

翱翔在Linux的天空

HumJb & HaHa

首页 | 博文目录 | 关于我



humjb_1983

博客访问：9493
博文数量：80
博客积分：0
博客等级：民兵
技术积分：685
用户组：普通用户
注册时间：2014-02-20 08:27

加关注 短消息
论坛 加好友

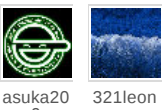
文章分类

- 全部博文 (80)
 - 硬件相关 (5)
 - 虚拟化 (13)
 - 其他 (1)
 - Linux其他方面 (3)
 - Linux内核 (57)
 - 未分配的博文 (1)

文章存档

2014年 (80)

我的朋友



asuka20 321leon

最近访客



arm-linu 码出一片 pisming
jeppeter 刘一痕 SCvsCS

KVM基本原理和架构一-概念和术语 2014-03-05 14:02:11

分类： LINUX

- 1 KVM：Kernel-based Virtual Machine，基于kernel的虚拟机
 - 1 VMM：Virtual Machine Monitor，虚拟机监控器，在KVM虚拟化环境中，即KVM内核。
 - 1 VM：Virtual Machine，虚拟机，即Guest机。
 - 1 全虚拟化：Guest OS无需做任何改动即可运行。有两种实现方式：软件辅助实现和硬件辅助实现。
 - ü 软件辅助实现的全虚拟化：一般采用优先级压缩(Ring Compression)和二进制翻译(Binary Translation)实现。由于一些特权指令必须运行于特权级Ring0，如LGDT，通常OS kernel运行于Ring0中，这样虚拟化环境中，VMM就没有Ring可用了。优先级压缩大致原理为：使VMM运行于Ring0，Guest OS kernel运行于Ring1，而Guest OS user运行于Ring3。优先级压缩能正确处理大部分的特权指令，但是由于x86架构存在虚拟化漏洞(部分敏感指令不能被捕获)，采用二进制翻译方法来处理这些虚拟化漏洞：VMM扫描并修改Guest的二进制代码，将无法捕获的敏感指令转换为支持虚拟化的指令。
 - ü 硬件辅助实现的全虚拟化：依赖硬件虚拟化特性(比如Intel VT-x)，基本思想是：在硬件架构上加入了足够的虚拟化功能，使其可以截获Guest OS中所有敏感指令，并进行模拟。
 - 1 半虚拟化：Guest OS需要进行修改，对不能“陷入”的敏感指令进行替换，以解决敏感指令的捕获问题。
 - 1 VT-x：Intel为CPU虚拟化提供了VT-x(Intel Virtualization for x86)技术(AMD SVM也提供了类似的功能)，VT-x引入了一套新的工作模式，即VMX(Virtual Machine eXtension)，该模式下的CPU具有两种操作模式：根模式(VMX root operation)和非根模式(VMX non-root operation)。每种操作模式下，都有独立的Ring0-Ring3。
 - ü VMM运行于根模式，对KVM虚拟化环境来说，即KVM主机运行于根模式，主机中的kernel运行于根模式中的Ring0，主机中用户态程序(比如Qemu-kvm)运行于根模式中的Ring3。
 - ü Guest运行于非根模式，Guest中的kernel运行于非根模式中的Ring0，Guest中的用户态程序运行于非根模式中的Ring3。
- 此外，VT-x还引入了一组新的指令，包括VMLAUNCH/VMRESUME(用于发起VM-Entry)、VMREAD/VMWRITE(用于配置VMCS)等。
- 1 VM-Exit/VM-Entry：在VT-x环境中，非根模式下(即Guest中)，执行敏感指令触发的“陷入”称为VM-Exit。VM-Exit发生时，CPU自动从非根模式切换到根模式下，然后就可以由VMM对触发VM-Exit的敏感指令做进一步处理或模拟。
- 与VM-Exit对应的是VM-Entry，此操作由VMM发起，通常是VMM调度某个Guest(或VCPU)运行时引起，此时CPU自动由根模式切换为非根模式。
- 1 VMCS：在VT-x环境中，引入了VMCS(Virtual Machine Control Structure，虚拟机控制结构)，以便更好的支持CPU虚拟化。VMCS用于保存虚拟CPU需要的相关状态，比如：CPU在根模式和非根模式下的寄存器的值、VM-Exit的原因等，当CPU发生VM-Exit和VM-Entry时，会自动查询和更新VMCS，VMM也可以通过配置VMCS控制CPU的行为。

阅读(111) | 评论(0) | 转发(0) |



风铃之音



embedde



chrxxy

订阅

推荐博文

- 云计算-Azure-3.负载均衡集...
- 读书与写论文的引导书——leo...
- 在framework层添加自己的jar...
- tcpdump工具浅析
- python json ajax django四星...
- Solaris文件管理和目录管理...
- Solaris退出系统,改变系统运...
- 监控Data Guard实时同步...
- Oracle的告警日志之v\$diag_al...
- 使用AWR生成报表

热词专题

- Debian设置
- 欢迎kkkkkkkybbb在ChinaUnix...
- 虚拟机ping不通win7宿主机...
- 安装oracle
- 关于STM32的SPI的问题

上一篇: TCP: time wait bucket table overflow原因
下一篇: KVM基本原理和架构二-基本原理

0

相关热门文章

caller和callee的区别	linux 常见服务端口	C语言 如何在一个整型左边补0...
什么是高端装修 高端装修的特...	【ROOTFS搭建】busybox的httpd...	python无法爬取阿里巴巴的数据...
什么是高端装修 高端装修的特...	xmanager 2.0 for linux配置	linux-2.6.28 和linux-2.6.32....
Lucene实践：全文检索的基本原...	什么是shell	linux su - username -c 命...
各年龄段的美白方法	linux socket的bug??	我不得不在这里问一下网站使用...

给主人留下些什么吧！~~

评论热议

请登录后评论。

[登录](#) [注册](#)

[关于我们](#) | [关于IT168](#) | [联系方式](#) | [广告合作](#) | [法律声明](#) | [免费注册](#)

Copyright 2001-2010 ChinaUnix.net All Rights Reserved 北京皓辰网域网络信息技术有限公司. 版权所有

感谢所有关心和支持过ChinaUnix的朋友们

京ICP证041476号 京ICP证060528号