



教学操作系统Nanos: 设计与实现

蒋炎岩

南京大学计算机软件研究所
南京大学软件新技术国家重点实验室



提纲



- Nanos背景
- Nanos简介
- 实验概述
- 系统演示
- 总结



Nanos背景



Nanos背景



■ “理解操作系统的唯一途径就是构造真实的操作系统”

- 开源工具为设计操作系统实验提供了可能
- 国外高校普遍有自己的教学操作系统与配套实验
 - Stanford: Pintos (11KLOC)
 - MIT: xv6, Jos (7KLOC, IA32 SMP)

■ 教学操作系统的要求

- 结合教学内容，辅助原理课程教学
- 补充原理课程难以覆盖的细节
- 养成重视理论联系实际，自己动手和操作的良好习惯
- 为学生解析实用操作系统打好基础



国内操作系统实验课程概况



- 国内部分院校已经开始在操作系统实验课中设置有深度/难度实验项目
- 在Linux操作系统上开展实验
 - 在Linux内核代码基础上分析、修改、替换和系统编程浙江大学[宋06]、西安邮电学院[陈09]
- 在国外实验方案基础上改进
 - 清华大学基于MIT xv6操作系统的实验方案[陈09a]
- 开发独立的教学操作系统
 - 北京航空航天大学基于MIPS的教学操作系统[王09]



设计实验方案



■ 基本目标

- 学生在一学期的课程中构造一个微型操作系统
- 巩固基本原理，理解操作系统的构造方法
- 培养动手和操作、分析和解决问题的能力

■ 特点：小而精

- 强调从理论到实际，理论与实践紧密结合，“学中干、干中学”
- 突出原理、简化细节
- 覆盖原理课程中的重点、难点
- 初步形成操作系统整体概念和框架



Nanos与南京大学操作系统实验



Nanos历史



- 系统设计受到了xv6, Minix和Linux的影响
- 2009年试验项目MiniOS
 - 2010年大学计算机课程报告论坛论文发表
 - 微内核, 8086, Minix交叉编译
- 2010年重构
 - 增加内核信号量、加强了消息机制
 - 2011年大学计算机课程报告论坛论文发表
- 2012年2月 第二次重构
 - 重构到IA32保护模式



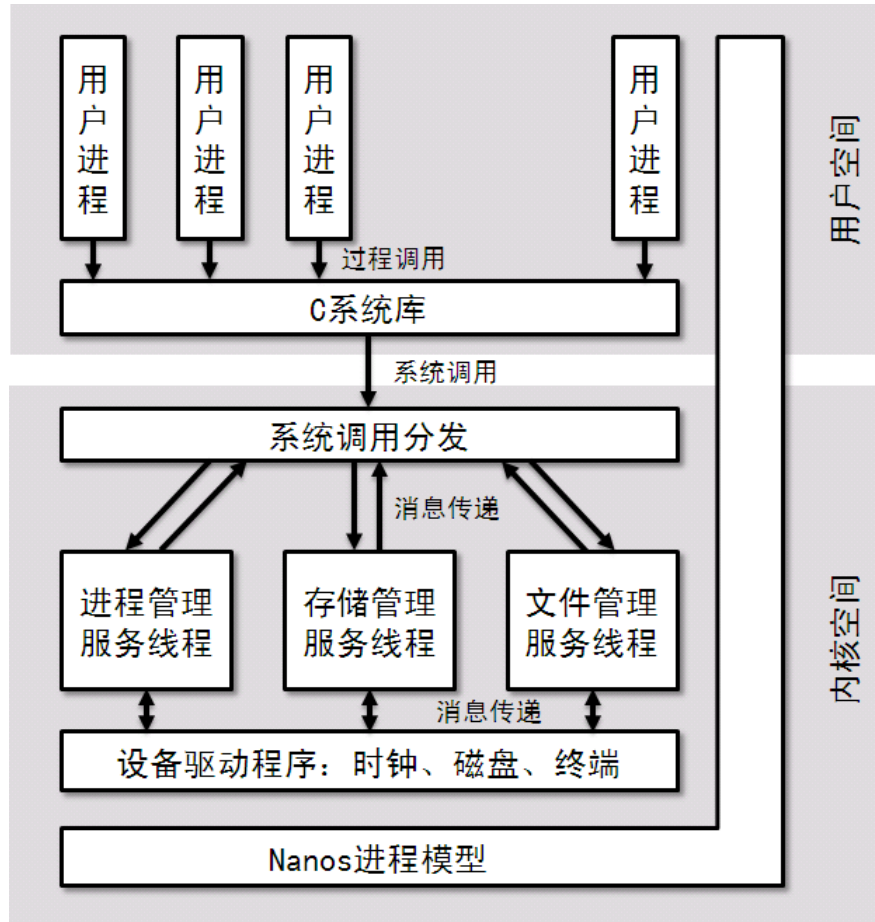
Nanos概述



- Nanos用户接口
 - 18个简化的POSIX.1系统调用
 - Nanofs文件系统
 - ELF32二进制文件格式
 - (部分)二进制级兼容Linux系统调用
- Nanos特性
 - x86保护模式
 - 抢占多任务
 - 分页存储管理
 - 简化的Unix文件系统



Nanos系统架构





Nanos内核



- Nanos内核
 - 系统调用由服务器执行
 - 服务器线程运行在内核空间
- 优点
 - 清晰地刻画了模块的边界
 - 同时简化实现的复杂性



Nanos功能和服务



- 核心功能
 - 进程模型和管理、内核信号量、半同步消息机制
- 设备驱动程序
 - 时钟(timer)、磁盘(ide)、终端(tty)
- 系统调用服务线程
 - 进程管理 fork, execve, exit, getpid, alarm, waitpid
 - 存储管理 sbrk
 - 文件管理 open, close, read, write, lseek, dup, chdir, mkdir, readdir, link, unlink



南京大学模拟类操作系统实验概述



前导课程知识



- 程序设计
 - 熟练掌握C语言
 - 理解堆栈和指针
- 汇编语言
 - 理解机器执行代码的原理
- 计算机体系结构
 - 中断/异常、I/O、MMU等概念



南京大学操作系统实验概述



■ 实验平台

- GNU/Linux
- GCC, gdb, qemu, git

■ 实验方案(框架代码量/学生作业量)

- Lab0: 游戏与系统启动 (300/500)
- Lab1: 内核线程调度 (400/300)
- Lab2: 内核同步与通信 (700/500)
- Lab3: 用户进程与存储管理 (-/1000)
- Lab4: 文件系统 (-/2000)
- 未设置设备管理实验



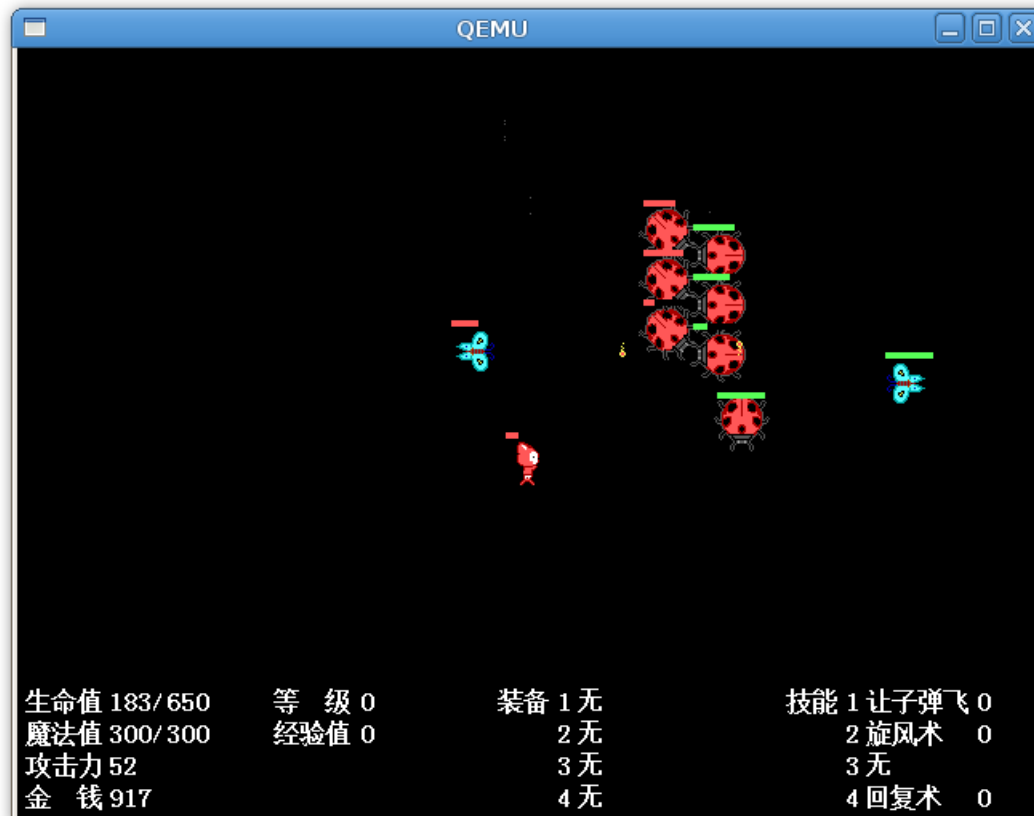
Lab0: 游戏与系统启动



- 阅读理解框架代码
 - 理解点火及系统引导过程
 - 理解基本中断管理
 - 理解基本绘图功能
- 实现游戏
 - 中断驱动的游戏逻辑更新



Lab0: 学生作业





Lab1: 内核线程调度



- 利用时钟中断完成线程上下文切换
- 编写“共享计数器”代码
 - 观察实验结果
 - 讨论产生实验结果的原因
- 编写“计算 π ”代码
 - 讨论如何处理浮点数



Lab1: 学生作业

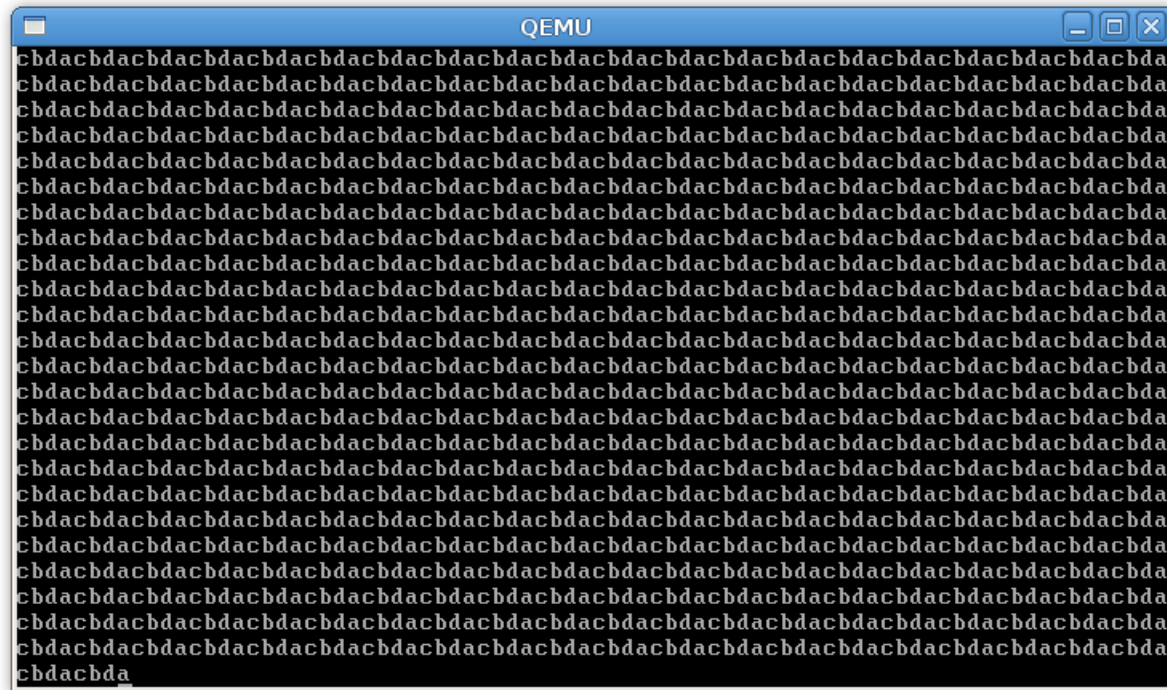




Lab2: 内核同步与通信



- 实现当前线程的阻塞
- 实现内核中的信号量
 - 实现单处理器中的互斥锁
 - 实现原子PV操作
- 利用信号量实现半同步的消息机制
 - 异步的send、同步的receive
 - 能够把中断包装成消息
- 添加设备驱动程序(框架代码)





Lab3: 用户进程与存储管理



- 为用户进程分配地址空间、实现分页存储管理
 - sbrk系统调用
 - 用户空间的malloc
- 实现简易的文件系统
 - 连续存储，数字表示文件名
- 实现进程的创建、替换和删除
 - fork, exec, exit



Lab3: 学生作业



```
QEMU
1
Father Process
Child Process

#2
I'm running happily~~
I'm running happily~~
I'm running happily~~
I'm running happily~~
I'm running happily~~
but now I'll exit 000000

#3
Hello, I'm the child process
I will exit soon
will exit soon
exit soon
soonHello I'm the father process
My child process's exit number is:7
#4
I will sleep 5 seconds
Happy to see you again
You can change the sleep time by give an arguement

#
```



Lab4: 文件系统



- 实现一个简化的Unix文件系统
 - iNode与文件系统管理
 - 文件描述符管理
 - 操作系统内的文件管理系统调用
- 实现Shell展示结果
 - 实现基本工具集(cat, ls, wc, ...)
 - 展示运行结果
- 选做：实现匿名管道



Lab4: 学生作业



```
QEMU
SYSTEM BOOT FINISHED
Use cat help1.txt and cat help2.txt to get help ^(^^)/
#ls
bin      hello      help1.txt  help2.txt  root

#cat help1.txt
1. ls [-iall] [name1 name2 name3 ...]
   -i I_node number printed before file or directory's name
   -a All entries are listed, even "." and ".."
   -l Long listing:name type size

2. cat [file1 file2 file3 ...]
   Display these files on the terminal

3. echo args      output args and a newline to the terminal

4. head [-n] [file1 file2 file3 ...]
   print the first n lines of every file, like : head -10 help.txt

5. tail [-n] [file1 file2 file3 ...]
   print the last  n lines of every file, the same as head

6. link name1 name2

#
```



实验支持和辅助材料



- 框架代码
 - github.com/NJUOS
- 课程网站
 - cslab.nju.edu.cn/opsystem
- 参考资料(见课程网站)
- 实验教材
 - 2011 《操作系统实验讲义》
 - 将会编写新实验讲义







框架代码



- 使用github公开发布
 - Organization on github: NJUOS
 - 学生使用git clone下载代码
 - 在本地使用git追踪版本(学习版本管理兼预防抄袭)

Lab1 /

name
 boot
 kernel
 .gitignore
 README.markdown



课程网站



■ 通知、讲义、作业、资料发布平台

The Course Wiki

← → ↻ cslab.nju.edu.cn/opsystem/#MainPage

您好，请[登录](#) / [注册](#)

[EDIT THIS PAGE](#)

2012操作系统实验

- 课时：每周三2pm-4pm，基础实验楼乙-124
- 联系方式：蒋炎岩 nju.jiangyy@gmail.com 计算机系楼813

通知公告

- 课前请阅读[为什么要做操作系统实验](#)和[怎样写实验报告](#)以及一篇[转载的短文](#)。
- 已经发布Lab1框架代码。
- 在完成实验的过程中将会遇到大量的问题。其中的绝大部分都可以通过阅读实验讲义、查阅搜索引擎或教科书解决。其他较为基本的问题可以与周围的同学讨论。



抄袭检测工具



- 针对《操作系统实验》特点开发
- 二进制代码相似度匹配
 - 函数切分、提取指令类型
 - 带窗口的最长公共子序列判定函数相似度
 - 二部图最优匹配判定全局相似度



系统演示



总结



探索中的操作系统实验



■ 初步成果

- Nanos的设计和实现
- 基于Nanos的操作系统实验方案

■ 未来工作

- 改进Nanos架构、扩充Nanos功能
- 设计和完善实验内容(项目)
- 更新实验教材和补充读物



学生反馈



有人说，OS 实验是 CSer 的大姨妈。哈哈，虽然有点恶俗，但是却很恰当，好像正好是每个月都有那么一个 lab，每次到 lab 的 deadline 的前几天，大家都是蓬头垢面精神憔悴两眼无神自言自语的。甚至有次躺到床上睡下了还突然坐起来开心地说我想通了。那几天总是全神贯注地倾入在 lab 中，吃饭在想，走路在想，连睡觉都在想，不得安心。而提交完实验后就仿佛获得了新生一样。

OS 实验就让我们这样痛并快乐着过完了一个学期，说得矫情点，每次的 lab 就好像一次破茧，过程是艰难的，结果是喜悦的，在茧里痛苦地挣扎却不断地强壮我们的翅膀，以后才能飞得更高更远。



学生反馈



所以的知识的获得，都是一件让人幸福与喜悦的事情，而当知识不再是书本，不再是皮皮踢，而是我手下码动的代码，屏幕上的黑框框的时候，我对它的记忆会比任何时候都来得深刻，而如果现在再让我考一遍操作系统期末考的话，我相信自己会比那个熬了好几个通宵背书的娃答题更胸有成竹，真正概念的掌握不是死记硬背来的，而是实践理解来的，没有什么比实践更让人印象深刻了。

操作系统当之无愧是让我这学期最印象深刻的一门课，我知道以后还会有孩子不可避免的走入它的世界，登上它的舞台，我由衷的希望他们可以坚持下去，享受它带给自己的蜕变与涅槃。在这里计算机的孩子才会真正长大。



学生反馈



OS实验激发了我们学习操作系统课程的兴趣，在完成整个实验的过程中，我们获益良多。“没有尝试去实现操作系统，你就不会理解操作系统。”这句话对我们不再是一句口号，而成为了实实在在的感受：真正操作系统实现与纸上谈兵相差甚远。我们学会了查阅、整理和利用资料，我们真正地得到了较大规模开发的亲身体验，这是一次宝贵的经历。没有OS实验，我们不会如此明显地感受到我们的不足以及和牛人的巨大差距，我们更不会发现在我们自己身上隐藏着的潜力。



存在的问题



- 前导课程强度不足
 - 学生对体系认识不够深入
 - 部分学生程序设计功底不足
- 改革带来两极分化
 - 好学生“吃不饱”，渴望更多有难度的实验和前沿知识
 - 差学生“下地狱”，完成率随实验推移逐步下降



谢谢！





参考文献



- [宋06] 宋广华,段健平,李善平. 边学边干, 教研结合——谈“操作系统”课程教学改革[J].电气电子教学学报,2006,28(5):1-3.
- [王09] 王雷.操作系统实验设计[J].计算机教育,2009.17(3):54-56.
- [陈09]陈莉君,梁琛,王小银.基于“Linux内核分析”的教学内容研究与实践 [J].计算机教育,2009.14(2):81-82.
- [陈09a] 陈渝,向勇.“操作系统”课程实验教学探讨[J].计算机教育,2009.14(2):135-136.



参考文献



- Linux内核 <http://kernel.org/>
- Minix3 <http://minix3.org/>
- MIT 6.828 <http://pdos.csail.mit.edu/6.828/>
- Stanford Pintos <http://scs.stanford.edu/10wi-cs140/pintos/pintos.html/>
- Qemu <http://qemu.org/>