

汇编语言与逆向技术

第3章 IA-32处理器

王志

zwang@nankai.edu.cn updated on 2022-09-28

南开大学 网络空间安全学院 2022-2023学年



本章知识点

- 计算机体系结构
- IA-32处理器体系结构
- IA-32的内存管理
- Hello World程序





计算机体系结构



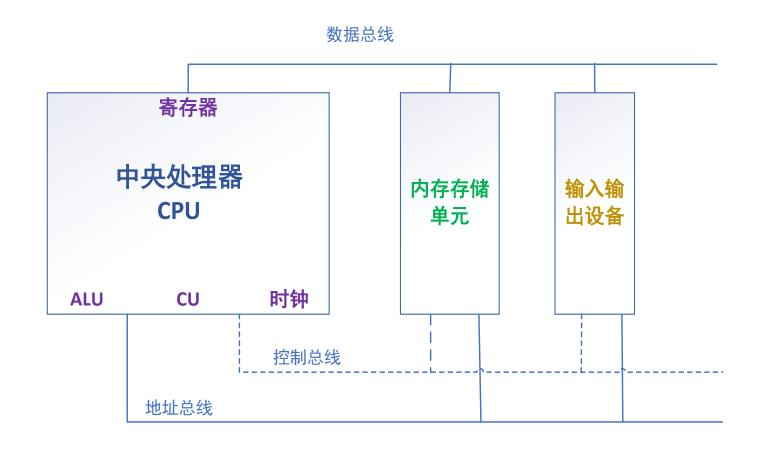
计算机基本概念

- 计算机基本结构
- 指令执行周期
- 内存的读取
- 程序是如何运行的



龙公允能日新月异

计算机基本结构





计算机基本结构

- 中央处理器(CPU,Central Processor Unit)进行计算和逻辑操作的 地方
 - 寄存器 (Register)
 - 时钟 (clock)
 - 控制单元 (CU, Control Unit)
 - 算数逻辑单元(ALU,Arithmetic Logic Unit)





CPU

- 寄存器: 数据存储, 数量有限
- •时钟:同步CPU的内部操作
- 控制单元: 控制机器指令的执行步骤
- 算数逻辑单元: 算术运算、逻辑运算



CPU时钟

- ·每个时钟周期CPU完成一步操作
- •时钟频率=1/时钟周期
- ·时钟频率反映了CPU速度的快慢

The World's Fastest Desktop Processor

13th Gen Intel® Core™ i9-13900K

Fastest P-Cores

5.8GHz

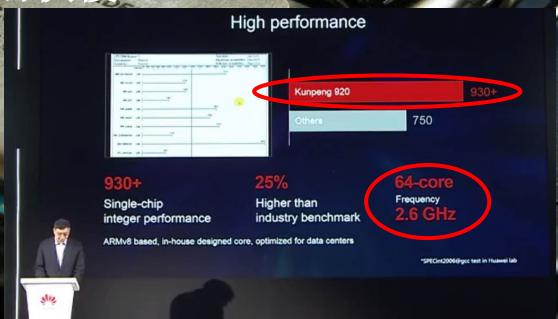
Double E-Cores

24C/32T

Larger L2 Caches

2MB per P-core 4MB per E-core cluster

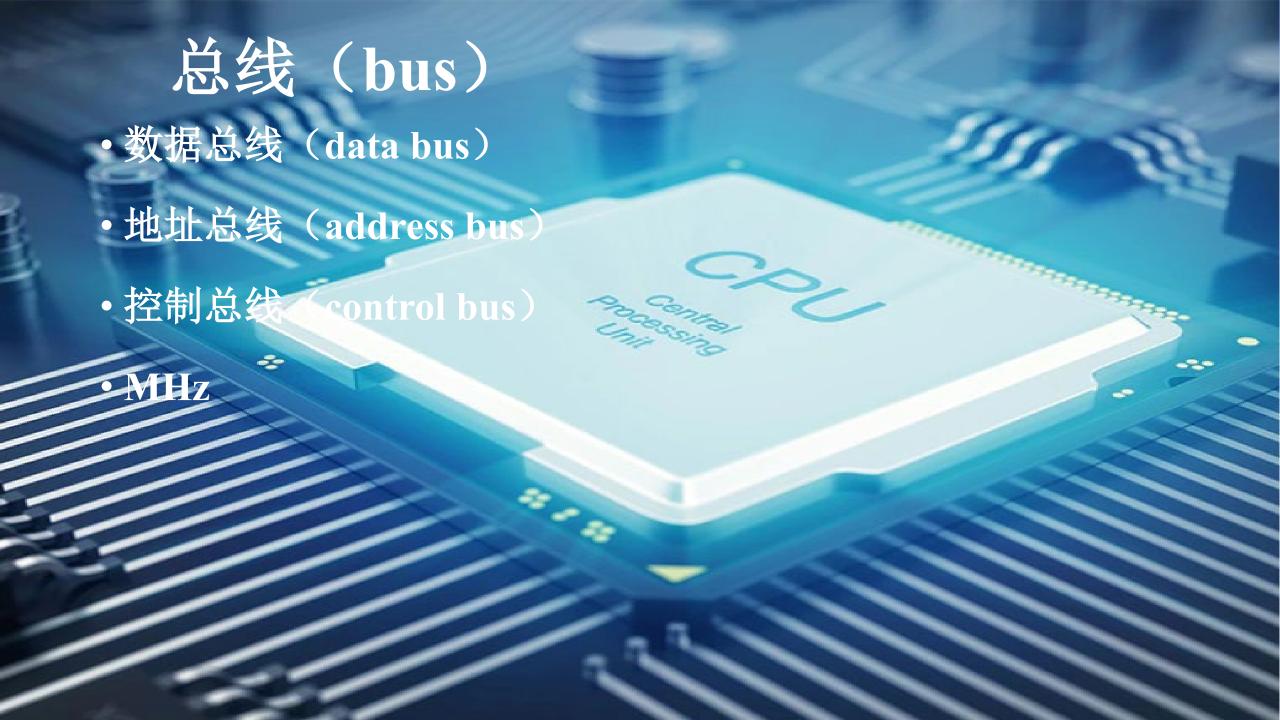
Delivering up to 15% ST and 41% MT Performance



内存存储单元

- · Memory storage unit存放指令和数据的地方
- ·核心频率133MHz~200MHz

THE PERSON NAMED IN	DDR SDRAM Standard	Internal rate (MHz)	Bus clock (MHz)	Prefetch	Data rate (MT/s)	Transfer rate (GB/s)	Voltage (V)
	SDRAM	100-166	100-166	1n	100-166	0.8-1.3	3.3
	DDR	133-200	133-200	2n	266-400	2.1-3.2	2.5/2.6
	DDR2	133-200	266-400	4n	533-800	4.2-6.4	1.8
	DDR3	133-200	533-800	8n	1066-1600	8.5-14.9	1.35/1.5
	DDR4	133-200	1066-1600	8n	2133-3200	17-21.3	1.2



指令执行周期

- 单条机器指令的执行包括一系列操作
 - 取指令: 指令指针IP
 - •解码:控制单元CU确定执行什么操作
 - 取操作数: 从内存读操作数
 - 执行: 算数逻辑单元ALU
 - 存储输出操作数: 向内存写入





IA-32处理器体系结构

IA-32处理器体系结构

- IA-32 (Intel Architecture 32-bit) 英特尔32位体系结构
 - 1985年 80386 CPU首先使用
 - 32位内存地址
 - 32位数据操作数



工作模式

- 实地址模式(Real-Address Mode)
 - 16位, 8086程序设计环境
- 保护模式 (Protected Mode)
 - 32位, IA-32程序设计环境
 - 虚拟8086模式: 执行8086程序





实地址模式

- 16位的8086程序设计环境
 - 20条地址线
 - 存储空间1MB(2的20次方)





保护模式

- IA-32 CPU的存储管理和保护机制
 - 多任务操作系统
 - •程序有独立的4GB内存存储空间(2的32次方)





地址空间

- IA-32 CPU 4GB 地址空间
 - 32位的寻址上限
- 8086只有1MB地址空间





寄存器 (Register)

- · 寄存器是CPU内部的高速存储单元
 - 比内存的访问速度快很多
 - 优化循环结构执行速度,把循环计数变量放到寄存器中。









段寄存器

• CS: Code Segement, 代码段寄存器

• SS: Stack Segment, 栈段寄存器

• DS: Data Segment,数据段寄存器

• ES: Extra(Data) Segment, 数据段寄存器

• FS: Data Segment,数据段寄存器

• GS: Data Segment, 数据段寄存器

```
00261DD7
8B55 E0
mov edx,dword ptr ss:[ehp-0x20]

00261DDA
8B02
mov eax,dword ptr ds:[edx]

00261DDC
8945 D8
mov dword ptr ss:[ebp-0x28],eax
```





EFLAGS寄存器

ID X ID Flag (CPUID support) DF C Direction Flag

VIP X Virtual Interrupt Pending IF X Interrupt Enable Flag

VIF X Virtual Interrupt Flag TF X Trap Flag

C X Alignment Check SF S Sign Flag

VM X Virtual 8086 Mode ZF S Zero Flag

RF X Resume Flag AF S Auxiliary Carry Flag

NT X Nested Task PF S Parity Flag

IOPL X I/O Privilege Level CF S Carry Flag

OF S Overflow Flag

Bit Positions shown as "0" or "1" are Intel reserved.

S = Status Flag C = Control Flag X = System Flag





零标志

• 零标志(ZF): 若算数或者逻辑运算结果为0则将其置1,反之清零

- xor eax, eax
- jz





进位标志

- 进位标志(CF): 在无符号算术运算的结果最高有效位(most-significant bit)发生进位或借位则将其置1,反之清零。
 - add eax, 0xffffffff
 - jc





溢出标志

- 溢出标志(OF): 在有符号算术运算的结果是较大的正数或较小的负数,并且目的操作数无法容纳时,将该位置1,反之清零。
- 这个标志为带符号整型运算指示溢出状态。





符号标志

- · 符号标志(SF): 该标志被设置为有符号整型的最高有效位。
- 0表示算术或者逻辑运算结果为正
- 1表示算数或者逻辑运算结果为负





奇偶标志

- 奇偶标志 (PF): 如果结果的最低有效字节(least-significant byte) 包含偶数个1位则该位置1,否则清零。
- 数据校验



辅助进位标志

- 辅助进位标志(AC): 如果算术操作在结果的第3位发生进位或借位则将该标志置1,否则清零。
- 这个标志在BCD(binary-code decimal)算术运算中被使用。





控制标志

- •方向标志(DF)
 - 控制串指令(MOVS, CMPS, SCAS, LODS以及STOS)
 - 设置DF标志使得串指令自动递减(从高地址向低地址方向处理字符串), 清除该标志则使得串指令自动递增
 - · STD以及CLD指令分别用于设置以及清除DF标志。





系统标志

- TF: 将该位设置为1以允许单步调试模式,清零则禁用该模式。
- 调试器的单步调试功能





指令指针

- · 指令指针寄存器(EIP)存放下一条机器指令的内存地址。
- 跳转指令可以修改指令指针寄存器





IA-32内存管理



IA-32内存管理

- IA-32保护模式的内存管理比实地址模式要复杂
 - 多任务
 - 多用户
 - 段模式、页模式
 - 页模式也是基于段模式的,通常称为段页式





平坦模式 (FLAT)

- ·每个程序有独立的4GB虚拟地址空间
 - 数据
 - 指令
 - 数据 ←→ 指令
- 虚拟地址到物理地址的转换是透明的





段管理

- •一般保护模式的程序有3个段
 - 代码段, CS
 - 数据段, DS
 - 堆栈段, SS
- 段是一块内存空间





段管理

- GDT(Global Descriptor Table)全局描述符表
 - 整个系统只有一个GDT(64bit)
 - Intel提供了一个寄存器GDTR用来存放GDT的入口地址
- LDT(Local Descriptor Table)局部描述符表
 - 每个程序都有自己的LDT
 - IA-32为LDT的入口地址也提供了一个寄存器LDTR
 - 因为在任何时刻只能有一个任务在运行,所以LDTR也只需要有一个





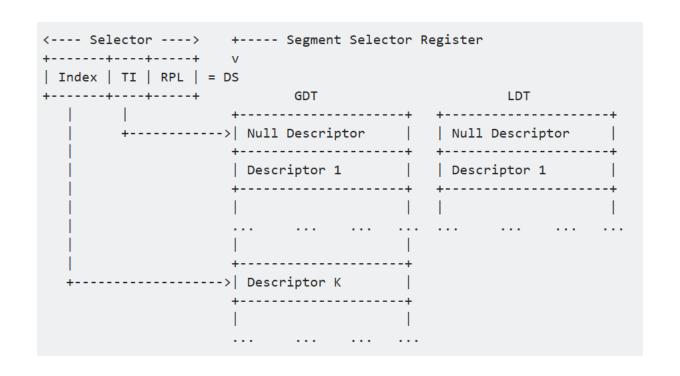
段寄存器

- index(13bits),段描述符在表中的索引
- TI(1bit),0是GDT,1是LDT
- RPL (2bits), Request Privilege Level, 权限
 - Ring 0, Kernel Mode
 - Ring 3, User Mode

Nankai University



段描述符

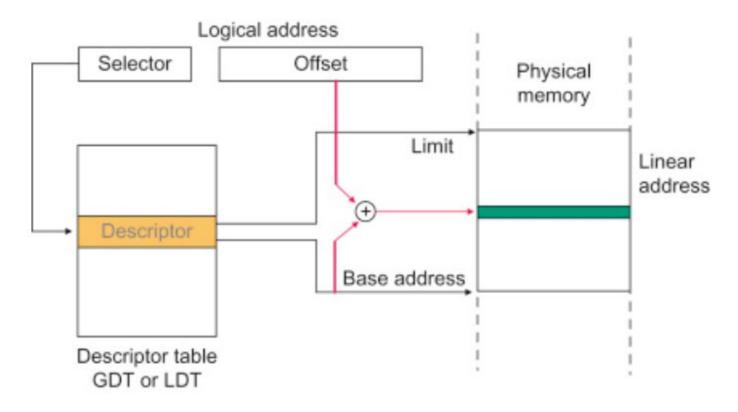






允公允铭日新月异

段模式





分页机制

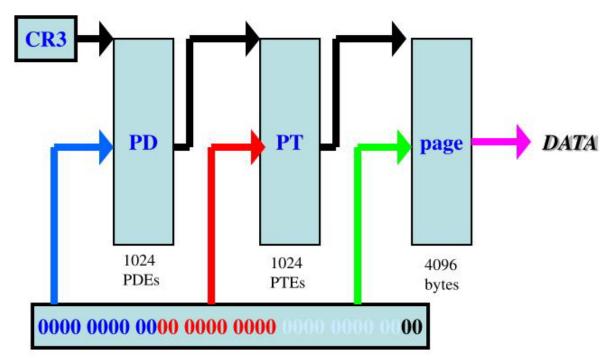
- 段又被分割成内存页 (page)
- 内存页统一为4096字节的内存块
- 提高内存的利用率,减少内存碎片
- 页交换,不使用的内存页被交换到硬盘上
 - 虚拟内存空间大于实际的物理内存空间
 - 页交换降低程序执行速度





分页机制

Virtual Address Translation



页目录表(PDT)的每一项元素 称为页目录表项(PDE) 每个页目录表项指向一个页表 (PTT)

每个页表的大小为4KB,即一个 页表可以存储**1024个**页表项(PTE)

页表(PTT)的每一个元素称为 页表项(PTE) 页表项(PTE)所指向的才是真 正的物理页





实验1 Hello World程序

汇编、链接和运行程序

- · 源文件: 用文本编辑器编写的asm文本文件
- 汇编: 汇编器把汇编源文件翻译成机器语言, 生成目标文件(.obj)
- 链接: 链接器从库中复制所需的过程,并将其同目标文件合并在
 - 一起生成可执行文件(.exe)





hello.asm

.model flat, stdcall option casemap :none include \masm32\include\windows.inc include \masm32\include\kernel32.inc include \masm32\include\masm32.inc includelib \masm32\lib\kernel32.lib includelib \masm32\lib\masm32.lib





hello.asm

.data

HelloWorld db "Hello World!", 0

.code

start:

invoke StdOut, addr HelloWorld

invoke ExitProcess, 0

end start





- .386
 - 允许汇编80386处理器的非特权指令,禁用其后处理器引入的汇编指令
- .model 初始化程序的内存模式
 - flat: 平坦模式, 4GB内存空间
 - stdcall:调用约定, stdcall是Win32 API函数的调用约定





- option casemap: none
 - 大小写敏感
- include ...inc 函数的常量和声明
- includelib ...lib 链接库





- .DATA
 - 定义已初始化数据段的开始
- .CODE
 - 定义代码段的开始
- start: , 指令标号, 标记指令地址





允公允帐日新月异

- StdOut, masm32.inc中定义的函数,将内存数据输出到命令行窗口上
- ExitProcess, Kernel32.inc 中定义的函数,退出程序执行





- END start
 - 标记模块的结束
 - 指定程序的入口点





编译

- \masm\bin\ml /c /Zd /coff hello.asm
- ml 程序可以用来汇编并链接一个或多个汇编语言源文件
- ml的命令行选项是大小写敏感的





编译

- /c Assemble without linking
 - 只编译、不链接
- /Zd Add line number debug info
 - 在目标文件中生成行号信息
- /coff generate COFF format object file
 - 生成Microsoft公共目标文件格式(common object file format)的文件





链接

- \masm32\bin\link /SUBSYSTEM:CONSOLE hello.obj
- link.exe 链接器,将obj文件合并,生成可执行文件
- /SUBSYSTEM:CONSOLE, 生成命令行程序





汇编语言与逆向技术

第3章 IA-32处理器

王志

zwang@nankai.edu.cn updated on 2022-09-28

南开大学 网络空间安全学院 2022-2023学年