S3C2410X。S3C2410X 是使用 ARM920T 核、采用 0.18um 工艺 CMOS 标准宏单元和存储编译器开发而成的。由于采用了由 ARM 公司设计的 16/32 位 ARM920T RISC 处理器，因此 S3C2410X 实现了 MMU 和独立的 16KB 指令和 16KB 数据哈佛结构的缓存，且每个缓存均为 8 个字长度的流水线。它的低功耗、精简而出色的全静态设计特别适用于对成本和功耗敏感的领域。

S3C2410X 提供全面的、通用的片上外设，大大降低系统的成本，下面列举了 S3C2410X 的主要片上功能。

- 1.8V ARM920T 内核供电，1.8V/2.5V/3.3V 存储器供电；
- 16KB 指令和 16KB 数据缓存的 MMU 内存管理单元；
- 外部存储器控制（SDRAM 控制和芯片选择逻辑）；
- 提供 LCD 控制器（最大支持 4K 色的 STN 或 256K 色 TFT 的 LCD），并带有 1 个通道的 LCD 专用 DMA 控制器；
- 提供 4 通道 DMA，具有外部请求引脚；

- 提供 3 通道 UART（支持 IrDA1.0，16 字节发送 FIFO 及 16 字节接收 FIFO）/2 通道 SPI 接口；
- 提供 1 个通道多主 IIC 总线控制器/1 通道 IIS 总线控制器；
- 兼容 SD 主机接口 1.0 版及 MMC 卡协议 2.11 版；
- 提供 2 个主机接口的 USB 口/1 个设备 USB 口（1.1 版本）；
- 4 通道 PWM 定时器/1 通道内部计时器；

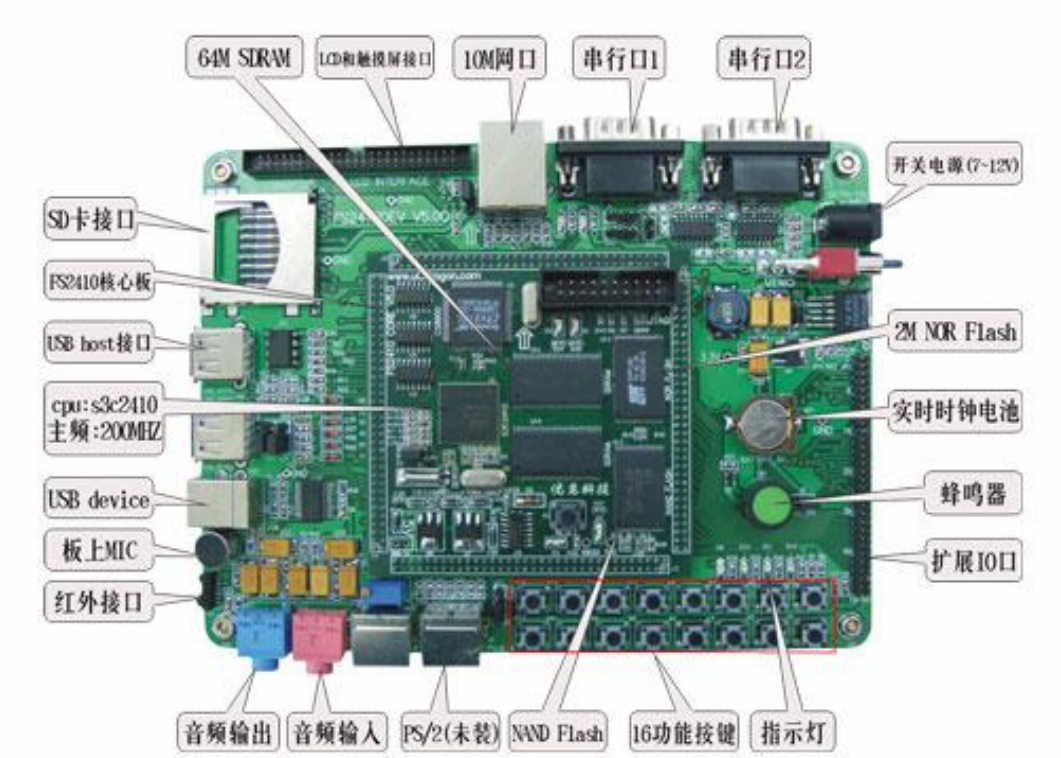


图 4.10　优龙 FS2410 开发板实物图

- 提供看门狗定时器；
- 提供 117 个通用 I/O 口/24 通道外部中断源；
- 提供电源控制不同模式：正常、慢速、空闲及电源关闭模式；
- 提供带触摸屏接口的 8 通道 10 位 ADC；
- 提供带日历功能的实时时钟控制器（RTC）；
- 具有 PLL 的片上时钟发生器。

S3C2410X 系统结构图如图 4.11 所示。

下面依次对 S3C2410X 的系统管理器、NAND Flash 引导装载器、缓冲存储器、时钟和电源管理及中断控制进行简要讲解，要注意，其中所有模式的选择都是通过相关寄存器特定值的设定来实现的，因此，当读者需要对此进行修改时，请参阅三星公司提供 S3C2410X 用户手册。

1．系统管理器

S3C2410X 支持小/大端模式，它将系统的存储空间分为 8 个组（bank），其中每个 bank 有 128MB，总共为 1GB。每个组可编程的数据总线宽度为 8/16/32 位，其中 bank0~bank5 具有固定的 bank 起始地址和结束地址，用于 ROM 和 SRAM。而 bank6 和 bank7 是大小可变的，用于 ROM、SRAM 或 SDRAM。这里，所有的存储器 bank 都具有可编程的操作周期，并且支持掉电时的 SDRAM 自刷新模式和多种类型的引导 ROM。

2．NAND Flash 引导装载器

S3C2410X 支持从 NAND flash 存储器启动，其中，开始的 4KB 为内置缓冲存储器，它

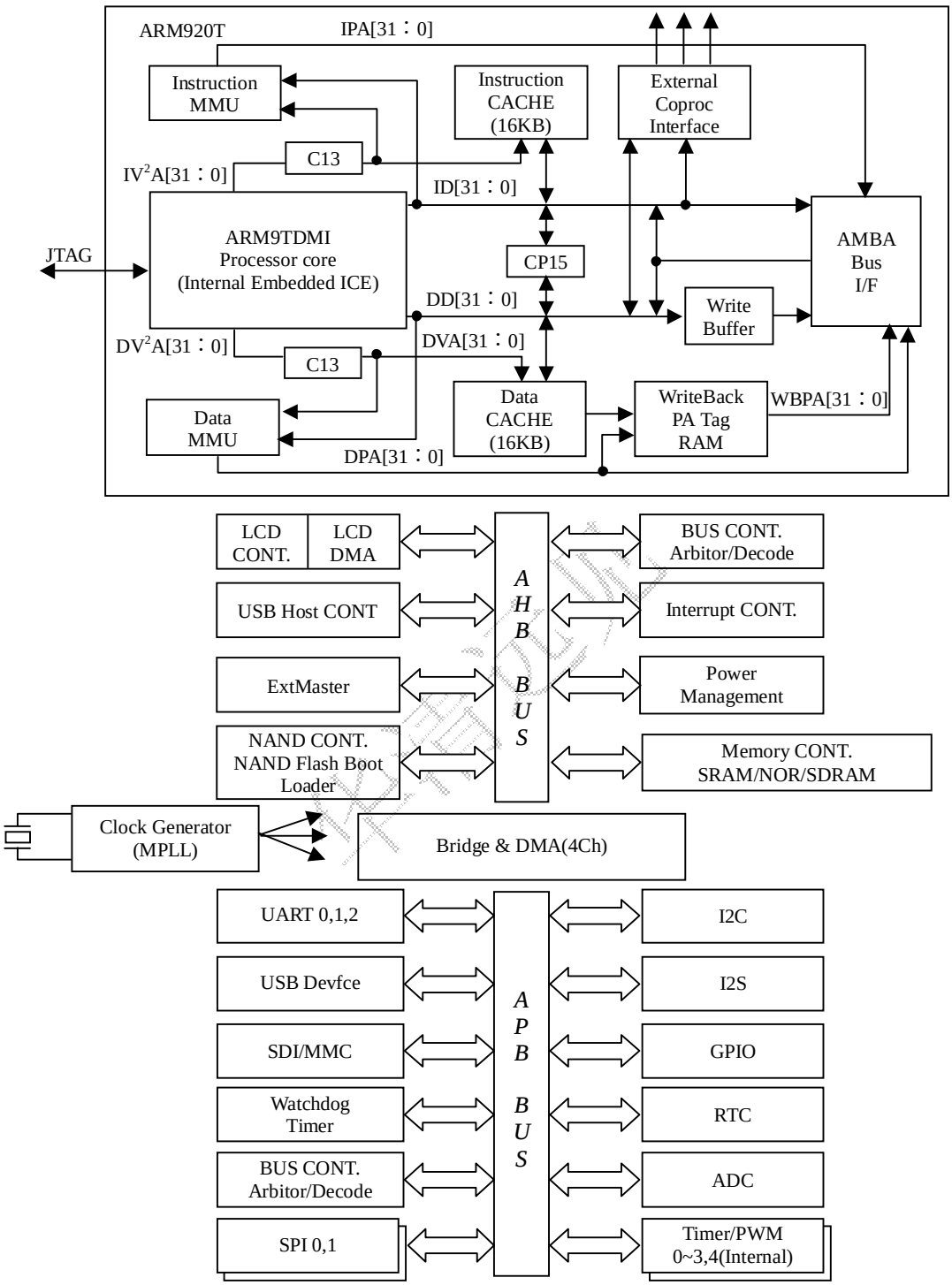


图 4.11　S3C2410X 系统结构图

在启动时将被转载（装在 or 转载）到 SDRAM 中并执行引导，之后该 4KB 可以用作其他

用途。

小知识

Flash 是一种非易失闪存技术。Intel 于 1988 年首先开发出 NOR Flash 技术之后，彻底改变了原先由 EPROM 和 EEPROM 一统天下的局面。紧接着，1989 年东芝公司发布了 NAND Flash 结构，强调降低每比特的成本、更高的性能，并且像磁盘一样可以通过接口轻松升级。

NOR Flash 的特点是芯片内执行（Execute In Place），这样应用程序可以直接在 flash 闪存内运行，而不必再把代码读到系统 RAM 中。NOR Flash 的传输效率很高，在 1～4MB 的小容量时具有很高的成本效益，但是很低的写入和擦除速度大大影响了它的性能。

NAND Flash 结构能提供极高的单元密度，可以达到高存储密度，NAND 读和写操作采用 512 字节的块，单元尺寸几乎是 NOR 器件的一半，同时由于生产过程更为简单，大大降低了生产的成本。NAND 闪存中每个块的最大擦写次数是 100 万次，是 NOR Flash 的 10 倍，这些都使得 NAND Flash 越来越受到人们的欢迎。

同时，S3C2410X 也支持从外部 nGCS0 片选的 NorFlash 启动，如在优龙的开发板上将 JP1 跳线去掉就可从 NorFlash 启动（默认从 NAND Flash 启动）。在这两种启动模式下，各片选的存储空间分配是不同的，如图 4.12 所示。

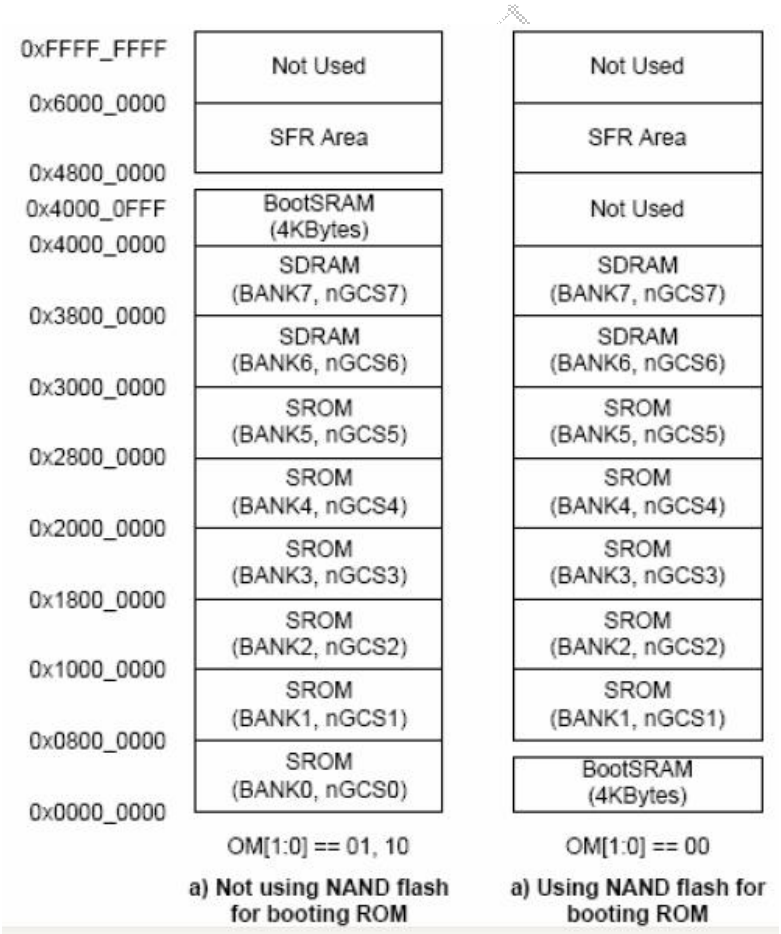


图 4.12　S3C2410 两种启动模式地址映射

3．缓冲存储器

S3C2410X 是带有指令缓存（16KB）和数据缓存（16KB）的联合缓存装置，一个缓冲区能够保持 16 字的数据和 4 个地址。

4．时钟和电源管理

S3C2410X 采用独特的时钟管理模式，它具有 PLL（相位锁定环路，用于稳定频率）的芯片时钟发生器，而在此，PLL 又分为 UPLL 和 MPLL。其中 UPLL 时钟发生器用于主/从 USB 操作，MPLL 时钟发生器用于产生主时钟，使其能在在极限频率 203MHz（1.8V）上运行。

S3C2410X 的电源管理模式又分为正常，慢速，空闲和掉电 4 种模式。其中慢速模式为不带 PLL 的低频时钟模式，空闲模式始终为 CPU 停止模式，掉电模式为所有外围设备全部掉电仅内核电源供电模式。

另外，S3C2410X 对片内的各个部件采用独立的供电方式。

- 1.8V 的内核供电。
- 3.3V 的存储器独立供电（通常对 SDRAM 采用 3.3V，对移动 SDRAM 采用 1.8/2.5V）。
- 3.3V 的 VDDQ。
- 3.3V 的 I/O 独立供电。

由于在嵌入式中电源管理非常关键，它直接涉及到功耗等各方面的系统性能，而 S3C2410X 的电源管理中独立的供电方式和多种模式可以有效地处理系统的不同状态，从而达到最优的配置。

5．中断控制

中断处理在嵌入式的开发中非常重要，尤其对于从单片机转入到嵌入式的读者来说，对比单片机中简单的中断模式而言，ARM 中的中断处理要复杂得多。如果读者无相关基础，建议先熟悉相关的基础概念再进行下一步学习。

首先给出了一般的中断处理流程，如图 4.13 所示。

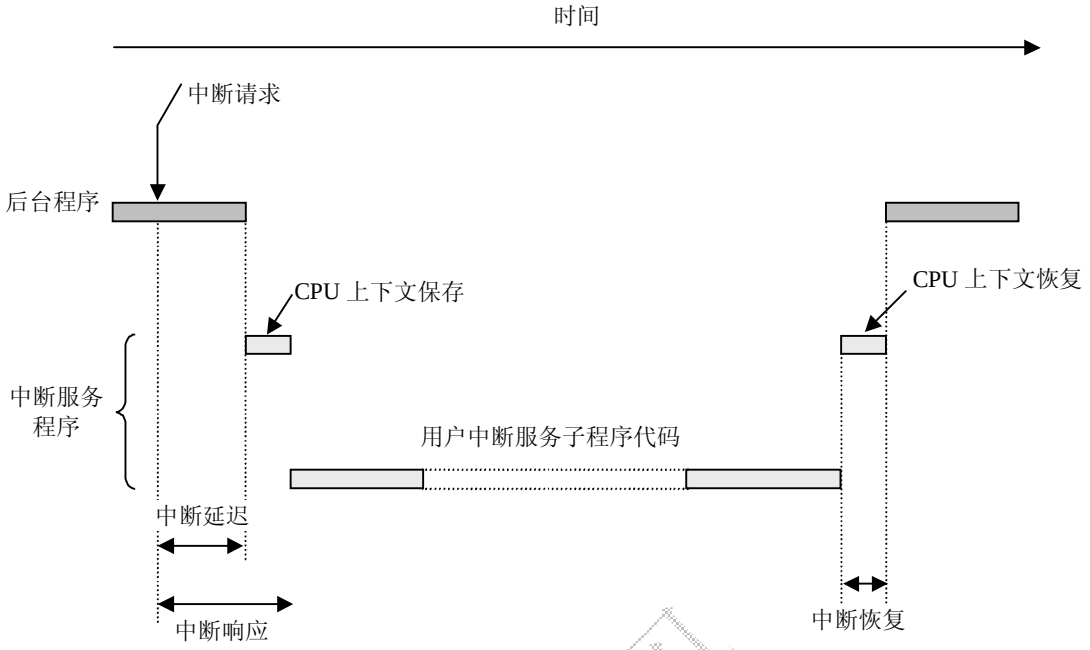


图 4.13　一般中断处理流程

S3C2410X 包括 55 个中断源，其中有 1 个看门狗定时器中断、5 个定时器中断、9 个通用异步串行口中断、24 个外部中断、4 个 DMA 中断、2 个 RTC（实时时钟控制器）中断、2 个 USB 中断、1 个 LCD 中断和 1 个电池故障。其中，对外部中断源具有电平/边沿两种触发模式。另外，对于非常紧急的中断可以支持使用快速中断请求（FIQ）。

S3C2410X 的中断处理流程（该图摘自 S3C2410X 用户手册）如图 4.14 所示。

图中的 SUBSRCPND、SRCPND、SUBMASK、MASK 和 MODE 都代表寄存器，其中 SUBSRCPND 和 SRCPND 寄存器表明有哪些中断被触发了和是否正在等待处理（pending）；SUBMASK（INTSUBMSK 寄存器）和 MASK（INTMSK 寄存器）用于屏蔽某些中断。

图中的“Request sources（with sub –register）”表示的是 INT_RXD0、INT_TXD0 等 11 个中断源，它们不同于“Request sources（without sub –register）”的操作在于：

（1）“Request sources（without sub –register）”中的中断源被触发之后，SRCPND 寄存器中相应位被置 1，如果此中断没有被 INTMSK 寄存器屏蔽、或者是快中断（FIQ）的话，它将被进一步处理。

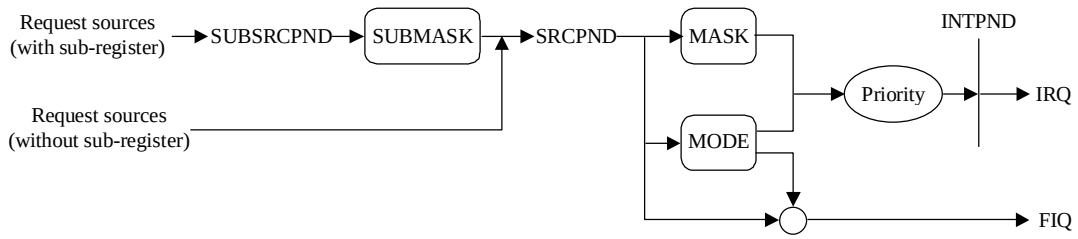


图 4.14　S3C2410X 中断处理流程

（2）对于“Request sources（with sub –register）”中的中断源被触发之后，SUBSRCPND 寄存器中的相应位被置 1，如果此中断没有被 SUBMSK 寄存器屏蔽的话，它在 SRCPND 寄存器中的相应位也被置 1。在此之后的两者的处理过程一样了。

接下来，在 SRCPND 寄存器中，被触发的中断的相应位被置 1，等待处理。

（1）如果被触发的中断中有快中断（FIQ）——MODE（INTMOD 寄存器）中为 1 的位对应的中断，则 CPU 的 FIQ 中断函数被调用。注意：FIQ 只能分配一个，即 INTMOD 中只能有一位被设为 1。

（2）对于一般中断 IRQ，可能同时有几个中断被触发，未被 INTMSK 寄存器屏蔽的中断经过比较后，选出优先级最高的中断，然后 CPU 调用 IRQ 中断处理函数。中断处理函数可以通过读取 INTPND（标识最高优先级的寄存器）寄存器来确定中断源是哪个，也可以读 INTOFFSET 寄存器来确定中断源。

4.3 嵌入式软件开发流程

4.3.1 嵌入式系统开发概述

由嵌入式系统本身的特性所影响，嵌入式系统开发与通用系统的开发有很大的区别。嵌入式系统的开发主要分为系统总体开发、嵌入式硬件开发和嵌入式软件开发 3 大部分，其总体流程图如图 4.15 所示。

在系统总体开发中，由于嵌入式系统与硬件依赖程序非常紧密，往往某些需求只能通过特定的硬件才能实现，因此需要进行处理器选型，以更好地满足产品的需求。另外，对于有些硬件和软件都可以实现的功能，就需要在成本和性能上做出抉择。往往通过硬件实现会增加产品的成品，但能大大提高产品的性能和可靠性。

再次，开发环境的选择对于嵌入式系统的开发也有很大的影响。这里的开发环境包括嵌入式操作系统的选择以及开发工具的选择等。本书在 4.1.5 节对各种不同的嵌入式操作系统进行了比较，读者可以以此为依据进行相关的选择。比如，对开发成本和进度限制较大的产品可以选择嵌入式 Linux，对实时性要求非常高的产品可以选择 Vxworks 等。

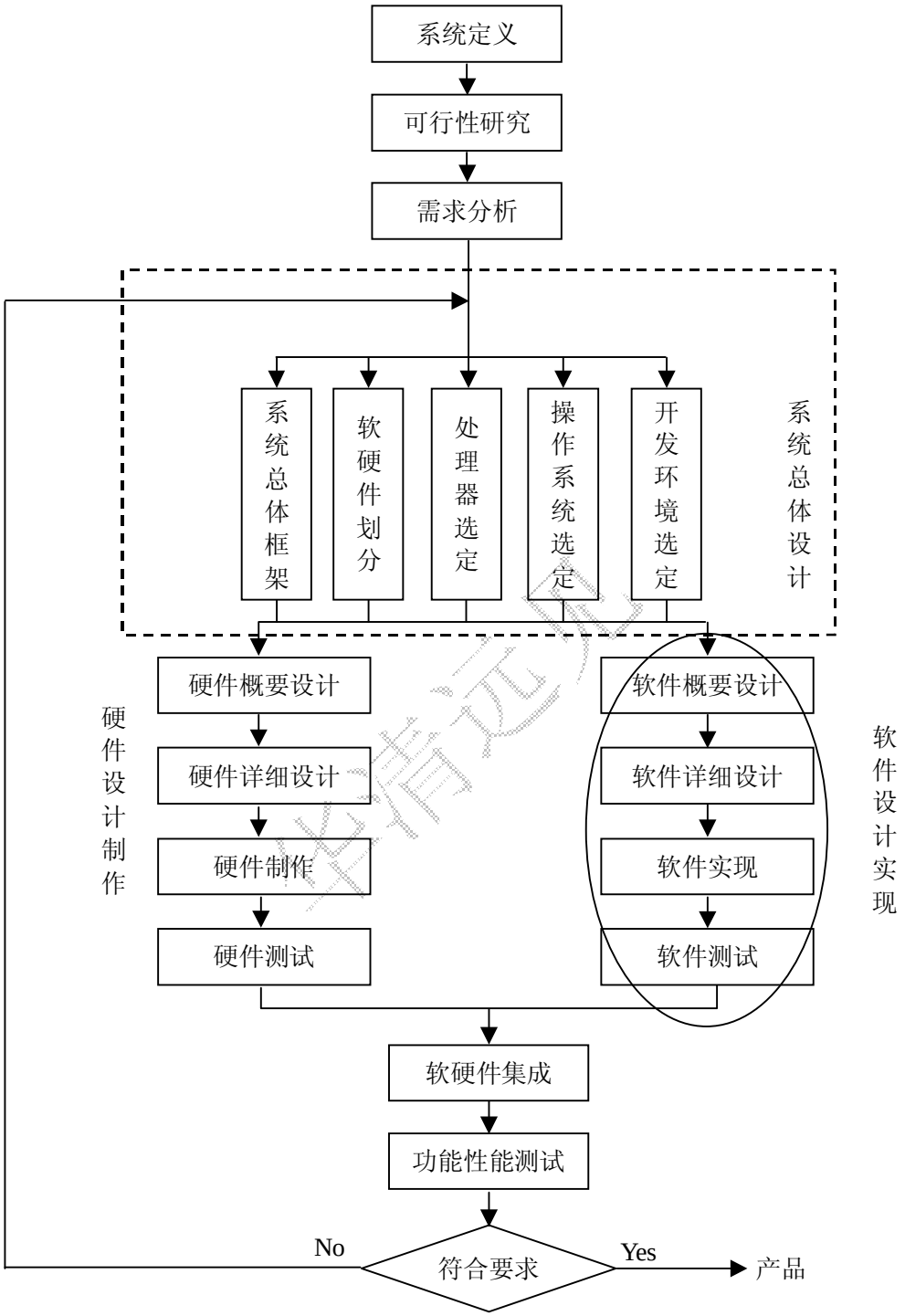


图 4.15　嵌入式系统开发流程图

由于本书主要讨论嵌入式软件的应用开发，因此对硬件开发不做详细讲解，而主要讨论嵌入式软件开发的流程。

4.3.2 嵌入式软件开发概述

嵌入式软件开发总体流程为图 4.15 中“软件设计实现”部分所示，它同通用计算机软件开发一样，分为需求分析、软件概要设计、软件详细设计、软件实现和软件测试。其中嵌入式软件需求分析与硬件的需求分析合二为一，故没有分开画出。

由于在嵌入式软件开发的工具非常多，为了更好地帮助读者选择开发工具，下面首先对嵌入式软件开发过程中所使用的工具做一简单归纳。

嵌入式软件的开发工具根据不同的开发过程而划分，如比在需求分析阶段，可以选择 IBM 的 Rational Rose 等软件，而在程序开发阶段可以采用 CodeWarrior（下面要介绍的 ADS 的一个工具）等，在调试阶段所用的 Multi-ICE 等。同时，不同的嵌入式操作系统往往会有配套的开发工具，比如 Vxworks 有集成开发环境 Tornado，WinCE 的集成开发环境 WinCE Platform 等。此外，不同的处理器可能还有针对的开发工具，比如 ARM 的常用集成开发工具 ADS 等。在这里，大多数软件都有比较高的使用费用，但也可以大大加快产品的开发进度，用户可以根据需求自行选择。图 4.16 是嵌入式开发的不同阶段的常用软件。

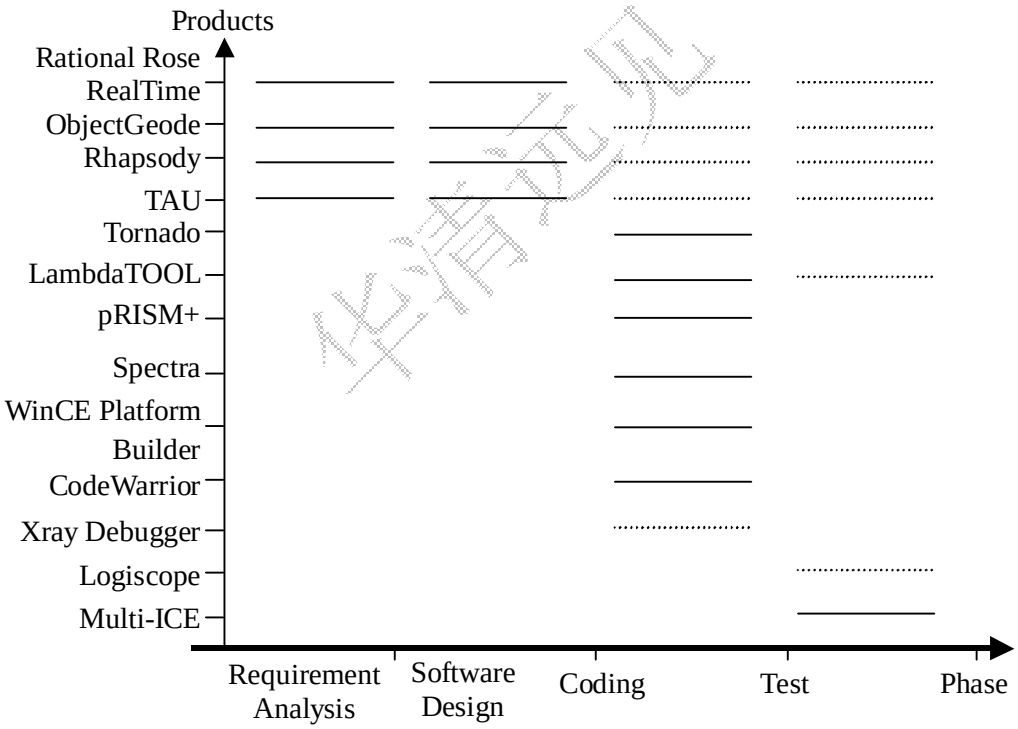


图 4.16 嵌入式开发不同阶段的常用软件

嵌入式系统的软件开发与通常软件开发的区别主要在于软件实现部分，其中又可以分为编译和调试两部分，下面分别对这两部分进行讲解。

1．交叉编译

嵌入式软件开发所采用的编译为交叉编译。所谓交叉编译就是在一个平台上生成可以在

另一个平台上执行的代码。在第 3 章中已经提到，编译的最主要的工作就在将程序转化成运行该程序的 CPU 所能识别的机器代码，由于不同的体系结构有不同的指令系统。因此，不同的 CPU 需要有相应的编译器，而交叉编译就如同翻译一样，把相同的程序代码翻译称不同的 CPU 对应语言。要注意的是，编译器本身也是程序，也要在与之对应的某一个 CPU 平台上运行。嵌入式系统交叉编译环境如图 4.17 所示。

图 4.17　交叉编译环境

小知识　与交叉编译相对应，平时常用的编译称本地编译。

这里一般把进行交叉编译的主机称为宿主机，也就是普通的通用计算机，而把程序实际的运行环境称为目标机，也就是嵌入式系统环境。由于一般通用计算机拥有非常丰富的系统资源、使用方便的集成开发环境和调试工具等，而嵌入式系统的系统资源非常紧缺，没有相关的编译工具，因此，嵌入式系统的开发需要借助宿主机（通用计算机）来编译出目标机的可执行代码。

由于编译的过程包括编译、链接等几个阶段，因此，嵌入式的交叉编译也包括交叉编译、交叉链接等过程，通常 ARM 的交叉编译器为 arm-elf-gcc，交叉链接器为 arm-elf-ld，交叉编译过程如图 4.18 所示。

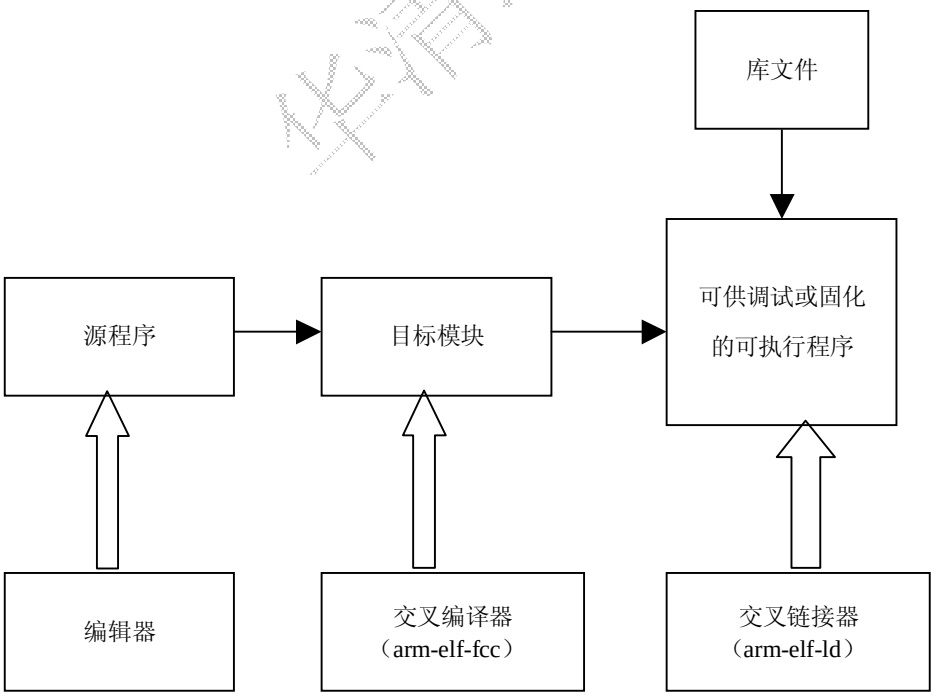


图 4.18　嵌入式交叉编译过程

2．交叉调试

嵌入式软件经过编译和链接后即进入调试阶段，调试是软件开发过程中必不可少的一个环节，嵌入式软件开发过程中的交叉调试与通用软件开发过程中的调试方式有很大的差别。在常见软件开发中，调试器与被调试的程序往往运行在同一台计算机上，调试器是一个单独运行着的进程，它通过操作系统提供的调试接口来控制被调试的进程。而在嵌入式软件开发中，调试时采用的是在宿主机和目标机之间进行的交叉调试，调试器仍然运行在宿主机的通用操作系统之上，但被调试的进程却是运行在基于特定硬件平台的嵌入式操作系统中，调试器和被调试进程通过串口或者网络进行通信，调试器可以控制、访问被调试进程，读取被调试进程的当前状态，并能够改变被调试进程的运行状态。

嵌入式系统的交叉调试有多种方法，主要可分为软件方式和硬件方式两种。它们一般都具有如下一些典型特点。

- 调试器和被调试进程运行在不同的机器上，调试器运行在 PC 机或者工作站上（宿主机），而被调试的进程则运行在各种专业调试板上（目标机）。
- 调试器通过某种通信方式（串口、并口、网络、JTAG 等）控制被调试进程。
- 在目标机上一般会具备某种形式的调试代理，它负责与调试器共同配合完成对目标机上运行着的进程的调试。这种调试代理可能是某些支持调试功能的硬件设备，也可能是某些专门的调试软件（如 gdbserver）。
- 目标机可能是某种形式的系统仿真器，通过在宿主机上运行目标机的仿真软件，整个调试过程可以在一台计算机上运行。此时物理上虽然只有一台计算机，但逻辑上仍然存在着宿主机和目标机的区别。

下面分别就软件调试桩方式和硬件片上调试两种方式进行详细介绍。

（1）软件方式

软件方式调试主要是通过插入调试桩的方式来进行的。调试桩方式进行调试是通过目标操作系统和调试器内分别加入某些功能模块，二者互通信息来进行调试。该方式的典型调试器有 Gdb 调试器。

GDB 的交叉调试器分为 GdbServer 和 GdbClient，其中的 GdbServer 就作为调试桩在安装在目标板上，GdbClient 就是驻于本地的 Gdb 调试器。它们的调试原理图如图 4.19 所示。

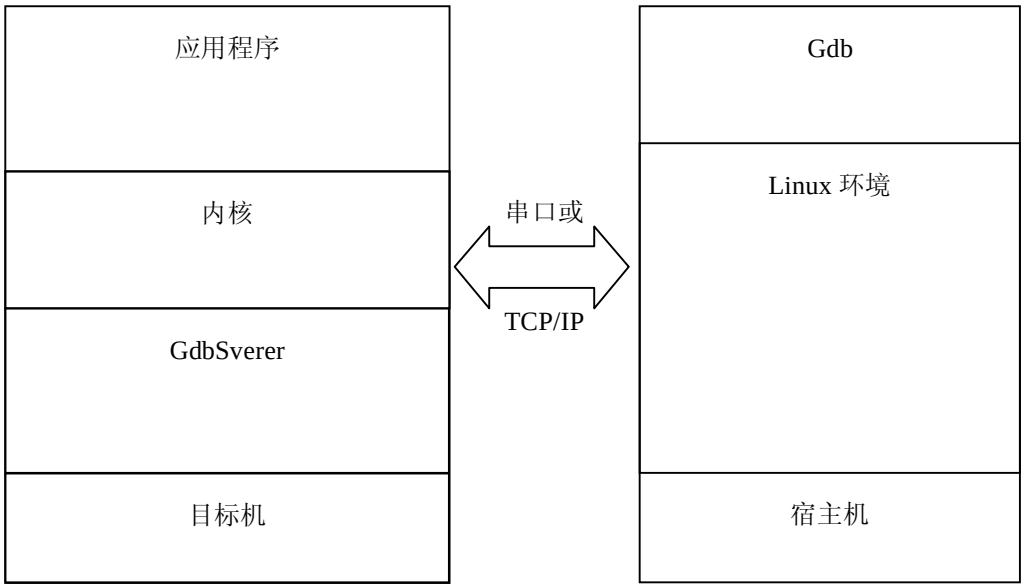


图 4.19　Gdb 远程调试原理图

Gdb 调试桩的工作流程。

- 首先，建立调试器（本地 Gdb）与目标操作系统的通信连接，可通过串口、网卡、并口等多种方式。
- 然后，在目标机上开启 Gdbserver 进程，并监听对应端口。
- 在宿主机上运行调试器 Gdb，这时，Gdb 就会自动寻找远端的通信进程，也就是 Gdbserver 的所在进程。
- 在宿主机上的 Gdb 通过 Gdbserver 请求对目标机上的程序发出控制命令。这时，Gdbserver 将请求转化为程序的地址空间或目标平台的某些寄存器的访问，这对于没有虚拟存储器的简单的嵌入式操作系统而言，是十分容易的。
- Gdbserver 把目标操作系统的所有异常处理转向通信模块，并告知宿主机上 Gdb 当前异常。
- 宿主机上的 Gdb 向用户显示被调试程序产生了哪一类异常。

这样就完成了调试的整个过程。这个方案的实质是用软件接管目标机的全部异常处理及部分中断处理，并在其中插入调试端口通信模块，与主机的调试器进行交互。但是它只能在目标机系统初始化完毕、调试通信端口初始化完成后才能起作用，因此，一般只能用于调试运行于目标操作系统之上的应用程序，而不宜用来调试目标操作系统的内核代码及启动代码。而且，它必须改变目标操作系统，因此，也就多了一个不用于正是发布的调试版。

（2）硬件调试

相对于软件调试而言，使用硬件调试器可以获得更强大的调试功能和更优秀的调试性能。硬件调试器的基本原理是通过仿真硬件的执行过程，让开发者在调试时可以随时了解到系统的当前执行情况。目前嵌入式系统开发中最常用到的硬件调试器是 ROMMonitor、ROMEmulator、In-CircuitEmulator 和 In-CircuitDebugger。

- 采用 ROMMonitor 方式进行交叉调试需要在宿主机上运行调试器，在目标机上运行

ROM 监视器（ROMMonitor）和被调试程序，宿主机通过调试器与目标机上的 ROM 监视器遵循远程调试协议建立通信连接。ROM 监视器可以是一段运行在目标机 ROM 上的可执行程序，也可以是一个专门的硬件调试设备，它负责监控目标机上被调试程序的运行情况，能够与宿主机端的调试器一同完成对应用程序的调试。

在使用这种调试方式时，被调试程序首先通过 ROM 监视器下载到目标机，然后在 ROM 监视器的监控下完成调试。

优点：ROM 监视器功能强大，能够完成设置断点、单步执行、查看寄存器、修改内存空间等各项调试功能。

确定：同软件调试一样，使用 ROM 监视器目标机和宿主机必须建立通信连接。

其原理图如图 4.20 所示。

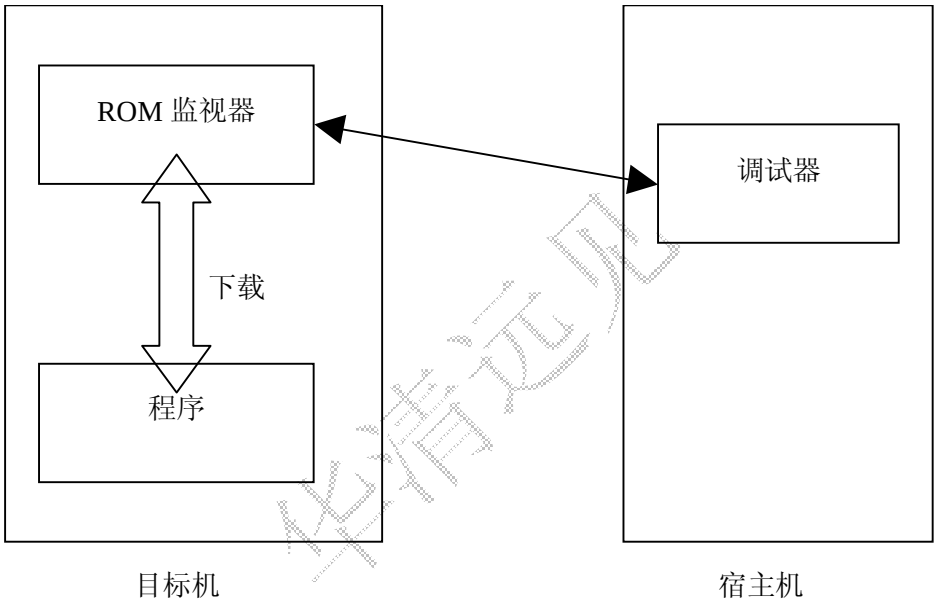


图 4.20　ROMMonitor 调试方式

- 采用 ROMEmulator 方式进行交叉调试时需要使用 ROM 仿真器，并且它通常被插入到目标机上的 ROM 插槽中，专门用于仿真目标机上的 ROM 芯片。

在使用这种调试方式时，被调试程序首先下载到 ROM 仿真器中，因此等效于下载到目标机的 ROM 芯片上，然后在 ROM 仿真器中完成对目标程序的调试。

优点：避免了每次修改程序后都必须重新烧写到目标机的 ROM 中。

缺点：ROM 仿真器本身比较昂贵，功能相对来讲又比较单一，只适应于某些特定场合。

其原理图如图 4.21 所示。

- 采用 In-CircuitEmulator（ICE）方式进行交叉调试时需要使用在线仿真器，它是目前最为有效的嵌入式系统的调试手段。它是仿照目标机上的 CPU 而专门设计的硬件，可以完全仿真处理器芯片的行为。仿真器与目标板可以通过仿真头连接，与宿主机可以通过串口、并口、网线或 USB 口等连接方式。由于仿真器自成体系，所以调试时既可以连接目标板，也可以不连接目标板。

在线仿真器提供了非常丰富的调试功能。在使用在线仿真器进行调试的过程中，可以按

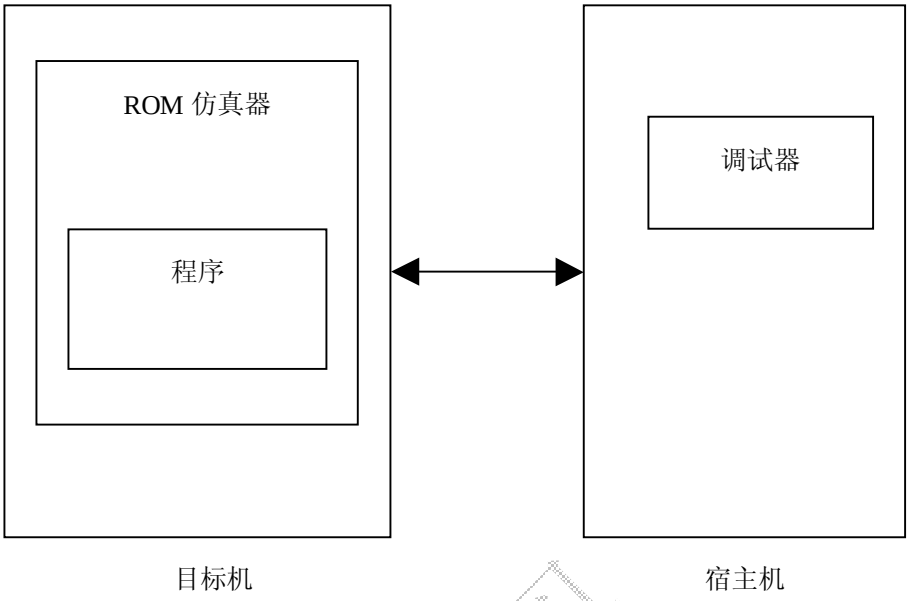


图 4.21　ROMEmulator 调试方式

顺序单步执行，也可以倒退执行，还可以实时查看所有需要的数据，从而给调试过程带来了很多的便利。嵌入式系统应用的一个显著特点是与现实世界中的硬件直接相关，并存在各种异变和事先未知的变化，从而给微处理器的指令执行带来各种不确定因素，这种不确定性在目前情况下只有通过在线仿真器才有可能发现。

优点：功能强大，软硬件都可做到完全实时在线调试。

确定：价格昂贵。

其原理图如图 4.22 所示。

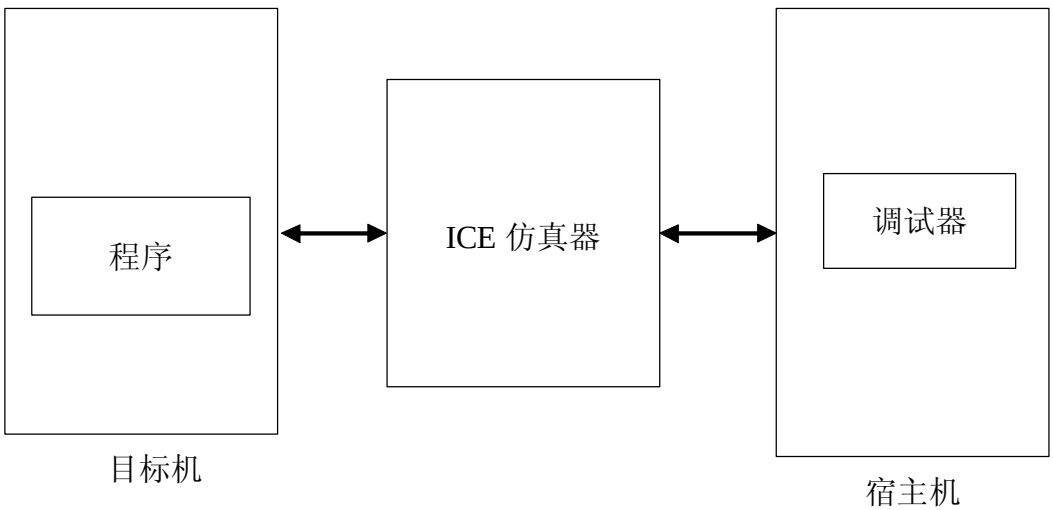


图 4.22　ICE 调试方式

- 采用 In-CircuitDebugger（ICD）方式进行交叉调试时需要使用在线调试器。由于 ICE 的价格非常昂贵，并且每种 CPU 都需要一种与之对应的 ICE，使得开发成本非常高。一个比较好的解决办法是让 CPU 直接在其内部实现调试功能，并通过在开发板上引出的调试端口发送调试命令和接收调试信息，完成调试过程。如在采用非常广泛的 ARM 处理器的 JTAG 端口技术就是由此而诞生的。

JTAG 是 1985 年指定的检测 PCB 和 IC 芯片的一个标准。1990 年被修改成为 IEEE 的一个标准，即 IEEE1149.1。JTAG 标准所采用的主要技术为边界扫描技术，它的基本思想就是在靠近芯片的输入输出管脚上增加一个移位寄存器单元。因为这些移位寄存器单元都分布在芯片的边界上（周围），所以被称为边界扫描寄存器（Boundary-Scan Register Cell）。

当芯片处于调试状态时候，这些边界扫描寄存器可以将芯片和外围的输入输出隔离开来。通过这些边界扫描寄存器单元，可以实现对芯片输入输出信号的观察和控制。对于芯片的输入脚，可通过与之相连的边界扫描寄存器单元把信号（数据）加载到该管脚中去；对于芯片的输出管脚，可以通过与之相连的边界扫描寄存器单元“捕获”（CAPTURE）该管脚的输出信号。这样，边界扫描寄存器提供了一个便捷的方式用于观测和控制所需要调试的芯片。

现在较为高档的微处理器都带有 JTAG 接口，包括 ARM7、ARM9、StrongARM、DSP 等，通过 JTAG 接口可以方便地对目标系统进行测试，同时，还可以实现的 Flash 的编程，是非常受人欢迎的。

优点：连接简单，成本低。

缺点：特性受制于芯片厂商。

其原理图如图 4.23 所示。

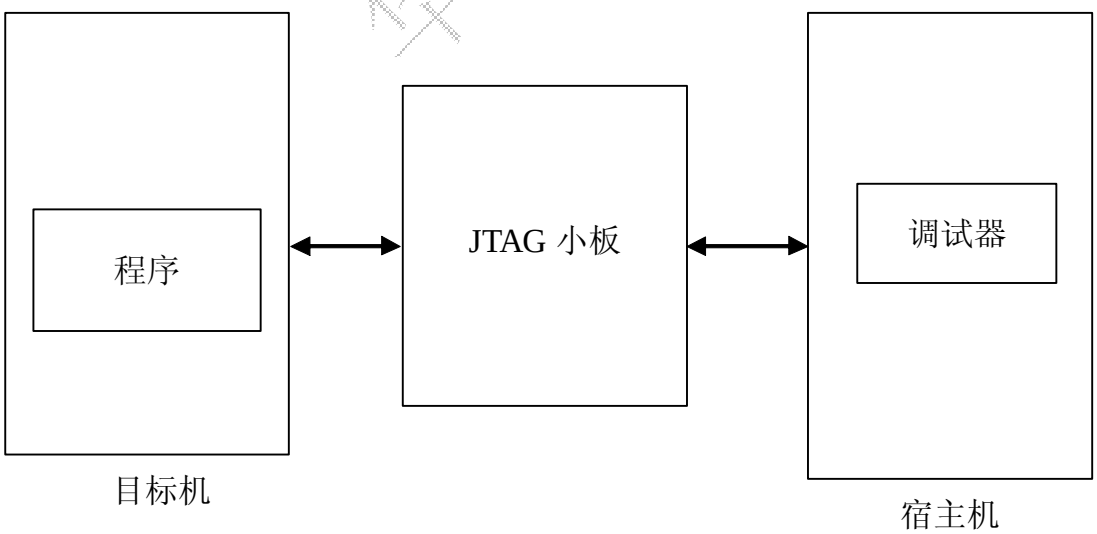


图 4.23　JTAG 调试方式

4.4 实验内容——使用 JTAG 烧写 NAND Flash

1．实验目的

通过使用 JTAG 烧写 Flash 的实验，了解嵌入式硬件环境，熟悉 JTAG 的使用，为今后的进一步学习打下良好的基础。（本书以优龙的 FS2410 及 Flash 烧写工具为例进行讲解，不同厂商的开发板都会提供相应的 Flash 烧写工具，并有相应的说明文档，请读者在了解基本原理之后查阅相关手册）

2．实验内容

（1）熟悉开发板的硬件布局。
（2）连接 JTAG 口。
（2）安装 giveio（用于烧写 Flash）驱动。
（3）打开 SJF2410_BIOS.BAT（Flash 烧写程序）进行烧写。

3．实验步骤

（1）熟悉开发板硬件设备请参阅本章 4.2 节的 FS2410 实物图。
（2）用 20 针的排线将 20 针的 JTAG 接口与 JTAG 小板的 JP3 接口相连。
（3）安装 givieo 驱动，如图 4.24 所示。

图 4.24　安装 givieo

单击 Install 按钮，出现如图 4.25 所示，就表明驱动安装成功。

图 4.25　givieo 驱动安装完成

（4）打开 SJF2410_BIOS.BAT，如图 4.26 所示。

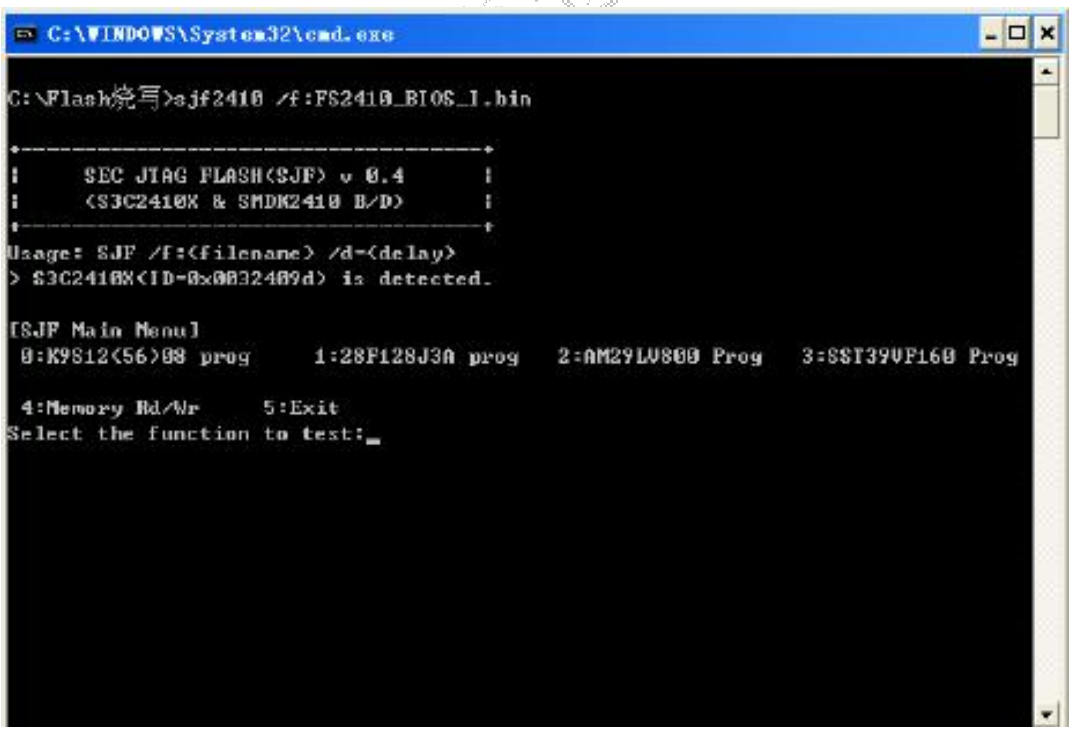


图 4.26　打开 SJF2410_BIOS.BAT

（5）在“Select the function to test:”输入“0”，表示对 K9S1208（FS2410 的 NAND Flash 的芯片型号）进行烧写，如图 4.27 所示。

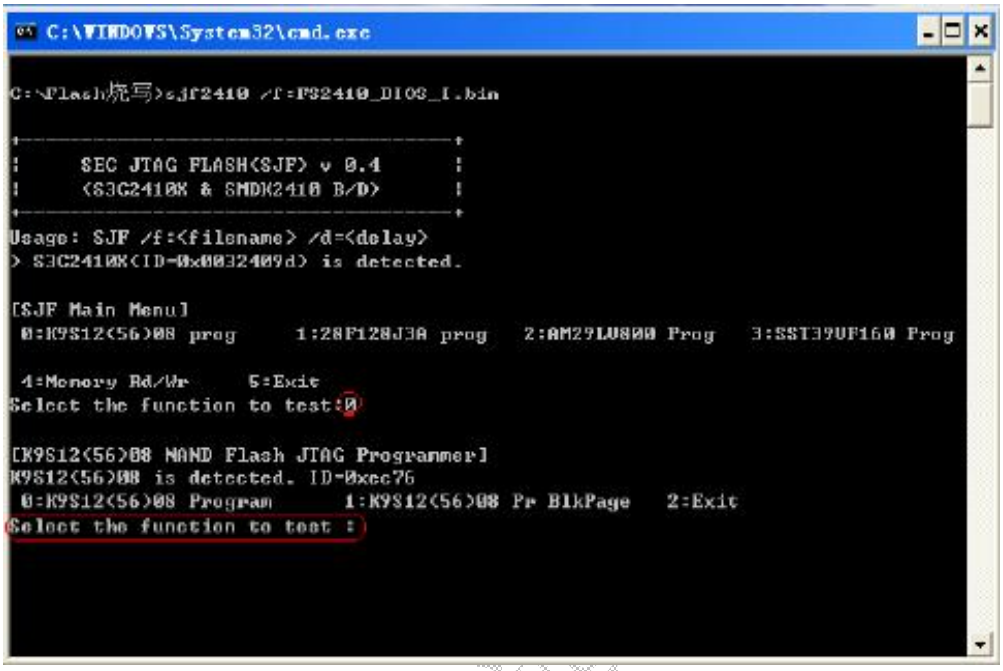


图 4.27　选择烧写对应芯片

（6）在接下来的“Select the function to test:”里输入“0”，表示烧写类型为程序。再在接下来的“Input the target block”里输入希望的偏移地址，在此处写为“0”，如图 4.28 所示。

（7）接下来，Flash 完成烧写过程，如图 4.29 所示。

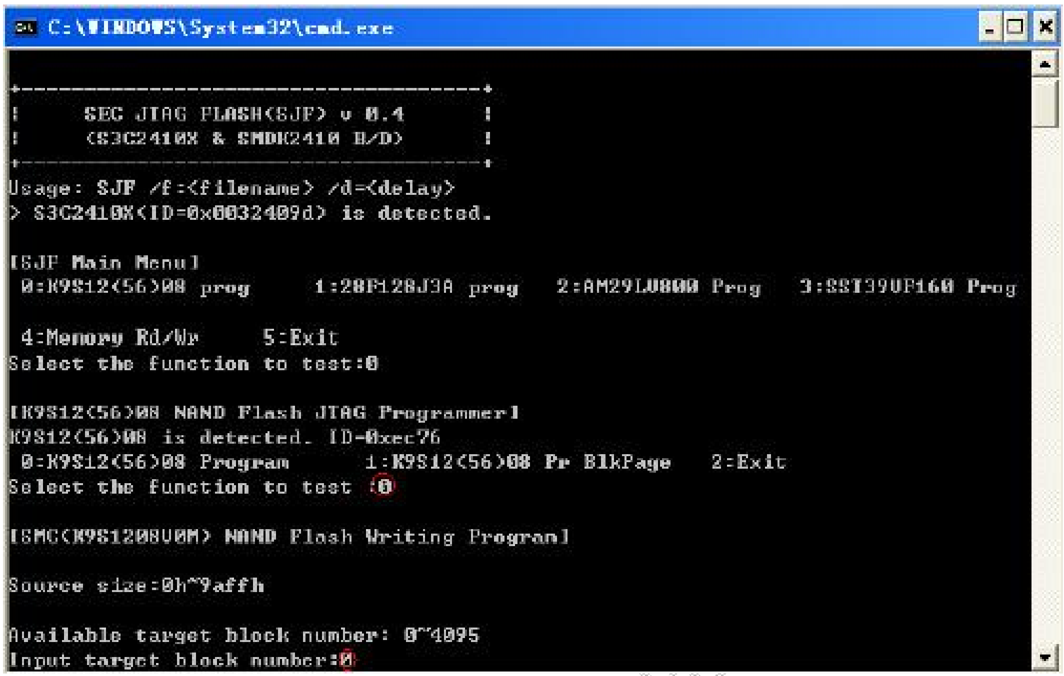


图 4.28 选择烧写类型及偏移地址

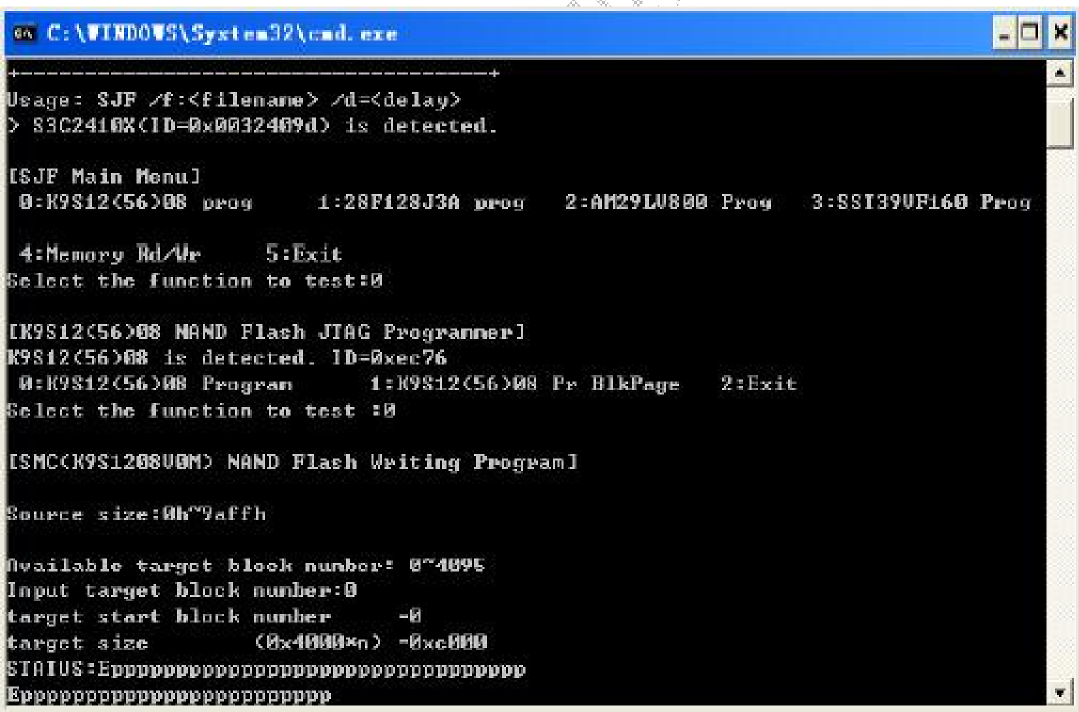


图 4.29 Flash 烧写过程

4．实验结果

系统烧写完成后，程序会自动退出，这样就完成了对 Flash 的烧写。

本 章 小 结

本章讲解了嵌入式中的基本概念，包括嵌入式系统的含义、它的发展历史、嵌入式系统的特点以及嵌入式系统的体系结构。在这里，重点要掌握嵌入式系统和通用计算机的区别，希望读者能对以上几方面都能对其进行一一比较，以加深对嵌入式系统的理解。

接下来对 ARM 体系进行了概括性地讲解，希望读者能重点掌握 ARM9 的特性，有条件的读者希望能结合实际开发板进行学习，没有开发板的读者也可参看图中的实物图，以获得感性的认识。另外，不同的硬件平台都会有一定的区别，但其主要原理是一样的，对于某些细节的不同处理请读者参阅不同厂商的用户手册。

本章的最后讲解了嵌入式软件开发的流程，其中重点讲解了交叉编译和交叉调试，这些概念初次学习会感觉比较枯燥，但这些概念又是非常重要的，在后面的具体开发中会经常涉及到，希望读者对这些内容能够认真消化。

最后安排的一个实验希望有条件的读者能动手做做，当然在做之前一定认真阅读不同厂商提供的用户手册。

思考与练习

1．从各方面比较嵌入式系统与通用计算器的区别。
2．ARM9 有哪些优于 ARM7 的特性？
3．什么是交叉编译？为什么要进行交叉编译？
4．嵌入式开发的常用调试手段有那几种？说出它们各自的优缺点。

“黑色经典”系列之《嵌入式 Linux 应用程序开发详解》

第 5 章 嵌入式 Linux 开发环境的搭建

本章目标

在了解了嵌入式开发的基本概念之后，本章主要学习如何搭建嵌入式 Linux 开发的环境，通过本章的学习，读者能够掌握以下内容。

- 掌握嵌入式交叉编译环境的搭建
- 掌握嵌入式主机通信环境的配置
- 学会制作交叉编译工具链
- 学会配置 Linux 下的 minicom 和 Windows 下的超级终端
- 学会在 Linux 下和 Windows 下配置 TFTP 服务
- 学会配置 NFS 服务
- 学会编译 Linux 内核
- 学会搭建 Linux 的根文件系统
- 熟悉嵌入式 Linux 的内核相关代码的分布情况
- 掌握 Bootloader 的原理
- 了解 U-Boot 的代码结构和编译方法

5.1 嵌入式开发环境的搭建

5.1.1 嵌入式交叉编译环境的搭建

交叉编译的概念在第 4 章中已经详细讲述过，搭建交叉编译环境是嵌入式开发的第一步，也是必备一步。搭建交叉编译环境的方法很多，不同的体系结构、不同的操作内容甚至是不同版本的内核，都会用到不同的交叉编译器，而且，有些交叉编译器经常会有部分的 BUG，这都会导致最后的代码无法正常地运行。因此，选择合适的交叉编译器对于嵌入式开发是非常重要的。

交叉编译器完整的安装一般涉及到多个软件的安装（读者可以从 ftp://gcc.gnu.org/pub/下载），包括 binutils、gcc、glibc 等软件。其中，binutils 主要用于生成一些辅助工具，如 objdump、as、ld 等；gcc 是用来生成交叉编译器，主要生成 arm-linux-gcc 交叉编译工具（应该说，生成此工具后已经搭建起了交叉编译环境，可以编译 Linux 内核了，但由于没有提供标准用户函数库，用户程序还无法编译）；glibc 主要是提供用户程序所使用的一些基本的函数库。这样，交叉编译环境就完全搭建起来了。

上面所述的搭建交叉编译环境比较复杂，很多步骤都涉及到对硬件平台的选择。因此，现在提供开发板的公司一般会在附赠的光盘中提供该公司测试通过的交叉编译器，而且很多公司把以上安装步骤全部写入脚本文件或者以发行包的形式提供，这样就大大方便了用户的使用。如优龙的开发光盘里就随带了 2.95.3 和 3.3.2 两个版本的交叉编译器，其中前一个版本是用于编译 Linux2.4 内核的，而后一个版本是用于编译 Linux2.6 版本内核的。由于这是厂商测试通过的编译器，因此可靠性会比较高，而且与开发板能够很好地吻合。所以推荐初学者直接使用厂商提供的编译器。当然，由于时间滞后的原因，这个编译器往往不是最新版本的，若需要更新时希望读者另外查找相关资料学习。本书就以优龙自带的 cross-3.3.2 为例进行讲解（具体的名称不同厂商可能会有区别）。

安装交叉编译器的具体步骤在第 2 章的实验二中已经进行了详细地讲解了，在此仅回忆关键步骤，对于细节请读者参见第 2 章的实验二。

在/usr/local/arm 下解压 cross-3.3.2.bar.bz2。

```
[root@localhost arm]# tar –jxvf cross-3.3.2.bar.bz2
[root@localhost arm]# ls
3.3.2  cross-3.3.2.tar.bz2
[root@localhost arm]# cd ./3.3.2
[root@localhost arm]# ls
arm-linux  bin  etc  include  info  lib  libexec  man  sbin  share  VERSIONS
[root@localhost bin]# which arm-linux*
/usr/local/arm/3.3.2/bin/arm-linux-addr2line
/usr/local/arm/3.3.2/bin/arm-linux-ar
/usr/local/arm/3.3.2/bin/arm-linux-as
```

```
/usr/local/arm/3.3.2/bin/arm-linux-c++
/usr/local/arm/3.3.2/bin/arm-linux-c++filt
/usr/local/arm/3.3.2/bin/arm-linux-cpp
/usr/local/arm/3.3.2/bin/arm-linux-g++
/usr/local/arm/3.3.2/bin/arm-linux-gcc
/usr/local/arm/3.3.2/bin/arm-linux-gcc-3.3.2
/usr/local/arm/3.3.2/bin/arm-linux-gccbug
/usr/local/arm/3.3.2/bin/arm-linux-gcov
/usr/local/arm/3.3.2/bin/arm-linux-ld
/usr/local/arm/3.3.2/bin/arm-linux-nm
/usr/local/arm/3.3.2/bin/arm-linux-objcopy
/usr/local/arm/3.3.2/bin/arm-linux-objdump
/usr/local/arm/3.3.2/bin/arm-linux-ranlib
/usr/local/arm/3.3.2/bin/arm-linux-readelf
/usr/local/arm/3.3.2/bin/arm-linux-size
/usr/local/arm/3.3.2/bin/arm-linux-strings
/usr/local/arm/3.3.2/bin/arm-linux-strip
```

可以看到，在/usr/local/arm/3.3.2/bin/下已经安装了很多交叉编译工具。用户可以查看 arm 文件夹下的 VERSION 文件，显示如下：

```
Versions
  gcc-3.3.2
  glibc-2.3.2
  binutils-head
Tool chain binutils configuration:
../binutils-head/configure …

../glibc-2.3.2/configure …
Tool chain gcc configuration
../gcc-3.3.2/configure …
```

可以看到，这个优龙公司提供的交叉编译工具确实集成了 binutils、gcc、glibc 这几个软件，而每个软件也都有比较复杂的配置信息，读者可以查看 Version 文件了解相关信息。

5.1.2 超级终端和 Minicom 配置及使用

前文已知，嵌入式系统开发的程序运行环境是在硬件开发板上的，那么如何把开发板上的信息显示给开发人员呢？最常用的就是通过串口线输出到宿主机的显示器上，这样，开发人员就可以看到系统的运行情况了。在 Windows 和 Linux 中都有不少串口通信软件，可以很方便地对串口进行配置，其中最主要的配置参数就是波特率、数据位、停止位、奇偶校验位

和数据流控制位等，但是它们一定要根据实际情况进行相应配置。下面介绍 Windows 中典型的串口通信软件“超级终端”和在 Linux 下的“Minicom”。

1．超级终端

首先，打开 Windows 下的“开始”→“附件”→“通讯”→“超级终端”，这时会出现如图 5.1 所示的新建超级终端界面，在“名称”处可随意输入该连接的名称。

接下来，将“连接时使用”的方式改为“COM1”，即通过串口 1，如图 5.2 所示。

接下来就到了最关键的一步——设置串口连接参数。要注意，每块开发板的连接参数有可能会有差异，其中的具体数据在开发商提供的用户手册中会有说明。如优龙的这款 FS2410 采用的是波特率：115200，数据为 8 位，无奇偶校验位，停止位 1，无硬件流，其对应配置如图 5.3 所示。

这样，就基本完成了配置，最后一步“单击”确定就可以了。这时，读者可以把开发板的串口线和 PC 机相连，若配置正确，在开发板上电后在超级终端的窗口里应能显示类似如图 5.4 的串口信息。

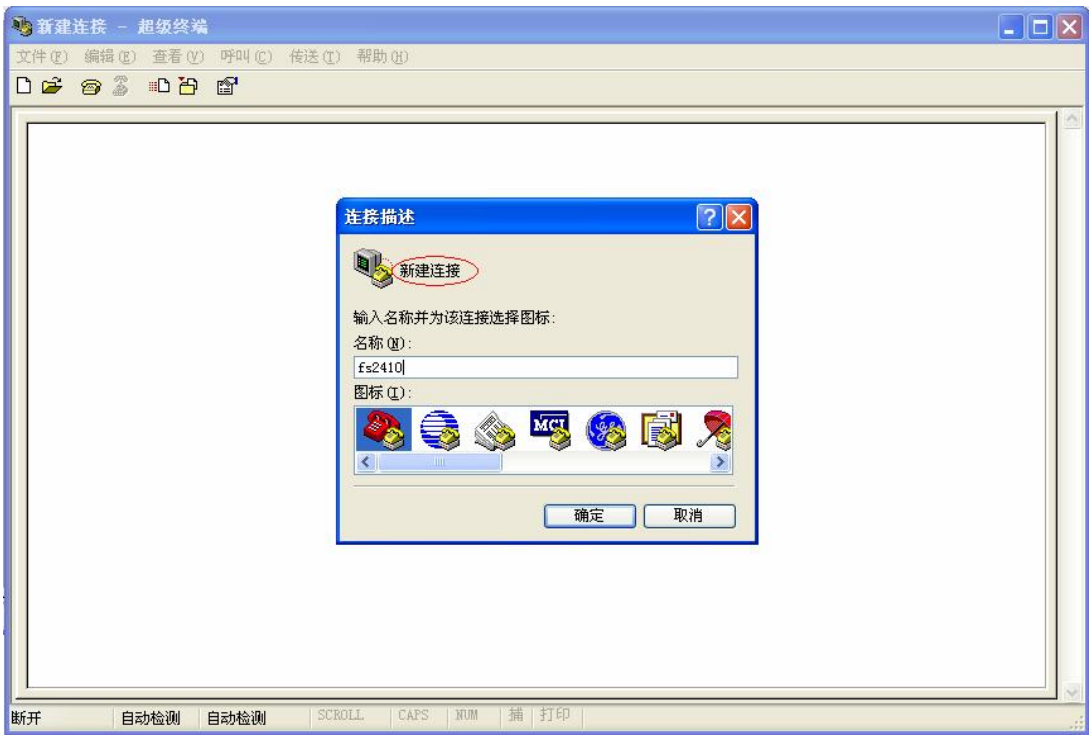


图 5.1　新建超级终端界面

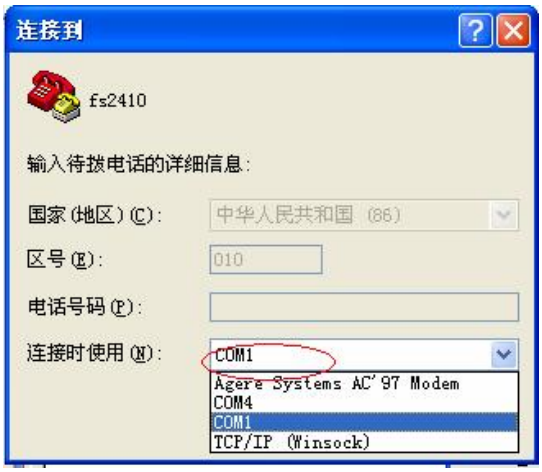


图 5.2　选择连接时使用方式

图 5.3　配置串口相关参数

注意　要分清开发板上的串口 1，串口 2，如在优龙的开发板上标有“UART1”、“UATR2”，否则串口无法打印出信息。

2．Minicom

Minicom 是 Linux 下串口通信的软件，它的使用完全依靠键盘的操作，虽然没有“超级终端”那么易用，但是使用习惯之后读者将会体会到它的高效与便利。下面主要讲解如何对 Minicom 进行串口参数的配置。

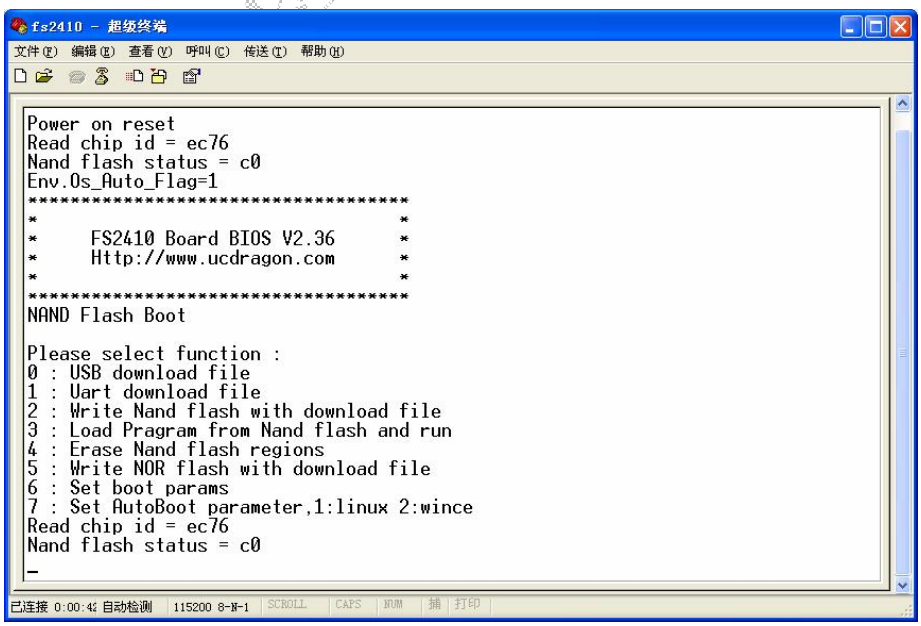


图 5.4　串口相关信息

首先在命令行中键入“minicom”，这就启动了 minicom 软件。Minicom 在启动时默认会进行初始化配置，如图 5.5 所示。

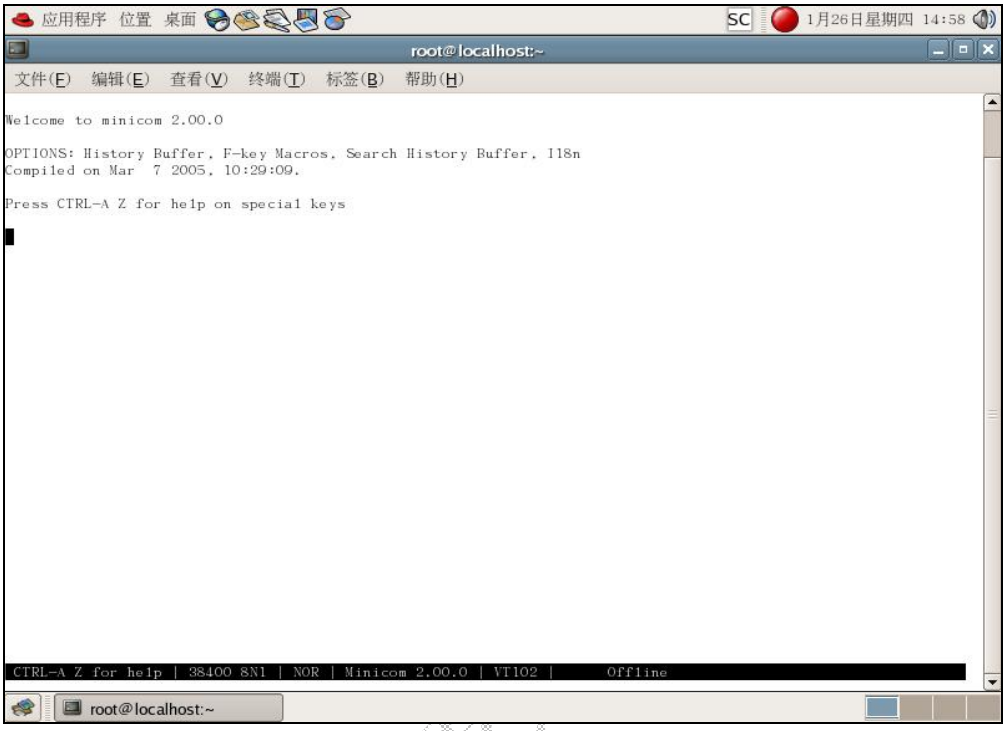


图 5.5 minicom 启动

注意 在 Minicom 的使用中，经常会遇到三个键的操作，如“CTRL-A Z”，这表示先同时按下 CTRL 和“A”（大写），然后松开此二键再按下“Z”。

正如图 5.5 中的提示，接下来可键入 CTRL+A Z，来查看 minicom 的帮助，如图 5.6 所示。

按照帮助所示，可键入“O”（代表 Configure Minicom）来配置 minicom 的串口参数，当然也可以直接键入“CTRL-A O”来进行配置。如图 5.7 所示。

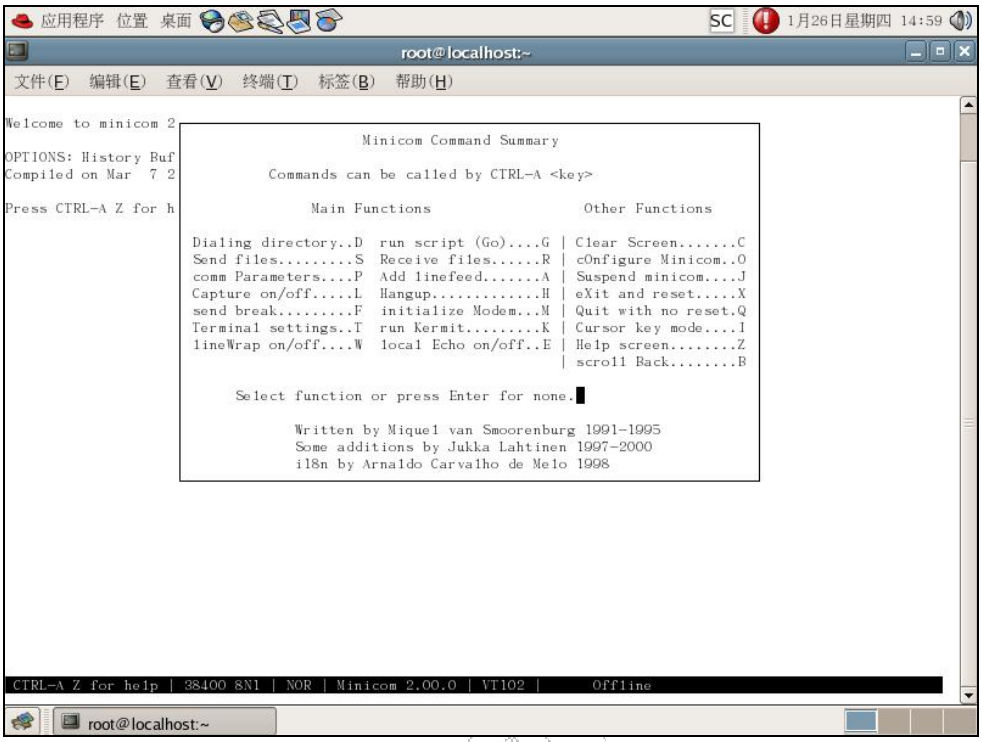


图 5.6 minicom 帮助

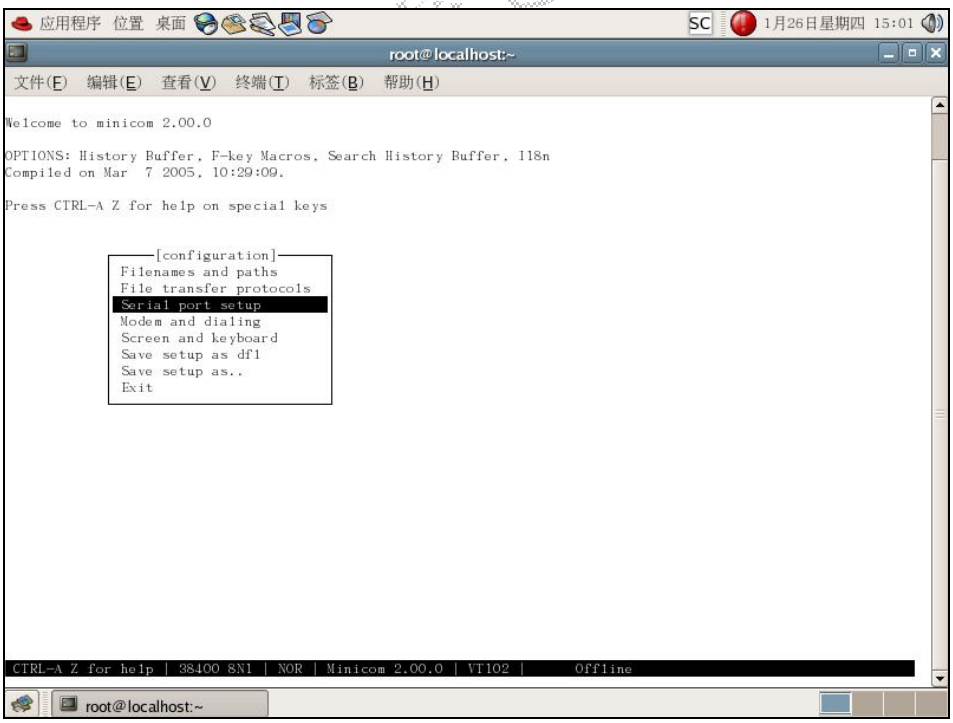


图 5.7 minicom 配置界面

在这个配置框中选择“Serial port setup”子项，进入如图 5.8 所示配置界面。

上面列出的配置是 minicom 启动是的默认配置，用户可以通过键入每一项前的大写字母，分别对每一项进行更改。图 5.9 所示为在“Change which setting 中”键入了“A”，此时光标转移到第 A 项的对应处。

注意 在 minicom 中“ttyS0”对应“COM1”，“ttyS1”对应“COM2”。

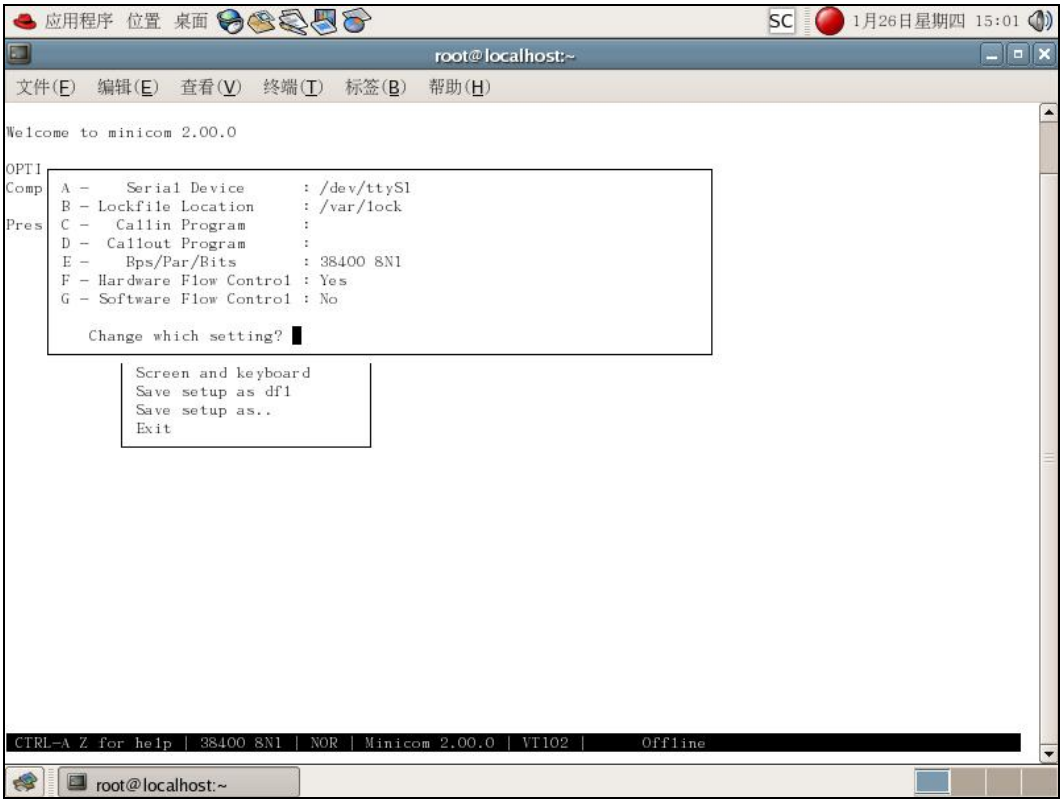


图 5.8　minicom 串口属性配置界面

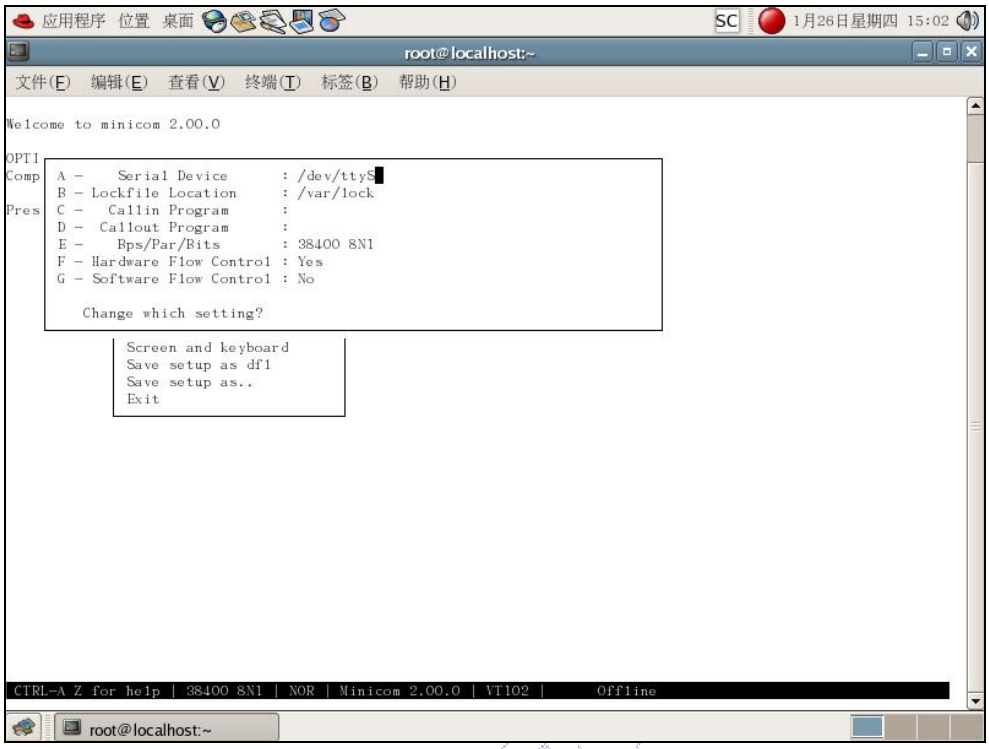


图 5.9　minicom 串口号配置

接下来，要对波特率、数据位和停止位进行配置，键入“E”，进入如图 5.10 所示的配置界面。

在该配置界面中，可以键入相应波特率、停止位等对应的字母，即可实现配置，配置完成后按回车键就退出了该配置界面，在上层界面中显示如图 5.11 所示配置信息，要注意与图 5.8 进行对比，确定相应参数是否已被重新配置。

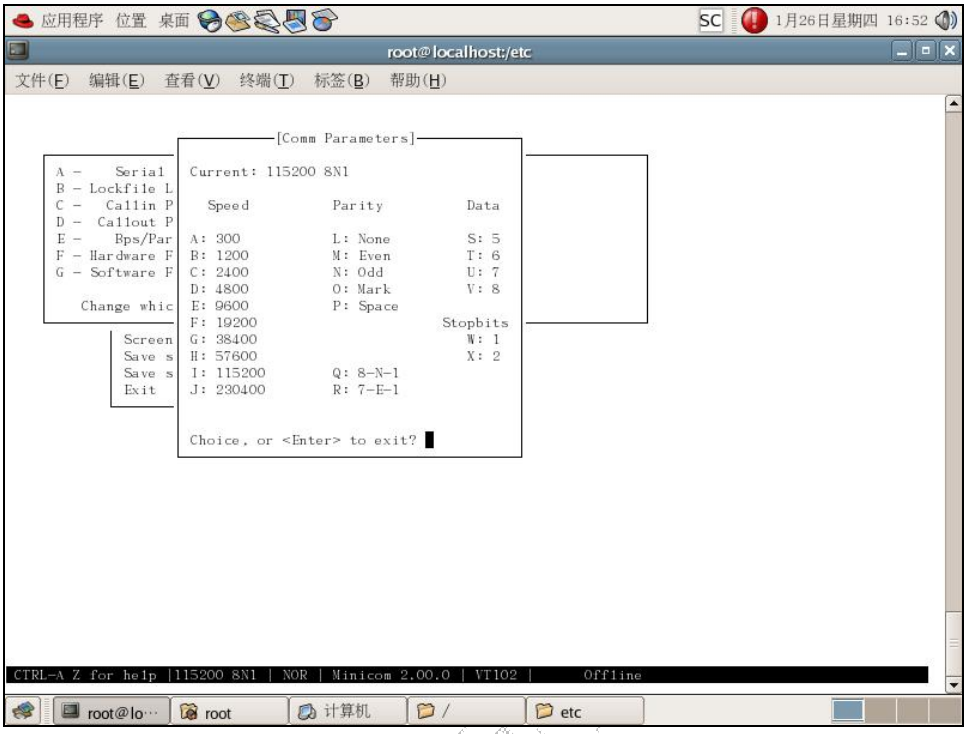


图 5.10 minicom 波特率等配置界面

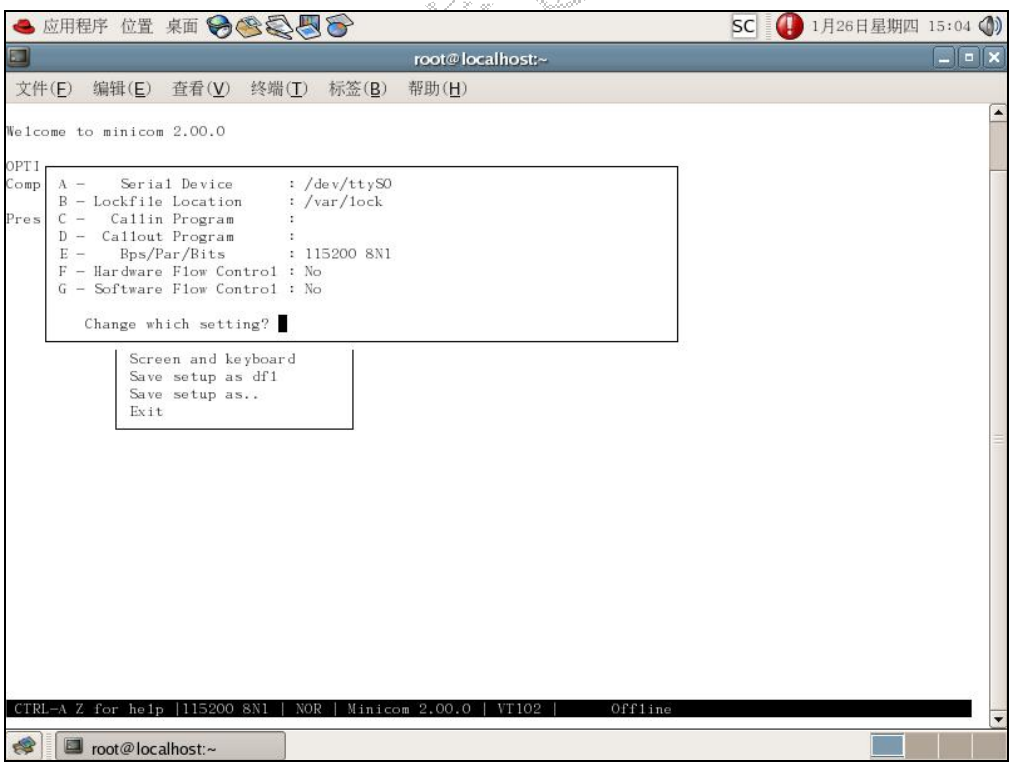


图 5.11 minicom 配置完成后界面

在确认配置正确后，可键入回车返回上级配置界面，并将其保存为默认配置，如图 5.12 所示。

之后，可重新启动 minicom 使刚才配置生效，在连上开发板的串口线之后，就可在 minicom 中打印出正确的串口信息，如图 5.13 所示。

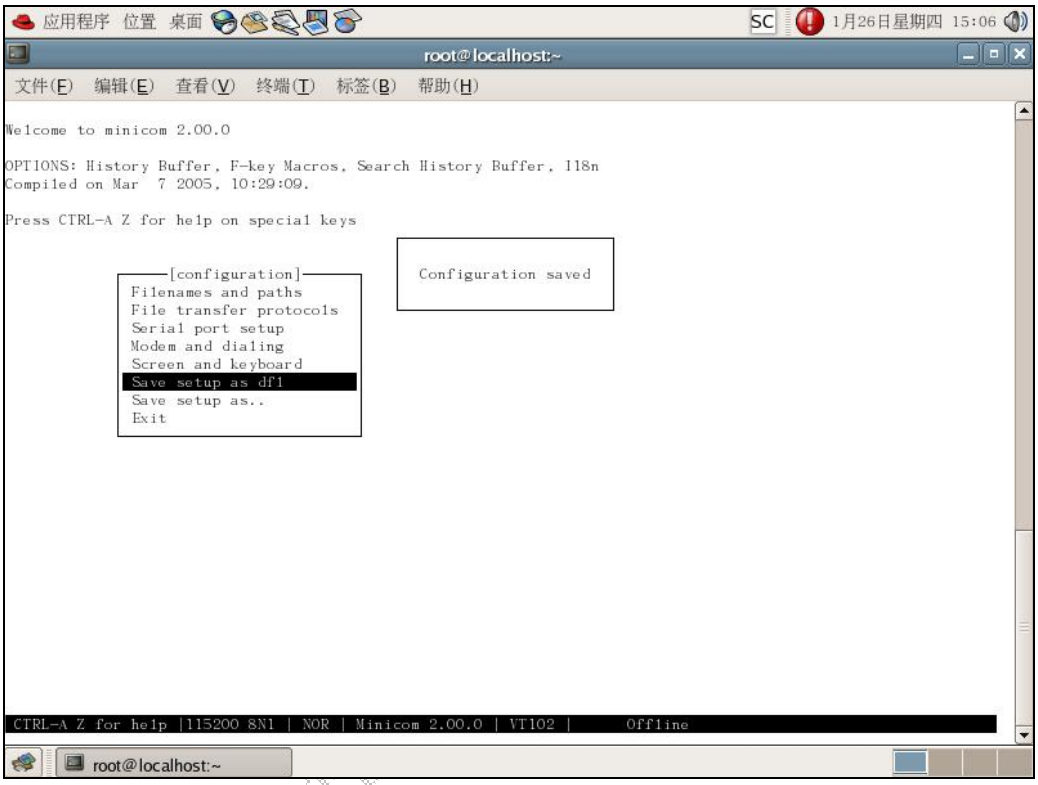


图 5.12　minicom 保存配置信息

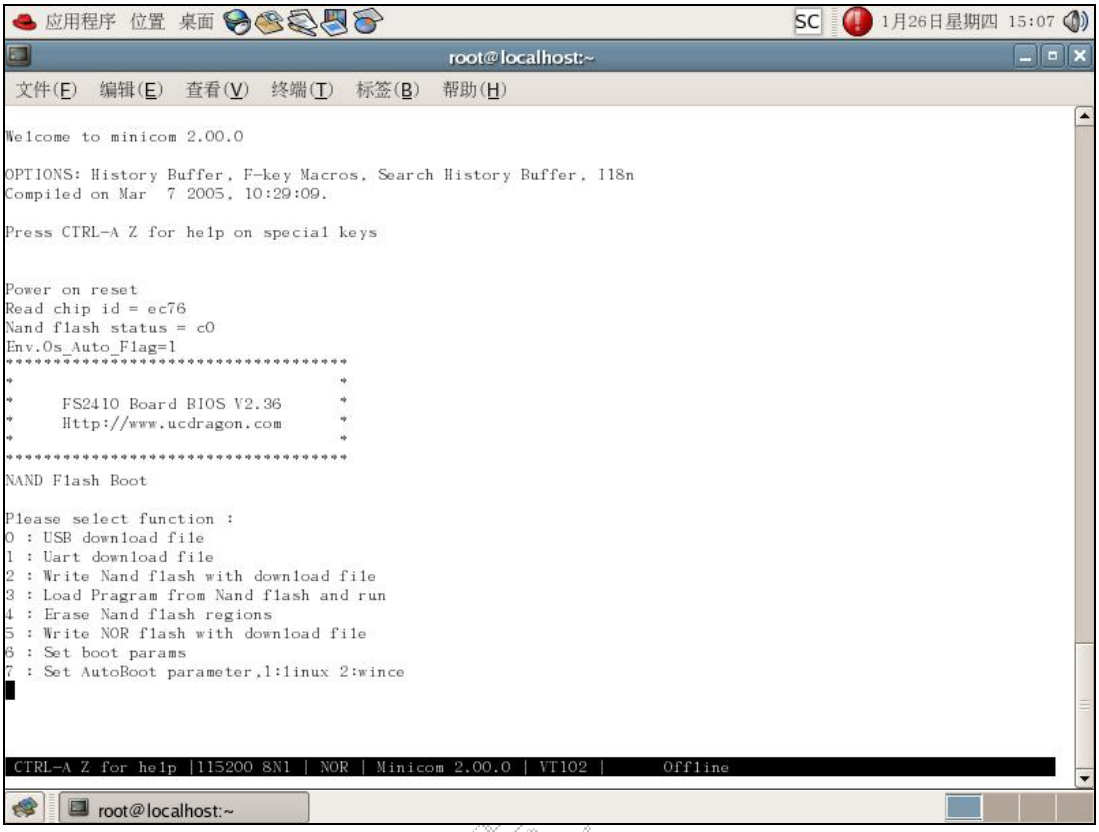


图 5.13 minicom 显示串口信息

到此为止，读者已经能将开发板的系统情况通过串口打印到宿主机上了，这样，就能很好地了解硬件的运行状况。

小知识 通过串口打印信息是一个很常见的手段，很多其他情况如路由器等也是通过配置串口的波特率这些参数来显示对应信息的。

5.1.3 下载映像到开发板

正如第 4 章中所述，嵌入式开发的运行环境是目标板，而开发环境是宿主机。因此，需要把宿主机中经过编译之后的可执行文件下载到目标板上去。要注意的是，这里所说的下载是下载到目标机中的 SDRAM。然后，用户可以选择直接从 SDRAM 中运行或写入到 Flash 中再运行。运行常见的下载方式有网络下载（如 tftp、ftp 等方式）、串口下载、USB 下载等，本书主要讲解网络下载中的 tftp 方式和串口下载方式。

1．tftp

Tftp 协议是简单文件传输协议，它可以看作是一个 FTP 协议的简化版本，与 FTP 协议相比，它的最大区别在于没有用户管理的功能。它的传输速度快，可以通过防火墙，使用方便快捷，因此在嵌入式的文件传输中广泛使用。

同 FTP 一样，tftp 分为客户端和服务器端两种。通常，首先在宿主机上开启 tftp 服务器

端服务，设置好 tftp 的根目录内容（也就是供客户端下载的文件），接着，在目标板上开启 tftp 的客户端程序（现在很多开发板都已经提供了该项功能）。这样，把目标板和宿主机用直连线相连之后，就可以通过 tftp 协议传输可执行文件了。

下面分别讲述在 Linux 下和 Windows 下的配置方法。

（1）Linux 下 tftp 服务配置

Linux 下 tftp 的服务器服务是由 xinetd 所设定的，默认情况下是处于关闭状态。

首先，要修改 tftp 的配置文件，开启 tftp 服务，如下所示：

```
[root@sunq tftpboot]# vi /etc/xinetd.d/tftp
# default: off
# description: The tftp server serves files using the trivial file transfer \
#          protocol.  The tftp protocol is often used to boot diskless \
#          workstations, download configuration files to network-aware printers, \
#          and to start the installation process for some operating systems.
service tftp
{
            socket_type                 = dgram
            protocol                    = udp
            wait                        = yes
            user                        = root
            server                      = /usr/sbin/in.tftpd
            server_args                 = -s /tftpboot
            disable                     = no
            per_source                  = 11
            cps                         = 100 2
            flags                       = IPv4
}
```

在这里，主要要将“disable=yes”改为“no”，另外，从“server_args”可以看出，tftp 服务器端的默认根目录为“/tftpboot”，用户若需要可以更改为其他目录。

接下来，重启 xinetd 服务，使刚才的更改生效，如下所示：

```
[root@sunq tftpboot]# service xinetd restart
关闭 xinetd:                                           [  确定  ]
启动 xinetd:                                           [  确定  ]
```

接着，使用命令“netstat -au”以确认 tftp 服务是否已经开启，如下所示：

```
[root@sunq tftpboot]# netstat -au
Active Internet connections (servers and established)
Proto Recv-Q Send-Q Local Address               Foreign Address      State
```

```
udp        0      0 *:32768                 *:*
udp        0      0 *:831                   *:*
udp        0      0 *:tftp                  *:*
udp        0      0 *:sunrpc                *:*
udp        0      0 *:ipp                   *:*
```

这时，用户就可以把所需要的传输文件放到“/tftpboot”目录下，这样，主机上的 tftp 服务就可以建立起来了。

接下来，用直连线（注意：不可以使用网线）把目标板和宿主机连起来，并且将其配置成一个网段的地址，再在目标板上启动 tftp 客户端程序（注意：不同的开发板所使用的命令可能会不同，读者可以查看帮助来获得确切的命令名及格式），如下所示：

```
=>tftpboot 0x30200000 zImage
TFTP from server 192.168.1.1; our IP address is 192.168.1.100
Filename 'zImage'.
Load address: 0x30200000
Loading: #################################################################
         #################################################################
         ##############################################
done
Bytes transferred = 881988 (d7544 hex)
```

可以看到，此处目标板使用的 IP 为“192.168.1.100”，宿主机使用的 IP 为“192.168.1.1”，下载到目标板的地址为 0x30200000，文件名为“zImage”。

（2）Windows 下 tftp 服务配置

在 Windows 下配置为 tftp 服务器端需要下载 tftp 服务器软件，常见的为 tftpd32。

首先，单击 tftpd32 下方的设置按钮，进入设置界面，如图 5.14 所示，在这里，主要配置 tftp 服务器端地址，也就是本机的地址。

接下来，重新启动 tftpd32 软件使刚才的配置生效，这样服务器端的配置就完成了，这时，就可以用直连线连接目标机和宿主机，且在目标机上开启 tftp 服务进行文件传输，这时，tftp 服务器端如图 5.15 和图 5.16 所示。

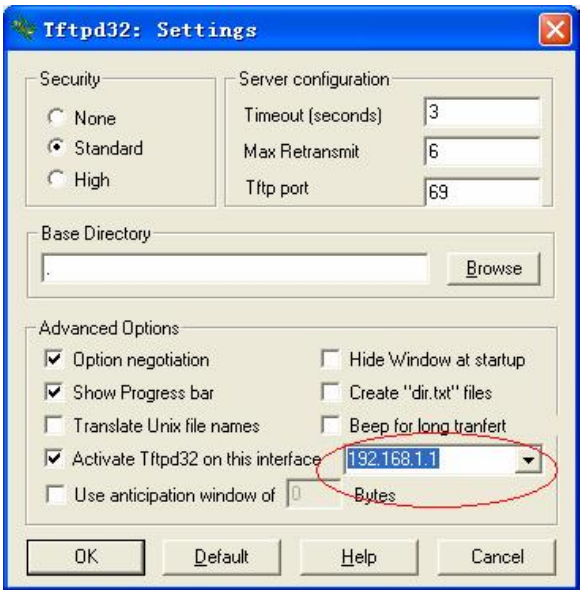


图 5.14　tftpd32 配置界面

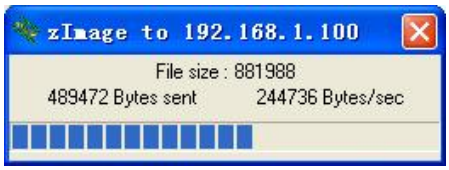


图 5.15　tftp 文件传输

图 5.16　tftp 服务器端显示情况

小知识　tftp 是一个很好的文件传输协议，它的简单易用吸引了广大用户。但它同时也存在着较大的安全隐患。由于 tftp 不需要用户的身份认证，因此给了黑客的可乘之机。读者一定还记得在 2003 年 8 月 12 日爆发的全球冲击波（Worm.Blaster）病毒，这种病毒就是模拟一个 tftp 服务器，并启动一个攻击传播线程，不断地随机生成攻击地址进行入侵。因此在使用 tftp 时一定要设置一个单独的目录作为 tftp 服务的根目录，如上文所述的“/tftpboot”等。

2．串口下载

使用串口下载需要配合特定的下载软件，如优龙公司提供的 DNW 软件等，一般在

Windows 下进行操作。虽然串口下载的速度没有网络下载快，但由于它很方便，不需要额外的连线和设置 IP 等操作，因此也广受用户的青睐。下面就以 DNW 软件为例，介绍串口下载的方式。

与其他串口通信的软件一样，在 DNW 中也要设置“波特率”、“端口号”等。打开“Configuration”下的“Options”界面，如图 5.17 所示。

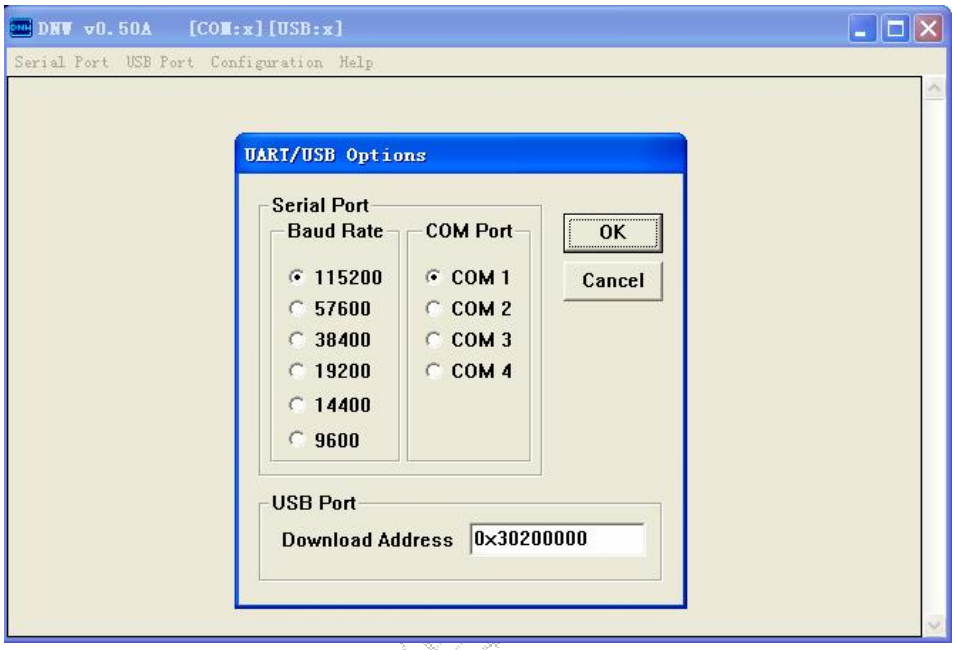


图 5.17 DNW 配置界面

在配置完之后，单击“Serial Port”下的“Connect”，再将开发板上电，选择“串口下载”，接着再在“Serial Port”下选择“Transmit”，这时，就可以进行文件传输了，如图 5.18 和图 5.19 所示。这里 DNW 默认串口下载的地址为 0x30200000。

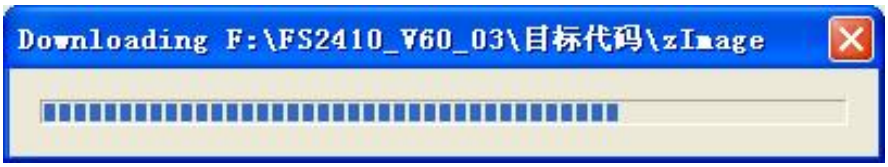


图 5.18 DNW 串口下载图

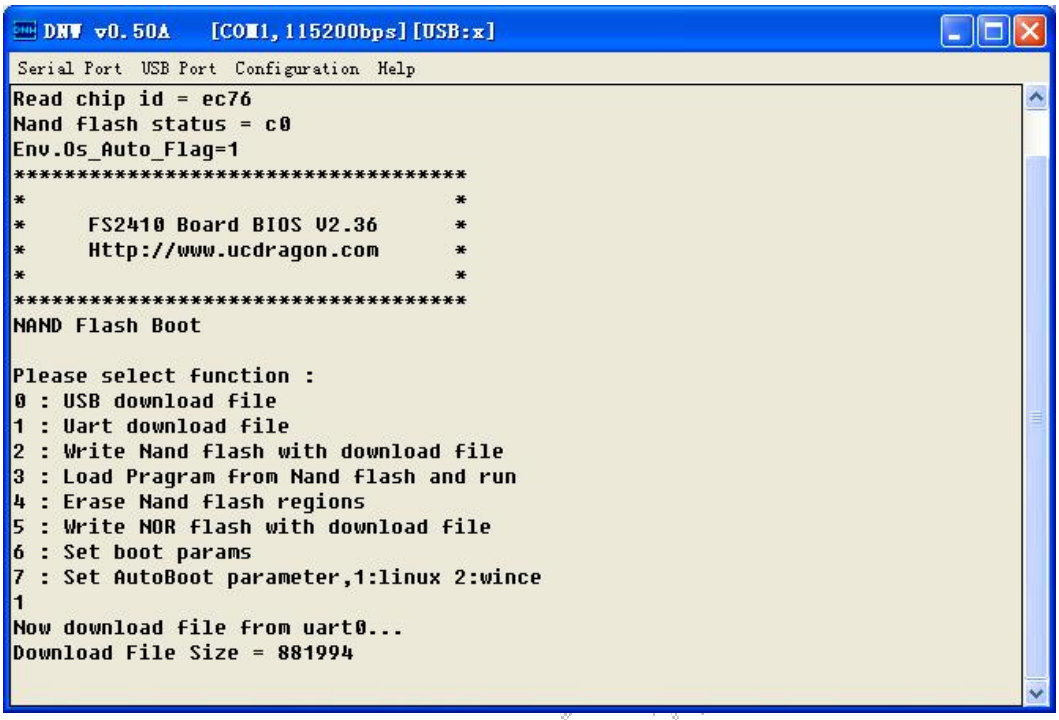


图 5.19　DNW 串口下载情形图

5.1.4　编译嵌入式 Linux 内核

在做完了前期的准备工作之后，在这一步，读者就可以编译嵌入式移植 Linux 的内核了。在这里，本书主要介绍嵌入式 Linux 内核的编译过程，在下一节会进一步介绍嵌入式 Linux 中体系结构相关的内核代码，读者在此之后就可以尝试嵌入式 Linux 操作系统的移植。

编译嵌入式 Linux 内核都是通过 make 的不同命令来实现的，它的执行配置文件就是在第 3 章中讲述的 Makefile。Linux 内核中不同的目录结构里都有相应的 Makefile，而不同的 Makefile 又通过彼此之间的依赖关系构成统一的整体，共同完成建立依存关系、建立内核等功能。

内核的编译根据不同的情况会有不同的步骤，但其中最主要分别为 3 个步骤：内核配置、建立依存关系、建立内核，其他的为一些辅助功能，如清除文件等。读者在实际编译时若出现错误等情况，可以考虑采用其他辅助功能。下面分别讲述这 3 步主要的步骤。

（1）内核配置

第一步内核配置中的选项主要是用户用来为目标板选择处理器架构的选项，不同的处理器架构会有不同的处理器选项，比如 ARM 就有其专用的选项如“Multimedia capabilities port drivers”等。因此，在此之前，必须确保在根目录中 Makefile 里“ARCH”的值已设定了目标板的类型，如：

```
ARCH        := arm
```

接下来就可以进行内核配置了，内核支持 4 种不同的配置方法，这几种方法只是与用

户交互的界面不同，其实现的功能是一样的。每种方法都会通过读入了一个默认的配置文件——根目录下“.config”隐藏文件（用户也可以手动修改该文件，但不推荐使用）。当然，用户也可以自己加载其他配置文件，也可以将当前的配置保存为其他名字的配置文件。这 4 种方式如下。

- make config：基于文本的最为传统的配置界面，不推荐使用。
- make menuconfig：基于文本选单的配置界面，字符终端下推荐使用。
- make xconfig：基于图形窗口模式的配置界面，Xwindow 下推荐使用。
- make oldconfig：自动读入“.config”配置文件，并且只要求用户设定前次没有设定过的选项。

在这 4 种模式中，make menuconfig 使用最为广泛，下面就以 make menuconfig 为例进行讲解，如图 5.20 所示。

从该图中可以看出，Linux 内核允许用户对其各类功能逐项配置，一共有 18 类配置选项，这里就不对这 18 类配置选项进行一一讲解了，需要的读者可以参见相关选项的 help。在 menuconfig 的配置界面中是纯键盘的操作，用户可使用上下键和“Tab”键移动光标以进入相关子项，图 5.21 所示为进入了“System Type”子项的界面，该子项是一个重要的选项，主要用来选择处理器的类型。

可以看到，每个选项前都有个括号，可以通过按空格键或“Y”键表示包含该选项，按“N”表示不包含该选项。

另外，读者可以注意到，这里的括号有 3 种，即中括号、尖括号或圆括号。读者可以用空格键选择相应的选项时可以发现中括号里要么是空，要么是“*”；尖括号里可以是空，“*”和“M”，分别表示包含选项、不包含选项和编译成模块；圆括号的内容是要求用户在所提供的几个选项中选择一项。

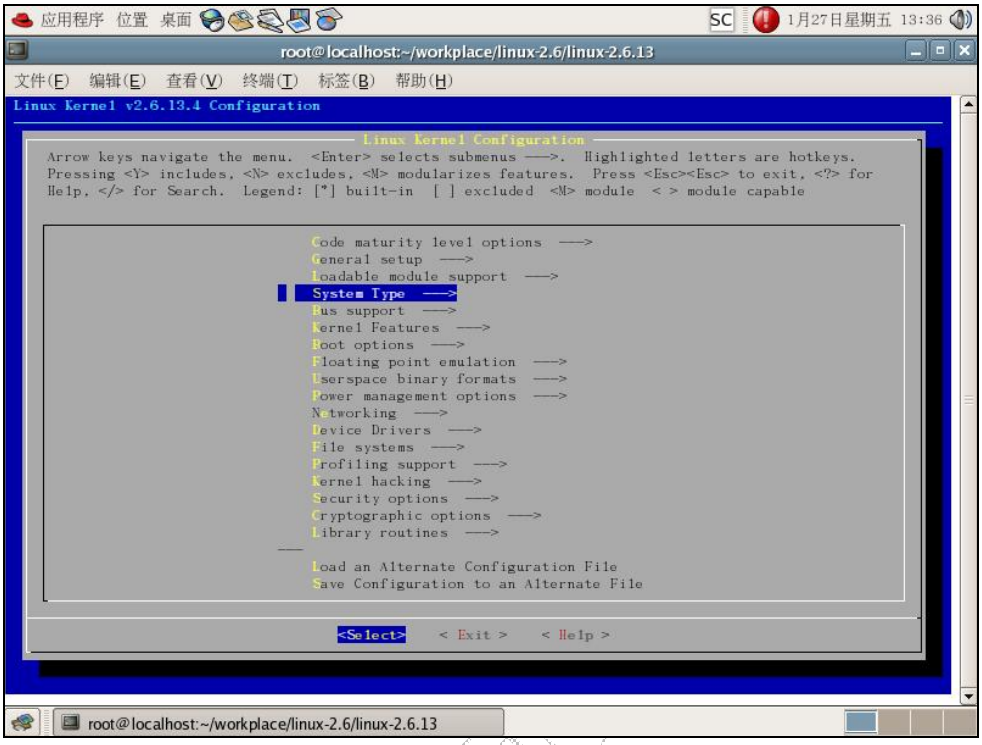


图 5.20　make menuconfig 配置界面

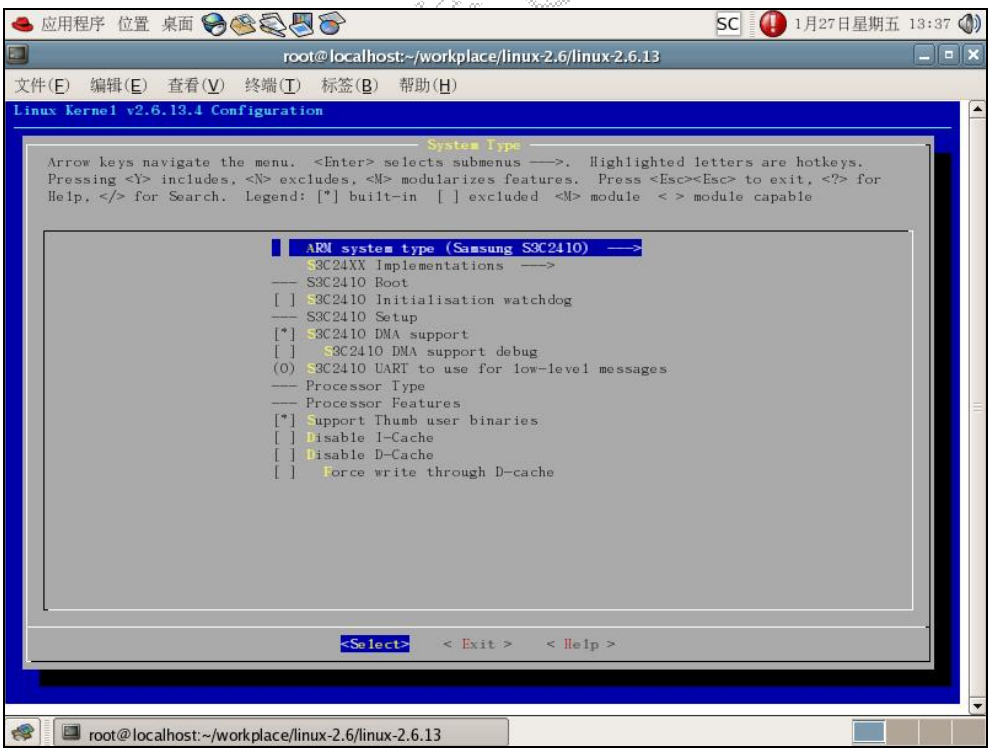


图 5.21　System Type 子项

此外，要注意 2.6 和 2.4 内核在串口命名上的一个重要区别，在 2.4 内核中“COM1”对应的是“ttyS0”，而在 2.6 内核中“COM1”对应“ttySAC0”，因此在启动参数的子项要格外注意，如图 5.22 所示，否则串口打印不出信息。

一般情况下，使用厂商提供的默认配置文件都能正常运行，所以用户初次使用时可以不用对其进行额外的配置，在以后使用需要其他功能时再另行添加，这样可以大大减少出错的几率，有利于错误定位。在完成配置之后，就可以保存退出，如图 5.23 所示。

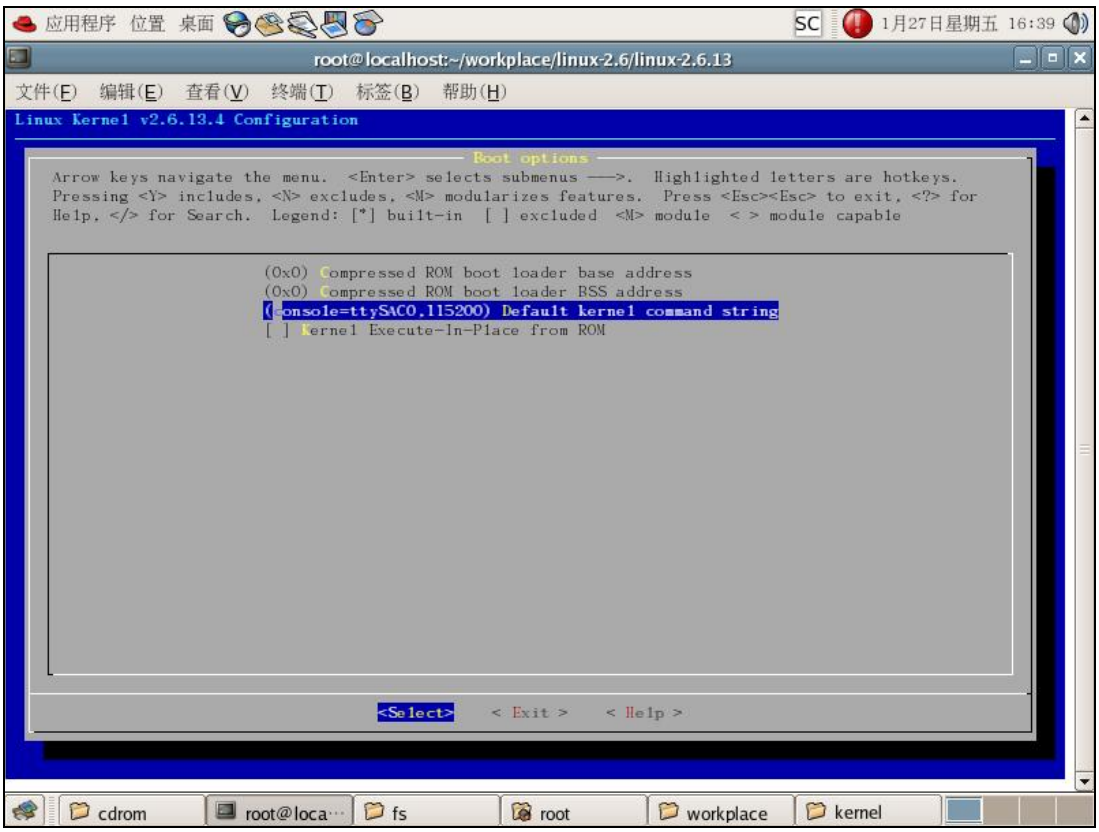


图 5.22 启动参数配置子项

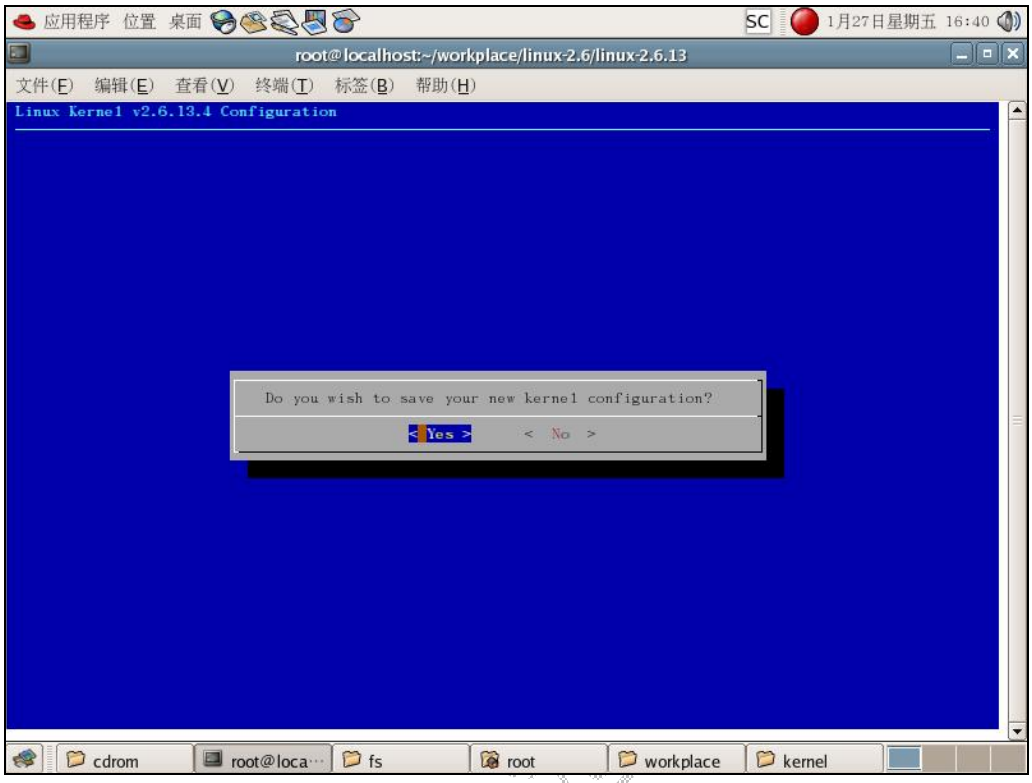


图 5.23　保存退出

（2）建立依赖关系

由于内核源码树中的大多数文件都与一些头文件有依赖关系，因此要顺利建立内核，内核源码树中的每个 Makefile 就必须知道这些依赖关系。建立依赖关系往往发生在第一次编译内核的时候，它会在内核源码树中每个子目录产生一个“.depend”文件。运行“make dep”即可。

（3）建立内核

建立内核可以使用“make zImage”或“make bzImage”，这里建立的为压缩的内核映像。通常在 Linux 中，内核映像分为压缩的内核映像和未压缩的内核映像。其中，压缩的内核映像通常名为 zImage，位于“arch/$（ARCH）/boot”目录中。而未压缩的内核映像通常名为 vmlinux，位于源码树的根目录中。

到这一步就完成了内核源代码的编译，之后，读者可以使用上一小节所讲述的方法把内核压缩文件下载到开发板上运行。

小知识　在嵌入式 Linux 的源码树中通常有以下几个配置文件，“.config”、“autoconf.h”、“config.h”，其中“.config”文件是 make menuconfig 默认的配置文件，位于源码树的根目录中。“autoconf.h”和“config.h”是以宏的形式表示了内核的配置，当用户使用 make menuconfig 做了一定的更改之后，系统自动会在“autoconf.h”和“config.h”中做出相应的更改。它们位于源码树的“/include/linux/”下。

5.1.5 Linux 内核目录结构

Linux 内核的目录结构如图 5.24 所示。

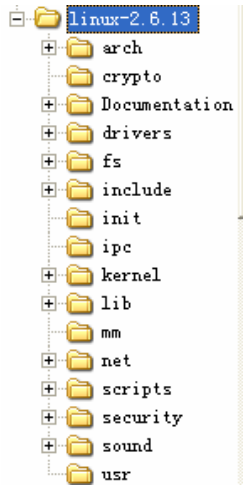


图 5.24 Linux 内核目录结构

- /include 子目录包含了建立内核代码时所需的大部分包含文件，这个模块利用其他模块重建内核。
- /init 子目录包含了内核的初始化代码，这是内核工作的开始的起点。
- /arch 子目录包含了所有硬件结构特定的内核代码。如：arm、i386、alpha。
- /drivers 子目录包含了内核中所有的设备驱动程序，如块设备和 SCSI 设备。
- /fs 子目录包含了所有的文件系统的代码，如：ext2，vfat 等。
- /net 子目录包含了内核的连网代码。
- /mm 子目录包含了所有内存管理代码。
- /ipc 子目录包含了进程间通信代码。
- /kernel 子目录包含了主内核代码。

5.1.6 制作文件系统

读者把上一节中所编译的内核压缩映像下载到开发板后会发现，系统在进行了一些初始化的工作之后，并不能正常启动，如图 5.25 所示。

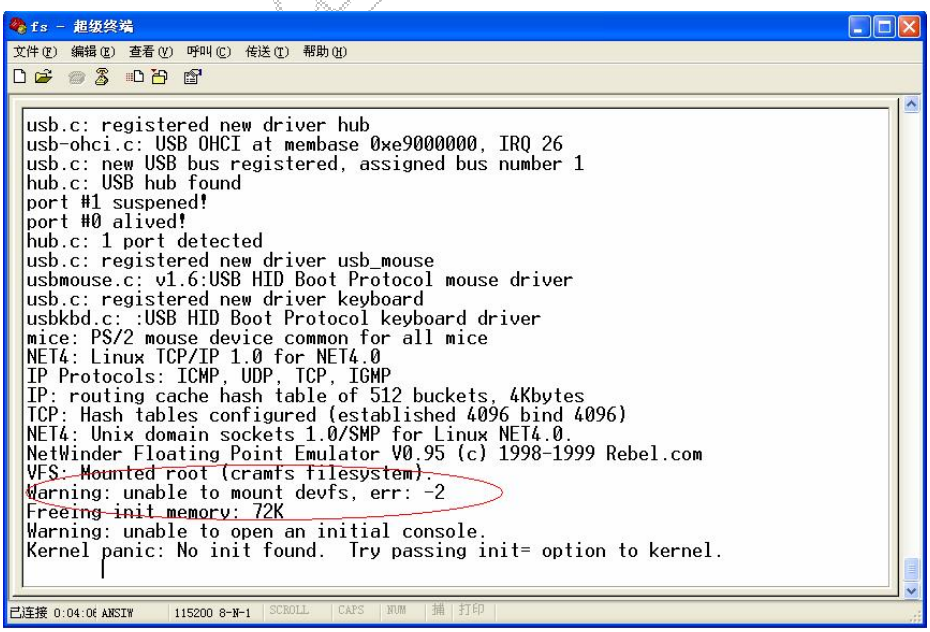


图 5.25 系统启动错误

可以看到，系统启动时发生了加载文件系统的错误。要记住，上一节所编译的仅仅是内核，文件系统和内核是完全独立的两个部分。读者可以回忆一下第 2 章讲解的 Linux 启动过程的分析（嵌入式 Linux 是 Linux 裁减后的版本，其精髓部分是一样的），其中在 head.S 中就加载了根文件系统。因此，加载根文件系统是 Linux 启动中不可缺少的一部分。本节就来讲解嵌入式 Linux 中文件系统和根文件系统的制作方法。

制作文件系统的方法有很多，可以从零开始手工制作，也可以在现有的基础上添加部分内容加载到目标板上去。由于完全手工制作工作量比较大，而且也很容易出错，因此，本节将主要介绍把现有的文件系统加载到目标板上的方法，主要包括制作文件系统镜像和用 NFS 加载文件系统的方法。

1．制作文件系统镜像

读者已经知道，Linux 支持多种文件系统，同样，嵌入式 Linux 也支持多种文件系统。虽然在嵌入式中，由于资源受限的原因，它的文件系统和 Linux 的文件系统有较大的区别（前者往往是只读文件系统），但是，它们的总体架构是一样的，都是采用目录树的结构。在嵌入式中常见的文件系统有 cramfs、romfs、jffs、yaffs 等，这里就以制作 cramfs 文件系统为例进行讲解。cramfs 文件系统是一种经压缩的、极为简单的只读文件系统，因此非常适合嵌入式系统。要注意的是，不同的文件系统都有相应的制作工具，但是其主要的原理和制作方法是类似的。

制作 cramfs 文件系统需要用到的工具是 mkcramfs，下面就来介绍使用 mkcramfs 制作文件系统映像的方法。这里假设用户已经有了一个 cramfs 文件系统，在目录“/root/workplace/fs/guo”里，如下所示：

```
[root@localhost guo]# ls
bin  dev  etc  home  lib  linuxrc  proc  Qtopia  ramdisk  sbin  tmp  usr  var
```

接下来就可以使用 mkcramfs 工具了，格式为：mkcramfs dir name，如下所示。

```
[root@localhost fs]# ./mkcramfs guo FS2410XP_camare_demo4.cramfs
-21.05% (-64 bytes)       Tongatapu
-21.03% (-49 bytes)       Truk
-21.03% (-49 bytes)       Wake
-22.41% (-52 bytes)       Wallis
-21.95% (-54 bytes)       Yap
-17.19% (-147 bytes)      WET
-47.88% (-8158 bytes)     zone.tab
-55.24% (-17421 bytes)    usb-storage.o
-54.18% (-16376 bytes)    usbvideo.o
-54.07% (-2736 bytes)     videodev.o
Everything: 27628 kilobytes
Super block: 76 bytes
CRC: e3a6d7ca
```

可以看到，mkcramfs 在制作文件镜像的时候对文件进行了压缩。

读者可以先在本机上通过 mount 进行验证，如下所示：

```
[root@localhost fs]# mkdir sunq
[root@localhost fs]# mount -o loop FS2410XP_camare_demo4.cramfs ./sunq
[root@localhost fs]# ls sunq
bin  dev  etc  home  lib  linuxrc  proc  Qtopia  ramdisk  sbin  tmp  usr  var
```

接下来，就可以烧入到开发板的相应部分了。

2．NFS 文件系统

NFS 为 Network FileSystem 的简称，最早是由 Sun 公司提出发展起来的，其目的就是让不同的机器、不同的操作系统之间可以彼此共享文件。NFS 可以让不同的主机通过网络将远端的 NFS 服务器共享出来的文件安装到自己的系统中，从客户端看来，使用 NFS 的远端文件就像是使用本地文件一样。在嵌入式中使用 NFS 会使应用程序的开发变得十分方便，并且不用反复地进行烧写镜像文件。

NFS 的使用分为服务器端和客户端，其中服务器端是提供要共享的文件，而客户端则通过挂载“mount”这一动作来实现对共享文件的访问操作。下面主要介绍 NFS 服务器端的使用。

NFS 服务器端是通过读入它的配置文件“/etc/exports”来决定所共享的文件目录的。下面首先讲解这个配置文件的书写规范。

在这个配置文件中，每一行都代表一项要共享的文件目录以及所指定的客户端对其的操作权限。客户端可以根据相应的权限，对该目录下的所有目录文件进行访问。配置文件中每一行的格式如下：

```
[共享的目录] [主机名称或 IP] [参数 1，参数 2…]
```

在这里，主机名或 IP 是可供共享的客户端主机名或 IP，若对所有的 IP 都可以访问，则可用“*”表示。

这里的参数有很多中组合方式，常见的参数如表 5.1 所示。

表 5.1　　常见参数

选　　项	参 数 含 义
rw	可读写的权限
ro	只读的权限
no_root_squash	NFS 客户端分享目录使用者的权限，即如果客户端使用的是 root 用户，那么对于这个共享的目录而言，该客户端就具有 root 的权限
sync	资料同步写入到内存与硬盘当中
async	资料会先暂存于内存当中，而非直接写入硬盘

如在本例中，配置文件“/etc/exports”的代码如下：

```
[root@localhost fs]# cat /etc/exports
/root/workplace    *(rw,no_root_squash)
```

在设定完配置文件之后，需要启动 nfs 服务和 portmap 服务，这里的 portmap 服务是允许 NFS 客户端查看 NFS 服务在用的端口，在它被激活之后，就会出现一个端口号为 111 的 sun RPC（远端过程调用）的服务。这是 NFS 服务中必须实现的一项，因此，也必须把它开启。如下所示：

```
[root@localhost fs]# service portmap start
启动 portmap:                                [确定]
[root@localhost fs]# service nfs start
启动 NFS 服务:                               [确定]
关掉 NFS 配额:                               [确定]
启动 NFS 守护进程:                           [确定]
启动 NFS mountd:                             [确定]
```

可以看到，在启动 NFS 服务的时候启动了 mountd 进程。这是 NFS 挂载服务，用于处理 NFSD 递交过来的客户端请求。另外还会激活至少两个以上的系统守护进程，然后就开始监听客户端的请求，用 cat/var/log/messages 可以看到操作是否成功。这样，就启动了 NFS 的服务，另外还有两个命令，可以方便 NFS 的使用。

其一是 exportfs，它可以重新扫描“/etc/exports”，使用户在修改了“/etc/exports”配置文件不需要每次重启 NFS 服务。其格式为：

exportfs [选项]

exportfs 的常见选项如表 5.2 所示。

表 5.2　　常见选项

选　项	参 数 含 义
-a	全部挂载（或卸载）/etc/exports 中的设定文件目录
-r	重新挂载/etc/exports 中的设定文件目录
-u	卸载某一目录
-v	在 export 的时候，将共享的目录显示到屏幕上

另外一个是 showmount 命令，它用于当前的挂载情况。其格式为：

```
showmount [选项] hostname
```

showmount 的常见选项如表 5.3 所示。

表 5.3　　常见选项

选　项	参 数 含 义
-a	在屏幕上显示目前主机与客户端所连上来的使用目录状态
-e	显示 hostname 中/etc/exports 里设定的共享目录

5.2 U-Boot 移植

5.2.1 Bootloader 介绍

1．概念

简单地说，Bootloader 就是在操作系统内核运行之前运行的一段程序，它类似于 PC 机中的 BIOS 程序。通过这段程序，可以完成硬件设备的初始化，并建立内存空间的映射图的功能，从而将系统的软硬件环境带到一个合适的状态，为最终调用系统内核做好准备。

通常，Bootloader 是严重地依赖于硬件实现的，特别是在嵌入式中。因此，在嵌入式世界里建立一个通用的 Bootloader 几乎是不可能的。尽管如此，仍然可以对 Bootloader 归纳出一些通用的概念来指导用户特定的 Bootloader 设计与实现。

（1）Bootloader 所支持的 CPU 和嵌入式开发板

每种不同的 CPU 体系结构都有不同的 Bootloader。有些 Bootloader 也支持多种体系结构的 CPU，如后面要介绍的 U-Boot 就同时支持 ARM 体系结构和 MIPS 体系结构。除了依赖于 CPU 的体系结构外，Bootloader 实际上也依赖于具体的嵌入式板级设备的配置。

（2）Bootloader 的安装媒介

系统加电或复位后，所有的 CPU 通常都从某个由 CPU 制造商预先安排的地址上取指令。而基于 CPU 构建的嵌入式系统通常都有某种类型的固态存储设备（比如 ROM、EEPROM 或 FLASH 等）被映射到这个预先安排的地址上。因此在系统加电后，CPU 将首先执行 Bootloader 程序。

（3）Bootloader 的启动过程分为单阶段和多阶段两种。通常多阶段的 Bootloader 能提供更为复杂的功能，以及更好的可移植性。

（4）Bootloader 的操作模式。大多数 Bootloader 都包含两种不同的操作模式：“启动加载”模式和“下载”模式，这种区别仅对于开发人员才有意义。

- 启动加载模式：这种模式也称为“自主”模式。也就是 Bootloader 从目标机上的某个固态存储设备上将操作系统加载到 RAM 中运行，整个过程并没有用户的介入。这种模式是嵌入式产品发布时的通用模式。
- 下载模式：在这种模式下，目标机上的 Bootloader 将通过串口连接或网络连接等通信手段从主机（Host）下载文件，比如：下载内核映像和根文件系统映像等。从主机下载的文件通常首先被 Bootloader 保存到目标机的 RAM 中，然后再被 Bootloader 写到目标机上的 FLASH 类固态存储设备中。Bootloader 的这种模系统是在更新时使用。工作于这种模式下的 Bootloader 通常都会向它的终端用户提供一个简单的命令行接口。

（5）Bootloader 与主机之间进行文件传输所用的通信设备及协议，最常见的情况就是，目标机上的 Bootloader 通过串口与主机之间进行文件传输，传输协议通常是 xmodem/ymodem/zmodem 协议中的一种。但是，串口传输的速度是有限的，因此通过以太网连接并借助 TFTP 协议来下载文件是个更好的选择。

2．Bootloader 启动流程

Bootloader 的启动流程一般分为两个阶段：stage1 和 stage2，下面分别对这两个阶段进行讲解：

（1）Bootloader 的 stage1

在 stage1 中 Bootloader 主要完成以下工作。

- 基本的硬件初始化，包括屏蔽所有的中断、设置 CPU 的速度和时钟频率、RAM 初始化、初始化 LED、关闭 CPU 内部指令和数据 cache 灯。
- 为加载 stage2 准备 RAM 空间，通常为了获得更快的执行速度，通常把 stage2 加载到 RAM 空间中来执行，因此必须为加载 Bootloader 的 stage2 准备好一段可用的 RAM 空间范围。
- 拷贝 stage2 到 RAM 中，在这里要确定两点：①stage2 的可执行映像在固态存储设备的存放起始地址和终止地址；②RAM 空间的起始地址。
- 设置堆栈指针 sp，这是为执行 stage2 的 C 语言代码做好准备。

（2）Bootloader 的 stage2

在 stage2 中 Bootloader 主要完成以下工作。

- 用汇编语言跳转到 main 入口函数

由于 stage2 的代码通常用 C 语言来实现，目的是实现更复杂的功能和取得更好的代码可读性和可移植性。但是与普通 C 语言应用程序不同的是，在编译和链接 Bootloader 这样的程序时，不能使用 glibc 库中的任何支持函数。

- 初始化本阶段要使用到的硬件设备，包括初始化串口、初始化计时器等。在初始化这些设备之前、可以输出一些打印信息。
- 检测系统的内存映射，所谓内存映射就是指在整个 4GB 物理地址空间中有指出哪些地址范围被分配用来寻址系统的 RAM 单元。
- 加载内核映像和根文件系统映像，这里包括规划内存占用的布局和从 Flash 上拷贝数据。
- 设置内核的启动参数。

5.2.2 U-Boot 概述

1．U-Boot 简介

U-Boot（UniversalBootloader），是遵循 GPL 条款的开放源码项目。它是从 FADSROM、8xxROM、PPCBOOT 逐步发展演化而来。其源码目录、编译形式与 Linux 内核很相似，事实上，不少 U-Boot 源码就是相应的 Linux 内核源程序的简化，尤其是一些设备的驱动程序，这从 U-Boot 源码的注释中能体现这一点。但是 U-Boot 不仅仅支持嵌入式 Linux 系统的引导，而且还支持 NetBSD、VxWorks、QNX、RTEMS、ARTOS、LynxOS 嵌入式操作系统。其目前要支持的目标操作系统是 OpenBSD、NetBSD、FreeBSD、4.4BSD、Linux、SVR4、Esix、Solaris、Irix、SCO、Dell、NCR、VxWorks,LynxOS、pSOS、QNX、RTEMS、ARTOS。这是 U-Boot 中 Universal 的一层含义，另外一层含义则是 U-Boot 除了支持 PowerPC 系列的处理器外，还能支持 MIPS、x86、ARM、NIOS、XScale 等诸多常用系列的处理器。这两个特点正

是 U-Boot 项目的开发目标，即支持尽可能多的嵌入式处理器和嵌入式操作系统。就目前为止，U-Boot 对 PowerPC 系列处理器支持最为丰富，对 Linux 的支持最完善。

2．U-Boot 特点

U-Boot 的特点如下。

- 开放源码；
- 支持多种嵌入式操作系统内核，如 Linux、NetBSD、VxWorks、QNX、RTEMS、ARTOS、LynxOS；
- 支持多个处理器系列，如 PowerPC、ARM、x86、MIPS、XScale；
- 较高的可靠性和稳定性；
- 高度灵活的功能设置，适合 U-Boot 调试，操作系统不同引导要求，产品发布等；
- 丰富的设备驱动源码，如串口、以太网、SDRAM、FLASH、LCD、NVRAM、EEPROM、RTC、键盘等；
- 较为丰富的开发调试文档与强大的网络技术支持。

3．U-Boot 主要功能

U-Boot 可支持的主要功能列表。

- 系统引导：支持 NFS 挂载、RAMDISK（压缩或非压缩）形式的根文件系统。支持 NFS 挂载，并从 FLASH 中引导压缩或非压缩系统内核。
- 基本辅助功能：强大的操作系统接口功能；可灵活设置、传递多个关键参数给操作系统，适合系统在不同开发阶段的调试要求与产品发布，尤其对 Linux 支持最为强劲；支持目标板环境参数多种存储方式，如 FLASH、NVRAM、EEPROM；CRC32 校验，可校验 FLASH 中内核、RAMDISK 镜像文件是否完好。
- 设备驱动：串口、SDRAM、FLASH、以太网、LCD、NVRAM、EEPROM、键盘、USB、PCMCIA、PCI、RTC 等驱动支持。
- 上电自检功能：SDRAM、FLASH 大小自动检测；SDRAM 故障检测；CPU 型号。
- 特殊功能：XIP 内核引导。

5.2.3 U-Boot 源码导读

1．U-Boot 源码结构

U-Boot 源码结构如图 5.26 所示。

- board：和一些已有开发板有关的文件，比如 Makefile 和 U-Boot.lds 等都和具体开发板的硬件和地址分配有关。
- common：与体系结构无关的文件，实现各种命令的 C 文件。

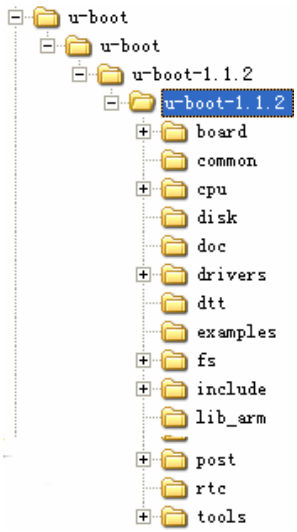


图 5.26 U-Boot 源码结构

- cpu：CPU 相关文件，其中的子目录都是以 U-BOOT 所支持的 CPU 为名，比如有子目录 arm926ejs、mips、mpc8260 和 nios 等，每个特定的子目录中都包括 cpu.c 和 interrupt.c，start.S。其中 cpu.c 初始化 CPU、设置指令 Cache 和数据 Cache 等；interrupt.c 设置系统的各种中断和异常，比如快速中断、开关中断、时钟中断、软件中断、预取中止和未定义指令等；start.S 是 U-BOOT 启动时执行的第一个文件，它主要是设置系统堆栈和工作方式，为进入 C 程序奠定基础。
- disk：disk 驱动的分区处理代码。
- doc：文档。
- drivers：通用设备驱动程序，比如各种网卡、支持 CFI 的 Flash、串口和 USB 总线等。
- fs：支持文件系统的文件，U-BOOT 现在支持 cramfs、fat、fdos、jffs2 和 registerfs。
- include：头文件，还有对各种硬件平台支持的汇编文件，系统的配置文件和对文件系统支持的文件。
- net：与网络有关的代码，BOOTP 协议、TFTP 协议、RARP 协议和 NFS 文件系统的实现。
- lib_arm：与 ARM 体系结构相关的代码。
- tools：创建 S-Record 格式文件和 U-BOOT images 的工具。

2．U-Boot 重要代码

（1）cpu/arm920t/start.S

这是 U-Boot 的起始位置。在这个文件中设置了处理器的状态、初始化中断和内存时序等，从 Flash 中跳转到定位好的内存位置执行。

```
    .globl_start
    _start:     b        reset
          ldr   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

    _undefined_instruction: .word 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
```

```
_TEXT_BASE:
    .word   TEXT_BASE

.globl _armboot_start
_armboot_start:
    .word _start

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

.globl _bss_end
_bss_end:
     .word _end
reset:
      /*
       * set the cpu to SVC32 mode
      */
      mrs   r0,cpsr
      bic   r0,r0,#0x1f
      orr   r0,r0,#0xd3
      msr   cpsr,r0
relocate:                 /* relocate U-Boot to RAM     */
      adr   r0, _start     /* r0 <- current position of code   */
      ldr   r1, _TEXT_BASE     /* test if we run from flash or RAM */
      cmp    r0, r1              /* don't reloc during debug       */
      beq    stack_setup

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

      /* 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 /* bdinfo                    */

clear_bss:
      ldr   r0, _bss_start     /* find start of bss segment       */
      ldr   r1, _bss_end       /* stop here                     */
      mov   r2, #0x00000000    /* clear                         */

clbss_l:str r2, [r0]       /* clear loop...                 */
      add   r0, r0, #4
      cmp   r0, r1
      bne   clbss_l
      ldr   pc, _start_armboot

_start_armboot: .word start_armboot
```

（2）interrupts.c

这个文件是处理中断的，如打开和关闭中断等。

```
#ifdef CONFIG_USE_IRQ
/* enable IRQ interrupts */
void enable_interrupts (void)
{
     unsigned long temp;
     __asm__ __volatile__("mrs %0, cpsr\n"
                          "bic %0, %0, #0x80\n"
                          "msr cpsr_c, %0"
                          : "=r" (temp)
                          :
                          : "memory");
}
```

```
/*
 * disable IRQ/FIQ interrupts
 * returns true if interrupts had been enabled before we disabled them
 */
int disable_interrupts (void)
{
     unsigned long old,temp;
     __asm__ __volatile__("mrs %0, cpsr\n"
                          "orr %1, %0, #0xc0\n"
                          "msr cpsr_c, %1"
                          : "=r" (old), "=r" (temp)
                          :
                          : "memory");
     return (old & 0x80) == 0;
}
#else
void enable_interrupts (void)
{
     return;
}
int disable_interrupts (void)
{
     return 0;
}
#endif
void show_regs (struct pt_regs *regs)
{
     unsigned long flags;
     const char *processor_modes[] = {
     "USER_26", "FIQ_26",   "IRQ_26",   "SVC_26",
     "UK4_26",  "UK5_26",   "UK6_26",   "UK7_26",
     "UK8_26",  "UK9_26",   "UK10_26",  "UK11_26",
     "UK12_26", "UK13_26",  "UK14_26",  "UK15_26",
     "USER_32", "FIQ_32",   "IRQ_32",   "SVC_32",
     "UK4_32",  "UK5_32",   "UK6_32",   "ABT_32",
     "UK8_32",  "UK9_32",   "UK10_32",  "UND_32",
     "UK12_32", "UK13_32",  "UK14_32",  "SYS_32",
     };
```

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

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

（3）cpu.c

这个文件是对处理器进行操作，如下所示：

```
int cpu_init (void)
{
     /*
      * setup up stacks if necessary
     */
#ifdef CONFIG_USE_IRQ
     DECLARE_GLOBAL_DATA_PTR;

     IRQ_STACK_START=_armboot_start - CFG_MALLOC_LEN - CFG_GBL_DATA_SIZE - 4;
     FIQ_STACK_START = IRQ_STACK_START - CONFIG_STACKSIZE_IRQ;
#endif
     return 0;
}
int cleanup_before_linux (void)
{
     /*
      * this function is called just before we call linux
      * it prepares the processor for linux
      *
      * we turn off caches etc ...
      */
```

```
    unsigned long i;

    disable_interrupts ();

    /* turn off I/D-cache */
    asm ("mrc p15, 0, %0, c1, c0, 0":"=r" (i));
    i &= ~(C1_DC | C1_IC);
    asm ("mcr p15, 0, %0, c1, c0, 0": :"r" (i));

    /* flush I/D-cache */
    i = 0;
    asm ("mcr p15, 0, %0, c7, c7, 0": :"r" (i));
    return (0);
}

OUTPUT_ARCH(arm)
ENTRY(_start)
SECTIONS
{
    . = 0x00000000;
       . = ALIGN(4);
    .text      :
    {
      cpu/arm920t/start.o  (.text)
      *(.text)
    }

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

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

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

    __u_boot_cmd_start = .;
    .u_boot_cmd : { *(.u_boot_cmd) }
```

```
        __u_boot_cmd_end = .;

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

（4）memsetup.S

这个文件是用于配置开发板参数的，如下所示：

```
    //memsetup.c
        /* memory control configuration */
        /* make r0 relative the current location so that it */
        /* reads SMRDATA out of FLASH rather than memory ! */
        ldr     r0, =SMRDATA
        ldr r1, _TEXT_BASE
        sub r0, r0, r1
        ldr r1, =BWSCON /* Bus Width Status Controller */
        add     r2, r0, #52
    0:
        ldr     r3, [r0], #4
        str     r3, [r1], #4
        cmp     r2, r0
        bne     0b

        /* everything is fine now */
        mov pc, lr


        .ltorg
```

5.2.4 U-Boot 移植主要步骤

1．配置主板

阅读 Makefile 文件，在 Makefile 文件中添加两行，如下所示：

```
    s3c2410_config: unconfig @./mkconfig $(@:_config=) arm arm920t s3c2410
```

其中 ARM 是 CPU 的种类，arm920t 是 ARM CPU 对应的代码目录，s3c2410 是自己主板对应的目录。

（1）在 board 目录中建立 s3c2410 目录，并复制 smdk2410 目录中的内容（cp smdk2410/* s3c2410）。

- 在 include/configs/目录下复制 smdk2410.h（cp smdk2410.h s3c2410.h）。
- 修改 ARM 编译器的目录名及前缀。
- 完成之后，可以测试配置。

#make s3c2410_config

#make

（2）修改程序连接地址

在 board/s3c2410 中有一个 config.mk 文件，它是用于设置程序连接的起始地址，因为会在 U-Boot 中增加功能，所以留下 6MB 的空间，并修改 33F80000 为 33A00000。

为了以后能用 U-Boot 的 GO 命令执行修改过的用 loadb 或 tftp 下载的 U-Boot：在 board/s3c2410 的 memsetup.S 中标记符”0:”上加入 5 句：

```
mov r3, pc
ldr r4, =0x3FFF0000
and r3, r3, r4 //以上 3 句得到实际起动的内存地址
aad r0, r0, r3 //用 GO 命令调试 u-boot 时，启动地址在 RAM
add r2, r2, r3 //把初始化内存信息的地址，加上实际起动地址
```

2．设置 FLASH 和 SDRAM 时序

在 include/configs/s3c2410 中加入以下几句：

```
#define CONFIG_DRIVER_DM9000 1 /* we have a CS8900 on-board */
#define DM9000_BASE 0x08000000
#define DM9000_BUS16 1 /* the Linux driver does accesses as shorts */
#define CONFIG_ETHADDR 08:00:3e:26:0a:5b
#define CONFIG_NETMASK 255.255.255.0
#define CONFIG_IPADDR 192.168.2.120
#define CONFIG_SERVERIP 192.168.2.122
```

5.2.5 U-Boot 常见命令

- ？：得到所有命令列表。
- help：help usb，列出 USB 功能的使用说明。
- ping：测试与对方主机的网络是否正常，这里只能开发板 PING 别的机器。
- setenv：设置环境变量。
- saveenv：保存环境变量。在设置好环境变量以后，要保存变量值。
- ftp：tftp 32000000 vmlinux，把 server（IP=环境变量中设置的 serverip）中/tftpdroot/下的 vmlinux 通过 TFTP 读入到物理内存 32000000 处。
- kgo：启动没有压缩的 linux 内核，kgo 32000000。
- bootm：启动 UBOOT TOOLS 制作的压缩 LINUX 内核，bootm 3200000。
- protect：对 FLASH 进行写保护或取消写保护，protect on 1:0-3（就是对第一块 FLASH

的 0-3 扇区进行保护），protect off 1:0-3 取消写保护。

- erase：删除 FLASH 的扇区，erase 1:0-2（就是对每一块 FLASH 的 0-2 扇区进行删除）。
- cp：在内存中复制内容，cp 32000000 0 40000（把内存中 0x32000000 开始的 0x40000 字节复制到 0x0 处）。
- mw：对 RAM 中的内容写操作，如 mw 32000000 ff 10000 是指把内存 0x32000000 开始的 0x10000 字节设为 0xFF。
- md：修改 RAM 中的内容，如 md 32000000 指内存的起始地址为。
- usb：usb start 是指启动 usb 功能；usb info 是指列出设备；usb scan 是指扫描 usb storage（U 盘）设备。
- fatls：列出 DOS FAT 文件系统，如：fatls usb 0 是指列出第一块 U 盘中的文件。
- fatload：读入 FAT 中的一个文件，如：fatload usb 0:0 32000000 aa.txt。
- flinfo：列出 flash 的信息。
- loadb：准备用 KERMIT 协议接收来自 kermit 或超级终端传送的文件。
- nfs：nfs 32000000 192.168.0.2:aa.txt，把 192.168.0.2（LINUX 的 NFS 文件系统）中的 NFS 文件系统中的 aa.txt 读入内存 0x32000000 处。

5.3 实验内容——移植 Linux 内核

1．实验目的

通过移植 Linux 内核，熟悉嵌入式开发环境的搭建和 Linux 内核的编译配置。由于具体步骤在前面已经详细讲解过了，因此，相关部分请读者查阅本章前面内容。

2．实验内容

首先在 Linux 环境下配置 minicom，使之能够正常显示串口的信息。然后再编译配置 Linux2.6 内核，并下载到开发板上运行。

3．实验步骤

（1）设置 minicom，按键“CTRL-A O”配置相应参数。
（2）连接开发板与主机，查看串口是否有正确输出。
（3）查看 Linux 内核顶层的 Makefile，确定相关参数是否正确。
（4）运行“make menuconfig”，进行相应配置。
（5）运行“make dep”。
（6）运行“make zImage”。
（7）将生成映像通过 tftp 或串口下载到开发板中。
（8）启动运行内核。

4．实验结果

开发板能够正确运行新生成的内核映像。

本 章 小 结

本章详细讲解了嵌入式 Linux 开发环境的搭建，包括 minicom 和超级终端的配置，如何下载映像文件到开发板，如何移植嵌入式 Linux 内核以及如何移植 U-Boot。这些都是操作性很强的内容，而且在嵌入式的开发中也是必不可少的一部分，因此希望读者确实掌握。

思考与练习

1．适当更改 Linux 内核配置，再进行编译下载查看结果。
2．配置 NFS 服务。

“黑色经典”系列之《嵌入式 Linux 应用程序开发详解》

第 6 章 文件 I/O 编程

本章目标

在搭建起嵌入式开发环境之后，从本章开始，读者将真正开始学习嵌入式 Linux 的应用开发。由于嵌入式 Linux 是经 Linux 裁减而来的，它的系统调用及用户编程接口 API 与 Linux 基本是一致的，因此，在以后的章节中，笔者将首先介绍 Linux 中相关内容的基本编程开发，主要讲解与嵌入式 Linux 中一致的部分，然后再将程序移植到嵌入式的开发板上运行。因此，没有开发板的读者也可以先在 Linux 上开发相关应用程序，这对以后进入嵌入式 Linux 的实际开发是十分有帮助的。本章主要讲解文件 I/O 相关开发，经过本章的学习，读者将会掌握以下内容。

- 掌握 Linux 中系统调用的基本概念
- 掌握 Linux 中用户编程接口（API）及系统命令的相互关系
- 掌握文件描述符的概念
- 掌握 Linux 下文件相关的不带缓存 I/O 函数的使用
- 掌握 Linux 下设备文件读写方法
- 掌握 Linux 中对串口的操作
- 熟悉 Linux 中标准文件 I/O 函数的使用

6.1 Linux 系统调用及用户编程接口（API）

由于本章是讲解 Linux 编程开发的第 1 章，因此希望读者更加明确 Linux 系统调用和用户编程接口（API）的概念。在了解了这些之后，会对 Linux 以及 Linux 的应用编程有更深入地理解。

6.1.1 系统调用

所谓系统调用是指操作系统提供给用户程序调用的一组“特殊”接口，用户程序可以通过这组“特殊”接口来获得操作系统内核提供的服务。例如用户可以通过进程控制相关的系统调用来创建进程、实现进程调度、进程管理等。

在这里，为什么用户程序不能直接访问系统内核提供的服务呢？这是由于在 Linux 中，为了更好地保护内核空间，将程序的运行空间分为内核空间和用户空间（也就是常称的内核态和用户态），它们分别运行在不同的级别上，在逻辑上是相互隔离的。因此，用户进程在通常情况下不允许访问内核数据，也无法使用内核函数，它们只能在用户空间操作用户数据，调用用户空间的函数。

但是，在有些情况下，用户空间的进程需要获得一定的系统服务（调用内核空间程序），这时操作系统就必须利用系统提供给用户的“特殊接口”——系统调用规定用户进程进入内核空间的具体位置。进行系统调用时，程序运行空间需要从用户空间进入内核空间，处理完后再返回到用户空间。

Linux 系统调用部分是非常精简的系统调用（只有 250 个左右），它继承了 UNIX 系统调用中最基本和最有用的部分。这些系统调用按照功能逻辑大致可分为进程控制、进程间通信、文件系统控制、系统控制、存储管理、网络管理、socket 控制、用户管理等几类。

6.1.2 用户编程接口（API）

前面讲到的系统调用并不是直接与程序员进行交互的，它仅仅是一个通过软中断机制向内核提交请求，以获取内核服务的接口。在实际使用中程序员调用的通常是用户编程接口——API，也就是本书后面要讲到的 API 函数。但并不是所有的函数都一一对应一个系统调用，有时，一个 API 函数会需要几个系统调用来共同完成函数的功能，甚至还有一些 API 函数不需要调用相应的系统调用（因此它所完成的不是内核提供的服务）。

在 Linux 中，用户编程接口（API）遵循了在 UNIX 中最流行的应用编程界面标准——POSIX 标准。POSIX 标准是由 IEEE 和 ISO/IEC 共同开发的标准系统。该标准基于当时现有的 UNIX 实践和经验，描述了操作系统的系统调用编程接口（实际上就是 API），用于保证应用程序可以在源代码一级上在多种操作系统上移植运行。这些系统调用编程接口主要是通过 C 库（libc）实现的。

6.1.3 系统命令

以上讲解了系统调用、用户编程接口（API）的概念，分析了它们之间的相互关系，那么，读者在第 2 章中学到的那么多的 Shell 系统命令与它们之间又是怎样的关系呢？

系统命令相对 API 更高了一层，它实际上一个可执行程序，它的内部引用了用户编程接口（API）来实现相应的功能。它们之间的关系如下图 6.1 所示。

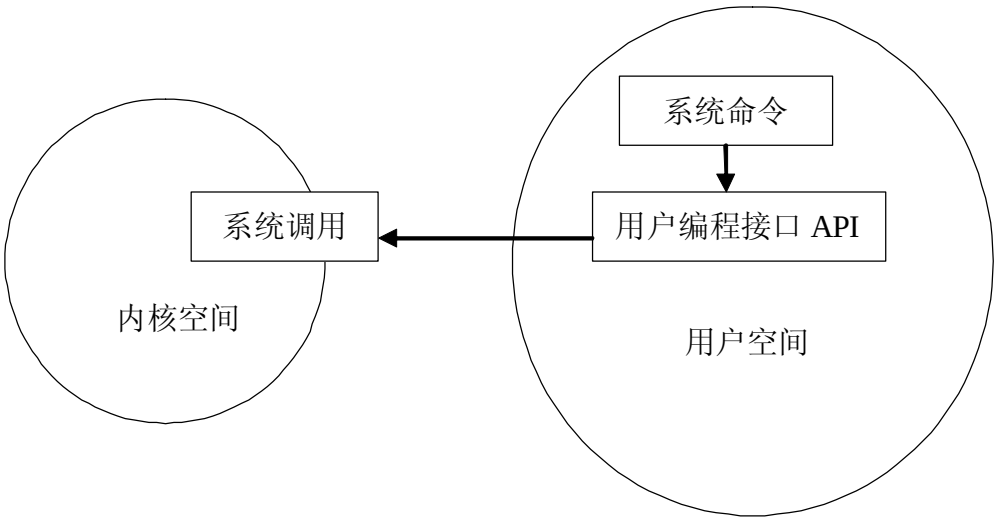


图 6.1　系统调用、API 及系统命令之间的关系

6.2　Linux 中文件及文件描述符概述

正如第 1 章中所述，在 Linux 中对目录和设备的操作都等同于文件的操作，因此，大大简化了系统对不同设备的处理，提高了效率。Linux 中的文件主要分为 4 种：普通文件、目录文件、链接文件和设备文件。

那么，内核如何区分和引用特定的文件呢？这里用到的就是一个重要的概念——文件描述符。对于 Linux 而言，所有对设备和文件的操作都使用文件描述符来进行的。文件描述符是一个非负的整数，它是一个索引值，并指向内核中每个进程打开文件的记录表。当打开一个现存文件或创建一个新文件时，内核就向进程返回一个文件描述符；当需要读写文件时，也需要把文件描述符作为参数传递给相应的函数。

通常，一个进程启动时，都会打开 3 个文件：标准输入、标准输出和标准出错处理。这 3 个文件分别对应文件描述符为 0、1 和 2（也就是宏替换 STDIN_FILENO、STDOUT_FILENO 和 STDERR_FILENO，鼓励读者使用这些宏替换）。

基于文件描述符的 I/O 操作虽然不能移植到类 Linux 以外的系统上去（如 Windows），但它往往是实现某些 I/O 操作的惟一途径，如 Linux 中低级文件操作函数、多路 I/O、TCP/IP 套接字编程接口等。同时，它们也很好地兼容 POSIX 标准，因此，可以很方便地移植到任何 POSIX 平台上。基于文件描述符的 I/O 操作是 Linux 中最常用的操作之一，希望读者能够很好地掌握。

6.3　不带缓存的文件 I/O 操作

本节主要介绍不带缓存的文件 I/O 操作，主要用到 5 个函数：open、read、write、lseek 和 close。这里的不带缓存是指每一个函数都只调用系统中的一个函数。这些函数虽然不是

ANSI C 的组成部分，但是是 POSIX 的组成部分。

6.3.1 open 和 close

（1）open 和 close 函数说明

open 函数是用于打开或创建文件，在打开或创建文件时可以指定文件的属性及用户的权限等各种参数。

close 函数是用于关闭一个打开文件。当一个进程终止时，它所有已打开的文件都由内核自动关闭，很多程序都使用这一功能而不显示地关闭一个文件。

（2）open 和 close 函数格式

open 函数的语法格式如表 6.1 所示。

表 6.1 open 函数语法要点

所需头文件	#include <sys/types.h> // 提供类型 pid_t 的定义 #include <sys/stat.h> #include <fcntl.h>	
续表		
函数原型	int open(const char *pathname，flags，int perms)	
函数传入值	pathname	被打开的文件名（可包括路径名）
	flag：文件打开的方式	O_RDONLY：只读方式打开文件
		O_WRONLY：可写方式打开文件
		O_RDWR：读写方式打开文件
		O_CREAT：如果该文件不存在，就创建一个新的文件，并用第三个参数为其设置权限
		O_EXCL：如果使用 O_CREAT 时文件存在，则可返回错误消息。这一参数可测试文件是否存在
		O_NOCTTY：使用本参数时，如文件为终端，那么终端不可以作为调用 open()系统调用的那个进程的控制终端
		O_TRUNC：如文件已经存在，并且以只读或只写成功打开，那么会先全部删除文件中原有数据
		O+APPEND：以添加方式打开文件，在打开文件的同时，文件指针指向文件的末尾
	perms	被打开文件的存取权限，为 8 进制表示法
函数返回值	成功：返回文件描述符 失败：–1	

在 open 函数中，flag 参数可通过“|”组合构成，但前 3 个函数不能相互组合。perms 是文件的存取权限，采用 8 进制表示法，相关内容读者可参见第 2 章。

close 函数的语法格式如下表 6.2 所示。

表 6.2 close 函数语法要点

所需头文件	#include <unistd.h>
函数原型	int close(int fd)
函数输入值	fd：文件描述符
函数返回值	0：成功 –1：出错

（3）open 和 close 函数使用实例

下面实例中的 open 函数带有 3 个 flag 参数：O_CREAT、O_TRUNC 和 O_WRONLY，这样就可以对不同的情况指定相应的处理方法。另外，这里对该文件的权限设置为 0600。其源码如下所示：

```
/*open.c*/
#include <unistd.h>
#include <sys/types.h>
#include <sys/stat.h>
#include <fcntl.h>
#include <stdlib.h>
#include <stdio.h>

int main(void)
{
      int fd;
/*调用 open 函数，以可读写的方式打开，注意选项可以用“|”符号连接*/
      if((fd = open("/tmp/hello.c", O_CREAT | O_TRUNC | O_WRONLY , 0600 ))<0){
            perror("open:");
            exit(1);
      }
      else{
            printf("Open file: hello.c %d\n",fd);
      }
      if( close(fd) < 0 ){
            perror("close:");
            exit(1);
      }
      else
            printf("Close hello.c\n");
      exit(0);
}
[root@(none) 1]# ./open
```

```
Open file: hello.c 3
Close hello.c
[root@(none) tmp]# ls -l |grep hello.c
-rw-------    1 root     root           0 Dec  4 00:59 hello.c
```

经过交叉编译后，将文件下载到目标板，则该可执行文件运行后就能在目录/tmp 下新建一个 hello.c 的文件，其权限为 0600。

注意 open 函数返回的文件描述符一定是最小的未用文件描述符。由于一个进程在启动时自动打开了 0、1、2 三个文件描述符，因此，该文件运行结果中返回的文件描述符为 3。读者可以尝试在调用 open 函数之前，加依据 close(0)，则此后在 open 函数时返回的文件描述符为 0（若关闭文件描述符 1，则在执行时会由于没有标准输出文件而无法输出）。

6.3.2 read、write 和 lseek

（1）read、write 和 lseek 函数作用

read 函数是用于将指定的文件描述符中读出数据。当从终端设备文件中读出数据时，通常一次最多读一行。

write 函数是用于向打开的文件写数据，写操作从文件的当前位移量处开始。若磁盘已满或超出该文件的长度，则 write 函数返回失败。

lseek 函数是用于在指定的文件描述符中将文件指针定位到相应的位置。

（2）read 和 write 函数格式

read 函数的语法格式如下表 6.3 所示。

表 6.3 read 函数语法要点

所需头文件	#include <unistd.h>
函数原型	ssize_t read(int fd,void *buf,size_t count)
函数传入值	fd：文件描述符
	buf：指定存储器读出数据的缓冲区
	count：指定读出的字节数
函数返回值	成功：读到的字节数 0：已到达文件尾 −1：出错

在读普通文件时，若读到要求的字节数之前已到达文件的尾部，则返回的字节数会小于希望读出的字节数。

write 函数的语法格式如下表 6.4 所示。

表 6.4 write 函数语法要点

所需头文件	#include <unistd.h>
函数原型	ssize_t write(int fd,void *buf,size_t count)

函数传入值	fd：文件描述符
	buf：指定存储器写入数据的缓冲区
	count：指定读出的字节数
函数返回值	成功：已写的字节数 –1：出错

在写普通文件时，写操作从文件的当前位移处开始。

lseek 函数的语法格式如下表 6.5 所示。

表 6.5　　lseek 函数语法要点

所需头文件	#include <unistd.h> #include <sys/types.h>	
函数原型	off_t lseek(int fd,off_t offset,int whence)	
函数传入值	fd：文件描述符	
	offset：偏移量，每一读写操作所需要移动的距离，单位是字节的数量，可正可负（向前移，向后移）	
续表		
	whence：当前位置的基点	SEEK_SET：当前位置为文件的开头，新位置为偏移量的大小
		SEEK_CUR：当前位置为文件指针的位置，新位置为当前位置加上偏移量
		SEEK_END：当前位置为文件的结尾，新位置为文件的大小加上偏移量的大小
函数返回值	成功：文件的当前位移 –1：出错	

（3）函数使用实例

该示例程序首先打开上一节中创建的文件，然后对此文件进行读写操作（记得要将文件打开属性改为可读写，将文件权限也做相应更改）。接着，写入“Hello! I'm writing to this file!”，此时文件指针位于文件尾部。接着在使用 lseek 函数将文件指针移到文件开始处，并读出 10 个字节并将其打印出来。程序源代码如下所示：

```
/*write.c*/
#include <unistd.h>
#include <sys/types.h>
#include <sys/stat.h>
#include <fcntl.h>
#include <stdlib.h>
#include <stdio.h>
#include <string.h>
#define MAXSIZE
```

```
int main(void)
{
     int i,fd,size,len;
     char *buf="Hello! I'm writing to this file!";
     char buf_r[10];
     len = strlen(buf);
/*首先调用 open 函数，并指定相应的权限*/
     if((fd = open("/tmp/hello.c", O_CREAT | O_TRUNC | O_RDWR,0666 ))<0){
            perror("open:");
            exit(1);
     }
     else
            printf("open file:hello.c %d\n",fd);
/*调用 write 函数，将 buf 中的内容写入到打开的文件中*/
     if((size = write( fd, buf, len)) < 0){
            perror("write:");
            exit(1);
     }
     else
            printf("Write:%s\n",buf);
/*调用 lsseek 函数将文件指针移到文件起始，并读出文件中的 10 个字节*/
     lseek( fd, 0, SEEK_SET );
     if((size = read( fd, buf_r, 10))<0){
            perror("read:");
            exit(1);
     }
     else
            printf("read form file:%s\n",buf_r);
     if( close(fd) < 0 ){
            perror("close:");
            exit(1);
     }
     else
            printf("Close hello.c\n");
     exit(0);
}
[root@(none) 1]# ./write
open file:hello.c 3
```

```
Write:Hello! I'm writing to this file!
read form file:Hello! I'm
Close hello.c
[root@(none) 1]# cat /tmp/hello.c
Hello! I'm writing to this file!
```

6.3.3 fcntl

（1）fcntl 函数说明

前面的这 5 个基本函数实现了文件的打开、读写等基本操作，这一节将讨论的是，在文件已经共享的情况下如何操作，也就是当多个用户共同使用、操作一个文件的情况，这时，Linux 通常采用的方法是给文件上锁，来避免共享的资源产生竞争的状态。

文件锁包括建议性锁和强制性锁。建议性锁要求每个上锁文件的进程都要检查是否有锁存在，并且尊重已有的锁。在一般情况下，内核和系统都不使用建议性锁。强制性锁是由内核执行的锁，当一个文件被上锁进行写入操作的时候，内核将阻止其他任何文件对其进行读写操作。采用强制性锁对性能的影响很大，每次读写操作都必须检查是否有锁存在。

在 Linux 中，实现文件上锁的函数有 lock 和 fcntl，其中 flock 用于对文件施加建议性锁，而 fcntl 不仅可以施加建议性锁，还可以施加强制锁。同时，fcntl 还能对文件的某一记录进行上锁，也就是记录锁。

记录锁又可分为读取锁和写入锁，其中读取锁又称为共享锁，它能够使多个进程都能在文件的同一部分建立读取锁。而写入锁又称为排斥锁，在任何时刻只能有一个进程在文件的某个部分上建立写入锁。当然，在文件的同一部分不能同时建立读取锁和写入锁。

注意 fcntl 是一个非常通用的函数，它还可以改变文件进程各方面的属性，在本节中，主要介绍它建立记录锁的方法，关于它其他用户感兴趣的读者可以参看 fcntl 手册。

（2）fcntl 函数格式

用于建立记录锁的 fcntl 函数格式如表 6.6 所示。

表 6.6 fcntl 函数语法要点

所需头文件	#include <sys/types.h> #include <unistd.h> #include <fcntl.h>	
函数原型	int fcnt1(int fd,int cmd,struct flock *lock)	
函数传入值	fd：文件描述符	
	cmd	F_DUPFD：复制文件描述符
		F_GETFD：获得 fd 的 close-on-exec 标志，若标志未设置，则文件经过 exec 函数之后仍保持打开状态
		F_SETFD：设置 close-on-exec 标志，该标志以参数 arg 的 FD_CLOEXEC 位决定
		F_GETFL：得到 open 设置的标志
		F_SETFL：改变 open 设置的标志

		F_GETFK：根据 lock 描述，决定是否上文件锁
		F_SETFK：设置 lock 描述的文件锁
		F_SETLKW：这是 F_SETLK 的阻塞版本（命令名中的 W 表示等待（wait））。如果存在其他锁，则调用进程睡眠；如果捕捉到信号则睡眠中断
		F_GETOWN：检索将收到 SIGIO 和 SIGURG 信号的进程号或进程组号
		F_SETOWN：设置进程号或进程组号
	Lock：结构为 flock，设置记录锁的具体状态，后面会详细说明	
函数返回值	成功：0 –1：出错	

这里，lock 的结构如下所示：

```
Struct flock{
short l_type;
off_t l_start;
short l_whence;
off_t l_len;
pid_t l_pid;
}
```

lock 结构中每个变量的取值含义如表 6.7 所示。

表 6.7 lock 结构变量取值

l_type	F_RDLCK：读取锁（共享锁）
	F_WRLCK：写入锁（排斥锁）
	F_UNLCK：解锁
l_stat	相对位移量（字节）
l_whence：相对位移量的起点（同 lseek 的 whence）。	SEEK_SET：当前位置为文件的开头，新位置为偏移量的大小
	SEEK_CUR：当前位置为文件指针的位置，新位置为当前位置加上偏移量
	SEEK_END：当前位置为文件的结尾，新位置为文件的大小加上偏移量的大小
l_len	加锁区域的长度

小技巧 为加锁整个文件，通常的方法是将 l_start 说明为 0，l_whence 说明为 SEEK_SET，l_len 说明为 0。

（3）fcntl 使用实例

下面首先给出了使用 fcntl 函数的文件记录锁函数。在该函数中，首先给 flock 结构体的对应位赋予相应的值。接着使用两次 fcntl 函数分别用于给相关文件上锁和判断文件是否可以上锁，这里用到的 cmd 值分别为 F_SETLK 和 F_GETLK。

这个函数的源代码如下所示：

```
/*lock_set 函数*/
void lock_set(int fd, int type)
{
      struct flock lock;
      lock.l_whence = SEEK_SET;//赋值 lock 结构体
      lock.l_start = 0;
      lock.l_len =0;
      while(1){
            lock.l_type = type;
/*根据不同的 type 值给文件上锁或解锁*/
            if((fcntl(fd, F_SETLK, &lock)) == 0){
            if( lock.l_type == F_RDLCK )
                  printf("read lock set by %d\n",getpid());
            else if( lock.l_type == F_WRLCK )
                  printf("write lock set by %d\n",getpid());
            else if( lock.l_type == F_UNLCK )
                  printf("release lock by %d\n",getpid());
            return;
            }
/*判断文件是否可以上锁*/
            fcntl(fd, F_GETLK,&lock);
/*判断文件不能上锁的原因*/
            if(lock.l_type != F_UNLCK){
/*/该文件已有写入锁*/
                  if( lock.l_type == F_RDLCK )
                        printf("read lock already set by %d\n",lock.l_pid);
/*该文件已有读取锁*/
                  else if( lock.l_type == F_WRLCK )
                        printf("write lock already set by %d\n",lock.l_pid);
                  getchar();
            }
      }
}
```

下面的实例是测试文件的写入锁，这里首先创建了一个 hello 文件，之后对其上写入锁，最后释放写入锁。代码如下所示：

```
/*fcntl_write.c 测试文件写入锁主函数部分*/
#include <unistd.h>
#include <sys/file.h>
```

```
#include <sys/types.h>
#include <sys/stat.h>
#include <stdio.h>
#include <stdlib.h>

int main(void)
{
     int fd;
/*首先打开文件*/
     fd=open("hello",O_RDWR | O_CREAT, 0666);
     if(fd < 0){
           perror("open");
           exit(1);
     }
/*给文件上写入锁*/
     lock_set(fd, F_WRLCK);
     getchar();
/*给文件接锁*/
     lock_set(fd, F_UNLCK);
     getchar();
     close(fd);
     exit(0);
}
```

为了能够使用多个终端，更好地显示写入锁的作用，本实例主要在 PC 机上测试，读者可将其交叉编译，下载到目标板上运行。下面是在 PC 机上的运行结果。为了使程序有较大的灵活性，笔者采用文件上锁后由用户键入一任意键使程序继续运行。建议读者开启两个终端，并且在两个终端上同时运行该程序，以达到多个进程操作一个文件的效果。在这里，笔者首先运行终端一，请读者注意终端二中的第一句。

终端一：

```
[root@localhost file]# ./fcntl_write
write lock set by 4994

release lock by 4994
```

终端二：

```
[root@localhost file]# ./fcntl_write
write lock already set by 4994
```

```
write lock set by 4997

release lock by 4997
```

由此可见，写入锁为互斥锁，一个时刻只能有一个写入锁存在。
接下来的程序是测试文件的读取锁，原理同上面的程序一样。

```
/*fcntl_read.c 测试文件读取锁主函数部分*/
#include <unistd.h>
#include <sys/file.h>
#include <sys/types.h>
#include <sys/stat.h>
#include <stdio.h>
#include <stdlib.h>

int main(void)
{
     int fd;
     fd=open("hello",O_RDWR | O_CREAT, 0666);
     if(fd < 0){
           perror("open");
           exit(1);
     }
/*给文件上读取锁*/
     lock_set(fd, F_RDLCK);
     getchar();
/*给文件接锁*/
     lock_set(fd, F_UNLCK);
     getchar();
     close(fd);
     exit(0);
}
```

同样开启两个终端，并首先启动终端一上的程序，其运行结果如下所示：
终端一：

```
[root@localhost file]# ./fcntl2
read lock set by 5009

release lock by 5009
```

终端二：

```
[root@localhost file]# ./fcntl2
read lock set by 5010

release lock by 5010
```

读者可以将此结果与写入锁的运行结果相比较，可以看出，读取锁为共享锁，当进程 5009 已设定读取锁后，进程 5010 还可以设置读取锁。

思考 如果在一个终端上运行设置读取锁，则在另一个终端上运行设置写入锁，会有什么结果呢？

6.3.4 select

（1）select 函数说明

前面的 fcntl 函数解决了文件的共享问题，接下来该处理 I/O 复用的情况了。

总的来说，I/O 处理的模型有 5 种。

- **阻塞 I/O 模型：**在这种模型下，若所调用的 I/O 函数没有完成相关的功能就会使进程挂起，直到相关数据到才会出错返回。如常见对管道设备、终端设备和网络设备进行读写时经常会出现这种情况。
- **非阻塞模型：**在这种模型下，当请求的 I/O 操作不能完成时，则不让进程睡眠，而且返回一个错误。非阻塞 I/O 使用户可以调用不会永远阻塞的 I/O 操作，如 open、write 和 read。如果该操作不能完成，则会立即出错返回，且表示该 I/O 如果该操作继续执行就会阻塞。
- **I/O 多路转接模型：**在这种模型下，如果请求的 I/O 操作阻塞，且它不是真正阻塞 I/O，而是让其中的一个函数等待，在这期间，I/O 还能进行其他操作。如本节要介绍的 select 函数和 poll 函数，就是属于这种模型。
- **信号驱动 I/O 模型：**在这种模型下，通过安装一个信号处理程序，系统可以自动捕获特定信号的到来，从而启动 I/O。这是由内核通知用户何时可以启动一个 I/O 操作决定的。
- **异步 I/O 模型：**在这种模型下，当一个描述符已准备好，可以启动 I/O 时，进程会通知内核。现在，并不是所有的系统都支持这种模型。

可以看到，select 的 I/O 多路转接模型是处理 I/O 复用的一个高效的方法。它可以具体设置每一个所关心的文件描述符的条件、希望等待的时间等，从 select 函数返回时，内核会通知用户已准备好的文件描述符的数量、已准备好的条件等。通过使用 select 返回值，就可以调用相应的 I/O 处理函数了。

（2）select 函数格式

Select 函数的语法格式如表 6.8 所示。

表 6.8 fcntl 函数语法要点

所需头文件	#include <sys/types.h> #include <sys/time.h>

	#include <unistd.h>	
函数原型	int select(int numfds,fd_set *readfds,fd_set *writefds，fd_set *exeptfds,struct timeval *timeout)	
函数传入值	numfds：需要检查的号码最高的文件描述符加 1	
	readfds：由 select()监视的读文件描述符集合	
	writefds：由 select()监视的写文件描述符集合	
	exeptfds：由 select()监视的异常处理文件描述符集合	
	timeout	NULL：永远等待，直到捕捉到信号或文件描述符已准备好为止
		具体值：struct timeval 类型的指针，若等待为 timeout 时间还没有文件描符准备好，就立即返回
		0：从不等待，测试所有指定的描述符并立即返回
函数返回值	成功：准备好的文件描述符 –1：出错	

思考 请读者考虑一下如何确定最高的文件描述符？

可以看到，select 函数根据希望进行的文件操作对文件描述符进行了分类处理，这里，对文件描述符的处理主要涉及到 4 个宏函数，如表 6.9 所示。

表 6.9 select 文件描述符处理函数

FD_ZERO(fd_set *set)	清除一个文件描述符集
FD_SET(int fd,fd_set *set)	将一个文件描述符加入文件描述符集中
FD_CLR(int fd,fd_set *set)	将一个文件描述符从文件描述符集中清除
FD_ISSET(int fd,fd_set *set)	测试该集中的一个给定位是否有变化

一般来说，在使用 select 函数之前，首先使用 FD_ZERO 和 FD_SET 来初始化文件描述符集，在使用了 select 函数时，可循环使用 FD_ISSET 测试描述符集，在执行完对相关后文件描述符后，使用 FD_CLR 来清楚描述符集。

另外，select 函数中的 timeout 是一个 struct timeval 类型的指针，该结构体如下所示：

```
struct timeval {
       long tv_sec; /* second */
       long tv_unsec; /* and microseconds*/
}
```

可以看到，这个时间结构体的精确度可以设置到微秒级，这对于大多数的应用而言已经足够了。

（3）使用实例

由于 Select 函数多用于 I/O 操作可能会阻塞的情况下，而对于可能会有阻塞 I/O 的管道、

网络编程，本书到现在为止还没有涉及。因此，本例主要表现了如何使用 select 函数，而其中的 I/O 操作是不会阻塞的。

本实例中主要实现将文件 hello1 里的内容读出，并将此内容每隔 10s 写入 hello2 中去。在这里建立了两个描述符集，其中一个描述符集 inset1 是用于读取文件内容，另一个描述符集 inset2 是用于写入文件的。两个文件描述符 fds[0]和 fds[1]分别指向这一文件描述符。在首先初始化完各文件描述符集之后，就开始了循环测试这两个文件描述符是否可读写，由于在这里没有阻塞，所以文件描述符处于准备就绪的状态。这时，就分别对文件描述符 fds[0]和 fsd[1]进行读写操作。该程序的流程图如图 6.2 所示。

```
/*select.c*/
#include <fcntl.h>
#include <stdio.h>
#include <unistd.h>
#include <stdlib.h>
#include <time.h>

int main(void)
{
     int fds[2];
     char buf[7];
     int i,rc,maxfd;
     fd_set inset1,inset2;
     struct timeval tv;
/*首先按一定的权限打开 hello1 文件*/
     if((fds[0] = open ("hello1", O_RDWR|O_CREAT,0666))<0)
           perror("open hello1");
/*再按一定的权限打开 hello2 文件*/
     if((fds[1] = open ("hello2", O_RDWR|O_CREAT,0666))<0)
           perror("open hello2");
     if((rc = write(fds[0],"Hello!\n",7)))
           printf("rc=%d\n",rc);
     lseek(fds[0],0,SEEK_SET);
/*取出两个文件描述符中的较大者*/
     maxfd = fds[0]>fds[1] ? fds[0] : fds[1];
/*初始化读集合 inset1，并在读集合中加入相应的描述集*/
     FD_ZERO(&inset1);
     FD_SET(fds[0],&inset1);
/*初始化写集合 inset2，并在写集合中加入相应的描述集*/
     FD_ZERO(&inset2);
```

```
        FD_SET(fds[1],&inset2);
        tv.tv_sec=2;
        tv.tv_usec=0;
    /*循环测试该文件描述符是否准备就绪，并调用 select 函数对相关文件描述符做对应操作*/
        while(FD_ISSET(fds[0],&inset1)||FD_ISSET(fds[1],&inset2)){
            if(select(maxfd+1,&inset1,&inset2,NULL,&tv)<0)
                perror("select");
            else{
                if(FD_ISSET(fds[0],&inset1)){
                    rc = read(fds[0],buf,7);
                    if(rc>0){
                        buf[rc]='\0';
                        printf("read: %s\n",buf);
                    }else
                        perror("read");
                }
                if(FD_ISSET(fds[1],&inset2)){
                    rc = write(fds[1],buf,7);
                    if(rc>0){
                        buf[rc]='\0';
                        printf("rc=%d,write: %s\n",rc,buf);
                    }else
                        perror("write");
                sleep(10);
                }
            }
        }
        exit(0);
    }
```

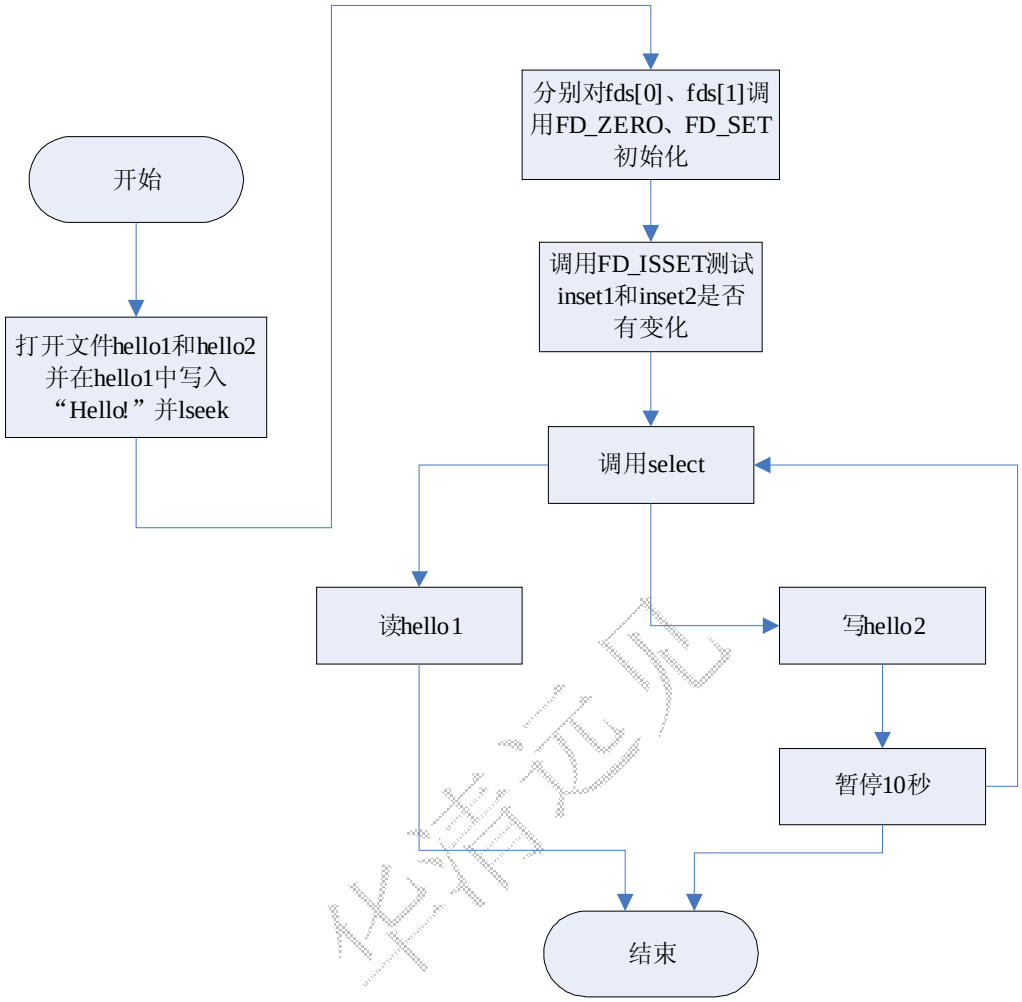


图 6.2 select 实例流程图

读者可以将以上程序交叉编译，并下载到开发板上运行。以下是运行结果：

```
[root@(none) 1]# ./select
rc=7
read: Hello!

rc=7,write: Hello!

rc=7,write:Hello!

rc=7,write:Hello!

...
```

```
[root@(none) 1]# cat hello1
Hello!
[root@(none) 1]# cat hello2
Hello!
Hello!
…
```

可以看到，使用 select 可以很好地实现 I/O 多路复用，在有阻塞的情况下更能够显示出它的作用。

6.4 嵌入式 Linux 串口应用开发

6.4.1 串口概述

用户常见的数据通信的基本方式可分为并行通信与串行通信两种。

- 并行通信是指利用多条数据传输线将一个资料的各位同时传送。它的特点是传输速度快，适用于短距离通信，但要求传输速度较高的应用场合。
- 串行通信是指利用一条传输线将资料一位位地顺序传送。特点是通信线路简单，利用简单的线缆就可实现通信，降低成本，适用于远距离通信，但传输速度慢的应用场合。

串口是计算机一种常用的接口，常用的串口有 RS-232-C 接口。它是于 1970 年由美国电子工业协会（EIA）联合贝尔系统、调制解调器厂家及计算机终端生产厂家共同制定的用于串行通讯的标准，它的全称是“数据终端设备（DTE）和数据通讯设备（DCE）之间串行二进制数据交换接口技术标准”。该标准规定采用一个 DB25 芯引脚的连接器或 9 芯引脚的连接器，其中 25 芯引脚的连接器如图 6.3 所示。

S3C2410X 内部具有 2 个独立的 UART 控制器，每个控制器都可以工作在 Interrupt（中断）模式或者 DMA（直接内存访问）模式。同时，每个 UART 均具有 16 字节的 FIFO（先入先出寄存器），支持的最高波特率可达到 230.4Kbps。UART 的操作主要可分为以下几个部分：资料发送、资料接收、产生中断、产生波特率、Loopback 模式、红外模式以及自动流控模式。

串口参数的配置读者在配置超级终端和 minicom 时也已经接触到过，一般包括波特率、起始位数量、数据位数量、停止位数量和流控协议。在此，可以将其配置为波特率 115200、起始位 1b、数据位 8b、停止位 1b 和无流控协议。

在 Linux 中，所有的设备文件一般都位于“/dev”下，其中串口一、串口二对应的设备名依次为“/dev/ttyS0”、“/dev/ttyS1”，可以查看在“/dev”下的文件以确认。在本章中已经提到过，在 Linux 下对设备的操作方法与对文件的操作方法是一样的，因此，对串口的读写就可以使用简单的“read”，“write”函数来完成，所不同的是只是需要对串口的其他参数另做配置，下面就来详细讲解串口应用开发的步骤。

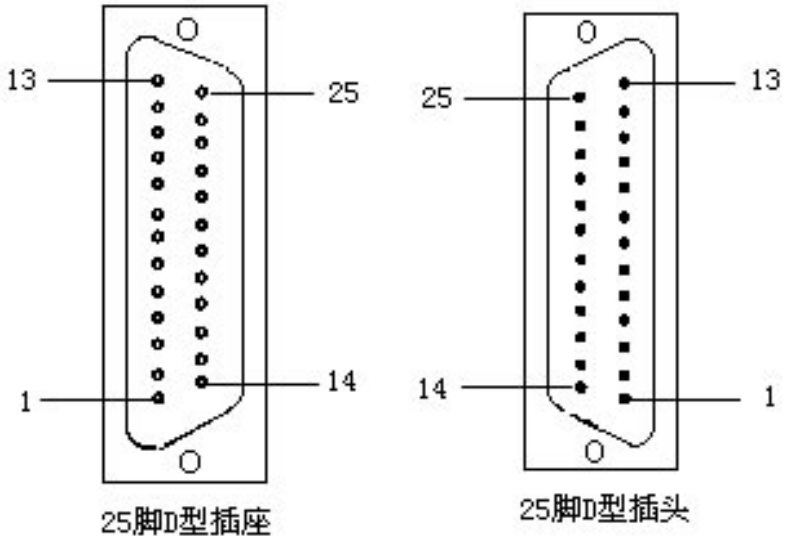


图 6.3　25 引脚串行接口图

6.4.2　串口设置详解

本节主要讲解设置串口的主要方法。

如前所述，设置串口中最基本的包括波特率设置，校验位和停止位设置。串口的设置主要是设置 struct termios 结构体的各成员值，如下所示：

```
#include<termios.h>
struct termio
{      unsigned short  c_iflag; /* 输入模式标志 */
       unsigned short  c_oflag;     /* 输出模式标志 */
       unsigned short  c_cflag;     /* 控制模式标志*/
       unsigned short  c_lflag;     /*本地模式标志 */
       unsigned char  c_line;       /* line discipline */
       unsigned char  c_cc[NCC];    /* control characters */
};
```

在这个结构中最为重要的是 c_cflag，通过对它的赋值，用户可以设置波特率、字符大小、数据位、停止位、奇偶校验位和硬件流控等。另外 c_iflag 和 c_cc 也是比较常用的标志。在此主要对这 3 个成员进行详细说明。

c_cflag 支持的常量名称如表 6.10 所示。其中设置波特率为相应的波特率前加上‘B’，由于数值较多，本表没有全部列出。

表 6.10　　c_cflag 支持的常量名称

CBAUD	波特率的位掩码
B0	0 波特率（放弃 DTR）

…	…

续表

B1800	1800 波特率
B2400	2400 波特率
B4800	4800 波特率
B9600	9600 波特率
B19200	19200 波特率
B38400	38400 波特率
B57600	57600 波特率
B115200	115200 波特率
EXTA	外部时钟率
EXTB	外部时钟率
CSIZE	数据位的位掩码
CS5	5 个数据位
CS6	6 个数据位
CS7	7 个数据位
CS8	8 个数据位
CSTOPB	2 个停止位（不设则是 1 个停止位）
CREAD	接收使能
PARENB PARODD	校验位使能 使用奇校验而不使用偶校验
HUPCL	最后关闭时挂线（放弃 DTR）
CLOCAL	本地连接（不改变端口所有者）
LOBLK	块作业控制输出
CNET_CTSRTS	硬件流控制使能

在这里，对于 c_cflag 成员不能直接对其初始化，而要将其通过“与”、“或”操作使用其中的某些选项。

输入模式 c_iflag 成员控制端口接收端的字符输入处理。c_iflag 支持的变量名称，如表 6.11 所示。

表 6.11　　c_iflag 支持的常量名称

INPCK	奇偶校验使能
IGNPAR	忽略奇偶校验错误
PARMRK	奇偶校验错误掩码

ISTRIP	除去奇偶校验位

续表

IXON	启动出口硬件流控
IXOFF	启动入口软件流控
IXANY	允许字符重新启动流控
IGNBRK	忽略中断情况
BRKINT	当发生中断时发送 SIGINT 信号
INLCR	将 NL 映射到 CR
IGNCR	忽略 CR
ICRNL	将 CR 映射到 NL
IUCLC	将高位情况映射到低位情况
IMAXBEL	当输入太长时回复 ECHO

c_cc 包含了超时参数和控制字符的定义。c_cc 所支持的常用变量名称，如表 6.12 所示。

表 6.12 c_cc 支持的常量名称

VINTR	中断控制，对应键为 CTRL+C
VQUIT	退出操作，对应键为 CRTL+Z
VERASE	删除操作，对应键为 Backspace（BS）
VKILL	删除行，对应键为 CTRL+U
VEOF	位于文件结尾，对应键为 CTRL+D
VEOL	位于行尾，对应键为 Carriage return（CR）
VEOL2	位于第二行尾，对应键为 Line feed（LF）
VMIN	指定了最少读取的字符数
VTIME	指定了读取每个字符的等待时间

下面就详细讲解设置串口属性的基本流程。

1．保存原先串口配置

首先，为了安全起见和以后调试程序方便，可以先保存原先串口的配置，在这里可以使用函数 tcgetattr（fd，&oldtio）。该函数得到与 fd 指向对象的相关参数，并将它们保存于 oldtio 引用的 termios 结构中。该函数还可以测试配置是否正确、该串口是否可用等。若调用成功，函数返回值为 0，若调用失败，函数返回值为–1，其使用如下所示：

```
if  ( tcgetattr( fd,&oldtio)  !=  0) {
      perror("SetupSerial 1");
```

```
        return -1;
    }
```

2．激活选项有 CLOCAL 和 CREAD

CLOCAL 和 CREAD 分别用于本地连接和接受使能，因此，首先要通过位掩码的方式激活这两个选项。

```
    newtio.c_cflag  |=  CLOCAL | CREAD;
```

3．设置波特率

设置波特率有专门的函数，用户不能直接通过位掩码来操作。设置波特率的主要函数有：cfsetispeed 和 cfsetospeed。这两个函数的使用很简单，如下所示：

```
    cfsetispeed(&newtio, B115200);
    cfsetospeed(&newtio, B115200);
```

一般地，用户需将输入输出函数的波特率设置成一样的。这几个函数在成功时返回 0，失败时返回–1。

4．设置字符大小

与设置波特率不同，设置字符大小并没有现成可用的函数，需要用位掩码。一般首先去除数据位中的位掩码，再重新按要求设置。如下所示：

```
    options.c_cflag &= ~CSIZE;/* mask the character size bits */
    options.c_cflag |= CS8;
```

5．设置奇偶校验位

设置奇偶校验位需要用到两个 termio 中的成员：c_cflag 和 c_iflag。首先要激活 c_cflag 中的校验位使能标志 PARENB 和是否要进行偶校验，同时还要激活 c_iflag 中的奇偶校验使能。如使能奇校验时，代码如下所示：

```
    newtio.c_cflag |= PARENB;
    newtio.c_cflag |= PARODD;
    newtio.c_iflag |= (INPCK | ISTRIP);
```

而使能偶校验时代码为：

```
    newtio.c_iflag |= (INPCK | ISTRIP);
    newtio.c_cflag |= PARENB;
    newtio.c_cflag &= ~PARODD;
```

6．设置停止位

设置停止位是通过激活 c_cflag 中的 CSTOPB 而实现的。若停止位为 1，则清除 CSTOPB，若停止位为 0，则激活 CSTOPB。下面是停止位是 1 时的代码：

```
newtio.c_cflag &=  ~CSTOPB;
```

7．设置最少字符和等待时间

在对接收字符和等待时间没有特别要求的情况下，可以将其设置为 0，如下所示：

```
newtio.c_cc[VTIME]  = 0;
newtio.c_cc[VMIN] = 0;
```

8．处理要写入的引用对象

由于串口在重新设置之后，在此之前要写入的引用对象要重新处理，这时就可调用函数 tcflush（fd，queue_selector）来处理要写入引用的对象。对于尚未传输的数据，或者收到的但是尚未读取的数据，其处理方法取决于 queue_selector 的值。

这里，queue_selector 可能的取值有以下几种。

- TCIFLUSH：刷新收到的数据但是不读。
- TCOFLUSH：刷新写入的数据但是不传送。
- TCIOFLUSH：同时刷新收到的数据但是不读，并且刷新写入的数据但是不传送。

如在本例中所采用的是第一种方法：

```
tcflush(fd,TCIFLUSH);
```

9．激活配置

在完成全部串口配置之后，要激活刚才的配置并使配置生效。这里用到的函数是 tcsetattr，它的函数原型是：

tcsetattr(fd,OPTION,&newtio);

这里的 newtio 就是 termios 类型的变量，OPTION 可能的取值有以下三种：

- TCSANOW：改变的配置立即生效。
- TCSADRAIN：改变的配置在所有写入 fd 的输出都结束后生效。
- TCSAFLUSH：改变的配置在所有写入 fd 引用对象的输出都被结束后生效，所有已接受但未读入的输入都在改变发生前丢弃。

该函数若调用成功则返回 0，若失败则返回–1。

如下所示：

```
if((tcsetattr(fd,TCSANOW,&newtio))!=0)
{
      perror("com set error");
```

```
        return -1;
    }
```

下面给出了串口配置的完整的函数。通常，为了函数的通用性，通常将常用的选项都在函数中列出，这样可以大大方便以后用户的调试使用。该设置函数如下所示：

```
int set_opt(int fd,int nSpeed, int nBits, char nEvent, int nStop)
{
    struct termios newtio,oldtio;
/*保存测试现有串口参数设置，在这里如果串口号等出错，会有相关的出错信息*/
    if  ( tcgetattr( fd,&oldtio)  !=  0) {
        perror("SetupSerial 1");
        return -1;
    }
    bzero( &newtio, sizeof( newtio ) );
/*步骤一，设置字符大小*/
    newtio.c_cflag  |=  CLOCAL | CREAD;
    newtio.c_cflag &= ~CSIZE;
/*设置停止位*/
    switch( nBits )
    {
    case 7:
        newtio.c_cflag |= CS7;
        break;
    case 8:
        newtio.c_cflag |= CS8;
        break;
    }
/*设置奇偶校验位*/
    switch( nEvent )
    {
    case 'O': //奇数
        newtio.c_cflag |= PARENB;
        newtio.c_cflag |= PARODD;
        newtio.c_iflag |= (INPCK | ISTRIP);
        break;
    case 'E': //偶数
        newtio.c_iflag |= (INPCK | ISTRIP);
        newtio.c_cflag |= PARENB;
        newtio.c_cflag &= ~PARODD;
```

```
                break;
        case 'N':  //无奇偶校验位
                newtio.c_cflag &= ~PARENB;
                break;
        }
/*设置波特率*/
switch( nSpeed )
        {
        case 2400:
                cfsetispeed(&newtio, B2400);
                cfsetospeed(&newtio, B2400);
                break;
        case 4800:
                cfsetispeed(&newtio, B4800);
                cfsetospeed(&newtio, B4800);
                break;
        case 9600:
                cfsetispeed(&newtio, B9600);
                cfsetospeed(&newtio, B9600);
                break;
        case 115200:
                cfsetispeed(&newtio, B115200);
                cfsetospeed(&newtio, B115200);
                break;
        case 460800:
                cfsetispeed(&newtio, B460800);
                cfsetospeed(&newtio, B460800);
                break;
        default:
                cfsetispeed(&newtio, B9600);
                cfsetospeed(&newtio, B9600);
               break;
        }
/*设置停止位*/
        if( nStop == 1 )
                newtio.c_cflag &=  ~CSTOPB;
        else if ( nStop == 2 )
                newtio.c_cflag |=  CSTOPB;
/*设置等待时间和最小接收字符*/
```

```
        newtio.c_cc[VTIME]  = 0;
        newtio.c_cc[VMIN] = 0;
    /*处理未接收字符*/
        tcflush(fd,TCIFLUSH);
    /*激活新配置*/
        if((tcsetattr(fd,TCSANOW,&newtio))!=0)
        {
               perror("com set error");
               return -1;
        }
        printf("set done!\n");
        return 0;
    }
```

6.4.3 串口使用详解

在配置完串口的相关属性后，就可以对串口进行打开、读写操作了。它所使用的函数和普通文件读写的函数一样，都是 open、write 和 read。它们相区别的只是串口是一个终端设备，因此在函数的具体参数的选择时会有一些区别。另外，这里会用到一些附加的函数，用于测试终端设备的连接情况等。下面将对其进行具体讲解。

1．打开串口

打开串口和打开普通文件一样，使用的函数同打开普通文件一样，都是 open 函数。如下所示：

```
    fd = open( "/dev/ttyS0", O_RDWR|O_NOCTTY|O_NDELAY);
```

可以看到，这里除了普通的读写参数外，还有两个参数 O_NOCTTY 和 O_NDELAY。

- O_NOCTTY 标志用于通知 Linux 系统，这个程序不会成为对应这个端口的控制终端。如果没有指定这个标志，那么任何一个输入（诸如键盘中止信号等）都将会影响用户的进程。
- O_NDELAY 标志通知 Linux 系统，这个程序不关心 DCD 信号线所处的状态（端口的另一端是否激活或者停止）。如果用户指定了这个标志，则进程将会一直处在睡眠状态，直到 DCD 信号线被激活。

接下来可恢复串口的状态为阻塞状态，用于等待串口数据的读入。可用 fcntl 函数实现，如下所示：

```
    fcntl(fd, F_SETFL, 0);
```

再接着可以测试打开文件描述符是否引用一个终端设备，以进一步确认串口是否正确打开，如下所示：

```
    isatty(STDIN_FILENO);
```

该函数调用成功则返回 0，若失败则返回-1。

这时，一个串口就已经成功打开了。接下来就可以对这个串口进行读、写操作。

下面给出了一个完整的打开串口的函数，同样写考虑到了各种不同的情况。程序如下所示：

```
/*打开串口函数*/
int open_port(int fd,int comport)
{
	char *dev[]={"/dev/ttyS0","/dev/ttyS1","/dev/ttyS2"};
	long  vdisable;
	if (comport==1)//串口 1
	{    fd = open( "/dev/ttyS0", O_RDWR|O_NOCTTY|O_NDELAY);
		if (-1 == fd){
			perror("Can't Open Serial Port");
			return(-1);
		}
	}
	else if(comport==2)//串口 2
	{    fd = open( "/dev/ttyS1", O_RDWR|O_NOCTTY|O_NDELAY);
		if (-1 == fd){
			perror("Can't Open Serial Port");
			return(-1);
		}
	}
	else if (comport==3)//串口 3
	{
		fd = open( "/dev/ttyS2", O_RDWR|O_NOCTTY|O_NDELAY);
		if (-1 == fd){
			perror("Can't Open Serial Port");
			return(-1);
		}
	}
/*恢复串口为阻塞状态*/
	if(fcntl(fd, F_SETFL, 0)<0)
		printf("fcntl failed!\n");
	else
		printf("fcntl=%d\n",fcntl(fd, F_SETFL,0));
/*测试是否为终端设备*/
	if(isatty(STDIN_FILENO)==0)
```

```
            printf("standard input is not a terminal device\n");
        else
            printf("isatty success!\n");
        printf("fd-open=%d\n",fd);
        return fd;
}
```

2．读写串口

读写串口操作和读写普通文件一样，使用 read、write 函数即可。如下所示：

```
write(fd,buff,8);
read(fd,buff,8);
```

下面两个实例给出了串口读和写的两个程序的 main 函数部分，这里用到的函数有前面讲述到的 open_port 和 set_opt 函数。

```
/*写串口程序*/
#include <stdio.h>
#include <string.h>
#include <sys/types.h>
#include <errno.h>
#include <sys/stat.h>
#include <fcntl.h>
#include <unistd.h>
#include <termios.h>
#include <stdlib.h>

/*读串口程序*/
int main(void)
{
    int fd;
    int nread,i;
    char buff[]="Hello\n";

    if((fd=open_port(fd,1))<0){//打开串口
        perror("open_port error");
        return;
    }
    if((i=set_opt(fd,115200,8,'N',1))<0){//设置串口
        perror("set_opt error");
        return;
```

```
        }
        printf("fd=%d\n",fd);
        fd=3;
        nread=read(fd,buff,8);//读串口
        printf("nread=%d,%s\n",nread,buff);
        close(fd);
        return;
    }
```

读者可以将该程序在宿主机上运行，然后用串口线将目标板和宿主机连接起来，之后将目标板上电，这样就可以看到宿主机上有目标板的串口输出。

```
 [root@localhost file]# ./receive
fcntl=0
isatty success!
fd-open=3
set done
fd=3
nread=8,…
```

另外，读者还可以考虑一下如何使用 select 函数实现串口的非阻塞读写，具体实例会在后面的实验中给出。

6.5 标准 I/O 开发

本章前面几节所述的文件及 I/O 读写都是基于文件描述符的。这些都是基本的 I/O 控制，是不带缓存的。而本节所要讨论的 I/O 操作都是基于流缓冲的，它是符合 ANSI C 的标准 I/O 处理，这里有很多函数读者已经非常熟悉了（如 printf、scantf 函数等），因此本节中仅简要介绍最主要的函数。

标准 I/O 提供流缓冲的目的是尽可能减少使用 read 和 write 调用的数量。标准 I/O 提供了 3 种类型的缓冲存储。

- 全缓冲。在这种情况下，当填满标准 I/O 缓存后才进行实际 I/O 操作。对于驻在磁盘上的文件通常是由标准 I/O 库实施全缓冲的。在一个流上执行第一次 I/O 操作时，通常调用 malloc 就是使用全缓冲。
- 行缓冲。在这种情况下，当在输入和输出中遇到新行符时，标准 I/O 库执行 I/O 操作。这允许我们一次输出一个字符（如 fputc 函数），但只有写了一行之后才进行实际 I/O 操作。当流涉及一个终端时（例如标准输入和标准输出），典型地使用行缓冲。
- 不带缓冲。标准 I/O 库不对字符进行缓冲。如果用标准 I/O 函数写若干字符到不带缓冲的流中，则相当于用 write 系统的用函数将这些字符写全相比较的打开文件上。标准出错况 stderr 通常是不带缓存后，这就使得出错信息可以尽快显示出来，而不管它们是否含有一个新行字符。

在下面讨论具体函数时，请读者注意区分这 3 种不同的情况。

6.5.1 打开和关闭文件

1．打开文件

（1）函数说明

打开文件有三个标准函数，分别为：fopen、fdopen 和 freopen。它们可以以不同的模式打开，但都返回一个指向 FILE 的指针，该指针以将对应的 I/O 流相绑定了。此后，对文件的读写都是通过这个 FILE 指针来进行。其中 fopen 可以指定打开文件的路径和模式，fdopen 可以指定打开的文件描述符和模式，而 freopen 除可指定打开的文件、模式外，还可指定特定的 IO 流。

（2）函数格式定义

fopen 函数格式如表 6.13 所示。

表 6.13 fopen 函数语法要点

所需头文件	#include <stdio.h>
函数原型	FILE * fopen(const char * path,const char * mode)
函数传入值	path：包含要打开的文件路径及文件名
	mode：文件打开状态（后面会具体说明）
函数返回值	成功：指向 FILE 的指针 失败：NULL

这里的 mode 类似于 open 中的 flag，可以定义打开文件的具体权限等，表 6.14 说明了 fopen 中 mode 的各种取值。

表 6.14 mode 取值说明

r 或 rb	打开只读文件，该文件必须存在
r＋或 r＋b	打开可读写的文件，该文件必须存在
w 或 wb	打开只写文件，若文件存在则文件长度清为 0，即会擦些文件以前内容。若文件不存在则建立该文件
w+或 w＋b	打开可读写文件，若文件存在则文件长度清为 0，即会擦些文件以前内容。若文件不存在则建立该文件
a 或 ab	以附加的方式打开只写文件。若文件不存在，则会建立该文件；如果文件存在，写入的数据会被加到文件尾，即文件原先的内容会被保留
a+或 a＋b	以附加方式打开可读写的文件。若文件不存在，则会建立该文件；如果文件存在，写入的数据会被加到文件尾后，即文件原先的内容会被保留

注意在每个选项中加入 b 字符用来告诉函数库打开的文件为二进制文件，而非纯文字文件。不过在 Linux 系统中会自动识别不同类型的文件而将此符号忽略。

fdopen 函数格式如表 6.15 所示。

表 6.15 fdopen 函数语法要点

所需头文件	#include <stdio.h>
函数原型	FILE * fdopen(int fd,const char * mode)
函数传入值	fd：要打开的文件描述符
	mode：文件打开状态（后面会具体说明）
函数返回值	成功：指向 FILE 的指针 失败：NULL

freopen 函数格式如表 6.16 所示。

表 6.16 freopen 函数语法要点

所需头文件	#include <stdio.h>
函数原型	FILE * freopen(const char *path,const char * mode,FILE * stream)
函数传入值	path：包含要打开的文件路径及文件名
	mode：文件打开状态（后面会具体说明）
	stream：已打开的文件指针
函数返回值	成功：指向 FILE 的指针 失败：NULL

2．关闭文件

（1）函数说明

关闭标准流文件的函数为 fclose，这时缓冲区内的数据写入文件中，并释放系统所提供的文件资源。

（2）函数格式说明

freopen 函数格式如表 6.17 所示。

表 6.17 fclose 函数语法要点

所需头文件	#include <stdio.h>
函数原型	int fclose(FILE * stream)
函数传入值	stream：已打开的文件指针
函数返回值	成功：0 失败：EOF

3．使用实例

文件打开关闭的操作都比较简单，这里仅以 fopen 和 fclose 为例，代码如下所示：

```
/*fopen.c*/
#include <stdio.h>
main()
{
    FILE *fp;
```

```
    int c;
/*调用 fopen 函数*/
    if((fp=fopen("exist","w"))!=NULL){
        printf("open success!");
    }
    fclose(fp);
}
```

读者可以尝试用其他文件打开函数进行练习。

6.5.2 文件读写

1．读文件

（1）fread 函数说明

在文件流打开之后，可对文件流进行读写等操作，其中读操作的函数为 fread。

（2）fread 函数格式

fread 函数格式如表 6.18 所示。

表 6.18 fread 函数语法要点

所需头文件	#include <stdio.h>
函数原型	size_t fread(void * ptr,size_t size,size_t nmemb,FILE * stream)
函数传入值	ptr：存放读入记录的缓冲区
	size：读取的记录大小
	nmemb：读取的记录数
	stream：要读取的文件流
函数返回值	成功：返回实际读取到的 nmemb 数目 失败：EOF

2．写文件

（1）fwrite 函数说明

fwrite 函数是用于对指定的文件流进行写操作。

（2）fwrite 函数格式

fwrite 函数格式如表 6.19 所示。

表 6.19 fwrite 函数语法要点

所需头文件	#include <stdio.h>
函数原型	size_t fwrite(const void * ptr,size_t size,size_t nmemb,FILE * stream)
函数传入值	ptr：存放写入记录的缓冲区

	size：写入的记录大小
	nmemb：写入的记录数
	stream：要写入的文件流
函数返回值	成功：返回实际写入到的 nmemb 数目 失败：EOF

这里仅以 fwrite 为例简单说明：

```
/*fwrite.c*/
#include <stdio.h>
int main()
{
     FILE *stream;
     char s[3]={'a','b','c'};
/*首先使用 fopen 打开文件，之后再调用 fwrite 写入文件*/
     stream=fopen("what","w");
     i=fwrite(s,sizeof(char),nmemb,stream);
     printf("i=%d",i);
     fclose(stream);
}
```

运行结果如下所示：

```
[root@localhost file]# ./write
i=3
[root@localhost file]# cat what
abc
```

6.5.3 输入输出

文件打开之后，根据一次读写文件中字符的数目可分为字符输入输出、行输入输出和格式化输入输出，下面分别对这 3 种不同的方式进行讨论。

1．字符输入输出

字符输入输出函数一次仅读写一个字符。其中字符输入输出函数如表 6.20 和表 6.21 所示。

表 6.20 字符输出函数语法要点

所需头文件	#include <stdio.h>
函数原型	int getc(FILE * stream) int fgetc(FILE * stream) int getchar(void)
函数传入值	stream：要输入的文件流

函数返回值	成功：下一个字符 失败：EOF

表 6.21　　字符输入函数语法要点

所需头文件	#include <stdio.h>
函数原型	int putc(int c,FILE * stream) int fputc(int c,FILE * stream) int putchar(int c)
函数返回值	成功：字符 c 失败：EOF

这几个函数功能类似，其区别仅在于 getc 和 putc 通常被实现为宏，而 fgetc 和 fputc 不能实现为宏，因此，函数的实现时间会有所差别。

下面这个实例结合 fputc 和 fgetc，将标准输入复制到标准输出中去。

```
/*fput.c*/
#include<stdio.h>
main()
{
    int c;
/*把 fgetc 的结果作为 fputc 的输入*/
    fputc(fgetc(stdin),stdout);
}
```

运行结果如下所示：

```
[root@localhost file]# ./file
w（用户输入）
w（屏幕输出）
```

2．行输入输出

行输入输出函数一次操作一行。其中行输入输出函数如表 6.22 和表 6.23 所示。

表 6.22　　行输出函数语法要点

所需头文件	#include <stdio.h>
函数原型	char * gets(char *s) char fgets(char * s,int size,FILE * stream)
函数传入值	s：要输入的字符串 size：输入的字符串长度 stream：对应的文件流
函数返回值	成功：s

	失败：NULL

表 6.23 行输入函数语法要点

所需头文件	#include <stdio.h>
函数原型	int puts(const char *s) int fputs(const char * s,FILE * stream)
函数传入值	s：要输出的字符串 stream：对应的文件流
函数返回值	成功：s 失败：NULL

这里以 gets 和 puts 为例进行说明，本实例将标准输入复制到标准输出，如下所示：

```
/*gets.c*/
#include<stdio.h>
main()
{
    char s[80];
/*同上例，把 fgets 的结果作为 fputs 的输入*/
    fputs(fgets(s,80,stdin),stdout);
}
```

运行该程序，结果如下所示：

```
[root@www yul]# ./file2
This is stdin（用户输入）
This is stdin（屏幕输出）
```

3．格式化输入输出

格式化输入输出函数可以指定输入输出的具体格式，这里有读者已经非常熟悉的 printf、scanf 等函数，这里就简要介绍一下它们的格式。如下表 6.24～表 6.26 所示。

表 6.24 格式化输出函数 1

所需头文件	#include <stdio.h>
函数原型	int printf(const char *format,…) int fprintf(FILE *fp,const char *format,…) int sprintf(char *buf,const char *format,…)
函数传入值	format：记录输出格式 fp：文件描述符 buf：记录输出缓冲区

函数返回值	成功：输出字符数（sprintf 返回存入数组中的字符数） 失败：NULL

表 6.25　　格式化输出函数 2

所需头文件	#include <stdarg.h> #include <stdio.h>
函数原型	int vprintf(const char *format,va_list arg) int vfprintf(FILE *fp,const char *format,va_list arg) int vsprintf(char *buf,const char *format,va_list arg)
函数传入值	format：记录输出格式 fp：文件描述符 arg：相关命令参数
函数返回值	成功：存入数组的字符数 失败：NULL

表 6.26　格式化输入函数

所需头文件	#include <stdio.h>
函数原型	int scanf(const char *format,…) int fscanf(FILE *fp,const char *format,…) int sscanf(char *buf,const char *format,…)
函数传入值	format：记录输出格式 fp：文件描述符 buf：记录输入缓冲区
函数返回值	成功：输出字符数（sprintf 返回存入数组中的字符数） 失败：NULL

由于本节的函数用法比较简单，并且比较常用，因此就不再举例了，请读者需要用到时自行查找其用法。

6.6　实验内容

6.6.1　文件读写及上锁

1．实验目的

通过编写文件读写及上锁的程序，进一步熟悉 Linux 中文件 I/O 相关的应用开发，并且熟练掌握 open、read、write、fcntl 等函数的使用。

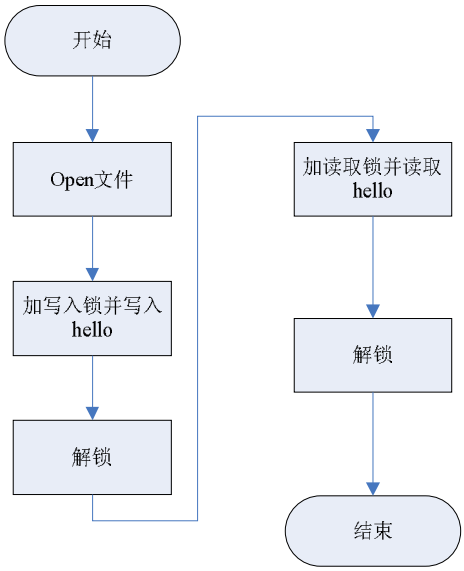
2．实验内容

该实验要求首先打开一个文件，然后将该文件上写入锁，并写入 hello 字符串。接着在解锁后再将该文件上读取锁，并读取刚才写入的内容。最后模拟多进程，同时读写一个文件时的情况。

3．实验步骤

（1）画出实验流程图

该实验流程图如图 6.4 所示。



6.4　实验 6.6.1 节流程图

（2）编写代码

该实验源代码如下所示，其中用到的 lock_set 函数可参见第 6.3.3 节。

```
/*expr1.c 实验一源码*/
#include <unistd.h>
#include <sys/file.h>
#include <sys/types.h>
#include <sys/stat.h>
#include <stdio.h>
#include <stdlib.h>
#include <string.h>
void lock_set(int fd,int type);

int main(void)
{
     int fd,nwrite,nread,len;
     char *buff="Hello\n";
     char buf_r[100];
     len=strlen(buff);
     fd=open("hello",O_RDWR | O_CREAT, 0666);
     if(fd < 0){
          perror("open");
          exit(1);
     }
/*加上写入锁*/
     lock_set(fd, F_WRLCK);
     if((nwrite=write(fd,buff,len))==len){
          printf("write success\n");
     }
     getchar();
/*解锁*/
     lock_set(fd, F_UNLCK);
     getchar();
/*加上读取锁*/
```

```
        lock_set(fd, F_RDLCK);
        lseek(fd,0,SEEK_SET);
        if((nread=read(fd,buf_r,len))==len){
              printf("read:%s\n",buf_r);
        }
        getchar();
    /*解锁*/
        lock_set(fd, F_UNLCK);
        getchar();
        close(fd);
        exit(0);
    }
```

（3）首先在宿主机上编译调试该程序，如下所示：

```
[root@localhost process]# gcc expr1.c –o expr1
```

（4）在确保没有编译错误后，使用交叉编译该程序，如下所示：

```
[root@localhost process]# arm-linux-gcc expr1.c –o expr2
```

（5）将生成的可执行程序下载到目标板上运行。

4．实验结果

此实验在目标板上的运行结果如下所示：

```
[root@(none) 1]# ./expr1
write lock set by 75
write success

release lock by 75

read lock set by 75
read:Hello


release lock by 75
```

另外，在本机上可以开启两个终端，同时运行该程序。实验结果会和这两个进程运行过程具体相关，希望读者能具体分析每种情况。下面列出其中一种情况：

终端一：

```
[root@localhost file]# ./expr1
write lock set by 3013
```

```
write success

release lock by 3013

read lock set by 3013
read:Hello


release lock by 3013
```

终端二：

```
[root@localhost file]# ./expr1
write lock already set by 3013

write lock set by 3014
write success

release lock by 3014

read lock set by 3014
read:Hello



release lock by 3014
```

6.6.2 多路复用式串口读写

1．实验目的

通过编写多路复用式串口读写，进一步理解 select 函数的作用，同时更加熟练掌握 Linux 设备文件的读写方法。

2．实验内容

完成串口读写操作，这里设定从串口读取消息时使用 select 函数，发送消息的程序不需要用 select 函数，只发送“Hello”消息由接收端接收。

3．实验步骤

（1）画出流程图

下图 6.5 是读串口的流程图，写串口的流程图与此类似。

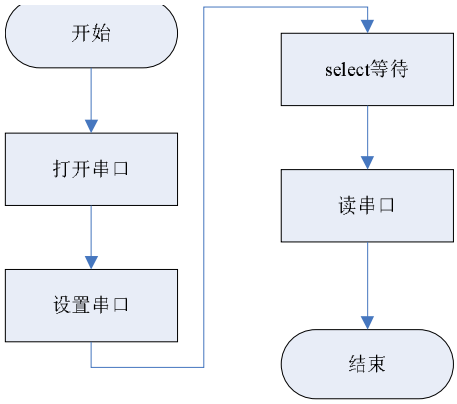


图 6.5　实验 6.6.2 节流程图

（2）编写代码

分别编写串口读写程序，该程序中用到的 open_port 和 set_opt 函数请参照 6.4 节所述。

写串口程序的代码如下所示：

```
/*写串口*/
#include <stdio.h>
#include <string.h>
#include <sys/types.h>
#include <errno.h>
#include <sys/stat.h>
#include <fcntl.h>
#include <unistd.h>
#include <termios.h>
#include <stdlib.h>
int main(void)
{
     int fd;
     int nwrite,i;
     char buff[]="Hello\n";
/*打开串口*/
     if((fd=open_port(fd,1))<0){
           perror("open_port error");
           return;
     }
/*设置串口*/
     if((i=set_opt(fd,115200,8,'N',1))<0){
           perror("set_opt error");
           return;
     }
     printf("fd=%d\n",fd);
/*向串口写入字符串*/
     nwrite=write(fd,buff,8);
     printf("nwrite=%d\n",nwrite);
     close(fd);
     return;
}
```

读串口程序的代码如下所示：

```
/*读串口*/
#include <stdio.h>
#include <string.h>
#include <sys/types.h>
#include <errno.h>
#include <sys/stat.h>
#include <fcntl.h>
#include <unistd.h>
#include <termios.h>
#include <stdlib.h>
int main(void)
{
     int fd;
     int nread,nwrite,i;
     char buff[8];
     fd_set rd;
/*打开串口*/
     if((fd=open_port(fd,1))<0){
           perror("open_port error");
           return;
     }
/*设置串口*/
     if((i=set_opt(fd,115200,8,'N',1))<0){
           perror("set_opt error");
           return;
     }
/*利用 select 函数来实现多个串口的读写*/
     FD_ZERO(&rd);
     FD_SET(fd,&rd);
     while(FD_ISSET(fd,&rd)){
     if(select(fd+1,&rd,NULL,NULL,NULL)<0)
                perror("select");
     else{
           while((nread = read(fd, buff, 8))>0)
           {
                printf("nread=%d,%s\n",nread,buff);
           }
     }
     close(fd);
```

```
        return;
    }
```

（3）接下来把第一个写串口程序交叉编译，再把第二个读串口程序在 PC 机上编译，分别得到可执行文件 write 和 read。

（4）将写串口程序下载到开发板上，然后连接 PC 和开发板的串口 1。首先运行读串口程序，再运行写串口程序。

4．实验结果

发送端的运行结果如下所示：

```
[root@(none) 1]# ./write
fcntl=0
isatty success!
fd-open=3
set done
fd=3
nwrite=8
```

接收端的运行结果如下所示：

```
[root@localhost file]# ./read
fcntl=0
isatty success!
fd-open=3
set done
fd=3
nread=8,Hello!
```

读者还可以尝试修改 select 函数选项，例如设定一个等待时间，查看程序的运行结果。

本 章 小 结

本章首先讲解了系统调用、用户函数接口（API）和系统命令之间的关系和区别，这也是贯穿本书的一条主线，本书的讲解就是从系统命令、用户函数接口（API）到系统调用为顺序一层层深入进行讲解的，希望读者能有一个较为深刻的认识。

接着，本章主要讲解了嵌入式 Linux 中文件 I/O 相关的开发，在这里主要讲解了不带缓存 I/O 函数的使用，这也是本章的重点。因为不带缓存 I/O 函数的使用范围非常广泛，在有很多情况下必须使用它，这也是学习嵌入式 Linux 开发的基础，因此读者一定要牢牢掌握相关知识。其中主要讲解了 open、close、read、write、lseek、fcntl 和 select 等函数，这几个函数包括了不带缓存 I/O 处理的主要部分，并且也体现了它的主要思想。

接下来，本章讲解了嵌入式 Linux 串口编程。这其实是 Linux 中设备文件读写的实例，由于它能很好地体现前面所介绍的内容，而且在嵌入式开发中也较为常见，因此对它进行了比较详细地讲解。

之后，本章简单介绍了标准 I/O 的相关函数，希望读者也能对它有一个总体的认识。

最后，本章安排了两个实验，分别是文件使用及上锁和多用复用串口读写。希望读者能够认真完成。

思考与练习

使用 select 函数实现 3 个串口的通信：串口 1 接收数据，串口 2 和串口 3 向串口 1 发送数据。

“黑色经典”系列之《嵌入式 Linux 应用程序开发详解》

第 7 章 进程控制开发

本章目标

文件是 Linux 中最常见最基础的操作对象，而进程则是系统调度的单位，在上一章学习了文件 I/O 控制之后，本章主要讲解进程控制开发部分，通过本章的学习，读者将会掌握以下内容。

- 掌握进程相关的基本概念
- 掌握 Linux 下的进程结构
- 掌握 Linux 下进程创建及进程管理
- 掌握 Linux 下进程创建相关的系统调用
- 掌握守护进程的概念
- 掌握守护进程的启动方法
- 掌握守护进程的输出及建立方法
- 学会编写多进程程序
- 学会编写守护进程

7.1 Linux 下进程概述

7.1.1 进程相关基本概念

1．进程的定义

进程的概念首先是在 60 年代初期由 MIT 的 Multics 系统和 IBM 的 TSS/360 系统引入的。经过了 40 多年的发展，人们对进程有过各种各样的定义。现列举较为著名的几种。

（1）进程是一个独立的可调度的活动（E. Cohen，D. Jofferson）

（2）进程是一个抽象实体，当它执行某个任务时，将要分配和释放各种资源（P. Denning）

（3）进程是可以并行执行的计算部分。（S. E. Madnick，J. T. Donovan）

以上进程的概念都不相同，但其本质是一样的。它指出了进程是一个程序的一次执行的过程。它和程序是有本质区别的，程序是静态的，它是一些保存在磁盘上的指令的有序集合，没有任何执行的概念；而进程是一个动态的概念，它是程序执行的过程，包括了动态创建、调度和消亡的整个过程。它是程序执行和资源管理的最小单位。因此，对系统而言，当用户在系统中键入命令执行一个程序的时候，它将启动一个进程。

2．进程控制块

进程是 Linux 系统的基本调度单位，那么从系统的角度看如何描述并表示它的变化呢？在这里，是通过进程控制块来描述的。进程控制块包含了进程的描述信息、控制信息以及资源信息，它是进程的一个静态描述。在 Linux 中，进程控制块中的每一项都是一个 task_struct 结构，它是在 include/linux/sched.h 中定义的。

3．进程的标识

在 Linux 中最主要的进程标识有进程号（PID，Process Idenity Number）和它的父进程号（PPID，parent process ID）。其中 PID 惟一地标识一个进程。PID 和 PPID 都是非零的正整数。

在 Linux 中获得当前进程的 PID 和 PPID 的系统调用函数为 getpid 和 getppid，通常程序获得当前进程的 PID 和 PPID 可以将其写入日志文件以做备份。getpid 和 getppid 系统调用过程如下所示：

```
/*process.c*/
#include<stdio.h>
#include<unistd.h>
#include <stdlib.h>
int main()
{
/*获得当前进程的进程 ID 和其父进程 ID*/
    printf("The PID of this process is %d\n",getpid());
    printf("The PPID of this process is %d\n",getppid());
```

```
}
```

使用 arm-linux-gcc 进行交叉编译，再将其下载到目标板上运行该程序，可以得到如下结果，该值在不同的系统上会有所不同：

```
[root@localhost process]# ./process
The PID of this process is 78
THe PPID of this process is 36
```

另外，进程标识还有用户和用户组标识、进程时间、资源利用情况等，这里就不做一一介绍，感兴趣的读者可以参见 W.Richard Stevens 编著的《Advanced Programming in the UNIX Environmen》。

4．进程运行的状态

进程是程序的执行过程，根据它的生命期可以划分成 3 种状态。

- 执行态：该进程正在，即进程正在占用 CPU。
- 就绪态：进程已经具备执行的一切条件，正在等待分配 CPU 的处理时间片。
- 等待态：进程不能使用 CPU，若等待事件发生则可将其唤醒。

它们之间转换的关系图如图 7.1 所示。

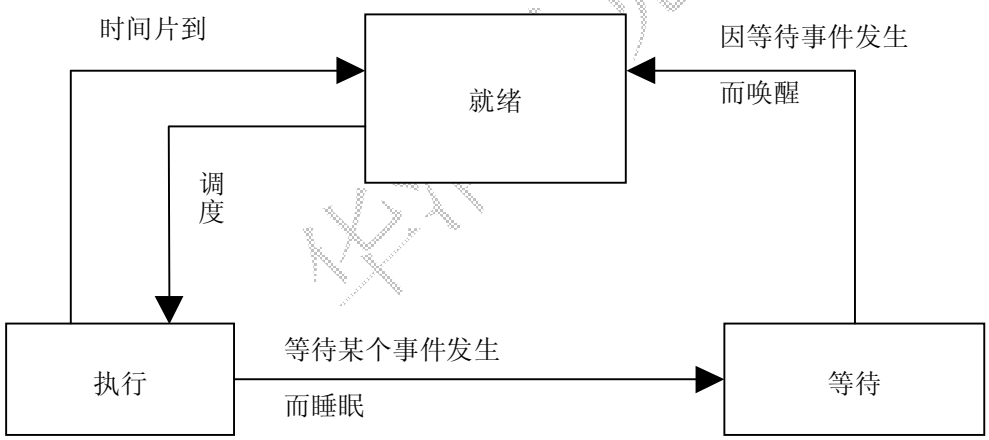


图 7.1　进程 3 种状态的转化关系

7.1.2　Linux 下的进程结构

Linux 系统是一个多进程的系统，它的进程之间具有并行性、互不干扰等特点。也就是说，进程之间是分离的任务，拥有各自的权利和责任。其中，每一个进程都运行在各自独立的虚拟地址空间，因此，即使一个进程发生异常，它也不会影响到系统中的其他进程。

Linux 中的进程包含 3 个段，分别为“数据段”、“代码段”和“堆栈段”。

- “数据段”存放的是全局变量、常数以及动态数据分配的数据空间（如 malloc 函数取得的空间）等。
- “代码段”存放的是程序代码的数据。
- “堆栈段”存放的是子程序的返回地址、子程序的参数以及程序的局部变量。如图

代码段	数据段	堆栈段

7.2 所示。

图 7.2　Linux 中进程结构示意图

7.1.3　Linux 下进程的模式和类型

在 Linux 系统中，进程的执行模式划分为用户模式和内核模式。如果当前运行的是用户程序、应用程序或者内核之外的系统程序，那么对应进程就在用户模式下运行；如果在用户程序执行过程中出现系统调用或者发生中断事件，那么就要运行操作系统（即核心）程序，进程模式就变成内核模式。在内核模式下运行的进程可以执行机器的特权指令，而且此时该进程的运行不受用户的干扰，即使是 root 用户也不能干扰内核模式下进程的运行。

用户进程既可以在用户模式下运行，也可以在内核模式下运行，如图 7.3 所示。

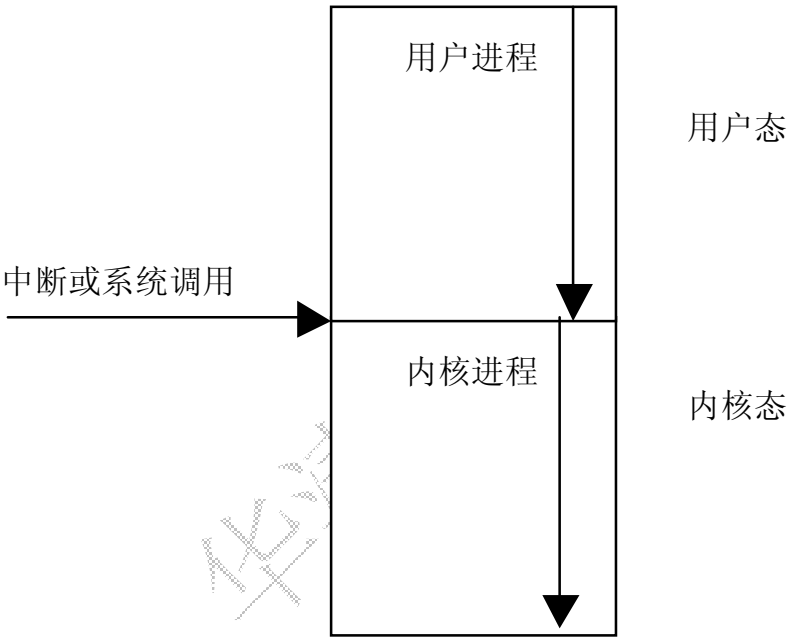


图 7.3　用户进程的两种运行模式

7.1.4　Linux 下的进程管理

Linux 下的进程管理包括启动进程和调度进程，下面就分别对这两方面进行简要讲解。

1．启动进程

Linux 下启动一个进程有两种主要途径：手工启动和调度启动。手工启动是由用户输入命令直接启动进程，而调度启动是指系统根据用户的设置自行启动进程。

（1）手工启动

手工启动进程又可分为前台启动和后台启动。

- 前台启动是手工启动一个进程的最常用方式。一般地，当用户键入一个命令如“ls -l”时，就已经启动了一个进程，并且是一个前台的进程。

- 后台启动往往是在该进程非常耗时，且用户也不急着需要结果的时候启动的。比如用户要启动一个需要长时间运行的格式化文本文件的进程。为了不使整个 shell 在格式化过程中都处于“瘫痪”状态，从后台启动这个进程是明智的选择。

（2）调度启动

有时，系统需要进行一些比较费时而且占用资源的维护工作，并且这些工作适合在深夜无人职守的时候进行，这时用户就可以事先进行调度安排，指定任务运行的时间或者场合，到时候系统就会自动完成这一切工作。

使用调度启动进程有几个常用的命令，如 at 命令在指定时刻执行相关进程，cron 命令可以自动周期性地执行相关进程，在需要使用时读者可以查看相关帮助手册。

2．调度进程

调度进程包括对进程的中断操作、改变优先级、查看进程状态等，在 Linux 下可以使用相关的系统命令实现其操作，下表列出了 Linux 中常见的调用进程的系统命令，读者在需要的时候可以自行查找其用法。

表 7.1　　Linux 中进程调度常见命令

选　　项	参 数 含 义
Ps	查看系统中的进程
Top	动态显示系统中的进程
Nice	按用户指定的优先级运行
Renice	改变正在运行进程的优先级
Kill	终止进程（包括后台进程）
crontab	用于安装、删除或者列出用于驱动 cron 后台进程的任务。
Bg	将挂起的进程放到后台执行

7.2　Linux 进程控制编程

进程创建

1．fork()

在 Linux 中创建一个新进程的惟一方法是使用 fork 函数。fork 函数是 Linux 中一个非常重要的函数，和读者以往遇到的函数也有很大的区别，它执行一次却返回两个值。希望读者能认真地学习这一部分的内容。

（1）fork 函数说明

fork 函数用于从已存在进程中创建一个新进程。新进程称为子进程，而原进程称为父进

程。这两个分别带回它们各自的返回值，其中父进程的返回值是子进程的进程号，而子进程则返回 0。因此，可以通过返回值来判定该进程是父进程还是子进程。

使用 fork 函数得到的子进程是父进程的一个复制品，它从父进程处继承了整个进程的地址空间，包括进程上下文、进程堆栈、内存信息、打开的文件描述符、信号控制设定、进程优先级、进程组号、当前工作目录、根目录、资源限制、控制终端等，而子进程所独有的只有它的进程号、资源使用和计时器等。因此可以看出，使用 fork 函数的代价是很大的，它复制了父进程中的代码段、数据段和堆栈段里的大部分内容，使得 fork 函数的执行速度并不很快。

（2）fork 函数语法

表 7.2 列出了 fork 函数的语法要点。

表 7.2　　fork 函数语法要点

所需头文件	#include <sys/types.h> // 提供类型 pid_t 的定义 #include <unistd.h>
函数原型	pid_t fork(void)
函数返回值	0：子进程
	子进程 ID（大于 0 的整数）：父进程
	–1：出错

（3）fork 函数使用实例

```
/*fork.c*/
#include <sys/types.h>
#include <unistd.h>
#include <stdio.h>
#include <stdlib.h>

int main(void)
{
      pid_t result;
/*调用 fork 函数，其返回值为 result*/
      result = fork();
/*通过 result 的值来判断 fork 函数的返回情况，首先进行出错处理*/
      if(result == -1){
            perror("fork");
            exit;
      }
/*返回值为 0 代表子进程*/
      else if(result == 0){
      printf("The  return  value  is  %d\nIn  child  process!!\nMy  PID  is
```

```
%d\n",result,getpid());
    }
    /*返回值大于 0 代表父进程*/
          else
          {
          printf("The  return  value  is  %d\nIn  father  process!!\nMy  PID  is
%d\n",result,getpid());
          }
    }
    [root@localhost process]# arm-linux-gcc fork..c –o fork
```

将可执行程序下载到目标板上，运行结果如下所示：

```
    The return valud s 76
    In father process!!
    My PID is 75
    The return value is :0
    In child process!!
    My PID is 76
```

从该实例中可以看出，使用 fork 函数新建了一个子进程，其中的父进程返回子进程的 PID，而子进程的返回值为 0。

（4）函数使用注意点

fork 函数使用一次就创建一个进程，所以若把 fork 函数放在了 if else 判断语句中则要小心，不能多次使用 fork 函数。

小知识

由于 fork 完整地拷贝了父进程的整个地址空间，因此执行速度是比较慢的。为了加快 fork 的执行速度，有些 UNIX 系统设计者创建了 vfork。vfork 也能创建新进程，但它不产生父进程的副本。它是通过允许父子进程可访问相同物理内存从而伪装了对进程地址空间的真实拷贝，当子进程需要改变内存中数据时才拷贝父进程。这就是著名的“写操作时拷贝”（copy-on-write）技术。

现在很多嵌入式 Linux 系统的 fork 函数调用都采用 vfork 函数的实现方式，实际上 uClinux 所有的多进程管理都通过 vfork 来实现。

2．exec 函数族

（1）exec 函数族说明

fork 函数是用于创建一个子进程，该子进程几乎拷贝了父进程的全部内容，但是，这个新创建的进程如何执行呢？这个 exec 函数族就提供了一个在进程中启动另一个程序执行的方法。它可以根据指定的文件名或目录名找到可执行文件，并用它来取代原调用进程的数据段、代码段和堆栈段，在执行完之后，原调用进程的内容除了进程号外，其他全部被新的进

程替换了。另外，这里的可执行文件既可以是二进制文件，也可以是 Linux 下任何可执行的脚本文件。

在 Linux 中使用 exec 函数族主要有两种情况：

- 当进程认为自己不能再为系统和用户做出任何贡献时，就可以调用任何 exec 函数族让自己重生；
- 如果一个进程想执行另一个程序，那么它就可以调用 fork 函数新建一个进程，然后调用任何一个 exec，这样看起来就好像通过执行应用程序而产生了一个新进程。（这种情况非常普遍）

（2）exec 函数族语法

实际上，在 Linux 中并没有 exec()函数，而是有 6 个以 exec 开头的函数族，它们之间语法有细微差别，本书在下面会详细讲解。

下表 7.3 列举了 exec 函数族的 6 个成员函数的语法。

表 7.3 exec 函数族成员函数语法

所需头文件	#include <unistd.h>
函数原型	int execl(const char *path, const char *arg, ...)
	int execv(const char *path, char *const argv[])
	int execle(const char *path, const char *arg, ..., char *const envp[])
	int execve(const char *path, char *const argv[], char *const envp[])
	int execlp(const char *file, const char *arg, ...)
	int execvp(const char *file, char *const argv[])
函数返回值	-1：出错

这 6 个函数在函数名和使用语法的规则上都有细微的区别，下面就可执行文件查找方式、参数表传递方式及环境变量这几个方面进行比较。

- 查找方式

读者可以注意到，表 7.3 中的前 4 个函数的查找方式都是完整的文件目录路径，而最后 2 个函数（也就是以 p 结尾的两个函数）可以只给出文件名，系统就会自动从环境变量“$PATH”所指出的路径中进行查找。

- 参数传递方式

exec 函数族的参数传递有两种方式：一种是逐个列举的方式，而另一种则是将所有参数整体构造指针数组传递。

在这里是以函数名的第 5 位字母来区分的，字母为“l”（list）的表示逐个列举的方式，其语法为 char *arg；字母为“v”（vertor）的表示将所有参数整体构造指针数组传递，其语法为*const argv[]。读者可以观察 execl、execle、execlp 的语法与 execv、execve、execvp 的区别。它们具体的用法在后面的实例讲解中会举例说明。

这里的参数实际上就是用户在使用这个可执行文件时所需的全部命令选项字符串（包括该可执行程序命令本身）。要注意的是，这些参数必须以 NULL 表示结束，如果使用逐个列举方式，那么要把它强制转化成一个字符指针，否则 exec 将会把它解释为一个整型参数，如

果一个整型数的长度 char *的长度不同，那么 exec 函数就会报错。

- 环境变量

exec 函数族可以默认系统的环境变量，也可以传入指定的环境变量。这里以“e”（Enviromen）结尾的两个函数 execle、execve 就可以在 envp[]中指定当前进程所使用的环境变量。

下表 7.4 再对这 4 个函数中函数名和对应语法做一总结，主要指出了函数名中每一位所表明的含义，希望读者结合此表加以记忆。

表 7.4　　exec 函数名对应含义

前 4 位	统一为：exec	
第 5 位	l：参数传递为逐个列举方式	execl、execle、execlp
	v：参数传递为构造指针数组方式	execv、execve、execvp
第 6 位	e：可传递新进程环境变量	execle、execve
	p：可执行文件查找方式为文件名	execlp、execvp

（3）exec 使用实例

下面的第一个示例说明了如何使用文件名的方式来查找可执行文件，同时使用参数列表的方式。这里用的函数是 execlp。

```
/*execlp.c*/
#include <unistd.h>
#include <stdio.h>
#include <stdlib.h>

int main()
{
     if(fork()==0){
/*调用 execlp 函数，这里相当于调用了“ps -ef”命令*/
         if(execlp("ps","ps","-ef",NULL)<0)
               perror("execlp error!");
     }
}
```

在该程序中，首先使用 fork 函数新建一个子进程，然后在子进程里使用 execlp 函数。读者可以看到，这里的参数列表就是在 shell 中使用的命令名和选项。并且当使用文件名的方式进行查找时，系统会在默认的环境变量 PATH 中寻找该可执行文件。读者可将编译后的结果下载到目标板上，运行结果如下所示：

```
[root@(none) 1]# ./execlp
  PID TTY    Uid       Size State Command
    1       root      1832    S    init
```

```
    2        root         0    S    [keventd]
    3        root         0    S    [ksoftirqd_CPU0]
    4        root         0    S    [kswapd]
    5        root         0    S    [bdflush]
    6        root         0    S    [kupdated]
    7        root         0    S    [mtdblockd]
    8        root         0    S    [khubd]
   35        root      2104    S    /bin/bash /usr/etc/rc.local
   36        root      2324    S    /bin/bash
   41        root      1364    S    /sbin/inetd
   53        root     14260    S    /Qtopia/qtopia-free-1.7.0/bin/qpe -qws
   54        root     11672    S    quicklauncher
   65        root         0    S    [usb-storage-0]
   66        root         0    S    [scsi_eh_0]
   83        root      2020    R    ps -ef
[root@(none) /]# env
PATH=/Qtopia/qtopia-free-1.7.0/bin:/usr/bin:/bin:/usr/sbin:/sbin
…
```

此程序的运行结果与在 Shell 中直接键入命令“ps -ef”是一样的，当然，在不同的系统不同时刻都可能会有不同的结果。

接下来的示例 2 使用完整的文件目录来查找对应的可执行文件。注意目录必须以“/”开头，否则将其视为文件名。

```
/*execl.c*/
#include <unistd.h>
#include <stdio.h>
#include <stdlib.h>

int main()
{
     if(fork()==0){
/*调用 execl 函数，注意这里要给出 ps 程序所在的完整路径*/
          if(execl("/bin/ps","ps","-ef",NULL)<0)
               perror("execl error!");
     }
}
```

同样下载到目标板上运行，运行结果同上例，如下所示：

```
[root@(none) 1]# ./execl
```

```
  PID TTY       Uid       Size State Command
    1           root      1832    S   init
    2           root         0    S   [keventd]
    3           root         0    S   [ksoftirqd_CPU0]
    4           root         0    S   [kswapd]
    5           root         0    S   [bdflush]
    6           root         0    S   [kupdated]
…
```

示例 3 利用函数 execle，将环境变量添加到新建的子进程中去，这里的“env”是查看当前进程环境变量的命令，如下所示：

```
/*execle*/
#include <unistd.h>
#include <stdio.h>
#include <stdlib.h>

int main()
{
/*命令参数列表，必须以 NULL 结尾*/
      char *envp[]={"PATH=/tmp","USER=sunq",NULL};
      if(fork()==0){
/*调用 execle 函数，注意这里也要指出 env 的完整路径*/
          if(execle("/bin/env","env",NULL,envp)<0)
               perror("execle error!");
      }
}
```

下载到目标板后的运行结果如下所示：

```
[root@(none) 1]# ./execle
PATH=/tmp
USER=sunq
```

最后一个示例使用 execve 函数，通过构造指针数组的方式来传递参数，注意参数列表一定要以 NULL 作为结尾标识符。其代码和运行结果如下所示：

```
#include <unistd.h>
#include <stdio.h>
#include <stdlib.h>

int main()
```

```
{
/*命令参数列表，必须以 NULL 结尾*/
      char *arg[]={"env",NULL};
      char *envp[]={"PATH=/tmp","USER=sunq",NULL};
      if(fork()==0){
            if(execve("/bin/env",arg,,envp)<0)
                  perror("execve error!");
      }
}
```

下载到目标板后的运行结果如下所示：

```
[root@(none) 1]# ./execve
PATH=/tmp
USER=sunq
```

（4）exec 函数族使用注意点

在使用 exec 函数族时，一定要加上错误判断语句。因为 exec 很容易执行失败，其中最常见的原因有：

- 找不到文件或路径，此时 errno 被设置为 ENOENT；
- 数组 argv 和 envp 忘记用 NULL 结束，此时 errno 被设置为 EFAULT；
- 没有对应可执行文件的运行权限，此时 errno 被设置为 EACCES。

小知识 事实上，这 6 个函数中真正的系统调用只有 execve，其他 5 个都是库函数，它们最终都会调用 execve 这个系统调用。

3．exit 和_exit

（1）exit 和_exit 函数说明

exit 和_exit 函数都是用来终止进程的。当程序执行到 exit 或_exit 时，进程会无条件地停止剩下的所有操作，清除包括 PCB 在内的各种数据结构，并终止本进程的运行。但是，这两个函数还是有区别的，这两个函数的调用过程如图 7.4 所示。

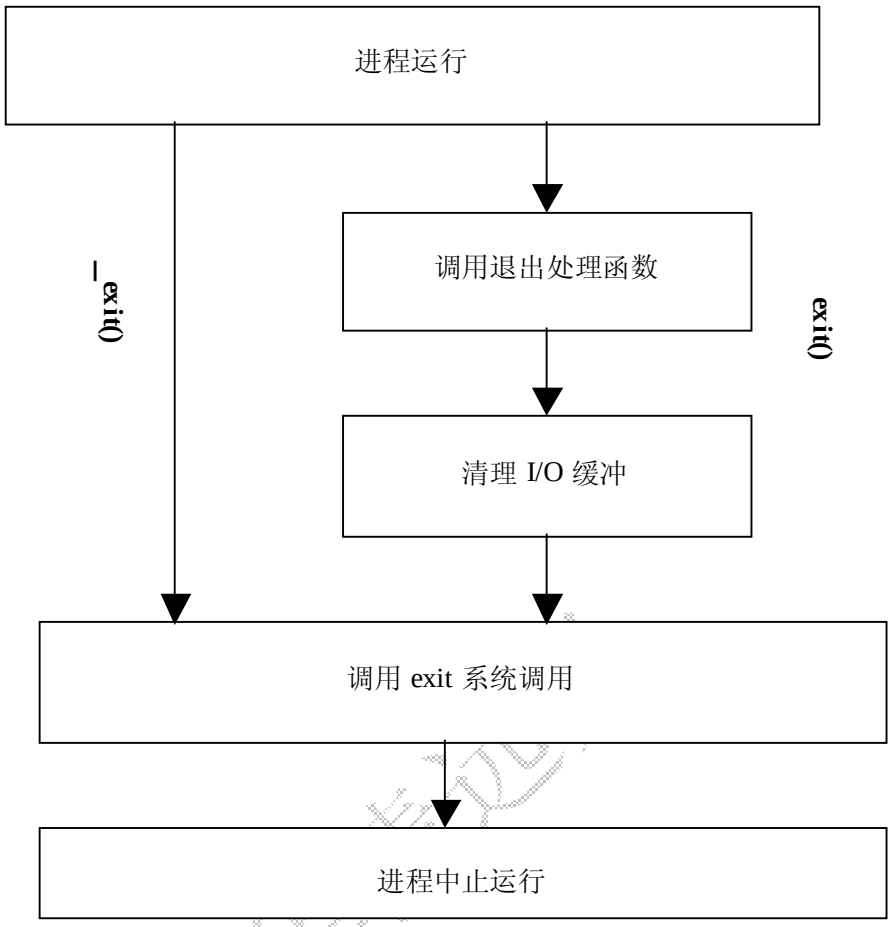


图 7.4　exit 和_exit 函数流程图

从图中可以看出，_exit()函数的作用是：直接使进程停止运行，清除其使用的内存空间，并清除其在内核中的各种数据结构；exit()函数则在这些基础上作了一些包装，在执行退出之前加了若干道工序。exit()函数与_exit()函数最大的区别就在于 exit()函数在调用 exit 系统之前要检查文件的打开情况，把文件缓冲区中的内容写回文件，就是图中的“清理 I/O 缓冲”一项。

由于在 Linux 的标准函数库中，有一种被称作“缓冲 I/O（buffered I/O）”操作，其特征就是对应每一个打开的文件，在内存中都有一片缓冲区。每次读文件时，会连续读出若干条记录，这样在下次读文件时就可以直接从内存的缓冲区中读取；同样，每次写文件的时候，也仅仅是写入内存中的缓冲区，等满足了一定的条件（如达到一定数量或遇到特定字符等），再将缓冲区中的内容一次性写入文件。

这种技术大大增加了文件读写的速度，但也为编程带来了一点麻烦。比如有一些数据，认为已经写入了文件，实际上因为没有满足特定的条件，它们还只是保存在缓冲区内，这时用_exit()函数直接将进程关闭，缓冲区中的数据就会丢失。因此，若想保证数据的完整性，就一定要使用 exit()函数。

（2）exit 和_exit 函数语法

下表 7.5 列出了 exit 和_exit 函数的语法规范。

表 7.5 exit 和_exit 函数族语法

所需头文件	exit：#include <stdlib.h>
	_exit：#include <unistd.h>
	续表
函数原型	exit：void exit(int status)
	_exit：void _exit(int status)
函数传入值	status 是一个整型的参数，可以利用这个参数传递进程结束时的状态。一般来说，0 表示正常结束；其他的数值表示出现了错误，进程非正常结束。 在实际编程时，可以用 wait 系统调用接收子进程的返回值，从而针对不同的情况进行不同的处理

（3）exit 和_exit 使用实例：

这两个示例比较了 exit 和_exit 两个函数的区别。由于 printf 函数使用的是缓冲 I/O 方式，该函数在遇到“\n”换行符时自动从缓冲区中将记录读出。示例中就是利用这个性质来进行比较的。以下是示例 1 的代码：

```
/*exit.c*/
#include <stdio.h>
#include <stdlib.h>

int main()
{
      printf("Using exit...\n");
      printf("This is the content in buffer");
      exit(0);
}
[root@(none) 1]# ./exit
Using exit...
This is the content in buffer[root@(none) 1]#
```

读者从输出的结果中可以看到，调用 exit 函数时，缓冲区中的记录也能正常输出。

以下是示例 2 的代码：

/*_exit.c*/

```
#include <stdio.h>
#include <unistd.h>
int main()
{
      printf("Using _exit...\n");
```

```
        printf("This is the content in buffer");
        _exit(0);
}
[root@(none) 1]# ./_exit
Using _exit...
[root@(none) 1]#
```

读者从最后的结果中可以看到，调用_exit 函数无法输出缓冲区中的记录。

小知识 在一个进程调用了 exit 之后，该进程并不马上就完全消失，而是留下一个称为僵尸进程（Zombie）的数据结构。僵尸进程是一种非常特殊的进程，它几乎已经放弃了所有内存空间，没有任何可执行代码，也不能被调度，仅仅在进程列表中保留一个位置，记载该进程的退出状态等信息供其他进程收集，除此之外，僵尸进程不再占有任何内存空间。

4．wait 和 waitpid

（1）wait 和 waitpid 函数说明

wait 函数是用于使父进程（也就是调用 wait 的进程）阻塞，直到一个子进程结束或者该进程接到了一个指定的信号为止。如果该父进程没有子进程或者他的子进程已经结束，则 wait 就会立即返回。

waitpid 的作用和 wait 一样，但它并不一定要等待第一个终止的子进程，它还有若干选项，如可提供一个非阻塞版本的 wait 功能，也能支持作业控制。实际上 wait 函数只是 waitpid 函数的一个特例，在 Linux 内部实现 wait 函数时直接调用的就是 waitpid 函数。

（2）wait 和 waitpid 函数格式说明

表 7.6 列出了 wait 函数的语法规范。

表 7.6　　wait 函数族语法

所需头文件	#include <sys/types.h> #include <sys/wait.h>
函数原型	pid_t wait(int *status)
函数传入值	这里的 status 是一个整型指针，是该子进程退出时的状态 • status 若为空，则代表任意状态结束的子进程 • status 若不为空，则代表指定状态结束的子进程 另外，子进程的结束状态可由 Linux 中一些特定的宏来测定
函数返回值	成功：子进程的进程号 失败：–1

下表 7.7 列出了 waitpid 函数的语法规范。

表 7.7　　waitpid 函数语法

所需头文件	#include <sys/types.h> #include <sys/wait.h>
函数原型	pid_t waitpid(pid_t pid, int *status, int options)

函数传入值	pid	pid>0：只等待进程 ID 等于 pid 的子进程，不管已经有其他子进程运行结束退出了，只要指定的子进程还没有结束，waitpid 就会一直等下去
		pid=–1：等待任何一个子进程退出，此时和 wait 作用一样
		pid=0：等待其组 ID 等于调用进程的组 ID 的任一子进程
		pid<–1：等待其组 ID 等于 pid 的绝对值的任一子进程

续表

函数传入值	status	同 wait
	options	WNOHANG：若由 pid 指定的子进程不立即可用，则 waitpid 不阻塞，此时返回值为 0
		WUNTRACED：若实现某支持作业控制，则由 pid 指定的任一子进程状态已暂停，且其状态自暂停以来还未报告过，则返回其状态
		0：同 wait，阻塞父进程，等待子进程退出
函数返回值	正常：子进程的进程号	
	使用选项 WNOHANG 且没有子进程退出：0	
	调用出错：–1	

（3）waitpid 使用实例

由于 wait 函数的使用较为简单，在此仅以 waitpid 为例进行讲解。本例中首先使用 fork 新建一子进程，然后让其子进程暂停 5s（使用了 sleep 函数）。接下来对原有的父进程使用 waitpid 函数，并使用参数 WNOHANG 使该父进程不会阻塞。若有子进程退出，则 waitpid 返回子进程号；若没有子进程退出，则 waitpid 返回 0，并且父进程每隔一秒循环判断一次。该程的流程图如图 7.5 所示。

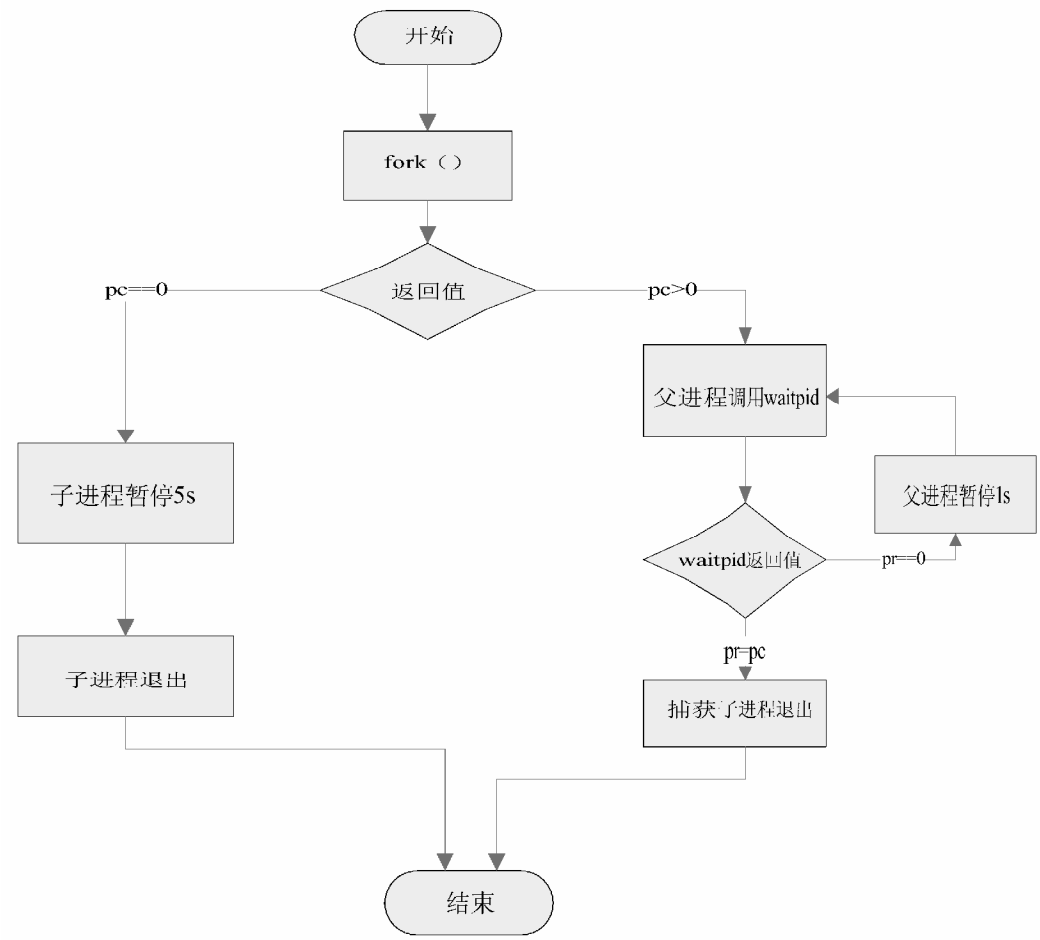


图 7.5　waitpid 示例程序流程图

该程序源代码如下所示：

```
/*waitpid.c*/
#include <sys/types.h>
#include <sys/wait.h>
#include <unistd.h>
#include <stdio.h>
#include <stdlib.h>

int main()
{
      pid_t pc,pr;
      pc=fork();
      if(pc<0)
            printf("Error fork.\n");
```

```
    /*子进程*/
        else if(pc==0){
    /*子进程暂停 5s*/
            sleep(5);
    /*子进程正常退出*/
            exit(0);
        }
    /*父进程*/
        else{
    /*循环测试子进程是否退出*/
            do{
    /*调用 waitpid，且父进程不阻塞*/
                pr=waitpid(pc,NULL,WNOHANG);
    /*若子进程还未退出，则父进程暂停 1s*/
                if(pr==0){
                    printf("The child process has not exited\n");
                    sleep(1);
                }
            }while(pr==0);
    /*若发现子进程退出，打印出相应情况*/
            if(pr==pc)
                printf("Get child %d\n",pr);
            else
                printf("some error occured.\n");
        }
    }
```

将该程序交叉编译，下载到目标板后的运行情况如下所示：

```
[root@(none) 1]# ./waitpid
The child process has not exited
The child process has not exited
The child process has not exited
The child process has not exited
The child process has not exited
Get child 75
```

可见，该程序在经过 5 次循环之后，捕获到了子进程的退出信号，具体的子进程号在不同的系统上会有所区别。

读者还可以尝试把“pr=**waitpid(pc,NULL,WNOHANG);**”这句改为“pr=waitpid(pc,NULL,

0);”和“pr=wait(NULL);”，运行的结果为：

```
[root@(none) 1]# ./waitpid
Get child 76
```

可见，在上述两种情况下，父进程在调用 waitpid 或 wait 之后就将自己阻塞，直到有子进程退出为止。

7.3 Linux 守护进程

7.3.1 守护进程概述

守护进程，也就是通常所说的 Daemon 进程，是 Linux 中的后台服务进程。它是一个生存期较长的进程，通常独立于控制终端并且周期性地执行某种任务或等待处理某些发生的事件。守护进程常常在系统引导装入时启动，在系统关闭时终止。Linux 系统有很多守护进程，大多数服务都是通过守护进程实现的，如本书在第二章中讲到的系统服务都是守护进程。同时，守护进程还能完成许多系统任务，例如，作业规划进程 crond、打印进程 lqd 等（这里的结尾字母 d 就是 Daemon 的意思）。

由于在 Linux 中，每一个系统与用户进行交流的界面称为终端，每一个从此终端开始运行的进程都会依附于这个终端，这个终端就称为这些进程的控制终端，当控制终端被关闭时，相应的进程都会自动关闭。但是守护进程却能够突破这种限制，它从被执行开始运转，直到整个系统关闭时才会退出。如果想让某个进程不因为用户或终端或其他的变化而受到影响，那么就必须把这个进程变成一个守护进程。可见，守护进程是非常重要的。

7.3.2 编写守护进程

编写守护进程看似复杂，但实际上也是遵循一个特定的流程。只要将此流程掌握了，就能很方便地编写出用户自己的守护进程。下面就分 4 个步骤来讲解怎样创建一个简单的守护进程。在讲解的同时，会配合介绍与创建守护进程相关的几个系统函数，希望读者能很好地掌握。

1．创建子进程，父进程退出

这是编写守护进程的第一步。由于守护进程是脱离控制终端的，因此，完成第一步后就会在 Shell 终端里造成一程序已经运行完毕的假象。之后的所有工作都在子进程中完成，而用户在 Shell 终端里则可以执行其他的命令，从而在形式上做到了与控制终端的脱离。

到这里，有心的读者可能会问，父进程创建了子进程，而父进程又退出之后，此时该子进程不就没有父进程了吗？守护进程中确实会出现这么一个有趣的现象，由于父进程已经先于子进程退出，会造成子进程没有父进程，从而变成一个孤儿进程。在 Linux 中，每当系统发现一个孤儿进程，就会自动由 1 号进程（也就是 init 进程）收养它，这样，原先的子进程就会变成 init 进程的子进程了。其关键代码如下所示：

```
/*父进程退出*/
pid=fork();
      if(pid>0){
            exit(0);
}
```

2．在子进程中创建新会话

这个步骤是创建守护进程中最重要的一步，虽然它的实现非常简单，但它的意义却非常重大。在这里使用的是系统函数 setsid，在具体介绍 setsid 之前，读者首先要了解两个概念：进程组和会话期。

- 进程组

进程组是一个或多个进程的集合。进程组由进程组 ID 来惟一标识。除了进程号（PID）之外，进程组 ID 也一个进程的必备属性。

每个进程组都有一个组长进程，其组长进程的进程号等于进程组 ID。且该进程 ID 不会因组长进程的退出而受到影响。

- 会话期

会话组是一个或多个进程组的集合。通常，一个会话开始于用户登录，终止于用户退出，在此期间该用户运行的所有进程都属于这个会话期，它们之间的关系如下图 7.6 所示。

接下来就可以具体介绍 setsid 的相关内容：

（1）setsid 函数作用

setsid 函数用于创建一个新的会话，并担任该会话组的组长。调用 setsid 有下面的 3 个作用。

- 让进程摆脱原会话的控制。
- 让进程摆脱原进程组的控制。
- 让进程摆脱原控制终端的控制。

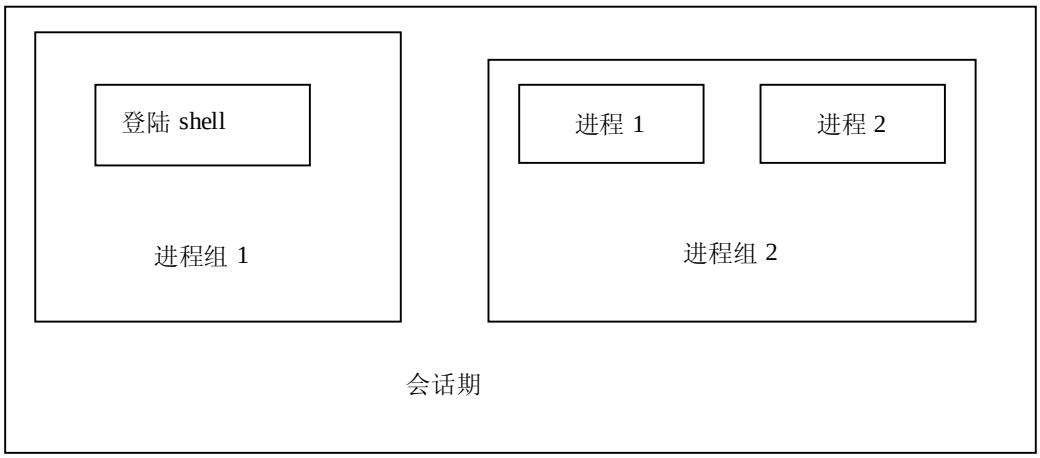


图 7.6　进程组、会话期关系图

那么，在创建守护进程时为什么要调用 setsid 函数呢？读者可以回忆一下创建守护进程的第一步，在那里调用了 fork 函数来创建子进程再将父进程退出。由于在调用 fork 函数时，子进程全盘拷贝了父进程的进会话期、进程组、控制终端等，虽然父进程退出了，但原先的会话期、进程组、控制终端等并没有改变，因此，还不是真正意义上独立开来，而 setsid 函数能够使进程完全独立出来，从而脱离所有其他进程的控制。

（2）setsid 函数格式

下表 7.8 列出了 setsid 函数的语法规范

表 7.8　　setsid 函数语法

所需头文件	#include <sys/types.h> #include <unistd.h>
函数原型	pid_t setsid(void)
函数返回值	成功：该进程组 ID 出错：–1

3．改变当前目录为根目录

这一步也是必要的步骤。使用 fork 创建的子进程继承了父进程的当前工作目录。由于在进程运行过程中，当前目录所在的文件系统（比如“/mnt/usb”等）是不能卸载的，这对以后的使用会造成诸多的麻烦（比如系统由于某种原因要进入单用户模式）。因此，通常的做法是让“/”作为守护进程的当前工作目录，这样就可以避免上述的问题，当然，如有特殊需要，也可以把当前工作目录换成其他的路径，如/tmp。改变工作目录的常见函数是 chdir。

4．重设文件权限掩码

文件权限掩码是指屏蔽掉文件权限中的对应位。比如，有一个文件权限掩码是 050，它就屏蔽了文件组拥有者的可读与可执行权限。由于使用 fork 函数新建的子进程继承了父进程

的文件权限掩码，这就给该子进程使用文件带来了诸多的麻烦。因此，把文件权限掩码设置为 0，可以大大增强该守护进程的灵活性。设置文件权限掩码的函数是 umask。在这里，通常的使用方法为 umask(0)。

5．关闭文件描述符

同文件权限掩码一样，用 fork 函数新建的子进程会从父进程那里继承一些已经打开了的文件。这些被打开的文件可能永远不会被守护进程读或写，但它们一样消耗系统资源，而且可能导致所在的文件系统无法卸下。

在上面的第二步之后，守护进程已经与所属的控制终端失去了联系。因此从终端输入的字符不可能达到守护进程，守护进程中用常规方法（如 printf）输出的字符也不可能在终端上显示出来。所以，文件描述符为 0、1 和 2 的 3 个文件（常说的输入、输出和报错这 3 个文件）已经失去了存在的价值，也应被关闭。通常按如下方式关闭文件描述符：

```
for(i=0;i<MAXFILE;i++)
      close(i);
```

这样，一个简单的守护进程就建立起来了，创建守护进程的流程图如图 7.7 所示。

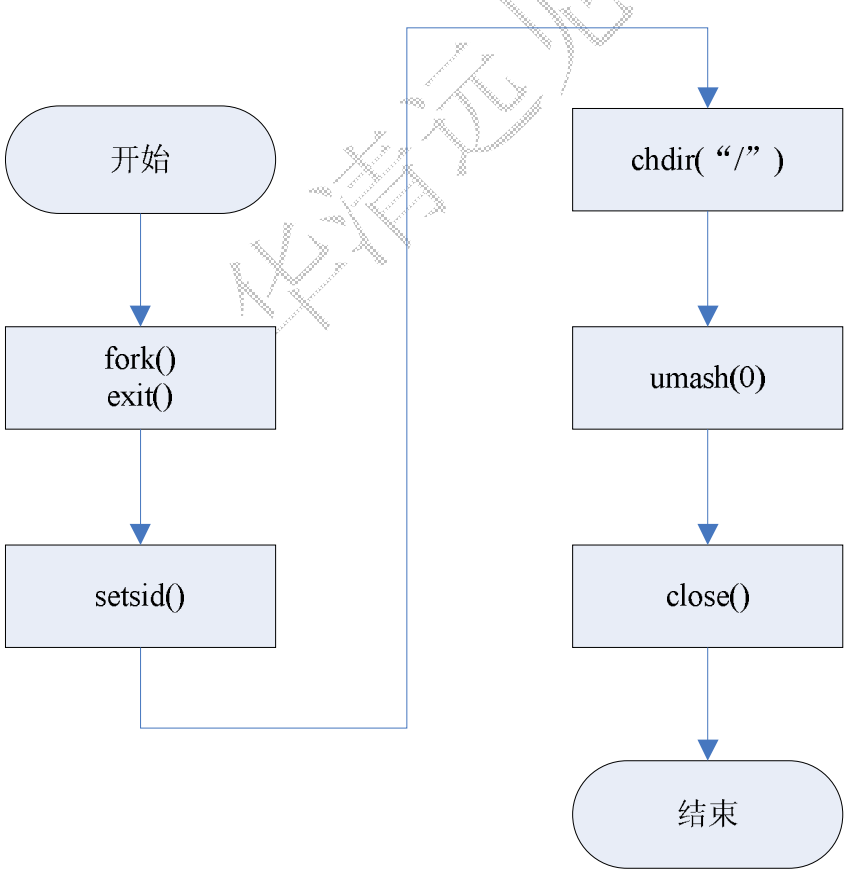


图 7.7　创建守护进程流程图

下面是实现守护进程的一个完整实例：

该实例首先建立了一个守护进程，然后让该守护进程每隔 10s 在/tmp/dameon.log 中写入一句话。

```
/*dameon.c 创建守护进程实例*/
#include<stdio.h>
#include<stdlib.h>
#include<string.h>
#include<fcntl.h>
#include<sys/types.h>
#include<unistd.h>
#include<sys/wait.h>

#define MAXFILE 65535
int main()
{
      pid_t pc;
      int i,fd,len;
      char *buf="This is a Dameon\n";
      len =strlen(buf);
      pc=fork(); //第一步
      if(pc<0){
            printf("error fork\n");
            exit(1);
      }else if(pc>0)
      exit(0);
/*第二步*/
      setsid();
/*第三步*/
      chdir("/");
/*第四步*/
      umask(0);
      for(i=0;i<MAXFILE;i++)
/*第五步*/
            close(i);
/*这时创建完守护进程，以下开始正式进入守护进程工作*/
      while(1){
            if((fd=open("/tmp/dameon.log",O_CREAT|O_WRONLY|O_APPEND,0600))<0){
               perror("open");
```

```
                exit(1);
                }
            write(fd, buf, len+1);
            close(fd);
            sleep(10);
        }
    }
```

将该程序下载到开发板中，可以看到该程序每隔 10s 就会在对应的文件中输入相关内容。并且使用 ps 可以看到该进程在后台运行。如下所示：

```
[root@(none) 1]# tail -f /tmp/dameon.log
This is a Dameon
This is a Dameon
This is a Dameon
This is a Dameon
…
[root@(none) 1]# ps -ef|grep daemon
   76          root         1272   S   ./daemon
   85          root         1520   S   grep daemon
```

7.3.3 守护进程的出错处理

读者在前面编写守护进程的具体调试过程中会发现，由于守护进程完全脱离了控制终端，因此，不能像其他进程的程序一样通过输出错误信息到控制终端来通知程序员即使使用 gdb 也无法正常调试。那么，守护进程的进程要如何调试呢？一种通用的办法是使用 syslog 服务，将程序中的出错信息输入到“/var/log/messages”系统日志文件中，从而可以直观地看到程序的问题所在。

注意 “/var/log/message”系统日志文件只能由拥有 root 权限的超级用户查看。

Syslog 是 Linux 中的系统日志管理服务，通过守护进程 syslogd 来维护。该守护进程在启动时会读一个配置文件“/etc/syslog.conf”。该文件决定了不同种类的消息会发送向何处。例如，紧急消息可被送向系统管理员并在控制台上显示，而警告消息则可记录到一个文件中。

该机制提供了 3 个 syslog 函数，分别为 openlog、syslog 和 closelog。下面就分别介绍这 3 个函数。

（1）syslog 函数说明

通常，openlog 函数用于打开系统日志服务的一个连接；syslog 函数是用于向日志文件中写入消息，在这里可以规定消息的优先级、消息输出格式等；closelog 函数是用于关闭系统日志服务的连接。

（2）syslog 函数格式

下表 7.9 列出了 openlog 函数的语法规范

表 7.9　　openlog 函数语法

所需头文件	#include <syslog.h>	
函数原型	void openlog (char *ident,int option ,int facility)	
函数传入值	Ident	要向每个消息加入的字符串，通常为程序的名称
	Option	LOG_CONS：如果消息无法送到系统日志服务，则直接输出到系统控制终端
		LOG_NDELAY：立即打开系统日志服务的连接。在正常情况下，直到发送到第一条消息时才打开连接
		LOG_PERROR：将消息也同时送到 stderr 上
		LOG_PID：在每条消息中包含进程的 PID

续表

函数传入值	facility：指定程序发送的消息类型	LOG_AUTHPRIV：安全/授权讯息
		LOG_CRON：时间守护进程（cron 及 at）
		LOG_DAEMON：其他系统守护进程
		LOG_KERN：内核信息
		LOG_LOCAL[0~7]：保留
		LOG_LPR：行打印机子系统
		LOG_MAIL：邮件子系统
		LOG_NEWS：新闻子系统
		LOG_SYSLOG：syslogd 内部所产生的信息
		LOG_USER：一般使用者等级讯息
		LOG_UUCP：UUCP 子系统

表 7.10 列出了 syslog 函数的语法规范。

表 7.10　　syslog 函数语法

所需头文件	#include <syslog.h>	
函数原型	void syslog(int priority, char *format, ...)	
函数传入值	priority：指定消息的重要性	LOG_EMERG：系统无法使用
		LOG_ALERT：需要立即采取措施
		LOG_CRIT：有重要情况发生
		LOG_ERR：有错误发生
		LOG_WARNING：有警告发生
		LOG_NOTICE：正常情况，但也是重要情况
		LOG_INFO：信息消息

		LOG_DEBUG：调试信息
	format	以字符串指针的形式表示输出的格式，类似 printf 中的格式

表 7.11 列出了 closelog 函数的语法规范。

表 7.11 closelog 函数语法

所需头文件	#include <syslog.h>
函数原型	void closelog(void)

（3）使用实例

这里将上一节中的示例程序用 syslog 服务进行重写，其中有区别的地方用加粗的字体表示，源代码如下所示：

```
/*syslog_dema.c 利用 syslog 服务的守护进程实例*/
#include<stdio.h>
#include<stdlib.h>
#include<string.h>
#include<fcntl.h>
#include<sys/types.h>
#include<unistd.h>
#include<sys/wait.h>
#include<syslog.h>

#define MAXFILE 65535
int main()
{
      pid_t pc,sid;
      int i,fd,len;
      char *buf="This is a Dameon\n";
      len =strlen(buf);
     pc=fork();
      if(pc<0){
            printf("error fork\n");
            exit(1);
      }else if(pc>0)
      exit(0);
/*打开系统日志服务，openlog*/
      openlog("demo_update",LOG_PID, LOG_DAEMON);
      if((sid=setsid())<0){
            syslog(LOG_ERR, "%s\n", "setsid");
```

```
            exit(1);
        }
        if((sid=chdir("/"))<0){
            syslog(LOG_ERR, "%s\n", "chdir");
            exit(1);
        }
        umask(0);
        for(i=0;i<MAXFILE;i++)
            close(i);
        while(1){
    /*打开守护进程的日志文件，并写入 open 的日志记录*/
            if((fd=open("/tmp/dameon.log",O_CREAT|O_WRONLY|O_APPEND, 0600))<0){
                syslog(LOG_ERR, "open");
                exit(1);
                }
            write(fd, buf, len+1);
            close(fd);
            sleep(10);
        }
        closelog();
        exit(0);
    }
```

读者可以尝试用普通用户的身份执行此程序，由于这里的 open 函数必须具有 root 权限，因此，syslog 就会将错误信息写入到“/var/log/messages”中，如下所示：

```
Jan 30 18:20:08  localhost demo_update[7996]: open
```

7.4 实验内容

7.4.1 编写多进程程序

1. 实验目的

通过编写多进程程序，使读者熟练掌握 fork、exec、wait、waitpid 等函数的使用，进一步理解在 Linux 中多进程编程的步骤。

2. 实验内容

该实验有 3 个进程，其中一个为父进程，其余两个是该父进程创建的子进程，其中一个

子进程运行“ls -l”指令，另一个子进程在暂停 5s 之后异常退出，父进程并不阻塞自己，并等待子进程的退出信息，待收集到该信息，父进程就返回。

3．实验步骤

（1）画出该实验流程图

该实验流程图如图 7.8 所示。

（2）实验源代码

具体代码设置如下：

```
/*exc.c 实验一源码*/
#include <stdio.h>
#include <stdlib.h>
#include <sys/types.h>
#include <unistd.h>
#include <sys/wait.h>
int main(void)
{
      pid_t child1,child2,child;
/*创建两个子进程*/
      child1 = fork();
      child2 = fork();
/*子进程 1 的出错处理*/
      if( child1 == -1 ){
            perror("child1 fork");
            exit(1);
      }
/*在子进程 1 中调用 execlp 函数*/
      else if( child1 == 0 ){
            printf("In child1: execute 'ls -l'\n");
            if(execlp("ls","ls","-l",NULL)<0)
                  perror("child1 execlp");
    }
/*子进程 2 的出错处理*/
      if( child2 == -1 ){
            perror("child2 fork");
            exit(1);
      }
/*在子进程 2 中使其暂停 5s*/
      else if( child2 == 0 ){
```

```
            printf("In child2: sleep for 5 seconds and then exit\n");
            sleep(5);
            exit(0);
      }
/*在父进程中等待子进程 2 的退出*/
      else{
            printf("In father process:\n");
            do{
                  child = waitpid( child2, NULL, WNOHANG );
                  if( child ==0 ){
                        printf("The child2 process has not exited!\n");
                        sleep(1);
                  }
            }while( child == 0 );
            if( child == child2 )
                  printf("Get child2\n");
            else
                  printf("Error occured!\n");
      }
}
```

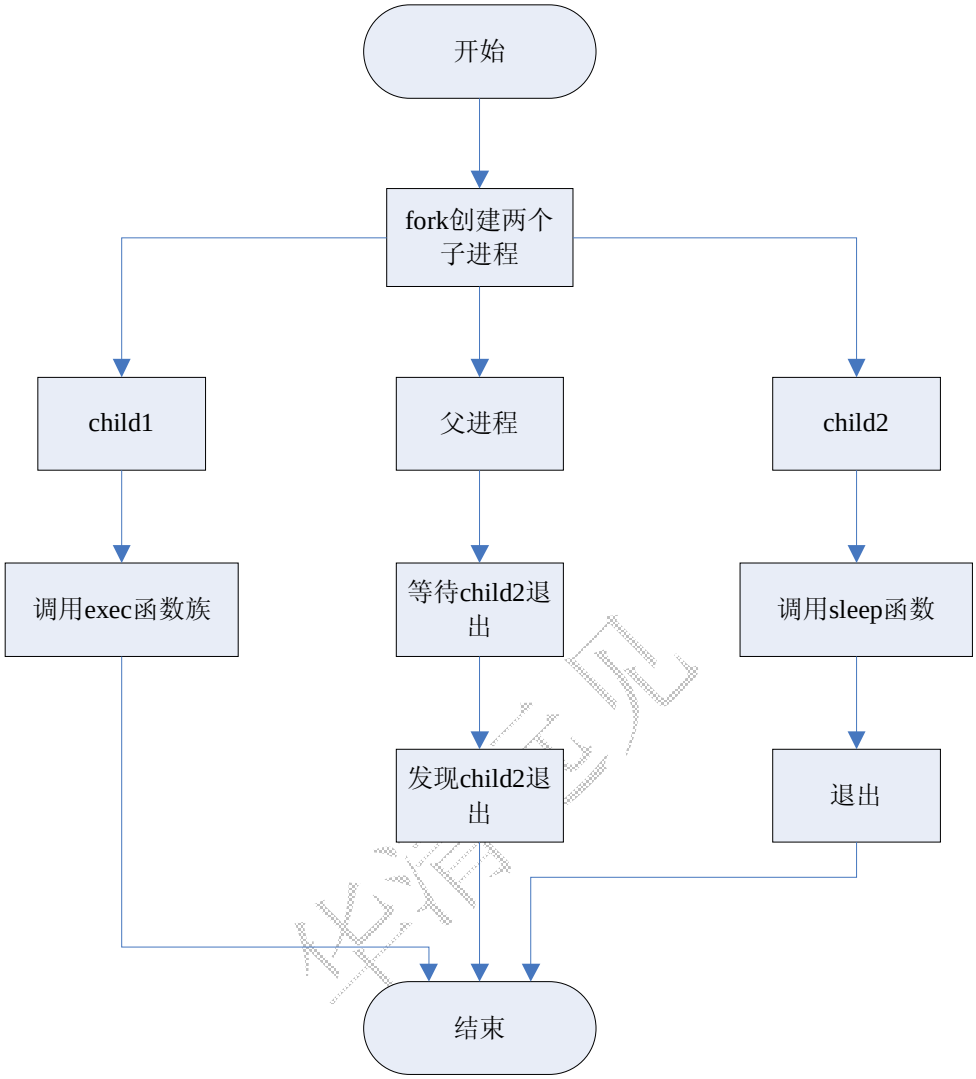


图 7.8　实验 7.5.1 流程图

（3）首先在宿主机上编译调试该程序：

```
[root@localhost process]# gcc exc.c -o exc
```

（4）在确保没有编译错误后，使用交叉编译该程序：

```
[root@localhost process]# arm-linux-gcc exc.c -o exc
```

（5）将生成的可执行程序下载到目标板上运行。

4．实验结果

在目标板上运行的结果如下所示（具体内容与各自的系统有关）：

```
[root@(none) 1]# ./exc
In child1: execute 'ls -l'
```

```
In child1: execute 'ls -l'
In child2: sleep for 5 seconds and then exit
total 57
-rwxr-xr-x   1 root    root       14443 Jan 31  2006 exc
-rwxr-xr-x   1 root    root       13512 Jan 29  2006 exit
-rwxr-xr-x   1 root    root       13956 Jan 29  2006 fork
-rwxr-xr-x   1 root    root       13999 Jan 30  2006 waitpid
total 57
-rwxr-xr-x   1 root    root       14443 Jan 31  2006 exc
-rwxr-xr-x   1 root    root       13512 Jan 29  2006 exit
-rwxr-xr-x   1 root    root       13956 Jan 29  2006 fork
-rwxr-xr-x   1 root    root       13999 Jan 30  2006 waitpid
In father process:
The child2 process has not exited!
The child2 process has not exited!
The child2 process has not exited!
The child2 process has not exited!
The child2 process has not exited!
Get child2
```

因为几个子进程的执行有竞争关系，因此，结果中的顺序没有完全按照程来所编写。读者可以思考怎样可以保证子进程的执行顺序呢？

7.4.2 编写守护进程

1．实验目的

通过编写一个完整的守护进程，使读者掌握守护进程编写和调试的方法，并且进一步熟悉编写多进程程序。

2．实验内容

在该实验中，读者首先建立起一个守护进程，然后在该守护进程中新建一个子进程，该子进程暂停 10s，然后自动退出，并由守护进程收集子进程退出的消息。在这里，子进程和守护进程的退出消息都在“/var/log/messages”中输出。子进程退出后，守护进程循环暂停，其间隔时间为 10s。

3．实验步骤

（1）画出该实验流程图

该程序流程图如图 7.9 所示。

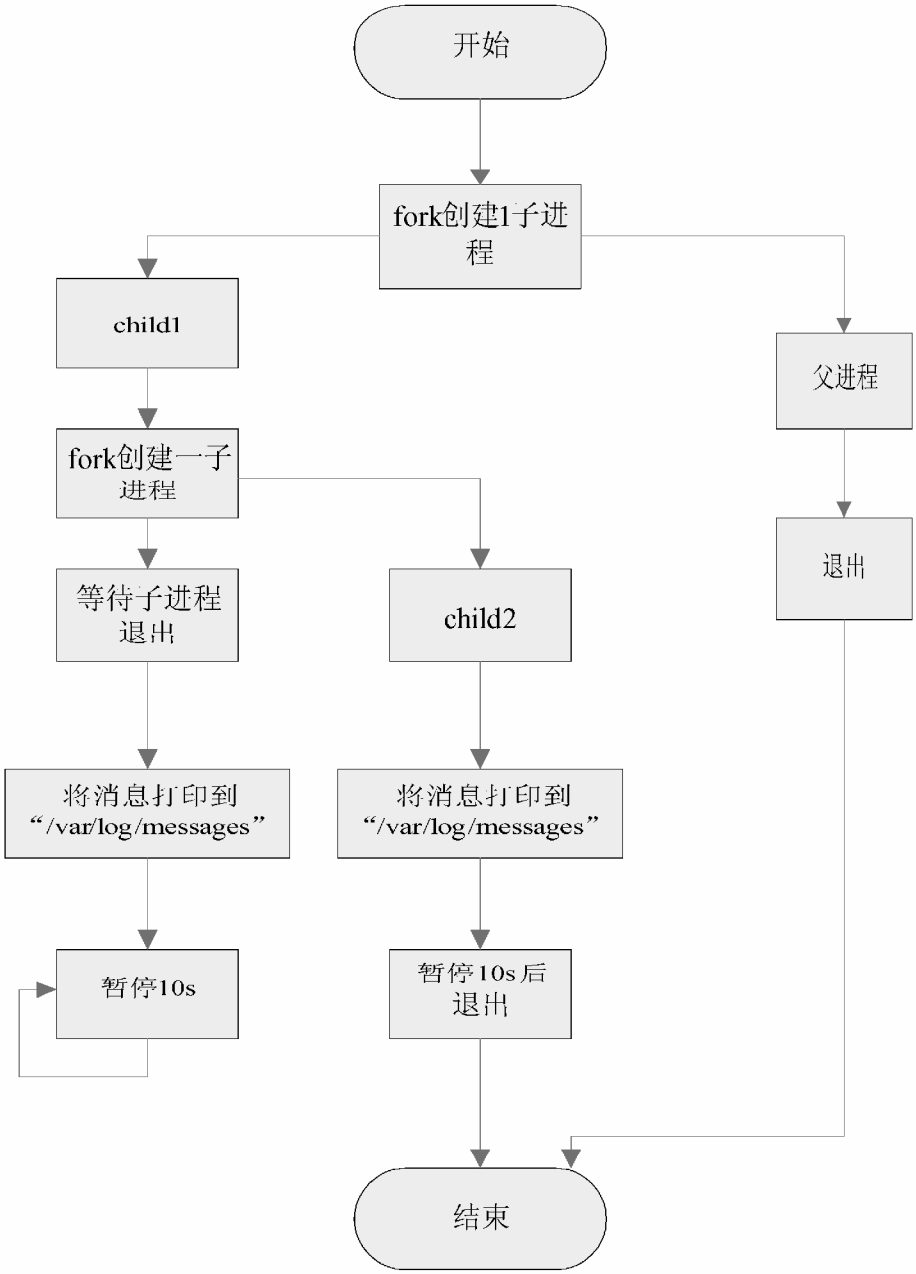


图 7.9　实验 7.7.2 流程图

（2）实验源代码

具体代码设置如下：

```
/*exc2.c 实验二源码*/
#include <stdio.h>
#include <stdlib.h>
#include <sys/types.h>
```

```
#include <unistd.h>
#include <sys/wait.h>
#include <syslog.h>
#define MAXFILE 65535
int main(void)
{
      pid_t child1,child2;
      int i;
      child1 = fork();
/*创建子进程 1*/
      if( child1 == –1 ){
            perror("child1 fork");
            exit(1);
      }
      else if( child1 > 0 )
            exit( 0 );
/*打开日志服务*/
      openlog("exc2_info", LOG_PID, LOG_DAEMON);
/*以下几步是编写守护进程的常规步骤*/
      setsid();
     chdir( "/" );
      umask( 0 );
      for( i = 0 ; i < MAXFILE ; i++ )
      {
            close( i );
      }
/*创建子进程 2*/
      child2 = fork();
      if( child2 == –1 ){
            perror("child2 fork");
            exit(1);
      }
      else if( child2 == 0 ){
/*在日志中写入字符串*/
            syslog( LOG_INFO, " child2 will sleep for 10s ");
            sleep(10);
            syslog( LOG_INFO, " child2 is going to exit! ");
            exit(0);
      }
```

```
        else{
            waitpid( child2, NULL, 0);
            syslog( LOG_INFO , " child1 noticed that child2 has exited " );
    /*关闭日志服务*/
            closelog();
            while(1){
                sleep(10);
            }
        }
    }
```

（3）由于有些嵌入式开发板没有 syslog 服务，读者可以在宿主机上编译运行。

```
    [root@localhost process]# gcc exc2.c –o exc2
```

（4）运行该程序。

（5）等待 10s 后，以 root 身份查看“/var/log/messages”文件。

（6）使用 ps –ef|grep exc2 查看该守护进程是否在运行。

4．实验结果

（1）在“/var/log/messages”中有类似如下的信息显示：

```
    Jan 31  13:59:11  localhost exc2_info[5517]: child2 will sleep for 10s
    Jan 31  13:59:21  localhost exc2_info[5517]: child2 is going to exit!
    Jan 31  13:59:21  localhost exc2_info[5516]: child1 noticed that child2 has
exited
```

读者可以从时间戳里清楚地看到 child2 确实暂停了 10s。

（2）使用命令 ps –ef|grep exc2 可看到如下结果：

```
    root    5516      1    0   13:59  ?      00:00:00  ./exc2
```

可见，exc2 确实一直在运行。

本 章 小 结

本章主要介绍进程的控制开发，首先给出了进程的基本概念，Linux 下进程的基本结构、模式与类型以及 Linux 进程管理。进程是 Linux 中程序运行和资源管理的最小单位，对进程的处理也是嵌入式 Linux 应用编程的基础，因此，读者一定要牢牢掌握。

接下来，本章具体讲解了进程控制编程，主要讲解了 fork 函数和 exec 函数族，并且举实例加以区别。Exec 函数族较为庞大，希望读者能够仔细比较它们之间的区别，认真体会并

理解。

最后，本章讲解了 Linux 守护进程的编写，包括守护进程的概念、编写守护进程的步骤以及守护进程的出错处理。由于守护进程非常特殊，因此，在编写时有不少的区别需要特别注意。守护进程的编写实际上涉及进程控制编程的很多部分，需要加以综合应用。

本章的实验安排了多进程编程和编写完整的守护进程两个部分。这两个实验都是较为综合性的，希望读者能够认真完成。

思考与练习

查阅资料，明确 Linux 中进程处理和嵌入式 Linux 中对进程的处理有什么区别？

“黑色经典”系列之《嵌入式 Linux 应用程序开发详解》

第 8 章 进程间通信

本章目标

在上一章中，读者已经学会了如何创建进程以及如何对进程进行基本的控制，而这些都只是停留在父子进程之间的控制，本章将要学习不同的进程间进行通信的方法，通过本章的学习，读者将会掌握如下内容。

- 掌握 Linux 中管道的基本概念
- 掌握 Linux 中管道的创建
- 掌握 Linux 中管道的读写
- 掌握 Linux 中有名管道的创建读写方法
- 掌握 Linux 中消息队列的处理
- 掌握 Linux 共享内存的处理

8.1 Linux 下进程间通信概述

在上一章中，读者已经知道了进程是一个程序的一次执行的过程。这里所说的进程一般是指运行在用户态的进程，而由于处于用户态的不同进程之间是彼此隔离的，就像处于不同城市的人们，它们必须通过某种方式来提供通信，例如人们现在广泛使用的手机等方式。本章就是讲述如何建立这些不同的通话方式，就像人们有多种通信方式一样。

Linux 下的进程通信手段基本上是从 UNIX 平台上的进程通信手段继承而来的。而对 UNIX 发展做出重大贡献的两大主力 AT&T 的贝尔实验室及 BSD（加州大学伯克利分校的伯克利软件发布中心）在进程间的通信方面的侧重点有所不同。前者是对 UNIX 早期的进程间通信手段进行了系统的改进和扩充，形成了“system V IPC”，其通信进程主要局限在单个计算机内；后者则跳过了该限制，形成了基于套接口（socket）的进程间通信机制。而 Linux 则把两者的优势都继承了下来，如图 8.1 所示。

- UNIX 进程间通信（IPC）方式包括管道、FIFO、信号。

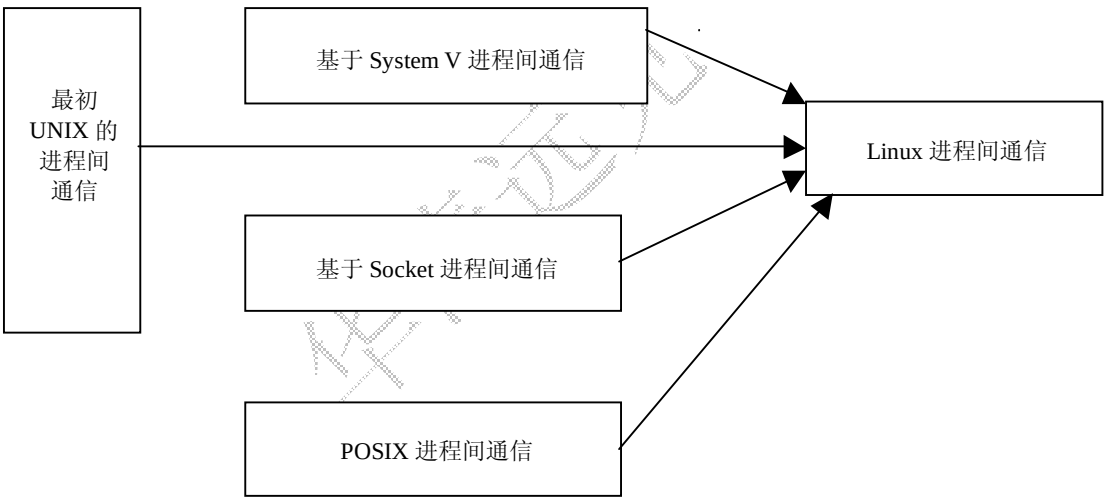


图 8.1 进程间通信发展历程

- System V 进程间通信（IPC）包括 System V 消息队列、System V 信号灯、System V 共享内存区。
- Posix 进程间通信（IPC）包括 Posix 消息队列、Posix 信号灯、Posix 共享内存区。

现在在 Linux 中使用较多的进程间通信方式主要有以下几种。

（1）管道（Pipe）及有名管道（named pipe）：管道可用于具有亲缘关系进程间的通信，有名管道，除具有管道所具有的功能外，它还允许无亲缘关系进程间的通信。

（2）信号（Signal）：信号是在软件层次上对中断机制的一种模拟，它是比较复杂的通信方式，用于通知接受进程有某事件发生，一个进程收到一个信号与处理器收到一个中断请求效果上可以说是一样的。

（3）消息队列：消息队列是消息的链接表，包括 Posix 消息队列 systemV 消息队列。它克服了前两种通信方式中信息量有限的缺点，具有写权限的进程可以向消息队列中按照一定的规则添加新消息；对消息队列有读权限的进程则可以从消息队列中读取消息。

（4）共享内存：可以说这是最有用的进程间通信方式。它使得多个进程可以访问同一块内存空间，不同进程可以及时看到对方进程中对共享内存中数据的更新。这种通信方式需要依靠某种同步机制，如互斥锁和信号量等。

（5）信号量：主要作为进程间以及同一进程不同线程之间的同步手段。

（6）套接字（Socket）：这是一种更为一般的进程间通信机制，它可用于不同机器之间的进程间通信，应用非常广泛。

本章会详细介绍前 4 种进程通信方式，对第 5 种通信方式将会在第 10 章中单独介绍。

8.2 管道通信

8.2.1 管道概述

细心的读者可能会注意到本书在第 2 章中介绍“ps”的命令时提到过管道，当时指出了管道是 Linux 中很重要的一种通信方式，它是把一个程序的输出直接连接到另一个程序的输入，这里仍以第 2 章中的“ps –ef|grep ntp”为例，描述管道的通信过程，如图 8.2 所示。

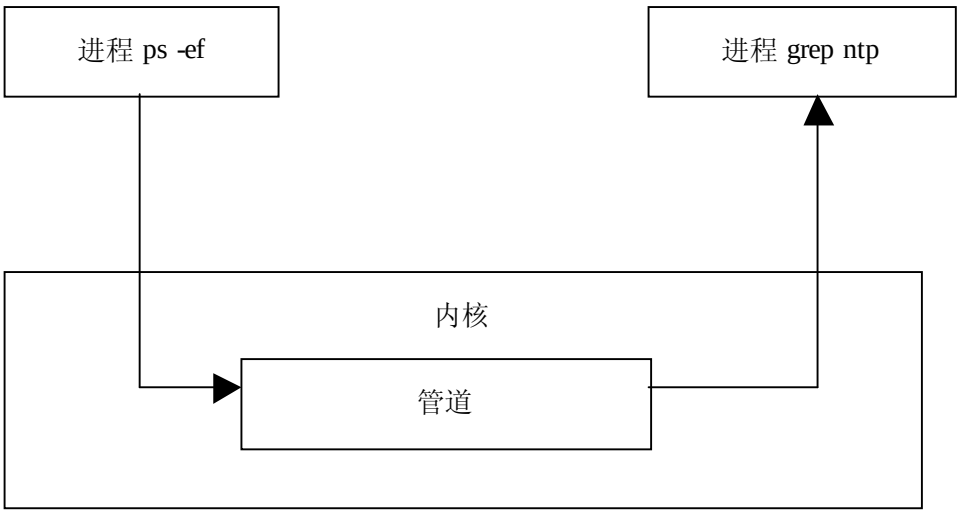


图 8.2　管道的通信过程

管道是 Linux 中进程间通信的一种方式。这里所说的管道主要指无名管道，它具有如下特点。

- 它只能用于具有亲缘关系的进程之间的通信（也就是父子进程或者兄弟进程之间）。
- 它是一个半双工的通信模式，具有固定的读端和写端。
- 管道也可以看成是一种特殊的文件，对于它的读写也可以使用普通的 read、write 等

函数。但是它不是普通的文件，并不属于其他任何文件系统，并且只存在于内存中。

8.2.2 管道创建与关闭

1．管道创建与关闭说明

管道是基于文件描述符的通信方式，当一个管道建立时，它会创建两个文件描述符 fds[0]和 fds[1]，其中 fds[0]固定用于读管道，而 fd[1]固定用于写管道，如图 8.3 所示，这样就构成了一个半双工的通道。

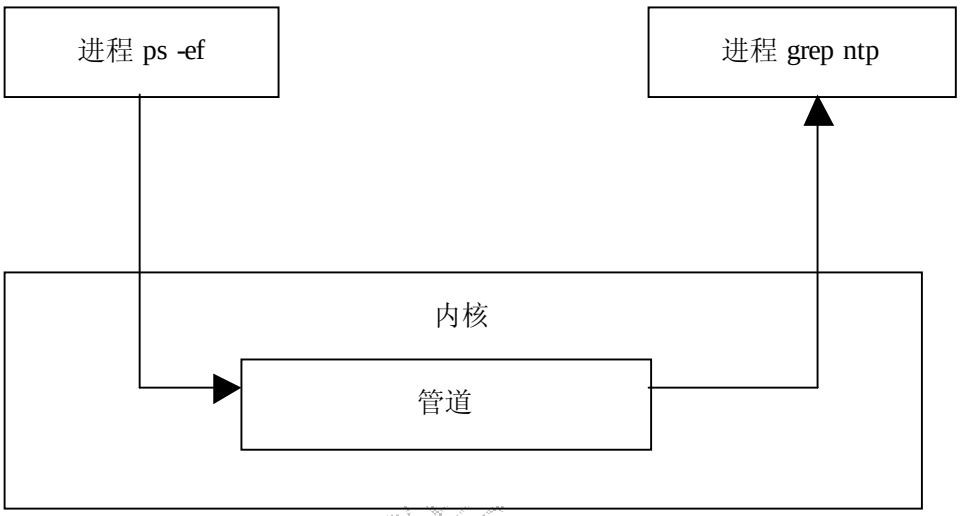


图 8.3 Linux 中管道与文件描述符的关系

管道关闭时只需将这两个文件描述符关闭即可，可使用普通的 close 函数逐个关闭各个文件描述符。

注意 一个管道共享了多对文件描述符时，若将其中的一对读写文件描述符都删除，则该管道就失效。

2．管道创建函数

创建管道可以通过调用 pipe 来实现，下表 8.1 列出了 pipe 函数的语法要点。

表 8.1 pipe 函数语法要点

所需头文件	#include <unistd.h>
函数原型	int pipe(int fd[2])
函数传入值	fd[2]：管道的两个文件描述符，之后就可以直接操作这两个文件描述符
函数返回值	成功：0
	出错：–1

3．管道创建实例

创建管道非常简单，只需调用函数 pipe 即可，如下所示：

```
/*pipe.c*/
#include <unistd.h>
#include <errno.h>
#include <stdio.h>
#include <stdlib.h>

int main()
{
      int pipe_fd[2];
/*创建一无名管道*/
      if(pipe(pipe_fd)<0)
      {
      printf("pipe create error\n");
      return -1;
      }
      else
            printf("pipe create success\n");
/*关闭管道描述符*/
      close(pipe_fd[0]);
      close(pipe_fd[1]);
}
```

程序运行后先成功创建一个无名管道，之后再将其关闭。

8.2.3 管道读写

1．管道读写说明

用 pipe 函数创建的管道两端处于一个进程中，由于管道是主要用于在不同进程间通信的，因此这在实际应用中没有太大意义。实际上，通常先是创建一个管道，再通过 fork()函数创建一子进程，该子进程会继承父进程所创建的管道，这时，父子进程管道的文件描述符对应关系就如图 8.4 所示。

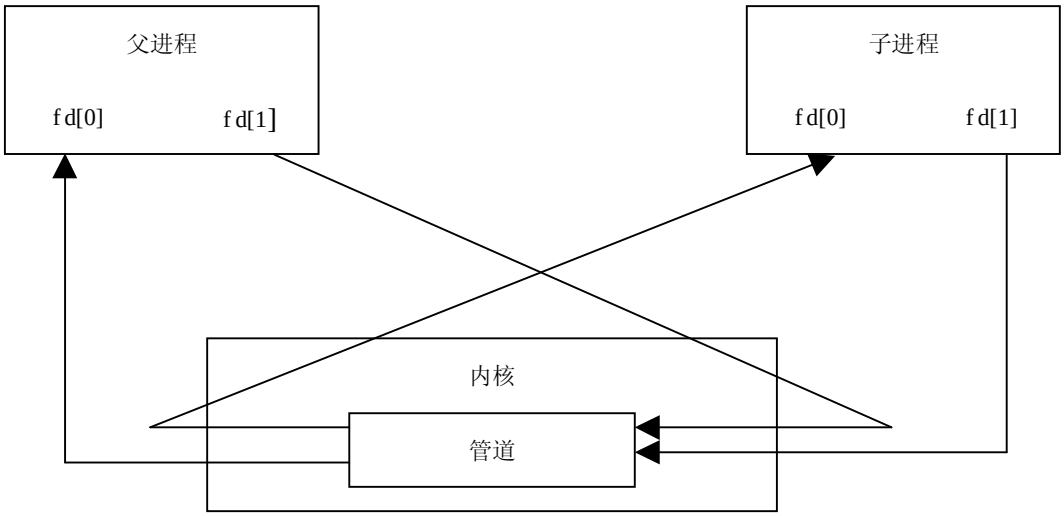


图 8.4　父子进程管道的文件描述符对应关系

这时的关系看似非常复杂，实际上却已经给不同进程之间的读写创造了很好的条件。这时，父子进程分别拥有自己的读写的通道，为了实现父子进程之间的读写，只需把无关的读端或写端的文件描述符关闭即可。例如在图 8.5 中把父进程的写端 fd[1]和子进程的读端 fd[0]关闭。这时，父子进程之间就建立起了一条“子进程写入父进程读”的通道。

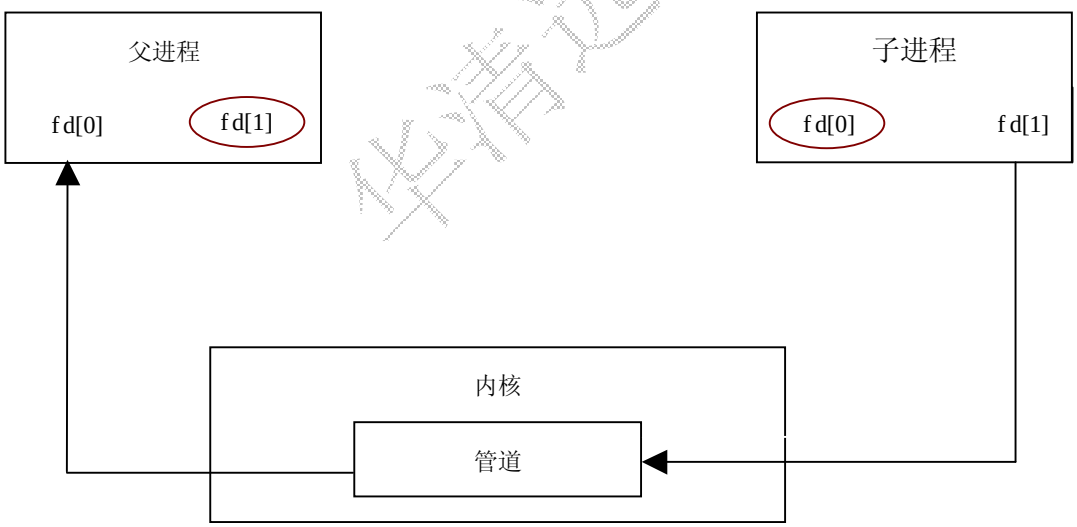


图 8.5　关闭父进程 fd[1]和子进程 fd[0]

同样，也可以关闭父进程的 fd[0]和子进程的 fd[1]，这样就可以建立一条“父进程写，子进程读”的通道。另外，父进程还可以创建多个子进程，各个子进程都继承了相应的 fd[0]和 fd[1]，这时，只需要关闭相应端口就可以建立其各子进程之间的通道。

想一想　为什么无名管道只能建立具有亲缘关系的进程之间?

2．管道读写实例

在本例中，首先创建管道，之后父进程使用 fork 函数创建子进程，之后通过关闭父进程的读描述符和子进程的写描述符，建立起它们之间的管道通信。

```
/*pipe_rw.c*/
#include <unistd.h>
#include <sys/types.h>
#include <errno.h>
#include <stdio.h>
#include <stdlib.h>
int main()
{
      int pipe_fd[2];
      pid_t pid;
      char buf_r[100];
      char* p_wbuf;
      int r_num;
      memset(buf_r,0,sizeof(buf_r));
/*创建管道*/
      if(pipe(pipe_fd)<0)
      {
      printf("pipe create error\n");
      return –1;
      }
/*创建一子进程*/
      if((pid=fork())==0)
      {
            printf("\n");
/*关闭子进程写描述符，并通过使父进程暂停 2 秒确保父进程已关闭相应的读描述符*/
            close(pipe_fd[1]);
            sleep(2);
/*子进程读取管道内容*/
            if((r_num=read(pipe_fd[0],buf_r,100))>0){
                  printf("%d numbers read from the pipe is %s\n",r_num,buf_r);
            }
/*关闭子进程读描述符*/
            close(pipe_fd[0]);
            exit(0);
      }
```

```
        else if(pid>0)
        {
/*/关闭父进程读描述符，并分两次向管道中写入Hello Pipe*/
            close(pipe_fd[0]);
            if(write(pipe_fd[1],"Hello",5)!= -1)
                printf("parent write1 success!\n");
            if(write(pipe_fd[1]," Pipe",5)!= -1)
                printf("parent write2 success!\n");
/*关闭父进程写描述符*/
            close(pipe_fd[1]);
            sleep(3);
/*收集子进程退出信息*/
            waitpid(pid,NULL,0);
            exit(0);
        }
}
```

将该程序交叉编译，下载到开发板上的运行结果如下所示：

```
[root@(none) 1]# ./pipe_rw2
parent write1 success!
parent write2 success!

10 numbers read from the pipe is Hello Pipe
```

3．管道读写注意点

- 只有在管道的读端存在时向管道中写入数据才有意义。否则，向管道中写入数据的进程将收到内核传来的 SIFPIPE 信号（通常 Broken pipe 错误）。
- 向管道中写入数据时，linux 将不保证写入的原子性，管道缓冲区一有空闲区域，写进程就会试图向管道写入数据。如果读进程不读取管道缓冲区中的数据，那么写操作将会一直阻塞。
- 父子进程在运行时，它们的先后次序并不能保证，因此，在这里为了保证父进程已经关闭了读描述符，可在子进程中调用 sleep 函数。

8.2.4 标准流管道

1．标准流管道函数说明

与 Linux 中文件操作有基于文件流的标准 I/O 操作一样，管道的操作也支持基于文件流的模式。这种基于文件流的管道主要是用来创建一个连接到另一个进程的管道，这里的“另一个进程”也就是一个可以进行一定操作的可执行文件，例如，用户执行“cat popen.c”或者自己编写的程序“hello”等。由于这一类操作很常用，因此标准流管道就将一系列的创建

过程合并到一个函数 popen 中完成。它所完成的工作有以下几步。

- 创建一个管道。
- fork 一个子进程。
- 在父子进程中关闭不需要的文件描述符。
- 执行 exec 函数族调用。
- 执行函数中所指定的命令。

这个函数的使用可以大大减少代码的编写量，但同时也有一些不利之处，例如，它没有前面管道创建的函数灵活多样，并且用 popen 创建的管道必须使用标准 I/O 函数进行操作，但不能使用前面的 read、write 一类不带缓冲的 I/O 函数。

与之相对应，关闭用 popen 创建的流管道必须使用函数 pclose 来关闭该管道流。该函数关闭标准 I/O 流，并等待命令执行结束。

2．函数格式

popen 和 pclose 函数格式如表 8.2 和表 8.3 所示。

表 8.2　　popen 函数语法要点

所需头文件	#include <stdio.h>	
函数原型	FILE *popen(const char *command, const char *type)	
函数传入值	Command：指向的是一个以 null 结束符结尾的字符串，这个字符串包含一个 shell 命令，并被送到/bin/sh 以-c 参数执行，即由 shell 来执行	
	type：	“r”：文件指针连接到 command 的标准输出，即该命令的结果产生输出 “w”：文件指针连接到 command 的标准输入，即该命令的结果产生输入
函数返回值	成功：文件流指针	
	出错：–1	

表 8.3　　pclose 函数语法要点

所需头文件	#include <stdio.h>
函数原型	int pclose(FILE *stream)
函数传入值	stream：要关闭的文件流
函数返回值	成功：返回 popen 中执行命令的终止状态
	出错：–1

3．函数使用实例

在该实例中，使用 popen 来执行“ps -ef”命令。可以看出，popen 函数的使用能够使程序变得短小精悍。

```
/*popen.c*/
```

```
#include <stdio.h>
#include <unistd.h>
#include <stdlib.h>
#include <fcntl.h>
#define BUFSIZE 1000
int main()
{
      FILE *fp;
      char *cmd = "ps -ef";
      char buf[BUFSIZE];

/*调用 popen 函数执行相应的命令*/
      if((fp=popen(cmd,"r"))==NULL)
            perror("popen");
      while((fgets(buf,BUFSIZE,fp))!=NULL)
            printf("%s",buf);
      pclose(fp);
      exit(0);
}
```

下面是该程序在目标板上的执行结果。

```
[root@(none) 1]# ./popen
  PID TTY        Uid       Size State Command
    1           root      1832   S   init
    2           root         0   S   [keventd]
    3           root         0   S   [ksoftirqd_CPU0]
    4           root         0   S   [kswapd]
    5           root         0   S   [bdflush]
    6           root         0   S   [kupdated]
    7           root         0   S   [mtdblockd]
    8           root         0   S   [khubd]
   35           root      2104   S   /bin/bash /usr/etc/rc.local
   36           root      2324   S   /bin/bash
   41           root      1364   S   /sbin/inetd
   53           root     14260   S   /Qtopia/qtopia-free-1.7.0/bin/qpe -qws
   54           root     11672   S   quicklauncher
   55           root         0   S   [usb-storage-0]
   56           root         0   S   [scsi_eh_0]
   74           root      1284   S   ./popen
```

```
   75          root      1836   S   sh -c ps -ef
   76          root      2020   R   ps –ef
```

8.2.5 FIFO

1．有名管道说明

前面介绍的管道是无名管道，它只能用于具有亲缘关系的进程之间，这就大大地限制了管道的使用。有名管道的出现突破了这种限制，它可以使互不相关的两个进程实现彼此通信。该管道可以通过路径名来指出，并且在文件系统中是可见的。在建立了管道之后，两个进程就可以把它当作普通文件一样进行读写操作，使用非常方便。不过值得注意的是，FIFO 是严格地遵循先进先出规则的，对管道及 FIFO 的读总是从开始处返回数据，对它们的写则把数据添加到末尾，它们不支持如 lseek()等文件定位操作。

有名管道的创建可以使用函数 mkfifo()，该函数类似文件中的 open()操作，可以指定管道的路径和打开的模式。

小知识 用户还可以在命令行使用“mknod 管道名 p”来创建有名管道。

在创建管道成功之后，就可以使用 open、read、write 这些函数了。与普通文件的开发设置一样，对于为读而打开的管道可在 open 中设置 O_RDONLY，对于为写而打开的管道可在 open 中设置 O_WRONLY，在这里与普通文件不同的是阻塞问题。由于普通文件的读写时不会出现阻塞问题，而在管道的读写中却有阻塞的可能，这里的非阻塞标志可以在 open 函数中设定为 O_NONBLOCK。下面分别对阻塞打开和非阻塞打开的读写进行一定的讨论。

对于读进程

- 若该管道是阻塞打开，且当前 FIFO 内没有数据，则对读进程而言将一直阻塞直到有数据写入。
- 若该管道是非阻塞打开，则不论 FIFO 内是否有数据，读进程都会立即执行读操作。

对于写进程

- 若该管道是阻塞打开，则写进程而言将一直阻塞直到有读进程读出数据。
- 若该管道是非阻塞打开，则当前 FIFO 内没有读操作，写进程都会立即执行读操作。

2．mkfifo 函数格式

表 8.4 列出了 mkfifo 函数的语法要点。

表 8.4 mkfifo 函数语法要点

所需头文件	#include <sys/types.h> #include <sys/state.h>
函数原型	int mkfifo(const char *filename,mode_t mode)
函数传入值	filename：要创建的管道

续表

函数传入值	mode：	O_RDONLY：读管道
		O_WRONLY：写管道
		O_RDWR：读写管道
		O_NONBLOCK：非阻塞
		O_CREAT：如果该文件不存在，那么就创建一个新的文件，并用第三的参数为其设置权限
		O_EXCL：如果使用 O_CREAT 时文件存在，那么可返回错误消息。这一参数可测试文件是否存在
函数返回值	成功：0	
	出错：−1	

表 8.5 再对 FIFO 相关的出错信息做一归纳，以方便用户差错。

表 8.5　　FIFO 相关的出错信息

EACCESS	参数 filename 所指定的目录路径无可执行的权限
EEXIST	参数 filename 所指定的文件已存在
ENAMETOOLONG	参数 filename 的路径名称太长
ENOENT	参数 filename 包含的目录不存在
ENOSPC	文件系统的剩余空间不足
ENOTDIR	参数 filename 路径中的目录存在但却非真正的目录
EROFS	参数 filename 指定的文件存在于只读文件系统内

3．使用实例

下面的实例包含了两个程序，一个用于读管道，另一个用于写管道。其中在写管道的程序里创建管道，并且作为 main 函数里的参数由用户输入要写入的内容。读管道读出了用户写入管道的内容，这两个函数用的是非阻塞读写管道。

```
/*fifo_write.c*/
#include <sys/types.h>
#include <sys/stat.h>
#include <errno.h>
#include <fcntl.h>
#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#define FIFO "/tmp/myfifo"

main(int argc,char** argv)
```

```
/*参数为即将写入的字节数*/
{
      int fd;
      char w_buf[100];
      int nwrite;
      if(fd== -1)
            if(errno==ENXIO)
                  printf("open error; no reading process\n");
/*打开 FIFO 管道，并设置非阻塞标志*/
      fd=open(FIFO_SERVER,O_WRONLY|O_NONBLOCK,0);
      if(argc==1)
            printf("Please send something\n");
      strcpy(w_buf,argv[1]);
/*向管道中写入字符串*/
      if((nwrite=write(fd,w_buf,100))== -1)
      {
            if(errno==EAGAIN)
                  printf("The FIFO has not been read yet.Please try later\n");
      }
      else
            printf("write %s to the FIFO\n",w_buf);
}

/*fifl_read.c*/
#include <sys/types.h>
#include <sys/stat.h>
#include <errno.h>
#include <fcntl.h>
#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#define FIFO "/tmp/myfifo"

main(int argc,char** argv)
{
      char buf_r[100];
      int  fd;
    int  nread;
/*创建有名管道，并设置相应的权限*/
```

```
        if((mkfifo(FIFO,O_CREAT|O_EXCL)<0)&&(errno!=EEXIST))
              printf("cannot create fifoserver\n");
        printf("Preparing for reading bytes...\n");
        memset(buf_r,0,sizeof(buf_r));
    /*打开有名管道，并设置非阻塞标志*/
        fd=open(FIFO,O_RDONLY|O_NONBLOCK,0);
        if(fd== -1)
        {
              perror("open");
              exit(1);
        }
        while(1)
        {
              memset(buf_r,0,sizeof(buf_r));
              if((nread=read(fd,buf_r,100))== -1){
                    if(errno==EAGAIN)
                          printf("no data yet\n");
              }
              printf("read %s from FIFO\n",buf_r);
              sleep(1);
        }
        pause();
        unlink(FIFO);
    }
```

为了能够较好地观察运行结果，需要把这两个程序分别在两个终端里运行，在这里首先启动读管道程序。由于这是非阻塞管道，因此在建立管道之后程序就开始循环从管道里读出内容。在启动了写管道程序后，读进程能够从管道里读出用户的输入内容，程序运行结果如下所示。

终端一:

```
[root@localhost FIFO]# ./read
Preparing for reading bytes…
read   from FIFO
read   from FIFO
read   from FIFO
read   from FIFO
read   from FIFO
read hello from FIFO
read   from FIFO
```

```
read   from FIFO
read FIFO from FIFO
read   from FIFO
read   from FIFO
…
```

终端二：

```
[root@localhost]# ./write hello
write hello to the FIFO
[root@localhost]# ./read FIFO
write FIFO to the FIFO
```

8.3 信号通信

8.3.1 信号概述

信号是 UNIX 中所使用的进程通信的一种最古老的方法。它是在软件层次上对中断机制的一种模拟，是一种异步通信方式。信号可以直接进行用户空间进程和内核进程之间的交互，内核进程也可以利用它来通知用户空间进程发生了哪些系统事件。它可以在任何时候发给某一进程，而无需知道该进程的状态。如果该进程当前并未处于执行态，则该信号就由内核保存起来，直到该进程恢复执行再传递给它为止；如果一个信号被进程设置为阻塞，则该信号的传递被延迟，直到其阻塞被取消时才被传递给进程。

细心的读者是否还记得，在第 2 章 kill 命令中曾讲解到“–l”选项，这个选项可以列出该系统所支持的所有信号列表。在笔者的系统中，信号值在 32 之前的则有不同的名称，而信号值在 32 以后的都是用“SIGRTMIN”或“SIGRTMAX”开头的，这就是两类典型的信号。前者是从 UNIX 系统中继承下来的信号，为不可靠信号（也称为非实时信号）；后者是为了解决前面“不可靠信号”的问题而进行了更改和扩充的信号，称为“可靠信号”（也称为实时信号）。那么为什么之前的信号不可靠呢？这里首先要介绍一下信号的生命周期。

一个完整的信号生命周期可以分为 3 个重要阶段，这 3 个阶段由 4 个重要事件来刻画的：信号产生、信号在进程中注册、信号在进程中注销、执行信号处理函数，如图 8.6 所示。相邻两个事件的时间间隔构成信号生命周期的一个阶段。要注意这里的信号处理有多种方式，一般是由内核完成的，当然也可以由用户进程来完成，故在此没有明确画出。

一个不可靠信号的处理过程是这样的：如果发现该信号已经在进程中注册，那么就忽略该信号。因此，若前一个信号还未注销又产生了相同的信号就会产生信号丢失。而当可靠信号发送给一个进程时，不管该信号是否已经在进程中注册，都会被再注册一次，因此信号就不会丢失。所有可靠信号都支持排队，而不可靠信号则都不支持排队。

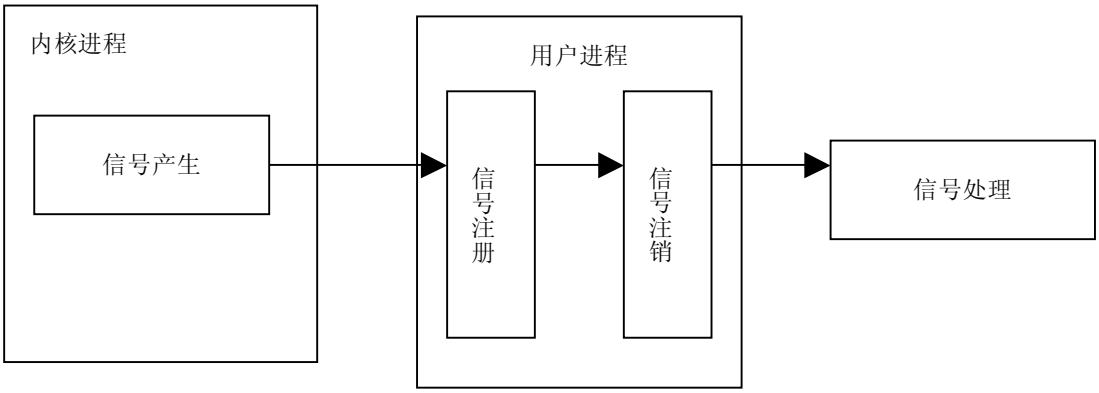


图 8.6 信号生命周期

> **注意** 这里信号的产生、注册、注销等是指信号的内部实现机制，而不是信号的函数实现。因此，信号注册与否，与本节后面讲到的发送信号函数（如 kill()等）以及信号安装函数（如 signal()等）无关，只与信号值有关。

用户进程对信号的响应可以有 3 种方式。

- 忽略信号，即对信号不做任何处理，但是有两个信号不能忽略，即 SIGKILL 及 SIGSTOP。
- 捕捉信号，定义信号处理函数，当信号发生时，执行相应的处理函数。
- 执行缺省操作，Linux 对每种信号都规定了默认操作。

Linux 中的大多数信号是提供给内核的，表 8.6 列出了 Linux 中最为常见信号的含义及其默认操作。

表 8.6 常见信号的含义及其默认操作

信 号 名	含 义	默 认 操 作
SIGHUP	该信号在用户终端连接（正常或非正常）结束时发出，通常是在终端的控制进程结束时，通知同一会话内的各个作业与控制终端不再关联	终止
SIGINT	该信号在用户键入 INTR 字符（通常是 Ctrl-C）时发出，终端驱动程序发送此信号并送到前台进程中的每一个进程	终止
SIGQUIT	该信号和 SIGINT 类似，但由 QUIT 字符（通常是 Ctrl-\）来控制	终止
SIGILL	该信号在一个进程企图执行一条非法指令时（可执行文件本身出现错误，或者试图执行数据段、堆栈溢出时）发出	终止
SIGFPE	该信号在发生致命的算术运算错误时发出。这里不仅包括浮点运算错误，还包括溢出及除数为 0 等其他所有的算术的错误	终止
SIGKILL	该信号用来立即结束程序的运行，并且不能被阻塞、处理和忽略	终止

SIGALRM	该信号当一个定时器到时的时候发出	终止
SIGSTOP	该信号用于暂停一个进程，且不能被阻塞、处理或忽略	暂停进程

续表

信 号 名	含 义	默 认 操 作
SIGTSTP	该信号用于交互停止进程，用户可键入 SUSP 字符时（通常是 Ctrl+Z）发出这个信号	停止进程
SIGCHLD	子进程改变状态时，父进程会收到这个信号	忽略
SIGABORT		

8.3.2 信号发送与捕捉

发送信号的函数主要有 kill()、raise()、alarm()以及 pause()，下面就依次对其进行介绍。

1．kill()和 raise()

（1）函数说明

kill 函数同读者熟知的 kill 系统命令一样，可以发送信号给进程或进程组（实际上，kill 系统命令只是 kill 函数的一个用户接口）。这里要注意的是，它不仅可以中止进程（实际上发出 SIGKILL 信号），也可以向进程发送其他信号。

与 kill 函数所不同的是，raise 函数允许进程向自身发送信号。

（2）函数格式

表 8.7 列出了 kill 函数的语法要点。

表 8.7 kill 函数语法要点

所需头文件	#include <signal.h> #include <sys/types.h>	
函数原型	int kill(pid_t pid,int sig)	
函数传入值	pid：	正数：要发送信号的进程号
		0：信号被发送到所有和 pid 进程在同一个进程组的进程
		–1：信号发给所有的进程表中的进程（除了进程号最大的进程外）
	sig：信号	
函数返回值	成功：0	
	出错：–1	

表 8.8 列出了 raise 函数的语法要点。

表 8.8 raise 函数语法要点

所需头文件	#include <signal.h> #include <sys/types.h>

函数原型	int raise(int sig)
函数传入值	sig：信号
函数返回值	成功：0
	出错：–1

（3）函数实例

下面这个示例首先使用 fork 创建了一个子进程，接着为了保证子进程不在父进程调用 kill 之前退出，在子进程中使用 raise 函数向子进程发送 SIGSTOP 信号，使子进程暂停。接下来再在父进程中调用 kill 向子进程发送信号，在该示例中使用的是 SIGKILL，读者可以使用其他信号进行练习。

```
/*kill.c*/
#include <stdio.h>
#include <stdlib.h>
#include <signal.h>
#include <sys/types.h>
#include <sys/wait.h>

int main()
{
      pid_t pid;
      int ret;
/*创建一子进程*/
      if((pid=fork())<0){
            perror("fork");
            exit(1);
      }
      if(pid == 0){
/*在子进程中使用 raise 函数发出 SIGSTOP 信号*/
            raise(SIGSTOP);
            exit(0);
      }
      else{
/*在父进程中收集子进程发出的信号，并调用 kill 函数进行相应的操作*/
            printf("pid=%d\n",pid);
            if((waitpid(pid,NULL,WNOHANG))==0){
                  if((ret=kill(pid,SIGKILL))==0)
                        printf("kill %d\n",pid);
                  else{
```

```
                    perror("kill");
                }
            }
        }
    }
```

该程序运行结果如下所示：

```
[root@(none) tmp]# ./kill
pid=78
kill 78
```

2．alarm()和 pause()

（1）函数说明

alarm 也称为闹钟函数，它可以在进程中设置一个定时器，当定时器指定的时间到时，它就向进程发送 SIGALARM 信号。要注意的是，一个进程只能有一个闹钟时间，如果在调用 alarm 之前已设置过闹钟时间，则任何以前的闹钟时间都被新值所代替。

pause 函数是用于将调用进程挂起直至捕捉到信号为止。这个函数很常用，通常可以用于判断信号是否已到。

（2）函数格式

表 8.9 列出了 alarm 函数的语法要点。

表 8.9　　alarm 函数语法要点

所需头文件	#include <unistd.h>
函数原型	unsigned int alarm(unsigned int seconds)
函数传入值	seconds：指定秒数
函数返回值	成功：如果调用此 alarm()前，进程中已经设置了闹钟时间，则返回上一个闹钟时间的剩余时间，否则返回 0
	出错：–1

表 8.10 列出了 pause 函数的语法要点。

表 8.10　　pause 函数语法要点

所需头文件	#include <unistd.h>
函数原型	int pause(void)
函数返回值	–1，并且把 error 值设为 EINTR

（3）函数实例

该实例实际上已完成了一个简单的 sleep 函数的功能，由于 SIGALARM 默认的系统动作为终止该进程，因此在程序调用 pause 之后，程序就终止了。如下所示：

```
/*alarm.c*/
```

```
#include <unistd.h>
#include <stdio.h>
#include <stdlib.h>

int main()
{
      int ret;
/*调用 alarm 定时器函数*/
      ret=alarm(5);
      pause();
      printf("I have been waken up.\n",ret);
}
[root@(none) tmp]#./alarm
Alarm clock
```

想一想 用这种形式实现的 sleep 功能有什么问题？

8.3.3 信号的处理

在了解了信号的产生与捕获之后，接下来就要对信号进行具体的操作了。从前面的信号概述中读者也可以看到，特定的信号是与一定的进程相联系的。也就是说，一个进程可以决定在该进程中需要对哪些信号进行什么样的处理。例如，一个进程可以选择忽略某些信号而只处理其他一些信号，另外，一个进程还可以选择如何处理信号。总之，这些都是与特定的进程相联系的。因此，首先就要建立其信号与进程之间的对应关系，这就是信号的处理。

注意 请读者区分信号的注册与信号的处理之间的差别，前者信号是主动方，而后者进程是主动方。信号的注册是在进程选择了特定信号处理之后特定信号的主动行为。

信号处理的主要方法有两种，一种是使用简单的 signal 函数，另一种是使用信号集函数组。下面分别介绍这两种处理方式。

1．signal()

（1）函数说明

使用 signal 函数处理时，只需把要处理的信号和处理函数列出即可。它主要是用于前 32 种非实时信号的处理，不支持信号传递信息，但是由于使用简单、易于理解，因此也受到很多程序员的欢迎。

（2）函数格式

Signal 函数的语法要点如表 8.11 所示。

表 8.11 signal 函数语法要点

所需头文件	#include <signal.h>	
函数原型	void (*signal(int signum, void (*handler)(int)))(int)	
函数传入值	signum：指定信号	
	handler：	SIG_IGN：忽略该信号
		SIG_DFL：采用系统默认方式处理信号
		自定义的信号处理函数指针

续表

函数返回值	成功：以前的信号处理配置
	出错：−1

这里需要对这个函数原型进行说明。这个函数原型非常复杂。可先用如下的 typedef 进行替换说明：

```
typedef void sign(int);
sign *signal(int, handler *);
```

可见，首先该函数原型整体指向一个无返回值带一个整型参数的函数指针，也就是信号的原始配置函数。接着该原型又带有两个参数，其中的第二个参数可以是用户自定义的信号处理函数的函数指针。

（3）使用实例

该示例表明了如何使用 signal 函数捕捉相应信号，并做出给定的处理。这里，my_func 就是信号处理的函数指针。读者还可以将其改为 SIG_IGN 或 SIG_DFL 查看运行结果。

```
/*mysignal.c*/
#include <signal.h>
#include <stdio.h>
#include <stdlib.h>
/*自定义信号处理函数*/
void my_func(int sign_no)
{
      if(sign_no==SIGINT)
            printf("I have get SIGINT\n");
    else if(sign_no==SIGQUIT)
            printf("I have get SIGQUIT\n");
}
int main()
{
      printf("Waiting for signal SIGINT or SIGQUIT \n ");
/*发出相应的信号，并跳转到信号处理函数处*/
      signal(SIGINT, my_func);
```

```
        signal(SIGQUIT, my_func);
        pause();
        exit(0);
}
[root@www yul]# ./mysignal
Waiting for signal SIGINT or SIGQUIT
I have get SIGINT
[root@www yul]# ./mysignal
Waiting for signal SIGINT or SIGQUIT
I have get SIGQUIT
```

2．信号集函数组

（1）函数说明

使用信号集函数组处理信号时涉及一系列的函数，这些函数按照调用的先后次序可分为以下几大功能模块：创建信号集合、登记信号处理器以及检测信号。

其中，创建信号集合主要用于创建用户感兴趣的信号，其函数包括以下几个。

- sigemptyset：初始化信号集合为空。
- sigfillset：初始化信号集合为所有信号的集合。
- sigaddset：将指定信号加入到信号集合中去。
- sigdelset：将指定信号从信号集中删去。
- sigismember：查询指定信号是否在信号集合之中。

登记信号处理器主要用于决定进程如何处理信号。这里要注意的是，信号集里的信号并不是真正可以处理的信号，只有当信号的状态处于非阻塞状态时才真正起作用。因此，首先就要判断出当前阻塞能不能传递给该信号的信号集。这里首先使用 sigprocmask 函数判断检测或更改信号屏蔽字，然后使用 sigaction 函数用于改变进程接收到特定信号之后的行为。

检测信号是信号处理的后续步骤，但不是必须的。由于内核可以在任何时刻向某一进程发出信号，因此，若该进程必须保持非中断状态且希望将某些信号阻塞，这些信号就处于“未决”状态（也就是进程不清楚它的存在）。所以，在希望保持非中断进程完成相应的任务之后，就应该将这些信号解除阻塞。Sigpending 函数就允许进程检测“未决”信号，并进一步决定对它们作何处理。

（2）函数格式

首先介绍创建信号集合的函数格式，表 8.12 列举了这一组函数的语法要点。

表 8.12　　创建信号集合函数语法要点

所需头文件	#include <signal.h>
函数原型	int sigemptyset(sigset_t *set)
	int sigfillset(sigset_t *set)

	int sigaddset(sigset_t *set,int signum)
	int sigdelset(sigset_t *set,int signum)
	int sigismember(sigset_t *set,int signum)
函数传入值	set：信号集
	signum：指定信号值
函数返回值	成功：0（sigismember 成功返回 1，失败返回 0）
	出错：−1

表 8.13 列举了 sigprocmask 的语法要点。

表 8.13　　sigprocmask 函数语法要点

所需头文件	#include <signal.h>	
函数原型	int sigprocmask(int how,const sigset_t *set,sigset_t *oset)	
函数传入值	how：决定函数的操作方式	SIG_BLOCK：增加一个信号集合到当前进程的阻塞集合之中
		SIG_UNBLOCK：从当前的阻塞集合之中删除一个信号集合
		SIG_SETMASK：将当前的信号集合设置为信号阻塞集合
	set：指定信号集	
	oset：信号屏蔽字	
函数返回值	成功：0（sigismember 成功返回 1，失败返回 0）	
	出错：−1	

此处，若 set 是一个非空指针，则参数 how 表示函数的操作方式；若 how 为空，则表示忽略此操作。

表 8.14 列举了 sigaction 的语法要点。

表 8.14　　sigaction 函数语法要点

所需头文件	#include <signal.h>
函数原型	int sigaction(int signum,const struct sigaction *act,struct sigaction *oldact)
函数传入值	signum：信号的值，可以为除 SIGKILL 及 SIGSTOP 外的任何一个特定有效的信号
	act：指向结构 sigaction 的一个实例的指针，指定对特定信号的处理
	oldact：保存原来对相应信号的处理
函数返回值	成功：0
	出错：−1

这里要说明的是 sigaction 函数中第 2 个和第 3 个参数用到的 sigaction 结构。这是一个看似非常复杂的结构，希望读者能够慢慢阅读此段内容。

首先给出了 sigaction 的定义，如下所示：

```
struct sigaction {
```

```
    void (*sa_handler)(int signo);
    sigset_t sa_mask;
    int sa_flags;
    void (*sa_restore)(void);
    }
```

sa_handler 是一个函数指针，指定信号关联函数，这里除可以是用户自定义的处理函数外，还可以为 SIG_DFL（采用缺省的处理方式）或 SIG_IGN（忽略信号）。它的处理函数只有一个参数，即信号值。

sa_mask 是一个信号集，它可以指定在信号处理程序执行过程中哪些信号应当被阻塞，在调用信号捕获函数之前，该信号集要加入到信号的信号屏蔽字中。

sa_flags 中包含了许多标志位，是对信号进行处理的各个选择项。它的常见可选值如下表 8.15 所示。

表 8.15 常见信号的含义及其默认操作

选　　项	含　　义
SA_NODEFER\SA_NOMASK	当捕捉到此信号时，在执行其信号捕捉函数时，系统不会自动阻塞此信号
SA_NOCLDSTOP	进程忽略子进程产生的任何 SIGSTOP、SIGTSTP、SIGTTIN 和 SIGTTOU 信号
SA_RESTART	可让重启的系统调用重新起作用
SA_ONESHOT\SA_RESETHAND	自定义信号只执行一次，在执行完毕后恢复信号的系统默认动作

最后，表 8.16 列举了 sigpending 函数的语法要点。

表 8.16 sigpending 函数语法要点

所需头文件	#include <signal.h>
函数原型	int sigpending(sigset_t *set)
函数传入值	set：要检测的信号集
函数返回值	成功：0
	出错：−1

总之，在处理信号时，一般遵循如图 8.7 所示的操作流程。

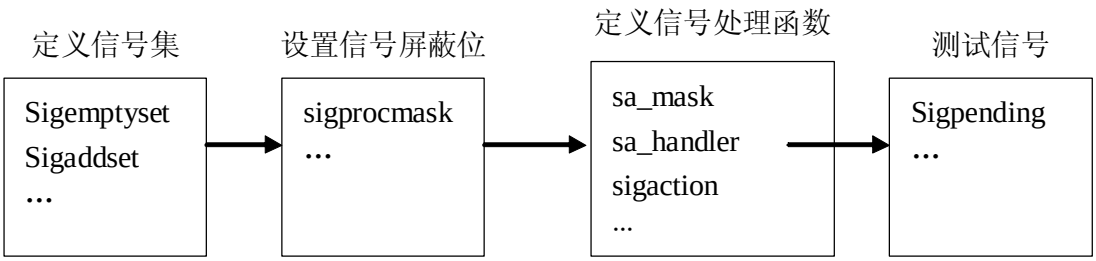


图 8.7　信号操作一般处理流程

（3）使用实例

该实例首先把 SIGQUIT、SIGINT 两个信号加入信号集，然后将该信号集设为阻塞状态，并在该状态下使程序暂停 5 秒。接下来再将信号集设置为非阻塞状态，再对这两个信号分别操作，其中 SIGQUIT 执行默认操作，而 SIGINT 执行用户自定义函数的操作。源代码如下所示：

```
/*sigaction.c*/
#include <sys/types.h>
#include <unistd.h>
#include <signal.h>
#include <stdio.h>
#include <stdlib.h>

/*自定义的信号处理函数*/
void my_func(int signum)
{
      printf("If you want to quit,please try SIGQUIT\n");
}
int main()
{
      sigset_t set,pendset;
      struct sigaction action1,action2;
/*初始化信号集为空*/
      if(sigemptyset(&set)<0)
            perror("sigemptyset");
/*将相应的信号加入信号集*/
      if(sigaddset(&set,SIGQUIT)<0)
            perror("sigaddset");
      if(sigaddset(&set,SIGINT)<0)
            perror("sigaddset");
/*设置信号集屏蔽字*/
      if(sigprocmask(SIG_BLOCK,&set,NULL)<0)
            perror("sigprocmask");
      else
      {
            printf("blocked\n");
            sleep(5);
      }
```

```
        if(sigprocmask(SIG_UNBLOCK,&set,NULL)<0)
            perror("sigprocmask");
        else
            printf("unblock\n");
    /*对相应的信号进行循环处理*/
        while(1){
            if(sigismember(&set,SIGINT)){
                sigemptyset(&action1.sa_mask);
                action1.sa_handler=my_func;
                sigaction(SIGINT,&action1,NULL);
            }else if(sigismember(&set,SIGQUIT)){
                sigemptyset(&action2.sa_mask);
                action2.sa_handler = SIG_DFL;
                sigaction(SIGTERM,&action2,NULL);
            }
        }
    }
```

该程序的运行结果如下所示，可以看见，在信号处于阻塞状态时，所发出的信号对进程不起作用。读者需等待 5 秒，在信号接触阻塞状态之后，用户发出的信号才能正常运行。这里 SIGINT 已按照用户自定义的函数运行。

```
[root@(none) tmp]# ./sigaction
blocked
unblock
If you want to quit,please try SIGQUIT
Quit
```

8.4 共享内存

8.4.1 共享内存概述

可以说，共享内存是一种最为高效的进程间通信方式。因为进程可以直接读写内存，不需要任何数据的拷贝。为了在多个进程间交换信息，内核专门留出了一块内存区。这段内存区可以由需要访问的进程将其映射到自己的私有地址空间。因此，进程就可以直接读写这一内存区而不需要进行数据的拷贝，从而大大提高了效率。当然，由于多个进程共享一段内存，因此也需要依靠某种同步机制，如互斥锁和信号量等。其原理示意图如图 8.8 所示。

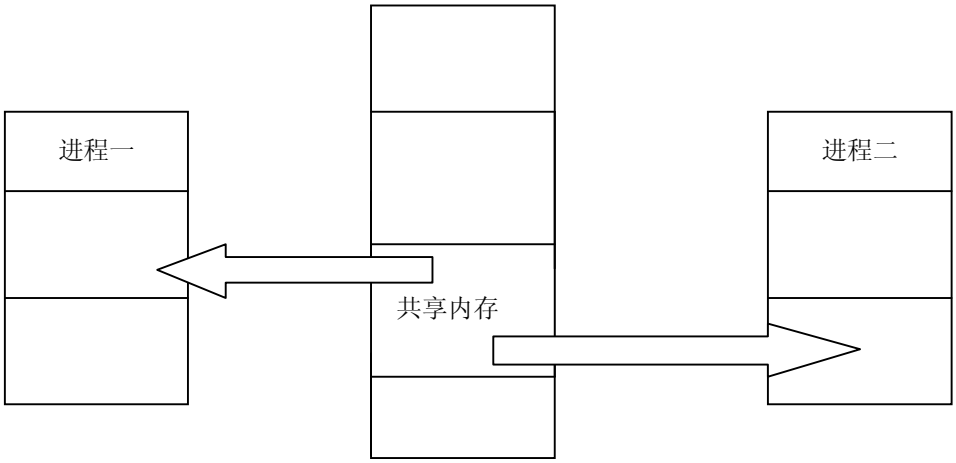


图 8.8　共享内存原理示意图

8.4.2　共享内存实现

1．函数说明

共享内存的实现分为两个步骤，第一步是创建共享内存，这里用到的函数是 shmget，也就是从内存中获得一段共享内存区域。第二步映射共享内存，也就是把这段创建的共享内存映射到具体的进程空间去，这里使用的函数是 shmat。到这里，就可以使用这段共享内存了，也就是可以使用不带缓冲的 I/O 读写命令对其进行操作。除此之外，当然还有撤销映射的操作，其函数为 shmdt。这里就主要介绍这 3 个函数。

2．函数格式

表 8.17 列举了 shmget 函数的语法要点。

表 8.17　　shmget 函数语法要点

所需头文件	#include <sys/types.h> #include <sys/ipc.h> #include <sys/shm.h>
函数原型	int shmget(key_t key,int size,int shmflg)
函数传入值	Key：IPC_PRIVATE
	Size：共享内存区大小
	Shmflg：同 open 函数的权限位，也可以用八进制表示法
函数返回值	成功：共享内存段标识符
	出错：−1

表 8.18 列举了 shmat 函数的语法要点。

表 8.18　　shmat 函数语法要点

所需头文件	#include <sys/types.h> #include <sys/ipc.h> #include <sys/shm.h>	
函数原型	char *shmat(int shmid,const void *shmaddr,int shmflg)	
函数传入值	shmid：要映射的共享内存区标识符	
	shmaddr：将共享内存映射到指定位置（若为 0 则表示把该段共享内存映射到调用进程的地址空间）	
	Shmflg	SHM_RDONLY：共享内存只读
		默认 0：共享内存可读写
函数返回值	成功：被映射的段地址	
	出错：–1	

表 8.19 列举了 shmdt 函数的语法要点。

表 8.19 shmdt 函数语法要点

所需头文件	#include <sys/types.h> #include <sys/ipc.h> #include <sys/shm.h>
函数原型	int shmdt(const void *shmaddr)
函数传入值	Shmaddr：被映射的共享内存段地址
函数返回值	成功：0
	出错：–1

3．使用实例

该实例说明了如何使用基本的共享内存函数，首先是创建一个共享内存区，之后将其映射到本进程中，最后再解除这种映射关系。这里要介绍的一个命令是 ipcs，这是用于报告进程间通信机制状态的命令。它可以查看共享内存、消息队列等各种进程间通信机制的情况，这里使用了 system 函数用于调用 shell 命令“ipcs”。程序源代码如下所示：

```
/*shmadd.c*/
#include <sys/types.h>
#include <sys/ipc.h>
#include <sys/shm.h>
#include <stdio.h>
#include <stdlib.h>

#define BUFSZ 2048
int main()
{
      int shmid;
      char *shmadd;
```

```
/*创建共享内存*/
      if((shmid=shmget(IPC_PRIVATE,BUFSZ,0666))<0){
            perror("shmget");
            exit(1);
      }
      else
            printf("created shared-memory: %d\n",shmid);
            system("ipcs -m");
/*映射共享内存*/
      if((shmadd=shmat(shmid,0,0))<(char *)0){
            perror("shmat");
            exit(1);
      }
      else
            printf("attached shared-memory\n");
/*显示系统内存情况*/
    system("ipcs -m");
/*删除共享内存*/
      if((shmdt(shmadd))<0){
            perror("shmdt");
            exit(1);
      }
      else
            printf("deleted shared-memory\n");
      system("ipcs -m");
      exit(0);
}
```

下面是运行结果。从该结果可以看出，nattch 的值随着共享内存状态的变化而变化，共享内存的值根据不同的系统会有所不同。

```
created shared-memory: 229383

------ Shared Memory Segments --------
key        shmid      owner     perms     bytes     nattch    status
0x00000000 229383     root      666       2048      0

attached shared-memory

------ Shared Memory Segments --------
```

```
key        shmid      owner      perms      bytes      nattch     status
0x00000000 229383     root       666        2048       1

deleted shared-memory

------ Shared Memory Segments --------
key        shmid      owner      perms      bytes      nattch     status
0x00000000 229383     root       666        2048       0
```

8.5 消息队列

8.5.1 消息队列概述

顾名思义，消息队列就是一个消息的列表。用户可以从消息队列种添加消息、读取消息等。从这点上看，消息队列具有一定的 FIFO 的特性，但是它可以实现消息的随机查询，比 FIFO 具有更大的优势。同时，这些消息又是存在于内核中的，由“队列 ID”来标识。

8.5.2 消息队列实现

1．函数说明

消息队列的实现包括创建或打开消息队列、添加消息、读取消息和控制消息队列这四种操作。其中创建或打开消息队列使用的函数是 msgget，这里创建的消息队列的数量会受到系统消息队列数量的限制；添加消息使用的函数是 msgsnd 函数，它把消息添加到已打开的消息队列末尾；读取消息使用的函数是 msgrcv，它把消息从消息队列中取走，与 FIFO 不同的是，这里可以指定取走某一条消息；最后控制消息队列使用的函数是 msgctl，它可以完成多项功能。

2．函数格式

表 8.20 列举了 msgget 函数的语法要点。

表 8.20　　msgget 函数语法要点

所需头文件	#include <sys/types.h> #include <sys/ipc.h> #include <sys/shm.h>
函数原型	int msgget(key_t key,int flag)
函数传入值	Key：返回新的或已有队列的队列 ID,IPC_PRIVATE
	Flag：
函数返回值	成功：消息队列 ID
	出错：–1

表 8.21 列举了 msgsnd 函数的语法要点。

表 8.21　　msgsnd 函数语法要点

所需头文件	#include <sys/types.h> #include <sys/ipc.h> #include <sys/shm.h>	
函数原型	int msgsnd(int msqid,const void *prt,size_t size,int flag)	
函数传入值	msqid：消息队列的队列 ID	

	prt：指向消息结构的指针。该消息结构 msgbuf 为：
	struct msgbuf{
	long mtype;//消息类型
	char mtext[1];//消息正文
	}	

	size：消息的字节数，不要以 null 结尾	
	flag：	IPC_NOWAIT 若消息并没有立即发送而调用进程会立即返回
		0：msgsnd 调用阻塞直到条件满足为止
函数返回值	成功：0	
	出错：−1	

表 8.22 列举了 msgrcv 函数的语法要点。

表 8.22　　msgrcv 函数语法要点

所需头文件	#include <sys/types.h> #include <sys/ipc.h> #include <sys/shm.h>	
函数原型	int msgrcv(int msgid,struct msgbuf *msgp,int size,long msgtype,int flag)	
函数传入值	msqid：消息队列的队列 ID	
	msgp：消息缓冲区	
	size：消息的字节数，不要以 null 结尾	
	Msgtype：	0：接收消息队列中第一个消息
		大于 0：接收消息队列中第一个类型为 msgtyp 的消息
		小于 0：接收消息队列中第一个类型值不小于 msgtyp 绝对值且类型值又最小的消息
	flag：	MSG_NOERROR：若返回的消息比 size 字节多，则消息就会截短到 size 字节，且不通知消息发送进程
		IPC_NOWAIT 若消息并没有立即发送而调用进程会立即返回
		0：msgsnd 调用阻塞直到条件满足为止
函数返回值	成功：0	
	出错：−1	

表 8.23 列举了 msgrcv 函数的语法要点。

表 8.23　　msgrcv 函数语法要点

所需头文件	#include <sys/types.h> #include <sys/ipc.h> #include <sys/shm.h>

函数原型	int msgrcv(int msgid,struct msgbuf *msgp,int size,long msgtype,int flag)	
函数传入值	msqid：消息队列的队列 ID	
	msgp：消息缓冲区	
	size：消息的字节数，不要以 null 结尾	
	Msgtype：	0：接收消息队列中第一个消息
		大于 0：接收消息队列中第一个类型为 msgtyp 的消息
		小于 0：接收消息队列中第一个类型值不小于 msgtyp 绝对值且类型值又最小的消息
	flag：	MSG_NOERROR：若返回的消息比 size 字节多，则消息就会截短到 size 字节，且不通知消息发送进程
		IPC_NOWAIT 若消息并没有立即发送而调用进程会立即返回
		0：msgsnd 调用阻塞直到条件满足为止
函数返回值	成功：0	
	出错：−1	

表 8.24 列举了 msgctl 函数的语法要点。

表 8.24　　msgctl 函数语法要点

所需头文件	#include <sys/types.h> #include <sys/ipc.h> #include <sys/shm.h>	
函数原型	int msgctl (int msgqid, int cmd, struct msqid_ds *buf)	
函数传入值	msqid：消息队列的队列 ID	
	cmd：	IPC_STAT：读取消息队列的数据结构 msqid_ds，并将其存储在 buf 指定的地址中
		IPC_SET：设置消息队列的数据结构 msqid_ds 中的 ipc_perm 元素的值。这个值取自 buf 参数
		IPC_RMID：从系统内核中移走消息队列
	Buf：消息队列缓冲区	
函数返回值	成功：0	
	出错：−1	

3．使用实例

这个实例体现了如何使用消息队列进行进程间通信，包括消息队列的创建、消息发送与读取、消息队列的撤销等多种操作。注意这里使用了函数 fotk，它可以根据不同的路径和关键表示产生标准的 key。程序源代码如下所示：

```
/*msg.c*/
#include <sys/types.h>
```

```
#include <sys/ipc.h>
#include <sys/msg.h>
#include <stdio.h>
#include <stdlib.h>
#include <unistd.h>
#include <string.h>
#define BUFSZ 512
struct message{
      long msg_type;
      char msg_text[BUFSZ];
};

int main()
{
      int qid;
      key_t key;
      int len;
      struct message msg;
/*根据不同的路径和关键表示产生标准的 key*/
      if((key=ftok(".",'a'))== -1){
            perror("ftok");
            exit(1);
      }
/*创建消息队列*/
      if((qid=msgget(key,IPC_CREAT|0666))== -1){
            perror("msgget");
            exit(1);
      }
      printf("opened queue %d\n",qid);
      puts("Please enter the message to queue:");
      if((fgets((&msg)->msg_text,BUFSZ,stdin))==NULL){
            puts("no message");
            exit(1);
      }
      msg.msg_type = getpid();
      len = strlen(msg.msg_text);
/*添加消息到消息队列*/
      if((msgsnd(qid,&msg,len,0))<0){
            perror("message posted");
```

```
            exit(1);
        }
    /*读取消息队列*/
        if(msgrcv(qid,&msg,BUFSZ,0,0)<0){
            perror("msgrcv");
            exit(1);
        }
        printf("message is:%s\n",(&msg)->msg_text);
    /*从系统内核中移走消息队列。*/
        if((msgctl(qid,IPC_RMID,NULL))<0){
            perror("msgctl");
            exit(1);
        }
        exit(0);
    }
```

以下是程序的运行结果。

```
[root@(none) tmp]# ./msg
opened queue 262146
Please enter the message to queue:
hello
message is:hello
```

8.6 实验内容

8.6.1 管道通信实验

1. 实验目的

通过编写有名管道多路通信实验，读者可进一步熟练掌握管道的创建、读写等操作，同时，也复习使用 select 函数实现管道的通信。

2. 实验内容

在该实验中，要求创建两个管道，首先读出管道一中的数据，再把从管道一中读入的数据写入到管道二中去。这里的 select 函数采用阻塞形式，也就是首先在程序中实现将数据写入管道一，并通过 select 函数实现将管道一的数据读出，并写入管道二，接着该程序一直等待用户输入管道一的数据并将其即时读出。

3．实验步骤

（1）画出流程图。

该实验流程图如图 8.9 所示。

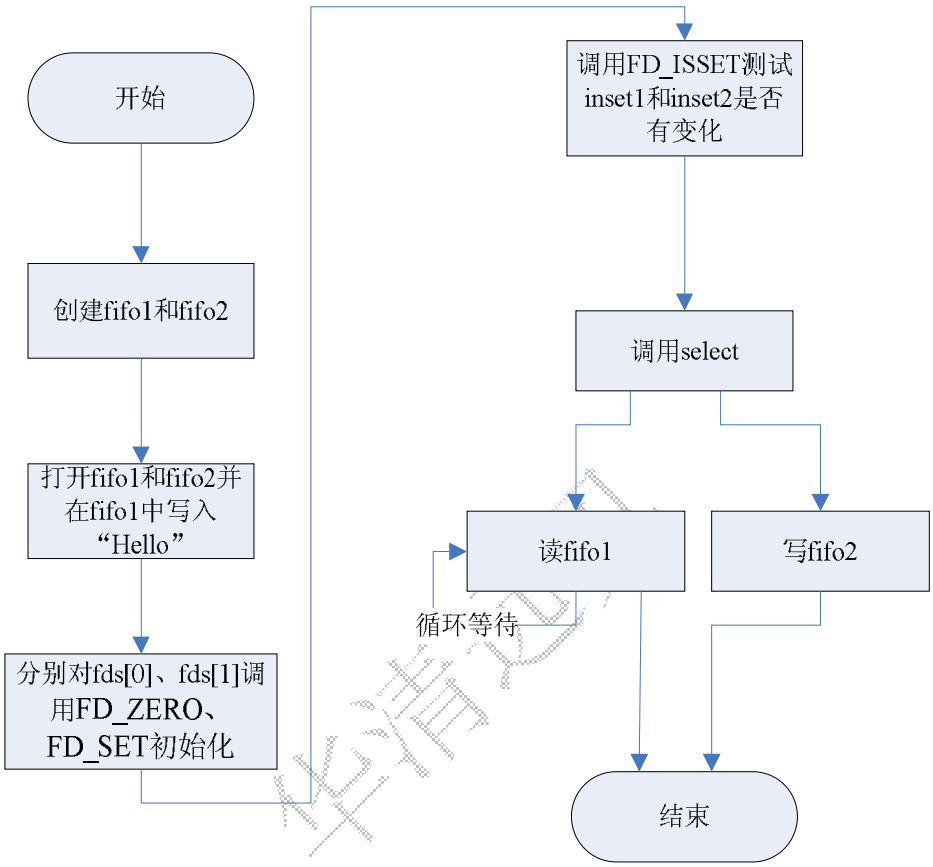


图 8.9　8.6.1 实验流程图

（2）编写代码。

该实验源代码如下所示。

```
/*exec.c*/
#include <sys/stat.h>
#include <sys/types.h>
#include <errno.h>
#include <fcntl.h>
#include <stdio.h>
#include <unistd.h>
#include <stdlib.h>
#include <time.h>

int main(void)
```

```
{
      int fds[2];
      char buf[7];
      int i,rc,maxfd;
      fd_set inset1,inset2;
      struct timeval tv;
/*创建两个有名管道*/
      if((mkfifo("fifo1",O_CREAT|O_EXCL)<0)&&(errno!=EEXIST))
            printf("cannot create fifoserver\n");
      if((mkfifo("fifo2",O_CREAT|O_EXCL)<0)&&(errno!=EEXIST))
            printf("cannot create fifoserver\n");
/*打开有名管道*/
      if((fds[0] = open ("fifo1", O_RDWR|O_NONBLOCK,0))<0)
            perror("open fifo1");
      if((fds[1] = open ("fifo2", O_RDWR|O_NONBLOCK,0))<0)
            perror("open fifo2");
      if((rc = write(fds[0],"Hello!\n",7)))
            printf("rc=%d\n",rc);
      lseek(fds[0],0,SEEK_SET);
      maxfd = fds[0]>fds[1] ? fds[0] : fds[1];
/*/初始化描述集，并将文件描述符加入到相应的描述集*/
      FD_ZERO(&inset1);
      FD_SET(fds[0],&inset1);
      FD_ZERO(&inset2);
      FD_SET(fds[1],&inset2);
/*循环测试该文件描述符是否准备就绪，并调用 select 函数*/
      while(FD_ISSET(fds[0],&inset1)||FD_ISSET(fds[1],&inset2)){
            if(select(maxfd+1,&inset1,&inset2,NULL,NULL)<0)
                  perror("select");
            else{
/*对相关文件描述符做对应操作*/
                  if(FD_ISSET(fds[0],&inset1)){
                        rc = read(fds[0],buf,7);
                        if(rc>0){
                              buf[rc]='\0';
                              printf("read: %s\n",buf);
                  }else
                        perror("read");
                  }
```

```
            if(FD_ISSET(fds[1],&inset2)){
                  rc = write(fds[1],buf,7);
                  if(rc>0){
                        buf[rc]='\0';
                        printf("rc=%d,write: %s\n",rc,buf);
                  }else
                        perror("write");
                  }
            }
      }
      exit(0);
}
```

（3）编译运行该程序。

（4）另开一终端，键入“cat >fifo1”，接着在该管道中键入相关内容，并观察实验结果。

4．实验结果

在一终端中键入“hello”、“why”、“sunq”：

```
[root@(none) tmp]# cat >fifo1
hello
why
sunq
```

在另一终端中显示如下结果：

```
[root@(none) tmp]#./exp1
rc=7
read: Hello!

rc=7,write: Hello!

read: hello

read: why

read: sunq
```

8.6.2 共享内存实验

1. 实验目的

通过编写共享内存实验，读者就可以进一步了解共享内存的具体步骤，同时也进一步加深了对共享内存的理解。由于共享内存涉及同步机制，关于这方面的知识本书现在还没有涉及，因此，现在只在一个进程中对共享内存进行操作。

2. 实验内容

该实现要求利用共享内存实现文件的打开、读写操作。

3. 实验步骤

（1）画出流程图
该实验流程图如图 8.10 所示。

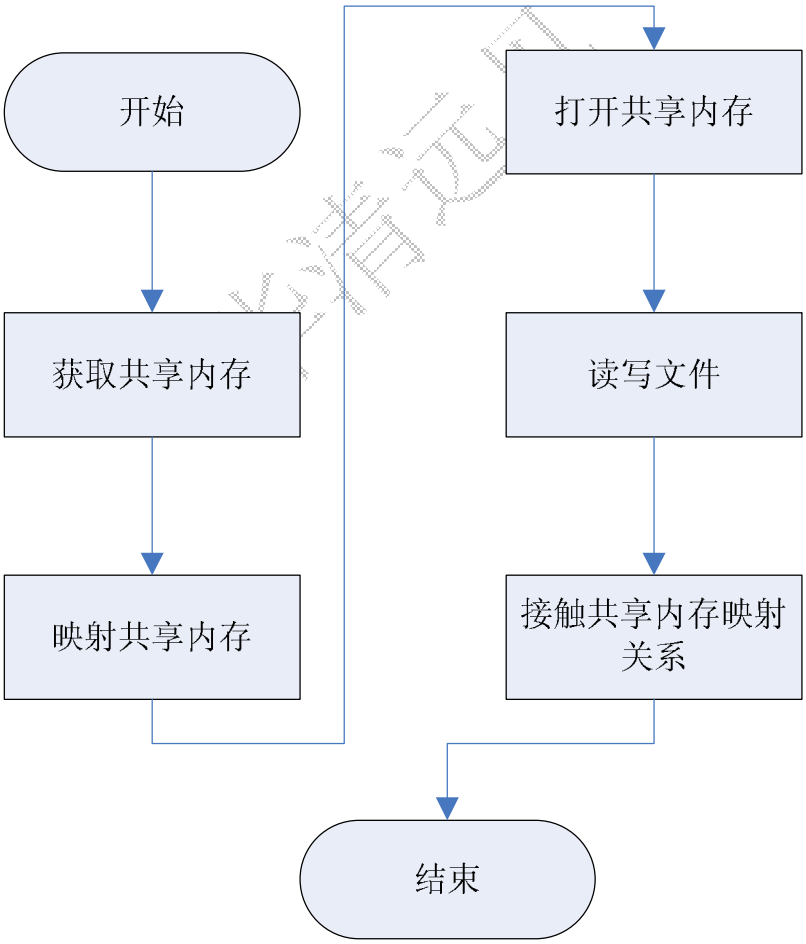


图 8.10 实验 8.6.2 流程图

（2）编写代码

```
/*exec2.c*/
#include <sys/types.h>
#include <sys/ipc.h>
#include <sys/shm.h>
#include <stdio.h>
#include <stdlib.h>
#include <fcntl.h>
#define BUFSZ 2048
int main()
{
      int shmid,i,fd,nwrite,nread;
      char *shmadd;
      char buf[5];
/*创建共享内存*/
      if((shmid=shmget(IPC_PRIVATE,BUFSZ,0666))<0){
            perror("shmget");
            exit(1);
      }
      else
            printf("created shared-memory: %d\n",shmid);
/*映射共享内存*/
      if((shmadd=shmat(shmid,0,0))<(char *)0){
            perror("shmat");
            exit(1);
      }
      else
            printf("attached shared-memory\n");
      shmadd="Hello";
      if((fd = open("share",O_CREAT | O_RDWR,0666))<0){
            perror("open");
            exit(1);
      }
      else
            printf("open success!\n");
      if((nwrite=write(fd,shmadd,5))<0){
            perror("write");
            exit(1);
      }
```

```
        else
                printf("write success!\n");
        lseek( fd, 0, SEEK_SET );
        if((nread=read(fd,buf,5))<0){
                perror("read");
                exit(1);
        }
        else
                printf("read %d form file:%s\n",nread,buf);
    /*删除共享内存*/
        if((shmdt(shmadd))<0){
                perror("shmdt");
                exit(1);
        }
        else
                printf("deleted shared-memory\n");
        exit(0);
    }
```

4. 实验结果

```
[root@(none) tmp]# ./shm
created shared-memory: 1245222
attached shared-memory
open fd=3!
write success!
read 5 form file:Hello
deleted shared-memory
```

本 章 小 结

本章详细讲解了 Linux 中进程间通信的几种机制，包括管道通信、信号通信、消息队列、共享内存机制等，并且讲解了进程间通信的演进。

接下来详细对管道通信、信号通信、消息队列和共享内存机制进行了详细的讲解。其中，管道通信又分为有名管道和无名管道。信号通信中要着重掌握如何对信号进行适当的处理，如采用信号集等方式。

消息队列和共享内存也是很好的进程间通信的手段，其中共享内存具有很高的效率。

本章的实验安排了管道通信实验和共享内存实现，具体的实验数据根据系统的不同可能

会有所区别，希望读者认真完成。

思考与练习

1．通过自定义信号完成进程间的通信。
2．编写一个简单的管道程序实现文件传输。

“黑色经典”系列之《嵌入式 Linux 应用程序开发详解》

第 9 章 多线程编程

本章目标

在前两章中，读者主要学习了有关进程控制和进程间通信的开发，这些都是 Linux 中开发的基础。在这一章中将学习轻量级进程——线程的开发，由于线程的高效性和可操作性，在大型程序开发中运用得非常广泛，希望读者能够很好地掌握。

- 掌握 Linux 中线程的基本概念
- 掌握 Linux 中线程的创建及使用
- 掌握 Linux 中线程属性的设置
- 能够独立编写多线程程序
- 能够处理多线程中的变量问题
- 能够处理多线程中的同步文件

9.1 Linux 下线程概述

9.1.1 线程概述

前面已经提到，进程是系统中程序执行和资源分配的基本单位。每个进程都拥有自己的数据段、代码段和堆栈段，这就造成了进程在进行切换等操作时都需要有比较负责的上下文切换等动作。为了进一步减少处理机的空转时间支持多处理器和减少上下文切换开销，进程在演化中出现了另一个概念——线程。它是一个进程内的基本调度单位，也可以称为轻量级进程。线程是在共享内存空间中并发的多道执行路径，它们共享一个进程的资源，如文件描述和信号处理。因此，大大减少了上下文切换的开销。

同进程一样，线程也将相关的变量值放在线程控制表内。一个进程可以有多个线程，也就是有多个线程控制表及堆栈寄存器，但却共享一个用户地址空间。要注意的是，由于线程共享了进程的资源和地址空间，因此，任何线程对系统资源的操作都会给其他线程带来影响，因此，多线程中的同步就是非常重要的问题了。在多线程系统中，进程与进程的关系如表 9.1 所示。

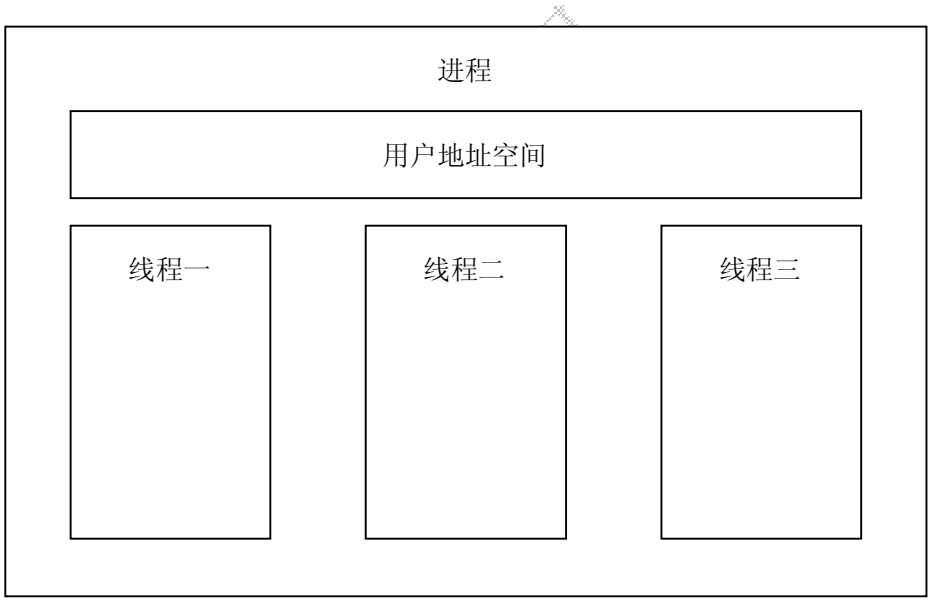


图 9.1 进程与线程关系

9.1.2 线程分类

线程按照其调度者可以分为用户级线程和核心级线程两种。

（1）用户级线程

用户级线程主要解决的是上下文切换的问题，它的调度算法和调度过程全部由用户自行选择决定，在运行时不需要特定的内核支持。在这里，操作系统往往会提供一个用户空间的线程库，该线程库提供了线程的创建、调度、撤销等功能，而内核仍然仅对进程进行管理。

如果一个进程中的某一个线程调用了一个阻塞的系统调用，那么该进程包括该进程中的其他所有线程也同时被阻塞。这种用户级线程的主要缺点是在一个进程中的多个线程的调度中无法发挥多处理器的优势。

（2）核心级线程

这种线程允许不同进程中的线程按照同一相对优先调度方法进行调度，这样就可以发挥多处理器的并发优势。

现在大多数系统都采用用户级线程与核心级线程并存的方法。一个用户级线程可以对应一个或几个核心级线程，也就是“一对一”或“多对一”模型。这样既可满足多处理机系统的需要，也可以最大限度地减少调度开销。

9.1.3 Linux 线程技术的发展

在 Linux 中，线程技术也经过了一代代的发展过程。

在 Linux2.2 内核中，并不存在真正意义上的线程。当时 Linux 中常用的线程 pthread 实际上是通过进程来模拟的，也就是说 Linux 中的线程也是通过 fork 创建的“轻”进程，并且线程的个数也很有限，最多只能有 4096 个进程/线程同时运行。

Linux2.4 内核消除了这个线程个数的限制，并且允许在系统运行中动态地调整进程数上限。当时采用的是 LinuxThread 线程库，它对应的线程模型是“一对一”线程模型，也就是一个用户级线程对应一个内核线程，而线程之间的管理在内核外函数库中实现。这种线程模型得到了广泛应用。但是，LinuxThread 也由于 Linux 内核的限制以及实现难度等原因，并不是完全与 POSIX 兼容。另外，它的进程 ID、信号处理、线程总数、同步等各方面都还有诸多的问题。

为了解决以上问题，在 Linux2.6 内核中，进程调度通过重新编写，删除了以前版本中效率不高的算法。内核线程框架也被重新编写，开始使用 NPTL（Native POSIX Thread Library）线程库。这个线程库有以下几点设计目标：POSIX 兼容性、多处理器结构的应用、低启动开销、低链接开销、与 LinuxThreads 应用的二进制兼容性、软硬件的可扩展能力、与 C++集成等。这一切都使得 Linux2.6 内核的线程机制更加完备，能够更好地完成其设计目标。与 LinuxThreads 不同，NPTL 没有使用管理线程，核心线程的管理直接放在核内进行，这也带了性能的优化。由于 NPTL 仍然采用 1：1 的线程模型，NPTL 仍然不是 POSIX 完全兼容的，但就性能而言相对 LinuxThreads 已经有很大程度上的改进。

9.2 Linux 线程实现

9.2.1 线程基本操作

这里要讲的线程相关操作都是用户空间线程的操作。在 Linux 中，一般 Pthread 线程库是一套通用的线程库，是由 POSIX 提出的，因此具有很好的可移植性。

1．线程创建和退出

（1）函数说明

创建线程实际上就是确定调用该线程函数的入口点，这里通常使用的函数是 pthread_create。在线程创建以后，就开始运行相关的线程函数，在该函数运行完之后，该线程也就退出了，这也是线程退出一种方法。另一种退出线程的方法是使用函数 pthread_exit，这是线程的主动行为。这里要注意的是，在使用线程函数时，不能随意使用 exit 退出函数进行出错处理，由于 exit 的作用是使调用进程终止，往往一个进程包含多个线程，因此，在使用 exit 之后，该进程中的所有线程都终止了。因此，在线程中就可以使用 pthread_exit 来代替进程中的 exit。

由于一个进程中的多个线程是共享数据段的，因此通常在线程退出之后，退出线程所占用的资源并不会随着线程的终止而得到释放。正如进程之间可以用 wait()系统调用来同步终止并释放资源一样，线程之间也有类似机制，那就是 pthread_join()函数。pthread_join 可以用于将当前线程挂起，等待线程的结束。这个函数是一个线程阻塞的函数，调用它的函数将一直等待到被等待的线程结束为止，当函数返回时，被等待线程的资源就被收回。

（2）函数格式

表 9.1 列出了 pthread_create 函数的语法要点。

表 9.1 pthread_create 函数语法要点

所需头文件	#include <pthread.h>
函数原型	int pthread_create ((pthread_t *thread, pthread_attr_t *attr, void *(*start_routine)(void *), void *arg))
函数传入值	thread：线程标识符
	attr：线程属性设置（具体设定在 9.2.2 会进行讲解）
	续表
	start_routine：线程函数的起始地址
	arg：传递给 start_routine 的参数
函数返回值	成功：0
	出错：−1

表 9.2 列出了 pthread_exit 函数的语法要点。

表 9.2 pthread_exit 函数语法要点

所需头文件	#include <pthread.h>
函数原型	void pthread_exit(void *retval)
函数传入值	Retval：pthread_exit()调用者线程的返回值，可由其他函数如 pthread_join 来检索获取

表 9.3 列出了 pthread_join 函数的语法要点。

表 9.3 pthread_join 函数语法要点

所需头文件	#include <pthread.h>
函数原型	int pthread_join ((pthread_t th, void **thread_return))

函数传入值	th：等待线程的标识符
	thread_return：用户定义的指针，用来存储被等待线程的返回值（不为 NULL 时）
函数返回值	成功：0
	出错：−1

（3）函数使用

以下实例中创建了两个线程，其中第一个线程是在程序运行到中途时调用 pthread_exit 函数退出，第二个线程正常运行退出。在主线程中收集这两个线程的退出信息，并释放资源。从这个实例中可以看出，这两个线程是并发运行的。

```
/*thread.c*/
#include <stdio.h>
#include <pthread.h>
/*线程一*/
void thread1(void)
{
     int i=0;
     for(i=0;i<6;i++){
          printf("This is a pthread1.\n");
          if(i==2)
               pthread_exit(0);
          sleep(1);
          }
}
/*线程二*/
void thread2(void)
{
     int i;
     for(i=0;i<3;i++)
          printf("This is a pthread2.\n");
     pthread_exit(0);
}

int main(void)
{
     pthread_t id1,id2;
     int i,ret;
/*创建线程一*/
    ret=pthread_create(&id1,NULL,(void *) thread1,NULL);
```

```
        if(ret!=0){
                printf ("Create pthread error!\n");
        exit (1);
        }
    /*创建线程二*/
        ret=pthread_create(&id2,NULL,(void *) thread2,NULL);
        if(ret!=0){
                printf ("Create pthread error!\n");
                exit (1);
        }
    /*等待线程结束*/
        pthread_join(id1,NULL);
        pthread_join(id2,NULL);
        exit (0);
    }
```

以下是程序运行结果：

```
 [root@(none) tmp]# ./thread
This is a pthread1.
This is a pthread2.
This is a pthread2.
This is a pthread2.
This is a pthread1.
This is a pthread1.
```

2．修改线程属性

（1）函数说明

读者是否还记得 pthread_create 函数的第二个参数——线程的属性。在上一个实例中，将该值设为 NULL，也就是采用默认属性，线程的多项属性都是可以更改的。这些属性主要包括绑定属性、分离属性、堆栈地址、堆栈大小、优先级。其中系统默认的属性为非绑定、非分离、缺省 1M 的堆栈、与父进程同样级别的优先级。下面首先对绑定属性和分离属性的基本概念进行讲解。

- 绑定属性

前面已经提到，Linux 中采用“一对一”的线程机制，也就是一个用户线程对应一个内核线程。绑定属性就是指一个用户线程固定地分配给一个内核线程，因为 CPU 时间片的调度是面向内核线程（也就是轻量级进程）的，因此具有绑定属性的线程可以保证在需要的时候总有一个内核线程与之对应。而与之相对的非绑定属性就是指用户线程和内核线程的关系不是始终固定的，而是由系统来控制分配的。

- 分离属性

分离属性是用来决定一个线程以什么样的方式来终止自己。在非分离情况下，当一个线

程结束时，它所占用的系统资源并没有被释放，也就是没有真正的终止。只有当 pthread_join() 函数返回时，创建的线程才能释放自己占用的系统资源。而在分离属性情况下，一个线程结束时立即释放它所占有的系统资源。这里要注意的一点是，如果设置一个线程的分离属性，而这个线程运行又非常快，那么它很可能在 pthread_create 函数返回之前就终止了，它终止以后就可能将线程号和系统资源移交给其他的线程使用，这时调用 pthread_create 的线程就得到了错误的线程号。

这些属性的设置都是通过一定的函数来完成的，通常首先调用 pthread_attr_init 函数进行初始化，之后再调用相应的属性设置函数。设置绑定属性的函数为 pthread_attr_setscope，设置线程分离属性的函数为 pthread_attr_setdetachstate，设置线程优先级的相关函数为 pthread_attr_getschedparam（获取线程优先级）和 pthread_attr_setschedparam（设置线程优先级）。在设置完这些属性后，就可以调用 pthread_create 函数来创建线程了。

（2）函数格式

表 9.4 列出了 pthread_attr_init 函数的语法要点。

表 9.4　　pthread_attr_init 函数语法要点

所需头文件	#include <pthread.h>
函数原型	int pthread_attr_init(pthread_attr_t *attr)
函数传入值	attr：线程属性
函数返回值	成功：0
	出错：−1

表 9.5 列出了 pthread_attr_setscope 函数的语法要点。

表 9.5　　pthread_attr_setscope 函数语法要点

所需头文件	#include <pthread.h>	
函数原型	int pthread_attr_setscope(pthread_attr_t *attr, int scope)	
函数传入值	attr：线程属性	
	scope	PTHREAD_SCOPE_SYSTEM：绑定
		PTHREAD_SCOPE_PROCESS：非绑定
函数返回值	成功：0	
	出错：−1	

表 9.6 列出了 pthread_attr_setdetachstate 函数的语法要点。

表 9.6　　pthread_attr_setdetachstate 函数语法要点

所需头文件	#include <pthread.h>	
函数原型	int pthread_attr_setscope(pthread_attr_t *attr, int detachstate)	
函数传入值	attr：线程属性	
	detachstate	PTHREAD_CREATE_DETACHED：分离

		PTHREAD _CREATE_JOINABLE：非分离
函数返回值	成功：0	
	出错：–1	

表 9.7 列出了 pthread_attr_getschedparam 函数的语法要点。

表 9.7 pthread_attr_getschedparam 函数语法要点

所需头文件	#include <pthread.h>
函数原型	int pthread_attr_getschedparam (pthread_attr_t *attr, struct sched_param *param)
函数传入值	attr：线程属性
	param：线程优先级
函数返回值	成功：0
	出错：–1

表 9.8 列出了 pthread_attr_setschedparam 函数的语法要点。

表 9.8 pthread_attr_setschedparam 函数语法要点

所需头文件	#include <pthread.h>
函数原型	int pthread_attr_setschedparam (pthread_attr_t *attr, struct sched_param *param)
函数传入值	attr：线程属性
	param：线程优先级
函数返回值	成功：0
	出错：–1

3．使用实例

该实例将上一节中的第一个线程设置为分离属性，并将第二个线程设置为始终运行状态，这样就可以在第二个线程运行过程中查看内存值的变化。

其源代码如下所示：

```
/*pthread.c*/
#include <stdio.h>
#include <pthread.h>
#include <time.h>
/*线程一*/
void thread1(void)
{
      int i=0;
      for(i=0;i<6;i++){
            printf("This is a pthread1.\n");
            if(i==2)
```

```
                pthread_exit(0);
            sleep(1);
        }
}
/*线程二*/
void thread2(void)
{
        int i;
        while(1){
        for(i=0;i<3;i++)
            printf("This is a pthread2.\n");
        sleep(1);}
        pthread_exit(0);
}

int main(void)
{
        pthread_t id1,id2;
        int i,ret;
        pthread_attr_t attr;
/*初始化线程*/
        pthread_attr_init(&attr);
/*设置线程绑定属性*/
        pthread_attr_setscope(&attr, PTHREAD_SCOPE_SYSTEM);
/*设置线程分离属性*/
        pthread_attr_setdetachstate(&attr,PTHREAD_CREATE_DETACHED);
/*创建线程*/
        ret=pthread_create(&id1,&attr,(void *) thread1,NULL);
        if(ret!=0){
            printf ("Create pthread error!\n");
        exit (1);
        }
        ret=pthread_create(&id2,NULL,(void *) thread2,NULL);
        if(ret!=0){
            printf ("Create pthread error!\n");
            exit (1);
        }
        pthread_join(id2,NULL);
        return (0);
```

```
}
```

接下来可以在线程一运行前后使用“free”命令查看内存使用情况。以下是运行结果：

```
[root@(none) tmp]# ./thread3
This is a pthread1.
This is a pthread2.
This is a pthread2.
This is a pthread2.
This is a pthread1.
This is a pthread2.
This is a pthread2.
This is a pthread2.
This is a pthread1.
This is a pthread2.
…
[root@www yul]# free
             total       used       free     shared    buffers     cached
Mem:         1028428     570212     458216         48      204292      93196
-/+ buffers/cache:  272724     755704
Swap:       1020116          0     1020116
[root@www yul]# free
             total       used       free     shared    buffers     cached
Mem:         1028428     570220     458208         48     204296      93196
-/+ buffers/cache:     272728     755700
Swap:        1020116          0    1020116
[root@www yul]# free
             total       used       free     shared    buffers     cached
Mem:         1028428     570212     458216         48     204296      93196
-/+ buffers/cache:    272720     755708
Swap:       1020116          0    1020116
```

可以看到，线程一在运行结束后就收回了系统资源，释放了内存。

9.2.2 线程访问控制

由于线程共享进程的资源和地址空间，因此在对这些资源进行操作时，必须考虑到线程间资源访问的惟一性问题，这里主要介绍 POSIX 中线程同步的方法，主要有互斥锁和信号量的方式。

1．mutex 互斥锁线程控制

（1）函数说明

mutex 是一种简单的加锁的方法来控制对共享资源的存取。这个互斥锁只有两种状态，也就是上锁和解锁，可以把互斥锁看作某种意义上的全局变量。在同一时刻只能有一个线程掌握某个互斥上的锁，拥有上锁状态的线程能够对共享资源进行操作。若其他线程希望上锁一个已经上锁了的互斥锁，则该线程就会挂起，直到上锁的线程释放掉互斥锁为止。可以说，这把互斥锁使得共享资源按序在各个线程中操作。

互斥锁的操作主要包括以下几个步骤。

- 互斥锁初始化：pthread_mutex_init
- 互斥锁上锁：pthread_mutex_lock
- 互斥锁判断上锁：pthread_mutex_trylock
- 互斥锁接锁：pthread_mutex_unlock
- 消除互斥锁：pthread_mutex_destroy

其中，互斥锁可以分为快速互斥锁、递归互斥锁和检错互斥锁。这三种锁的区别主要在于其他未占有互斥锁的线程在希望得到互斥锁时的是否需要阻塞等待。快速锁是指调用线程会阻塞直至拥有互斥锁的线程解锁为止。递归互斥锁能够成功地返回并且增加调用线程在互斥上加锁的次数，而检错互斥锁则为快速互斥锁的非阻塞版本，它会立即返回并返回一个错误信息。

（2）函数格式

表 9.9 列出了 pthread_mutex_init 函数的语法要点。

表 9.9　　pthread_mutex_init 函数语法要点

所需头文件	#include <pthread.h>	
函数原型	int pthread_mutex_init(pthread_mutex_t *mutex, const pthread_mutexattr_t *mutexattr)	
函数传入值	Mutex：互斥锁	

续表

函数传入值	Mutexattr	PTHREAD_MUTEX_INITIALIZER：创建快速互斥锁
		PTHREAD_RECURSIVE_MUTEX_INITIALIZER_NP：创建递归互斥锁
		PTHREAD_ERRORCHECK_MUTEX_INITIALIZER_NP：创建检错互斥锁
函数返回值	成功：0	
	出错：−1	

表 9.10 列出了 pthread_mutex_lock 等函数的语法要点。

表 9.10　　pthread_mutex_lock 等函数语法要点

所需头文件	#include <pthread.h>
函数原型	int pthread_mutex_lock(pthread_mutex_t *mutex,) int pthread_mutex_trylock(pthread_mutex_t *mutex,)

	int pthread_mutex_unlock(pthread_mutex_t *mutex,) int pthread_mutex_destroy(pthread_mutex_t *mutex,)
函数传入值	Mutex：互斥锁
函数返回值	成功：0
	出错：−1

（3）使用实例

该实例使用互斥锁来实现对变量 lock_var 的加一、打印操作。

```
/*mutex.c*/
#include <stdio.h>
#include <stdlib.h>
#include <unistd.h>
#include <pthread.h>
#include <errno.h>

pthread_mutex_t mutex = PTHREAD_MUTEX_INITIALIZER;
int lock_var;
time_t end_time;

void pthread1(void *arg);
void pthread2(void *arg);

int main(int argc, char *argv[])
{
     pthread_t id1,id2;
     pthread_t mon_th_id;
     int ret;

     end_time = time(NULL)+10;
/*互斥锁初始化*/
     pthread_mutex_init(&mutex,NULL);
/*创建两个线程*/
     ret=pthread_create(&id1,NULL,(void *)pthread1, NULL);
     if(ret!=0)
           perror("pthread cread1");
     ret=pthread_create(&id2,NULL,(void *)pthread2, NULL);
     if(ret!=0)
           perror("pthread cread2");
```

```
        pthread_join(id1,NULL);
        pthread_join(id2,NULL);
        exit(0);
}

void pthread1(void *arg)
{
        int i;
        while(time(NULL) < end_time){
/*互斥锁上锁*/
            if(pthread_mutex_lock(&mutex)!=0){
            perror("pthread_mutex_lock");
            }
            else
                   printf("pthread1:pthread1 lock the variable\n");
            for(i=0;i<2;i++){
                  sleep(1);
                  lock_var++;
            }
/*互斥锁接锁*/
            if(pthread_mutex_unlock(&mutex)!=0){
                  perror("pthread_mutex_unlock");
            }
            else
                  printf("pthread1:pthread1 unlock the variable\n");
        sleep(1);
        }
}

void pthread2(void *arg)
{
        int nolock=0;
        int ret;
        while(time(NULL) < end_time){
/*测试互斥锁*/
            ret=pthread_mutex_trylock(&mutex);
            if(ret==EBUSY)
                  printf("pthread2:the variable is locked by pthread1\n");
            else{
```

```
                if(ret!=0){
                        perror("pthread_mutex_trylock");
                        exit(1);
                }
                else
                        printf("pthread2:pthread2 got lock.The variable is
%d\n",lock_var);
    /*互斥锁接锁*/
                if(pthread_mutex_unlock(&mutex)!=0){
                      perror("pthread_mutex_unlock");
                }
                else
                      printf("pthread2:pthread2 unlock the variable\n");
            }
            sleep(3);
         }
    }
```

该实例的运行结果如下所示：

```
 [root@(none) tmp]# ./mutex2
pthread1:pthread1 lock the variable
pthread2:the variable is locked by pthread1
pthread1:pthread1 unlock the variable
pthread:pthread2 got lock.The variable is 2
pthread2:pthread2 unlock the variable
pthread1:pthread1 lock the variable
pthread1:pthread1 unlock the variable
pthread:pthread2 got lock.The variable is 4
pthread2:pthread2 unlock the variable
pthread1:pthread1 lock the variable
…
```

2．信号量线程控制

（1）信号量说明

在第 8 章中已经讲到，信号量也就是操作系统中所用到的 PV 原语，它广泛用于进程或线程间的同步与互斥。信号量本质上是一个非负的整数计数器，它被用来控制对公共资源的访问。这里先来简单复习一下 PV 原语的工作原理。

PV 原语是对整数计数器信号量 sem 的操作。一次 P 操作使 sem 减一，而一次 V 操作使

sem 加一。进程（或线程）根据信号量的值来判断是否对公共资源具有访问权限。当信号量 sem 的值大于等于零时，该进程（或线程）具有公共资源的访问权限；相反，当信号量 sem 的值小于零时，该进程（或线程）就将阻塞直到信号量 sem 的值大于等于 0 为止。

PV 原语主要用于进程或线程间的同步和互斥这两种典型情况。若用于互斥，几个进程（或线程）往往只设置一个信号量 sem，它们的操作流程如图 9.2 所示。

当信号量用于同步操作时，往往会设置多个信号量，并安排不同的初始值来实现它们之间的顺序执行，它们的操作流程如图 9.3 所示。

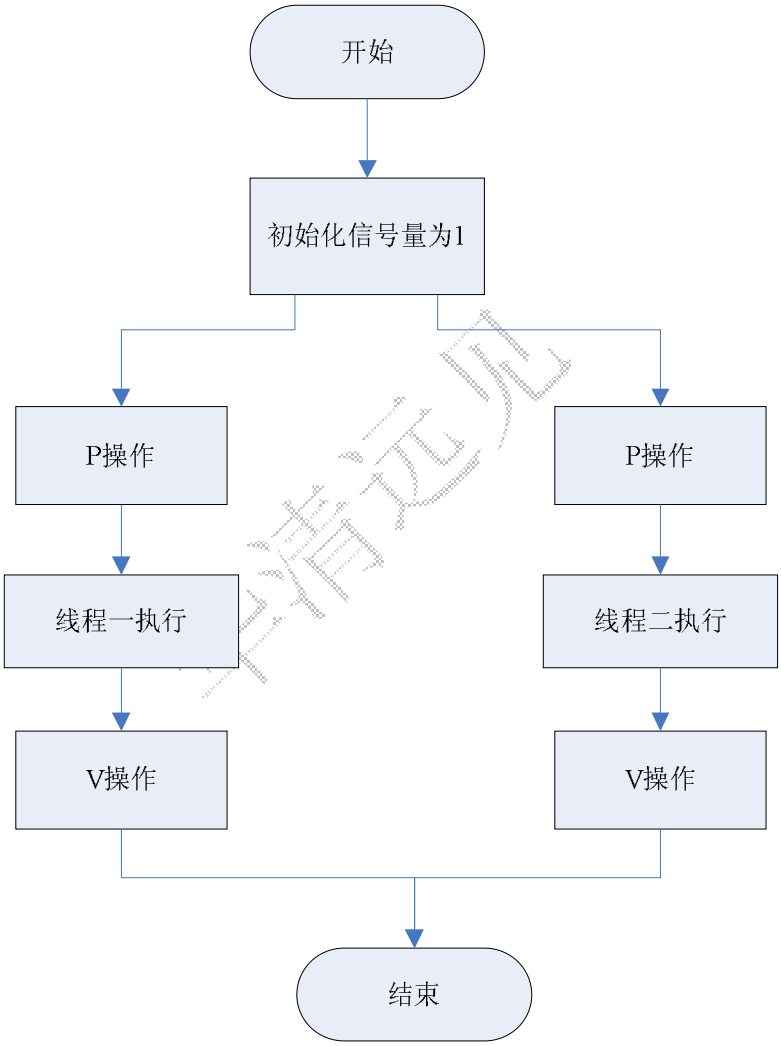


图 9.2　信号量互斥操作

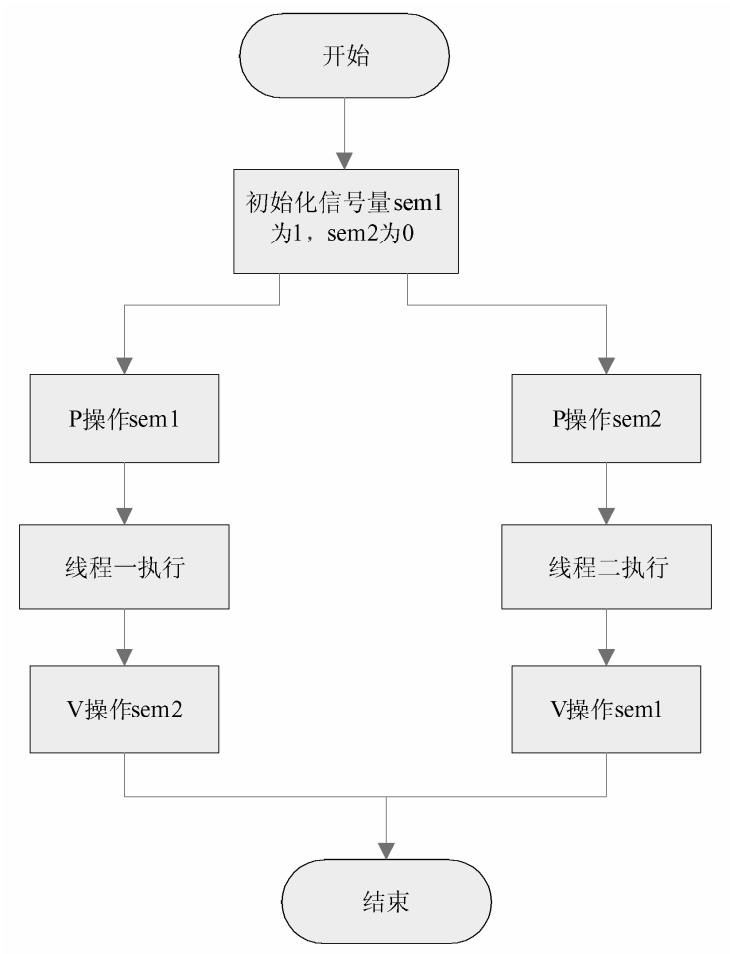


图 9.3　信号量同步操作

（2）函数说明

Linux 实现了 POSIX 的无名信号量，主要用于线程间的互斥同步。这里主要介绍几个常见函数。

- sem_init 用于创建一个信号量，并能初始化它的值。
- sem_wait 和 sem_trywait 相当于 P 操作，它们都能将信号量的值减一，两者的区别在于若信号量小于零时，sem_wait 将会阻塞进程，而 sem_trywait 则会立即返回。
- sem_post 相当于 V 操作，它将信号量的值加一同时发出信号唤醒等待的进程。
- sem_getvalue 用于得到信号量的值。
- sem_destroy 用于删除信号量。

（3）函数格式

表 9.11 列出了 sem_init 函数的语法要点。

表 9.11　　sem_init 函数语法要点

所需头文件	#include <semaphore.h>

函数原型	int sem_init(sem_t *sem,int pshared,unsigned int value)
函数传入值	sem：信号量
	pshared：决定信号量能否在几个进程间共享。由于目前 Linux 还没有实现进程间共享信号量，所以这个值只能够取 0
	value：信号量初始化值
函数返回值	成功：0
	出错：−1

表 9.12 列出了 sem_wait 等函数的语法要点。

表 9.12　　sem_wait 等函数语法要点

所需头文件	#include <pthread.h>
函数原型	int sem_wait(sem_t *sem) int sem_trywait(sem_t *sem) int sem_post(sem_t *sem) int sem_getvalue(sem_t *sem) int sem_destroy(sem_t *sem)
函数传入值	sem：信号量
函数返回值	成功：0
	出错：−1

（4）使用实例

下面实例 1 使用信号量实现了上一实例中对 lock_var 的操作，在这里使用的是互斥操作，也就是只使用一个信号量来实现。代码如下所示：

```
/*sem_mutex.c*/
#include <stdio.h>
#include <stdlib.h>
#include <unistd.h>
#include <pthread.h>
#include <errno.h>
#include <sys/ipc.h>
#include <semaphore.h>
int lock_var;
time_t end_time;
sem_t sem;

void pthread1(void *arg);
void pthread2(void *arg);
```

```
int main(int argc, char *argv[])
{
     pthread_t id1,id2;
     pthread_t mon_th_id;
     int ret;
     end_time = time(NULL)+30;
/*初始化信号量为 1*/
     ret=sem_init(&sem,0,1);
     if(ret!=0)
     {
           perror("sem_init");
     }
/*创建两个线程*/
     ret=pthread_create(&id1,NULL,(void *)pthread1, NULL);
     if(ret!=0)
           perror("pthread cread1");
     ret=pthread_create(&id2,NULL,(void *)pthread2, NULL);
     if(ret!=0)
           perror("pthread cread2");
           pthread_join(id1,NULL);
           pthread_join(id2,NULL);

     exit(0);
}

void pthread1(void *arg)
{
     int i;
     while(time(NULL) < end_time){
/*信号量减一，P 操作*/
         sem_wait(&sem);
         for(i=0;i<2;i++){
               sleep(1);
               lock_var++;
               printf("lock_var=%d\n",lock_var);
         }
         printf("pthread1:lock_var=%d\n",lock_var);
/*信号量加一，V 操作*/
         sem_post(&sem);
```

```
            sleep(1);
        }
    }

    void pthread2(void *arg)
    {
        int nolock=0;
        int ret;
        while(time(NULL) < end_time){
    /*信号量减一，P 操作*/
            sem_wait(&sem);
            printf("pthread2:pthread1 got lock;lock_var=%d\n",lock_var);
    /*信号量加一，V 操作*/
            sem_post(&sem);
            sleep(3);
        }
    }
```

程序运行结果如下所示：

```
[root@(none) tmp]# ./sem_num
lock_var=1
lock_var=2
pthread1:lock_var=2
pthread2:pthread1 got lock;lock_var=2
lock_var=3
lock_var=4
pthread1:lock_var=4
pthread2:pthread1 got lock;lock_var=4
```

接下来是通过两个信号量来实现两个线程间的同步，仍然完成了以上实例中对 lock_var 的操作。代码如下所示：

```
/*sem_syn.c*/
#include <stdio.h>
#include <stdlib.h>
#include <unistd.h>
#include <pthread.h>
#include <errno.h>
#include <sys/ipc.h>
#include <semaphore.h>
```

```
int lock_var;
time_t end_time;
sem_t sem1,sem2;

void pthread1(void *arg);
void pthread2(void *arg);

int main(int argc, char *argv[])
{
     pthread_t id1,id2;
     pthread_t mon_th_id;
     int ret;
     end_time = time(NULL)+30;
/*初始化两个信号量，一个信号量为 1，一个信号量为 0*/
     ret=sem_init(&sem1,0,1);
     ret=sem_init(&sem2,0,0);
     if(ret!=0)
     {
            perror("sem_init");
     }
/*创建两个线程*/
     ret=pthread_create(&id1,NULL,(void *)pthread1, NULL);
     if(ret!=0)
            perror("pthread cread1");
     ret=pthread_create(&id2,NULL,(void *)pthread2, NULL);
     if(ret!=0)
            perror("pthread cread2");
     pthread_join(id1,NULL);
     pthread_join(id2,NULL);
     exit(0);
}

void pthread1(void *arg)
{
     int i;
     while(time(NULL) < end_time){
/*P 操作信号量 2*/
            sem_wait(&sem2);
```

```
            for(i=0;i<2;i++){
                  sleep(1);
                  lock_var++;
                  printf("lock_var=%d\n",lock_var);
            }
            printf("pthread1:lock_var=%d\n",lock_var);
/*V 操作信号量 1*/
            sem_post(&sem1);
            sleep(1);
      }
}

void pthread2(void *arg)
{
      int nolock=0;
      int ret;
      while(time(NULL) < end_time){
/*P 操作信号量 1*/
      sem_wait(&sem1);
      printf("pthread2:pthread1 got lock;lock_var=%d\n",lock_var);
/*V 操作信号量 2*/
      sem_post(&sem2);
      sleep(3);
      }
}
```

从以下结果中可以看出，该程序确实实现了先运行线程二，再运行线程一。

```
[root@(none) tmp]# ./sem_num
pthread2:pthread1 got lock;lock_var=0
lock_var=1
lock_var=2
pthread1:lock_var=2
pthread2:pthread1 got lock;lock_var=2
lock_var=3
lock_var=4
pthread1:lock_var=4
```

9.3 实验内容——“生产者消费者”实验

1．实验目的

“生产者消费者”问题是一个著名的同时性编程问题的集合。通过编写经典的“生产者消费者”问题的实验，读者可以进一步熟悉 Linux 中多线程编程，并且掌握用信号量处理线程间的同步互斥问题。

2．实验内容

“生产者消费者”问题描述如下。

有一个有限缓冲区和两个线程：生产者和消费者。他们分别把产品放入缓冲区和从缓冲区中拿走产品。当一个生产者在缓冲区满时必须等待，当一个消费者在缓冲区空时页必须等待。它们之间的关系如下图所示：

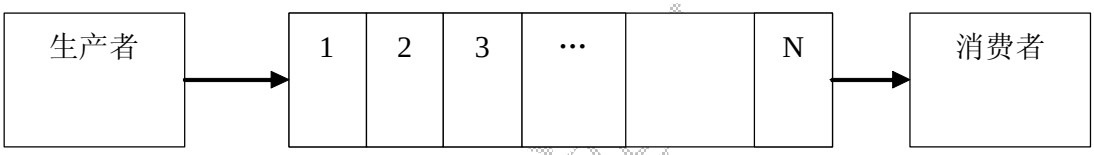


图 9.4 生产者消费者问题描述

这里要求用有名管道来模拟有限缓冲区，用信号量来解决生产者消费者问题中的同步和互斥问题。

3．实验步骤

（1）信号量的考虑

这里使用 3 个信号量，其中两个信号量 avail 和 full 分别用于解决生产者和消费者线程之间的同步问题，mutex 是用于这两个线程之间的互斥问题。其中 avail 初始化为 N（有界缓冲区的空单元数），mutex 初始化为 1，full 初始化为 0。

（2）画出流程图

本实验流程图如下图 9.5 所示。

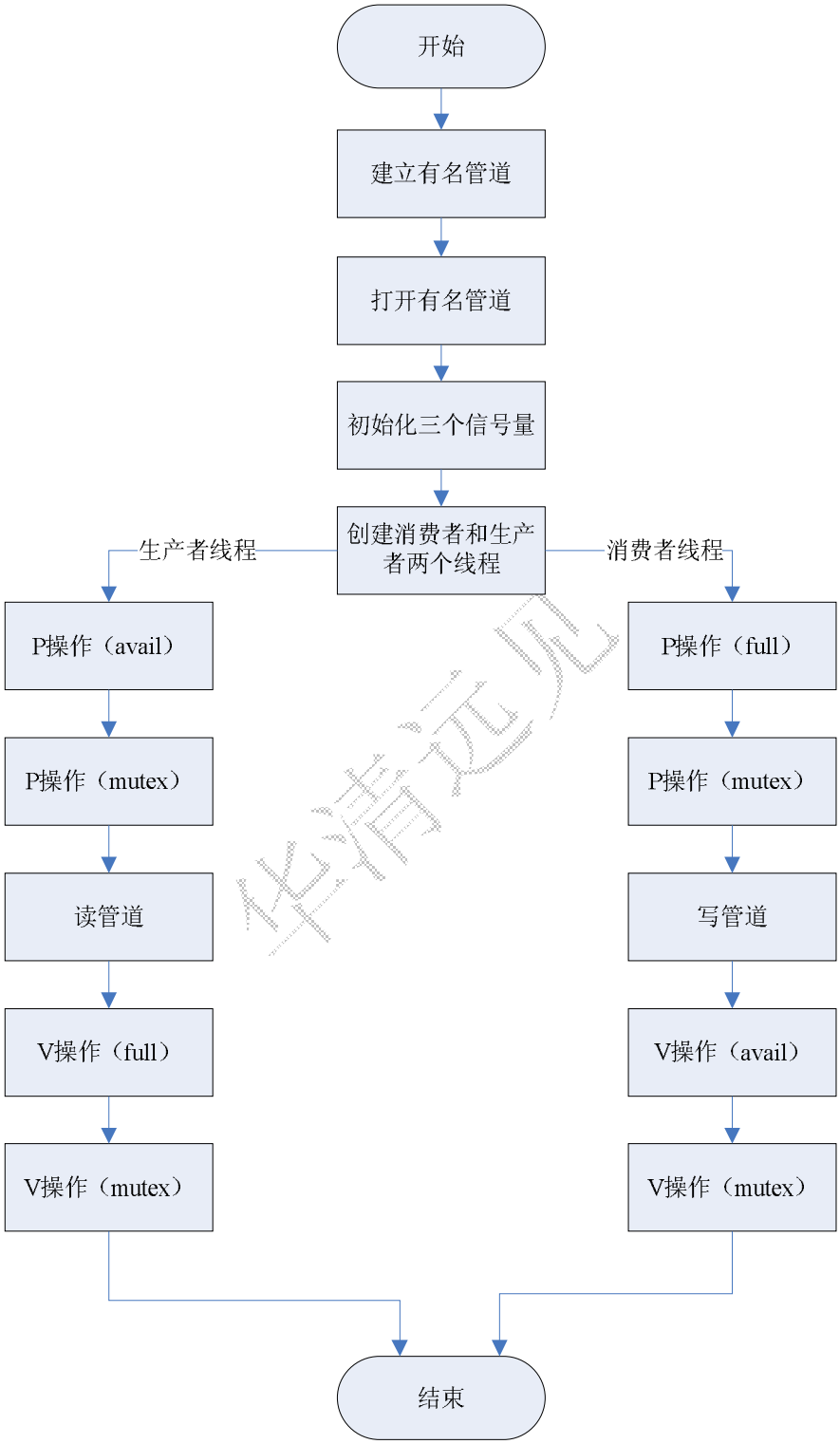


图 9.5 “生产者消费者”实验流程图

（3）编写代码

本实验代码如下：

```
/*product.c*/
#include <stdio.h>
#include <stdlib.h>
#include <unistd.h>
#include <pthread.h>
#include <errno.h>
#include <sys/ipc.h>
#include <semaphore.h>
#include <fcntl.h>
#define FIFO "myfifo"
#define N 5
int lock_var;
time_t end_time;
char buf_r[100];
sem_t mutex,full,avail;
int fd;
void pthread1(void *arg);
void pthread2(void *arg);

int main(int argc, char *argv[])
{
     pthread_t id1,id2;
     pthread_t mon_th_id;
     int ret;
     end_time = time(NULL)+30;
/*创建有名管道*/
     if((mkfifo(FIFO,O_CREAT|O_EXCL)<0)&&(errno!=EEXIST))
          printf("cannot create fifoserver\n");
     printf("Preparing for reading bytes...\n");
     memset(buf_r,0,sizeof(buf_r));
/*打开管道*/
     fd=open(FIFO,O_RDWR|O_NONBLOCK,0);
     if(fd==-1)
     {
          perror("open");
          exit(1);
     }
```

```
/*初始化互斥信号量为 1*/
     ret=sem_init(&mutex,0,1);
/*初始化 avail 信号量为 N*/
     ret=sem_init(&avail,0,N);
/*初始化 full 信号量为 0*/
     ret=sem_init(&full,0,0);
     if(ret!=0)
     {
           perror("sem_init");
     }
/*创建两个线程*/
     ret=pthread_create(&id1,NULL,(void *)productor, NULL);
     if(ret!=0)
           perror("pthread cread1");
     ret=pthread_create(&id2,NULL,(void *)consumer, NULL);
     if(ret!=0)
           perror("pthread cread2");
     pthread_join(id1,NULL);
     pthread_join(id2,NULL);

     exit(0);
}

/*生产者线程*/
void productor(void *arg)
{
     int i,nwrite;
     while(time(NULL) < end_time){
/*P 操作信号量 avail 和 mutex*/
          sem_wait(&avail);
          sem_wait(&mutex);
/*生产者写入数据*/
          if((nwrite=write(fd,"hello",5))==-1)
          {
               if(errno==EAGAIN)
                     printf("The FIFO has not been read yet.Please try later\n");
          }
          else
               printf("write hello to the FIFO\n");
```

```
/*V操作信号量full和mutex*/
          sem_post(&full);
          sem_post(&mutex);
          sleep(1);
     }
}

/*消费者线程*/
void consumer(void *arg)
{
     int nolock=0;
     int ret,nread;
     while(time(NULL) < end_time){
/*P操作信号量full和mutex*/
         sem_wait(&full);
         sem_wait(&mutex);
         memset(buf_r,0,sizeof(buf_r));
         if((nread=read(fd,buf_r,100))==-1){
               if(errno==EAGAIN)
                     printf("no data yet\n");
         }
         printf("read %s from FIFO\n",buf_r);
/*V操作信号量avail和mutex*/
         sem_post(&avail);
         sem_post(&mutex);
         sleep(1);
     }
}
```

4．实验结果

运行该程序，得到如下结果：

```
[root@(none) tmp]#./exec
Preparing for reading bytes...
write hello to the FIFO
read hello from FIFO
write hello to the FIFO
read hello from FIFO
write hello to the FIFO
```

```
read hello from FIFO
write hello to the FIFO
read hello from FIFO
```

本 章 小 结

本章首先介绍了线程的基本概念、线程的分类和线程的发展历史，可以看出，线程技术已有了很大的进展。

接下来讲解了 Linux 中线程基本操作的 API 函数，包括线程的创建及退出，修改线程属性的操作，对每种操作都给出了简短的实例并加以说明。

再接下来，本章讲解了线程的控制操作，由于线程的操作必须包括线程间的同步互斥操作，包括互斥锁线程控制和信号量线程控制。

最后，本章的实验是一个经典的生产者——消费者问题，可以使用线程机制很好地实现，希望读者能够认真地编程实验，进一步理解多线程的同步、互斥操作。

思考与练习

通过查找资料，查看主流的嵌入式操作系统（如嵌入式 Linux，Vxworks 等）是如何处理多线程操作的？

“黑色经典”系列之《嵌入式 Linux 应用程序开发详解》

第 10 章 嵌入式 Linux 网络编程

本章目标

本章将介绍嵌入式 Linux 网络编程的基础知识。由于网络在嵌入式中的应用非常广泛，基本上常见的应用都会与网络有关，因此，掌握这一部分的内容是非常重要的。经过本章的学习，读者将会掌握以下内容。

- 掌握 TCP/IP 协议的基触知识
- 掌握嵌入式 Linux 基础网络编程
- 掌握嵌入式 Linux 高级网络编程
- 分析理解 Ping 源代码
- 能够独立编写客户端、服务器端的通信程序
- 能够独立编写 NTP 协议实现程序

10.1 TCP/IP 协议概述

10.1.1 OSI 参考模型及 TCP/IP 参考模型

读者一定都听说过著名的 OSI 协议参考模型，它是基于国际标准化组织（ISO）的建议发展起来的，从上到下共分为 7 层：应用层、表示层、会话层、传输层、网络层、数据链路层及物理层。这个 7 层的协议模型虽然规定得非常细致和完善，但在实际中却得不到广泛的应用，其重要的原因之一就在于它过于复杂。但它仍是此后很多协议模型的基础，这种分层架构的思想在很多领域都得到了广泛的应用。

与此相区别的 TCP/IP 协议模型从一开始就遵循简单明确的设计思路，它将 TCP/IP 的 7 层协议模型简化为 4 层，从而更有利于实现和使用。TCP/IP 的协议参考模型和 OSI 协议参考模型的对应关系如下图 10.1 所示。

下面分别对者 TCP/IP 的 4 层模型进行简要介绍。

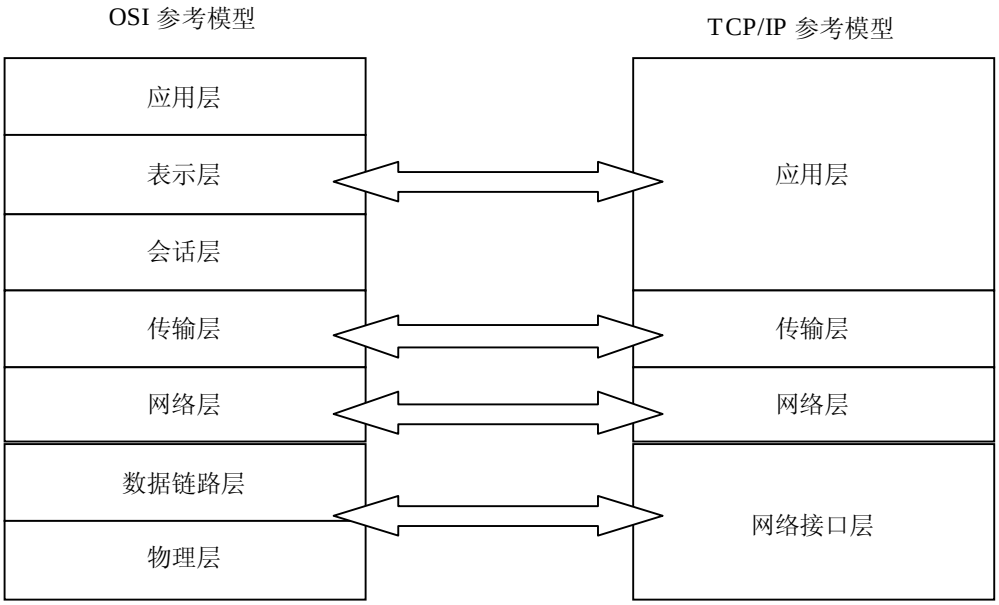


图 10.1 OSI 模型和 TCP/IP 参考模型对应关系

- 网络接口层：负责将二进制流转换为数据帧，并进行数据帧的发送和接收。要注意的是数据帧是独立的网络信息传输单元。
- 网络层：负责将数据帧封装成 IP 数据报，并运行必要的路由算法。
- 传输层：负责端对端之间的通信会话连接与建立。传输协议的选择根据数据传输方式而定。
- 应用层：负责应用程序的网络访问，这里通过端口号来识别各个不同的进程。

10.1.2 TCP/IP 协议族

虽然 TCP/IP 名称只包含了两个协议，但实际上，TCP/IP 是一个庞大的协议族，它包括了各个层次上的众多协议，图 10.2 列举了各层中一些重要的协议，并给出了各个协议在不同层次中所处的位置如下。

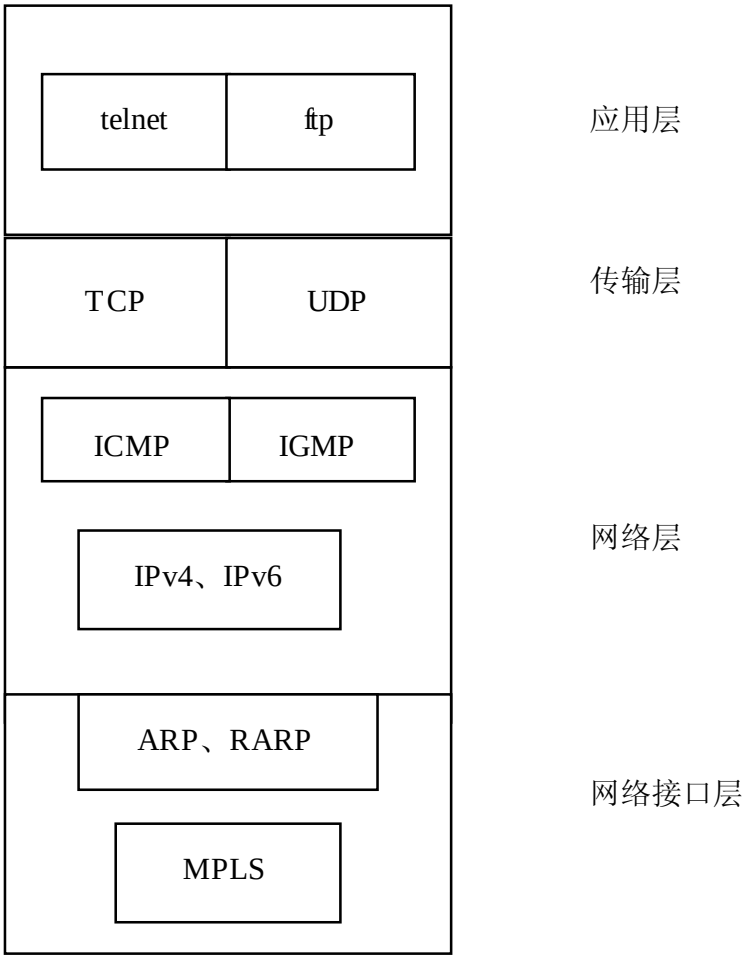


图 10.2 TCP/IP 协议族

- ARP：用于获得同一物理网络中的硬件主机地址。
- MPLS：多协议标签协议，是很有发展前景的下一代网络协议。
- IP：负责在主机和网络之间寻址和路由数据包。
- ICMP：用于发送报告有关数据包的传送错误的协议。
- IGMP：被 IP 主机用来向本地多路广播路由器报告主机组成员的协议。
- TCP：为应用程序提供可靠的通信连接。适合于一次传输大批数据的情况。并适用于要求得到响应的应用程序。
- UDP：提供了无连接通信，且不对传送包进行可靠的保证。适合于一次传输少量数据，

可靠性则由应用层来负责。

10.1.3 TCP 和 UDP

在此主要介绍在网络编程中涉及到的传输层 TCP 和 UDP 协议。

1．TCP

（1）概述

同其他任何协议栈一样，TCP 向相邻的高层提供服务。因为 TCP 的上一层就是应用层，因此，TCP 数据传输实现了从一个应用程序到另一个应用程序的数据传递。应用程序通过编程调用 TCP 并使用 TCP 服务，提供需要准备发送的数据，用来区分接收数据应用的目的地址和端口号。

通常应用程序通过打开一个 socket 来使用 TCP 服务，TCP 管理到其他 socket 的数据传递。可以说，通过 IP 的源/目的可以惟一地区分网络中两个设备的关联，通过 socket 的源/目的可以惟一地区分网络中两个应用程序的关联。

（2）三次握手协议

TCP 对话通过三次握手来初始化的。三次握手的目的是使数据段的发送和接收同步，告诉其他主机其一次可接收的数据量，并建立虚连接。

下面描述了这三次握手的简单过程。

- 初始化主机通过一个同步标志置位的数据段发出会话请求。
- 接收主机通过发回具有以下项目的数据段表示回复：同步标志置位、即将发送的数据段的起始字节的顺序号、应答并带有将收到的下一个数据段的字节顺序号。
- 请求主机再回送一个数据段，并带有确认顺序号和确认号。

图 10.3 就是这个流程的简单示意图。

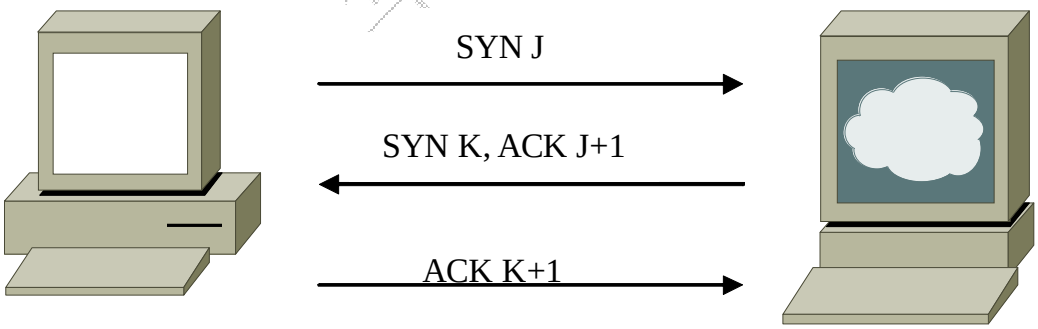


图 10.3　TCP 三次握手协议

TCP 实体所采用的基本协议是滑动窗口协议。当发送方传送一个数据报时，它将启动计时器。当该数据报到达目的地后，接收方的 TCP 实体向回发送一个数据报，其中包含有一个确认序号，它意思是希望收到的下一个数据报的顺序号。如果发送方的定时器在确认信息到达之前超时，那么发送方会重发该数据报。

（3）TCP 数据报头

图 10.4 给出了 TCP 数据报头的格式。

TCP 数据报头的含义如下所示。

- 源端口、目的端口：16 位长。标识出远端和本地的端口号。

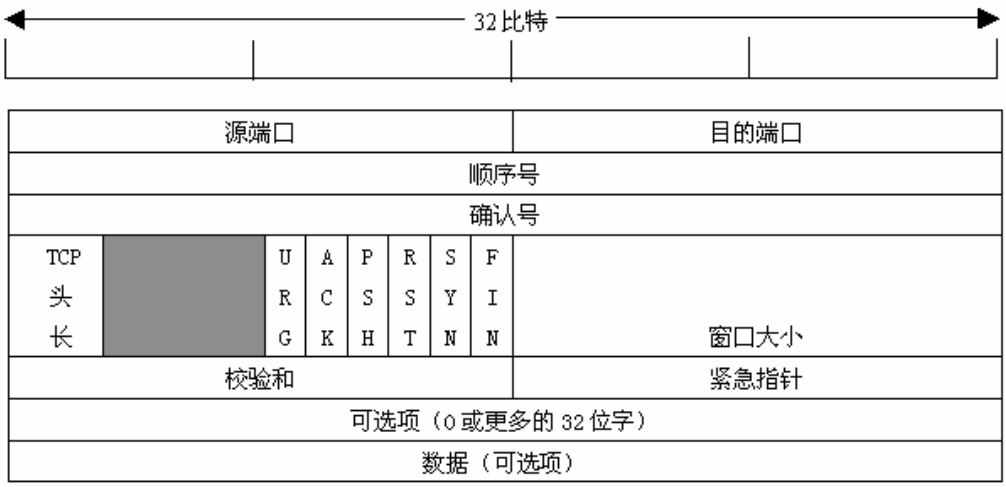


图 10.4　TCP 数据报头的格式

- 序号：32 位长。标识发送的数据报的顺序。
- 确认号：32 位长。希望收到的下一个数据报的序列号。
- TCP 头长：4 位长。表明 TCP 头中包含多少个 32 位字。
- 6 位未用。
- ACK：ACK 位置 1 表明确认号是合法的。如果 ACK 为 0，那么数据报不包含确认信息，确认字段被省略。
- PSH：表示是带有 PUSH 标志的数据。接收方因此请求数据报一到便可送往应用程序而不必等到缓冲区装满时才传送。
- RST：用于复位由于主机崩溃或其他原因而出现的错误的连接。还可以用于拒绝非法的数据报或拒绝连接请求。
- SYN：用于建立连接。
- FIN：用于释放连接。
- 窗口大小：16 位长。窗口大小字段表示在确认了字节之后还可以发送多少个字节。
- 校验和：16 位长。是为了确保高可靠性而设置的。它校验头部、数据和伪 TCP 头部之和。
- 可选项：0 个或多个 32 位字。包括最大 TCP 载荷，窗口比例、选择重发数据报等选项。

2．UDP

（1）概述

UDP 即用户数据报协议，它是一种无连接协议，因此不需要像 TCP 那样通过三次握手来建立一个连接。同时，一个 UDP 应用可同时作为应用的客户或服务器方。由于 UDP 协议并不需要建立一个明确的连接，因此建立 UDP 应用要比建立 TCP 应用简单得多。

UDP 协议从问世至今已经被使用了很多年，虽然其最初的光彩已经被一些类似协议所掩盖，但是在网络质量越来越高的今天，UDP 的应用得到了大大的增强。它比 TCP 协议更为高效，也能更好地解决实时性的问题。如今，包括网络视频会议系统在内的众多的客户/服务器模式的网络应用都使用 UDP 协议。

（2）UDP 数据包头

UDP 数据包头如下图 10.5 所示。

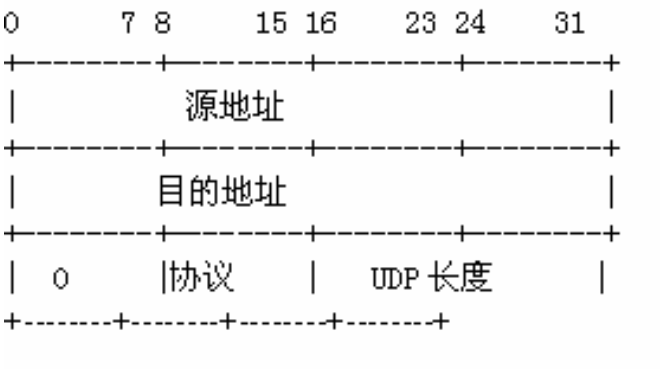


图 10.5 UDP 数据包头

- 源地址、目的地址：16 位长。标识出远端和本地的端口号。
- 数据报的长度是指包括报头和数据部分在内的总的字节数。因为报头的长度是固定的，所以该域主要用来计算可变长度的数据部分（又称为数据负载）。

3．协议的选择

协议的选择应该考虑到以下 3 个方面。

（1）对数据可靠性的要求

对数据要求高可靠性的应用需选择 TCP 协议，如验证、密码字段的传送都是不允许出错的，而对数据的可靠性要求不那么高的应用可选择 UDP 传送。

（2）应用的实时性

由于 TCP 协议在传送过程中要进行三次握手、重传确认等手段来保证数据传输的可靠性。使用 TCP 协议会有较大的时延，因此不适合对实时性要求较高的应用，如 VOIP、视频监控等。相反，UDP 协议则在这些应用中能发挥很好的作用。

（3）网络的可靠性

由于 TCP 协议的提出主要是解决网络的可靠性问题，它通过各种机制来减少错误发生的概率。因此，在网络状况不是很好的情况下需选用 TCP 协议（如在广域网等情况），但是若在网络状况很好的情况下（如局域网等）就不需要再采用 TCP 协议，选择 UDP 协议来减少网络负荷。

10.2 网络基础编程

10.2.1 socket 概述

1．socket 定义

在 Linux 中的网络编程是通过 socket 接口来进行的。人们常说的 socket 接口是一种特殊的 I/O，它也是一种文件描述符。每一个 socket 都用一个半相关描述{协议，本地地址、本地端口}来表示；一个完整的套接字则用一个相关描述{协议，本地地址、本地端口、远程地址、远程端口}。socket 也有一个类似于打开文件的函数调用，该函数返回一个整型的 socket 描述符，随后的连接建立、数据传输等操作都是通过 socket 来实现的。

2．socket 类型

常见的 socket 有 3 种类型如下。

（1）流式 socket（SOCK_STREAM）

流式套接字提供可靠的、面向连接的通信流；它使用 TCP 协议，从而保证了数据传输的正确性和顺序性。

（2）数据报 socket（SOCK_DGRAM）

数据报套接字定义了一种无连接的服务，数据通过相互独立的报文进行传输，是无序的，并且不保证是可靠、无差错的。它使用数据报协议 UDP。

（3）原始 socket

原始套接字允许对底层协议如 IP 或 ICMP 进行直接访问，它功能强大但使用较为不便，主要用于一些协议的开发。

10.2.2 地址及顺序处理

1．地址结构相关处理

（1）数据结构介绍

下面首先介绍两个重要的数据类型：sockaddr 和 sockaddr_in，这两个结构类型都是用来保存 socket 信息的，如下所示：

```
struct sockaddr {
      unsigned short sa_family; /*地址族*/
      char sa_data[14]; /*14 字节的协议地址，包含该 socket 的 IP 地址和端口号。*/
};
struct sockaddr_in {
      short int sa_family; /*地址族*/
      unsigned short int sin_port; /*端口号*/
      struct in_addr sin_addr; /*IP 地址*/
```

```
        unsigned char sin_zero[8]; /*填充 0 以保持与 struct sockaddr 同样大小*/
    };
```

这两个数据类型是等效的，可以相互转化，通常 sockaddr_in 数据类型使用更为方便。在建立 socketadd 或 sockaddr_in 后，就可以对该 socket 进行适当的操作了。

（2）结构字段

表 10.1 列出了该结构 sa_family 字段可选的常见值。

表 10.1

结构定义头文件	#include <netinet/in.h>
Sa_family	AF_INET：IPv4 协议
	AF_INET6：IPv6 协议
	AF_LOCAL：UNIX 域协议
	AF_LINK：链路地址协议
	AF_KEY：密钥套接字（socket）

对了解 sockaddr_in 其他字段的含义非常清楚，具体的设置涉及到其他函数，在后面会有详细讲解。

2．数据存储优先顺序

（1）函数说明

计算机数据存储有两种字节优先顺序：高位字节优先和低位字节优先。Internet 上数据以高位字节优先顺序在网络上传输，因此在有些情况下，需要对这两个字节存储优先顺序进行相互转化。这里用到了四个函数：htons、ntohs、htonl、ntohl。这四个地址分别实现网络字节序和主机字节序的转化，这里的 h 代表 host，n 代表 network，s 代表 short，l 代表 long。通常 16 位的 IP 端口号用 s 代表，而 IP 地址用 l 来代表。

（2）函数格式说明

表 10.2 列出了这 4 个函数的语法格式。

表 10.2　　htons 等函数语法要点

所需头文件	#include <netinet/in.h>
函数原型	uint16_t htons(unit16_t host16bit) uint32_t htonl(unit32_t host32bit) uint16_t ntohs(unit16_t net16bit) uint32_t ntohs(unit32_t net32bit)
函数传入值	host16bit：主机字节序的 16bit 数据
	host32bit：主机字节序的 32bit 数据
	net16bit：网络字节序的 16bit 数据
	net32bit：网络字节序的 32bit 数据

函数返回值	成功：返回要转换的字节序
	出错：–1

注意 调用该函数只是使其得到相应的字节序，用户不需清楚该系统的主机字节序和网络字节序是否真正相等。如果是相同不需要转换的话，该系统的这些函数会定义成空宏。

3．地址格式转化

（1）函数说明

通常用户在表达地址时采用的是点分十进制表示的数值（或者是以冒号分开的十进制 IPv6 地址），而在通常使用的 socket 编程中所使用的则是二进制值，这就需要将这两个数值进行转换。这里在 IPv4 中用到的函数有 inet_aton、inet_addr 和 inet_ntoa，而 IPv4 和 IPv6 兼容的函数有 inet_pton 和 inet_ntop。由于 IPv6 是下一代互联网的标准协议，因此，本书讲解的函数都能够同时兼容 IPv4 和 IPv6，但在具体举例时仍以 IPv4 为例。

这里 inet_pton 函数是将点分十进制地址映射为二进制地址，而 inet_ntop 是将二进制地址映射为点分十进制地址。

（2）函数格式

表 10.3 列出了 inet_pton 函数的语法要点。

表 10.3　　inet_pton 函数语法要点

所需头文件	#include <arpa/inet.h>	
函数原型	int inet_pton(int family,const char *strptr, void *addrptr)	
函数传入值	family	AF_INET：IPv4 协议
		AF_INET6：IPv6 协议
	strptr：要转化的值	
	addrptr：转化后的地址	
函数返回值	成功：0	
	出错：–1	

表 10.4 列出了 inet_ntop 函数的语法要点。

表 10.4　　inet_ntop 函数语法要点

所需头文件	#include <arpa/inet.h>	
函数原型	int inet_ntop(int family, void *addrptr, char *strptr, size_t len)	
函数传入值	family	AF_INET：IPv4 协议
		AF_INET6：IPv6 协议
	addrptr：转化后的地址	
	strptr：要转化的值	

	Len：转化后值的大小
函数返回值	成功：0
	出错：–1

4．名字地址转化

（1）函数说明

通常，人们在使用过程中都不愿意记忆冗长的 IP 地址，尤其到 IPv6 时，地址长度多达 128 位，那时就更加不可能一次次记忆那么长的 IP 地址了。因此，使用主机名将会是很好的选择。在 Linux 中，同样有一些函数可以实现主机名和地址的转化，最为常见的有 gethostbyname、gethostbyaddr、getaddrinfo 等，它们都可以实现 IPv4 和 IPv6 的地址和主机名之间的转化。其中 gethostbyname 是将主机名转化为 IP 地址，gethostbyaddr 则是逆操作，是将 IP 地址转化为主机名，另外 getaddrinfo 还能实现自动识别 IPv4 地址和 IPv6 地址。

gethostbyname 和 gethostbyaddr 都涉及到一个 hostent 的结构体，如下所示：

```
Struct hostent{
      char *h_name;/*正式主机名*/
      char **h_aliases;/*主机别名*/
      int h_addrtype;/*地址类型*/
      int h_length;/*地址长度*/
      char **h_addr_list;/*指向 IPv4 或 IPv6 的地址指针数组*/
}
```

调用该函数后就能返回 hostent 结构体的相关信息。

getaddrinfo 函数涉及到一个 addrinfo 的结构体，如下所示：

```
struct addrinfo{
      int ai_flags;/*AI_PASSIVE,AI_CANONNAME;*/
      int ai_family;/*地址族*/
      int ai_socktype;/*socket 类型*/
      int ai_protocol;/*协议类型*/
      size_t ai_addrlen;/*地址长度*/
      char *ai_canoname;/*主机名*/
      struct sockaddr *ai_addr;/*socket 结构体*/
      struct addrinfo *ai_next;/*下一个指针链表*/
}
```

hostent 结构体而言，addrinfo 结构体包含更多的信息。

（2）函数格式

表 10.5 列出了 gethostbyname 函数的语法要点。

表 10.5　　gethostbyname 函数语法要点

所需头文件	#include <netdb.h>
函数原型	Struct hostent *gethostbyname(const char *hostname)
函数传入值	Hostname：主机名
函数返回值	成功：hostent 类型指针
	出错：−1

调用该函数时可以首先对 addrinfo 结构体中的 h_addrtype 和 h_length 进行设置，若为 IPv4 可设置为 AF_INET 和 4；若为 IPv6 可设置为 AF_INET6 和 16；若不设置则默认为 IPv4 地址类型。

表 10.6 列出了 getaddrinfo 函数的语法要点。

表 10.6　　getaddrinfo 函数语法要点

所需头文件	#include <netdb.h>
函数原型	Int getaddrinfo(const char *hostname,const char *service,const struct addrinfo *hints,struct addrinfo **result)
函数传入值	Hostname：主机名
	service：服务名或十进制的串口号字符串
	hints：服务线索
	result：返回结果
函数返回值	成功：0
	出错：−1

在调用之前，首先要对 hints 服务线索进行设置。它是一个 addrinfo 结构体，表 10.7 列举了该结构体常见的选项值。

表 10.7　　addrinfo 结构体常见选项值

结构体头文件	#include <netdb.h>	
ai_flags	AI_PASSIVE：该套接口是用作被动地打开	
	AI_CANONNAME：通知 getaddrinfo 函数返回主机的名字	
family	AF_INET：IPv4 协议	
	AF_INET6：IPv6 协议	
	AF_UNSPE：IPv4 或 IPv6 均可	
ai_socktype	SOCK_STREAM：字节流套接字 socket（TCP）	
	SOCK_DGRAM：数据报套接字 socket（UDP）	
ai_protocol	IPPROTO_IP：IP 协议	
	IPPROTO_IPV4：IPv4 协议	4
	IPPROTO_IPV6：IPv6 协议	
	IPPROTO_UDP：UDP	
	IPPROTO_TCP：TCP	

（1）通常服务器端在调用 getaddrinfo 之前，ai_flags 设置 AI_PASSIVE，用于 bind 函数（用于端口和地址的绑定后面会讲到），主机名 nodename 通常会设置为 NULL。

（2）客户端调用 getaddrinfo 时，ai_flags 一般不设置 AI_PASSIVE，但是主机名 nodename 和服务名 servname（端口）则应该不为空。

注意

（3）即使不设置 ai_flags 为 AI_PASSIVE，取出的地址也并非不可以被 bind，很多程序中 ai_flags 直接设置为 0，即 3 个标志位都不设置，这种情况下只要 hostname 和 servname 设置的没有问题就可以正确 bind。

（3）使用实例

下面的实例给出了 getaddrinfo 函数用法的示例，在后面小节中会给出 gethostbyname 函数用法的例子。

```
/*getaddrinfo.c*/
#include <stdio.h>
#include <stdlib.h>
#include <errno.h>
#include <string.h>
#include <netdb.h>
#include <sys/types.h>
#include <netinet/in.h>
#include <sys/socket.h>
int main()
{
      struct addrinfo hints,*res=NULL;
      int rc;
      memset(&hints,0,sizeof(hints));
/*设置 addrinfo 结构体中各参数*/
      hints.ai_family=PF_UNSPEC;
      hints.ai_socktype=SOCK_DGRAM;
      hints.ai_protocol=IPPROTO_UDP;
/*调用 getaddinfo 函数*/
      rc=getaddrinfo("127.0.0.1","123",&hints,&res);
      if (rc != 0) {
            perror("getaddrinfo");
            exit(1);
      }
      else
            printf("getaddrinfo success\n");
}
```

运行结果如下所示：

```
[root@(none) tmp]# getaddrinfo success
```

10.2.3 socket 基础编程

（1）函数说明

进行 socket 编程的基本函数有 socket、bind、listen、accept、send、sendto、recv、recvfrom 这几个，其中对于客户端和服务器端以及 TCP 和 UDP 的操作流程都有所区别，这里先对每个函数进行一定的说明，再给出不同情况下使用的流程图。

- socket：该函数用于建立一个 socket 连接，可指定 socket 类型等信息。在建立了 socket 连接之后，可对 socketadd 或 sockaddr_in 进行初始化，以保存所建立的 socket 信息。
- bind：该函数是用于将本地 IP 地址绑定端口号的，若绑定其他地址则不能成功。另外，它主要用于 TCP 的连接，而在 UDP 的连接中则无必要。
- connect：该函数在 TCP 中是用于 bind 的之后的 client 端，用于与服务器端建立连接，而在 UDP 中由于没有了 bind 函数，因此用 connect 有点类似 bind 函数的作用。
- send 和 recv：这两个函数用于接收和发送数据，可以用在 TCP 中，也可以用在 UDP 中。当用在 UDP 时，可以在 connect 函数建立连接之后再用。
- sendto 和 recvfrom：这两个函数的作用与 send 和 recv 函数类型，也可以用在 TCP 和 UDP 中。当用在 TCP 时，后面的几个与地址有关参数不起作用，函数作用等同于 send 和 recv；当用在 UDP 时，可以用在之前没有使用 connect 的情况时，这两个函数可以自动寻找制定地址并进行连接。

服务器端和客户端使用 TCP 协议的流程图如图 10.6 所示。

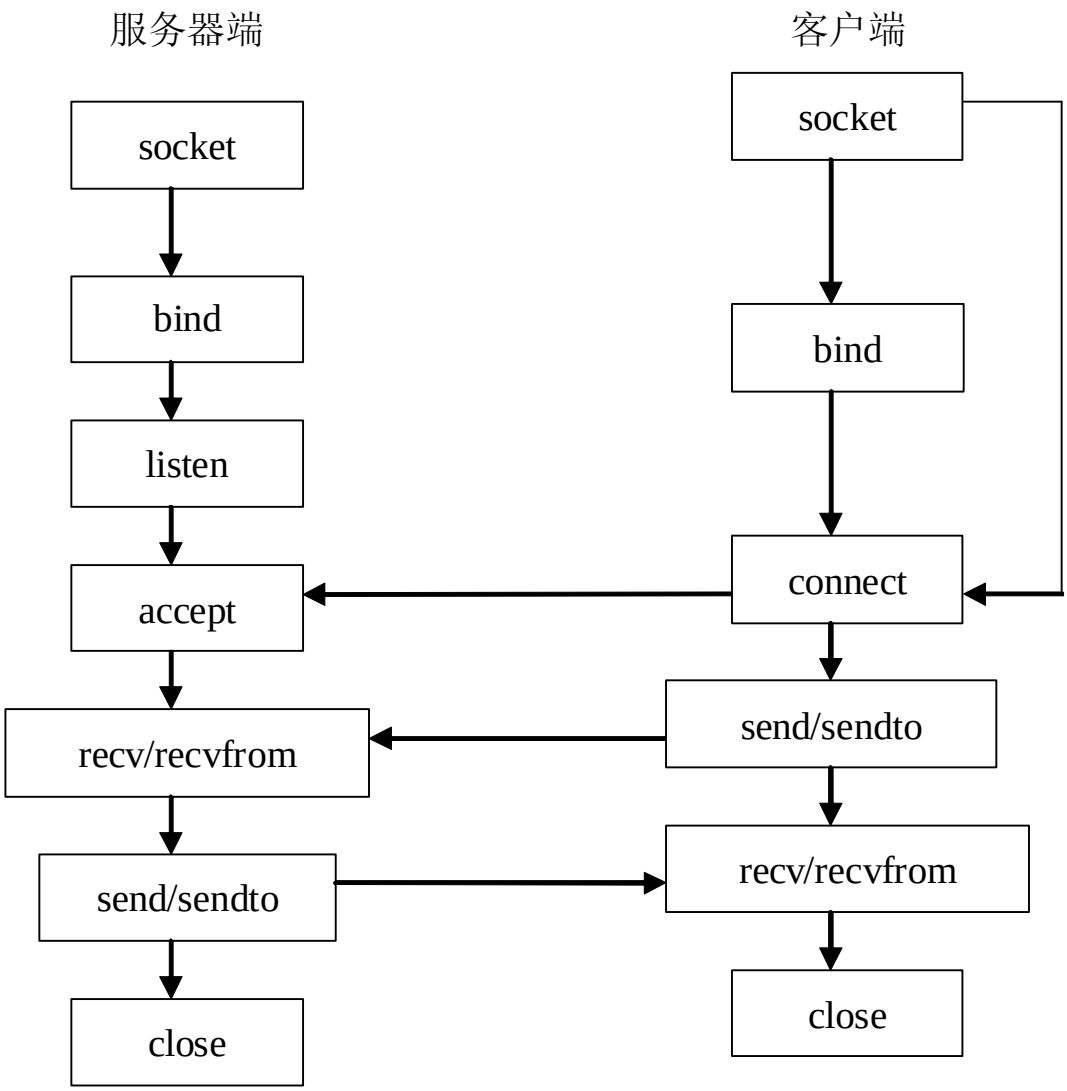


图 10.6　使用 TCP 协议 socket 编程流程图

服务器端和客户端使用 UDP 协议的流程图如图 10.7 所示。

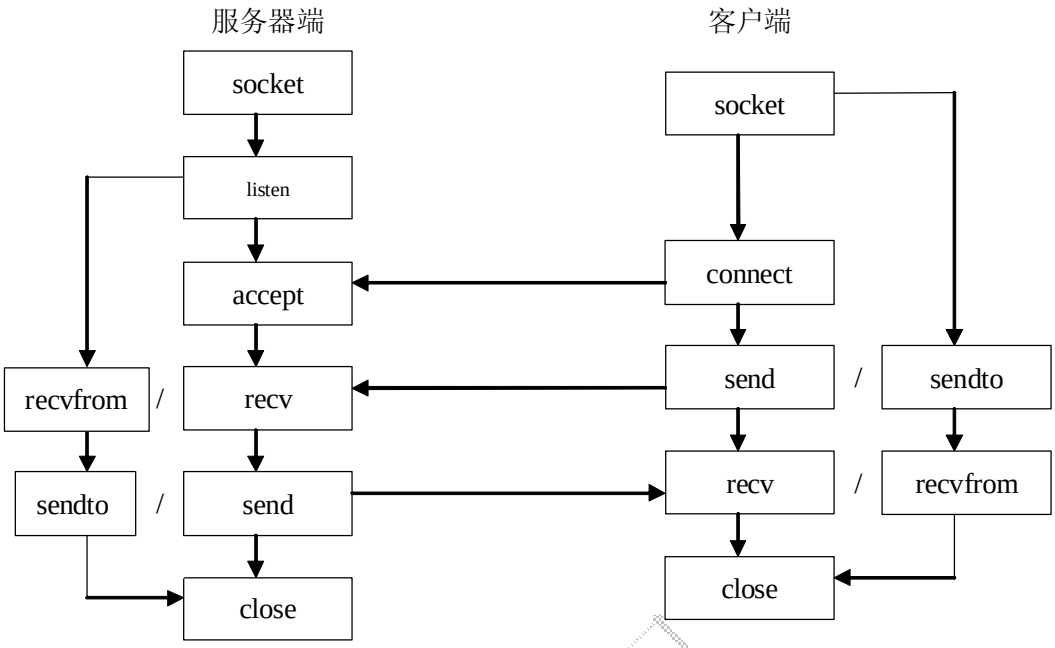


图 10.7　使用 UDP 协议 socket 编程流程图

（2）函数格式

表 10.8 列出了 socket 函数的语法要点。

表 10.8　　socket 函数语法要点

所需头文件	#include <sys/socket.h>	
函数原型	int socket(int family, int type, int protocol)	
函数传入值	family： 协议族	AF_INET：IPv4 协议
		AF_INET6：IPv6 协议
		AF_LOCAL：UNIX 域协议
		AF_ROUTE：路由套接字（socket）
		AF_KEY：密钥套接字（socket）
	type： 套接字类型	SOCK_STREAM：字节流套接字 socket
		SOCK_DGRAM：数据报套接字 socket
		SOCK_RAW：原始套接字 socket
	protoco：0（原始套接字除外）	
函数返回值	成功：非负套接字描述符	
	出错：−1	

表 10.9 列出了 bind 函数的语法要点。

表 10.9　　bind 函数语法要点

所需头文件	#include <sys/socket.h>
函数原型	int bind(int sockfd,struct sockaddr *my_addr, int addrlen)
函数传入值	socktd：套接字描述符
	my_addr：本地地址
	addrlen：地址长度
函数返回值	成功：0
	出错：−1

端口号和地址在 my_addr 中给出了，若不指定地址，则内核随意分配一个临时端口给该应用程序。

表 10.10 列出了 listen 函数的语法要点。

表 10.10 listen 函数语法要点

所需头文件	#include <sys/socket.h>
函数原型	int listen(int sockfd， int backlog)
函数传入值	socktd：套接字描述符
	Backlog：请求队列中允许的最大请求数，大多数系统缺省值为 20
函数返回值	成功：0
	出错：−1

表 10.11 列出了 accept 函数的语法要点。

表 10.11 accept 函数语法要点

所需头文件	#include <sys/socket.h>
函数原型	int accept(int sockfd, struct sockaddr *addr, socklen_t *addrlen)
函数传入值	socktd：套接字描述符
	addr：客户端地址
	addrlen：地址长度
函数返回值	成功：0
	出错：−1

表 10.12 列出了 connect 函数的语法要点。

表 10.12 connect 函数语法要点

所需头文件	#include <sys/socket.h>
函数原型	int connect(int sockfd, struct sockaddr *serv_addr, int addrlen)
函数传入值	socktd：套接字描述符
	serv_addr：服务器端地址
	addrlen：地址长度
函数返回值	成功：0

	出错：−1

表 10.13 列出了 send 函数的语法要点。

表 10.13　　send 函数语法要点

所需头文件	#include <sys/socket.h>
函数原型	int send(int sockfd, const void *msg, int len, int flags)
函数传入值	socktd：套接字描述符
	msg：指向要发送数据的指针
	len：数据长度
	flags：一般为 0
函数返回值	成功：发送的字节数
	出错：−1

表 10.14 列出了 recv 函数的语法要点。

表 10.14　　recv 函数语法要点

所需头文件	#include <sys/socket.h>
函数原型	int recv(int sockfd,void *buf,int len,unsigned int flags)

续表

函数传入值	socktd：套接字描述符
	buf：存放接收数据的缓冲区
	len：数据长度
	flags：一般为 0
函数返回值	成功：接收的字节数
	出错：−1

表 10.15 列出了 sendto 函数的语法要点。

表 10.15　　sendto 函数语法要点

所需头文件	#include <sys/socket.h>
函数原型	int sendto(int sockfd, const void *msg,int len,unsigned int flags,const struct sockaddr *to, int tolen)
函数传入值	socktd：套接字描述符
	msg：指向要发送数据的指针
	len：数据长度
	flags：一般为 0
	to：目地机的 IP 地址和端口号信息

	tolen：地址长度
函数返回值	成功：发送的字节数
	出错：–1

表 10.16 列出了 recvfrom 函数的语法要点。

表 10.16 recvfrom 函数语法要点

所需头文件	#include <sys/socket.h>
函数原型	int recvfrom(int sockfd,void *buf,int len,unsigned int flags,struct sockaddr *from,int *fromlen)
函数传入值	socktd：套接字描述符
	buf：存放接收数据的缓冲区
	len：数据长度
	flags：一般为 0
	from：源机的 IP 地址和端口号信息
	tolen：地址长度
函数返回值	成功：接收的字节数
	出错：–1

（3）使用实例

该实例分为客户端和服务器端，其中服务器端首先建立起 socket，然后调用本地端口的绑定，接着就开始与客户端建立联系，并接收客户端发送的消息。客户端则在建立 socket 之后调用 connect 函数来建立连接。

源代码如下所示：

```
/*server.c*/
#include <sys/types.h>
#include <sys/socket.h>
#include <stdio.h>
#include <stdlib.h>
#include <errno.h>
#include <string.h>
#include <unistd.h>
#include <netinet/in.h>
#define SERVPORT 3333
#define BACKLOG 10
#define MAX_CONNECTED_NO 10
#define MAXDATASIZE 5
```

```
    int main()
    {
          struct sockaddr_in server_sockaddr,client_sockaddr;
          int sin_size,recvbytes;
          int sockfd,client_fd;
          char buf[MAXDATASIZE];
    /*建立 socket 连接*/
          if((sockfd = socket(AF_INET,SOCK_STREAM,0))== -1){
                perror("socket");
                exit(1);
          }
          printf("socket success!,sockfd=%d\n",sockfd);
    /*设置 sockaddr_in 结构体中相关参数*/
          server_sockaddr.sin_family=AF_INET;
          server_sockaddr.sin_port=htons(SERVPORT);
          server_sockaddr.sin_addr.s_addr=INADDR_ANY;
          bzero(&(server_sockaddr.sin_zero),8);
    /*绑定函数 bind*/
          if(bind(sockfd,(struct sockaddr *)&server_sockaddr,sizeof(struct
sockaddr))== -1){
                perror("bind");
                exit(1);
          }
          printf("bind success!\n");
    /*调用 listen 函数*/
          if(listen(sockfd,BACKLOG)== -1){
                perror("listen");
                exit(1);
          }
          printf("listening....\n");
    /*调用 accept 函数，等待客户端的连接*/
          if((client_fd=accept(sockfd,(struct sockaddr *)&client_sockaddr,&sin_
size))== -1){
                perror("accept");
                exit(1);
          }
    /*调用 recv 函数接收客户端的请求*/
          if((recvbytes=recv(client_fd,buf,MAXDATASIZE,0))== -1){
                perror("recv");
```

```
            exit(1);
        }
        printf("received a connection :%s\n",buf);
        close(sockfd);
}

/*client.c*/
#include <stdio.h>
#include <stdlib.h>
#include <errno.h>
#include <string.h>
#include <netdb.h>
#include <sys/types.h>
#include <netinet/in.h>
#include <sys/socket.h>
#define SERVPORT 3333
#define MAXDATASIZE 100
main(int argc,char *argv[]){
        int sockfd,sendbytes;
        char buf[MAXDATASIZE];
        struct hostent *host;
        struct sockaddr_in serv_addr;
        if(argc < 2){
            fprintf(stderr,"Please enter the server's hostname!\n");
            exit(1);
      }
/*地址解析函数*/
        if((host=gethostbyname(argv[1]))==NULL){
            perror("gethostbyname");
            exit(1);
        }
/*创建 socket*/
        if((sockfd=socket(AF_INET,SOCK_STREAM,0))== -1){
            perror("socket");
            exit(1);
        }
/*设置 sockaddr_in 结构体中相关参数*/
        serv_addr.sin_family=AF_INET;
        serv_addr.sin_port=htons(SERVPORT);
```

```
        serv_addr.sin_addr=*((struct in_addr *)host->h_addr);
        bzero(&(serv_addr.sin_zero),8);
    /*调用 connect 函数主动发起对服务器端的连接*/
        if(connect(sockfd,(struct sockaddr *)&serv_addr,\
              sizeof(struct sockaddr))== -1){
              perror("connect");
              exit(1);
        }
    /*发送消息给服务器端*/
        if((sendbytes=send(sockfd,"hello",5,0))== -1){
              perror("send");
              exit(1);
        }
    close(sockfd);
    }
```

在运行时需要先启动服务器端，再启动客户端。这里可以把服务器端下载到开发板上，客户端在宿主机上运行，然后配置双方的 IP 地址，确保在双方可以通信（如使用 ping 命令验证）的情况下运行该程序即可。

```
[root@(none) tmp]# ./server
socket success!,sockfd=3
bind success!
listening....
received a connection :hello
[root@www yul]# ./client 59.64.128.1
```

10.3 网络高级编程

在实际情况中，人们往往遇到多个客户端连接服务器端的情况。由于之前介绍的如 connet、recv、send 都是阻塞性函数，若资源没有准备好，则调用该函数的进程将进入睡眠状态，这样就无法处理 I/O 多路复用的情况了。本节给出了两种解决 I/O 多路复用的解决方法，这两个函数都是之前学过的 fcntl 和 select（请读者先复习第 6 章中的相关内容）。可以看到，由于在 Linux 中把 socket 也作为一种特殊文件描述符，这给用户的处理带来了很大的方便。

1．fcntl

函数 fcntl 针对 socket 编程提供了如下的编程特性。

- 非阻塞 I/O：可将 cmd 设置为 F_SETFL，将 lock 设置为 O_NONBLOCK。
- 信号驱动 I/O：可将 cmd 设置为 F_SETFL，将 lock 设置为 O_ASYNC。

下面是用 fcntl 设置为非阻塞 I/O 的使用实例：

```
/*fcntl.c*/
#include <sys/types.h>
#include <sys/socket.h>
#include <sys/wait.h>
#include <stdio.h>
#include <stdlib.h>
#include <errno.h>
#include <string.h>
#include <sys/un.h>
#include <sys/time.h>
#include <sys/ioctl.h>
#include <unistd.h>
#include <netinet/in.h>
#include <fcntl.h>
#define SERVPORT 3333
#define BACKLOG 10
#define MAX_CONNECTED_NO 10
#define MAXDATASIZE 100
int main()
{
      struct sockaddr_in server_sockaddr,client_sockaddr;
      int sin_size,recvbytes,flags;
      int sockfd,client_fd;
      char buf[MAXDATASIZE];
      if((sockfd = socket(AF_INET,SOCK_STREAM,0))== -1){
            perror("socket");
            exit(1);
      }
      printf("socket success!,sockfd=%d\n",sockfd);
      server_sockaddr.sin_family=AF_INET;
      server_sockaddr.sin_port=htons(SERVPORT);
      server_sockaddr.sin_addr.s_addr=INADDR_ANY;
      bzero(&(server_sockaddr.sin_zero),8);
      if(bind(sockfd,(struct sockaddr *)&server_sockaddr,sizeof(struct
sockaddr))== -1){
            perror("bind");
            exit(1);
```

```
        }
        printf("bind success!\n");
        if(listen(sockfd,BACKLOG)== -1){
              perror("listen");
              exit(1);
        }
        printf("listening....\n");
    /*调用 fcntl 函数设置非阻塞参数*/
        if((flags=fcntl( sockfd, F_SETFL, 0))<0)
                    perror("fcntl F_SETFL");
              flag |= O_NONBLOCK;
              if(fcntl(fd,F_SETEL,flags)<0)
                    perror("fcntl");
        while(1){
              sin_size=sizeof(struct sockaddr_in);
              if((client_fd=accept(sockfd,(structsockaddr*)&client_sockaddr,
&sin_size))== -1){
                    perror("accept");
                    exit(1);
              }
              if((recvbytes=recv(client_fd,buf,MAXDATASIZE,0))== -1){
                    perror("recv");
                    exit(1);
              }
              if(read(client_fd,buf,MAXDATASIZE)<0){
                    perror("read");
                    exit(1);
              }
              printf("received a connection :%s",buf);
        close(client_fd);
        exit(1);
        }/*while*/
    }
```

运行该程序，结果如下所示：

```
    [root@(none) tmp]]# ./fcntl
    socket success!,sockfd=3
    bind success!
    listening....
```

```
accept: Resource temporarily unavailable
```

可以看到，当 accept 的资源不可用时，程序就会自动返回。

2．select

使用 fcntl 函数虽然可以实现非阻塞 I/O 或信号驱动 I/O，但在实际使用时往往会对资源是否准备完毕进行循环测试，这样就大大增加了不必要的 CPU 资源。在这里可以使用 select 函数来解决这个问题，同时，使用 select 函数还可以设置等待的时间，可以说功能更加强大。下面是使用 select 函数的服务器端源代码：

```
/*select_socket.c*/
#include <sys/types.h>
#include <sys/socket.h>
#include <sys/wait.h>
#include <stdio.h>
#include <stdlib.h>
#include <errno.h>
#include <string.h>
#include <sys/un.h>
#include <sys/time.h>
#include <sys/ioctl.h>
#include <unistd.h>
#include <netinet/in.h>
#define SERVPORT 3333
#define BACKLOG 10
#define MAX_CONNECTED_NO 10
#define MAXDATASIZE 100
int main()
{
      struct sockaddr_in server_sockaddr,client_sockaddr;
      int sin_size,recvbytes;
      fd_set readfd;
      fd_set writefd;
      int sockfd,client_fd;
      char buf[MAXDATASIZE];
      if((sockfd = socket(AF_INET,SOCK_STREAM,0))== -1){
            perror("socket");
            exit(1);
      }
      printf("socket success!,sockfd=%d\n",sockfd);
```

```
        server_sockaddr.sin_family=AF_INET;
        server_sockaddr.sin_port=htons(SERVPORT);
        server_sockaddr.sin_addr.s_addr=INADDR_ANY;
        bzero(&(server_sockaddr.sin_zero),8);
        if(bind(sockfd,(struct   sockaddr   *)&server_sockaddr,sizeof(struct
sockaddr))== -1){
              perror("bind");
              exit(1);
        }
        printf("bind success!\n");
        if(listen(sockfd,BACKLOG)== -1){
              perror("listen");
              exit(1);
        }
        printf("listening....\n");
   /*将调用 socket 函数的描述符作为文件描述符*/
        FD_ZERO(&readfd);
        FD_SET(sockfd,&readfd);
        while(1){
        sin_size=sizeof(struct sockaddr_in);
   /*调用 select 函数*/
        if(select(MAX_CONNECTED_NO,&readfd,NULL,NULL,(struct timeval *)0)>0){
              if(FD_ISSET(sockfd,&readfd)>0){
                    if((client_fd=accept(sockfd,(struct sockaddr *)&client_
sockaddr,&sin_size))== -1){
                          perror("accept");
                          exit(1);
                    }
                    if((recvbytes=recv(client_fd,buf,MAXDATASIZE,0))== -1){
                          perror("recv");
                          exit(1);
                    }
                    if(read(client_fd,buf,MAXDATASIZE)<0){
                          perror("read");
                          exit(1);
                    }
                    printf("received a connection :%s",buf);
              }/*if*/
              close(client_fd);
```

```
            }/*select*/
        }/*while*/
    }
```

运行该程序时，可以先启动服务器端，再反复运行客户端程序即可，服务器端运行结果如下所示：

```
[root@(none) tmp]# ./server2
socket success!,sockfd=3
bind success!
listening....
received a connection :hello
received a connection :hello
```

10.4 ping 源码分析

10.4.1 ping 简介

Ping 是网络中应用非常广泛的一个软件，它是基于 ICMP 协议的。下面首先对 ICMP 协议做一简单介绍。

ICMP 是 IP 层的一个协议，它是用来探测主机、路由维护、路由选择和流量控制的。ICMP 报文的最终报宿不是报宿计算机上的一个用户进程，而是那个计算机上的 IP 层软件。也就是说，当一个带有错误信息的 ICMP 报文到达时，IP 软件模块就处理本身问题，而不把这个 ICMP 报文传送给应用程序。

ICMP 报文类型有：回送(ECHO)回答（0）；报宿不可到达（3）；报源断开（4）；重定向（改变路由）（5）；回送（ECHO）请求（8）；数据报超时（11）；数据报参数问题（12）；时间印迹请求（13）；时间印迹回答（14）；信息请求（15）；信息回答（16）；地址掩码请求（17）；地址掩码回答（18）。

虽然每种报文都有不同的格式，但它们开始都有下面三段：

- 一个 8 位整数报文 TYPE（类型）段；
- 一个 8 位 CODE（代码码）段，提供更多的报文类型信息；
- 一个 16 位 CHECKSUM（校验和）段；

此外，报告差错的 ICMP 报文还包含产生问题数据报的网际报头及前 64 位数据。一个 ICMP 回送请求与回送回答报文的格式如表 10.17 所示。

表 10.17 ICMP 回送请求与回送回答报文格式

类型	CODE	校验和[CHECKSUM]
标识符		序列号

数据

10.4.2 ping 源码分析

下面的 ping.c 源码是在 busybox 里实现的源码。在这个完整的 ping.c 代码中有较多选项的部分代码，因此，这里先分析除去选项部分代码的函数实现部分流程，接下来再给出完整的 ping 代码分析。这样，读者就可以看到一个完整协议实现应该考虑到的各个部分 。

1．Ping 代码主体流程

Ping.c 主体流程图如下图 10.8 所示。另外，由于 ping 是 IP 层的协议，因此在建立 socket 时需要使用 SOCK_RAW 选项。在循环等待回应信息处，用户可以指定“-f”洪泛选项，这时就会使用 select 函数来指定在一定的时间内进行回应。

2．主要选项说明

Ping 函数主要有以下几个选项：

- d：调试选项（F_SO_DEBUG）
- f：洪泛选项（F_FLOOD）
- i：等待选项（F_INTERVAL）
- r：路由选项（F_RROUTE）
- l：广播选项（MULTICAST_NOLOOP）

对于这些选项，尤其是路由选项、广播选项和洪泛选项都会有不同的实现代码。

另外，ping 函数可以接受用户使用的 SIGINT 和 SIGALARM 信号来结束程序，它们分别指向了不同的结束代码，请读者阅读下面相关代码。

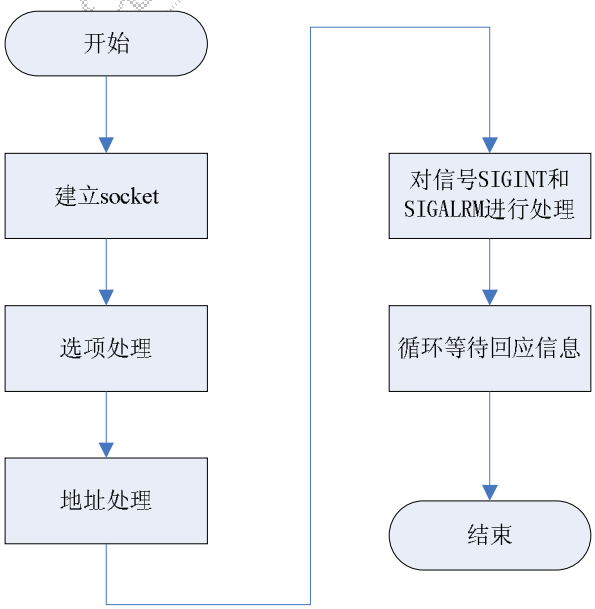


图 10.8 ping 主体流程图

3．源代码及注释

（1）主体代码

ping 代码的主体部分可以四部分，首先是一些头函数及宏定义：

```
#include <sys/param.h>
#include <sys/socket.h>
#include <sys/file.h>
#include <sys/time.h>
#include <sys/signal.h>
#include <netinet/in.h>
#include <netinet/ip.h>
#include <netinet/ip_icmp.h>
#include <arpa/inet.h>
#include <netdb.h>
#include <unistd.h>
#include <stdlib.h>
#include <string.h>
#include <stdio.h>
#include <ctype.h>
#include <errno.h>
#include <getopt.h>
#include <resolv.h>
#define F_FLOOD 0x001
#define F_INTERVAL  0x002
#define F_NUMERIC   0x004
#define F_PINGFILLED    0x008
#define F_QUIET 0x010
#define F_RROUTE    0x020
#define F_SO_DEBUG  0x040
#define F_SO_DONTROUTE  0x080
#define F_VERBOSE   0x100

/* 多播选项 */
int moptions;
#define MULTICAST_NOLOOP    0x001
#define MULTICAST_TTL       0x002
#define MULTICAST_IF        0x004
…
```

接下来的第 2 部分是建立 socket 并处理选项：

```
Int main(int argc, char *argv[])
{
      struct timeval timeout;
      struct hostent *hp;
      struct sockaddr_in *to;
      struct protoent *proto;
      struct in_addr ifaddr;
      int i;
      int ch, fdmask, hold, packlen, preload;
      u_char *datap, *packet;
      char *target, hnamebuf[MAXHOSTNAMELEN];
      u_char ttl, loop;
      int am_i_root;
…
      static char *null = NULL;

      /*__environ = &null;*/
      am_i_root = (getuid()==0);

      /*
      *建立 socket 连接，并且测试是否是 root 用户
      */
      if ((s = socket(AF_INET, SOCK_RAW, IPPROTO_ICMP)) < 0) {
            if (errno==EPERM) {
                  fprintf(stderr, "ping: ping must run as root\n");
            }
            else perror("ping: socket");
            exit(2);
      }
…
      preload = 0;
      datap = &outpack[8 + sizeof(struct timeval)];
      while ((ch = getopt(argc, argv, "I:LRc:dfh:i:l:np:qrs:t:v")) != EOF)
            switch(ch) {
            case 'c':
                  npackets = atoi(optarg);
                  if (npackets <= 0) {
```

```
                        (void)fprintf(stderr,
                            "ping: bad number of packets to transmit.\n");
                        exit(2);
                    }
                    break;
    /*调用选项*/
                case 'd':
                    options |= F_SO_DEBUG;
                    break;
    /*flood 选项*/
                case 'f':
                    if (!am_i_root) {
                        (void)fprintf(stderr,
                            "ping: %s\n", strerror(EPERM));
                        exit(2);
                    }
                    options |= F_FLOOD;
                    setbuf(stdout, NULL);
                    break;
    /*等待选项*/
                case 'i':       /* wait between sending packets */
                    interval = atoi(optarg);
                    if (interval <= 0) {
                        (void)fprintf(stderr,
                            "ping: bad timing interval.\n");
                    exit(2);
                }
                options |= F_INTERVAL;
                break;
            case 'l':
                if (!am_i_root) {
                    (void)fprintf(stderr,
                            "ping: %s\n", strerror(EPERM));
                        exit(2);
                    }
                    preload = atoi(optarg);
                    if (preload < 0) {
                        (void)fprintf(stderr,
                            "ping: bad preload value.\n");
```

```
                    exit(2);
                }
                break;
    …
            default:
                usage();
            }
        argc -= optind;
        argv += optind;

        if (argc != 1)
            usage();
        target = *argv;
```

接下来的第 3 部分是用于获取地址，这里主要使用了 inet_aton 函数，将点分十进制地址转化为二进制地址。当然，作为完整的 ping 程序有较完善的出错处理：

```
        memset(&whereto, 0, sizeof(struct sockaddr));
        to = (struct sockaddr_in *)&whereto;
        to->sin_family = AF_INET;
    /*地址转换函数*/
        if (inet_aton(target, &to->sin_addr)) {
            hostname = target;
        }
        else {
    #if 0
            char * addr = resolve_name(target, 0);
            if (!addr) {
                (void)fprintf(stderr,
                    "ping: unknown host %s\n", target);
                exit(2);
          }
            to->sin_addr.s_addr = inet_addr(addr);
            hostname = target;
    #else
    /*调用 gethostbyname 识别主机名*/
            hp = gethostbyname(target);
            if (!hp) {
                (void)fprintf(stderr,
                    "ping: unknown host %s\n", target);
```

```
                exit(2);
            }
            to->sin_family = hp->h_addrtype;
            if (hp->h_length > (int)sizeof(to->sin_addr)) {
                hp->h_length = sizeof(to->sin_addr);
            }
            memcpy(&to->sin_addr, hp->h_addr, hp->h_length);
             (void)strncpy(hnamebuf, hp->h_name, sizeof(hnamebuf) - 1);
            hostname = hnamebuf;
#endif
        }
```

接下来的一部分主要是对各个选项（如路由、多播）的处理，这里就不做介绍了。再接下来是 ping 函数的最主要部分，就是接收无限循环回应信息，这里主要用到了函数 recvfrom。另外，对用户中断信息也有相应的处理，如下所示：

```
        if (to->sin_family == AF_INET)
            (void)printf("PING %s (%s): %d data bytes\n", hostname,
                inet_ntoa(*(struct in_addr *)&to->sin_addr.s_addr),
                datalen);
        else
            (void)printf("PING %s: %d data bytes\n", hostname, datalen);
/*若程序接收到 SIGINT 或 SIGALRM 信号，调用相关的函数*/
        (void)signal(SIGINT, finish);
        (void)signal(SIGALRM, catcher);
…
/*循环等待客户端的回应信息*/
        for (;;) {
            struct sockaddr_in from;
            register int cc;
            int fromlen;

            if (options & F_FLOOD) {
/*形成 ICMP 回应数据包，在后面会有讲解*/
                pinger();
/*设定等待实践*/
                timeout.tv_sec = 0;
                timeout.tv_usec = 10000;
                fdmask = 1 << s;
/*调用 select 函数*/
```

```
                if (select(s + 1, (fd_set *)&fdmask, (fd_set *)NULL,
                    (fd_set *)NULL, &timeout) < 1)
                    continue;
            }
            fromlen = sizeof(from);
/*接收客户端信息*/
            if ((cc = recvfrom(s, (char *)packet, packlen, 0,
                (struct sockaddr *)&from, &fromlen)) < 0) {
                if (errno == EINTR)
                    continue;
                perror("ping: recvfrom");
                continue;
            }
            pr_pack((char *)packet, cc, &from);
            if (npackets && nreceived >= npackets)
                break;
        }
        finish(0);
        /* NOTREACHED */
        return 0;
}
```

（2）其他函数

下面的函数也是 ping 程序中用到的重要函数。首先 catcher 函数是用户在发送 SIGINT 时调用的函数，在该函数中又调用了 SIGALARM 信号的处理来结束程序。

```
static void
catcher(int ignore)
{
        int waittime;

        (void)ignore;
        pinger();
/*调用 catcher 函数*/
        (void)signal(SIGALRM, catcher);
        if (!npackets || ntransmitted < npackets)
            alarm((u_int)interval);
        else {
            if (nreceived) {
                waittime = 2 * tmax / 1000;
```

```
                    if (!waittime)
                          waittime = 1;
                    if (waittime > MAXWAIT)
                          waittime = MAXWAIT;
            }   else
                    waittime = MAXWAIT;
    /*调用 finish 函数，并设定一定的等待实践*/
                (void)signal(SIGALRM, finish);
                (void)alarm((u_int)waittime);
          }
    }
```

Pinger 函数也是一个非常重要的函数，用于形成 ICMP 回应数据包，其中 ID 是该进程的 ID，数据段中的前 8 字节用于存放时间间隔，从而可以计算 ping 程序从对端返回的往返时延差，这里的数据校验用到了后面定义的 in_cksum 函数。其代码如下所示：

```
    static void
    pinger(void)
    {
          register struct icmphdr *icp;
          register int cc;
          int i;

    /*形成 icmp 信息包，填写 icmphdr 结构体中的各项数据*/
          icp = (struct icmphdr *)outpack;
          icp->icmp_type = ICMP_ECHO;
          icp->icmp_code = 0;
          icp->icmp_cksum = 0;
          icp->icmp_seq = ntransmitted++;
          icp->icmp_id = ident;         /* ID */

          CLR(icp->icmp_seq % mx_dup_ck);

    /*设定等待实践*/
          if (timing)
                (void)gettimeofday((struct timeval *)&outpack[8],
                   (struct timezone *)NULL);

          cc = datalen + 8;         /* skips ICMP portion */
```

```
        /* compute ICMP checksum here */
        icp->icmp_cksum = in_cksum((u_short *)icp, cc);

        i = sendto(s, (char *)outpack, cc, 0, &whereto,
             sizeof(struct sockaddr));

        if (i < 0 || i != cc)  {
             if (i < 0)
                  perror("ping: sendto");
             (void)printf("ping: wrote %s %d chars, ret=%d\n",
                  hostname, cc, i);
        }
        if (!(options & F_QUIET) && options & F_FLOOD)
             (void)write(STDOUT_FILENO, &DOT, 1);
   }
```

pr_pack 是数据包显示函数，分别打印出 IP 数据包部分和 ICMP 回应信息。在规范的程序中通常将数据的显示部分独立出来，这样就可以很好地加强程序的逻辑性和结构性。

```
   void
   pr_pack(char *buf, int cc, struct sockaddr_in *from)
   {
        register struct icmphdr *icp;
        register int i;
        register u_char *cp,*dp;
   /*#if 0*/
        register u_long l;
        register int j;
        static int old_rrlen;
        static char old_rr[MAX_IPOPTLEN];
   /*#endif*/
        struct iphdr *ip;
        struct timeval tv, *tp;
        long triptime = 0;
        int hlen, dupflag;

        (void)gettimeofday(&tv, (struct timezone *)NULL);

        /* 检查 IP 数据包头信息 */
        ip = (struct iphdr *)buf;
```

```
        hlen = ip->ip_hl << 2;
        if (cc < datalen + ICMP_MINLEN) {
              if (options & F_VERBOSE)
                    (void)fprintf(stderr,
                      "ping: packet too short (%d bytes) from %s\n", cc,
                      inet_ntoa(*(struct in_addr *)&from->sin_addr.s_addr));
              return;
        }

        /* ICMP 部分显示 */
        cc -= hlen;
        icp = (struct icmphdr *)(buf + hlen);
        if (icp->icmp_type == ICMP_ECHOREPLY) {
              if (icp->icmp_id != ident)
                    return;           /* 'Twas not our ECHO */
              ++nreceived;
              if (timing) {
#ifndef icmp_data
                    tp = (struct timeval *)(icp + 1);
#else
                    tp = (struct timeval *)icp->icmp_data;
#endif
                    tvsub(&tv, tp);
                    triptime = tv.tv_sec * 10000 + (tv.tv_usec / 100);
                    tsum += triptime;
                    if (triptime < tmin)
                          tmin = triptime;
                    if (triptime > tmax)
                          tmax = triptime;
              }

              if (TST(icp->icmp_seq % mx_dup_ck)) {
                    ++nrepeats;
                    --nreceived;
                    dupflag = 1;
              } else {
                    SET(icp->icmp_seq % mx_dup_ck);
                    dupflag = 0;
              }
```

```
            if (options & F_QUIET)
                  return;

            if (options & F_FLOOD)
                  (void)write(STDOUT_FILENO, &BSPACE, 1);
            else {
                  (void)printf("%d bytes from %s: icmp_seq=%u", cc,
                    inet_ntoa(*(struct in_addr *)&from->sin_addr.s_addr),
                    icp->icmp_seq);
                  (void)printf(" ttl=%d", ip->ip_ttl);
                  if (timing)
                        (void)printf(" time=%ld.%ld ms", triptime/10,
                                    triptime%10);
                  if (dupflag)
                        (void)printf(" (DUP!)");
                  /* check the data */
#ifndef icmp_data
                  cp = ((u_char*)(icp + 1) + 8);
#else
                  cp = (u_char*)icp->icmp_data + 8;
#endif
                  dp = &outpack[8 + sizeof(struct timeval)];
                  for (i = 8; i < datalen; ++i, ++cp, ++dp) {
                        if (*cp != *dp) {
      (void)printf("\nwrong data byte #%d should be 0x%x but was 0x%x",
         i, *dp, *cp);
                              cp = (u_char*)(icp + 1);
                              for (i = 8; i < datalen; ++i, ++cp) {
                                    if ((i % 32) == 8)
                                          (void)printf("\n\t");
                                    (void)printf("%x ", *cp);
                              }
                              break;
                        }
                  }
            }
      } else {
            /* We've got something other than an ECHOREPLY */
```

```
            if (!(options & F_VERBOSE))
                  return;
            (void)printf("%d bytes from %s: ", cc,
            pr_addr(from->sin_addr.s_addr));
            pr_icmph(icp);
      }

/*#if 0*/
      /*显示其他 IP 选项 */
      cp = (u_char *)buf + sizeof(struct iphdr);

      for (; hlen > (int)sizeof(struct iphdr); --hlen, ++cp)
            switch (*cp) {
            case IPOPT_EOL:
                  hlen = 0;
                  break;
            case IPOPT_LSRR:
                  (void)printf("\nLSRR: ");
                  hlen -= 2;
                  j = *++cp;
                  ++cp;
                  if (j > IPOPT_MINOFF)
                        for (;;) {
                              l = *++cp;
                              l = (l<<8) + *++cp;
                              l = (l<<8) + *++cp;
                              l = (l<<8) + *++cp;
                              if (l == 0)
                                    (void)printf("\t0.0.0.0");
                        else
                              (void)printf("\t%s", pr_addr(ntohl(l)));
                        hlen -= 4;
                        j -= 4;
                        if (j <= IPOPT_MINOFF)
                              break;
                        (void)putchar('\n');
                  }
                  break;
            case IPOPT_RR:
```

```
                j = *++cp;     /* get length */
                i = *++cp;     /* and pointer */
                hlen -= 2;
                if (i > j)
                     i = j;
                i -= IPOPT_MINOFF;
                if (i <= 0)
                     continue;
                if (i == old_rrlen
                   && cp == (u_char *)buf + sizeof(struct iphdr) + 2
                   && !memcmp((char *)cp, old_rr, i)
                   && !(options & F_FLOOD)) {
                     (void)printf("\t(same route)");
                     i = ((i + 3) / 4) * 4;
                     hlen -= i;
                     cp += i;
                     break;
                }
                old_rrlen = i;
                memcpy(old_rr, cp, i);
                (void)printf("\nRR: ");
                for (;;) {
                     l = *++cp;
                     l = (l<<8) + *++cp;
                     l = (l<<8) + *++cp;
                     l = (l<<8) + *++cp;
                     if (l == 0)
                          (void)printf("\t0.0.0.0");
                     else
                          (void)printf("\t%s", pr_addr(ntohl(l)));
                     hlen -= 4;
                     i -= 4;
                     if (i <= 0)
                          break;
                     (void)putchar('\n');
                }
                break;
           case IPOPT_NOP:
                (void)printf("\nNOP");
```

```
                    break;
              default:
                    (void)printf("\nunknown option %x", *cp);
                    break;
              }
/*#endif*/
       if (!(options & F_FLOOD)) {
              (void)putchar('\n');
              (void)fflush(stdout);
       }
}
```

in_cksum 是数据校验程序，如下所示：

```
static int
in_cksum(u_short *addr, int len)
{
       register int nleft = len;
       register u_short *w = addr;
       register int sum = 0;
       u_short answer = 0;

       /*这里的算法很简单，就采用 32bit 的加法*/
       while (nleft > 1)  {
              sum += *w++;
              nleft -= 2;
       }

       if (nleft == 1) {
              *(u_char *)(&answer) = *(u_char *)w ;
              sum += answer;
       }

       /*把高 16bit 加到低 16bit 上去*/
       sum = (sum >> 16) + (sum & 0xffff);
       sum += (sum >> 16);
       answer = ~sum;
       return(answer);
}
```

Finish 程序是 ping 程序的结束程序，主要是打印出来一些统计信息，如下所示：

```
    static void
    finish(int ignore)
    {
          (void)ignore;
          (void)signal(SIGINT, SIG_IGN);
          (void)putchar('\n');
          (void)fflush(stdout);
          (void)printf("--- %s ping statistics ---\n", hostname);
          (void)printf("%ld packets transmitted, ", ntransmitted);
          (void)printf("%ld packets received, ", nreceived);
          if (nrepeats)
               (void)printf("+%ld duplicates, ", nrepeats);
          if (ntransmitted)
               if (nreceived > ntransmitted)
                    (void)printf("-- somebody's printing up packets!");
               else
                    (void)printf("%d%% packet loss",
                       (int) (((ntransmitted - nreceived) * 100) /
                       ntransmitted));
          (void)putchar('\n');
          if (nreceived && timing)
               (void)printf("round-trip min/avg/max = %ld.%ld/%lu.%ld/%ld.%ld
ms\n",
                    tmin/10, tmin%10,
                    (tsum / (nreceived + nrepeats))/10,
                    (tsum / (nreceived + nrepeats))%10,
                    tmax/10, tmax%10);

          if (nreceived==0) exit(1);
          exit(0);
    }

    #ifdef notdef
    static char *ttab[] = {
          "Echo Reply",     /* ip + seq + udata */
          "Dest Unreachable",   /* net, host, proto, port, frag, sr + IP */
          "Source Quench",  /* IP */
```

```
        "Redirect",       /* redirect 类型, gateway, + IP  */
        "Echo",
        "Time Exceeded",  /*传输超时*/
        "Parameter Problem",  /* IP 参数问题 */
        "Timestamp",      /* id + seq + three timestamps */
        "Timestamp Reply",    /* " */
        "Info Request",       /* id + sq */
        "Info Reply"      /* " */
};
#endif
```

pr_icmph 函数是用于打印 ICMP 的回应信息，如下所示：

```
static void
pr_icmph(struct icmphdr *icp)
{
      switch(icp->icmp_type) {
/*ICMP 回应*/
      case ICMP_ECHOREPLY:
            (void)printf("Echo Reply\n");
            /* XXX ID + Seq + Data */
            break;
/*ICMP 终点不可达*/
      case ICMP_DEST_UNREACH:
            switch(icp->icmp_code) {
            case ICMP_NET_UNREACH:
                  (void)printf("Destination Net Unreachable\n");
                  break;
            case ICMP_HOST_UNREACH:
                  (void)printf("Destination Host Unreachable\n");
                  break;
            case ICMP_PROT_UNREACH:
                  (void)printf("Destination Protocol Unreachable\n");
                  break;
            …
            default:
                  (void)printf("Dest Unreachable, Unknown Code: %d\n",
                     icp->icmp_code);
                  break;
            }
```

```
            /* Print returned IP header information */
#ifndef icmp_data
            pr_retip((struct iphdr *)(icp + 1));
#else
            pr_retip((struct iphdr *)icp->icmp_data);
#endif
            break;
            …
            default:
                  (void)printf("Redirect, Bad Code: %d", icp->icmp_code);
                  break;
            }
            (void)printf("(New addr: %s)\n",
                        inet_ntoa(icp->icmp_gwaddr));
#ifndef icmp_data
            pr_retip((struct iphdr *)(icp + 1));
#else
            pr_retip((struct iphdr *)icp->icmp_data);
#endif
            break;
      case ICMP_ECHO:
            (void)printf("Echo Request\n");
            /* XXX ID + Seq + Data */
            break;
      case ICMP_TIME_EXCEEDED:
            switch(icp->icmp_code) {
            case ICMP_EXC_TTL:
                  (void)printf("Time to live exceeded\n");
                  break;
            case ICMP_EXC_FRAGTIME:
                  (void)printf("Frag reassembly time exceeded\n");
                  break;
            default:
                  (void)printf("Time exceeded, Bad Code: %d\n",
                     icp->icmp_code);
                  break;
            }
      …
      default:
```

```
                (void)printf("Bad ICMP type: %d\n", icp->icmp_type);
        }
    }
```

pr_iph 函数是用于打印 IP 数据包头选项，如下所示：

```
    static void
    pr_iph(struct iphdr *ip)
    {
        int hlen;
        u_char *cp;

        hlen = ip->ip_hl << 2;
        cp = (u_char *)ip + 20;       /* point to options */
        (void)printf("Vr HL TOS  Len  ID Flg  off TTL Pro cks  Src  Dst Data\n");
        (void)printf(" %1x  %1x  %02x %04x %04x",
           ip->ip_v, ip->ip_hl, ip->ip_tos, ip->ip_len, ip->ip_id);
        (void)printf("   %1x %04x", ((ip->ip_off) & 0xe000) >> 13,
           (ip->ip_off) & 0x1fff);
        (void)printf("  %02x  %02x %04x", ip->ip_ttl, ip->ip_p, ip->ip_sum);
        (void)printf(" %s ", inet_ntoa(*((struct in_addr *) &ip->ip_src)));
        (void)printf(" %s ", inet_ntoa(*((struct in_addr *) &ip->ip_dst)));
        /* dump and option bytes */
        while (hlen-- > 20) {
            (void)printf("%02x", *cp++);
        }
        (void)putchar('\n');
    }
```

pr_addr 是用于将 ascii 主机地址转换为十进制点分形式并打印出来，这里使用的函数是 inet_ntoa，如下所示：

```
    static char *
    pr_addr(u_long l)
    {
        struct hostent *hp;
        static char buf[256];

        if ((options & F_NUMERIC) ||
           !(hp = gethostbyaddr((char *)&l, 4, AF_INET)))
             (void)sprintf(buf, /*sizeof(buf),*/ "%s",inet_ntoa(*(struct
```

```
in_addr *)&l));
          else
               (void)sprintf(buf, /*sizeof(buf),*/ "%s (%s)", hp->h_name,inet_
ntoa(*(struct in_addr *)&l));
          return(buf);
     }
```

Usage 函数是用于显示帮助信息，如下所示：

```
     static void
     usage(void)
     {
          (void)fprintf(stderr,
             "usage: ping [-LRdfnqrv] [-c count] [-i wait] [-l preload]\n\t[-p
pattern] [-s packetsize] [-t ttl] [-I interface address] host\n");
          exit(2);
     }
```

10.5 实验内容——NTP 协议实现

1．实验目的

通过实现 NTP 协议的练习，进一步掌握 Linux 下网络编程，并且提高协议的分析与实现能力，为参与完成综合性项目打下良好的基础。

2．实验内容

Network Time Protocol（NTP）协议是用来使计算机时间同步化的一种协议，它可以使计算机对其服务器或时钟源（如石英钟，GPS 等）做同步化，它可以提供高精确度的时间校正（LAN 上与标准间差小于 1 毫秒，WAN 上几十毫秒），且可用加密确认的方式来防止恶毒的协议攻击。

NTP 提供准确时间，首先要有准确的时间来源，这一时间应该是国际标准时间 UTC。NTP 获得 UTC 的时间来源可以是原子钟、天文台、卫星，也可以从 Internet 上获取。这样就有了准确而可靠的时间源。时间是按 NTP 服务器的等级传播。按照距离外部 UTC 源的远近将所有服务器归入不同的 Stratun（层）中。Stratum-1 在顶层，有外部 UTC 接入，而 Stratum-2 则从 Stratum-1 获取时间，Stratum-3 从 Stratum-2 获取时间，以此类推，但 Stratum 层的总数限制在 15 以内。所有这些服务器在逻辑上形成阶梯式的架构相互连接，而 Stratum-1 的时间服务器是整个系统的基础。

进行网络协议实现时最重要的是了解协议数据格式。NTP 数据包有 48 个字节，其中 NTP 包头 16 字节，时间戳 32 个字节。其协议格式如图 10.9 所示。

2	5	8	16	24	32bit
LI	VN	Mode	Stratum	Poll	Precision
Root Delay					
Root Dispersion					
Reference Identifier					
Reference timestamp (64)					
Originate Timestamp (64)					
Receive Timestamp (64)					
Transmit Timestamp (64)					
Key Identifier (optional) (32)					
Message digest (optional) (128)					

图 10.9　NTP 协议数据格式

其协议字段的含义如下所示。

- LI：跳跃指示器，警告在当月最后一天的最终时刻插入的迫近闰秒（闰秒）。
- VN：版本号。
- Mode：模式。该字段包括以下值：0—预留；1—对称行为；3—客户机；4—服务器；5—广播；6—NTP 控制信息。
- Stratum：对本地时钟级别的整体识别。
- Poll：有符号整数表示连续信息间的最大间隔。
- Precision：有符号整数表示本地时钟精确度。
- Root Delay：有符号固定点序号表示主要参考源的总延迟，很短时间内的位 15 到 16 间的分段点。
- Root Dispersion：无符号固定点序号表示相对于主要参考源的正常差错，很短时间内的位 15 到 16 间的分段点。
- Reference Identifier：识别特殊参考源。
- Originate Timestamp：这是向服务器请求分离客户机的时间，采用 64 位时标格式。
- Receive Timestamp：这是向服务器请求到达客户机的时间，采用 64 位时标格式。
- Transmit Timestamp：这是向客户机答复分离服务器的时间，采用 64 位时标格式。
- Authenticator（Optional）：当实现了 NTP 认证模式时，主要标识符和信息数字域就包括已定义的信息认证代码（MAC）信息。

由于 NTP 协议中涉及到比较多的时间相关的操作，为了简化实现过程，本实验仅要求实现 NTP 协议客户端部分的网络通信模块，也就是构造 NTP 协议字段进行发送和接收，最后与时间相关的操作不需进行处理。

3．实验步骤

（1）画出流程图

简易 NTP 客户端实现流程图如图 10.10 所示。

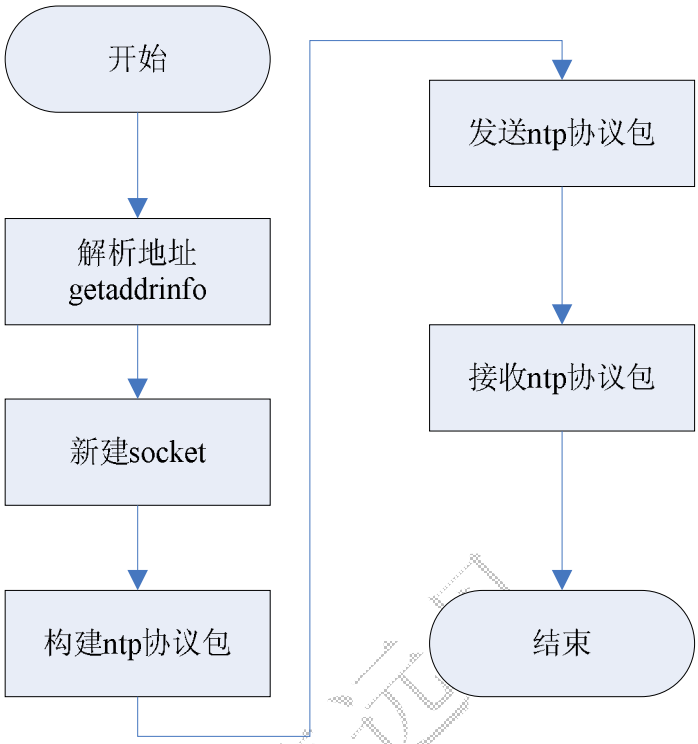


图 10.10　简易 NTP 客户端流程图

（2）编写程序

具体代码如下：

```
#include <sys/socket.h>
#include <sys/wait.h>
#include <stdio.h>
#include <stdlib.h>
#include <errno.h>
#include <string.h>
#include <sys/un.h>
#include <sys/time.h>
#include <sys/ioctl.h>
#include <unistd.h>
#include <netinet/in.h>
#include <string.h>
#include <netdb.h>
struct NTPPacket
{
```

```
    char Leap_Ver_Mode;
    /*client=0*/
    char Startum;
    char Poll;
    char Precision;
    double RootDelay;
    double Dispersion;
    char RefIdentifier[4];
    char RefTimeStamp[8];
    char OriTimeStamp[8];
    char RecvTimeStamp[8];
    char TransTimeStamp[8];
    };

    #define NTPPORT       123
    #define TIMEPORT      37
    #define NTPV1   "NTP/V1"
    #define NTPV2   "NTP/V2"
    #define NTPV3   "NTP/V3"
    #define NTPV4   "NTP/V4"
    #define TIME    "TIME/UDP"
    double SecondBef1970;
    struct sockaddr_in sin;
    struct addrinfo       hints, *res=NULL;
    int rc,sk;
    char Protocol[32];
    /*构建 NTP 协议包*/
    int ConstructPacket(char *Packet)
    {
    char Version=1;
    long SecondFrom1900;
    long Zero=0;
    int Port;
    time_t timer;
    strcpy(Protocol,NTPV1);
    /*判断协议版本*/
    if(strcmp(Protocol,NTPV1)||strcmp(Protocol,NTPV2)||strcmp(Protocol,NTPV3)
||strcmp(Protocol,NTPV4))
    {
```

```
        Port=NTPPORT;
        Version=Protocol[6]-0x30;
        Packet[0]=(Version<<3)|3;      //LI--Version--Mode
        Packet[1]=0;          //Startum
        Packet[2]=0;          //Poll interval
        Packet[3]=0;          //Precision
/*包括 Root delay、Root disperse 和 Ref Indentifier */
        memset(&Packet[4],0,12);
/*包括 Ref timestamp、Ori timastamp 和 Receive Timestamp */
        memset(&Packet[16],0,24);
        time(&timer);
        SecondFrom1900=SecondBef1970+(long)timer;
        SecondFrom1900=htonl(SecondFrom1900);
        memcpy(&Packet[40],&SecondFrom1900,4);
        memcpy(&Packet[44],&Zero,4);
        return 48;
        }
else   //time/udp
{
        Port=TIMEPORT;
        memset(Packet,0,4);
        return 4;
        }
return 0;
}

/*计算从 1900 年到现在一共有多少秒*/
long GetSecondFrom1900(int End)
{
int Ordinal=0;
int Run=0;
long Result;
int i;
        for(i=1900;i<End;i++)
        {
                if(((i%4==0)&&(i%100!=0))||(i%400==0)) Run++;
                else Ordinal++;
        }
        Result=(Run*366+Ordinal*365)*24*3600;
```

```
        return Result;
    }

    /*获取 NTP 时间*/
    long GetNtpTime(int sk,struct addrinfo *res)
    {
    char Content[256];
    int  PacketLen;
    fd_set PendingData;
    struct timeval BlockTime;
    int  FromLen;
    int  Count=0;
    int  result,i;
    int  re;
    struct NTPPacket RetTime;
        PacketLen=ConstructPacket(Content);
        if(!PacketLen)
            return 0;
    /*客户端给服务器端发送 NTP 协议数据包*/
        if((result=sendto(sk,Content,PacketLen,0,res->ai_addr,res->ai_addrl-
en))<0)
            perror("sendto");
        else
            printf("sendto success result=%d \n",result);
        for(i=0;i<5;i++)
        {
            printf("in for\n");
    /*调用 select 函数，并设定超时时间为 1s*/
            FD_ZERO(&PendingData);
            FD_SET(sk, &PendingData);
            BlockTime.tv_sec=1;
            BlockTime.tv_usec=0;
            if(select(sk+1,&PendingData,NULL,NULL,&BlockTime)>0)
            {
                FromLen=sizeof(sin);
    /*接收服务器端的信息*/
                if((Count=recvfrom(sk,Content,256,0,res->ai_addr,&(res->
ai_addrlen)))<0)
                    perror("recvfrom");
```

```
                    else
                          printf("recvfrom success,Count=%d \n",Count);
                    if(Protocol==TIME)
                    {
                          memcpy(RetTime.TransTimeStamp,Content,4);
                          return 1;
                    }
                    else if(Count>=48&&Protocol!=TIME)
                    {
                          RetTime.Leap_Ver_Mode=Content[0];
                          RetTime.Startum=Content[1];
                          RetTime.Poll=Content[2];
                          RetTime.Precision=Content[3];
                          memcpy((void *)&RetTime.RootDelay,&Content[4],4);
                          memcpy((void *)&RetTime.Dispersion,&Content[8],4);
                          memcpy((void *)RetTime.RefIdentifier,&Content[12],4);
                          memcpy((void *)RetTime.RefTimeStamp,&Content[16],8);
                          memcpy((void *)RetTime.OriTimeStamp,&Content[24],8);
                          memcpy((void *)RetTime.RecvTimeStamp,&Content[32],8);
                          memcpy((void *)RetTime.TransTimeStamp,&Content[40],8);
                          return 1;
                    }
              }
        }
        close(sk);
        return 0;
}

int main()
{
        memset(&hints,0,sizeof(hints));
        hints.ai_family=PF_UNSPEC;
        hints.ai_socktype=SOCK_DGRAM;
        hints.ai_protocol=IPPROTO_UDP;
/*调用 getaddrinfo 函数，获取地址信息*/
        rc=getaddrinfo("200.205.253.254","123",&hints,&res);
        if (rc != 0) {
              perror("getaddrinfo");
              return;
```

```
        }
        sk = socket(res->ai_family, res->ai_socktype,res->ai_protocol);
        if (sk <0 ) {
              perror("socket");
        }
        else
        {
              printf("socket success!\n");
        }
    /*调用取得 NTP 时间函数*/
        GetNtpTime(sk,res);
    }
```

本 章 小 结

本章首先概括地讲解了 OSI 分层结构以及 TCP/IP 协议各层的主要功能，介绍了常见的 TCP/IP 协议族，并且重点讲解了网络编程中需要用到的 TCP 和 UDP 协议，为嵌入式 Linux 的网络编程打下良好的基础。

接着本章介绍了 socket 的定义及其类型，并逐个介绍常见的 socket 基础函数，包括地址处理函数、数据存储转换函数，这些函数都是最为常用的函数，要在理解概念的基础上熟练掌握。

接下来介绍的是网络编程中的基础函数，这也是最为常见的几个函数，这里要注意 TCP 和 UDP 在处理过程中的不同。同时，本章还介绍了较为高级的网络编程，包括调用 fcntl 和 select 函数，这两个函数在之前都已经讲解过，但在这里会有特殊的用途。

最后，本章以 ping 函数为例，讲解了常见协议的实现过程，读者可以看到一个成熟的协议是如何实现的。

本章的实验安排了实现一个较为简单的 NTP 客户端程序，主要实现了其中数据收发的主要功能，至于其他时间调整相关的功能在这里就不详细介绍了。

思考与练习

实现一个小型模拟的路由器，就是接收从某个 IP 地址的连接，再把该请求转发到另一个 IP 地址的主机上去。

“黑色经典”系列之《嵌入式 Linux 应用程序开发详解》

第 11 章 嵌入式 Linux 设备驱动开发

本章目标

本书从第 6 章到第 10 章详细讲解了嵌入式 Linux 应用程序的开发，这些都是处于用户空间的内容。本章将进入到 Linux 的内核空间，初步介绍嵌入式 Linux 设备驱动的开发。驱动的开发流程相对于应用程序的开发是全新的，与读者以前的编程习惯完全不同，希望读者能尽快地熟悉现在环境。经过本章的学习，读者将会掌握以下内容。

- Linux 设备驱动的基本概念
- Linux 设备驱动程序的基本功能
- Linux 设备驱动的运作过程
- 常见设备驱动接口函数
- 掌握 LCD 设备驱动程序编写步骤
- 掌握键盘设备驱动程序编写步骤
- 能够独立定制 Linux 服务

11.1 设备驱动概述

11.1.1 设备驱动简介及驱动模块

操作系统是通过各种驱动程序来驾驭硬件设备的，它为用户屏蔽了各种各样的设备，驱动硬件是操作系统最基本的功能，并且提供统一的操作方式。设备驱动程序是内核的一部分，硬件驱动程序是操作系统最基本的组成部分，在 Linux 内核源程序中也占有 60%以上。因此，熟悉驱动的编写是很重要的。

在第 2 章中已经提到过，Linux 内核中采用可加载的模块化设计（LKMs，Loadable Kernel Modules），一般情况下编译的 Linux 内核是支持可插入式模块的，也就是将最基本的核心代码编译在内核中，其他的代码可以选择在内核中，或者编译为内核的模块文件。

常见的驱动程序也是作为内核模块动态加载的，比如声卡驱动和网卡驱动等，而 Linux 最基础的驱动，如 CPU、PCI 总线、TCP/IP 协议、APM（高级电源管理）、VFS 等驱动程序则直接编译在内核文件中。有时也把内核模块叫做驱动程序，只不过驱动的内容不一定是硬件罢了，比如 ext3 文件系统的驱动。因此，加载驱动时就是加载内核模块。

这里，首先列举一些模块相关命令。

- lsmod 列出当前系统中加载的模块，其中左边第一列是模块名，第二列是该模块大小，第三列则是该模块使用的数量。如下所示：

```
[root@www root]# lsmod
Module                  Size  Used by
autofs                 12068   0  (autoclean) (unused)
eepro100               18128   1
iptable_nat            19252   0  (autoclean) (unused)
ip_conntrack           18540   1  (autoclean) [iptable_nat]
iptable_mangle          2272   0  (autoclean) (unused)
iptable_filter          2272   0  (autoclean) (unused)
ip_tables              11936   5  [iptable_nat iptable_mangle iptable_filter]
usb-ohci               19328   0  (unused)
usbcore                54528   1  [usb-ohci]
ext3                   67728   2
jbd                    44480   2  [ext3]
aic7xxx               114704   3
sd_mod                 11584   3
scsi_mod               98512   2  [aic7xxx sd_mod]
```

- rmmod 是用于将当前模块卸载。
- insmod 和 modprobe 是用于加载当前模块，但 insmod 不会自动解决依存关系，而 modprobe 则可以根据模块间依存关系以及/etc/modules.conf 文件中的内容自动插入模块。
- mknod 是用于创建相关模块。

11.1.2 设备文件分类

本书在前面也提到过，Linux 的一个重要特点就是将所有的设备都当做文件进行处理，这一类特殊文件就是设备文件，它们可以使用前面提到的文件、I/O 相关函数进行操作，这样就大大方便了对设备的处理。它通常在/dev 下面存在一个对应的逻辑设备节点，这个节点以文件的形式存在。

Linux 系统的设备文件分为三类：块设备文件、字符设备文件和网络设备文件。

- 块设备文件通常指一些需要以块（如 512 字节）的方式写入的设备，如 IDE 硬盘、SCSI 硬盘、光驱等。
- 字符型设备文件通常指可以直接读写，没有缓冲区的设备，如并口、虚拟控制台等。
- 网络设备文件通常是指网络设备访问的 BSD socket 接口，如网卡等。

对这三种设备文件编写驱动程序时会有一定的区别，本书在后面会有相关内容的讲解。

11.1.3 设备号

设备号是一个数字，它是设备的标志。就如前面所述，一个设备文件（也就是设备节点）可以通过 mknod 命令来创建，其中指定了主设备号和次设备号。主设备号表明某一类设备，一般对应着确定的驱动程序；次设备号一般是用于区分标明不同属性，例如不同的使用方法，不同的位置，不同的操作等，它标志着某个具体的物理设备。高字节为主设备号和底字节为次设备号。例如，在系统中的块设备 IDE 硬盘的主设备号是 3，而多个 IDE 硬盘及其各个分区分别赋予次设备号 1、2、3……

11.1.4 驱动层次结构

Linux 下的设备驱动程序是内核的一部分，运行在内核模式，也就是说设备驱动程序为内核提供了一个 I/O 接口，用户使用这个接口实现对设备的操作。

图 11.1 显示了典型的 Linux 输入/输出系统中各层次结构和功能。

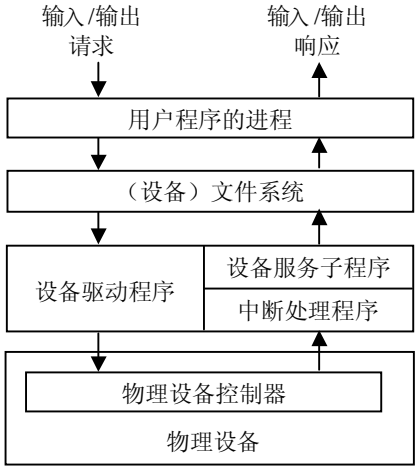


图 11.1 Linux 输入/输出系统层次结构和功能

Linux 设备驱动程序包含中断处理程序和设备服务子程序两部分。

设备服务子程序包含了所有与设备操作相关的处理代码。它从面向用户进程的设备文件系统中接受用户命令，并对设备控制器执行操作。这样，设备驱动程序屏蔽了设备的特殊性，使用户可以像对待文件一样操作设备。

设备控制器需要获得系统服务时有两种方式：查询和中断。因为 Linux 下的设备驱动程序是内核的一部分，在设备查询期间系统不能运行其他代码,查询方式的工作效率比较低，所以只有少数设备如软盘驱动程序采取这种方式，大多设备以中断方式向设备驱动程序发出输入/输出请求。

11.1.5 设备驱动程序与外界的接口

每种类型的驱动程序，不管是字符设备还是块设备都为内核提供相同的调用接口，因此内核能以相同的方式处理不同的设备。Linux 为每种不同类型的设备驱动程序维护相应的数据结构，以便定义统一的接口并实现驱动程序的可装载性和动态性。Linux 设备驱动程序与外界的接口可以分为如下三个部分。

- 驱动程序与操作系统内核的接口：这是通过数据结构 file_operations（在本书后面会有详细介绍）来完成的。
- 驱动程序与系统引导的接口：这部分利用驱动程序对设备进行初始化。
- 驱动程序与设备的接口：这部分描述了驱动程序如何与设备进行交互，这与具体设备密切相关。

它们之间的相互关系如下图 11.2 所示。

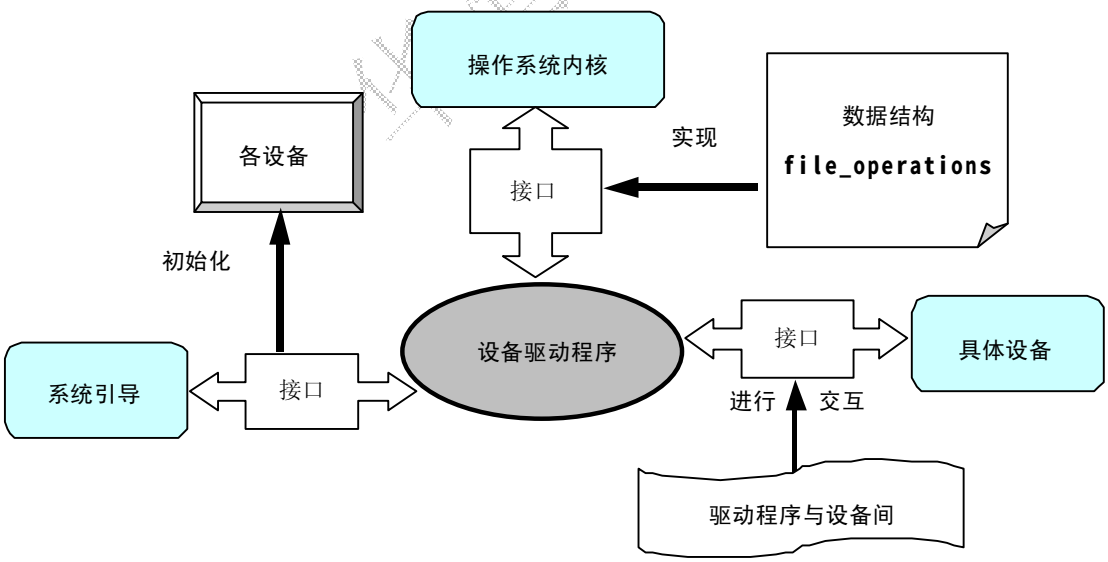


图 11.2 设备驱动程序与外界的接口

11.1.6 设备驱动程序的特点

综上所述，Linux 中的设备驱动程序有如下特点。

（1）内核代码：设备驱动程序是内核的一部分，如果驱动程序出错，则可能导致系统崩溃。

（2）内核接口：设备驱动程序必须为内核或者其子系统提供一个标准接口。比如，一个终端驱动程序必须为内核提供一个文件 I/O 接口；一个 SCSI 设备驱动程序应该为 SCSI 子系统提供一个 SCSI 设备接口，同时 SCSI 子系统也必须为内核提供文件的 I/O 接口及缓冲区。

（3）内核机制和服务：设备驱动程序使用一些标准的内核服务，如内存分配等。

（4）可装载：大多数的 Linux 操作系统设备驱动程序都可以在需要时装载进内核，在不需要时从内核中卸载。

（5）可设置：Linux 操作系统设备驱动程序可以集成为内核的一部分，并可以根据需要把其中的某一部分集成到内核中，这只需要在系统编译时进行相应的设置即可。

（6）动态性：在系统启动且各个设备驱动程序初始化后，驱动程序将维护其控制的设备。如果该设备驱动程序控制的设备不存在也不影响系统的运行，那么此时的设备驱动程序只是多占用了一点系统内存罢了。

11.2 字符设备驱动编写

字符设备驱动编写流程

1．流程说明

在上一节中已经提到，设备驱动程序可以使用模块的方式动态加载到内核中去。加载模块的方式与以往的应用程序开发有很大的不同。以往在开发应用程序时都有一个 main 函数作为程序的入口点，而在驱动开发时却没有 main 函数，模块在调用 insmod 命令时被加载，此时的入口点是 init_module 函数，通常在该函数中完成设备的注册。同样，模块在调用 rmmod 函数时被卸载，此时的入口点是 cleanup_module 函数，在该函数中完成设备的卸载。在设备完成注册加载之后，用户的应用程序就可以对该设备进行一定的操作，如 read、write 等，而驱动程序就是用于实现这些操作，在用户应用程序调用相应入口函数时执行相关的操作，init_module 入口点函数则不需要完成其他如 read、write 之类功能。

上述函数之间的关系如图 11.3 所示。

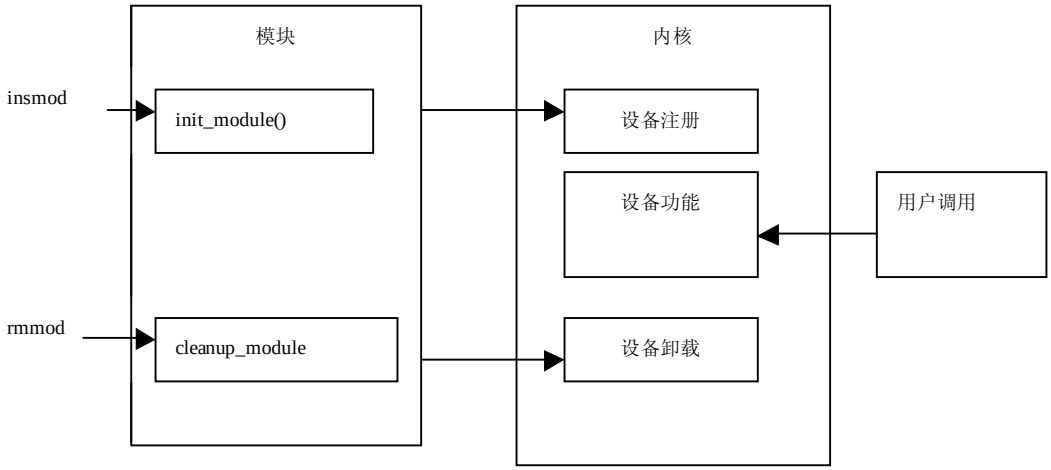


图 11.3　设备驱动程序流程图

2．重要数据结构

用户应用程序调用设备的一些功能是在设备驱动程序中定义的，也就是设备驱动程序的入口点，它是一个在<linux/fs.h>中定义的 struct file 结构，这是一个内核结构，不会出现在用户空间的程序中，它定义了常见文件 I/O 函数的入口。如下所示：

```
struct file_operations {
loff_t (*llseek) (struct file *, loff_t, int);
ssize_t (*read) (struct file *filp, char *buff, size_t count, loff_t *offp);
ssize_t (*write) (struct file *filp, const char *buff, size_t count, loff_t *offp);
int (*readdir) (struct file *, void *, filldir_t);
unsigned int (*poll) (struct file *, struct poll_table_struct *);
int (*ioctl) (struct inode *, struct file *, unsigned int, unsigned
long);
int (*mmap) (struct file *, struct vm_area_struct *);
int (*open) (struct inode *, struct file *);
int (*flush) (struct file *);
int (*release) (struct inode *, struct file *);
int (*fsync) (struct file *, struct dentry *);
int (*fasync) (int, struct file *, int);
int (*check_media_change) (kdev_t dev);
int (*revalidate) (kdev_t dev);
int (*lock) (struct file *, int, struct file_lock *);
};
```

这里定义的很多函数读者在第 6 章中已经见到过了，当时是调用这些函数，而在这里我们将学习如何实现这些函数。当然，每个设备的驱动程序不一定要实现其中所有的函数操作，

若不需要定义实现时，则只需将其设为 NULL 即可。

其中，struct inode 提供了关于设备文件/dev/driver（假设此设备名为 driver）的信息。struct file 提供关于被打开的文件信息，主要用于与文件系统对应的设备驱动程序使用。struct file 较为重要，这里列出了它的定义：

```
struct file {
mode_t f_mode;/*标识文件是否可读或可写，FMODE_READ 或 FMODE_WRITE*/
dev_t f_rdev; /* 用于/dev/tty */
off_t f_pos; /* 当前文件位移 */
unsigned short f_flags; /* 文件标志，如 O_RDONLY、O_NONBLOCK 和 O_SYNC */
unsigned short f_count; /* 打开的文件数目 */
unsigned short f_reada;
struct inode *f_inode; /*指向 inode 的结构指针 */
struct file_operations *f_op;/* 文件索引指针 */
};
```

3．设备驱动程序主要组成

（1）设备注册

设备注册使用函数 register_chrdev，调用该函数后就可以向系统申请主设备号，如果 register_chrdev 操作成功，设备名就会出现在/proc/devices 文件里。

register_chrdev 函数格式如表 11.1 所示。

表 11.1　　register_chrdev 等函数语法要点

所需头文件	#include <linux/fs.h>
函数原型	int register_chrdev(unsigned int major, const char *name,struct file_operations *fops)
函数传入值	major：设备驱动程序向系统申请的主设备号，如果为 0 则系统为此驱动程序动态地分配一个主设备号
	name：设备名
	fops：对各个调用的入口点
函数返回值	成功：如果是动态分配主设备号，此返回所分配的主设备号。且设备名就会出现在/proc/devices 文件里
	出错：–1

（2）设备解除注册

在关闭设备时，通常需要解除原先的设备注册，此时可使用函数 unregister_chrdev，此后该设备就会从/proc/devices 里消失。

unregister_chrdev 函数格式如下表 11.2 所示：

表 11.2　　unregister_chrdev 等函数语法要点

所需头文件	#include <linux/fs.h>

函数原型	int unregister_chrdev(unsigned int major, const char *name)
函数传入值	major：设备的主设备号，必须和注册时的主设备号相同。
	name：设备名
函数返回值	成功：0，且设备名从/proc/devices 文件里消失。
	出错：–1

（3）打开设备

打开设备的接口函数是 open，根据设备的不同，open 函数完成的功能也有所不同，但通常情况下在 open 函数中要完成如下工作。

- 递增计数器。
- 检查特定设备的特殊情况。
- 初始化设备。
- 识别次设备号。

其中递增计数器是用于设备计数的。由于设备在使用时通常会打开较多次数，也可以由不同的进程所使用，所以若有一进程想要关闭该设备，则必须保证其他设备没有使用该设备。因此使用计数器就可以很好地完成这项功能。

这里，实现计数器操作的是用在<linux/module.h>中定义的 3 个宏如下。

- MOD_INC_USE_COUNT：计数器加一。
- MOD_DEC_USE_COUNT：计数器减一。
- MOD_IN_USE：计数器非零时返回真。

另外，当有多个物理设备时，就需要识别次设备号来对各个不同的设备进行不同的操作，在有些驱动程序中并不需要用到。

注意 虽然这是对设备文件执行的第一个操作，但却不是驱动程序一定要声明的操作。若这个函数的入口为 NULL，那么设备的打开操作将永远成功，但系统不会通知驱动程序。

（4）释放设备

释放设备的接口函数是 release。要注意释放设备和关闭设备是完全不同的。当一个进程释放设备时，其他进程还能继续使用该设备，只是该进程暂时停止对该设备的使用；而当一个进程关闭设备时，其他进程必须重新打开此设备才能使用。

释放设备时要完成的工作如下。

- 递减计数器 MOD_DEC_USE_COUNT。
- 在最后一次释放设备操作时关闭设备。

（5）读写设备

读写设备的主要任务就是把内核空间的数据复制到用户空间，或者从用户空间复制到内核空间，也就是将内核空间缓冲区里的数据复制到用户空间的缓冲区中或者相反。这里首先解释一个 read 和 write 函数的入口函数，如表 11.3 所示。

表 11.3　　read、write 函数语法要点

所需头文件	#include <linux/fs.h>
函数原型	ssize_t (*read) (struct file *filp, char *buff, size_t count, loff_t *offp) ssize_t (*write) (struct file *filp, const char *buff, size_t count, loff_t *offp)
函数传入值	filp：文件指针
	buff：指向用户缓冲区
	count：传入的数据长度
	offp：用户在文件中的位置
函数返回值	成功：写入的数据长度

虽然这个过程看起来很简单，但是内核空间地址和应用空间地址是有很大区别的，其中之一就是用户空间的内存是可以被换出的，因此可能会出现页面失效等情况。所以就不能使用诸如 memcpy 之类的函数来完成这样的操作。在这里就要使用 copy_to_user 或 copy_from_user 函数，它们就是用来实现用户空间和内核空间的数据交换的。

copy_to_user 和 copy_from_user 的格式如表 11.4 所示。

表 11.4　　copy_to_user/copy_from_user 函数语法要点

所需头文件	#include <asm/uaccess.h>
函数原型	Unsigned long copy_to_user(void *to, const void *from, unsigned long count) Unsigned long copy_from_user(void *to, const void *from, unsigned long count)
函数传入值	To：数据目的缓冲区
	From：数据源缓冲区
	count：数据长度
函数返回值	成功：写入的数据长度 失败：-EFAULT

要注意，这两个函数不仅实现了用户空间和内核空间的数据转换，而且还会检查用户空间指针的有效性。如果指针无效，那么就不进行复制。

（6）获取内存

在应用程序中获取内存通常使用函数 malloc，但在设备驱动程序中动态开辟内存可以有基于内存地址和基于页面为单位两类。其中，基于内存地址的函数有 kmalloc，注意的是，kmalloc 函数返回的是物理地址，而 malloc 等返回的是线性地址，因此在驱动程序中不能使用 malloc 函数。与 malloc()不同，kmalloc()申请空间有大小限制。长度是 2 的整次方，并且不会对所获取的内存空间清零。

基于页为单位的内存有函数族有如下。

- get_zeroed_page：获得一个已清零页面。
- get_free_page：获得一个或几个连续页面。
- get_dma_pages：获得用于 DMA 传输的页面。

与之相对应的释放内存用也有 kfree 或 free_pages 族。

表 11.5 给出了 kmalloc 函数的语法格式。

表 11.5 kmalloc 函数语法要点

所需头文件	#include <linux/malloc.h>	
函数原型	void *kmalloc(unsigned int len,int flags)	
函数传入值	Len：希望申请的字节数	
	flags	GFP_KERNEL：内核内存的通常分配方法，可能引起睡眠
		GFP_BUFFER：用于管理缓冲区高速缓存
		GFP_ATOMIC：为中断处理程序或其他运行于进程上下文之外的代码分配内存，且不会引起睡眠
		GFP_USER：用户分配内存，可能引起睡眠
		GFP_HIGHUSER：优先高端内存分配
		_GFP_DMA：DMA 数据传输请求内存
		_GFP_HIGHMEN：请求高端内存
函数返回值	成功：写入的数据长度 失败：-EFAULT	

表 11.6 给出了 kfree 函数的语法格式。

表 11.6 kfree 函数语法要点

所需头文件	#include <linux/malloc.h>
函数原型	void kfree(void * obj)
函数传入值	obj：要释放的内存指针
函数返回值	成功：写入的数据长度 失败：-EFAULT

表 11.7 给出了基于页的分配函数 get_free_ page 族函数的语法格式。

表 11.7 get_free_ page 类函数语法要点

所需头文件	#include <linux/malloc.h>
函数原型	unsigned long get_zeroed_page(int flags) unsigned long __get_free_page(int flags) unsigned long __get_free_page(int flags,unsigned long order) unsigned long __get_dma_page(int flags,unsigned long order)
函数传入值	flags：同 kmalloc
	order：要请求的页面数，以 2 为底的对数
函数返回值	成功：写入的数据长度 失败：-EFAULT

表 11.8 给出了基于页的内存释放函数 free_ page 族函数的语法格式。

表 11.8　　free_page 类函数语法要点

所需头文件	#include <linux/malloc.h>
函数原型	unsigned long free_page(unsigned long addr) unsigned long free_page(unsigned long addr)
函数传入值	flags：同 kmalloc
	order：要请求的页面数，以 2 为底的对数
函数返回值	成功：写入的数据长度 失败：-EFAULT

（7）打印信息

就如同在编写用户空间的应用程序，打印信息有时是很好的调试手段，也是在代码中很常用的组成部分。但是与用户空间不同，在内核空间要用函数 printk 而不能用平常的函数 printf。printk 和 printf 很类似，都可以按照一定的格式打印消息，所不同的是，printk 还可以定义打印消息的优先级。

表 11.9 给出了 printk 函数的语法格式。

表 11.9　　printk 类函数语法要点

所需头文件	#include <linux/kernel>	
函数原型	int printk(const char * fmt,…)	
函数传入值	fmt： 日志级别	KERN_EMERG：紧急时间消息
		KERN_ALERT：需要立即采取动作的情况
		KERN_CRIT：临界状态，通常涉及严重的硬件或软件操作失败
		KERN_ERR：错误报告
		KERN_WARNING：对可能出现的问题提出警告
		KERN_NOTICE：有必要进行提示的正常情况
		KERN_INFO：提示性信息
		KERN_DEBUG：调试信息
	…：如 printf 一样的格式说明	
函数返回值	成功：0 失败：–1	

这些不同优先级的信息可以输出到控制台上、/var/log/messages 里。其中，对输出给控制台的信息有一个特定的优先级 console_loglevel。若优先级小于这个整数值时，则消息才能显示到控制台上，否则，消息会显示在/var/log/messages 里。若不加任何优先级选项，则消息默认输出到/var/log/messages 文件中。

注意　要开启 klogd 和 syslogd 服务，消息才能正常输出。

4．proc 文件系统

/proc 文件系统是一个伪文件系统，它是一种内核和内核模块用来向进程发送信息的机

制。这个伪文件系统让用户可以和内核内部数据结构进行交互，获取有关进程的有用信息，在运行时通过改变内核参数改变设置。与其他文件系统不同，/proc 存在于内存之中而不是硬盘上。读者可以通过“ls”查看/proc 文件系统的内容。

表 11.10 列出了/proc 文件系统的主要目录内容。

表 11.10 /proc 文件系统主要目录内容

目 录 名 称	目 录 内 容	目 录 名 称	目 录 内 容
apm	高级电源管理信息	locks	内核锁
cmdline	内核命令行	meminfo	内存信息
cpuinfo	关于 CPU 信息	misc	杂项
devices	设备信息（块设备/字符设备）	modules	加载模块列表
dma	使用的 DMA 通道	mounts	加载的文件系统
filesystems	支持的文件系统	partitions	系统识别的分区表
interrupts	中断的使用	rtc	实时时钟
ioports	I/O 端口的使用	slabinfo Slab	池信息
kcore	内核核心印象	stat	全面统计状态表
kmsg	内核消息	swaps	对换空间的利用情况
ksyms	内核符号表	version	内核版本
loadavg	负载均衡	uptime	系统正常运行时间

除此之外，还有一些是以数字命名的目录，它们是进程目录。系统中当前运行的每一个进程都有对应的一个目录在/proc 下，以进程的 PID 号为目录名，它们是读取进程信息的接口。进程目录的结构如表 11.11 所示。

表 11.11 /proc 中进程目录结构

目 录 名 称	目 录 内 容	目 录 名 称	目 录 内 容
cmdline	命令行参数	cwd	当前工作目录的链接
environ	环境变量值	exe	指向该进程的执行命令文件
fd	一个包含所有文件描述符的目录	maps	内存映像
mem	进程的内存被利用情况	statm	进程内存状态信息
stat	进程状态	root	链接此进程的 root 目录
status	进程当前状态，以可读的方式显示出来		

用户可以使用 cat 命令来查看其中的内容。

可以看到，/proc 文件系统体现了内核及进程运行的内容，在加载模块成功后，读者可以使用查看/proc/device 文件获得相关设备的主设备号。

11.3 LCD 驱动编写实例

11.3.1 LCD 工作原理

S3C2410LCD 控制器用于传输视频数据和产生必要的控制信号，如 VFRAME、VLINE、

VCLK、VM 等。除了控制信号，S3C2410 还有输出视频数据的端口 VD[23:0]，如图 11.4 所示。

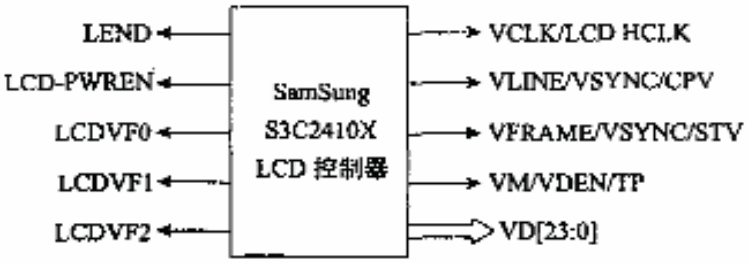


图 11.4　S3C2410 LCD 控制器

（1）寄存器介绍

LCD 的寄存器主要有：LCDCON1 寄存器、LCDCON2 寄存器、LCDCON3 寄存器、LCDC-ON4 寄存器和 LCDCON5 寄存器。

（2）控制流程

LCD 控制器由 REGBANK、LCDCDMA、VIDPRCS 和 LPC3600 组成（如图 11.5 所示）。REGBANK 有 17 个可编程寄存器组和 256*16 的调色板存储器，用来设定 LCD 控制器。LCDCDMA 是一个专用 DMA，自动从帧存储器传输视频数据到 LCD 控制器，用这个特殊的 DMA，视频数据可不经过 CPU 干涉就显示在屏幕上。IDPRCS 接受从 LCDCDMA 来的视频数据并在将其改变到合适数据格式后经 VD[23：0]将之送到 LCD 驱动器，如 4/8 单扫描或 4 双扫描显示模式。TIMEGEN 由可编程逻辑组成，以支持不同 LCD 驱动器的接口时序和速率的不同要求。TIMEGEN 产生 VFRAME、VLINE、VCLK、VM 信号等。

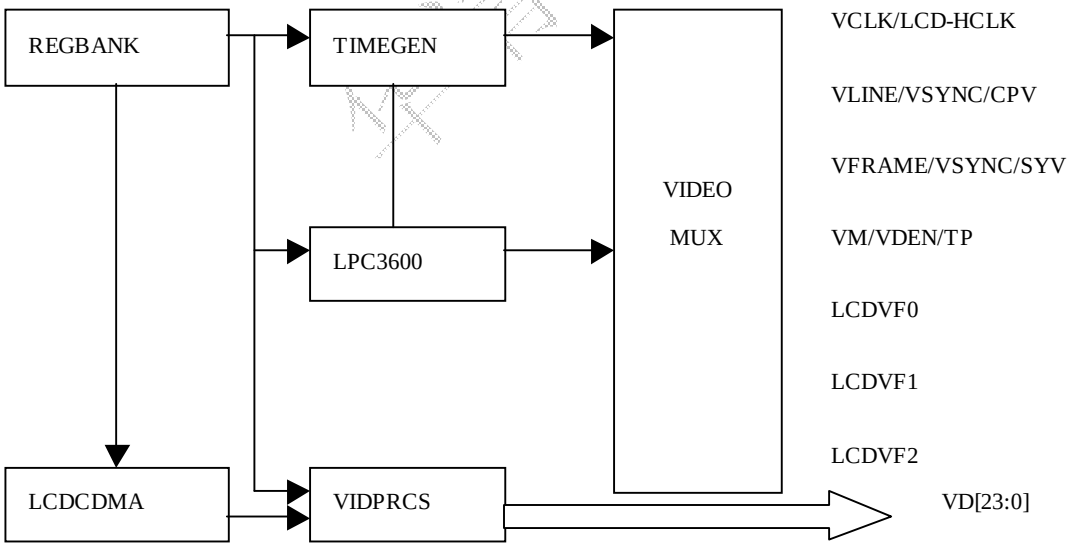


图 11.5　S3C2410 LCD 控制器内部方框图

（3）数据流描述

FIFO 存储器位于 LCDCDMA。当 FIFO 空或部分空时，LCDCDMA 要求从基于突发传输模式的帧存储器中取来数据，存入要显示的图像数据，而这个帧存储器是 LCD 控制器在 RAM 中开辟的一片缓冲区。当这个传送请求被存储控制器中的总线仲裁器接收到后，从系统存储器到内部 FIFO

就会成功传送 4 个字。FIFO 的总大小是 28 个字，其中低位 FIFOL 是 12 个字，高位 FIFOH 是 16 个字。S3C2410 有两个 FIFO 来支持双扫描显示模式。在单扫描模式下，只使用一个 FIFO（FIFOH）。

（4）TFT 控制器操作

S3C2410 支持 STN-LCD 和 TFT-LCD。TIMEGEN 产生 LCD 驱动器的控制信号，如 VSYNC、HSYNC、VCLK、VDEN 和 LEND 等。这些控制信号与 REGBANK 寄存器组中的 LCDCONl/2/3/4/5 寄存器的配置关系相当密切，基于 LCD 控制寄存器中的这些可编程配置，TIMEGEN 产生可编程控制信号来支持不同类型的 LCD 驱动器。

VSYNC 和 HSYNC 脉冲的产生依赖于 LCDCON2/3 寄存器的 HOZVAl 域和 LINEVAL 域的配置。HOZVAL 和 LINEVAL 的值由 LCD 屏的尺寸决定，如下公式：

HOZVAL=水平显示尺寸–1 （1）

LINEVAL=垂直显示尺寸–1 （2）

VCLK 信号的频率取决于 LCDCONl 寄存器中的 CLKVAL 域。VCLK 和 CLKVAL 的关系如下，其中 CLKVAl 的最小值是 0：

VCLK(Hz)＝HCLK/L(CLKVAL+1)×2] （3）

帧频率是 VSYNC 信号的频率，它与 LCDCONl 和 LCDCON2/3/4 寄存器的 VSYNC、VD-PD、VFPD、LINEVAL、HSYNC、HBPD、HFPD、HOZVAL 和 CLKVAL 都有关系。大多数 LCD 驱动器都需要与显示器相匹配的帧频率，帧频率计算公式如下：

FrameRate=1/{[(VSPW+1)+(VBPD+1)+(LINEVAL+1)+(VFPD+1)]*[(HSPW+1)+(HBPD+1)+(HFPD+l)+(HOZVAL+1)*[2*(CLKVAL+1)/(HCLK)]}

11.3.2 LCD 驱动实例

LCD 驱动代码如下所示：

```
#include <linux/kernel.h>
#include <linux/init. h>
#include <linux/module. h>
#include <linuxJfs. h>
#include <linux/delay.h>
#include <asm/fcntl.h>
#include <asm/unistd.h>
#include <asm/io.h>
#include <asm/uaccess.h>
#include "lcdexp.h"
static unsigned char*, lcd base;
/* LCD 配置函数 */
static void setup_lcd(void)
{
/*在设置 LCD 寄存器之前关闭 LCD*/
     LCDEN[12] = 0;
```

```
        SYSCON1 &= ~0x00001000;
    /* 设置 LCD 控制寄存器
        * Video Buffer Size[0:12]: 320'240'12 / 128 = 0xlclf
        * Line Length[13:18]: 320 / 16 -1 = 0x13
        * Pixel Prescale[19:24]: 0x01
        * ACPmscale[25:29]: 0x13
        * GSEN[30]: =1,Enables gray scale output to LCD
        * GSMD[31]: =1,4 bpp ( 16-gray scale )
        */
        LCDCON = 0xe60f7clf;
        /* 设置 LCD Palette 寄存器 */
        PALLSW = 0x76543210;
        PALMSW = 0xfedcba98;
        /*
        * 设置 LCD frame buffer Sets 的起始位置
        * 这样, frame buffer 就从 0xc0000000 起始
        */
        FBADDR = 0xc;
    /*使能 LCD 使之改变 LCD 的配置*/
        LCDEN[12] = 1;
        SYSCON1 = 0x00001000;
        return;
    }

    /*在 LCD 中画一个点
     * x,y: 亮点的坐标
     * color:点的颜色
     */
    static void lcd__pixel_set(int x, int y, COLOR color)
    {
          unsigned char* fb_ptr;
          COLOR pure_color = OxO000;
          /* if the dot is out of the LCD, return */
          If (x<0 II x>=32011 y<0 II y>=240){
    /*计算点的地址 */
          fb_.ptr =  lcd base + (x+y*320)*12/8;
          /*把版面上的点映射到帧缓冲中（frame buffer）*/
          if (x & 0xl ) (
                pure_color = ( color & 0x000f ) < < 4;
```

```
                *fb_ptr &= 0x0f;
                *fb_ptr I= pure_color;
                pure_color = ( color & 0x0ff0 ) >> 4;
                *(fb_ptr+l) = 0xff & pure_color;
        } else {
                pure_color = color & 0x00ff;
                *fb_ptr = 0xff & pure_color;
                pure_color = (color & 0x0f00 ) >> 8;
                *(fb_ptr+l) &= OxfO;
                *(fb_ptr+l) |= pure_color;
                }
        return;
}
/*
        把所有 LCD 图片清零
*/
void clear_lcd(void)
{
        int x;
        int y;
        for (y=0;y<240; y++) {
              for (x=0; x<320; x++) {
                    //lcd_disp.x = x;
                    //lcd_disp.y = y;
                    Lcd_plxel_set(x,y,0x0000);
               }
        }
        Return;
}

/* (start x, start_y): 矩形最左边的坐标
 * (end_x, end_y): 矩形最右边的坐标
 */
static void draw_rectangle(int start_x,int start_y,int end_x,int end_y,COLOR
color)
{
        draw__vline(start_x, start_y, end_y, color);
        draw_vline(end_x, start_y, end_y, color);
        draw_hline(start_x, end_x, start_y, color),
```

```
        draw_hline(start_x, end_x, end_y, color);
        return;
    }
    /*
     * (start x, start_y):矩形最上边的坐标
     * (end_x, end_y): 矩形最下边的坐标
     */
    static void draw_full_rectangle(int start_x, int start_y,int end_x,int
end_y,COLOR color)
    {
        int i = 0;
        int tmp= 0;
        tmp= end_x - start_x;
        for ( i=0;i<tmp;++i ) {
            draw_vline(start_x+i, start_y,end_y, color);
        }
        return;
    }
    /*显示一个 ascii 符号
     * x,y: 符号起始坐标
     * codes: 要显示的字节数
    */
    static void write_en(int x, int y, unsigned char* codes, COLOR color)
    {
        inti = 0;
        /* total 16 bytes codes*/
        for (i=0;i<16;++i) {
            intj = 0;
        x += 8;
        for ( j=0;j<8;++j ){
            --x;
            if ((codes[i]>>j) 8, 0x1 ) {
                lcd_pixel_set(x, y, color);
            }
        }
        /*移到下一行，x 轴不变，y 轴加一*/
        ++y;
        }
        return;
```

```
    }
    /*显示一个中文字符
     * x,y: 字符起始点
     * codes: 要显示的字节数
     * color: 要显示的字符颜色
     */
    static void write_cn(int x, iht y, unsigned char* codes, COLOR color)
    {
           int i;
           /* total 2'16 bytes codes */
           for(i=0;i< 16;i++) {
                 int j = 0;
                 for (j=0;j<2;++j) {
                       int k = 0;
                       x += 8(j+1);
                       for ( k=O;k<8;++k ){
                             --x;
                             if ( ( codes[2*i+j] >> k) &0xl ) {
                                    Icd_pixel_set(x,y,color);
                             }
                       }
                 }
                 x-= 8;
                ++y;
           }
           return;
    }
    static int lcdexp_open(struct inode *node, struct file *file)
    {
           return 0;
    }
    static int lcdexp_read(struct file *file, char *buff, size_t count, Ioff_t *offp)
    {
           return 0;
    }
    static int lcdexp_write(struct file *file, const char *buff, size_t count,
Ioff_t *offp)
    {
           return 0;
```

```
    }
    /*lcd ioctl 驱动函数，分类处理 lcd 的各项动作，在每种情况下都会调用前述的关键函数*/
    static int lcdex_ioctl(struct inode *inode, struct file *file, unsigned int
cmd, unsigned long
    arg)
    {
           switch ( cmd ) {
           case LCD_Clear:/*lcd 清屏*/
           {
                 clear_lcd();
                 break;
           }
           case LCD_Pixel_Set /*lcd 象素设置*/
           {
                 struct lcd_display    pixel_display;
                 if(copy_from_user(&pixel_display,(struct Icd_display*)arg,sizeof
(struct lcd_display)))     {
                       printk("copy_frorn_user error!\n"),
                       return -1;
                 }
                 lcd_pixel_set(pixeLdisplay.xl, pixel_display.yl, pixel_display.
color);
                 break;
           }
           case LCD_Big_Pixel_Set:/*lcd 高级象素设置*/
           {
                struct lcd_display   b_pixel_display;
                if(copy_from_user(&b_pixel_display,(struct Icd_display*)arg,sizeof
(struct lcd_display)))                {
                printk("copy_from_user error!\n");
                return -1;
           }
           lcd_big_pixel_set(b_pixel_display.xl, b_pixel display.y1, b_pixel_display.
color);
           break;
        }
        case LCD_Draw_Vline:/*lcd 中显示水平线*/
        {
              struct lcd_display    vline_display;
```

```
             if(copy_from_user(&vline_display,(struct Icd_display*)arg,sizeof
(struct lcd_display)))
             {
                  printk("copy_from_user error!\n");
                  return -1;
             }
             draw_vline(vline_display.xl, vline_display.yl, vline_display.y2,
vline_display.color);
        }
        case LCD_Draw_HLine:/*lcd 中显示垂直线*/
        {
             struct lcddisplay    hline_display;
             if ( copy_from_user(Ehline_display,(struct Icd_display*)arg,sizeof
(struct lcd_display)))     {
                   printk("copy_from_user error!\n");
                   return -1;
             }
        draw_hline(hline_display.xl,   hline   display.x2,   hline_display.yl,
hline_display.color);
        break;
        }
        Case LCD_Draw_Vdashed:/*lcd 中显示水平随意线*/
        {
             struct lcd-_display    vdashed display;
             if(copy_from_user(&vdashed_display,(structlcd_display*)arg,sizeof
(struct lcd_display)))     {
                  printk("copy_from_user error!\n");
                  return -1;
             }
             draw hdashed(hdashed-display.xl, hdashed_display.x2, hdashed_
display.yl,
    vdashed_display.color);
             break;
        }
        Case LCD_Draw_HDdashed:/*lcd 中显示垂直随意线*/
        {
             struct lcd_display    hdashed display;
             if(copy_from_user(&hdashed_display,(structlcd_display*)arg,sizeof
(struct lcd_display)))     {
```

```
                printk("copy_from_user error!\n");
                return -1;
            }
            draw hdashed(hdashed-display.xl, hdashed_display.x2, hdashed_
display.yl,
    vdashed_display.color);
            break;
        }
        case LCD_Draw_Rectangle:/*lcd 中显示矩阵*/
        {
            struct/cd-display    rect display;
             if ( copy_from_user(&rect_display,(struct Icd_display*)arg,sizeof
(struct lcd_display))) {
                   printk("copy_from_user error!\n");
                   return -1;
             }
             draw_rectangle(rect_display.xl,rect_display.yl,rect_display.x2,
rect_display.y2,rect_display.color);
        break;
        }
        case LCD_Draw_Full_Rectangle:/*lcd 中显示填充矩阵*/
        {
            Struct xlcd_display    frect_display;
            if ( copy_from_user(&frect_display,(struct Icd_display*)arg,sizeof
(struct lcd_display)))                {
                 printk("copy_from_user error!\n");
            }
            draw_full_rectangle(frect_display.xl,
frect_display.yl,frect_display.x2,frect_display.y2, rect_display.color);
            break;
        }
        case LCD Write_EN:/*lcd 英文显示*/
        {
            Struct lcd_display    en_display;
             if ( copy_from_user(&en_display,(struct Icd_display*)arg,sizeof
(struct lcd_display)))         {
                   printk("copy_from_user error!\n");
             }
             return -1;
```

```
            write_en(en_display.xl, en_.display.yl, en_display.buf, en_display.
color);
            break;
        }
        case LCD Write_CN:/*lcd 中文显示*/
        {
            struct lcd_display    cn_display;
            if ( copy_from_user(&cn_display,(struct Icd_display*)arg,sizeof
(struct lcd_display))) {
                printk("copy_ffom_user errod\n");
                return -1;
            }
            write_cn(cn_display.xl, cn_display.yl, cn_display.buf, cn_display.
color);
            break;
    }
        default:
            printk("unknown cmd\n");
            break;
    return 0;
    }

    static struct file_operations lcdexp_fops = {
          open:             lcdexp_open,
          read:              lcdexp_read,
          ioctl:            lcdexp_joctl,
          write:             lcdexp_write,
          release:          lcdexp_release,
    };
    int lcdexp_init(void)
    {
         int result;
         lcd base = (unsigned char*)0xc0000000;
         result = register_chrdev(DEV_MA)OR,"lcdexp",&lcdexp_fops);
         if ( result < 0 ) {
              printk( KERN_INFO "lcdexp:register Icdexp failed !\n'
              return result;
         setup_lcd();
         for ( i=0;i<320*240*12/8;i++ )
```

```
            lcd base++ = 0x77;
        _lcd_base = (unsigned char*)0xc0000000;
        printk("LCD  .............  support.\n");
return 0;
}
static void _exit lcdexp_exit(void)
{
        /* clear LCD */
        unregister_chrdev(DEV_MA.lOR,"lcdexp");
}
module_init(lcdexp_init);
module_exit(lcclexp_exit);
```

11.4 块设备驱动编写

11.4.1 块设备驱动程序描述符

块设备文件通常指一些需要以块（如 512 字节）的方式写入的设备，如 IDE 硬盘、SCSI 硬盘、光驱等。它的驱动程序的编写过程与字符型设备驱动程序的编写有很大的区别。块设备驱动程序描述符是一个包含在<linux/blkdev.h>中的 blk_dev_struct 类型的数据结构，其定义如下所示：

```
struct blk_dev_struct {
        request_queue_t  request_queue;
        queue_proc   *queue;
        void  *date;
};
```

在这个结构中，请求队列 request_queue 是主体，包含了初始化之后的 I/O 请求队列。对于函数指针 queue，当其为非 0 时，就调用这个函数来找到具体设备的请求队列，这是为考虑具有同一主设备号的多种同类设备而设的一个域，该指针也在初始化时就设置好。指针 data 是辅助 queue 函数找到特定设备的请求队列，保存一些私有的数据。

所有块设备的描述符都存放在 blk_dev 表 struct blk_dev_struct blk_dev[MAX_BLKDEV]中；每个块设备都对应着数组中的一项，可以使用主设备号进行检索。每当用户进程对一个块设备发出一个读写请求时，首先调用块设备所公用的函数 generic_file_read()和 generic_file_write()。如果数据存在在缓冲区中或缓冲区还可以存放数据，那么就同缓冲区进行数据交换。否则，系统会将相应的请求队列结构添加到其对应项的 blk_dev_struct 中，如下图 11.6 所示。

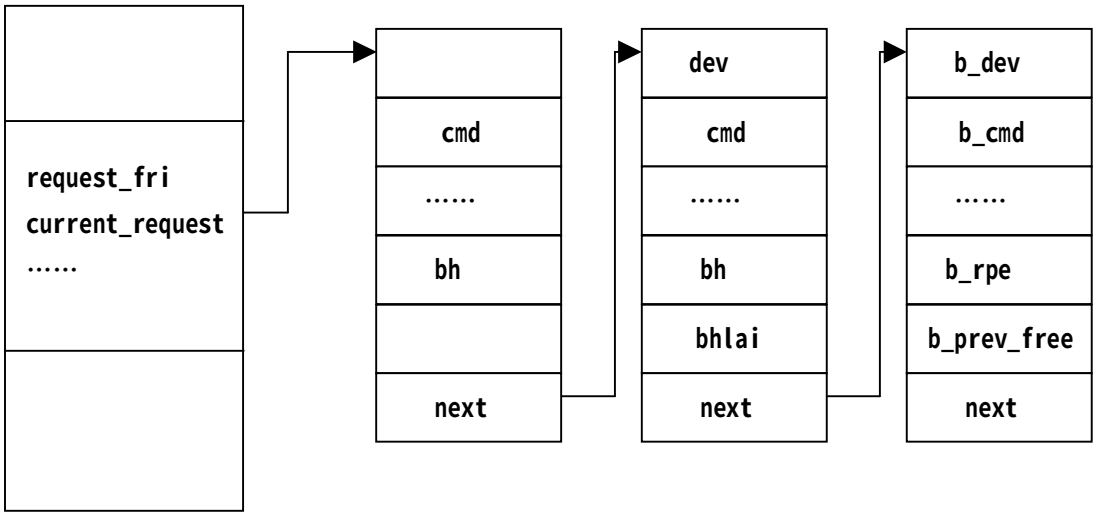


图 11.6　块设备请求队列

11.4.2　块设备驱动编写流程

1．流程说明

块设备驱动程序的编写流程同字符设备驱动程序的编写流程很类似，也包括了注册和使用两部分。但与字符驱动设备所不同的是，块设备驱动程序包括一个 request 请求队列。它是当内核安排一次数据传输时在列表中的一个请求队列，用以最大化系统性能为原则进行排序。在后面的读写操作时会详细讲解这个函数，下图 11.7 给出了块设备驱动程序的流程图，请读者注意与字符设备驱动程序的区别。

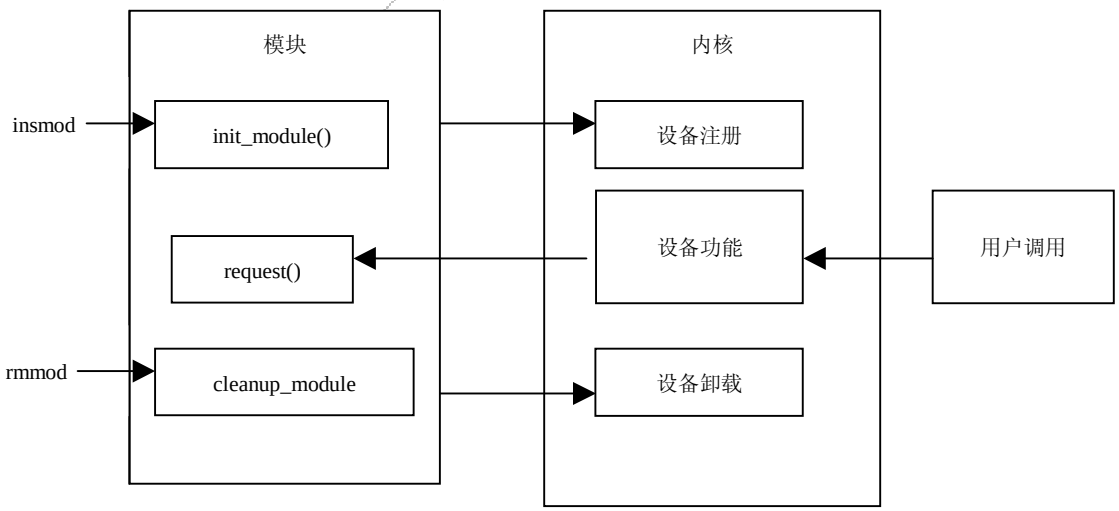


图 11.7　块设备驱动程序流程图

2．重要数据结构

由于块设备驱动程序的绝大部分都与设备无关的，故内核的开发者通过把大部分相同的代码放在一个头文件<linux/blk.h>中来简化驱动程序的代码。从而每个块设备驱动程序都必须包含这个头文件。先给出块设备驱动程序要用到的数据结构定义：

```
struct device_struct {
       const char *name;
       struct file_operations *chops;
};
static struct device_struct blkdevs[MAX_BLKDEV];
struct sbull_dev {
       void **data;
       int quantum; // the current quantum size
       int qset; // the current array size
       unsigned long size;
       unsigned int access_key; // used by sbulluid and sbullpriv
       unsigned int usage; // lock the device while using it
       unsigned int new_msg;
       struct sbull_dev *next; // next listitem
};
```

与字符设备驱动程序一样，块设备驱动程序也包含一个 file_operation 结构，其结构定义一般如下所示：

```
struct file_operation blk_fops = {
       NULL，//seek
       block_read，//内核函数
       block_write，//内核函数
       NULL，//readdir
       NULL，//poll
       sbull_ioctl，// ioctl
       NULL，//mmap
       sbull_open，//open
       NULL，//flush
       sbull_release，//release
       block_fsync，//内核函数
       NULL，//fasync
       sbull_check_media_change，//check media change
       NULL，//revalidate
       NULL，//lock
```

```
};
```

从上面结构中可以看出，所有的块驱动程序都调用内核函数 block_read()、block_write()，block_fsync()函数，所以在块设备驱动程序入口中不包含这些函数，只需包括 ioctl()、open()和 release()函数即可。

（1）设备初始化

块设备的初始化过程要比字符设备复杂，它既需要像字符设备一样在引导内核时完成一定的工作，还需要在内核编译时增加一些内容。块设备驱动程序初始化时，由驱动程序的 init()完成。

块设备驱动程序初始化的工作主要包括：

- 检查硬件是否存在；
- 登记主设备号；
- 将 fops 结构的指针传递给内核；
- 利用 register_blkdev()函数对设备进行注册：

```
if(register_blkdev(sbull_MAJOR, “sbull”, &sbull_fops)) {
      printk( “Registering block device major: %d failed\n”, sbull_MAJOR);
      return-EIO;
};
```

- 将 request()函数的地址传递给内核：

```
blk_dev[sbull_MAJOR].request_fn = DEVICE_REQUEST;
```

- 将块设备驱动程序的数据容量传递给缓冲区：

```
#define  sbull_HARDS_SIZE  512
#define  sbull_BLOCK_SIZE  1024
static  int  sbull_hard = sbull_HARDS_SIZE;
static  int  sbull_soft = sbull_BLOCK_SIZE;
hardsect_size[sbull_MAJOR] = &sbull_hard;
blksize_size[sbull_MAJOR] = &sbull_soft;
```

在块设备驱动程序内核编译时，应把下列宏加到 blk.h 文件中：

```
#define  MAJOR_NR  sbull_MAJOR
#define  DEVICE_NAME  “sbull”
#define  DEVICE_REQUEST  sbull_request
#define  DEVICE_NR(device)  (MINOR(device))
#define  DEVICE_ON(device)
#define  DEVICE_OFF(device)
```

（2）request 操作

Request 操作涉及一个重要的数据结构如下。

```
struct request {
      kdev_t rq_dev;
```

```
        int cmd;    // 读或写
        int errors;
        unsigned long sector;
        char *buffer;
        struct request *next;
};
```

对于具体的块设备，函数指针 request_fn 当然是不同的。块设备的读写操作都是由 request()函数完成。所有的读写请求都存储在 request 结构的链表中。request()函数利用 CURRENT 宏检查当前的请求：

```
#define CURRENT (blk_dev[MAJOR_NR].current_request)
接下来看一看 sbull_request 的具体使用:
extern struct request *CURRENT;
void sbull_request(void) {
     unsigned long offset, total;
Begin:
     INIT_REQUEST:
          offset = CURRENT -> sector * sbull_hard;
          total = CURRENT -> current_nr_sectors * sbull_hard;
/*超出设备的边界*/
    if(total + offset > sbull_size * 1024) {
/*请求错误*/
         end_request(0);
         goto Begin;
     }
     if(CURRENT -> cmd == READ) {
          memcpy(CURRENT -> buffer, sbull_storage + offset, total);
     }
     else if(CURRENT -> cmd == WRITE) {
          memcpy(sbull_storage + offset, CURRENT -> buffer, total);
          }
          else {
               end_request(0);
          }
/*成功*/
     end_request(1);
/*当请求做完时让 INIT_REQUEST 返回*/
     goto Begin;
}
```

request()函数从 INIT_REQUEST 宏命令开始（它也在 blk.h 中定义），它对请求队列进行检查，保证请求队列中至少有一个请求在等待处理。如果没有请求（即 CURRENT = 0），则 INIT_REQUEST 宏命令将使 request()函数返回，任务结束。

假定队列中至少有一个请求，request()函数现在应处理队列中的第一个请求，当处理完请求后，request()函数将调用 end_request()函数。如果成功地完成了读写操作，那么应该用参数值 1 调用 end_request()函数；如果读写操作不成功，那么以参数值 0 调用 end_request()函数。如果队列中还有其他请求，那么将 CURRENT 指针设为指向下一个请求。执行 end_request()函数后，request()函数回到循环的起点，对下一个请求重复上面的处理过程。

（3）打开操作

打开操作要完成的流程图如下图 11.8 所示。

典型实现代码如下所示：

```
int sbull_open(struct inode *inode, struct file *filp) {
      int num = MINOR(inode -> i_rdev);
      if(num >= sbull -> size)
           return   ENODEV;
      sbull -> size = sbull -> size + num;
      if(!sbull -> usage) {
            check_disk_change(inode -> i_rdev);
            if(!*(sbull -> data))
                  return -ENOMEM;
            }
      sbull -> usage++;
      MOD_INC_USE_COUNT;
      return 0;
}
```

（4）释放设备操作

释放设备操作要完成的流程图如图 11.9 所示。

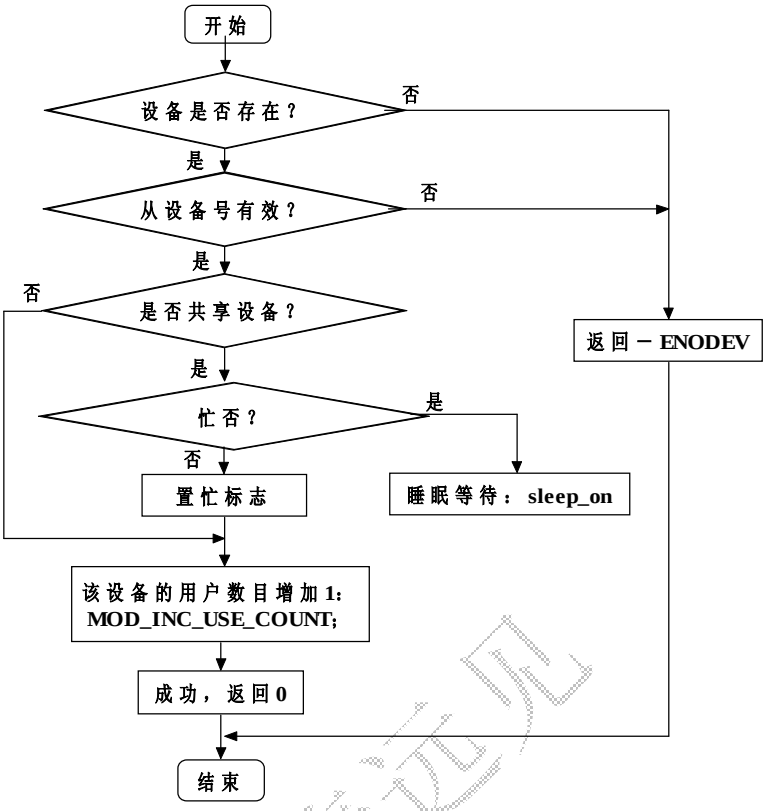


图 11.8　块设备打开操作流程图

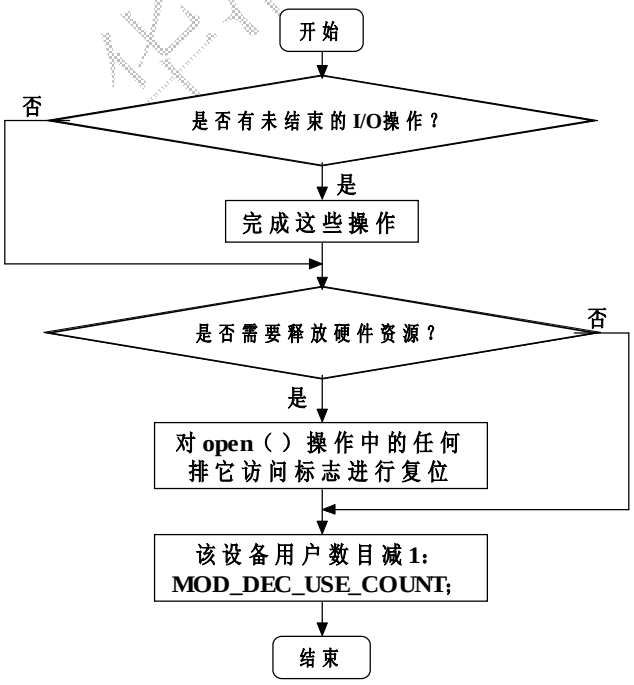


图 11.9 释放设备操作流程图

典型实现代码如下所示：

```
void sbull_release(struct inode *inode, struct file *filp) {
      sbull -> size = sbull -> size + MINOR(inode -> i_rdev);
      sbull -> usage--;
      MOD_DEC_USE_COUNT;
      printk("This blkdev is in release!\n");
      return 0;
}
```

（5）ioctl 操作

ioctl 操作要完成的流程图如图 11.10 所示。

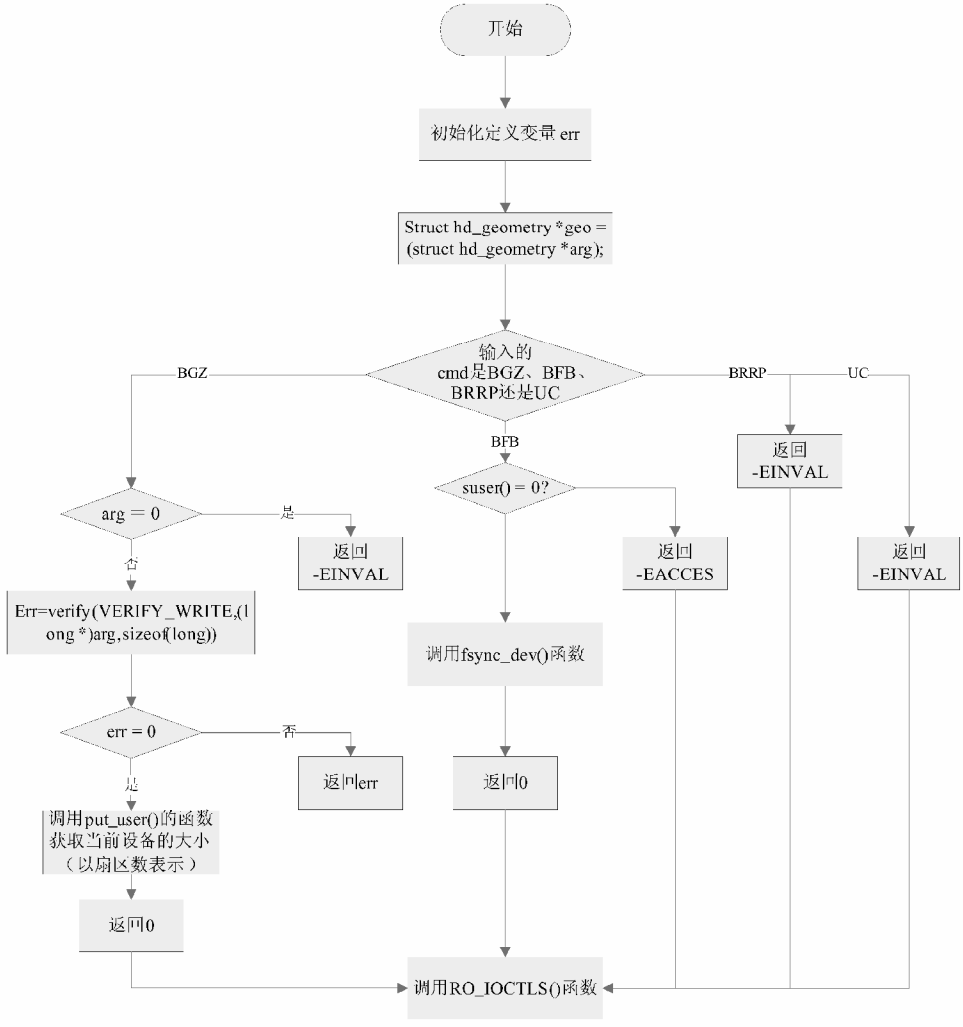


图 11.10 ioctl 操作要完成的流程图

其典型实现代码如下所示：

```
    #include <linux/ioctl.h>
    #include <linux/fs.h>
    int  sbull_ioctl(struct inode *inode，struct file *filp，unsigned int cmd，
unsigned long arg) {
    int err;
    struct hd_geometry *geo = (struct hd_geometry *)arg;
    PDEBUG("ioctl  0x%x  0x%lx\n"，cmd，arg);
    switch(cmd) {
          case  BLKGETSIZE:
    / *返回设备大小*/
          if(!arg)
                return –EINVAL;   // NULL pointer: not valid
          err = verify_area(VERIFY_WRITE，(long *)arg，sizeof(long));
          if(err)
          return err;
          put_user(1024*sbull_sizes[MINOR(inode -> i_rdev)/sbull_hardsects
[MINOR(inode -> i_rdev)]，(long*)arg];
          return 0;
          case  BLKFLSBUF: // flush
          if(!suser())
                return –EACCES;   // only root
          fsync_dev(inode -> i_rdev);
                return 0;
          case  BLKRRPART:   // re-read partition table: can't do it
               return –EINVAL;
          RO_IOCTLS(inode -> i_rdev，arg);
    // the default RO operations，宏 RO_IOCTLS(kdev_t dev，unsigned long where)
在 blk.h 中定义
          }
    return –EINVAL;   // unknown command
    }
```

11.5　中断编程

前面所讲述的驱动程序中都没有涉及到中断处理，而实际上，有很多 Linux 的驱动都是通过中断的方式来进行内核和硬件的交互。

这是驱动程序申请中断和释放中断的调用。在 include/linux/sched.h 里声明。

request_irq()调用的定义：

```
int request_irq(unsigned int irq,
void (*handler)(int irq, void *dev_id, struct pt_regs *regs),
unsigned long irqflags,const char * devname,oid *dev_id);
```

irq 是要申请的硬件中断号。在 Intel 平台，范围是 0～15。handler 是向系统登记的中断处理函数。这是一个回调函数，中断发生时，系统调用这个函数，传入的参数包括硬件中断号，device id，寄存器值。dev_id 就是下面的 request_irq 时传递给系统的参数 dev_id。irqflags 是中断处理的一些属性。比较重要的有 SA_INTERRUPT，标明中断处理程序是快速处理程序（设置 SA_INTERRUPT）还是慢速处理程序（不设置 SA_INTERRUPT）。快速处理程序被调用时屏蔽所有中断。慢速处理程序不屏蔽。还有一个 SA_SHIRQ 属性，设置了以后运行多个设备共享中断。dev_id 在中断共享时会用到。一般设置为这个设备的 device 结构本身或者 NULL。中断处理程序可以用 dev_id 找到相应的控制这个中断的设备，或者用 irq2dev_map 找到中断对应的设备。void free_irq(unsigned int irq,void *dev_id);

11.6 键盘驱动实现

11.6.1 键盘工作原理

1．原理简介

在键盘产生按键动作之后，键盘上的扫描芯片（一般为 8048）获得键盘的扫描码，并将其发送到主机端。在主机端的处理过程为是端口读取扫描码之后，对键盘模式作一个判断，如果是 RAW 模式，则直接将键盘扫描码发送给应用程序；如果是其他模式，则就将扫描码转化成为键盘码，然后再判断模式以决定是否将键盘码直接发送给应用程序；如果是 XLATE 或 Unicod。模式，则将键盘码再次转化成为符号码，然后根据对符号码解析，获得相应的处理函数，并将其送到 TY 设备的缓存中。模式判断的对应关系如图 11.11 所示。

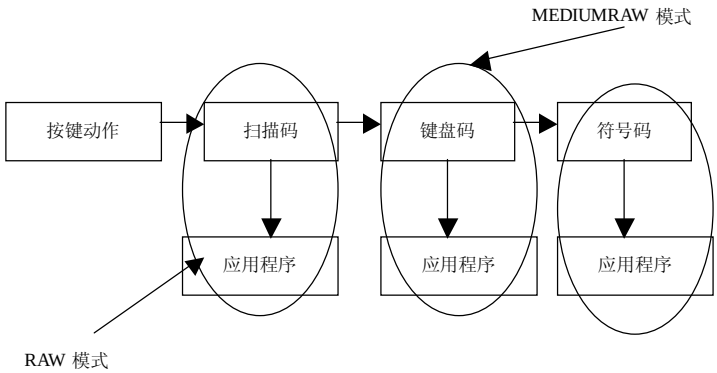


图 11.11 模式判断的对应关系

键盘模式有 4 种，这 4 种模式的对应关系如图 1 所示。

- Scancode mode（RAW）模式：将键盘端口上读出的扫描码放入缓冲区，通过参数 s 可以设置。
- Keycode mode（MEDIUMRAW）模式：将扫描码过滤为键盘码放入缓冲区，通过参数 k 可以设置。
- ASCII mode（XLATE）模式：识别各种键盘码的组合，转换为 TTY 终端代码放入缓冲区，通过参数 a 可以设置。
- UTF-8 mode（Unicode）模式：Unicode 模式基本上与 XLATE 相同，只不过可以通过数字小键盘 I 旬接枪入 Unicode 代码，通过参数 u 可以设置。

2．扫描码

一个基本按键的扫描码由 3 个字节组成：1 个字节的接通扫描码和 2 个字节的断开扫描码。其中第 1 和第 2 个字节相同，中间字节是断开标志 FOH。例如 B 键的接通扫描码是 32H，断开扫描码是 FOH 32H，B 键被按下时，32H 被发送出去，如果移植按住不放，则键盘将以按键重复率不停地发送 32H，直到该键释放，才发出断开扫描码 FOH 32H。扫描码与按键的位置有关，与该键的 ASCII 码并无对应关系。键盘上还有部分扩展键（功能键和控制键等），这些键的扫描码由 5 个字节组成，与基本键的扫描码相比，接通扫描码与断开扫描码前各多了一个固定值字节 EOH。例如 Home 键的接通扫描码是 EOH 70H，断开扫描码是 EOH FOH 70H。还有两个特殊键，PrintScreen 键的接通扫描码是 EOH 12H EOH 7CH；断开扫描码是 EOH FOH 7CH EOH FOH 77H，无断开扫描树。

3．键盘码

由前面的分析可见，单单一个键的按下与断开，键盘最多要产生一系列多达 6 个字节的扫描码序列，而内核必须解析扫描码序列从而定位某个键被按下与释放的事件。为达到这个目的，每一个键被分配一个键盘码 k（k 的范围 1-127）0 如果按键按下产生的键盘码为 k，则释放该键产生的键盘码为 k+128。按照键盘码的分配规则，对于产生单个扫描码范围 OxOl～Ox58 的键，其键盘码与扫描码相同。而对于 0x59～0x71 范围的键，可以查表获得其扫描码与键盘码对应。

4．符号码

符号码（keysym）最终是用来标志一个按键事件的惟一值；根据上面的分析，它由键盘码经过 Keymap 表映射而来。它包括 2 个字节，高 8 位表示 type，根据 type 的不同，我们最终选择不同的处理函数来处理不同类型的事件，type 相同的事件由同一个函数来处理。Type 包括一般键、方向键、字母键、函数键等。在<linux/keyboard.h>中可以找到 13 种键类型的宏定义。

5．Keymap 表

在使用键盘时常常使用组合键，而组合键的意义通常是系统另外赋予的，所以从键盘码向 TY 输入符的转换需要借助 Keymap 表来索引。

```
intshift_ final = shift state ^ kbd-Aockstate;
ushort'key map=key maps[shift finall];
```

```
keysym=key_map[keycodel];
```

由于共计有 8 个修饰符（modifier），即 Shift，AitGr，Control，Alt，ShiftL，ShiftR，CtrIL 和 CtrlR，因此共有 256 张可能的 Keymap，而在实际使用时，内核缺省只分配 7 张 Keymap:plain, Shift, AltR, Ctrl, Ctrl+Shift, AItL 和 Ctrl+AitL。

Keymap 表是一张二维表，结构图如图 11.12 所示。通过这张 Keymap 表，就能完成键盘码到符号码的转化，获得相应的符号码。

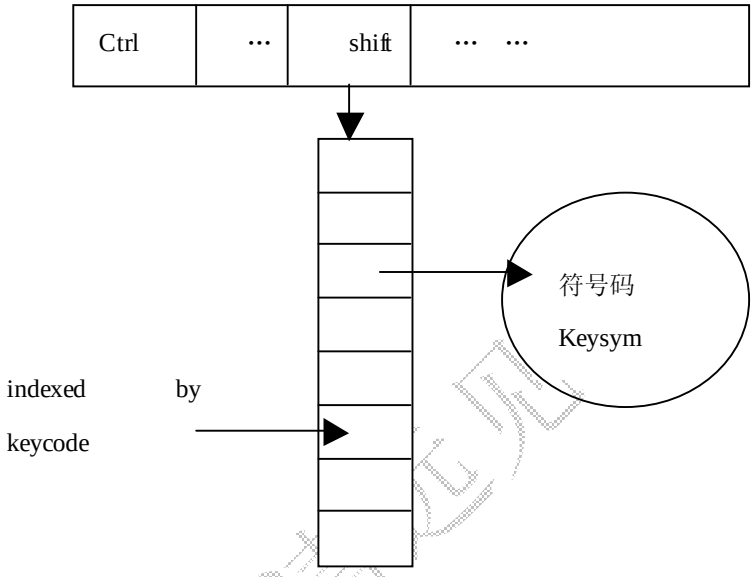


图 11.12 二维表结构图

11.6.2 键盘驱动综述

Linux 中的大多数驱动程序都采用了层次型的体系结构，键盘驱动程序也不例外。在 Linux 中，键盘驱动被划分成两层来实现。其中，上层是一个通用的键盘抽象层，完成键盘驱动中不依赖于底层具体硬件的一些功能，并且负责为底层提供服务；下层则是硬件处理层，与具体硬件密切相关，主要负责对硬件进行直接操作。键盘驱动程序的上层公共部分都在 driver/keyboard.c 中。在 keyboard.c 中，不涉及底层操作，也不涉及到任何体系结构，主要负责键盘初始化、键盘 tasklet 的挂入、按键盘后的处理、Keymap 表的装入、Scancode 的转化、与 TTY 设备的通信。

在 pc_keyb.c 中，主要负责一些底层操作，跟具体的体系结构相关，它完成的功能有：键盘的 I/O 端口和中断号的分配，键盘的硬件初始化，扫描码到键盘码的转化，键盘中断处理。

11.6.3 键盘驱动流程

（1）初始化

kbd_init()函数是键盘代码执行的入口点。kbd_init()在对键盘的工作模式及其他参数进行配置后，调用 kbd_init hw()函数。对于上层来说，此函数是一个统一的接口，对于不同体系结构或同体系下的不同开发板，它们的 kbd_init_hw()的实现代码是不同的（通过

CONFIG_ARCHXXX 的值来确定），它就是进行键盘的硬件初始化功能。然后将 keyboard tasklet 加入到 tasklet 链表中。至此键盘驱动的初始化工作已经完成。

键盘驱动的初始化代码是 keyboard.c 中的 kbd_init，其源码及分析如下所示：

```
int __init kbd_init(void)
{
      int i;
      struct kbd_struct kbd0;
/*维护 tty/console 对象，承担 tty 对外的输入和输出*/
      extern struct tty_driver console_driver;
/*缺省不亮灯*/
      kbd0.ledflagstate = kbd0.default_ledflagstate = KBD_DEFLEDS;
/*用于显示 flag*/
      kbd0.ledmode = LED_SHOW_FLAGS;
/*表示用 key_map 的第一个表，没有 lock 键*/
      kbd0.lockstate = KBD_DEFLOCK;
/*没有粘键*/
      kbd0.slockstate = 0;
      kbd0.modeflags = KBD_DEFMODE;
      kbd0.kbdmode = VC_XLATE;
 /*为每个控制台分配一个 KBD 结构*/
      for (i = 0 ; i < MAX_NR_CONSOLES ; i++)
            kbd_table[i] = kbd0;
/*维护当前各个控制台的 tty_struct 表*/
      ttytab = console_driver.table;
      kbd_init_hw();
/*把 keyboard_tasklet 挂到 CPU 的运行队列中去*/
      tasklet_enable(&keyboard_tasklet);
      tasklet_schedule(&keyboard_tasklet);

/*注册电源管理的 KEB 设备*/
      pm_kbd = pm_register(PM_SYS_DEV, PM_SYS_KBC, pm_kbd_request_override);

      return 0;
}
```

Kbd_init_hw()，包含了为 keyboard 分配 1/O 端口、分配中断号及对应处理函数、为进行基本保证测试（BAT）初始化寄存器，然后调用在 pc_keyb.c 的 intialize_kbd()进行硬件初始化，这是一个非常重要的函数，它的初始化过程的流程如图 11.13 所示。

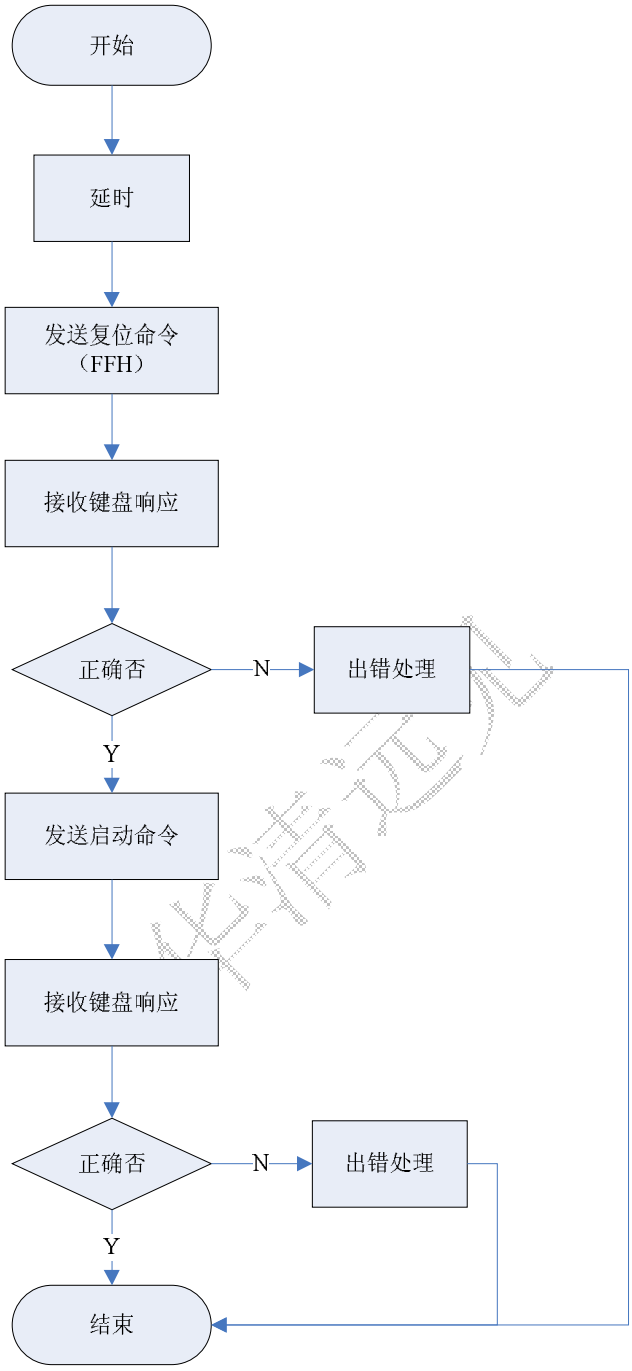


图 11.13　intialize_kbd()函数流程

```
static char * __init initialize_kbd(void)
{
       int status;
```

```
        /*
         * 测试键盘接口
         * 如果测试成功，那么将会有一个 x55 放在缓冲区中
         */
        kbd_write_command_w(KBD_CCMD_SELF_TEST);
        if (kbd_wait_for_input() != 0x55)
            return "Keyboard failed self test";

        /*
         * 启动一个键盘接口测试，这时会启动控制器来测试键盘的时钟和数据线，测试结果
        放在输入缓冲区中
         */
        kbd_write_command_w(KBD_CCMD_KBD_TEST);
        if (kbd_wait_for_input() != 0x00)
              return "Keyboard interface failed self test";

        /*
         * 通过启动键盘时钟使能键盘
         */
        kbd_write_command_w(KBD_CCMD_KBD_ENABLE);

        /*
         * 重启键盘。如果读取时间超时，就会认为在该机器里没有键盘
         * 如果键盘要求再次发送，则使能键盘重发机制
         */
        do {
               kbd_write_output_w(KBD_CMD_RESET);
               status = kbd_wait_for_input();
               if (status == KBD_REPLY_ACK)
                      break;
               if (status != KBD_REPLY_RESEND)
                      return "Keyboard reset failed, no ACK";
        } while (1);

        if (kbd_wait_for_input() != KBD_REPLY_POR)
               return "Keyboard reset failed, no POR";

        /*
         * 设置键盘控制模式，在这期间，键盘应该设置为关闭状态
```

```
         */
        do {
               kbd_write_output_w(KBD_CMD_DISABLE);
               status = kbd_wait_for_input();
               if (status == KBD_REPLY_ACK)
                      break;
               if (status != KBD_REPLY_RESEND)
                      return "Disable keyboard: no ACK";
        } while (1);

        kbd_write_command_w(KBD_CCMD_WRITE_MODE);
        kbd_write_output_w(KBD_MODE_KBD_INT
                        | KBD_MODE_SYS
                        | KBD_MODE_DISABLE_MOUSE
                        | KBD_MODE_KCC);

        if (!(kbd_write_command_w_and_wait(KBD_CCMD_READ_MODE) & KBD_MODE_KCC))
        {
               /*
                * If the controller does not support conversion,
                * Set the keyboard to scan-code set 1.
                */
               kbd_write_output_w(0xF0);
               kbd_wait_for_input();
               kbd_write_output_w(0x01);
               kbd_wait_for_input();
        }

        if (kbd_write_output_w_and_wait(KBD_CMD_ENABLE) != KBD_REPLY_ACK)
               return "Enable keyboard: no ACK";

        /*
         *最后，把键盘读取率设置为最高
         */
        if (kbd_write_output_w_and_wait(KBD_CMD_SET_RATE) != KBD_REPLY_ACK)
               return "Set rate: no ACK";
        if (kbd_write_output_w_and_wait(0x00) != KBD_REPLY_ACK)
               return "Set rate: no 2nd ACK";
```

```
        return NULL;
}
```

（2）按键处理

按键处理是键盘驱动中最为重要的一部分。当有按键事件产生时，则调用键盘中断处理函数，也就是 keyboard interrupt()，它会调用到 handle_kbd_event()并调用 handle_scancode()函数。handle_scancode()这个函数完成按键处理的过程，它的功能是与 TY 设备通信，Keymap 表装入，按键处理。handle scancode()处理的结果就是把按键发给相应的处理函数，这些函数基本上都会调用 put_queue()函数。这个函数就是将处理函数的结果发送到 TY 或者 Console 进行显示。

下面是 handle_kbd_event 和 handle_scancode 函数源代码：

```
static unsigned char handle_kbd_event(void)
{
        unsigned char status = kbd_read_status();
        unsigned int work = 10000;

        while ((--work > 0) && (status & KBD_STAT_OBF)) {
                unsigned char scancode;

                scancode = kbd_read_input();

                /*错误字节必须被忽略*/
#if 1
                /* 忽略错误字节 */
                if (!(status & (KBD_STAT_GTO | KBD_STAT_PERR)))
#endif
                {
                        if (status & KBD_STAT_MOUSE_OBF)
                                handle_mouse_event(scancode);
                        else
                                handle_keyboard_event(scancode);
                }

                status = kbd_read_status();
        }

        if (!work)
                printk(KERN_ERR "pc_keyb: controller jammed (0x%02X).\n", status);
```

```
        return status;
}

static inline void handle_keyboard_event(unsigned char scancode)
{
#ifdef CONFIG_VT
        kbd_exists = 1;
        if (do_acknowledge(scancode))
                handle_scancode(scancode, !(scancode & 0x80));
#endif
        tasklet_schedule(&keyboard_tasklet);
}
```

（3）转化键盘扫描码

在完成键盘的初始化之后，就需要完成对键盘扫描码的转化。这里调用函数 pckdb_translate，实现了 scancode 和 keycode 之间的转换。

```
int pckbd_translate(unsigned char scancode, unsigned char *keycode,char raw_mode)
{
     static int prev_scancode;

     /* special prefix scancodes.. */
     if (scancode == 0xe0 || scancode == 0xe1) {
          prev_scancode = scancode;
          return 0;
     }

     /* 0xFF 很少被发送，故可以忽略，0x00 是错误码。*/
     if (scancode == 0x00 || scancode == 0xff) {
          prev_scancode = 0;
          return 0;
     }

     scancode &= 0x7f;

     if (prev_scancode) {
        /*
         通常是 0xe0，但是一个暂停键可以产生 e1 1d 45 e1 9d c5 字符
         */
        if (prev_scancode != 0xe0) {
```

```
                if (prev_scancode == 0xe1 && scancode == 0x1d) {
                prev_scancode = 0x100;
                return 0;
               } else if (prev_scancode == 0x100 && scancode == 0x45) {
                *keycode = E1_PAUSE;
                prev_scancode = 0;
                } else {
   #ifdef KBD_REPORT_UNKN
                if (!raw_mode)
                    printk(KERN_INFO "keyboard: unknown e1 escape sequence\n");
   #endif
                prev_scancode = 0;
                return 0;
                }
          } else {
               prev_scancode = 0;
               /*
                *  键盘保持了它自己的内部总线锁和锁状态。在总线锁中，状态 E0 AA 会生成代
码，而状态 E0 2A 会跟随停止码
               if (scancode == 0x2a || scancode == 0x36)
                    return 0;

               if (e0_keys[scancode])
               *keycode = e0_keys[scancode];
               else {
   #ifdef KBD_REPORT_UNKN
              if (!raw_mode)
                  printk(KERN_INFO "keyboard: unknown scancode e0 %02x\n",
                      scancode);
   #endif
              return 0;
              }
        }
      } else if (scancode >= SC_LIM) {
         *keycode = high_keys[scancode - SC_LIM];

         if (!*keycode) {
              if (!raw_mode) {
   #ifdef KBD_REPORT_UNKN
```

```
                printk(KERN_INFO "keyboard: unrecognized scancode (%02x)"
                     " - ignored\n", scancode);
#endif
                }
                return 0;
           }
         } else
           *keycode = scancode;
           return 1;
}
```

（4）按键处理

在完成键盘扫描码转换之后就可以开始进行按键处理，这里用到了 keyboard.c 中的重要函数 kbd_processkeycode。源码如下所示：

```
static void
kbd_processkeycode(unsigned char keycode, char up_flag, int autorepeat)
{
     char raw_mode = (kbd->kbdmode == VC_RAW);

     if (up_flag) {
          rep = 0;
          if(!test_and_clear_bit(keycode, key_down))
               up_flag = kbd_unexpected_up(keycode);
     } else {
          rep = test_and_set_bit(keycode, key_down);
          /*如果键盘自动重复，那么就要把它忽略，我们会使用自己的自动重复机制*/
          if (rep && !autorepeat)
               return;
     }

     if (kbd_repeatkeycode == keycode || !up_flag || raw_mode) {
          kbd_repeatkeycode = -1;
          del_timer(&key_autorepeat_timer);
     }

#ifdef CONFIG_MAGIC_SYSRQ      /* Handle the SysRq Hack */
     if (keycode == SYSRQ_KEY) {
          sysrq_pressed = !up_flag;
          goto out;
```

```
        } else if (sysrq_pressed) {
              if (!up_flag) {
                    handle_sysrq(kbd_sysrq_xlate[keycode], kbd_pt_regs, kbd, tty);
                    goto out;
              }
        }
    #endif

        /*
         *计算下一次需要自动重复的时间
         */
        if (!up_flag && !raw_mode) {
              kbd_repeatkeycode = keycode;
              if (vc_kbd_mode(kbd, VC_REPEAT)) {
                    if (rep)
                          key_autorepeat_timer.expires = jiffies + kbd_repeati-
nterval;
                    else
                          key_autorepeat_timer.expires = jiffies + kbd_repea-
ttimeout;
                    add_timer(&key_autorepeat_timer);
              }
        }

        if (kbd->kbdmode == VC_MEDIUMRAW) {
              /* soon keycodes will require more than one byte */
              put_queue(keycode + up_flag);
              raw_mode = 1;    /* Most key classes will be ignored */
        }

        if (!rep ||
              (vc_kbd_mode(kbd,VC_REPEAT) && tty &&
               (L_ECHO(tty) || (tty->driver.chars_in_buffer(tty) == 0)))) {
              u_short keysym;
              u_char type;

              /* the XOR below used to be an OR */
              int shift_final = (shift_state | kbd->slockstate) ^
                   kbd->lockstate;
```

```
            ushort *key_map = key_maps[shift_final];

            if (key_map != NULL) {
                  keysym = key_map[keycode];
                  type = KTYP(keysym);

             if (type >= 0xf0) {
                   type -= 0xf0;
                   if (raw_mode && !(TYPES_ALLOWED_IN_RAW_MODE & (1 << type)))
                    goto out;
                   if (type == KT_LETTER) {
                    type = KT_LATIN;
                    if (vc_kbd_led(kbd, VC_CAPSLOCK)) {
                         key_map = key_maps[shift_final ^ (1<<KG_SHIFT)];
                         if (key_map)
                             keysym = key_map[keycode];
                    }
                  }
                  (*key_handler[type])(keysym & 0xff, up_flag);
                  if (type != KT_SLOCK)
                     kbd->slockstate = 0;
             } else {
                  /* maybe only if (kbd->kbdmode == VC_UNICODE) ? */
                  if (!up_flag && !raw_mode)
                     to_utf8(keysym);
             }
        } else {
             /* 我们至少需要更新移动状态*/
#if 1
            compute_shiftstate();
            kbd->slockstate = 0; /* play it safe */
#else
            keysym = U(key_maps[0][keycode]);
            type = KTYP(keysym);
            if (type == KT_SHIFT)
               (*key_handler[type])(keysym & 0xff, up_flag);
#endif
        }
     }
```

```
        rep = 0;
    out:
        return;
    }
```

11.7 实验内容——skull 驱动

1. 实验目的

该实验是编写最简单的字符驱动程序，这里的设备也就是一段内存，实现简单的读写功能。读者可以了解到整个驱动的编写流程。

2. 实验内容

该实验要求实现对一段内存的打开、关闭、读写的操作，并要通过编写测试程序来测试驱动安装是否成功。

3. 实验步骤

（1）编写代码

这个简单的驱动程序的源代码如下所示：

```
#include <linux/module.h>
#include <linux/init.h>
#include <linux/fs.h>
#include <linux/kernel.h>
#include <linux/malloc.h>
#include <asm/uaccess.h>
#include <linux/errno.h>
/*全局变量*/
unsigned int fs_major =0;
static char *data;
/*关键数据类型，注意每行结尾是逗号*/
static struct file_operations chr_fops={
          read:   test_read,
          write:  test_write,
          open:   test_open,
          release:      test_release,
};
```

```
    /*函数声明*/
    static ssize_t test_read(struct file *file, char *buf, size_t count, loff_t *f_pos);
    static ssize_t test_write(struct file *file,const char *buffer, size_t
count,loff_t *f_pos);
    static int test_open(struct inode *inode, struct file *file);
    static int test_release(struct inode *inode,struct file *file);
    int init_module(void);
    void cleanup_module(void);
    /*读函数*/
    static ssize_t test_read(struct file *file, char *buf, size_t count, loff_t *f_pos)
    {
              int len;
              if(count<0)
                        return -EINVAL;
              len = strlen(data);
              if(len < count)
                         count = len;
              copy_to_user(buf,data,count+1);
              return count;
    }
    /*写函数*/
    static ssize_t test_write(struct file *file,const char *buffer, size_t
count,loff_t *f_pos)
    {
              if(count < 0)
                         return -EINVAL;
              kfree(data);
              data = (char *)kmalloc(sizeof(char)*(count+1),GFP_KERNEL);
              if(!data)
                         return -ENOMEM;
              copy_from_user(data,buffer,count+1);
              return count;
    }
    /*打开函数*/
    static int test_open(struct inode *inode, struct file *file)
    {
               MOD_INC_USE_COUNT;
               printk("This is open\n");
               return 0;
```

```
}
/*释放函数*/
static int test_release(struct inode *inode,struct file *file)
{
           MOD_DEC_USE_COUNT;
           printk("this is released\n");
           return 0;
}
/*模块注册入口*/
int init_module(void)
{
          int res;
          res=register_chrdev(0,"fs",&chr_fops);
          if(res<0)
          {
                    printk("can't get major name!\n");
                    return res;
          }
          if(fs_major == 0)
                   fs_major = res;
          return 0;
}
/*撤销模块入口*/
void cleanup_module(void)
{
unregister_chrdev(fs_major,"fs");
}
```

（2）编译代码

要注意在此处要加上-DMODULE -D__KERNEL__选项，如下所示：

```
arm-linux-gcc -DMODULE -D__KERNEL__ -c kernel.c
```

（3）加载模块

```
insmod ./kernel.o
```

（4）查看设备号

```
vi /proc/device
```

（5）映射为设备文件

接下来就要将相应的设备映射为设备文件，这里可以使用命令 mknod，如下所示：

```
mknod /dev/fs c 254 0
```

这里的/dev/fs 就是相应的设备文件，c 代表字符文件，254 代表主设备号（与/proc/devices 中一样），0 为次设备号。

（6）编写测试代码

最后一步是编写测试代码，也就是用户空间的程序，该程序调用设备驱动来测试驱动的正确性。上面的实例只实现了简单的读写功能，测试代码如下所示：

```
#include <stdio.h>
#include <stdlib.h>
#include <sys/stat.h>
#include <sys/types.h>
#include <unistd.h>
#include <fcntl.h>
#include <linux/ioctl.h>
int main()
{
          int fd,i,nwrite,nread;
          char *buf ="hello\n";
          char read_buf[6]={0};
          fd=open("/dev/fs",O_RDWR);
          if(fd<=0)
          {
                    perror("open");
                    exit(1);
          }
          else
                    printf("open success\n");
          nwrite = write(fd,buf,strlen(buf));
          if(nwrite<0)
          {
                    perror("write");
                    exit(1);
          }
          nread = read(fd,read_buf,6);
          if(nread<0)
          {
                    perror("read");
```

```
            exit(1);
        }
        else
                printf("read is %s\n",read_buf);
        close(fd);
        exit(0);
}
```

4．实验结果

在加载模块后可以查看/var/log/messages 是否有程序中相应的信息输出：

```
Feb 21 09:49:10  kernel: This is open
```

查看设备号时有类似如下信息：

```
254 fs
```

这代表 fs 设备的主设备号是 254。
最后运行测试程序，结果如下所示：

```
[root@(none) tmp]# ./testing
open success
read is hello
```

查看/var/log/messages，有输出信息如下所示：

```
Feb 21 12:57:06 kernel: This is open
Feb 21 12:57:06 kernel: this is released
Feb 21 09:43:40 kernel: Goodbye world
```

本 章 小 结

本章主要介绍了嵌入式 Linux 设备驱动程序的开发。首先介绍了设备驱动程序的概念及 Linux 对设备驱动的处理，这里要明确驱动程序在 Linux 中的定位。

接下来介绍了字符设备驱动程序的编写，这里详细介绍了字符设备驱动程序的编写流程、重要的数据结构、设备驱动程序的主要组成以及 proc 文件系统。接着又以 LCD 驱动为例介绍了一个较为大型的驱动程序的编写步骤。

再接下来，本章介绍了块设备驱动程序的编写，主要包括块设备驱动程序描述符和块设备驱动的编写流程。

最后，本章介绍了中断编程，并以键盘驱动为例进行讲解。

本章的实验安排的是 skull 驱动程序的编写，通过该实验，读者可以了解到编写驱动程序的整个流程。

思考与练习

将本章中所述的 lcd 驱动程序运行编译，并通过模块加载在开发板上测试实验。

嵌入式与移动开发系列

NITE 国家信息技术紧缺人才培养工程
National Information Technology Education Project
国家信息技术紧缺人才培养工程系列丛书

众多专家、厂商联合推荐 • 业界权威培训机构的经验总结

嵌入式Linux应用程序开发标准教程（第2版）

华清远见嵌入式培训中心 编著

提供36小时嵌入式专家讲座视频和教学课件

Embeded Linux Application Development

人民邮电出版社
POSTS & TELECOM PRESS

第 12 章 Qt 图形编程基础

本章目标

- 掌握嵌入式 GUI 的种类和特点
- 掌握 Qt 中的信号与槽的机制
- 掌握 Qt/Embedded 的安装和配置
- 掌握 Qt/Embedded 应用程序的基本流程

12.1 嵌入式 GUI 简介

目前的桌面机操作系统大多有着美观、操作方便、功能齐全的 GUI（图形用户界面），例如 KDE 或者 GNOME。GUI（图形用户界面）是指计算机与其使用者之间的对话接口，可以说，GUI 是当今计算机技术的重大成就。它的存在为使用者提供了友好便利的界面，并大大地方便了非专业用户的使用，使得人们从繁琐的命令中解脱出来，可以通过窗口、菜单方便地进行操作。

而在嵌入式系统中，GUI 的地位也越来越重要，但是不同于桌面机系统，嵌入式 GUI 要求简单、直观、可靠、占用资源小且反应快速，以适应系统硬件资源有限的条件。另外，由于嵌入式系统硬件本身的特殊性，嵌入式 GUI 应具备高度可移植性与可裁减性，以适应不同的硬件条件和使用需求。总体来讲，嵌入式 GUI 具备以下特点：

- 体积小；
- 运行时耗用系统资源小；
- 上层接口与硬件无关，高度可移植；
- 高可靠性；
- 在某些应用场合应具备实时性。

UNIX 环境下的图形视窗标准为 X Window System，Linux 是类 UNIX 系统，所以顶层运行的 GUI 系统是兼容 X 标准的 XFree86 系统。X 标准大致可以划分 X Server、Graphic Library（底层绘图函数库）、Toolkits、Window Manager 等几大部分。其好处是具有可扩展性、可移植性等优点，但对于嵌入式系统而言无疑太过庞大、累赘、低效。目前流行的嵌入式 GUI 与 X 思路不同，这些 GUI 一般不局限于 X 标准，更强调系统的空间和效率。

12.1.1 Qt/Embedded

表 12.1 归纳了 Qt/Embedded 的一些优缺点。

表 12.1　　Qt/Embedded 分析

Qt/Embedded 分析		
优点	以开发包形式提供	包括了图形设计器、Makefile 制作工具、字体国际化工具、Qt 的 C++类库等
	跨平台	支持 Microsoft Windows 95/98/2000、Microsoft Windows NT、MacOS X、Linux、Solaris、HP-UX、Tru64 (Digital UNIX)、Irix、FreeBSD、BSD/OS、SCO、AIX 等众多平台
	类库支持跨平台	Qt 类库封装了适应不同操作系统的访问细节，这正是 Qt 的魅力所在
	模块化	可以任意裁减
缺点	结构也过于复杂臃肿，很难进行底层的扩充、定制和移植	例如： • 尽管 Qt/Embedded 声称，它最小可以裁剪到几百 KB，但这时的 Qt/Embedded 库已经基本失去了使用价值 • 它提供的控件集沿用了 PC 风格，并不太适合许多手持设备的操作要求 • Qt/Embedded 的底层图形引擎只能采用 framebuffer，只是针对高端嵌入式图形领域的应用而设计的 • 由于该库的代码追求面面俱到，以增加它对多种硬件设备的支持，

		造成了其底层代码比较凌乱，各种补丁较多的问题

12.1.2 MiniGUI

提起国内的开源软件，就肯定会提到 MiniGUI，它由魏永明先生和众多志愿者开发，是一个基于 Linux 的实时嵌入式系统的轻量级图形用户界面支持系统。

MiniGUI 分为最底层的 GAL 层和 IAL 层，向上为基于标准 POSIX 接口中 pthread 库的 Mini-thread 架构和基于 Server/Client 的 Mini-Lite 架构。其中前者受限于 thread 模式对于整个系统的可靠性——进程中某个 thread 的意外错误可能导致整个进程的崩溃，该架构应用于系统功能较为单一的场合。Mini-Lite 应用于多进程的应用场合，采用多进程运行方式设计的 Server/Client 架构能够较好地解决各个进程之间的窗口管理、Z 序剪切等问题。MiniGUI 还有一种从 Mini-Lite 衍生出的 standalone 运行模式。与 Lite 架构不同的是，standalone 模式一次只能以窗口最大化的方式显示一个窗口。这在显示屏尺寸较小的应用场合具有一定的应用意义。

MiniGUI 的 IAL 层技术 SVGA lib、LibGGI、基于 framebuffer 的 native 图形引擎以及哑图形引擎等，对于 Trolltech 公司的 QVFB 在 X Window 下也有较好的支持。IAL 层则支持 Linux 标准控制台下的 GPM 鼠标服务、触摸屏、标准键盘等。

MiniGUI 下丰富的控件资源也是 MiniGUI 的特点之一。当前 MiniGUI 的最新版本是 1.3.3。在该版本的控件中已经添加了窗口皮肤、工具条等桌面 GUI 中的高级控件支持。对比其他系统，“Mini”是 MiniGUI 的特色，轻量、高性能和高效率的 MiniGUI 已经应用在电视机顶盒、实时控制系统、掌上电脑等诸多场合。

12.1.3 Microwindows、Tiny X 等

Microwindows Open Source Project 成立的宗旨在于针对体积小的装置，建立一套先进的视窗环境，在 Linux 桌面上通过交叉编译可以很容易地制作出 Microwindows 的程序。Microwindows 能够在没有任何操作系统或其他图形系统的支持下运行，它能对裸显示设备进行直接操作。这样，Microwindows 就显得十分小巧，便于移植到各种硬件和软件系统上。

然而 Microwindows 的免费版本进展一直很慢，几乎处于停顿状态，而且至今为止，国内没有任何一家对 Microwindows 提供全面技术支持、服务和担保的专业公司。

Tiny X Server 是 XFree86 Project 的一部分，由 Keith Pachard 发展起来的，而他本身就是 XFree86 专案的核心成员之一。一般的 X Server 都过于庞大，因此 Keith Packard 就以 XFree86 为基础，精简而成 Tiny X Server，它的体积可以小到几百 KB，非常适合应用于嵌入式环境。

就纯 X Window System 搭配 Tiny X Server 架构来说，其最大的优点就是具有很好的弹性开发机制，并能大大提高开发速度。因为与桌面的 X 架构相同，因此相对于很多以 Qt、GTK+、FLTK 等为基础开发的软件可以很容易地移植过来。

虽然移植方便，但是却有体积大的缺点，由于很多软件本来是针对桌面环境开发的，因此无形之中具备了桌面环境中很多复杂的功能。因此“调校”变成采用此架构最大的课题，有时候重新改写可能比调校所需的时间还短。

表 12.2 总结了常见 GUI 的参数比较。

表 12.2　　常见 GUI 参数比较

名称 / 参数	MiniGUI	OpenGUI	Qt/Embedded
API（完备性）	Win32（很完备）	私有（很完备）	Qt（C++）（很完备）
函数库的典型大小	300KB	300KB	600KB
移植性	很好	只支持 x86 平台	较好
授权条款	LGPL	LGPL	QPL/GPL
系统消耗	小	最小	最大
操作系统支持	Linux	Linux，DOS，QNX	Linux

12.2　Qt/Embedded 开发入门

12.2.1　Qt/Embedded 介绍

1．架构

Qt/Embedded 以原始 Qt 为基础，并做了许多出色的调整以适用于嵌入式环境。Qt/Embedded 通过 Qt API 与 Linux I/O 设施直接交互，成为嵌入式 Linux 端口。同 Qt/X11 相比，Qt/Embedded 很省内存，因为它不需要一个 X 服务器或是 Xlib 库，它在底层抛弃了 X lib，采用 framebuffer（帧缓冲）作为底层图形接口。同时，将外部输入设备抽象为 keyboard 和 mouse 输入事件。Qt/Embedde 的应用程序可以直接写内核缓冲帧，这避免开发者使用繁琐的 Xlib/Server 系统。图 12.1 所示比较了 Qt/Embedded 与 Qt/X11 的架构区别。

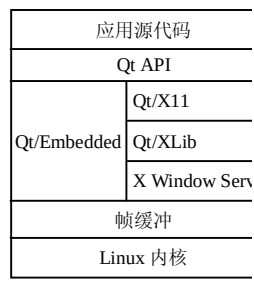


图 12.1　Qt/Embedded 与 X11 的 Linux 版本的比

使用单一的 API 进行跨平台的编程可以有很多好处。提供嵌入式设备和桌面计算机环境下应用的公司可以培训开发人员使用同一套工具开发包，这有利于开发人员之间共享开发经验与知识，也使得管理人员在分配开发人员到项目中的时候增加灵活性。更进一步来说，针对某个平台而开发的应用和组件也可以销售到 Qt 支持的其他平台上，从而以低廉的成本扩大产品的市场。

（1）窗口系统。

一个 Qt/Embedded 窗口系统包含了一个或多个进程，其中的一个进程可作为服务器。该服务进程会分配客户显示区域，以及产生鼠标和键盘事件。该服务进程还能够提供输入方法和一个用户接口给运行起来的客户应用程序。该服务进程其实就是一个有某些额外权限的客户进程。任何程序都可以在命令行上加上“-qws”的选项来把它作为一个服务器运行。

客户与服务器之间的通信使用共享内存的方法实现，通信量应该保持最小，例如

客户进程直接访问帧缓冲来完成全部的绘制操作，而不会通过服务器，客户程序需要负责绘制它们自己的标题栏和其他式样。这就是 Qt/Embedded 库内部层次分明的处理过程。客户可以使用 QCOP 通道交换消息。服务进程简单的广播 QCOP 消息给所有监听指定通道的应用进程，接着应用进程可以把一个插槽连接到一个负责接收的信号上，从而对消息做出响应。消息的传递通常伴随着二进制数据的传输，这是通过一个 QDataStream 类的序列化过程来实现的，有关这个类的描述，请读者参考相关资料。

QProcess 类提供了另外一种异步的进程间通信机制。它用于启动一个外部的程序并且通过写一个标准的输入和读取外部程序的标准输出和错误码来和它们通信。

（2）字体

Qt/Embedded 支持 4 种不同的字体格式：True Type 字体（TTF），Postscript Type1 字体，位图发布字体（BDF）和 Qt 的预呈现（Pre-rendered）字体（QPF）。Qt 还可以通过增加 Qfont-
Factory 的子类来支持其他字体，也可以支持以插件方式出现的反别名字体。

每个 TTF 或者 TYPE1 类型的字体首次在图形或者文本方式的环境下被使用时，这些字体的字形都会以指定的大小被预先呈现出来，呈现的结果会被缓冲。根据给定的字体尺寸（例如 10 或 12 点阵）预先呈现 TTF 或者 TYPE1 类型的字体文件并把结果以 QPF 的格式保存起来，这样可以节省内存和 CPU 的处理时间。QPF 文件包含了一些必要的字体，这些字体可以通过 makeqpf 工具取得，或者通过运行程序时加上“-savefonts”选项获取。如果应用程序中使用到的字体都是 QPF 格式，那么 Qt/Embedded 将被重新配置，并排除对 TTF 和 TYPE1 类型的字体的编译，这样就可以减少 Qt/Embedded 的库的大小和存储字体的空间。例如一个 10 点阵大小的包含所有 ASCII 字符的 QPF 字体文件的大小为 1300 字节，这个文件可以直接从物理存储格式映射成为内存存储格式。

Qt/Embedded 的字体通常包括 Unicode 字体的一部分子集，ASCII 和 Latin-1。一个完整的 16 点阵的 Unicode 字体的存储空间通常超过 1MB，我们应尽可能存储一个字体的子集，而不是存储所有的字，例如在一个应用中，仅仅需要以 Cappuccino 字体、粗体的方式显示产品的名称，但是却有一个包含了全部字形的字体文件。

（3）输入设备及输入法。

Qt/Embedded 3.0 支持几种鼠标协议：BusMouse、IntelliMouse,Microsoft 和 MouseMan.Qt/
Embedded 还支持 NECVr41XX 和 iPAQ 的触摸屏。通过从 QWSMouseHandler 或者 Qcalibra-
tedMouseHandler 派生子类，开发人员可以让 Qt/Embedded 支持更多的客户指示设备。

Qt/Embedded 支持标准的 101 键盘和 Vr41XX 按键，通过子类化 QWSKeyboardHandler 可以让 Qt/Embedded 支持更多的客户键盘和其他的非指示设备。

对于非拉丁语系字符（例如阿拉伯、中文、希伯来和日语）的输入法，需要把它写成过滤器的方式，并改变键盘的输入。输入法的作者应该对全部的 Qt API 的使用有完整的认识。在一个无键盘的设备上，输入法成了惟一的输入字符的手段。Qtopia 提供了 4 种输入方法：笔迹识别器、图形化的标准键盘、Unicode 键盘和基于字典方式

提取的键盘。

（4）屏幕加速

通过子类化 QScreen 和 QgfxRaster 可以实现硬件加速，从而为屏幕操作带来好处。Troll-
tech 提供了 Mach64 和 Voodoo3 视频卡的硬件加速的驱动例子，同时可以按照协议编写其他的驱动程序。

2．Qt 的开发环境

Qt/Embedded 的开发环境可以取代那些我们熟知的 UNIX 和 Windows 开发工具。它提供了几个跨平台的工具使得开发变得迅速和方便，尤其是它的图形设计器。UNIX 下的开发者可以在 PC 机或者工作站使用虚拟缓冲帧，从而可以模仿一个和嵌入式设备的显示终端大小，像素相同的显示环境。

嵌入式设备的应用可以在安装了一个跨平台开发工具链的不同的平台上编译。最通常的做法是在一个 UNIX 系统上安装跨平台的带有 libc 库的 GNU C++编译器和二进制工具。在开发的许多阶段，一个可替代的做法是使用 Qt 的桌面版本，例如通过 Qt/X11 或是 Qt/Windows 来进行开发。这样开发人员就可以使用他们熟悉的开发环境，例如微软公司的 Visual C++或者 Borland C++。在 UNIX 操作系统下，许多环境也是可用的，例如 Kdevelop，它也支持交互式开发。

如果 Qt/Embedded 的应用是在 UNIX 平台下开发的话，那么它就可以在开发的机器上以一个独立的控制台或者虚拟缓冲帧的方式来运行，对于后者来说，其实是有一个 X11 的应用程序虚拟了一个缓冲帧。通过指定显示设备的宽度、高度和颜色深度，虚拟出来的缓冲帧将和物理的显示设备在每个像素上保持一致。这样每次调试应用时开发人员就不用总是刷新嵌入式设备的 Flash 存储空间，从而加速了应用的编译、链接和运行周期。运行 Qt 的虚拟缓冲帧工具的方法是在 Linux 的图形模式下运行以下命令：

```
qvfb (回车)
```

当 Qt 嵌入式的应用程序要把显示结果输出到虚拟缓冲帧时，我们在命令行运行这个程序，并在程序名后加上-qws 的选项。例如：$> hello－qws。

3．Qt 的支撑工具

Qt 包含了许多支持嵌入式系统开发的工具，有两个最实用的工具是 qmake 和 Qt designer（图形设计器）。

- qmake 是一个为编译 Qt/Embedded 库和应用而提供的 Makefile 生成器。它能够根据一个工程文件（.pro）产生不同平台下的 Makefile 文件。qmake 支持跨平台开发和影子生成，影子生成是指当工程的源代码共享给网络上的多台机器时，每台机器编译链接这个工程的代码将在不同的子路径下完成，这样就不会覆盖别人的编译链接生成的文件。qmake 还易于在不同的配置之间切换。
- Qt 图形设计器可以使开发者可视化地设计对话框而不需编写代码。使用 Qt

图形设计器的布局管理可以生成能平滑改变尺寸的对话框。

qmake 和 Qt 图形设计器是完全集成在一起的。

12.2.2 Qt/Embedded 信号和插槽机制

1．机制概述

信号和插槽机制是 Qt 的核心机制，要精通 Qt 编程就必须对信号和插槽有所了解。信号和插槽是一种高级接口，应用于对象之间的通信，它是 Qt 的核心特性，也是 Qt 区别于其他工具包的重要地方。信号和插槽是 Qt 自行定义的一种通信机制，它独立于标准的 C/C++语言，因此要正确地处理信号和插槽，必须借助一个称为 moc（Meta Object Compiler）的 Qt 工具，该工具是一个 C++预处理程序，它为高层次的事件处理自动生成所需要的附加代码。

所谓图形用户接口的应用就是要对用户的动作做出响应。例如，当用户单击了一个菜单项或是工具栏的按钮时，应用程序会执行某些代码。大部分情况下，是希望不同类型的对象之间能够进行通信。程序员必须把事件和相关代码联系起来，这样才能对事件做出响应。以前的工具开发包使用的事件响应机制是易崩溃的，不够健壮的，同时也不是面向对象的。

以前，当使用回调函数机制把某段响应代码和一个按钮的动作相关联时，通常把那段响应代码写成一个函数，然后把这个函数的地址指针传给按钮，当那个按钮被单击时，这个函数就会被执行。对于这种方式，以前的开发包不能够确保回调函数被执行时所传递进来的函数参数就是正确的类型，因此容易造成进程崩溃。另外一个问题是，回调这种方式紧紧地绑定了图形用户接口的功能元素，因而很难进行独立的开发。

信号与插槽机制是不同的。它是一种强有力的对象间通信机制，完全可以取代原始的回调和消息映射机制。在 Qt 中信号和插槽取代了上述这些凌乱的函数指针，使得用户编写这些通信程序更为简洁明了。信号和插槽能携带任意数量和任意类型的参数，它们是类型完全安全的，因此不会像回调函数那样产生 core dumps。

所有从 QObject 或其子类（例如 Qwidget）派生的类都能够包含信号和插槽。当对象改变状态时，信号就由该对象发射（emit）出去了，这就是对象所要做的全部工作，它不知道另一端是谁在接收这个信号。这就是真正的信息封装，它确保对象被当作一个真正的软件组件来使用。插槽用于接收信号，但它们是普通的对象成员函数。一个插槽并不知道是否有任何信号与自己相连接。而且，对象并不了解具体的通信机制。

用户可以将很多信号与单个插槽进行连接，也可以将单个信号与很多插槽进行连接，甚至将一个信号与另外一个信号相连接也是可能的，这时无论第一个信号什么时候发射，系统都将立刻发射第二个信号。总之，信号与插槽构造了一个强大的部件编程机制。

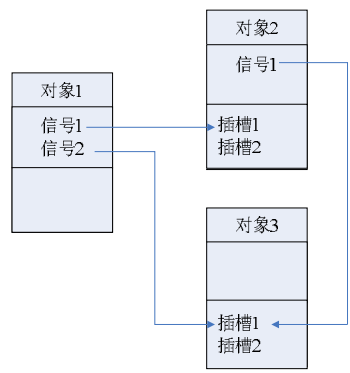


图 12.2 对象间信号与插槽的关

图 12.2 所示为对象间信号与插槽的关系。

2．信号与插槽实现实例

（1）信号。

当某个信号对其客户或所有者内部状态发生改变时，信号就被一个对象发射。只有定义了这个信号的类及其派生类才能够发射这个信号。当一个信号被发射时，与其相关联的插槽将被立刻执行，就像一个正常的函数调用一样。信号－插槽机制完全独立于任何 GUI 事件循环。只有当所有的槽返回以后发射函数（emit）才返回。如果存在多个槽与某个信号相关联，那么，当这个信号被发射时，这些槽将会一个接一个地执行，但是它们执行的顺序将会是随机的、不确定的，用户不能人为地指定哪个先执行、哪个后执行。

Qt 的 signals 关键字指出进入了信号声明区，随后即可声明自己的信号。例如，下面定义了 3 个信号：

```
signals:
void mySignal();
void mySignal(int x);
void mySignalParam(int x,int y);
```

在上面的定义中，signals 是 Qt 的关键字，而非 C/C++的。接下来的一行 void mySignal()定义了信号 mySignal，这个信号没有携带参数；接下来的一行 void mySignal(int x)定义了重名信号 mySignal，但是它携带一个整形参数，这有点类似于 C++中的虚函数。从形式上讲信号的声明与普通的 C++函数是一样的，但是信号却没有函数体定义。另外，信号的返回类型都是 void。信号由 moc 自动产生，它们不应该在.cpp 文件中实现。

（2）插槽。

插槽是普通的 C++成员函数，可以被正常调用，它们惟一的特殊性就是很多信号可以与其相关联。当与其关联的信号被发射时，这个插槽就会被调用。插槽可以有参数，但插槽的参数不能有缺省值。

插槽是普通的成员函数，因此与其他的函数一样，它们也有存取权限。插槽的存取权限决定了谁能够与其相关联。同普通的 C++成员函数一样，插槽函数也分为 3 种类型，即 public slots、private slots 和 protected slots。

- public slots：在这个区内声明的槽意味着任何对象都可将信号与之相连接。这对于组件编程非常有用，用户可以创建彼此互不了解的对象，将它们的信号与槽进行连接以便信息能够正确地传递。
- protected slots：在这个区内声明的槽意味着当前类及其子类可以将信号与之相连接。这适用于那些槽，它们是类实现的一部分，但是其界面接口却面向外部。
- private slots：在这个区内声明的槽意味着只有类自己可以将信号与之相连接。这适用于联系非常紧密的类。

插槽也能够被声明为虚函数，这也是非常有用的。插槽的声明也是在头文件中进行的。例如，下面声明了 3 个插槽：

```
public slots:
void mySlot();
void mySlot(int x);
void mySignalParam(int x,int y);
```

（3）信号与插槽关联。

通过调用 QObject 对象的 connect()函数可以将某个对象的信号与另外一个对象的插槽函数或信号相关联，当发射者发射信号时，接收者的槽函数或信号将被调用。

该函数的定义如下所示：

```
bool QObject::connect (const QObject * sender, const char * signal,const QObject * receiver, const char * member) [static]
```

这个函数的作用就是将发射者 sender 对象中的信号 signal 与接收者 receiver 中的 member 插槽函数联系起来。当指定信号 signal 时必须使用 Qt 的宏 SIGNAL()，当指定插槽函数时必须使用宏 SLOT()。如果发射者与接收者属于同一个对象的话，那么在 connect()调用中接收者参数可以省略。

- 信号与插槽相关联。

下例定义了两个对象：标签对象 label 和滚动条对象 scroll，并将 valueChanged()信号与标签对象的 setNum()插槽函数相关联，另外信号还携带了一个整型参数，这样标签总是显示滚动条所处位置的值。

```
QLabel *label = new QLabel;
QScrollBar *scroll = new QScrollBar;
QObject::connect(scroll,                SIGNAL(valueChanged(int)),label, SLOT(setNum(int)));
```

- 信号与信号相关联。

在下面的构造函数中，MyWidget 创建了一个私有的按钮 aButton，按钮的单击事件产生的信号 clicked()与另外一个信号 aSignal()进行关联。这样，当信号 clicked()被发射时，信号 aSignal()也接着被发射。如下所示：

```
class MyWidget : public QWidget
{
public:
MyWidget();
...
signals:
void aSignal();
...
private:
...
QPushButton *aButton;
```

```
};

MyWidget::MyWidget()
{
    aButton = new QPushButton(this);
    connect(aButton, SIGNAL(clicked()), SIGNAL(aSignal()));
}
```

（4）解除信号与插槽关联。

当信号与槽没有必要继续保持关联时，用户可以使用 disconnect()函数来断开连接。其定义如下所示：

```
bool QObject::disconnect (const QObject * sender, const char *
signal,const Object * receiver, const char * member) [static]
```

这个函数断开发射者中的信号与接收者中的槽函数之间的关联。

有 3 种情况必须使用 disconnect()函数。

- 断开与某个对象相关联的任何对象。

当用户在某个对象中定义了一个或者多个信号，这些信号与另外若干个对象中的槽相关联，如果想要切断这些关联的话，就可以利用这个方法，非常简洁。如下所示：

```
disconnect(myObject, 0, 0, 0)
或者
myObject->disconnect()
```

- 断开与某个特定信号的任何关联。

这种情况是非常常见的，其典型用法如下所示：

```
disconnect(myObject, SIGNAL(mySignal()), 0, 0)
或者
myObject->disconnect(SIGNAL(mySignal()))
```

- 断开两个对象之间的关联。

这也是非常常用的情况，如下所示：

```
disconnect(myObject, 0, myReceiver, 0)
或者
myObject->disconnect(myReceiver)
```

注意　在 disconnect()函数中 0 可以用作一个通配符，分别表示任何信号、任何接收对象、接收对象中的任何槽函数。但是发射者 sender 不能为 0，其他 3 个参数的值可以等于 0。

12.2.3　搭建 Qt/Embedded 开发环境

一般来说，用 Qt/Embedded 开发的应用程序最终会发布到安装有嵌入式 Linux 操

作系统的小型设备上，所以使用装有 Linux 操作系统的 PC 机或者工作站来完成 Qt/Embedded 开发当然是最理想的环境，此外 Qt/Embedded 也可以安装在 UNIX 或 Windows 系统上。这里就以在 Linux 操作系统中安装为例进行介绍。

这里需要有 3 个软件安装包：tmake 工具安装包、Qt/Embedded 安装包和 Qt 的 X11 版的安装包。

- tmake1.11 或更高版本：生成 Qt/Embedded 应用工程的 Makefile 文件。
- Qt/Embedded：Qt/Embedded 安装包。
- Qt 2.3.2 for X11：Qt 的 X11 版的安装包，产生 X11 开发环境所需要的两个工具。

注意

这些软件安装包都有许多不同的版本，由于版本的不同会导致这些软件在使用时可能引起的冲突，为此必须依照一定的安装原则，Qt/Embedded 安装包的版本必须比 Qt for X11 的安装包的版本新，这是因为 Qt for X11 的安装包中的两个工具 uic 和 designer 产生的源文件会和 Qt/Embedded 的库一起被编译链接，因此要本着“向前兼容”的原则，Qt for X11 的版本应比 Qt/Embedded 的版本旧。

1．安装 tmake

用户使用普通的解压缩即可，注意要将路径添加到全局变量中去，如下所示：

```
tar zxvf tmake-1.11.tar.gz
export TMAKEDIR=$PWD/tmake-1.11
export TMAKEPATH=$TMAKEDIR/lib/qws/linux-x86-g++
export PATH=$TMAKEDIR/bin:$PATH
```

2．安装 Qt/Embedded 2.3.7

这里使用常见的解压命令及安装命令即可，要注意这里的路径与不同的系统有关，读者要根据实际情况进行修改。另外，这里的 configure 命令带有参数“-qconfig –qvfb –depths 4816，32”分别为指定 Qt 嵌入式开发包生成虚拟缓冲帧工具 qvfb，并支持 4、8、16、32 位的显示颜色深度。另外读者也可以在 configure 的参数中添加“-system”、“-jpeg”或“gif”命令，使 Qt/Embedded 平台能支持 jpeg、gif 格式的图形。

Qt/Embedded 开发包有 5 种编译范围的选项，使用这些选项可控制 Qt 生成的库文件的大小。如命令 make sub-src 指定按精简方式编译开发包，也就是说有些 Qt 类未被编译。其他编译选项的具体用法可通过“./configure–help”命令查看。精简方式的安装步骤如下所示：

```
tar zxvf qt-embedded-2.3.7.tar.gz
cd qt-2.3.7
export QTDIR=$PWD
export QTEDIR=$QTDIR
export PATH=$QTDIR/bin:$PATH
```

```
export LD_LIBRARY_PATH=$QTDIR/lib:$LD_LIBRARY_PATH
./configure -qconfig local-qvfb -depths 4,8,16,32
make sub-src
```

3．安装 Qt/X11 2.3.2

与上一步类似，用户也可以在 configure 后添加一定的参数，如“-no-opengl”或“-no-xfs”，可以键入命令“./configure –help”来获得一些帮助信息。

```
tar xfz qt-x11-2.3.2.tar.gz
cd qt-2.3.2
export QTDIR=$PWD
export PATH=$QTDIR/bin:$PATH
export LD_LIBRARY_PATH=$QTDIR/lib:$LD_LIBRARY_PATH
./configure -no-opengl
make
make -C tools/qvfb
mv tools/qvfb/qvfb bin
cp bin/uic $QTEDIR/bin
```

12.2.4 Qt/Embedded 窗口部件

Qt 提供了一整套的窗口部件。它们组合起来可用于创建用户界面的可视元素。按钮、菜单、滚动条、消息框和应用程序窗口都是窗口部件的实例。因为所有的窗口部件既是控件又是容器，因此 Qt 的窗口部件不能任意地分为控件和容器。通过子类化已存在的 Qt 部件或少数时候必要的全新创建，自定义的窗口部件能很容易地创建出来。

窗口部件是 QWidget 或其子类的实例，用户自定义的窗口通过子类化得到，如图 12.3 所示。

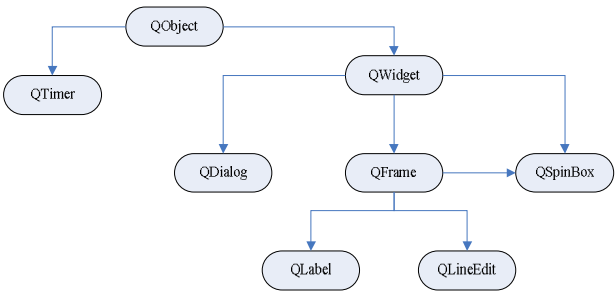


图 12.3　源自 QWidget 的类层次结构

一个窗口部件可包含任意数量的子部件。子部件在父部件的区域内显示。没有父部件的部件是顶级部件（比如一个窗口），通常在桌面的任务栏上有它们的入口。Qt 不在窗口部件上施加任何限制。任何部件都可以是顶级部件，任何部件都可以是其他部件的子部件。通过自动或手动（如果你喜欢）使用布局管理器可以设定子部件在父部件区域中的位置。如果父部件被停用、隐藏或删除，则同样的动作会应用于它的所

有子部件。

1．Hello 窗口实例

下面是一个显示“Hello Qt/Embedded!”的程序的完整代码：

```
#include <qapplication.h>
#include <qlabel.h>
int main(int argc, char **argv)
{
    QApplication app(argc, argv);
    QLabel *hello=new QLabel
("<font color=blue>Hello""<i>Qt Embedded!</i></font>",0);
    app.setMainWidget(hello);
    hello->show();
    return app.exec();
}
```

图 12.4 是该 Hello 窗口的运行效果图：

2．常见通用窗口组合

Qt 中还有一些常见的通用窗口，它们使用了 Windows 风格显示，图 12.5、12.6、12.7、12.8 分别描述了常见的一些通用窗口的组合使用。

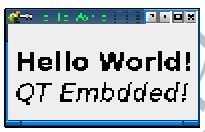


图 12.4　Hello 窗口运行效果图

图 12.5　使用 QHBox 排列一个标签和一个按钮

图 12.6　使用了 QButtonGroup 的两个单选框和两个复选框

图 12.7　QGroupBox 组合图示

图 12.8 使用了 QGroupBox 进行排列的日期类 QDateTimeEdit、一个行编辑框类 QLine-
Edit、一个文本编辑类 QTextEdit 和一个组合框类 QComboBox。

图 12.9 是以 QGrid 排列的一个 QDial、一个 QProgressBar、一个 QSpinBox、一个 QScrollBar、一个 QLCDNumber 和一个 QSlider。

图 12.10 是以 QGrid 排列的一个 QIconView、一个 QListView、一个 QListBox 和一个 QTable。

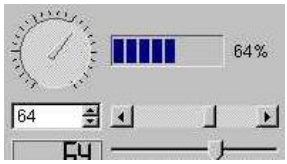


图 12.8　QGrid 组合图示 1

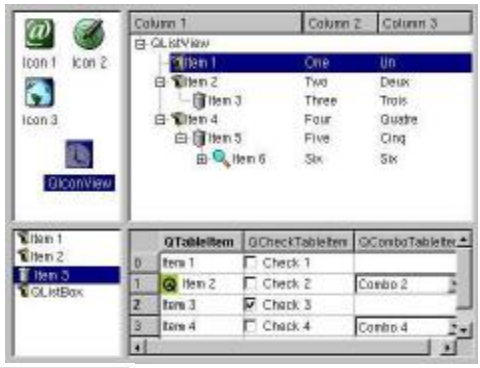


图 12.9　QGrid 组合图示 2

图 12.10　钟表部件图示

3．自定义窗口

开发者可以通过子类化 QWidget 或它的一个子类创建他们自己的部件或对话框。为了举例说明子类化，下面提供了数字钟部件的完整代码。

钟表部件是一个能显示当前时间并自动更新的 LCD。一个冒号分隔符随秒数的流逝而闪烁，如图 12.10 所示。

Clock 从 QLCDNumber 部件继承了 LCD 功能。它有一个典型部件类所拥有的典型构造函数，带有可选的 parent 和 name 参数（如果设置了 name 参数，测试和调试会更容易）。系统有规律地调用从 QObject 继承的 timerEvent()函数。

它在 clock.h 中定义如下所示：

```
#include <qlcdnumber.h>
class Clock:public QLCDNumber
{
public:
    Clock(QWidget *parent=0,const char *name=0);
protected:
    void timerEvent(QTimerEvent *event);
private:
    void showTime();
    bool showingColon;
};
```

构造函数 showTime()是用当前时间初始化钟表，并且告诉系统每 1000ms 调用一次 timerEvent()来刷新 LCD 的显示。在 showTime()中，通过调用 QLCDNumber::display()来显示当前时间。每次调用 showTime()来让冒号闪烁时，冒号就被空白代替。

clock.cpp 的源码如下所示：

```
#include <qdatetime.h>
#include "clock.h"
Clock::Clock(QWidget *parent,const char *name)
:QLCDNumber(parent,name),showingColon(true)
{
    showTime();
    startTimer(1000);
}
void Clock::timerEvent(QTimerEvent *)
{
    showTime();
}
void Clock::showTime()
{
    QString timer=QTime::currentTime().toString().left(5);
    if (!showingColon)
    {
    time[2]=' ';
    }
    display(time);
    showingColon=!showingColon;
}
```

文件 clock.h 和 clock.cpp 完整地声明并实现了 Clock 部件。

```
#include <qapplication.h>
#include "clock.h"
int main(int argc,char **argv)
{
   QApplication app(argc,argv);
   Clock *clock=new Clock;
   app.setMainWidget(clock);
   clock->show();
   return app.exec();
}
```

12.2.5 Qt/Embedded 图形界面编程

Qt 提供了所有可能的类和函数来创建 GUI 程序。Qt 既可用来创建“主窗口”式的程序，即一个有菜单栏，工具栏和状态栏作为环绕的中心区域；也可以用来创建“对