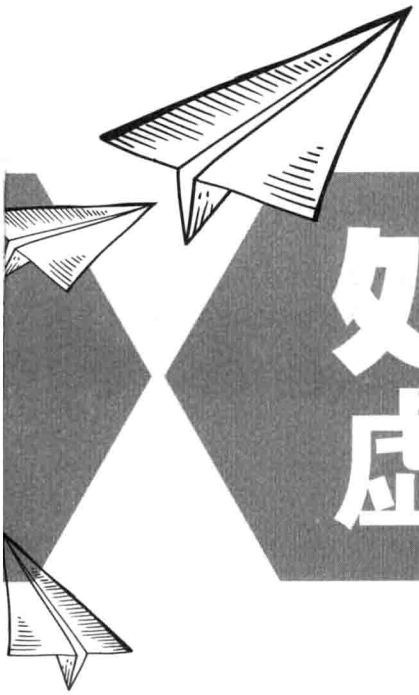
Broadview
www.broadview.com.cn

处理器虚拟化技术

邓志 著

電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn



处理器虚拟化技术

邓志 著

電子工業出版社
Publishing House of Electronics Industry
北京•BEIJING

内 容 简 介

本书针对在Intel处理器端的虚拟化技术（Intel Virtualization Technology for x86，即Intel VT-x）进行全面讲解。在Intel VT-x技术下实现了VMX（Virtual-Machine Extensions，虚拟机扩展）架构平台来支持对处理器的虚拟化管理。因此，VMX架构是Intel VT-x技术的核心。本书内容围绕VMX架构实现细节展开全面讲解。但Intel VT-d（Virtualization Technology for Directed I/O）和Intel VT-c（Virtualization Technology for Connectivity）技术并不在本书的描述范围。同时，也不针对AMD-v技术进行讨论。

全书共分为7章，书的整体结构也较为规整，可读性比较强。本书共提供14个例子，对VMX架构的一些特色功能进行辅助讲解。

读者阅读本书，可以学习Intel VT-x技术的VMX架构知识，并且对整个x86/x64体系有更深入的了解！可以说，不了解VMX架构，根本算不上对x86/x64体系熟悉，因为，在处理器的虚拟化技术里需要使用全方位的体系知识，对处理器在非常细节的地方进行虚拟化处理。

因此，本书适合有一定x86/x64体系知识基础或者想更深入学习x86/x64体系知识的读者。

图书在版编目（CIP）数据

处理器虚拟化技术 / 邓志著. —北京：电子工业出版社，2014.6
ISBN 978-7-121-23019-6

Ⅰ. ①处… Ⅱ. ①邓… Ⅲ. ①微处理器—虚拟技术Ⅳ. ①TP332

中国版本图书馆 CIP 数据核字(2014)第 080434 号

策划编辑：李　冰
责任编辑：李　冰　高洪霞
印　　刷：北京京科印刷有限公司
装　　订：三河市皇庄路通装订厂
出版发行：电子工业出版社
　　北京市海淀区万寿路 173 信箱　邮编 100036
开　　本：787×1092　1/16　印张：41.75　字数：1076 千字
印　　次：2014 年 6 月第 1 次印刷
印　　数：2000 册　定价：109.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线：（010）88258888。

前 言

虚拟化技术大概可以分为软件虚拟化和硬件虚拟化两大类。软件上的虚拟化实际上很多地方都可以看到，例如常见的系统虚拟机软件，Java 虚拟机和 Android 上的虚拟机可以算是广义上的软件虚拟机了。

在硬件虚拟化技术出现之前，虚拟机只能靠软件来模拟实现。硬件上的虚拟化技术可以算是比较热的一门技术了，但它并不是什么新鲜的技术。据资料显示，实际上 Intel 第一代 VT-x 技术 Vanderpool 在 2005 年就已经推出了，但 Intel 发展得很快，从处理器端到芯片组端，再到网络端都已经可以布署硬件虚拟化技术了。

《处理器虚拟化技术》一书围绕 Intel VT-x 处理器端的虚拟化技术而写，可以算是前作《x86/x64 体系探索及编程》的下册。尽管此书只围绕一个主题，但笔者认为它的知识面并不比《x86/x64 体系探索及编程》一书窄，甚至还要超出。因为，完整地虚拟化一个“虚拟处理器”，必须结合原来的 x86/x64 体系知识并扩展开来。编写本书所花的时间比前作还要多许多，从开始动笔到完稿用了 11 个月的时间，期间不可谓不艰辛。

本书内容

全书共 7 章。

第 1 章 系统平台

本书的代码使用汇编语言编写，运行在裸机平台上。本章主要介绍如何实现一个简易平台来支撑书中所有例子的运行。这个平台可以被编译生成 32 位或者 64 位环境。也实现了一些特色机制，例如，开启了多处理器环境，并且为了更好地、直观地了解代码的执行流程而加入了调试信息机制。

第 2 章 VMX 架构基础

VMX 架构是 Intel 处理器端虚拟化技术（Intel VT-x）的实现框架。本章对虚拟化技术进行概括讲解，全面了解 VMX 架构的基础知识。首先介绍了 VMX 架构下引入 VMX root operation 与 VMX non-root operation 模式，以及这两个模式之间的切换。然后介绍了

VMX 架构下的各种处理器能力，以及这些能力的检测。最后全面讲解新引进的 VMX 指令集，以及 VMX 指令产生的各种失败的情形。

第 3 章 VMCS 结构

VMCS 结构是 VMX 架构下最基本、最重要的数据结构。VMCS 结构内分为 6 个区域，每个区域有若干个字段。整个 VMX operation 模式的运行环境由 VMCS 结构内的这些字段来配置与管理，例如配置 guest 与 host 的运行环境。全面地讲解了这 6 个区域内每个字段的设置与使用。

第 4 章 VM-entry 处理

在 VMX 架构下，guest 运行在 VMX non-root operation 模式下，而 host 运行在 VMX root operation 模式下。如果需要运行 guest 软件，则处理器需要从 VMX root operation 切换到 VMX non-root operation 模式运行，这个行为被称为“VM-entry”。

本章详细地讲解处理器在执行 VM-entry 操作时的每个流程。例如，在 VM-entry 时会检查 VMCS 结构内字段配置是否合理，并且加载 guest 的运行环境。处理器也会根据 VMCS 内字段来控制 guest 软件的运行。

第 5 章 VM-exit 处理

在 guest 运行过程中，会因为某些事件而被迫返回 host 环境。处理器由 VMX non-root operation 模式切换回 VMX root operation 模式，这个行为被称为“VM-exit”。

本章详细地讲解处理器在执行 VM-exit 操作时的每个流程。例如，在 VM-exit 处理器的相应状态信息会得到更新，guest 的环境信息会保存在 VMCS 结构内的 guest-state 区域，并且从 host-state 区域加载 host 环境信息。最后转入 host 入口执行 VMM 的管理代码。

第 6 章 内存虚拟化

VMX 架构引进了 EPT 机制来实现对物理内存的虚拟化管理，使得每个 VM 在物理平台上拥有自己独立的物理内存区域，而不受 VMM 及其他 VM 的干扰。同样也避免了 VM 对 VMM 的干扰。

本章全面讲解了 EPT 机制的实现细节，也着重介绍了 VMX 架构下的 cache 管理。在随书例子中，我们实现了 EPT 机制来管理 VM 内存。

第 7 章 中断虚拟化

VMX 架构下也实现了对 local APIC 的虚拟化管理，引进了两个重要的页面：APIC-access page 页面与 virtual-APIC page 页面。因此而形成了 virtual-APIC 的概念。它是物理平台上的 local APIC 的 shadow。

在这一章里，我们将看到处理器如何处理 APIC-access page 与 virtual-APIC page 之间的关系，将对 APIC-access page 的访问转化为对 virtual-APIC page 的访问。从而实现访问和管理 virtual-APIC 组件。

本章也着重介绍了 VMM 应该如何处理 guest 产生的异常与任务切换。最后以实际例子来展现 VMM 如何监控 INT 指令及 NMI 与外部中断。

如何阅读本书

本书将理论结合实际地进行讲解，但每章内容还是会有侧重点。以章节的篇幅来说，第 2 章至第 5 章偏重于理论知识。而第 1 章、第 6 章和第 7 章偏于实际例子。

因此，如果读者并不想通篇阅读，可以选择偏重理论知识的章节来看。第 6 章与第 7 章介绍的重点知识主要是 EPT（扩展页表）与 local APIC 虚拟化。这两个内容也是必须掌握的。

在第 1 章中并不涉及 VMX 架构知识，而是通过代码的讲解来介绍基础平台。书中的全部例子将构建于这个基础平台之上。如果读者对此章不感兴趣，可以直接跳过。但是，如果读者对一些 OS 的基础元素感兴趣，大可详细加以阅读。

读者对象

读者最好有一些 x86/x64 体系知识基础，如果感到这方面的基础知识相对薄弱，可以阅读作者的另一本著作《x86/x64 体系探索及编程》或者 Intel 官方手册。

如果读者只是想了解 VMX 的理论知识，那么可以选择相应章节，对汇编知识的要求不高。如果读者想深入全面地了解整个 VMX 架构知识，希望能够仔细阅读所有章节（第 1 章可以略过），那么需要对汇编有一定了解。

随书例子

书中的每个章节有若干例子，全书共 14 个例子。读者请登录网址 http://www.mouseos.com/books/vt/index.html，选择相应的栏目下载源码包。

源码包解压后有 7 个 chapXX 目录，对应于每一章。还有 commom、inc 以及 lib 目录，存放所有例子共用的代码。每个 chapXX 目录下有多个 ex 实例目录，对应于每个例子。以第 1 章为例，chap01 目录下有 ex1-1 和 ex1-2 实例目录。

每个实例目录里有下面的文件。

- Bs：bochs 的配置文件。
- Build：一个 DOS 批处理的编译工具，用来编译源码。
- c.img：硬盘与 U 盘映像文件。
- demo.img：软盘映像文件。
- ex.asm：例子实体代码的汇编源文件。
- ex.inc：例子实体代码头文件。

在时间宽裕的情况下，我会在 Windows 平台上用 C 语言来实现代码示例。目前已经写了大部分代码，但为了保证代码的正确性，力求在所有代码完成后再放出来。

勘误反馈

由于作者学识有限，尽管在书写时已经力求认真细致，但难免有错漏之处。读者如果发现有不妥之处，敬请不吝指出，邮件请发至：mik@mouseos.com，或登录 www.mouseos.com 网站留言。

参考资料

- 《Intel® 64 and IA-32 Architectures Software Developer's Manual》volumes 1, 2A, 2B, 2C, 3A, 3B and 3C
- 《Intel Virutalization Technology for Directed I/O Architecture Specification》

致谢

感谢电子工业出版社博文视点公司的全体工作人员，特别感谢本书的策划编辑李冰，正是有你们的辛勤劳动，才使得本书可以顺利出版。

邓志

2014 年 5 月

目 录