



操作系统实验

Labs of Operating Systems

陈鹏飞
计算机学院
2021-03-09



实验安排

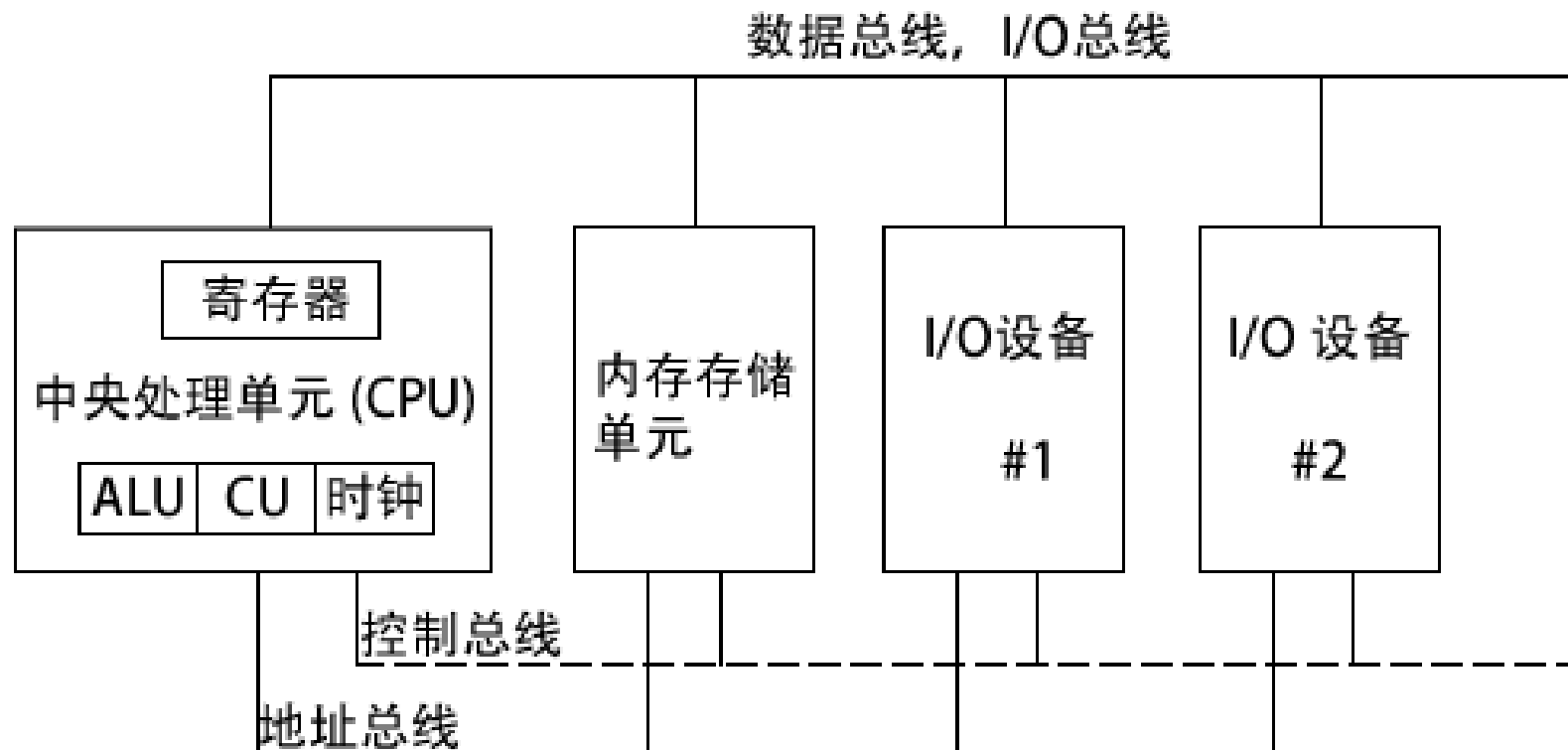
分成两次课程

实验-2： 实模式和保护模式下的OS启动

了解操作系统启动的原理，利用汇编语言实模式即（20位地址空间）的启动和保护模式即（32位地址空间）下的启动方法，并能够在此基础上利用汇编或者C程序实现简单的应用；

1. 回顾、学习32位汇编语言的基本语法；
2. 编写简单的汇编程序，进行中断、输入输出测试；
3. 实现实模式下OS启动；
4. 在实模式下利用汇编/C/Rust等实现简单的应用；
5. 实现保护模式下OS启动；
6. 在保护模式下利用汇编/C/Rust等实现简单的应用；
7. 比较实模式和保护模式的不同；

处理器架构



加载程序

在程序执行之前，需要用一种工具程序将其加载到内存，这种工具程序称为程序加载器（program loader）。加载后，操作系统必须将 CPU 向程序的入口，即程序开始执行的地址。以下步骤是对这一过程的详细分解。

- 1) 操作系统（OS）在当前磁盘目录下搜索程序的文件名；
- 2) 如果程序文件被找到，OS 就访问磁盘目录中的程序文件基本信息，包括文件大小，及其在磁盘驱动器上的物理位置；
- 3) OS 确定内存中下一个可使用的位置，将程序文件加载到内存；
- 4) OS 开始执行程序的第一条机器指令（程序入口）。当程序开始执行后，就成为一个进程（process）；
- 5) 进程自动运行。OS 的工作是追踪进程的执行，并响应系统资源的请求；
- 6) 进程结束后，就会从内存中移除；



IA-32处理器基本架构

x86 处理器有三个主要的操作模式：保护模式、实地址模式和系统管理模式；以及一个子模式：虚拟 8086 (virtual-8086) 模式，这是保护模式的特殊情况。

1) 保护模式 (Protected Mode)

保护模式是处理器的原生状态，在这种模式下，所有的指令和特性都是可用的。分配给程序的独立内存区域被称为段，而处理器会阻止程序使用自身段范围之外的内存。

2) 虚拟 8086 模式 (Virtual-8086 Mode)

保护模式下，处理器可以在一个安全环境中，直接执行实地址模式软件，如 MS-DOS 程序。换句话说，如果一个程序崩溃了或是试图向系统内存区域写数据，都不会影响到同一时间内执行的其他程序。现代操作系统可以同时执行多个独立虚拟 8086 会话。

3) 实地址模式 (Real-Address Mode)

实地址模式实现的是早期 Intel 处理器的编程环境，但是增加了一些其他的特性，如切换到其他模式的功能。当程序需要直接访问系统内存和硬件设备时，这种模式就很有用。

4) 系统管理模式 (System Management Mode)

系统管理模式 (SMM) 向操作系统提供了实现诸如电源管理和系统安全等功能的机制。这些功能通常是由计算机制造商实现的，他们为了一个特定的系统设置而定制处理器。

地址空间

- 在 32 位保护模式下，一个任务或程序最大可以寻址 4GB 的线性地址空间。从 P6 处理器开始，一种被称为扩展物理寻址 (extended physical addressing) 的技术使得可以被寻址的物理内存空间增加到 64GB。
- 在实地址模式下，IA-32 处理器使用 20 位的地址线，可以访问 $2^{20}=1\text{MB}$ 的内存，范围是 0x0000 到 0xFFFFF。但是，我们看到寄存器的访问模式只有 32 位，16 位和 8 位，形如 `eax`，`ax`，`ah`，`al`。那么我们如何才能使用 16 位的寄存器表示 20 位的地址空间呢？这在当时也给 Intel 工程师带来了极大的困扰，但是聪明的工程师想出来一种“段地址+偏移地址”的解决方案。段地址和偏移地址均为 16 位。此时，一个 1MB 中的地址，称为物理地址，按如下方式计算出来。

物理地址 = (段地址 \ll 4) + 偏移地址



寄存器

- 基本寄存器。寄存器是CPU内部的高速存储单元。IA-32处理器主要有8个通用寄存器eax, ebx, ecx, edx, ebp, esp, esi, edi、6个段寄存器cs, ss, ds, es, fs, gs、标志寄存器eflags、指令地址寄存器eip。
- 通用寄存器用于算术运算和数据传输。32位寄存器用于保护模式，为了兼容16位的实模式，每一个32位寄存器又可以拆分成16位寄存器和8位寄存器来访问。例如ax是eax的低16位，ah是ax高8位，al是ax的低8位。ebx, ecx, edx也有相同的访问模式。如下所示。

32 位	16 位	8 位 (高)	8 位 (低)
EAX	AX	AH	AL
EBX	BX	BH	BL
ECX	CX	CH	CL
EDX	DX	DH	DL

32 位	16 位	32 位	16 位
ESI	SI	EBP	BP
EDI	DI	ESP	SP

寄存器

- 某些通用寄存器有特殊用法：
乘除指令默认使用EAX。它常常被称为扩展累加器（extended accumulator）寄存器。
- CPU 默认使用 ECX 为循环计数器。
- ESP 用于寻址堆栈（一种系统内存结构）数据。它极少用于一般算术运算和数据传输，通常被称为扩展堆栈指针（extended stack pointer）寄存器。
- ESI 和 EDI 用于高速存储器传输指令，有时也被称为扩展源变址（extended source index）寄存器和扩展目的变址（extended destination index）寄存器。
- 高级语言通过 EBP 来引用堆栈中的函数参数和局部变量。除了高级编程，它不用于一般算术运算和数据传输。它常常被称为扩展帧指针（extended frame pointer）寄存器。

指令指针

指令指针（EIP）寄存器中包含下一条将要执行指令的地址。某些机器指令能控制EIP，使得程序分支转向到一个新位置。

EFLAGS 寄存器

EFLAGS（或 Flags）寄存器包含了独立的二进制位，用于控制 CPU 的操作，或是反映一些 CPU 操作的结果。有些指令可以测试和控制这些单独的处理器标志位。



汇编语言

- 采用Intel x86汇编语言格式;

```
mov ax, 3  
mov bx, 2  
add ax, bx
```

编译器采用nasm;

- 样例程序

```
1  section .text  
2  global main  
3  main:  
4  mov eax, 4;  
5  mov ebx, 1;  
6  mov ecx, msg;  
7  mov edx, 14;  
8  int 80h;  
9  mov eax, 1;  
10 int 80h;  
11 msg:  
12 db "hello world", 0ah, 0dh
```

1. 编译成可执行二进制过程:

```
nasm -f elf32 -o *.o *.s;  
gcc -m32 -g -o *.bin *.o ;
```

2. 编译成可装载的二进制镜像:

```
nasm -f bin -o *.o *.s;
```

3. 使用ld链接形成可执行二进制;

```
ld -m elf_i386 -o *.bin *.o;
```

操作系统启动

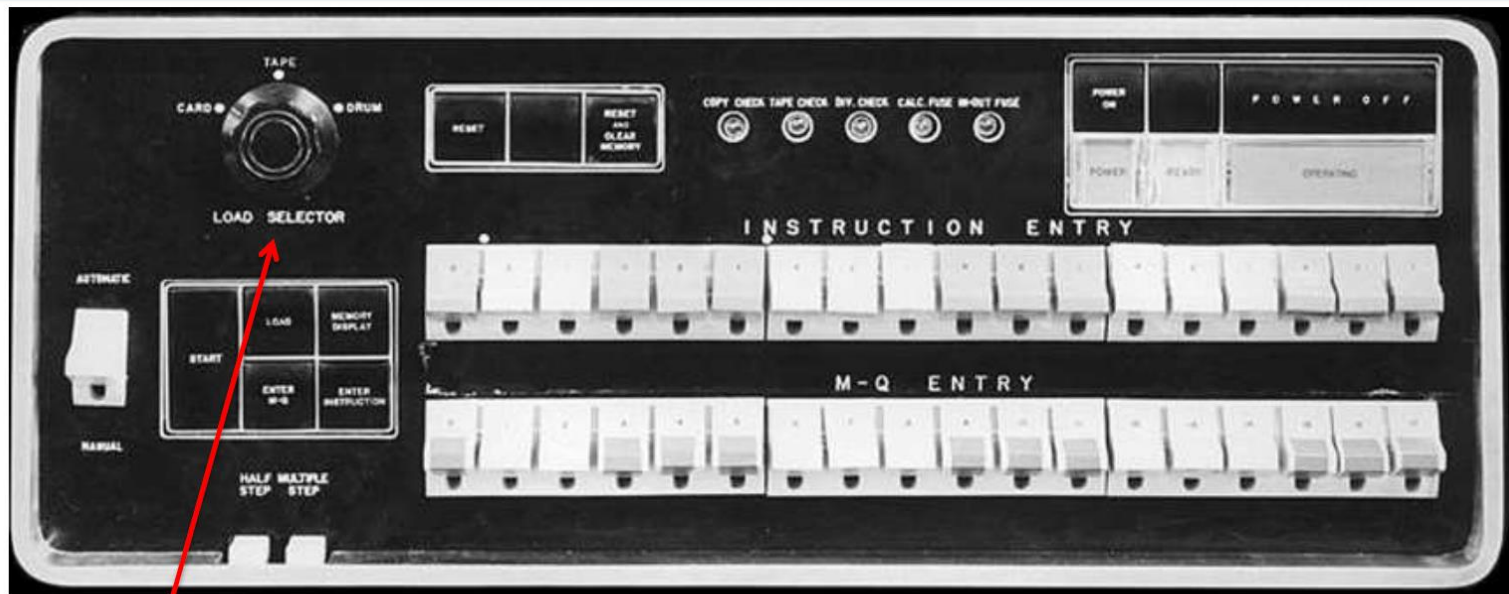
What runs first?

- Boot loader
 - A program that loads a bigger program (e.g., the OS)



操作系统启动

Booting



Load selector: Card, Tape, Drum



操作系统启动

Booting

GNU GRUB version 1.99~rc1

```
Ubuntu, with Linux 2.6.38-8-generic
Ubuntu, with Linux 2.6.38-8-generic (recovery mode)
Chainload to rEFIt
Chainload to ELILO
```

Use the ▲ and ▼ keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the commands
before booting or 'c' for a command-line.

操作系统启动

Multi-stage boot loader (chain loading)

- **First stage boot loader**
 - Often primitive enough that an operator could enter the code via front panel switches ... or it could sit in the first block of a disk
- **Second stage loader**
 - More sophisticated and included error checking
 - Second stage loader may give the user a choice:
 - Different operating systems
 - Boot a test program
 - Enable diagnostic modes (e.g., safe boot) in the OS

操作系统启动

Transfer of control

- When the boot loader finishes loading the OS, it transfers control to it
- The OS will initialize itself and load various modules as needed (for example, device drivers and various file systems)

操作系统启动

Intel/AMD PC Startup

- CPU reset at startup
- Start execution at `0xffffffff0`
 - Jump instruction to BIOS code in non-volatile memory
 - Near the top of 32-bit addressable memory map
 - *Reset vector*: jump to firmware initialization code
 - Processor starts in Real Mode
 - 20-bit address space (top 12 address lines held high)
 - Direct access to I/O, interrupts, and memory



操作系统启动

BIOS

- BIOS = Basic Input/Output System
- Found in Intel-based 16- and 32-bit PCs
- Code resident in ROM or non-volatile flash memory
- Background: CP/M (MS-DOS was almost a clone)
 - Console Command Processor (CCP): *user interface*
 - Basic Disk Operating System (BDOS): *generic code*
 - Basic Input/Output System (BIOS): *all the device interfaces*



操作系统启动

1、显示服务(Video Service——INT 10H)

00H — 设置显示器模式
0CH — 写图形象素
01H — 设置光标形状
0DH — 读图形象素
02H — 设置光标位置
0EH — 在Teletype模式下显示字符
03H — 读取光标信息
0FH — 读取显示器模式
04H — 读取光笔位置
10H — 颜色
05H — 设置显示页
11H — 字体

3、串行口服务(Serial Port Service——INT 14H)

00H — 初始化通信口
03H — 读取通信口状态
01H — 向通信口输出字符
04H — 扩充初始化通信口
02H — 从通信口读入字符
(1)、功能00H
功能描述：初始化通信口
入口参数：AH = 00H

4、键盘服务(Keyboard Service——INT 16H)

00H、10H — 从键盘读入字符
03H — 设置重复率
01H、11H — 读取键盘状态
04H — 设置键盘点击
02H、12H — 读取键盘标志
05H — 字符及其扫描码进栈
(1)、功能00H和10H
功能描述：从键盘读入字符
入口参数：AH = 00H——读键盘

2、直接磁盘服务(Direct Disk Service——INT 13H)

00H — 磁盘系统复位
0EH — 读扇区缓冲区
01H — 读取磁盘系统状态
0FH — 写扇区缓冲区
02H — 读扇区
10H — 读取驱动器状态
03H — 写扇区
11H — 校准驱动器
04H — 检验扇区
12H — 控制器RAM诊断
05H — 格式化磁道
13H — 控制器驱动诊断
06H — 格式化坏磁道
14H — 控制器内部诊断

操作系统启动

PC Startup

- BIOS executes:
 - Power-on self-test (POST)
 - Detect video card's BIOS – execute video initialization
 - Detect other device BIOS – initialize
 - Display start-up screen
 - Brief memory test
 - Set memory, drive parameters
 - Configure Plug & Play devices: PCIe, USB, SATA, SPI
 - Assign resources (DMA channels & IRQs)
 - Identify boot device:
 - Load block 0 (Master Boot Record) to `0x7c00` and jump there

操作系统启动

Booting Windows (NT/Windows 20xx,7,8)

- BIOS-based booting
 - The BIOS does *not* know file systems but can read disk blocks
- **MBR = Master Boot Record** = Block 0 of disk (512 bytes)
 - Small boot loader (chain loader, ≤ 440 bytes)
 - Disk signature (4 bytes)
 - Disk partition table (16 bytes per partition * 4)
- BIOS firmware loads and executes the contents of the MBR
- MBR code scans through partition table and loads the **Volume Boot Record (VBR)** for that partition
 - Identifies partition type & size
 - Contains **Instruction Program Loader** that executes startup code
 - IPL reads additional sectors to load **BOOTMGR (Windows 7, 8)**
 - The loader is called **NTLDR** for Windows NT, XP, 2003

操作系统启动

Booting other systems on a PC

- Example: GRUB (Grand Unified Boot Loader)
- MBR contains GRUB Stage 1
 - Or another boot loader that may boot GRUB Stage 1 from the Volume Boot Record
- Stage 1 loads Stage 2
 - Present user with choice of operating systems to boot
 - Optionally specify boot parameters
 - Load selected kernel and run the kernel
 - For Windows (which is not Multiboot compliant),
 - Run MBR code or Windows boot menu
 - **Multiboot specification:**
 - Free Software Foundation spec on loading multiple kernels using a single boot loader

操作系统启动

Good-bye BIOS: PCs and UEFI

- ~2005: Unified Extensible Firmware Interface (UEFI)
 - Originally called EFI; then changed to UEFI
You still see both names in use
- Created for 32- and 64-bit architectures
 - Including Macs, which also have BIOS support for Windows
- Goal:
 - Create a successor to the BIOS
 - no restrictions on running in 16-bit 8086 mode with 20-bit addressing



操作系统启动

UEFI Includes

- Preserved from BIOS:
 - Power management (Advanced Configuration & Power Interface, ACPI)
 - System management components from the BIOS
- Support for larger disks
 - BIOS only supported 4 partitions per disk, each up to 2.2 TB per partition
 - EFI supports max partition size of 9.4 ZB (9.4×10^{21} bytes)
- Pre-boot execution environment with direct access to all memory
- Device drivers, including the ability to interpret architecture-independent EFI Byte Code (EBC)
- Boot manager: lets you select and load an OS
 - *No need for a dedicated boot loader (but they may be present anyway)*
 - *Stick your files in the EFI boot partition and EFI can load them*
- Extensible: extensions can be loaded into non-volatile memory

操作系统启动

UEFI Booting

- No need for MBR code (*ignore block 0*)
- Read GUID Partition Table (GPT)
 - Describes layout of the partition table on a disk (blocks 1-33)
- EFI understands Microsoft FAT file systems
 - Apple's EFI knows HFS+ in addition
- Read programs stored as *files* in the EFI System Partition:
 - Windows 7/8, Windows 2008/2012 (64-bit Microsoft systems):
 - **Windows Boot Manager** (BOOTMGR) is in the EFI partition
 - NT (IA-64): IA64ldr
 - Linux: elilo.efi (ELILO = EFI Linux Boot Loader)
 - OS X: boot.efi

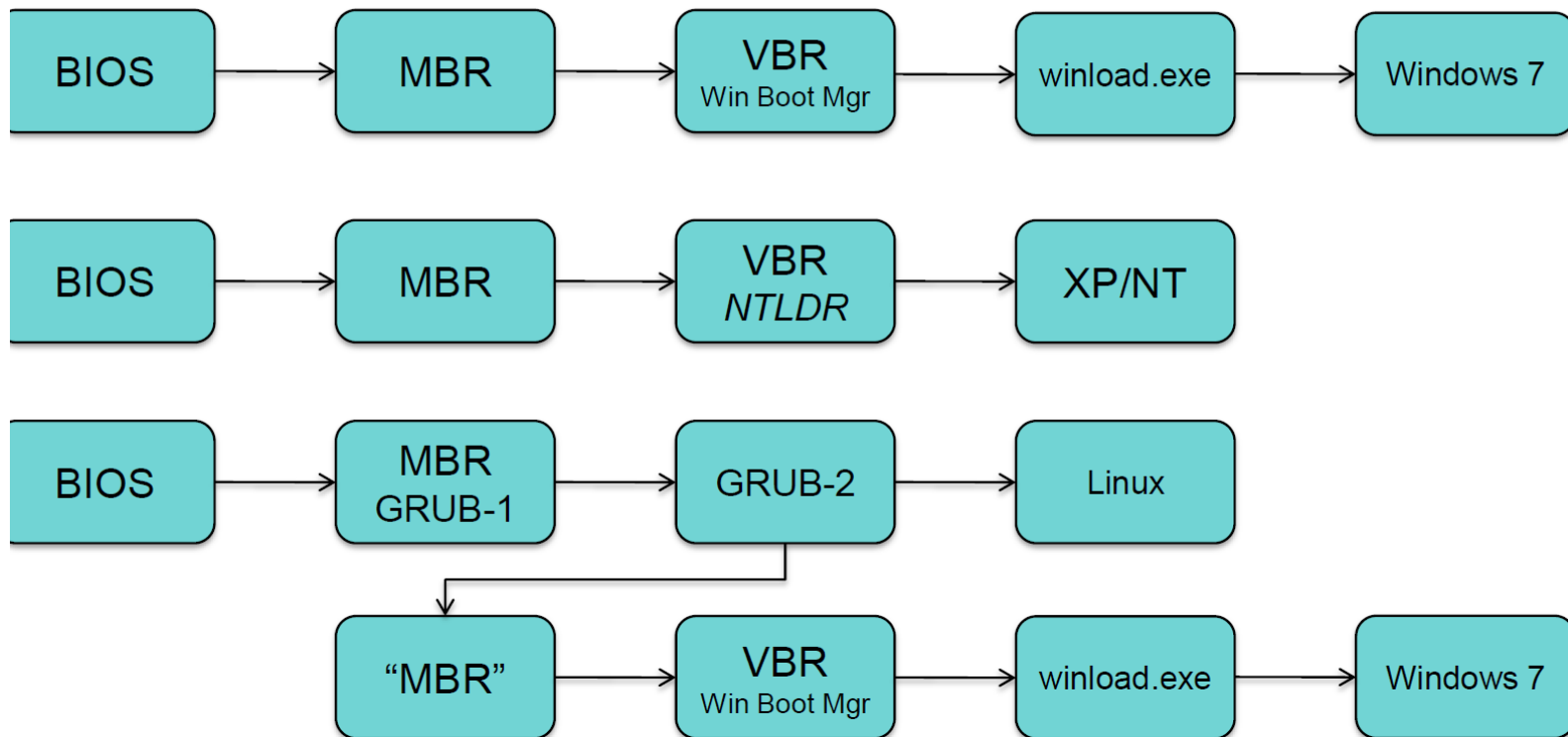
操作系统启动

Non-Intel Systems

- Power on: execute boot ROM code (typically NOR Flash)
 - Often embedded in the CPU ASIC
- Boot ROM code detects boot media
 - Loads first stage boot loader (sometimes to internal RAM)
 - Initialize RAM
 - Execute boot loader
- Second stage boot loader loads kernel into RAM
 - For Linux, typically GRUB for larger systems
 - uBoot for embedded systems
 - Set up network support, memory protection, security options

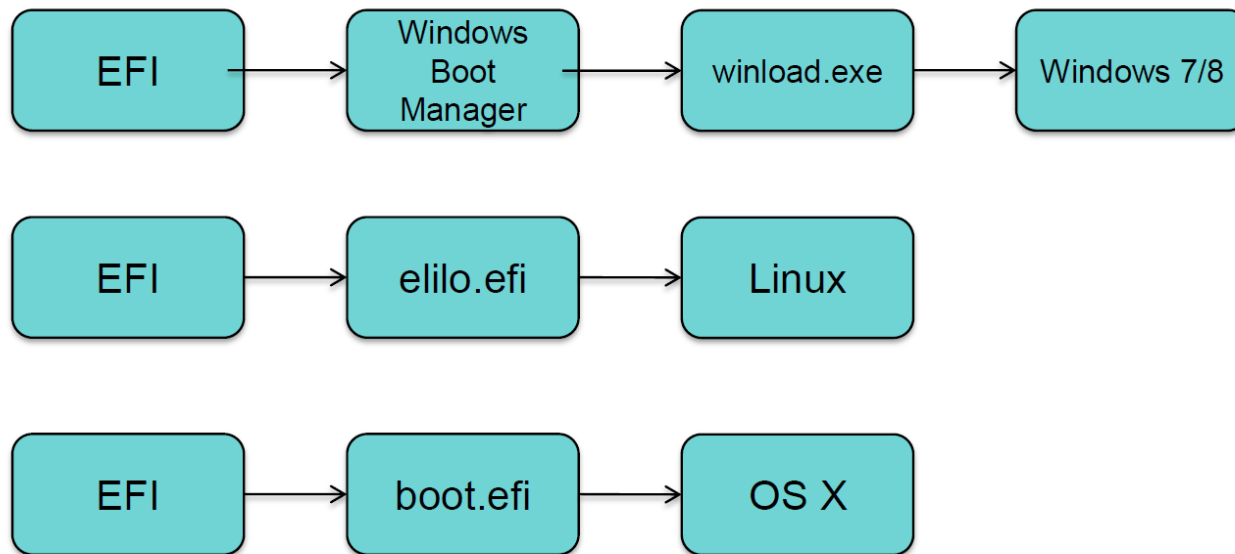
操作系统启动

Summary



操作系统启动

Summary





保护模式启动代码

```
9 ;
10 ;bits 16
11
12 [bits 16]
13 mov ah, 0x0e ;设置tele-type mode
14 mov al, 'H' ;设置待显示字符
15 int 0x10 ;screen相关的中断
16 mov al, 'e'
17 int 0x10
18 mov al, 'l'
19 int 0x10
20 mov al, 'l'
21 int 0x10
22 mov al, 'o'
23 int 0x10
24
25 jmp $ ;跳到当前地址，无限循环
26 times 510-($-$$) db 0 ;填充程序到512个字节
27 dw 0xaa55 ;让bios识别此扇区为可启动扇区的魔数
28
```

1. `nasm -f bin *.asm -o *.bin;`
2. `qemu-img create hd.img 10m;`
3. `dd if=*.bin of=hd.img bs=512 count=1 seek=0 conv=notrunc;`
4. `qemu-system-i386 -hda hd.img -serial null -parallel stdio ;`

实模式启动代码

Intel x86系列CPU有实模式和保护模式，实模式从8086开始就有，保护模式从80386开始引入。为了兼容，Intel x86系列CPU都支持实模式。现代操作系统都是运行在保护模式下（Intel x86系列CPU）。计算机启动时，默认的工作模式是实模式，为了让内核能运行在保护模式下，Bootloader需要从实模式切换到保护模式，切换步骤如下：

- 1. 准备好GDT(Global Descriptor Table)
- 2. 关中断
- 3. 加载GDT到GDTR寄存器
- 4. 开启A20，让CPU寻址大于1M
- 5. 开启CPU的保护模式，即把cr0寄存器第一个bit置1
- 6. 跳转到保护模式代码

GDT是Intel CPU保护模式运行的核心数据结构，所有保护模式操作的数据都从GDT表开始查找，[这里](#)有GDT的详细介绍。

31				16		15				0			
Base 0:15						Limit 0:15							
63		56		55 52		51 48		47		40 39		32	
Base 24:31				Flags		Limit 16:19		Access Byte				Base 16:23	



参考资料

- [x86汇编\(Intel汇编\)入门](#)
- 《Intel汇编语言程序设计》 第1-8章
- 《从实模式到保护模式》 第1-8章
- <http://c.biancheng.net/makefile/>
- How to write a simple operating system
- The little book about OS development

谢谢