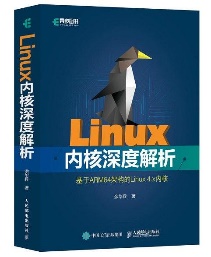
作者: 余华兵  
副标题: 基于ARM64架构的Linux4.x内核  
书名: Linux内核深度解析



# 第　1章 内核引导和初始化　1

处理器上电之后，首先执行引导程序，引导程序吧内核加载到内存，然后执行内核，内核初始化完成之后，启动用户空间的第一个进程

## 到哪里读取引导程序　1

从物理地址0读取第一条指令，而物理地址0是分配给NOR闪存的，NOR存放了引导程序。

## 引导程序　1

嵌入式设备通常使用u-boot作为引导程序

### 1．2．1　入口\_start　1

### 1．2．2　标号reset　2

### 1．2．3　函数\_main　4

### 1．2．4　函数run\_main\_loop　6

## 1．3　内核初始化　8

1．3．1　汇编语言部分　8

1．3．2　C语言部分　11

1．3．3　SMP系统的引导　12

## 1．4　init进程　15

# 第　2章 进程管理　17

2．1　进程　17

2．2　命名空间　18

2．3　进程标识符　20

2．4　进程关系　21

2．5　启动程序　23

2．5．1　创建新进程　23

2．5．2　装载程序　41

2．6　进程退出　48

2．6．1　线程组退出　49

2．6．2　终止进程　51

2．6．3　查询子进程终止原因　53

2．7　进程状态　55

2．8　进程调度　55

2．8．1　调度策略　55

2．8．2　进程优先级　56

2．8．3　调度类　57

2．8．4　运行队列　59

2．8．5　任务分组　60

2．8．6　调度进程　65

2．8．7　调度时机　75

2．8．8　带宽管理　85

2．9　SMP调度　93

2．9．1　进程的处理器亲和性　93

2．9．2　对调度器的扩展　96

2．9．3　限期调度类的处理器负载均衡　96

2．9．4　实时调度类的处理器负载均衡　98

2．9．5　公平调度类的处理器负载均衡　99

2．9．6　迁移线程　108

2．9．7　隔离处理器　110

2．10　进程的安全上下文　111

# 第3章　内存管理　113

3．1　概述　113

3．2　虚拟地址空间布局　115

3．2．1　虚拟地址空间划分　115

3．2．2　用户虚拟地址空间布局　115

3．2．3　内核地址空间布局　121

3．3　物理地址空间　122

3．4　内存映射　124

3．4．1　应用编程接口　125

3．4．2　数据结构　129

3．4．3　创建内存映射　133

3．4．4　虚拟内存过量提交策略　137

3．4．5　删除内存映射　139

3．5　物理内存组织　140

3．5．1　体系结构　140

3．5．2　内存模型　140

3．5．3　三级结构　141

3．6　引导内存分配器　144

3．6．1　bootmem分配器　144

3．6．2　memblock分配器　145

3．6．3　物理内存信息　148

3．7　伙伴分配器　151

3．7．1　基本的伙伴分配器　151

3．7．2　分区的伙伴分配器　152

3．7．3　根据可移动性分组　158

3．7．4　每处理器页集合　162

3．7．5　分配页　163

3．7．6　释放页　181

3．8　块分配器　184

3．8．1　编程接口　185

3．8．2　SLAB分配器　186

3．8．3　SLUB分配器　197

3．8．4　SLOB分配器　204

3．9　不连续页分配器　207

3．9．1　编程接口　207

3．9．2　数据结构　208

3．9．3　技术原理　209

3．10　每处理器内存分配器　210

3．10．1　编程接口　210

3．10．2　技术原理　212

3．11　页表　219

3．11．1　统一的页表框架　219

3．11．2　ARM64处理器的页表　222

3．12　页表缓存　226

3．12．1　TLB表项格式　226

3．12．2　TLB管理　226

3．12．3　地址空间标识符　228

3．12．4　虚拟机标识符　232

3．13　巨型页　233

3．13．1　处理器对巨型页的支持　233

3．13．2　标准巨型页　235

3．13．3　透明巨型页　245

3．14　页错误异常处理　257

3．14．1　处理器架构特定部分　257

3．14．2　用户空间页错误异常　266

3．14．3　内核模式页错误异常　283

3．15　反碎片技术　288

3．15．1　虚拟可移动区域　289

3．15．2　内存碎片整理　291

3．16　页回收　309

3．16．1　数据结构　310

3．16．2　发起页回收　317

3．16．3　计算扫描的页数　320

3．16．4　收缩活动页链表　321

3．16．5　回收不活动页　323

3．16．6　页交换　325

3．16．7　回收slab缓存　335

3．17　内存耗尽杀手　338

3．17．1　使用方法　338

3．17．2　技术原理　338

3．18　内存资源控制器　340

3．18．1　使用方法　340

3．18．2　技术原理　344

3．19　处理器缓存　370

3．19．1　缓存结构　370

3．19．2　缓存策略　372

3．19．3　缓存维护　374

3．19．4　SMP缓存一致性　378

3．19．5　利用缓存提高性能的编程技巧　383

3．20　连续内存分配器　384

3．20．1　使用方法　385

3．20．2　技术原理　386

3．21　userfaultfd　391

3．21．1　使用方法　391

3．21．2　技术原理　395

3．22　内存错误检测工具KASAN　401

3．22．1　使用方法　401

3．22．2　技术原理　402

# 第4章　中断、异常和系统调用　403

4．1　ARM64异常处理　403

4．1．1　异常级别　403

4．1．2　异常分类　404

4．1．3　异常向量表　405

4．1．4　异常处理　407

4．2　中断　411

4．2．1　中断控制器　412

4．2．2　中断域　413

4．2．3　中断控制器驱动初始化　415

4．2．4　Linux中断处理　422

4．2．5　中断线程化　428

4．2．6　禁止/开启中断　430

4．2．7　禁止/开启单个中断　431

4．2．8　中断亲和性　431

4．2．9　处理器间中断　432

4．3　中断下半部　434

4．3．1　软中断　435

4．3．2　小任务　441

4．3．3　工作队列　444

4．4　系统调用　457

4．4．1　定义系统调用　457

4．4．2　执行系统调用　459

# 第5章　内核互斥技术　463

5．1　信号量　464

5．2　读写信号量　465

5．3　互斥锁　466

5．4　实时互斥锁　467

5．5　原子变量　468

5．6　自旋锁　472

5．7　读写自旋锁　476

5．8　顺序锁　478

5．8．1　完整版的顺序锁　479

5．8．2　只提供序列号的顺序锁　481

5．9　禁止内核抢占　482

5．10　进程和软中断互斥　483

5．11　进程和硬中断互斥　483

5．12　每处理器变量　484

5．12．1　静态每处理器变量　484

5．12．2　动态每处理器变量　484

5．12．3　访问每处理器变量　485

5．13　每处理器计数器　485

5．14　内存屏障　487

5．14．1　编译器屏障　488

5．14．2　处理器内存屏障　489

5．14．3　MMIO写屏障　492

5．14．4　隐含内存屏障　493

5．14．5　ARM64处理器内存屏障　493

5．15　RCU　495

5．15．1　使用方法　496

5．15．2　技术原理　504

5．16　可睡眠RCU　533

5．16．1　使用方法　533

5．16．2　技术原理　534

5．17　死锁检测工具lockdep　542

5．17．1　使用方法　543

5．17．2　技术原理　543

# 第6章　文件系统　548

6．1　概述　548

6．1．1　用户空间层面　549

6．1．2　硬件层面　549

6．1．3　内核空间层面　550

6．2　虚拟文件系统的数据结构　552

6．2．1　超级块　552

6．2．2　挂载描述符　554

6．2．3　文件系统类型　555

6．2．4　索引节点　556

6．2．5　目录项　559

6．2．6　文件的打开实例和打开文件表　561

6．3　注册文件系统类型　563

6．4　挂载文件系统　564

6．4．1　系统调用mount　566

6．4．2　绑定挂载　567

6．4．3　挂载命名空间　568

6．4．4　挂载根文件系统　574

6．5　打开文件　580

6．5．1　编程接口　580

6．5．2　技术原理　582

6．6　关闭文件　591

6．7　创建文件　593

6．7．1　使用方法　593

6．7．2　技术原理　594

6．8　删除文件　595

6．8．1　使用方法　595

6．8．2　技术原理　595

6．9　设置文件权限　597

6．9．1　使用方法　597

6．9．2　技术原理　598

6．10　页缓存　599

6．10．1　地址空间　600

6．10．2　基数树　601

6．10．3　编程接口　602

6．11　读文件　602

6．11．1　编程接口　602

6．11．2　技术原理　603

6．12　写文件　606

6．12．1　编程接口　606

6．12．2　技术原理　607

6．13　文件回写　610

6．13．1　编程接口　610

6．13．2　技术原理　610

6．14　DAX　618

6．14．1　使用方法　618

6．14．2　技术原理　618

6．15　常用的文件系统类型　621