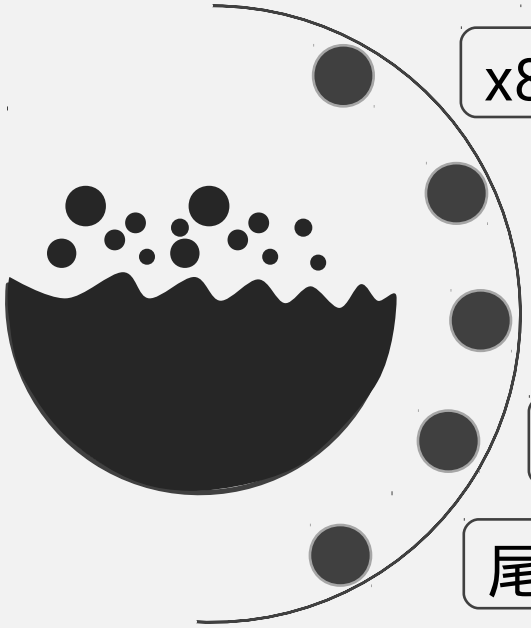
Linux x86 内核态调试技术

李盛秋

正片开始之前

您是否……

- 了解基本的 x86 汇编语言？
- 了解“用户态”、“内核态”？
- 了解“中断”？



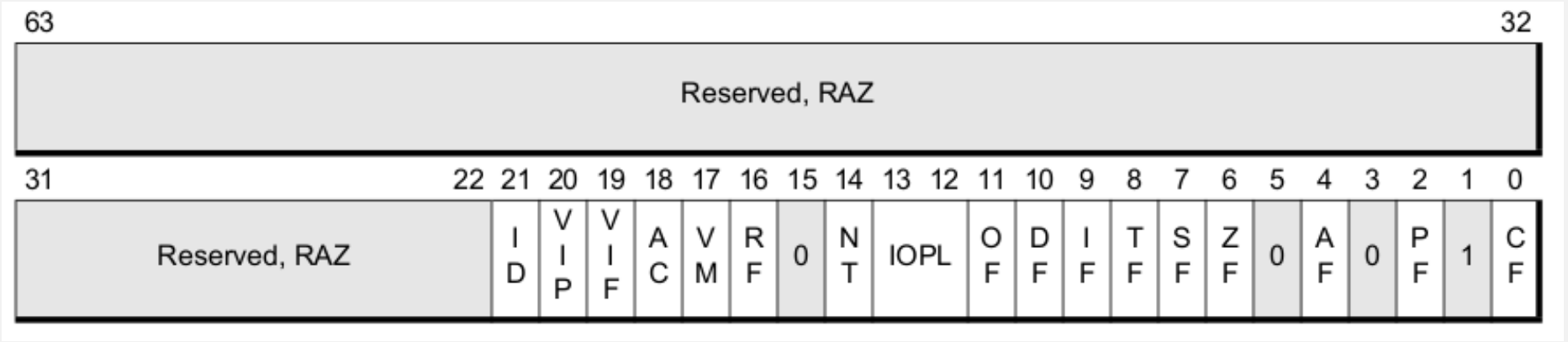
- x86 架构调试功能简介
- Linux 用户态程序调试方式
- 内核态程序调试原理
- 其他常见 Linux 内核态程序调试手段
- 尾声

x86 架构调试功能简介

标志寄存器调试位

标志寄存器用于标志当前程序的相关状态。

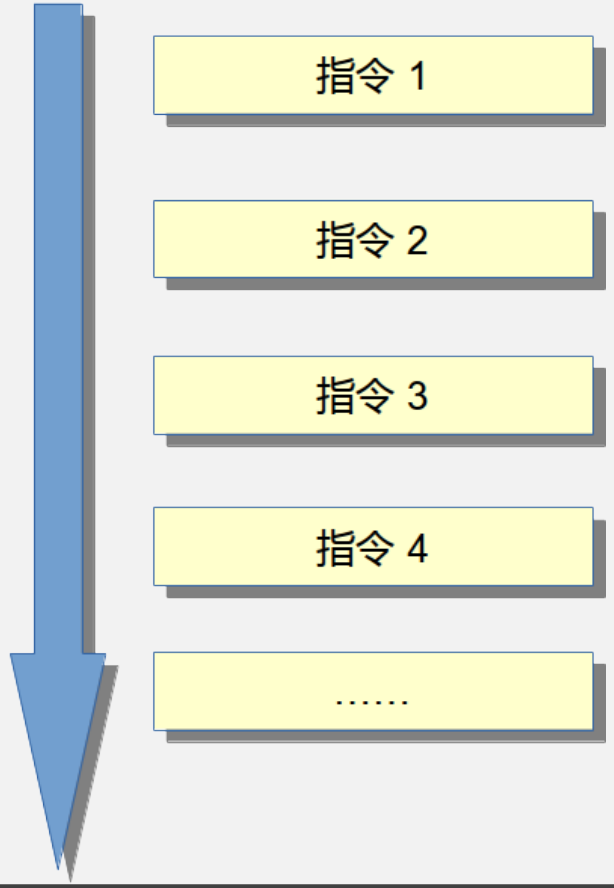
x86_64 标志寄存器 RFLAGS ：



TF ： Trap Flag 单步运行标志位

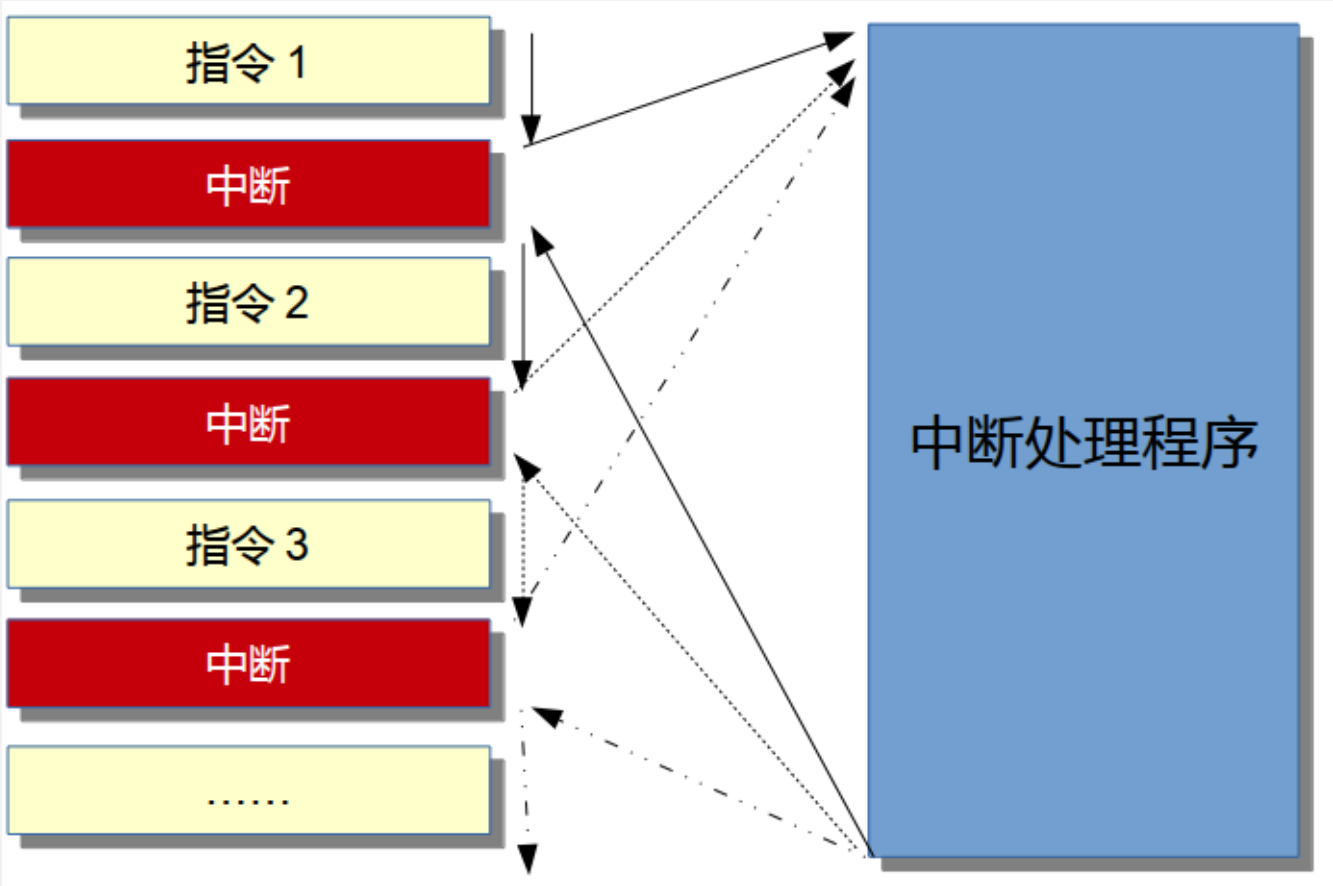
单步调试执行流程

TF = 0 ，正常执行



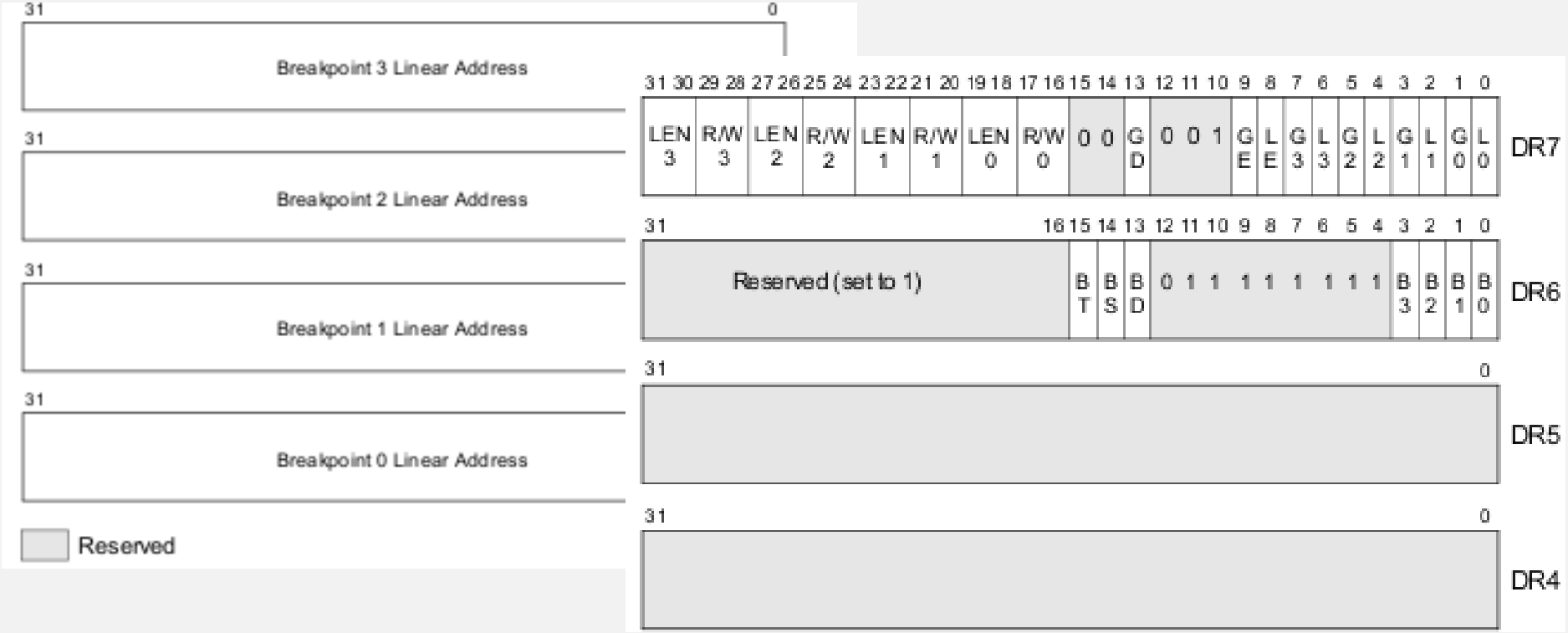
单步调试执行流程

TF = 1 ，单步执行



调试寄存器

调试寄存器用于在硬件层面上支持程序断点。
x86 调试寄存器 DR0 ~ DR7 ：



调试支持功能概览

1.调试异常（中断）
2.断点异常（中断）
3.断点地址寄存器（ DR0 ~ DR3 ）
4.断点状态寄存器（ DR6 ）
5.断点控制寄存器（ DR7 ）
6.TF(Trap Flag) ， EFLAGS
7.RF(Resume Flag) ， EFLAGS
8.T(Trap Flag) ， TSS
9.软件断点指令（ INT 3 ）
10.最近分支记录（ LBR ）

Linux 用户态程序调试方式

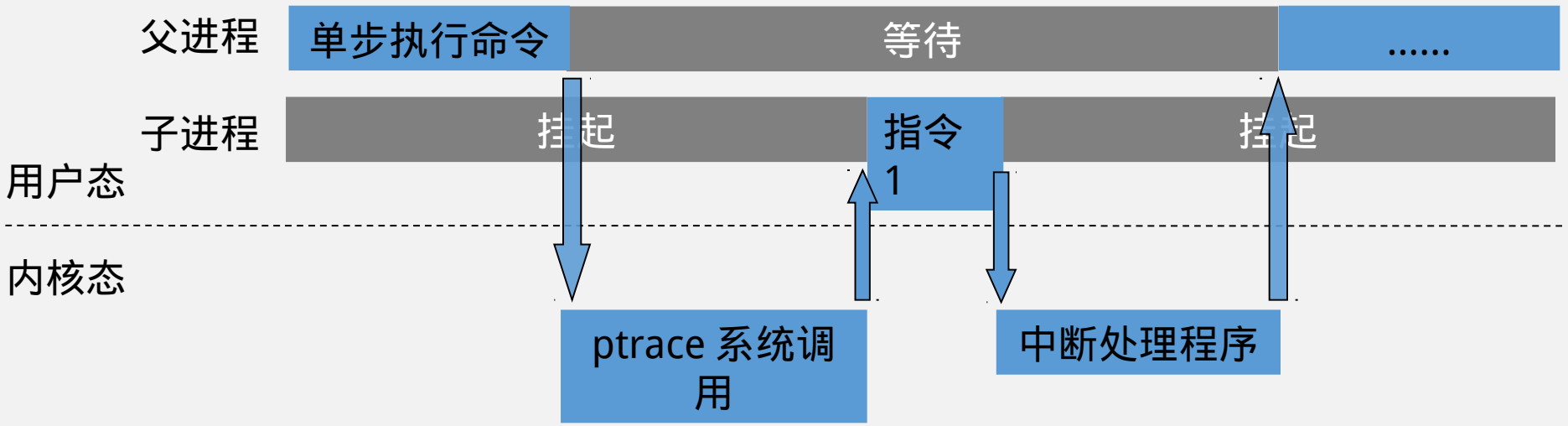
ptrace 系统调用

```
int ptrace(int request, int pid, int addr, int data);
```

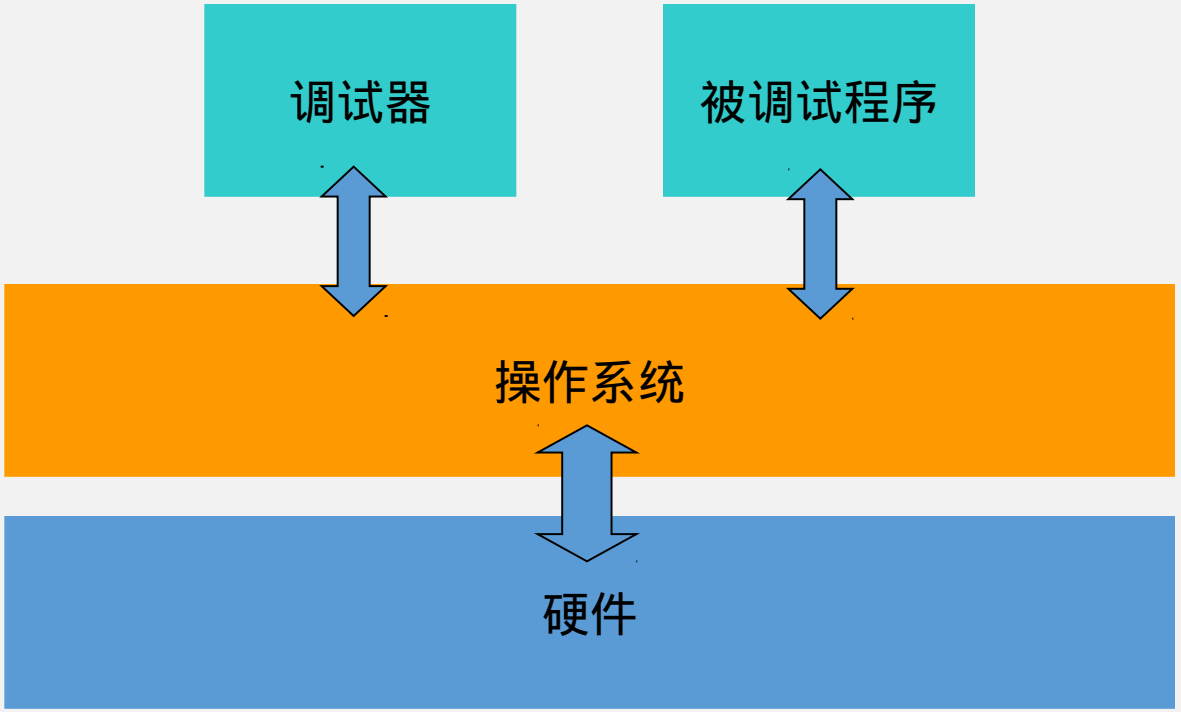
request:

- PTRACE_TRACEME
- PTRACE_SINGLESTEP
- PTRACE_GETREGS
- ……

用户态程序单步调试过程



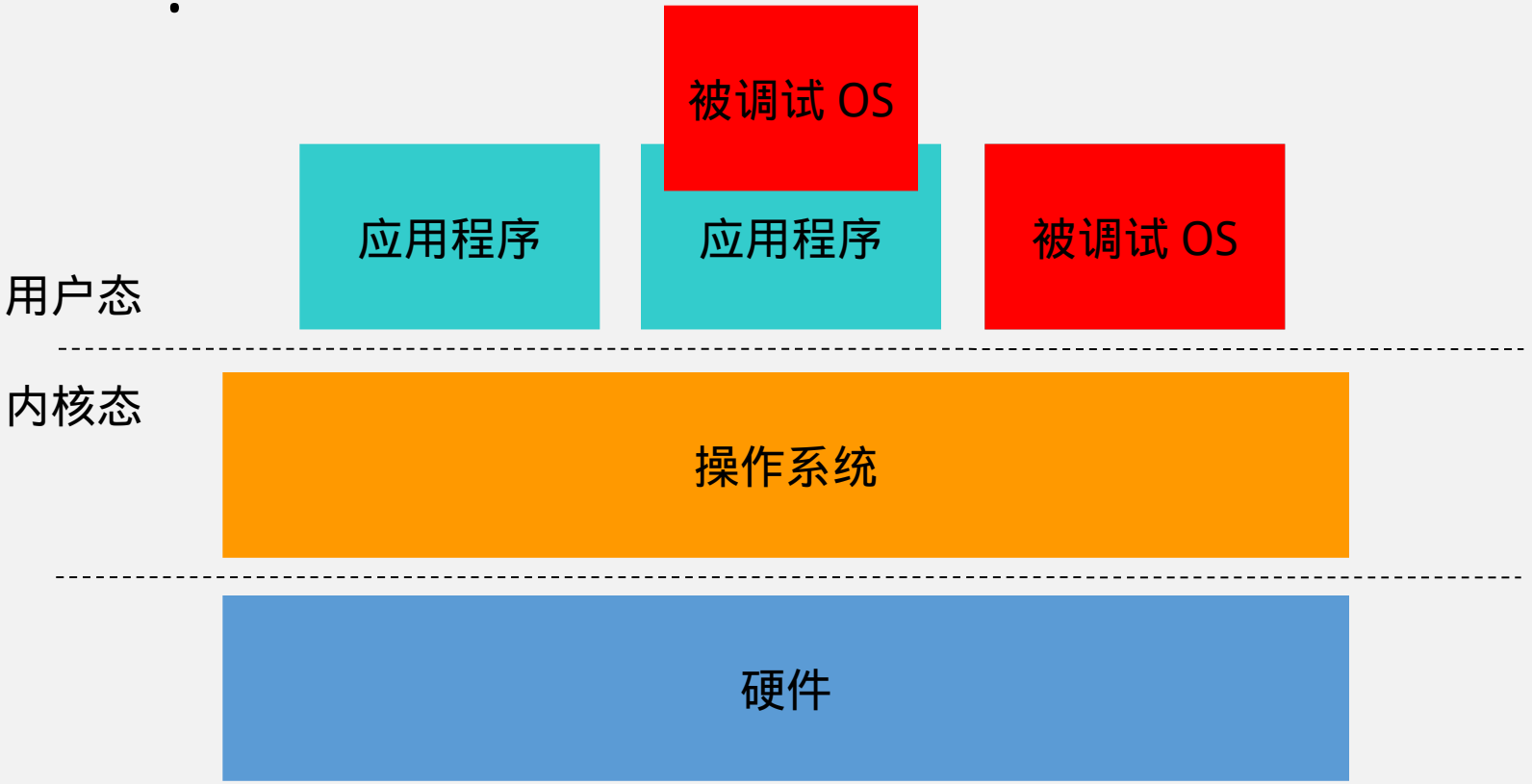
用户态程序单步调试过程



内核态程序调试原理

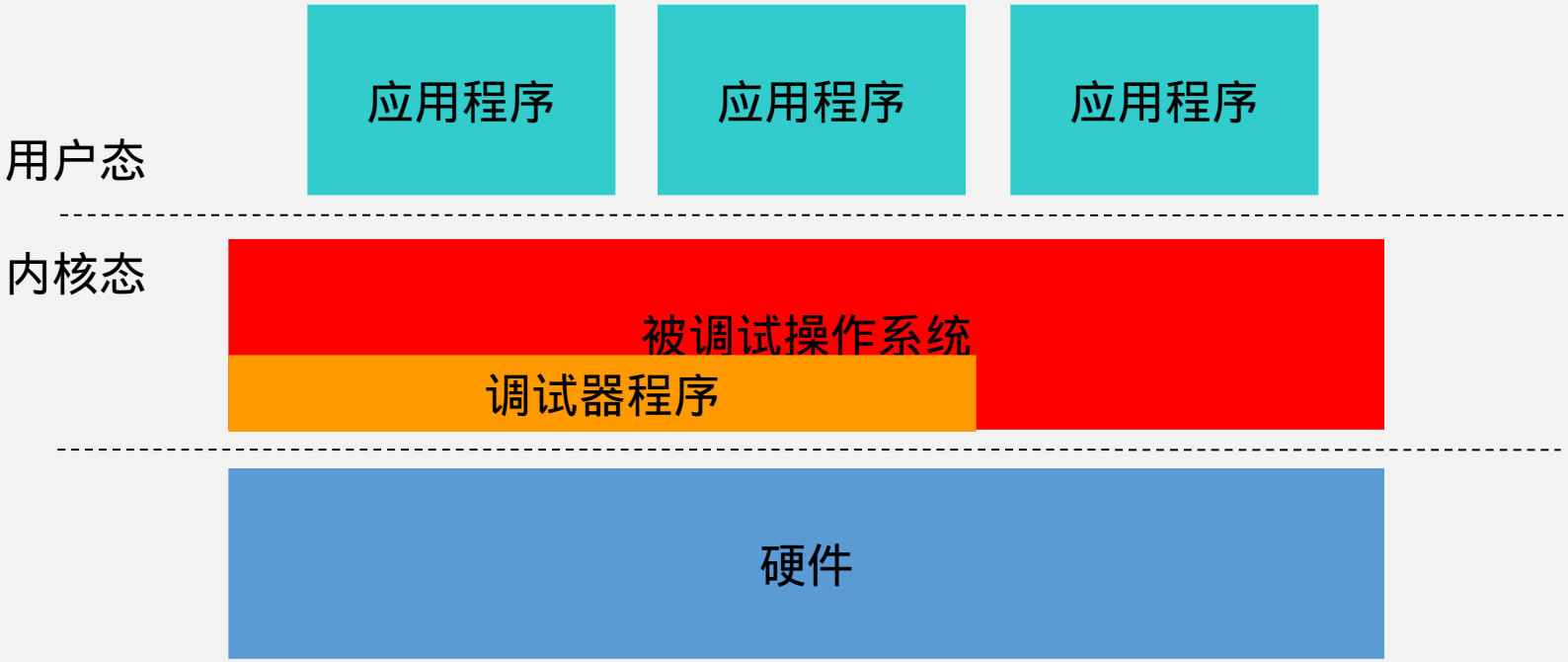
如何调试操作系统 / 内核态程序

?



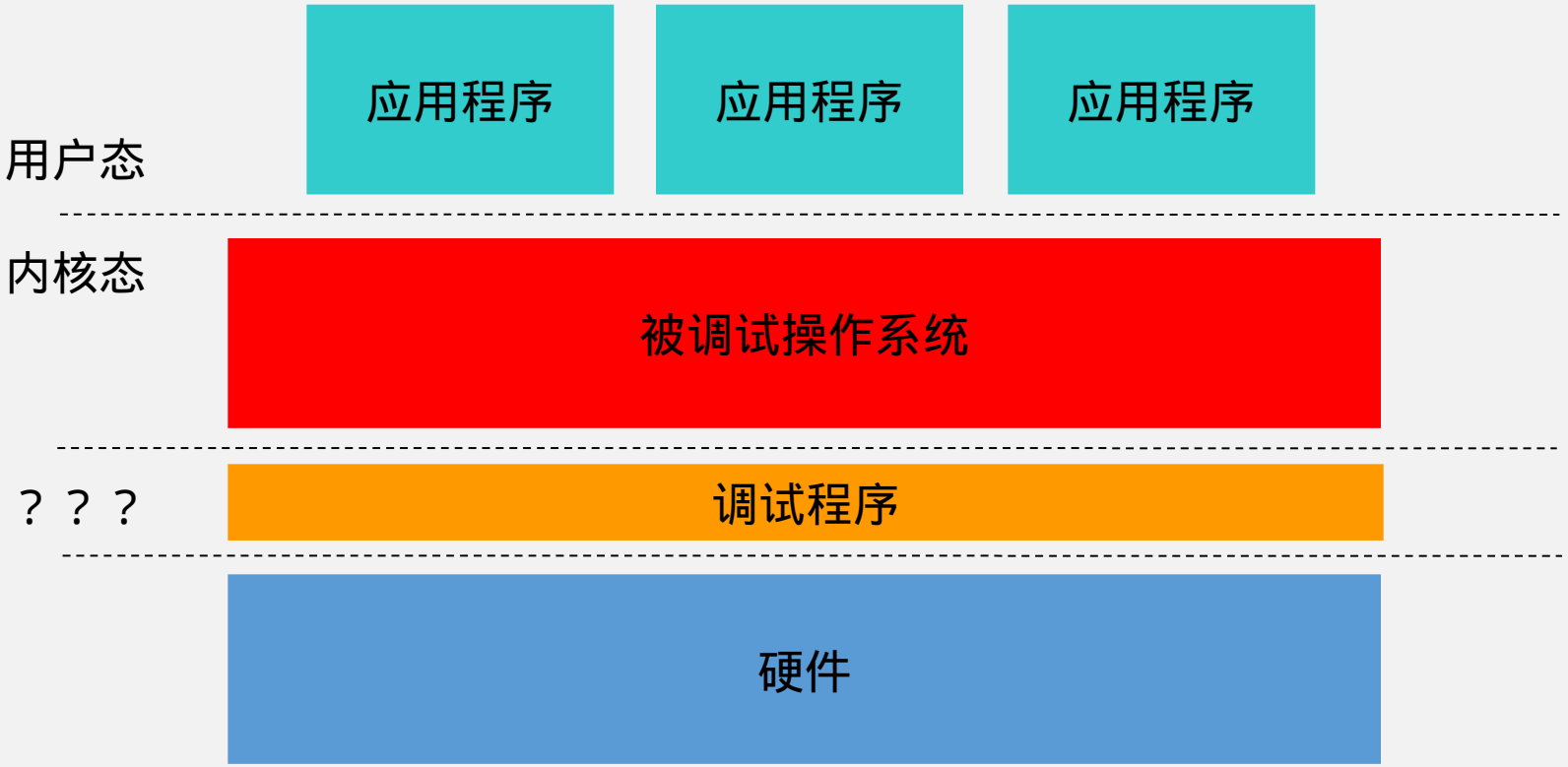
如何调试操作系统 / 内核态程序

?



如何调试操作系统 / 内核态程序

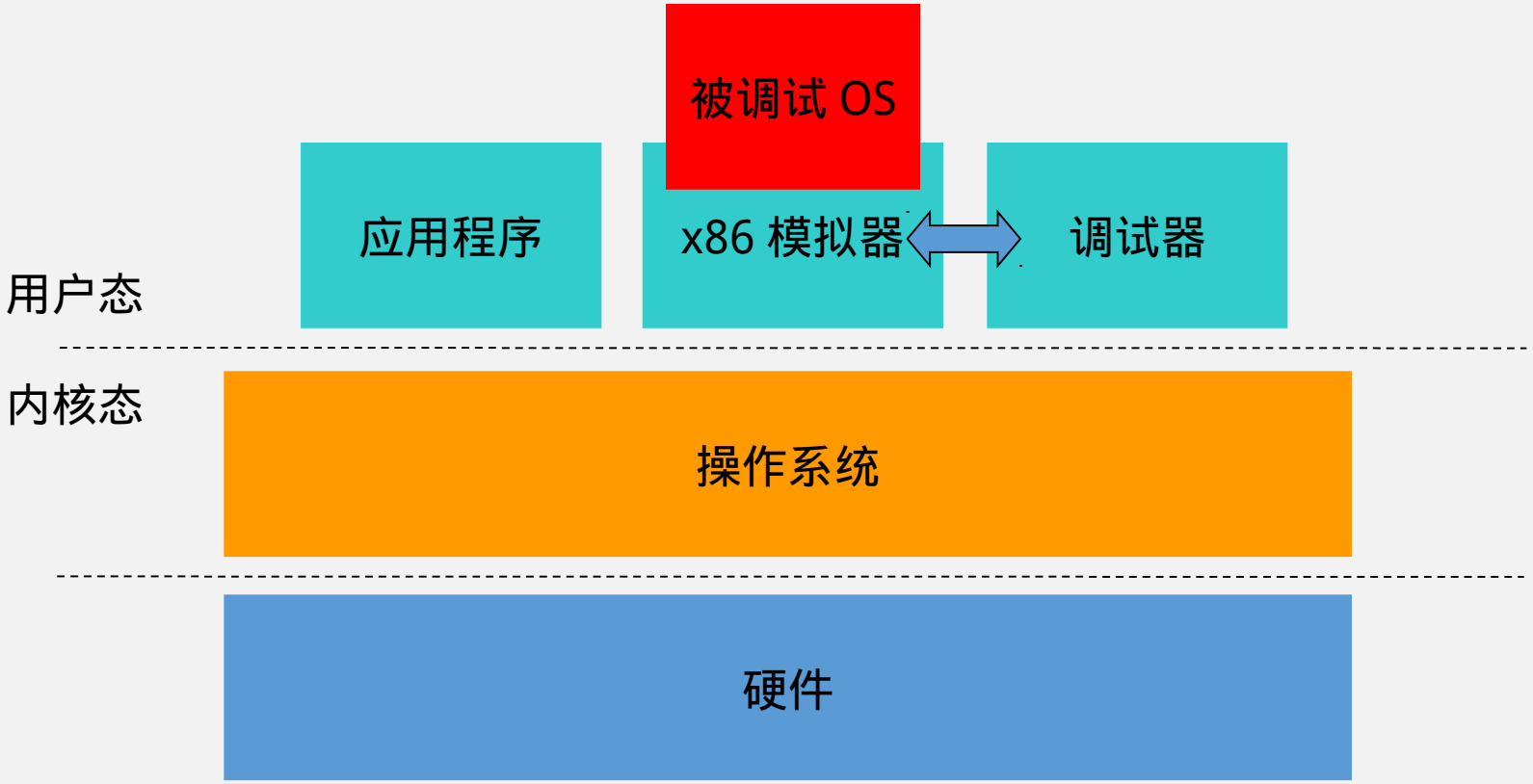
?



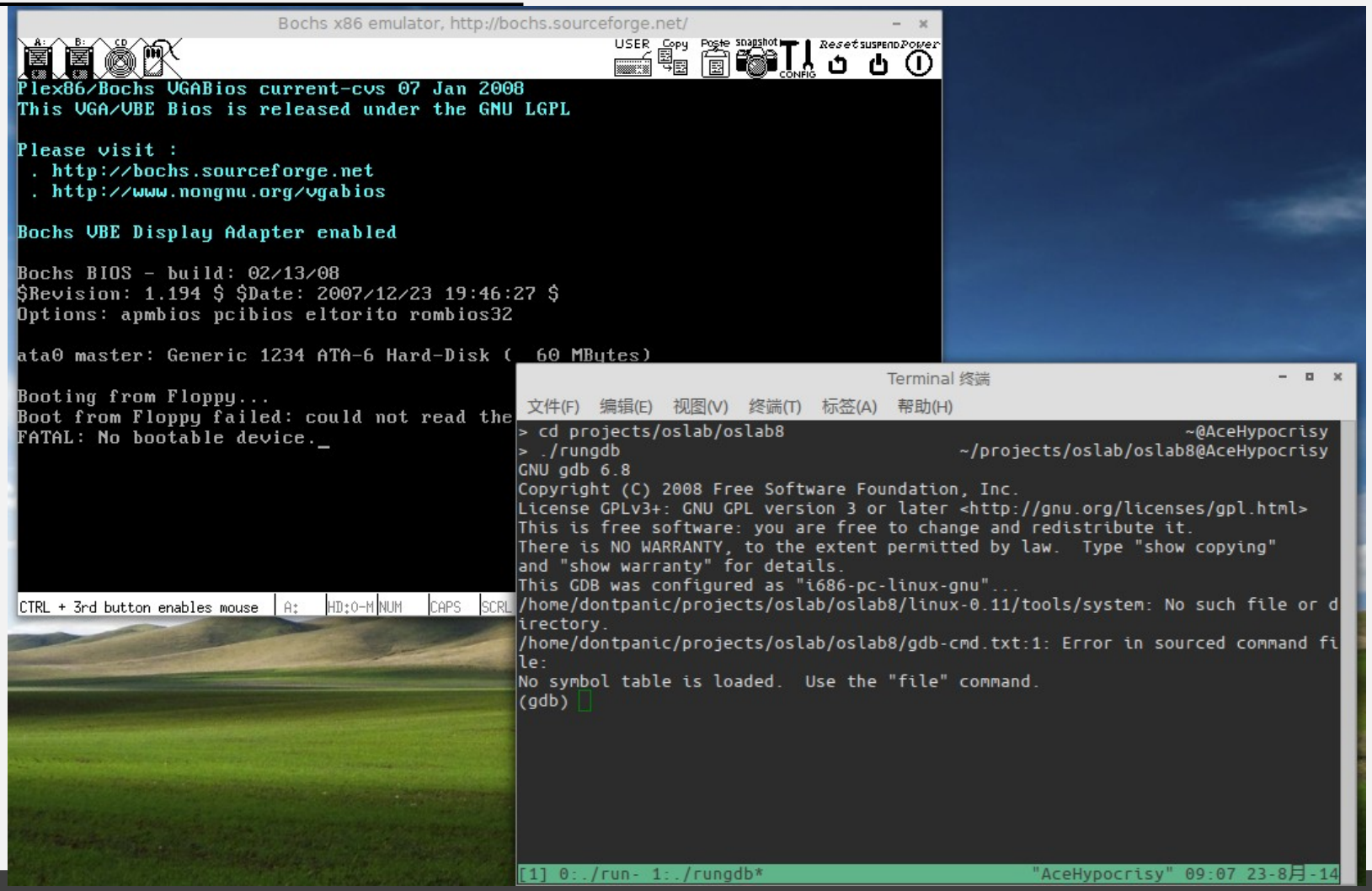
如何调试操作系统 / 内核态程序？

- 用户态
  - x86 模拟器、 User Mode Linux
- 内核态
  - Hook IDT
- x86 虚拟化技术

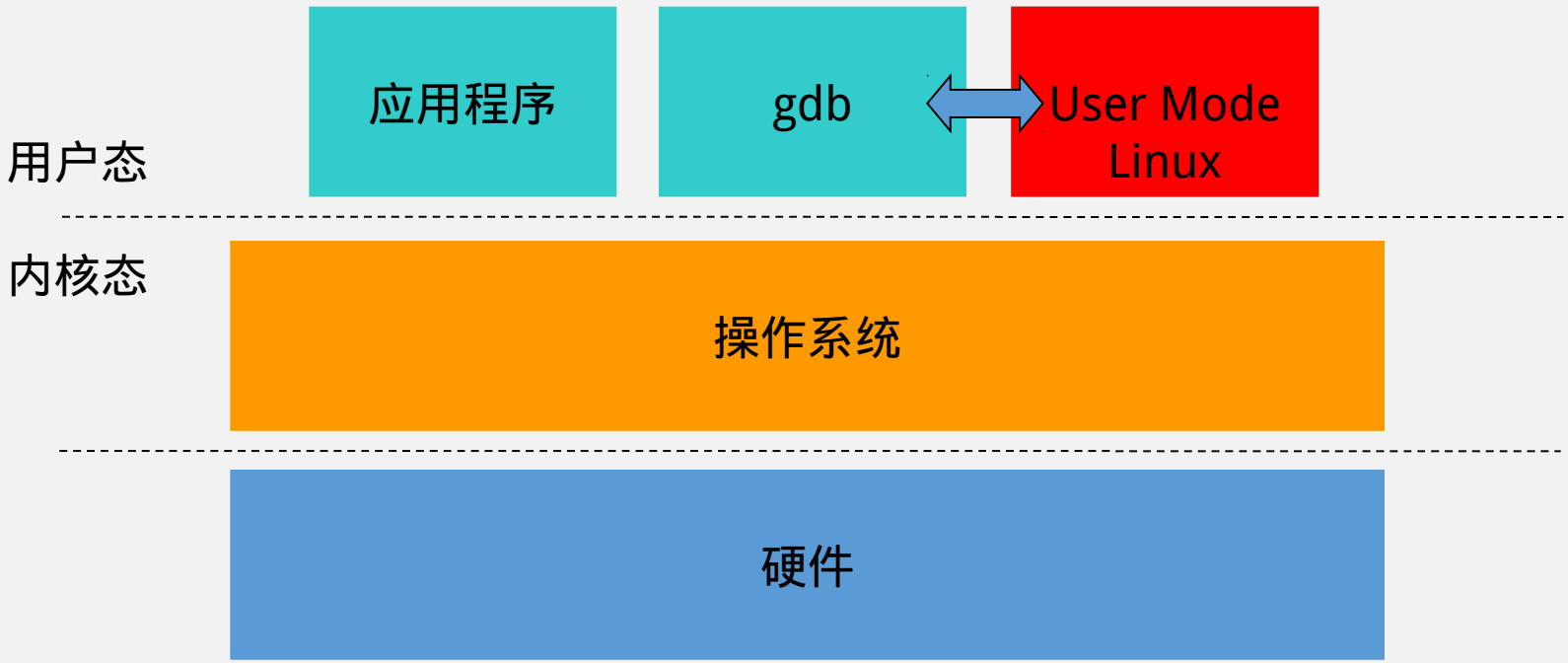
X86 模拟器



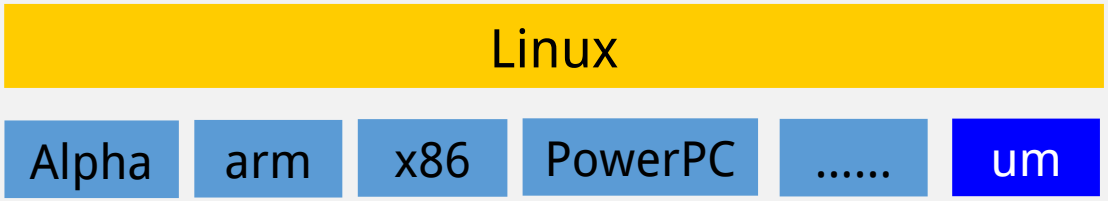
Bochs ：跨平台 X86 模拟器



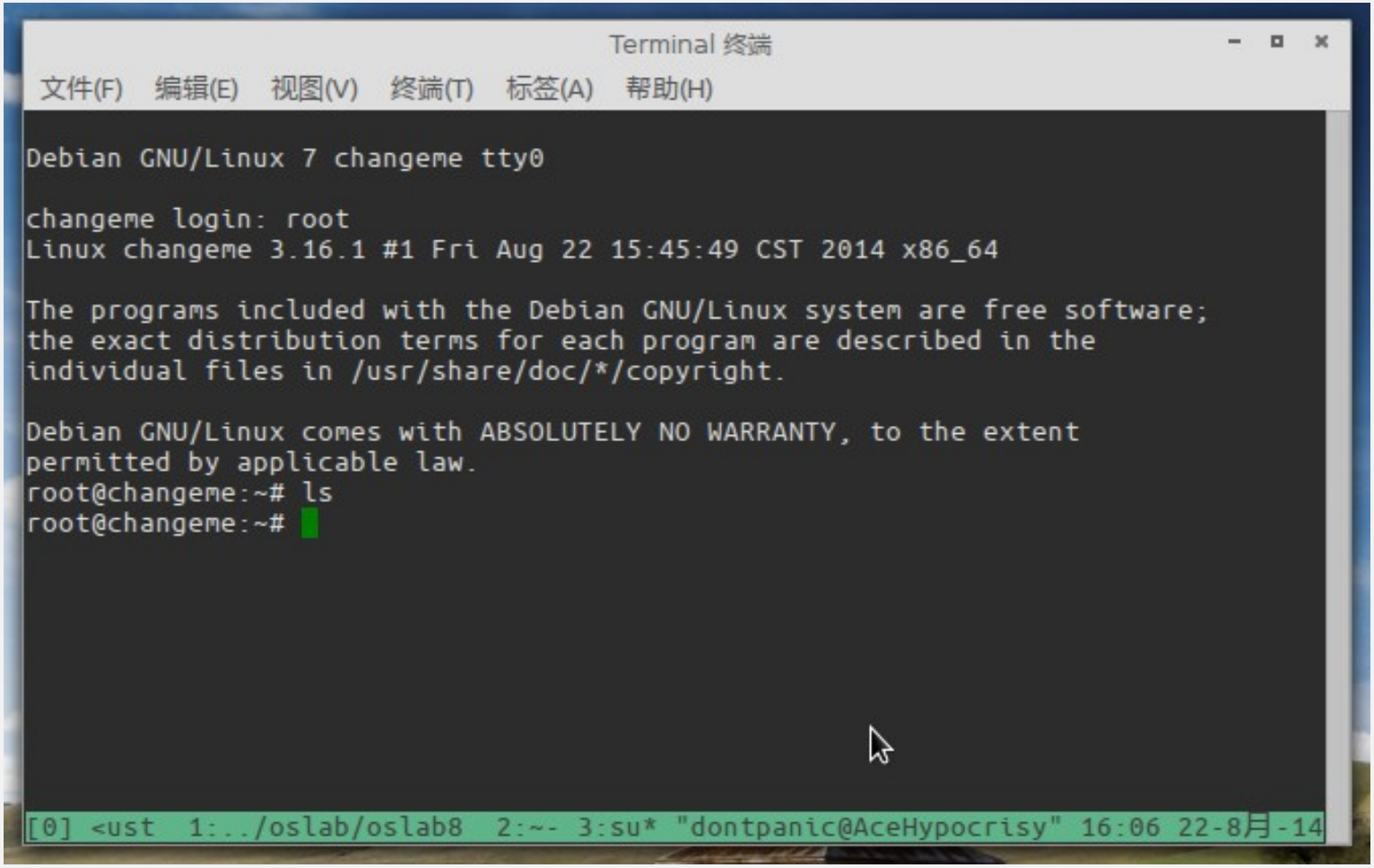
User Mode Linux

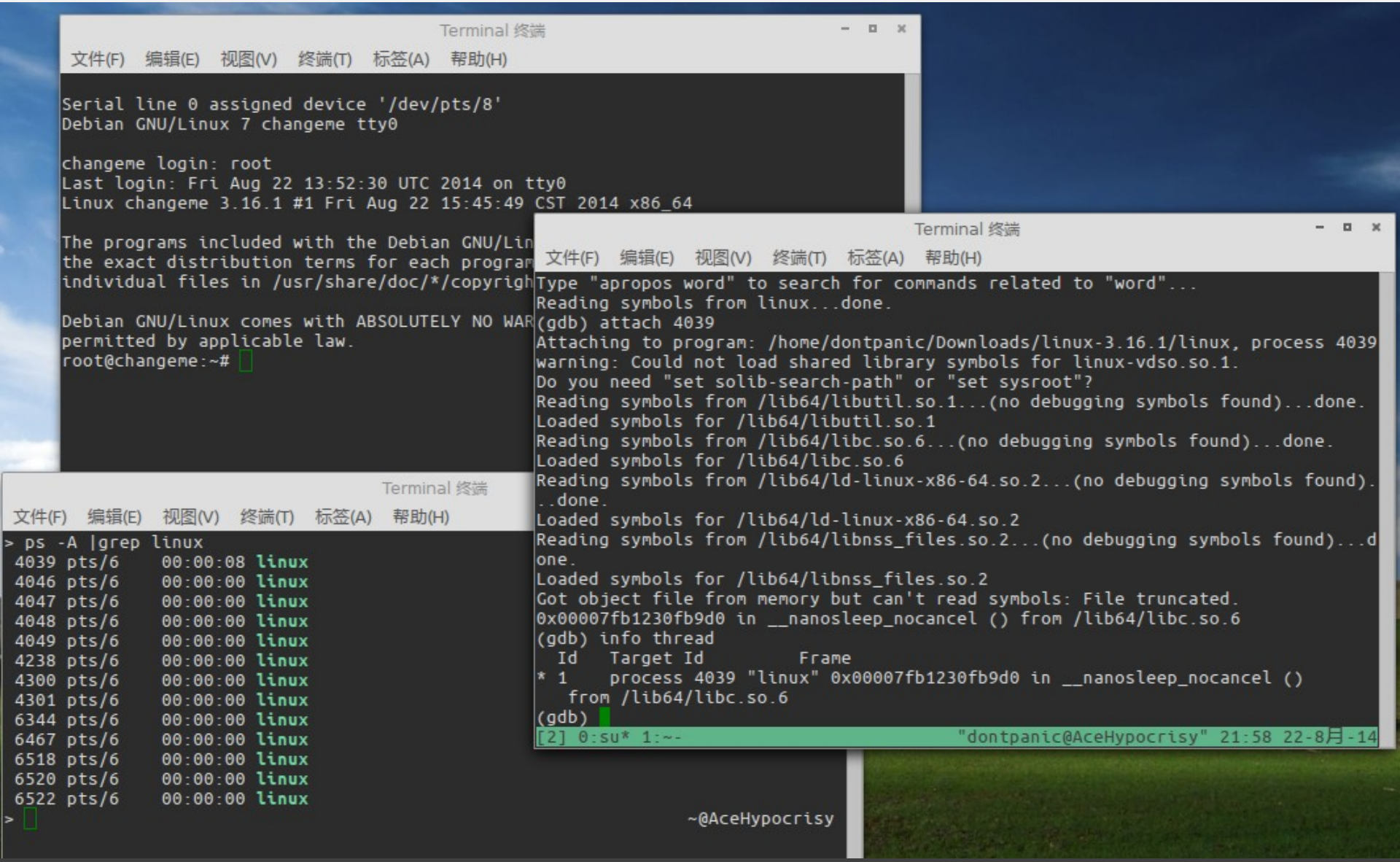


User Mode Linux

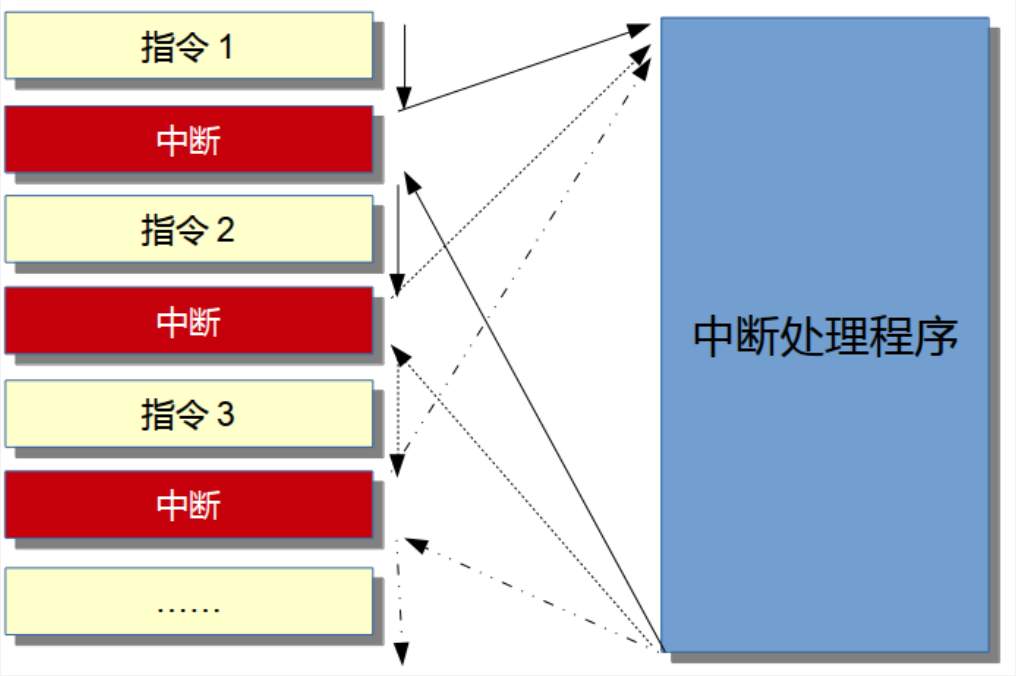


User Mode Linux



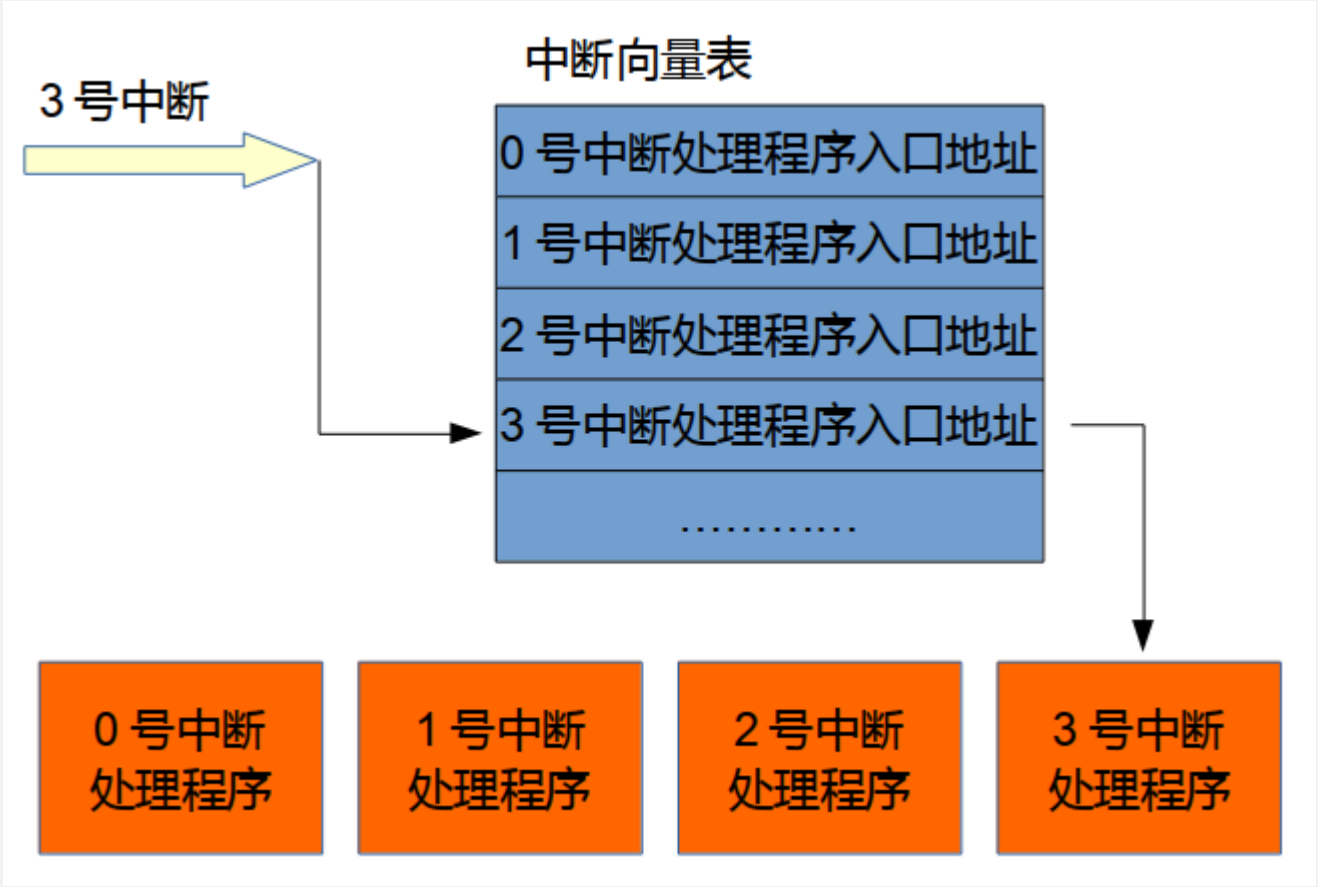

Terminal 终端
文件(F) 编辑(E) 视图(V) 终端(T) 标签(A) 帮助(H)
Serial line 0 assigned device '/dev/pts/8'
Debian GNU/Linux 7 changeme tty0
changeme login: root
Last login: Fri Aug 22 13:52:30 UTC 2014 on tty0
Linux changeme 3.16.1 #1 Fri Aug 22 15:45:49 CST 2014 x86_64
The programs included with the Debian GNU/Lin
the exact distribution terms for each progra
individual files in /usr/share/doc/*/copyrigh
Debian GNU/Linux comes with ABSOLUTELY NO WAR
permitted by applicable law.
root@changeme:~#
Terminal 终端
文件(F) 编辑(E) 视图(V) 终端(T) 标签(A) 帮助(H)
Type "apropos word" to search for commands related to "word"...
Reading symbols from linux...done.
(gdb) attach 4039
Attaching to program: /home/dontpanic/Downloads/linux-3.16.1/linux, process 4039
warning: Could not load shared library symbols for linux-vdso.so.1.
Do you need "set solib-search-path" or "set sysroot"?
Reading symbols from /lib64/libutil.so.1...(no debugging symbols found)...done.
Loaded symbols for /lib64/libutil.so.1
Reading symbols from /lib64/libc.so.6...(no debugging symbols found)...done.
Loaded symbols for /lib64/libc.so.6
Reading symbols from /lib64/ld-linux-x86-64.so.2...(no debugging symbols found)...done.
Loaded symbols for /lib64/ld-linux-x86-64.so.2
Reading symbols from /lib64/libnss_files.so.2...(no debugging symbols found)...done.
Loaded symbols for /lib64/libnss_files.so.2
Got object file from memory but can't read symbols: File truncated.
0x00007fb1230fb9d0 in __nanosleep_nocancel () from /lib64/libc.so.6
(gdb) info thread
Id Target Id Frame
* 1 process 4039 "linux" 0x00007fb1230fb9d0 in __nanosleep_nocancel ()
from /lib64/libc.so.6
(gdb)
[2] 0:su* 1:~- "dontpanic@AceHypocrisy" 21:58 22-8月-14
Terminal 终端
文件(F) 编辑(E) 视图(V) 终端(T) 标签(A) 帮助(H)
> ps -A |grep linux
4039 pts/6 00:00:08 linux
4046 pts/6 00:00:00 linux
4047 pts/6 00:00:00 linux
4048 pts/6 00:00:00 linux
4049 pts/6 00:00:00 linux
4238 pts/6 00:00:00 linux
4300 pts/6 00:00:00 linux
4301 pts/6 00:00:00 linux
6344 pts/6 00:00:00 linux
6467 pts/6 00:00:00 linux
6518 pts/6 00:00:00 linux
6520 pts/6 00:00:00 linux
6522 pts/6 00:00:00 linux
>
~@AceHypocrisy

Hook IDT

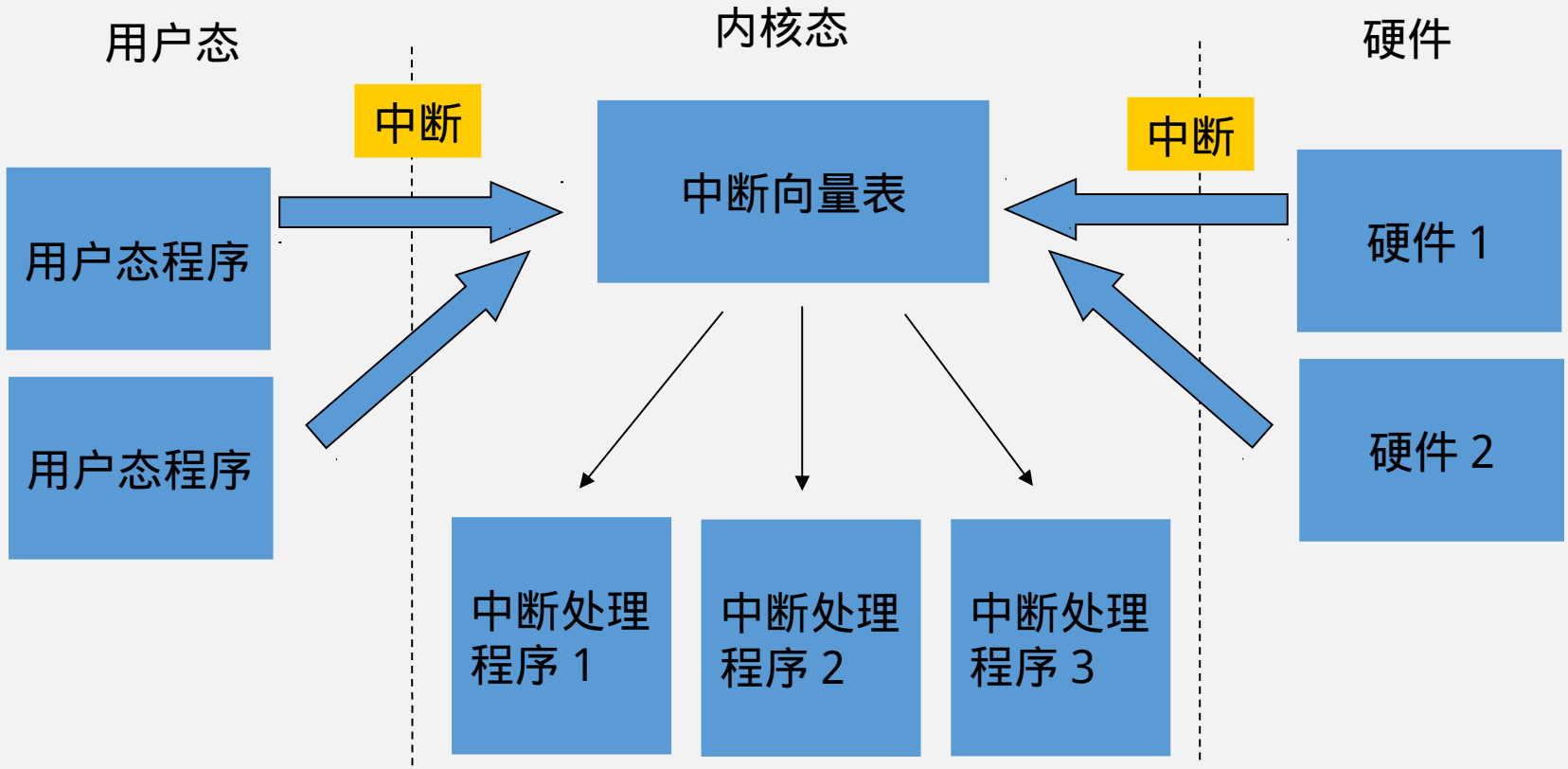


Hook IDT

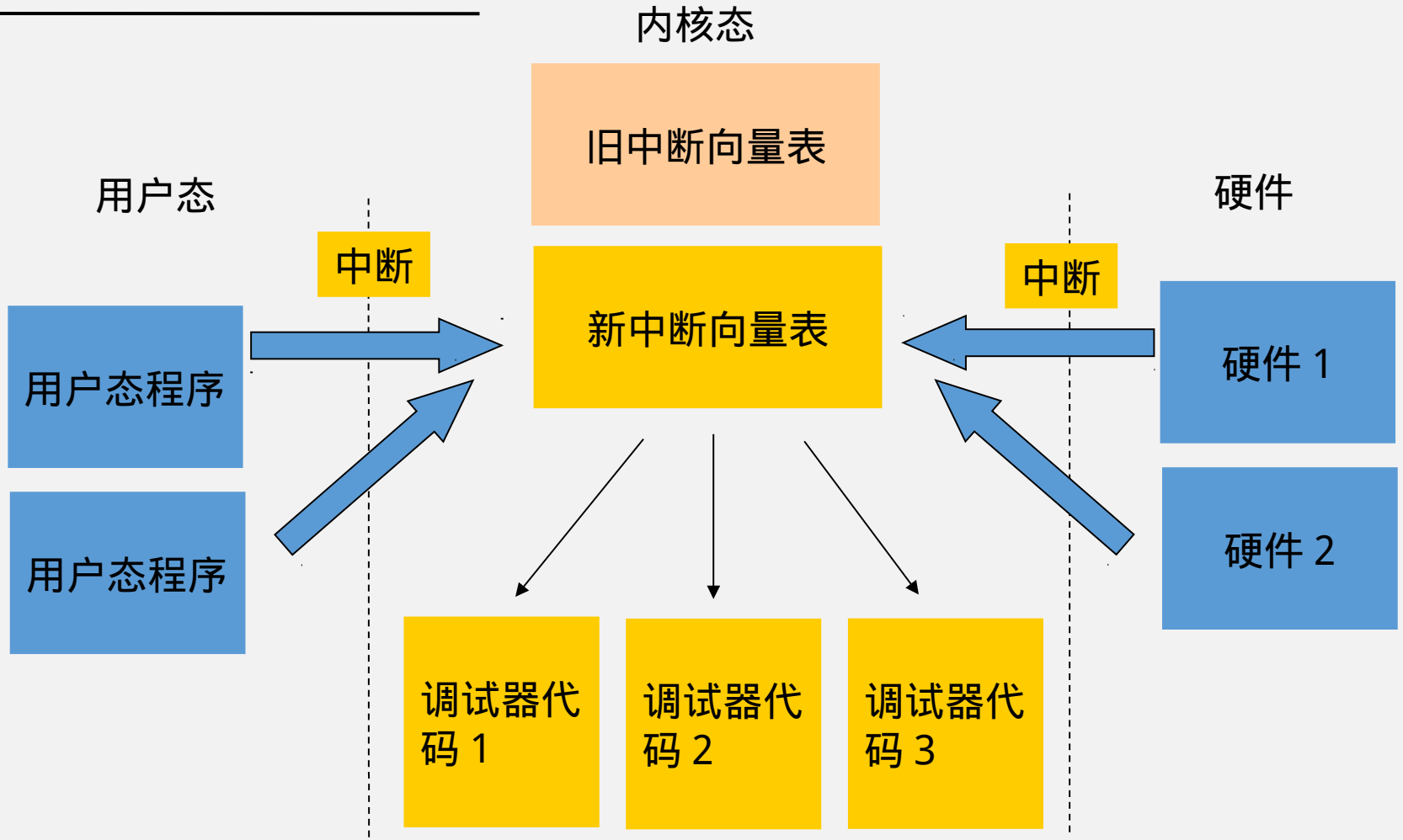
- IDT ：中断向量表



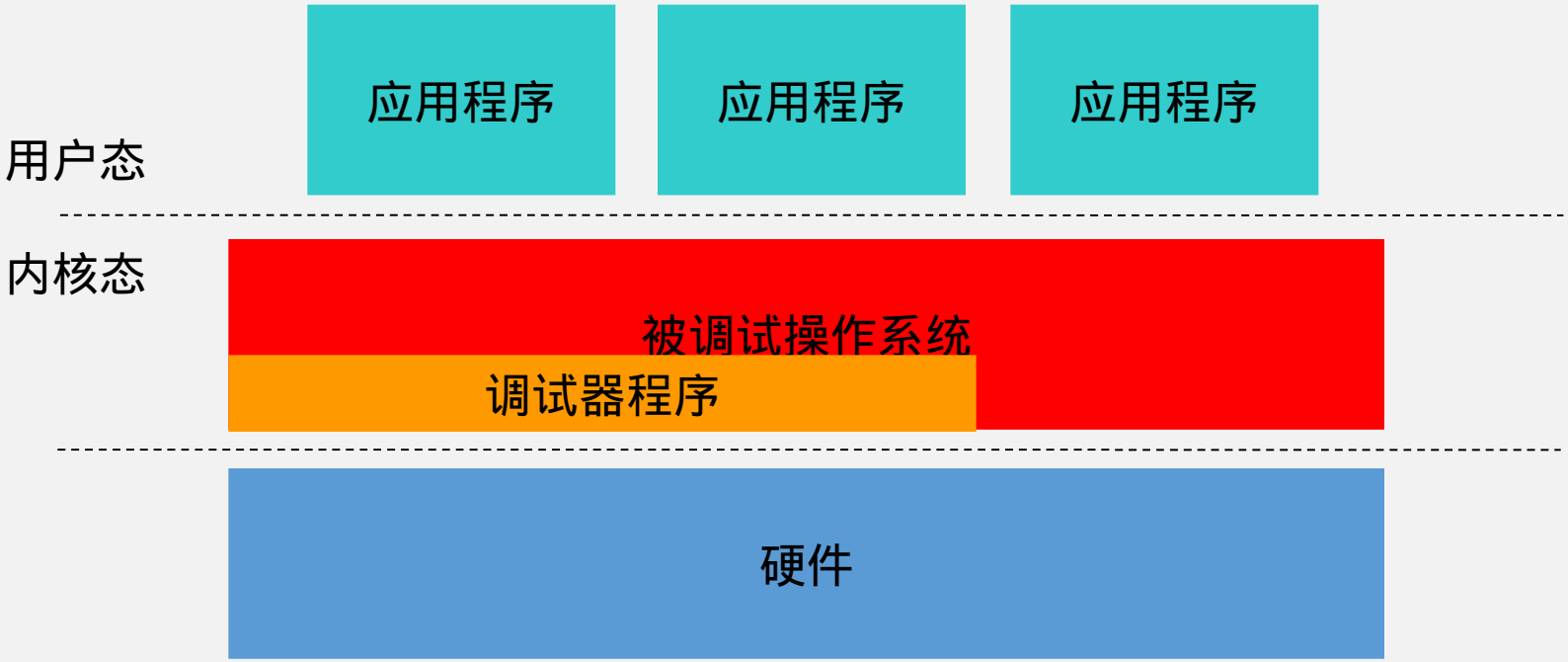
Hook IDT


用户态
内核态
硬件
中断
中断
用户态程序
用户态程序
中断向量表
硬件 1
硬件 2
中断处理程序 1
中断处理程序 2
中断处理程序 3

Hook IDT


内核态
旧中断向量表
用户态
硬件
中断
新中断向量表
中断
用户态程序
用户态程序
硬件 1
硬件 2
调试器代码 1
调试器代码 2
调试器代码 3

Hook IDT



LinICE: Linux Kernel Debugger

RedHat9 [Running] - Oracle VM VirtualBox

Machine View Devices Help

```
EAX=00005305    EBX=00000000    ECX=00000000    EDX=00000000    ESI=00000000
EDI=00000000    EBP=00000002    ESP=C034DF62    EIP=00009434    o d I S z A p C
CS=00B0   DS=00B8   SS=0068   ES=0000   FS=0000   GS=0000
    Data                                byte                          (0)
0068:00000000 01 00 00 00 53 FF 00 F0 53 FF 00 F0 53 FF 00 F0  ....S...S...S...
0068:00000010 53 FF 00 F0 53 FF 00 F0 DC E9 00 F0 53 FF 00 F0  S...S.......S...
0068:00000020 A5 FE 00 F0 87 E9 00 F0 53 FF 00 F0 46 E7 00 F0  ........S...F...
0068:00000030 46 E7 00 F0 46 E7 00 F0 57 EF 00 F0 53 FF 00 F0  F...F...W...S...
    Code
00B0:00009434  EB18            jmp     short 0000944E          (jump ahead)
00B0:00009436  EB16            jmp     short 0000944E
00B0:00009438  EB16            jmp     short 00009450
00B0:0000943A  B480            mov     ah, 80

Version: 2.6
BY USING THIS SOFTWARE YOU AGREE WITH THE TERMS OF THE LICENSE AGREEMENT.

LINICE: Init: lines 25; code on; set lowercase on; i3here on;
XWIN: Received XInitPacket.. Using buffer size of 1048576 (1024 Kb)
XWIN: User virtual address = 40114000
XWIN: Physical address     = E0000000
XWIN: Kernel address       = E18CD000 (000EA600)
XWIN: Desktop 800 x 600, stride=1600  bpp=16
:set
SET [setvariable] [ON | OFF] [value]
```

Trash

```
[dontpanic@localhost dontpanic]$ su -
Password:
[root@localhost root]# cd /home/dontpanic/linice/bin
[root@localhost bin]# ./linsym -i

Linice Debugger Symbol Translator/Loader Version 2.6
Linice and Linsym (C) 2005 by Goran Devic. All Rights Reserve

Linice installed.
[root@localhost bin]# ./xice
```

root@

Tue Feb 04
11:37 AM

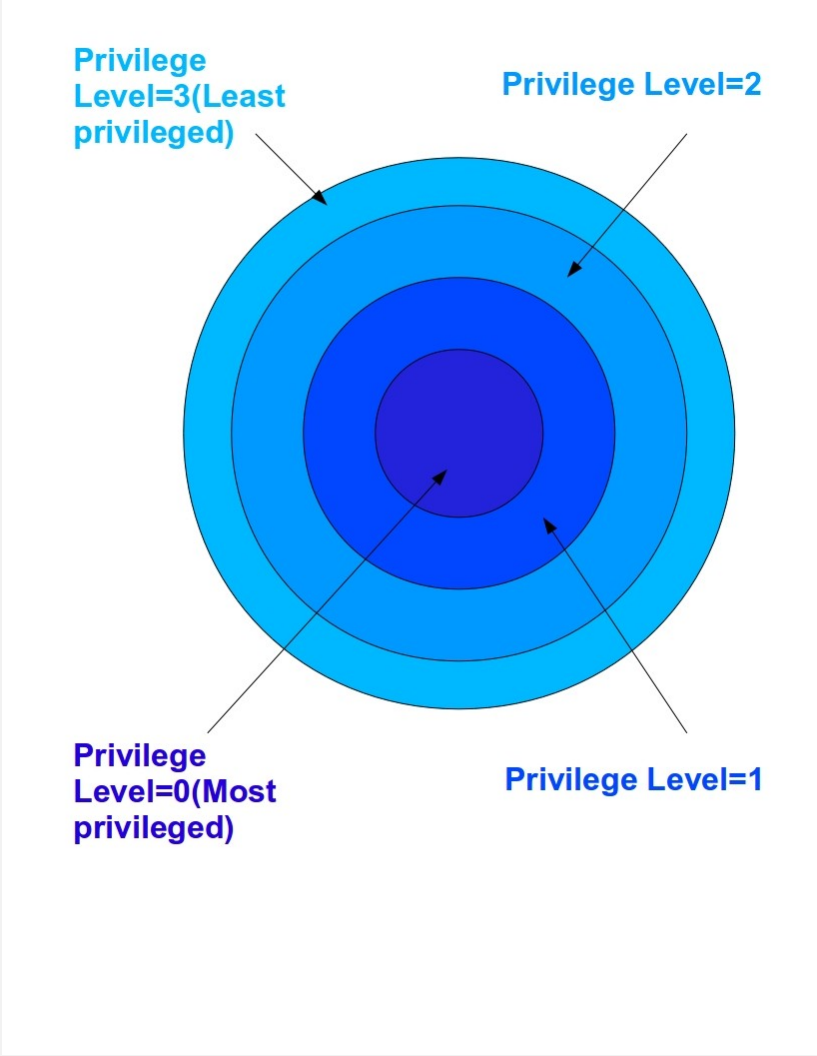
Right Ctrl

X86 虚拟化技术

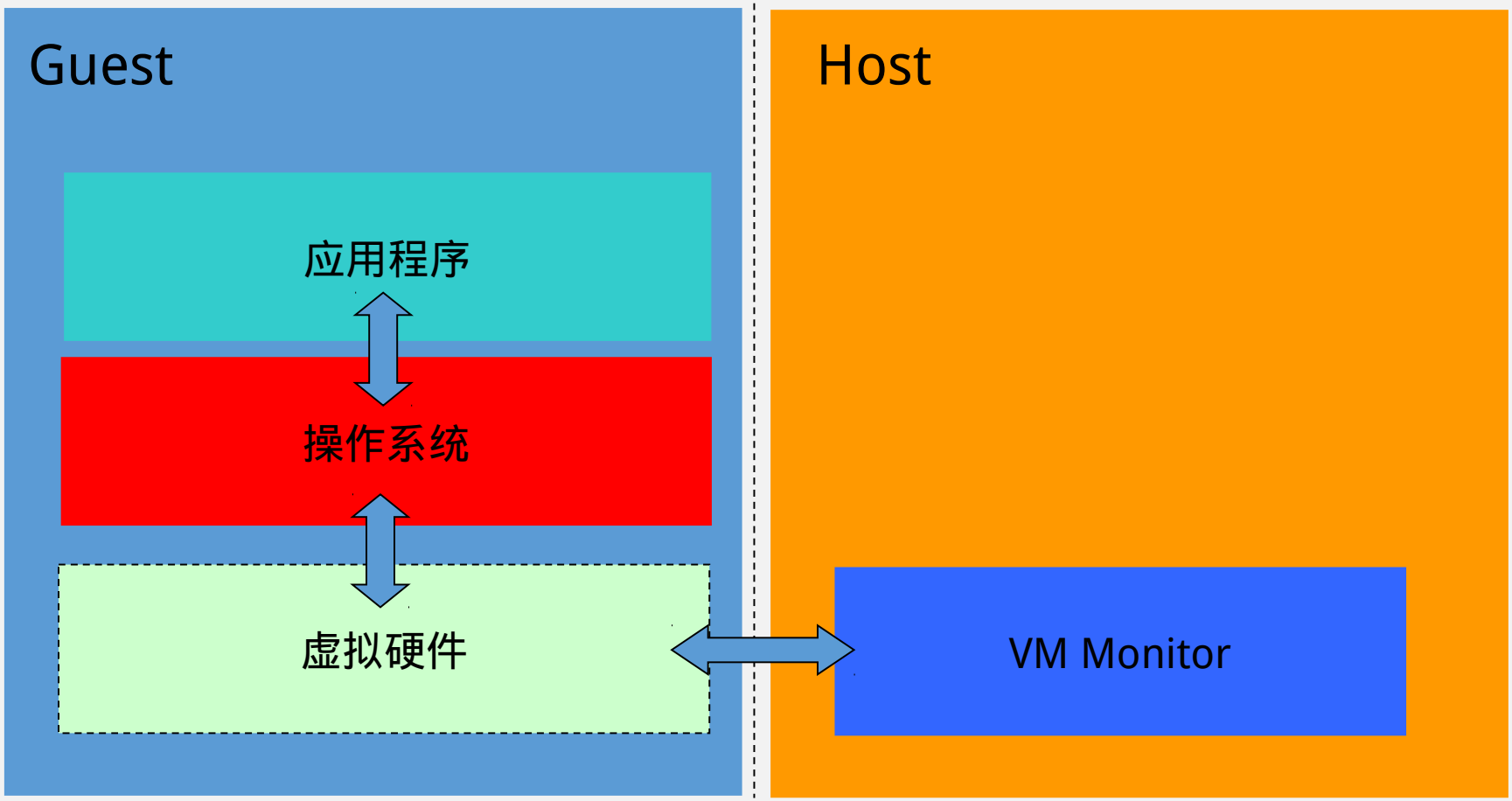
Privilege Level

Ring 0: Kernel
Ring 3: User Application

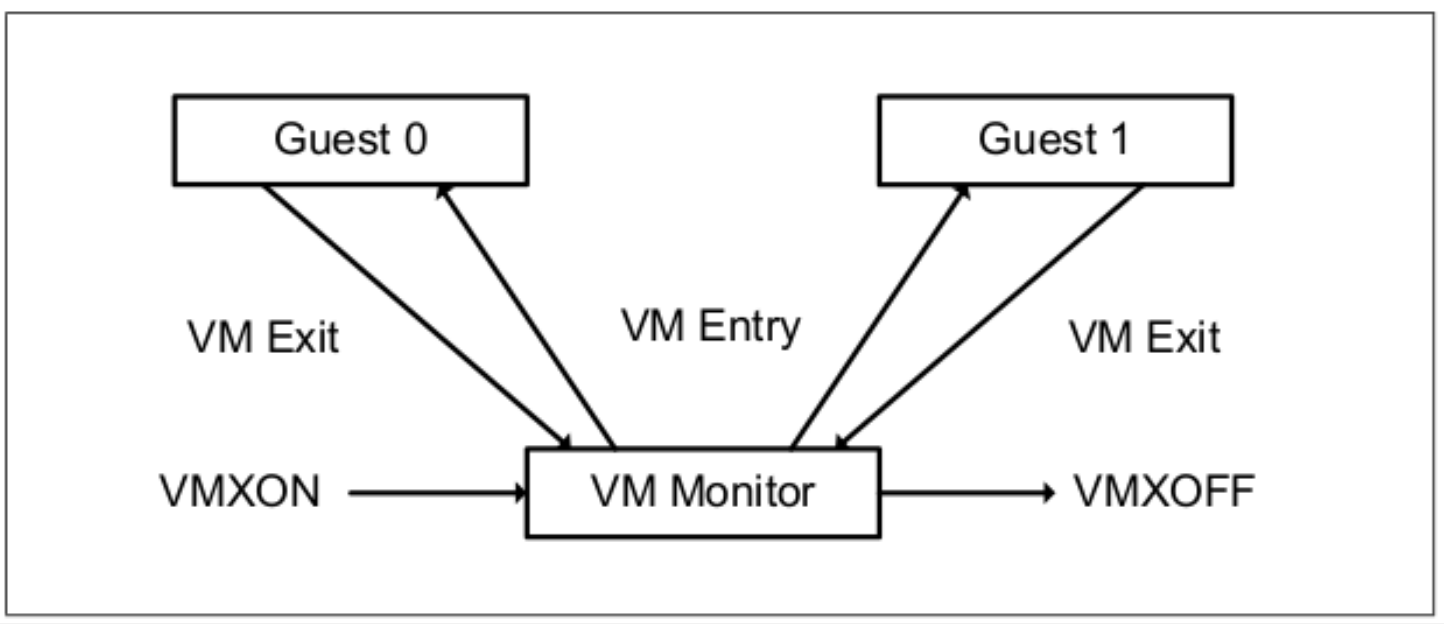
Ring 1/2: Not Used



X86 虚拟化技术



X86 虚拟化技术


Guest 0
Guest 1
VM Exit
VM Entry
VM Exit
VMXON
VM Monitor
VMXOFF

其他常见 Linux 内核态程序调试手

其他常见 Linux 内核态程序调试手段

- printk / dmesg
- Oops
- kprobe
- 内核转储

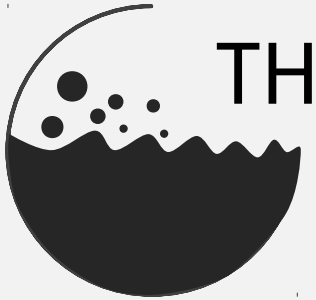

尾声

尚未涉及的细节

- ptrace 的工作机制
- 进程的挂起
- 多（核）CPU
- APIC
- Hook 内核函数
- 单机调试器的输入输出方式
- 中断处理程序的内容
- ……

参考资料

- Intel® 64 and IA-32 Architectures Software Developer's Manual, Volume 3
- AMD64 Architecture Programmer's Manual, Volume 2
- Linux Cross Reference, http://lxr.free-electrons.com
- LinICE Source, http://sourceforge.net/projects/linice/
- IBM Developer Networks, http://www.ibm.com/developerworks/cn/linux/



THANKS ！