

# 深度实践 KVM

核心技术、管理运维、性能优化与项目实施

肖力 汪爱伟 杨俊俊 赵德禄 著

---

## Inside Explore the KVM

Core Technology, Operation and Maintenance,  
Performance Optimization, Project Implementation

---

- 中国运维领域里划时代著作，国内顶尖KVM专家和运维专家多年大规模生产环境实践经验总结，国内数十位运维专家鼎力支持和推荐
- 从核心技术、管理运维、性能优化、项目实施等多个维度系统、深入讲解KVM虚拟化技术的工作原理、应用方法和技术生态，并针对生产环境中的各种难题给出了解决方案



机械工业出版社  
China Machine Press

需要整本电子书，联系我QQ：3411317522

需要整本电子书，联系我QQ：3411317522

# 深度实践 KVM

核心技术、管理运维、性能优化与项目实施

---

Inside Explore the KVM  
Core Technology, Operation and Maintenance,  
Performance Optimization, Project Implementation

---

肖力 汪爱伟 杨俊俊 赵德禄 著



机械工业出版社  
China Machine Press

需要整本电子书，联系我QQ：3411317522

# 需要整本电子书，联系我QQ：3411317522

图书在版编目（CIP）数据

深度实践 KVM：核心技术、管理运维、性能优化与项目实施 / 肖力等著. —北京：机械工业出版社，2015.9  
(Linux/Unix 技术丛书)

ISBN 978-7-111-51193-9

I. 深… II. 肖… III. 虚拟处理机 IV. TP338

中国版本图书馆 CIP 数据核字（2015）第 188798 号

## 深度实践 KVM

核心技术、管理运维、性能优化与项目实施

---

出版发行：机械工业出版社（北京市西城区百万庄大街 22 号 邮政编码：100037）

责任编辑：高婧雅

责任校对：董纪丽

印 刷：北京市荣盛彩色印刷有限公司

版 次：2015 年 9 月第 1 版第 1 次印刷

开 本：186mm×240mm 1/16

印 张：35

书 号：ISBN 978-7-111-51193-9

定 价：95.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

客服热线：（010）88379426 88361066

投稿热线：（010）88379604

购书热线：（010）68326294 88379649 68995259

读者信箱：hzit@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问：北京大成律师事务所 韩光 / 邹晓东

# 需要整本电子书，联系我QQ：3411317522

# 需要整本电子书，联系我QQ：3411317522

## Praise 本书赞誉

这本书干货太多，都是作者亲身经历和经验之分享，运维人说实话，做实事，拒绝浮夸，在这本书体现得很好，希望能引起你的共鸣。

——赫宝生，金山西山居运维总监

云计算是一个令人兴奋的领域，作者在书中系统、详尽且量化地总结了一系列的方法与准则，这些方法与准则都在盛大游戏中进行过大规模的应用实践，不仅实用且大大开拓了读者视野。

更难能可贵的是，作者肖力是一位很注重实践和乐意分享的人。他发自内心、长期热爱这份事业，这一点在盛大游戏任职期间表现得淋漓尽致，盛大游戏G云的发展也得益于肖力及其团队对技术的深刻理解与不断探索。我相信，正是这份质朴的热爱和执著的追求成就了这本极具含金量的专业书籍。感谢他为盛大游戏所做的贡献，同时，诚挚分享此书给所有相关从业人员，希望更多的朋友能精读此书，从中得到更多启发和借鉴，一起成为云计算的弄潮儿。

——陈桂新，盛大游戏G云负责人、技术保障高级总监

很高兴看到越来越多的朋友加入云计算的大军，但对技术人员来说，从理论、实现到支撑业务会遇到不少的挑战。肖力、俊俊等把G云在技术方面走过的路、踩过的坑总结并分享出来，对有志于加入这个领域的同学有很大的实用参考价值。

感谢本书作者肖力、俊俊、爱伟和德禄对盛大游戏G云做出的贡献，G云正是在你们的努力下一步步成长起来，并成为用户首选的、高性价比的定制云的。

——冯祯旺，盛大游戏G云COO、技术保障副总监

在这个风起云涌的年代，虚拟化是各家云公司的核心和门槛。可以说没有虚拟化就谈不

# 需要整本电子书，联系我QQ：3411317522



上 IaaS 的云服务。该书很好地诠释了时下最流行的 Hypervisor——KVM 的前生与今世，并且由浅入深，从实战的角度展现给读者怎么搭建、使用、运维大规模的 KVM 集群，非常值得读者深入研读。

——宋伟，金山云合伙人、技术 VP、政企事业部总经理

KVM 技术虽已出现多年，并在云计算浪潮中发挥重要作用，但相关技术深度实践的书籍并不多见。本书作者为国内 KVM 技术早期研究者及实践者之一，本书内容由浅入深、图文并茂，适合各种级别的 KVM 技术人员查阅学习。

——萧田国，触控科技运维总监、InfoQ 金牌专栏《高效运维最佳实践》作者

不管你是否承认云时代已经来临，VMware、Xen、KVM、Docker 等虚拟机技术层出不穷，KVM 作为内核级的虚拟化技术，一直受到很多工程师的青睐，不管是性能还是安全性都算是其中的佼佼者。

本书作者有多年 KVM 运维实践经验，把自己对 KVM 的理解及工作中遇到的问题，通过这本书深入浅出地告诉大家，特别是还介绍了 Ceph 和 OpenStack 等案例，希望大家看完这本书后可以对 KVM 有一个新的认识，对虚拟化技术有一个新的认识。

——南非蜘蛛，开源爱好者

认识肖力已经很长时间，从盛大游戏时期参与建设盛大 G 云的前身，到如今在西山居从事游戏虚拟化工作，肖力是我在业界见过为数不多的既有实操能力，又有宏观格局，同时具备海量运维经验的云计算从业者。作为国内最早一批 KVM 实践者，本书是肖力多年实战经验和思考的积累，站在应用者的角度剖析了利用 KVM 建设私有云平台的过程和经验，这本书不仅适合 KVM 爱好者，对于所有的云计算爱好者和从业人员都大有裨益，郑重推荐。

——吴磊，阿里云多媒体云总经理，兼通用业务总监

现在是思维横行的时代，我们每天在微博和朋友圈刷着各种信息，很容易变得浮躁。认识老肖快十年了吧，从盛大到西山居，一直在踏踏实实地研究虚拟化相关领域，可以说在老肖身上能看到互联网运维圈子里出来的人所固有的执着与沉静。

其实，现在这类技术书籍很多，但这本书的宝贵之处在于是一群互联网资深匠者，以真诚与严谨的方式，对待自己孩子般地去雕琢技术细节，里面珍贵的经验或许可以帮助我们少踩很多坑。而更重要的是，你在这里收获的不仅仅是技术，还有互联网资深老兵的经验和感悟。本书是适合从事互联网行业以及服务器领域的技术工程师们仔细研读的一部作品。

——李刚，浪潮互联网行业部技术总监

# 需要整本电子书，联系我QQ：3411317522 v

随着近些年互联网的发展，虚拟化技术的日益成熟，虚拟化和云也逐步被大家接受和应用。

伴随需求的增长，公有云厂商也不断崛起和扩大。海外的亚马逊云、RackSpace、谷歌云、微软云，都在拓展中国市场；国内的阿里云、金山云、盛大云、腾讯云、UCloud 也日渐壮大，还有百花齐放的私有云，甚至近两年流行的容器技术，这些现象向我们暗示着未来的技术走向。

虚拟化作为云服务的基石，其重要性不言而喻。或许你很早就知道 VMware，然后可能你开始听说 Xen/Citrix、Hyper-V。没错，然后 KVM 横空出世，一发不可收拾！

相信现在的运维、开发或测试人员，或多或少会用到虚拟化。无论你使用哪种虚拟化方案，KVM 作为当下主流的虚拟化方案都值得你去了解和学习。

肖力同学专注于虚拟化工作多年，有着极其丰富的实战经验。善于总结和分享的他，提笔撰写本书必定会成为 KVM 运维技术的经典之作。

本书涵盖知识面非常广，能很好地拆分讲解并实现由浅入深，可用“从入门到精通”来形容。开篇介绍了虚拟化选型和 KVM 的发展史，让读者很容易融入进来；接下来，很贴切地从“第一台虚拟机”开始动手实战；然后，分类解读各种虚拟化技术及应用场景、怎样做资源控制；而后接地气地介绍了物理机转换到虚拟机的实践以及桌面虚拟化实践。实现了从技术的基本介绍深入到“你想用，怎么做”的自然延伸。

第二篇介绍分布式文件系统，是云平台的另一个基础组件，和 KVM 密不可分。文中详细介绍了各大主流分布式存储系统与 KVM 的结合使用，分布式存储完善 KVM 的健壮性，同时保证数据的安全性。随后，高大上地介绍了主流的 KVM 管理平台，以及虚拟化项目实施的实践经验分享。本书内容丰富精彩、层次清晰，无论初学者还是“老司机”，都能从中受益。

——胡凯，猎豹移动运维总监

“基石”这个词，可以形象地描述开源虚拟化技术对互联网的价值，也代表着互联网基础设施技术的未来。但“开源”对普通的技术人员来说，意味着需要在缺乏路标的海洋中航行。

这本书难得之处在于，是由一群奋战在互联网第一线的牛人总结的经验宝典，本书的内容，让我不断回想起多年来和各个互联网公司的小伙伴们在虚拟化技术实验、部署、侦错、管理过程中不下几百次的各种研讨，历历在目。他们成功运营着成千上万的服务器，处理上亿客户的海量业务请求，他们所做的技术探索，在目前的互联网行业具有广泛的代表性和实用性。

而今天，有幸他们能够把自己的经验和体会全面地分享，全都是干货。我想，这就是互

# 需要整本电子书，联系我QQ：3411317522

联网所崇尚的开放精神吧。这本书的作者和读者，以及这一代在互联网开源技术领域勤奋工作的人们，将会构成中国互联网基础架构技术的基石。

——陆晴，英特尔互联网及云计算业务部经理

知其因，晓其理；漫步在云端，然不知处处都是虚无缥缈的浮夸；落地、上云都离不开底层技术的支持，一线大量实战经验的总结是对垂直领域最好的诠释。这本书恰恰是这一领域的佼佼者，深入、全面并包含大量实践，不仅能帮助我们晓其理，一步步学习虚拟机技术，而且能帮助我们快速将云落地。

——刘宇，西山居运维架构师

基于 KVM 的虚拟化方案在业界已应用多年，而且越来越多的云计算服务商使用该方案对外提供服务，这对运维人员来说是一个大的挑战，因为国内还没有一本关于 KVM 运维实践的权威的书籍，大部分的问题都是靠运维人员自己摸索，不成体系。这本书的出版对从事 KVM 虚拟化方案的运维人员来说是个福音，从内容看，作者很用心，而且对运维和从事虚拟化平台开发的人员也有很好的参考作用，相信此书的出版能推动我国虚拟化乃至云计算方面的运维水平和国际接轨。

——叶理灯，UCloud 云主机开发部经理

得知力哥要写一本关于 KVM 的书，我的第一感觉是国内做 KVM 人有福了。当年我和力哥一起搞 KVM 的时候，就深感他功底扎实，思维缜密，而且动手能力快得惊人。后来，我转向 PaaS 领域，力哥又不间断地在 KVM 上深耕数载，如今已是该领域的专家。相信无论是虚拟化领域的新手还是资深从业人员，都能从他的书中得到很多借鉴。

——刘晓光，Cloud Foundry 工程师

从力哥的博客和“KVM 虚拟化实践”的公众号就可以看出这是一位注重实践、乐于分享、很有互联网精神的人。

实践出真知，本书全面性佳、实战性强，相信对 KVM 有兴趣以及正在做 KVM 相关虚拟化、云计算相关的兄弟来说必定会备感亲切，喜爱有加，强烈推荐！

——秦洁，安居客运维负责人

在虚拟化道路上，选型 KVM 是一个深思熟虑的决定，通过小规模部署验证，最终大规模部署到生产环境中。在虚拟化的道路上，我们学习到很多，想要尽可能和大家去分享。本书将会给大家带来 KVM 在生产环境中的、资深运维角度所理解的最佳实践。

云计算改变了传统运维的思维方式，短短几分钟可以实现快速部署，扩容生产业务。想

要通过 KVM 虚拟化技术结合存储构建一个稳定、高效的生产集群，支撑业务弹性伸缩，这本书不可不读。

——杨树林，动视 SRE

说来有缘，肖兄的这本书还算是我牵的线。自知我在 KVM 方面并非专家，因此不敢多加点评。我和肖兄认识的时间并不长，但从肖兄的微信公众号分享的持续性可见他对知识积累的专注和专业，让我非常敬佩，也让我相信此书绝对能成为国内 KVM 领域的巅峰之作，对诸多同行们有所帮助。

——叶金荣，<http://imysql.com> 创始人，Oracle ACE (MySQL)

## 前 言 *Preface*

目前，国内新一代的公有云底层采用的都是 KVM，据笔者了解，国内原来一些采用 Xen 的公有云也逐步在向 KVM 迁移。许多企业的私有云项目招标中，已经明确要求使用 KVM。在全球范围内，OpenStack 就是云建设的标准，而当前 OpenStack 底层默认就是 KVM。经过几年的快速发展，KVM 技术已经非常成熟稳定，在任何 Linux 发行版中，KVM 都是标配。虚拟化已经成为一项必备的技术技能，而学习 Linux 就必须学习 KVM。

## 为什么写这本书

2009 年，当虚拟化在互联网业内鲜有使用时，我们虚拟化小组在国内已经在生产环境率先开始 KVM 虚拟化技术的探索。通过实施虚拟化，我们几年内为公司节省了上亿元的运营成本，帮助公司多个处于生命周期末的游戏项目实现扭亏为盈。

当时，KVM 还是一个小众产品，国内文档较少，我们只能通过钻研国外的英文资料甚至阅读源代码来学习。RedHat 是 KVM 的重要支持者，因此与其公司 RHEL 系统同源的 CentOS 成为了研究 KVM 的重要资料。从当时刚发布的 CentOS 5.5 版本开始，我们一路追随 CentOS 系统的升级，每有一个版本发布，我们都要测试评估新的 KVM 虚拟化特性，然后逐步在生产环境部署，这样的工作方法一直持续到最新的 CentOS 6.6 及 CentOS 7.1。这也是我们小组人员快速成长的一个阶段，我们对 KVM 技术的热爱一直保持了下来，我们在周围普及 KVM 技术的热情一直没有改变。

时至今日，KVM 经过几年的快速发展，已经成为了最炙手可热的虚拟化引擎，国内新一代的公有云全部选择了 KVM。而与 KVM 如火如荼的发展形成鲜明对比的是中文资料依然匮乏，尤其是介绍实践经验的资料。在我们周围和各种技术论坛上，经常碰到一些运维工程师，在实施一个 KVM 虚拟化项目时，不知道从哪里开始；也碰到对 KVM 技术感兴趣的运维人员在重复问一些相近的问题。于是，我们几个人萌生了一个想法，为什么不把我们的 KVM 项

目运维实践经验——记录下来，分享给大家。与此同时，我们也犹豫过，因为有许多问题我们也并未完全理解。但是，经过认真思考后的我们认为，即便在这种情况下，我们几年积累下来的实践经验对于想要了解和学习 KVM 虚拟化技术的工程师来说也是非常有意义的，毕竟实践更能检验真理。回想当初自己遇到问题的时候，我们也得到了一些论坛朋友、圈内专家的热心解答。互联网精神崇尚开放，这更加坚定了我们创作此书的决心。我们愿意将自己对 KVM 技术的理解和认识，以及在 KVM 虚拟化中解决问题的思路和方法分享出来，让大家少走一些我们走过的弯路，同时能够与所有运维工程师共同学习提升！

## 本书特色

本书是 KVM 一线项目运维实践经验总结，在 KVM CPU、内存、网络、磁盘方面拥有丰富的技术积累，在 KVM 虚拟化开源分布式与商业存储、KVM 桌面虚拟化方面有的大量实践，我们还会指出踩过的“技术坑”，这些都是弥足珍贵的经验。

本书对 Linux 系统和虚拟化有一定了解，但是对 KVM 不了解或者是了解不多，而准备在测试或者生产环境部署 KVM 虚拟化的读者，本书会介绍虚拟化项目如何评估、如何逐步实施，并介绍相关的技术与运维实践，希望对读者的生产环境虚拟化实施有所帮助。

对于正在做 KVM 虚拟化日常运维的读者，本书有相应的技术、实践经验、问题案例分享，便于这些读者碰到相似的问题时查询或者找到解决思路。

## 本书面向的读者

KVM 已经是 Linux 系统的标配，是 Linux 的必备技能，希望本书对广大 Linux 从业者能够有所帮助。

本书适用于以下读者：

- Linux 系统工程师。
- 从事 Linux 虚拟化开发的程序员。
- 虚拟化工程师。
- IDC 机房管理员。
- 有志于从事虚拟化的在校大学生。

## 如何阅读本书

本书分为五篇。

第一篇（第 1 ~ 8 章），介绍虚拟化选型及 KVM 虚拟机技术。实施虚拟化可带给我们哪

些收益，虚拟化引擎选择，KVM 虚拟化相关的 CPU、内存、磁盘、网络技术及生产环境应用场景，物理机转虚拟机，KVM 桌面虚拟化实践。

第二篇（第 9 ~ 10 章），介绍开源分布式文件系统在 KVM 虚拟化的应用。介绍一些常见的开源分布式文件系统如何在 KVM 环境使用，重点详细介绍最近一两年比较热的 Ceph 文件系统，包括配置、使用和常见故障处理。

第三篇（第 11 ~ 13 章），介绍开源的虚拟化管理平台在生产环境的实践经验。重点介绍 OpenStack、OpenNebula 在生产环境的实践及故障处理，还有 oVirt、CloudStack、Convirt、WebVirtMgr 几个管理平台的架构和理念。

第四篇（第 14 ~ 21 章），介绍 KVM 虚拟化项目实施经验，涵盖业务性能评估、压力模型建立、虚拟机性能测试、宿主机基准测试、业务测试环境测试验证、单机虚拟化技术、集群虚拟化技术、虚拟化正式环境上线等一套完整的虚拟化项目实施过程。

第五篇（第 22 章），介绍 KVM 虚拟化过程中比较典型的一些案例。主要介绍一些在生产环境中的案例及解决思路、方法，方便读者碰到类似的问题时找到解决方案。

根据本书内容，我们给出如下阅读建议：

- ☐ 如果是刚准备开始实施虚拟化的读者，请从头开始阅读本书。
- ☐ 如果对分布式文件系统在虚拟化中的应用感兴趣的读者，可以阅读第二篇的内容。
- ☐ 如果对虚拟化管理平台使用感兴趣的读者，可以阅读第三篇的内容。
- ☐ 如果是对虚拟化技术有一定了解，但是对不知道如何实施虚拟化的读者，建议阅读第四篇的内容。
- ☐ 如果在虚拟化实施碰到问题，需要寻找解决方法的读者，可以阅读第五篇的内容。

## 资源和勘误

由于时间紧张，笔者水平有限，书中难免有错误和不足之处，笔者运营着一个微信订阅号，名字为“KVM 虚拟化实践”，订阅号为“kvm\_virt”二维码为：



欢迎读者扫描关注，“KVM 虚拟化实践”会定时发布笔者对 KVM 虚拟化新的认识和经验总结。也欢迎读者通过这个微信订阅号和笔者交互，书中的勘误和更新也会通过这个订阅号发布。

## 致谢

感谢盛大游戏高级总监陈桂新，盛大游戏副总监、G 云 COO 冯祯旺对我们的支持，我们四个人都曾经或者现在在盛大游戏虚拟化小组工作，桂总、冯总一路上支持我们，坚定地支持我们沿着 KVM 的方向前进，即使中间碰到多次挫折。回首看，我们坚持了下来，也很高兴我们这些年的成果终于开花结果，孵化出了 G 云（[www.gicloud.com.cn](http://www.gicloud.com.cn)），一款专门针对游戏行业物理机和云主机的混合云。

感谢金山西山居运维总监赫宝生对本书写作的鼓励和支持。感谢金山西山居运维架构师刘宇对本书的帮助和建议。

感谢曾经在盛大游戏虚拟化小组工作过的刘晓光、秦洁、杨树林，感谢现在还在盛大游戏虚拟化小组工作的夏学峰、闫强，这本书也是我们所有人的经验总结。

感谢盛大游戏网络技术负责人、G 云网络架构师苏永华，金山西山居资深网络工程师李正奇在网络方面的支持和帮助。

感谢金山西山居资深系统工程师、Zabbix 权威专家薛群在宿主机监控方面的帮助。

特别感谢金山西山居高级系统工程师赵斌，提供了部分虚拟化生产环境实践案例及测试结果。

感谢盛大游戏资深工程师杨帆、季青、武文对我们对服务器知识方面的支持和帮助。

感谢国内顶尖的数据库专家叶金荣在本书成书过程中的支持和协助。

感谢机械出版社的杨福川和高婧雅，福川兄出版了一系列互联网技术图书，在认识福川兄之前，我们已经购买了多本福川兄出版的书，和福川兄的合作是我们的梦想。相信随着时间的推移，会有更多的人认识到福川兄对中国互联网技术巨大的贡献。高婧雅编辑的专业、细致、认真、耐心令我们非常敬佩，和高婧雅合作的大半年内，我们受益匪浅，很感谢、很庆幸有机会和杨兄、高编辑合作。

## 致我们的家人

感谢我的舅舅、舅妈、表哥、表姐，你们对我的人生观影响很大；感谢我的父亲、母亲，是你们一直在默默支持我；感谢我妻子和女儿一直以来的理解和支持，你们是我的动力。

——肖力，2015 年 6 月



感谢父母辛苦的培养，没有你们就没有今天的我。感谢老婆在我一无所有的时候默默陪在我的身边，还有可爱的儿子，你们是我一直前进的动力。

——汪爱伟，2015年6月

感谢生我养我的父亲母亲，这些年，你们辛苦了，儿子长大了，你们却老了。感谢所有一路和我走过来的人，你们的支持和信任，给了我前进的勇气和力量。感谢相信我，愿意和我一路走下去的人，无论明天怎样，我都希望一路有你。

——杨俊俊，2015年6月

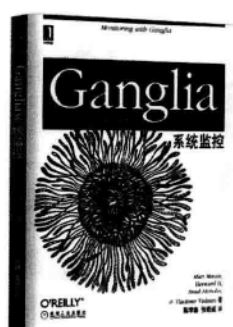
参与这本书的写作对我来说非常荣幸，首先要感谢肖力等前辈传授经验并耐心支持，其次还要感谢家人对我业余时间不能陪伴的理解！

——赵德禄，2015年6月

需要整本电子书，联系我QQ：3411317522

---

推荐阅读



---

需要整本电子书，联系我QQ：3411317522

需要整本电子书，联系我QQ：3411317522  
— 推荐阅读 —



### Cacti 实战

作者：刘钊 等 ISBN：978-7-111-50187-9 定价：69.00元



### Python 自动化运维：技术与最佳实践

作者：刘天斯 ISBN：978-7-111-48306-9 定价：69.00元



### Puppet 权威指南

作者：王冬生 ISBN：978-7-111-48598-8 定价：69.00元



### Linux 内核设计的艺术：图解Linux操作系统架构设计与实现原理（第2版）

作者：新设计团队 ISBN：978-7-111-42176-4 定价：89.00元



### Linux 内核揭秘：深入解析文件系统和设备驱动的架构与设计

作者：Hirokazu Takahashi 等 ISBN：978-7-111-41049-2 定价：79.00元

### Linux 内核探秘：深入解析文件系统和设备驱动的架构与设计

作者：高剑林 ISBN：978-7-111-44585-2 定价：59.00元

需要整本电子书，联系我QQ：3411317522

Contents 目 录

本书赞誉

前 言

第一篇 KVM 技术详解与实践

第 1 章 企业虚拟化选型与 KVM

介绍 ..... 2

1.1 KVM 的前世今生 ..... 2

1.2 KVM 与常用企业级虚拟化产品  
的 PK ..... 7

1.3 判断企业是否适合使用 KVM 的  
标准 ..... 9

1.4 本章小结 ..... 10

第 2 章 开始自己的第一台虚拟机 ..... 11

2.1 服务器 BIOS 设置 ..... 11

2.2 宿主机 CentOS 6.5、CentOS 7  
系统安装与配置技巧 ..... 13

2.3 第一台虚拟机安装 ..... 15

2.4 本章小结 ..... 21

第 3 章 CPU、内存虚拟化技术与  
应用场景 ..... 22

3.1 NUMA 技术与应用 ..... 22

3.1.1 KVM 虚拟机 NUMA 调优 ..... 24

3.1.2 CPU 绑定操作方法 ..... 27

3.2 CPU 热添加与应用 ..... 32

3.3 CPU host-passthrough 技术与应用 ..... 36

3.4 CPU Nested 技术与配置方法 ..... 38

3.5 KSM 技术与应用 ..... 39

3.6 内存气球技术详解与应用 ..... 40

3.7 内存限制技术与应用 ..... 43

3.8 巨型页内存技术与应用 ..... 44

3.9 本章小结 ..... 47

第 4 章 网络虚拟化技术与应用场景 ..... 48

4.1 半虚拟化网卡技术详解 ..... 48

4.1.1 半虚拟化网卡的配置 ..... 49

4.1.2 全虚拟化网卡、半虚拟化  
网卡性能比较 ..... 50

4.1.3 全虚拟化网卡、半虚拟化  
网卡的应用场景 ..... 63

4.2	MacVTap 和 vhost-net 技术原理与应用 .....	63
4.2.1	MacVTap 技术与应用 .....	64
4.2.2	vhost_net 技术 .....	66
4.3	网卡的中断与多队列 .....	66
4.4	网卡 PCI Passthrough 技术 .....	71
4.5	SR-IVO 虚拟化技术 .....	72
4.6	虚拟化软件交换机 Open vSwitch 的安装与配置 .....	75
4.6.1	Open vSwitch 安装 .....	75
4.6.2	Open vSwitch 典型配置 .....	77
4.7	多网卡绑定与建桥 .....	79
4.8	本章小结 .....	82

## 第5章 KVM 磁盘虚拟化技术与应用

5.1	磁盘虚拟化方式性能比较与应用场景 .....	83
5.1.1	QEMU 磁盘虚拟化方式概述 .....	83
5.1.2	IDE 与 Virtio 性能比较测试 .....	84
5.2	Virtio 磁盘缓存方式详解、性能比较与应用场景 .....	85
5.2.1	磁盘缓存详解 .....	86
5.2.2	缓存方式的性能比较测试 .....	87
5.2.3	缓存方式的应用场景 .....	89
5.3	磁盘镜像格式 .....	89
5.3.1	QEMU 支持的磁盘镜像格式 .....	89
5.3.2	裸设备使用 .....	96
5.3.3	KVM 虚拟机镜像管理利器 guestfish .....	97

5.3.4	raw、qcow2、裸盘、lvm 性能比较 .....	106
5.3.5	磁盘镜像格式的应用场景 .....	107
5.4	文件系统块对齐 .....	108
5.4.1	什么是块对齐 .....	108
5.4.2	块对齐与块不对齐性能比较 .....	109
5.4.3	生产环境如何配置块对齐 .....	110
5.5	SSD 在 KVM 虚拟化中的使用实践 .....	111
5.5.1	SSD 原理与写放大 .....	111
5.5.2	SSD 在 KVM 中的使用 .....	112
5.5.3	Flachcache 配置与性能测试 .....	117
5.5.4	DM-cache 配置 .....	120
5.5.5	LVM cache 配置与性能测试 .....	122
5.5.6	生产环境中 SSD 使用要点 .....	123
5.6	本章小结 .....	126

## 第6章 KVM 虚拟机的资源限制

6.1	哪些场景要做虚拟机的资源限制 .....	127
6.2	使用 CGroups .....	130
6.3	KVM 虚拟机 CPU 资源限制配置 .....	142
6.4	KVM 虚拟机网络资源限制 .....	144
6.5	KVM 虚拟机磁盘资源限制 .....	150
6.6	本章小结 .....	151

## 第7章 物理机转虚拟机实践

7.1	哪些应用场景适合做物理机转虚拟机 .....	152
7.2	P2V 的技术实施方案 .....	152
7.3	Windows 系统物理机转虚拟机的操作方法 .....	153

7.3.1 WinPE 的制作方法 .....	153
7.3.2 Windows 系统 P2V 操作方法 .....	155
7.4 Linux 物理服务器转换虚拟机 操作 .....	155
7.5 使用 virt-p2v 工具进行 P2V 转换 .....	157
7.6 P2V 经验总结 .....	160
7.7 关于 V2V .....	160
7.8 本章小结 .....	160

## 第 8 章 KVM 桌面虚拟化实践 .....

8.1 桌面虚拟化简介及 Spice 协议 .....	161
8.2 桌面虚拟化实践经验 .....	162
8.3 本章小结 .....	167

## 第二篇 分布式文件系统

## 第 9 章 几种常见开源文件系统在 KVM 中的应用 .....

9.1 演示环境基础配置 .....	170
9.2 DRBD 在 KVM 中的应用 .....	172
9.2.1 DRBD 介绍 .....	172
9.2.2 DRBD 搭建方法及在 KVM 中的应用 .....	173
9.2.3 DRBD 在 KVM 虚拟化中的 优化方案 .....	183
9.3 GlusterFS 在 KVM 虚拟化中的 应用 .....	184
9.3.1 GlusterFS 介绍 .....	184
9.3.2 GlusterFS 集群搭建与使用 .....	189
9.3.3 GlusterFS 在 KVM 虚拟化 中的应用 .....	195

9.3.4 GlusterFS 在 KVM 虚拟化 中的优化方案 .....	198
---	-----

## 9.4 Sheepdog 在 KVM 中的应用 .....

9.4.1 Sheepdog 介绍 .....	199
9.4.2 Sheepdog 搭建方法及 在 KVM 中的应用 .....	201

## 9.5 MooseFS 在 KVM 中的应用 .....

9.5.1 MooseFS 介绍 .....	208
9.5.2 MFS 搭建方法及在 KVM 虚拟化中的应用 .....	208

## 9.6 本章小结 .....

## 第 10 章 Ceph 在 KVM 虚拟化中的 应用与故障处理 .....

10.1 Ceph 简介 .....	214
10.2 Ceph 部署方法及在 KVM 中的 应用 .....	217
10.2.1 使用 ceph-deploy 工具部署 Ceph .....	217
10.2.2 挂载 CephFS .....	221
10.2.3 通过 Librbd 方式使用 Ceph RBD .....	223
10.2.4 使用内核方式挂载并使用 Ceph RBD .....	226
10.3 CRUSH 算法介绍及调优方法 .....	228
10.3.1 CRUSH 算法介绍 .....	228
10.3.2 CRUSH 算法调优方法 .....	230
10.4 多角度评估 Ceph 在 KVM 虚拟化环境中的应用 .....	234
10.5 Ceph 生产环境搭建案例 .....	240

10.5.1 使用 CephFS 集中备份虚拟机	240
10.5.2 使用 Librbd 将虚拟机运行在 Ceph RBD	241
10.5.3 在 CephFS 上创建 qcow2 文件作为云硬盘	242
10.6 Ceph 常见故障处理	242
10.6.1 Ceph 集群监控状态检查	243
10.6.2 Ceph 常见错误提示的含义与处理方法	244
10.7 本章小结	245

### 第三篇 KVM 虚拟化管理平台

#### 第 11 章 利用 OpenStack 管理

KVM	248
11.1 OpenStack 搭建	248
11.1.1 OpenStack 简介	248
11.1.2 控制节点的安装	249
11.1.3 计算节点的安装	266
11.2 OpenStack 实用技巧	269
11.2.1 RabbitMQ Cluster 搭建	269
11.2.2 利用 Haproxy 和 Keepalived 实现控制节点高可用	271
11.2.3 冷迁移虚拟机的方法	277
11.3 本章小结	277

#### 第 12 章 利用 OpenNebula 管理

KVM	278
12.1 OpenNebula 架构	278
12.1.1 OpenNebula 架构介绍	278

12.1.2 OpenNebula 组件介绍	279
12.2 在 KVM 中使用 OpenNebula	280
12.2.1 在 CentOS 6.5 上快速搭建 OpenNebula	280
12.2.2 使用 OpenNebula 创建第一台 KVM 虚拟机	285
12.2.3 使虚拟机硬盘创建在宿主机本地硬盘上	288
12.3 OpenNebula 常见故障处理	291
12.4 本章小结	293

#### 第 13 章 其他管理平台介绍

13.1 ConVirt 管理平台介绍	294
13.1.1 ConVirt 的理念	294
13.1.2 ConVirt 的架构	295
13.2 CloudStack 管理平台介绍	295
13.3 oVirt 管理平台介绍	297
13.4 WebVirtMgr 管理平台介绍	299
13.5 如何选择管理平台	303
13.6 本章小结	304

### 第四篇 KVM 虚拟化项目实施

#### 第 14 章 业务性能评估与压力

模型建立	306
14.1 Linux 系统性能数据采集原则与方法	306
14.2 Linux 系统压力模型建立	311
14.3 Windows 系统性能数据采集原则与方法	314
14.4 Windows 系统压力模型建立	317

14.5 本章小结 .....	319	16.5 本章小结 .....	375
<b>第 15 章 宿主机选型与基础性能测试</b> .....	320	<b>第 17 章 单机虚拟化技术与生产环境实践</b> .....	376
15.1 宿主机 CPU 选型原则 .....	320	17.1 单机虚拟化技术与应用场景 .....	376
15.2 宿主机内存选型注意事项 .....	322	17.1.1 单机虚拟化的优势和劣势 .....	377
15.3 宿主机硬盘选型 .....	324	17.1.2 单机虚拟化技术 .....	378
15.4 RAID 卡选型 .....	326	17.1.3 单机虚拟化应用场景 .....	380
15.5 服务器网卡选型 .....	327	17.2 单机虚拟化生产环境注意点 .....	381
15.6 宿主机性能基准测试实践 .....	328	17.2.1 单机虚拟化灾备与应急方案 .....	381
15.7 本章小结 .....	345	17.2.2 单机虚拟化生产环境部署实践 .....	382
<b>第 16 章 虚拟机镜像制作、配置与测试</b> .....	346	17.3 本章小结 .....	385
16.1 Windows 虚拟机制作方法 .....	346	<b>第 18 章 KVM 虚拟化集群技术与应用场景</b> .....	386
16.1.1 Windows 镜像制作 .....	346	18.1 虚拟化集群设计思路与架构 .....	386
16.1.2 Windows 虚拟机 sysprep 初始化封装 .....	361	18.2 虚拟化集群的实施流程与技术方案 .....	389
16.2 Linux 镜像制作方法 .....	362	18.3 虚拟化集群搭建 .....	392
16.2.1 RHEL/CentOS 镜像制作方法 .....	362	18.3.1 计算节点配置与管理 .....	392
16.2.2 Ubuntu、Debian 虚拟机配置注意点 .....	364	18.3.2 网络设计与配置 .....	403
16.3 虚拟机自动配置 IP 的实现办法 .....	365	18.3.3 虚拟机迁移管理 .....	404
16.3.1 通过 DHCP 给虚拟机配置 IP .....	365	18.3.4 虚拟化集群的备份与恢复 .....	407
16.3.2 通过 Libvirt 实现虚拟机自动配置 IP .....	365	18.4 Dell EqualLogic 存储的管理实践经验 .....	414
16.3.3 笔者生产环境虚拟机 IP 自动配置的方法 .....	365	18.5 本章小结 .....	419
16.4 虚拟机镜像测试 .....	375	<b>第 19 章 业务迁移到虚拟化环境流程</b> .....	420
		19.1 性能评估与测试环境测试 .....	420



19.2	上线前的检查工作 .....	428
19.3	小规模部署及逐步切换到 虚拟化环境 .....	429
19.4	本章小结 .....	430

## 第 20 章 宿主机自动化运维管理 .....

20.1	Puppet 简介与安装部署 .....	431
20.2	利用 Puppet 管理宿主机的 运维架构设计 .....	436
20.3	利用 Puppet 配置宿主机初始化 操作 .....	441
20.4	利用 Puppet 管理宿主机配置与 更新 .....	442
20.5	利用 Puppet 推送宿主机的 非一致性配置 .....	445
20.6	本章小结 .....	449

## 第 21 章 虚拟化监控、报警与应急 响应方案 .....

21.1	虚拟化监控内容 .....	450
21.1.1	宿主机 CPU 监控要点 .....	450
21.1.2	宿主机内存监控注意点 .....	455
21.1.3	宿主机网络监控注意点 .....	458
21.1.4	宿主机磁盘监控注意点 .....	460
21.1.5	宿主机层面监控虚拟机的 方法 .....	461
21.2	监控软件 Cacti、Zabbix .....	463
21.2.1	用 Cacti 监控宿主机 .....	464
21.2.2	用 Zabbix 监控宿主机 .....	471
21.3	应急方案 .....	487

21.3.1	应急方案制定注意要点 .....	487
21.3.2	单机虚拟化应急方案 .....	491
21.3.3	集群虚拟化应急方案 .....	495
21.4	本章小结 .....	497

## 第五篇 KVM 虚拟化典型案例

## 第 22 章 生产环境问题案例与 分析 .....

22.1	游戏在虚拟机上非常慢的 案例 .....	500
22.2	虚拟机流量过高引起网络风暴 的案例一 .....	502
22.3	虚拟机流量过高引起网络风暴 的案例二 .....	517
22.4	其他生产环境问题案例 .....	520
22.4.1	共享存储集群虚拟机迁移 故障案例 .....	520
22.4.2	宿主机异常关机后虚拟机 无法启动案例 .....	522
22.4.3	宿主机使用 CentOS 5.6 系统问题案例 .....	523
22.4.4	Windows 虚拟机网络闪断、 不通问题案例 .....	523
22.4.5	Windows 7 虚拟机只能使用 2 个 CPU 案例 .....	524
22.4.6	生产环境非常奇怪的一个 案例 .....	524
22.4.7	U 盘在虚拟机上挂载问题 案例 .....	526

22.4.8	HP ILO4 固件 Bug 引起 宿主机重启案例 .....	528	22.4.12	生产环境 Ubuntu 虚拟机 故障问题案例 .....	533
22.4.9	一次机房断电引起的问题 案例 .....	529	22.5	本章小结 .....	538
22.4.10	CPU 绑定问题案例 .....	531	<b>附录 笔者推荐的 KVM 学习 资料 .....</b>		539
22.4.11	生产环境存储故障导致 50 个虚拟机丢失案例 .....	532			

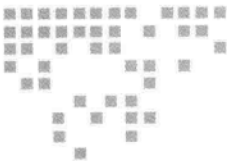
需要整本电子书，联系我QQ：3411317522

第一篇 *Part I*

## KVM 技术详解与实践

- 第 1 章 企业虚拟化选型与 KVM 介绍
- 第 2 章 开始自己的第一台虚拟机
- 第 3 章 CPU、内存虚拟化技术与应用场景
- 第 4 章 网络虚拟化技术与应用场景
- 第 5 章 KVM 磁盘虚拟化技术与应用场景
- 第 6 章 KVM 虚拟机的资源限制
- 第 7 章 物理机转虚拟机实践
- 第 8 章 KVM 桌面虚拟化实践

需要整本电子书，联系我QQ：3411317522



# 企业虚拟化选型与 KVM 介绍

虚拟化技术从 2008 年开始越来越热，经过一个大爆发的阶段，目前已经是企业 IT 环境的必备技术，在许多企业里面，虚拟机的数量已经远远大于物理机。同许多技术一样，虚拟化也分为开源和闭源技术。KVM 就是一种开源的虚拟化技术，本章将为读者介绍 KVM 技术的演进过程、KVM 的发展历史及 KVM 的应用场景。

## 1.1 KVM 的前世今生

如图 1-1 所示，根据 OpenStack 平台上 2013 年 10 月虚拟化引擎选择的调查统计数据，KVM 已经在 OpenStack 平台占到 71% 的份额。

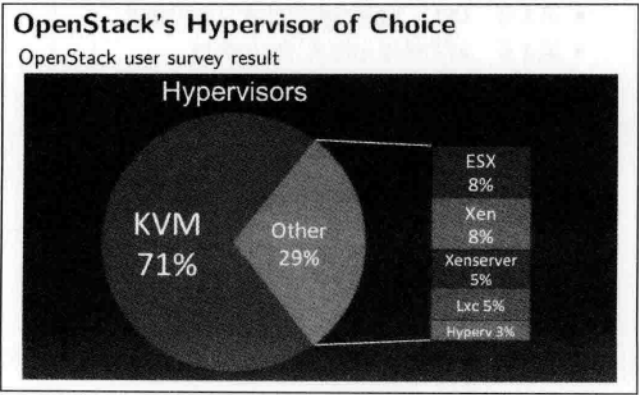


图 1-1 OpenStack 平台虚拟化引擎，用户调查结果

KVM 是一种年轻的虚拟化技术，在出生的时候就吸取了其他虚拟化技术的优点，所以 KVM 的架构简单，没有历史兼容性的包袱，性能表现优异。

本节介绍一下 KVM 的演进过程。

1. 虚拟化技术的演变过程

虚拟化技术的演变过程可以分为软件模拟、虚拟化层翻译、容器虚拟化三个大的阶段。

其中，虚拟化层翻译又可以分为：

- ❑ 软件捕获翻译，即软件全虚拟化。
- ❑ 改造虚拟机系统内核加虚拟化层翻译，即半虚拟化。
- ❑ 硬件支持的虚拟化层翻译，即硬件支持的全虚拟化。

下面分别介绍一下这几种技术方式。

(1) 软件模拟的技术方式

软件模拟是通过软件完全模拟 CPU、芯片组、磁盘、网卡等计算机硬件，如图 1-2 所示。

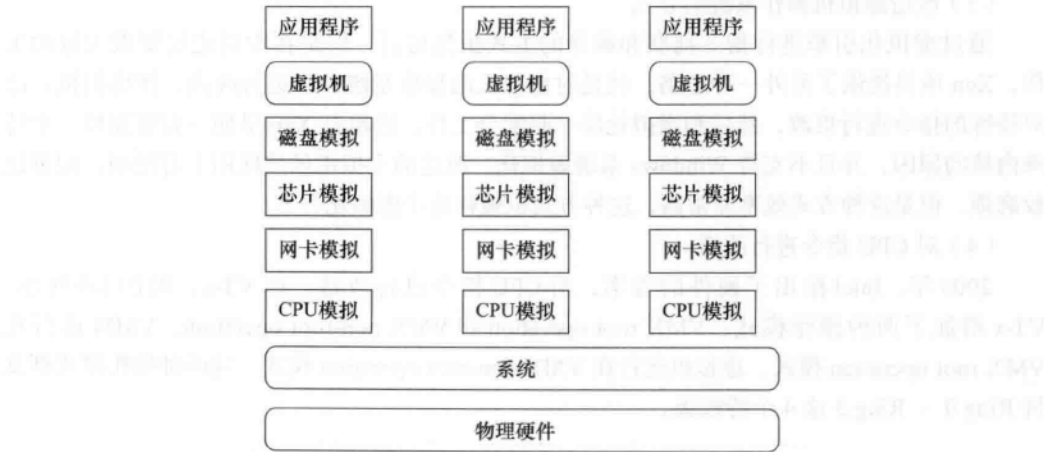


图 1-2 QEMU 的虚拟机架构

因为是软件模拟，所以理论上可以模拟任何硬件，甚至是不存在的硬件。但是因为这种方式全部是软件模拟硬件，所以非常低效，一般只用于研究测试的场景。采用这种技术的典型产品有 Bochs、QEMU 等。

(2) 虚拟化层翻译

先介绍一下 X86 平台的指令集权限划分。如图 1-3 所示，X86 平台指令集划分为 4 个特权模式：Ring 0、Ring 1、Ring 2、Ring 3。操作系统一般使用 Ring 0 级别，应用程序使用 Ring 3 级别，驱动程序使用 Ring 1 和 Ring 2 级别。X86 平台在虚拟化方面的一个难点就是如何将虚拟机越级的指令使用进行隔离。

VMware 公司找到了最早的解决方法，图 1-4 所示为对虚拟机指令的使用进行捕获和翻译的示意图。



图 1-3 X86 CPU 指令级别

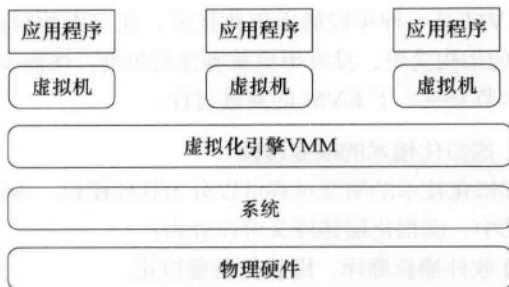


图 1-4 软件全虚拟化方案架构

通过虚拟化引擎，捕获虚拟机的指令，并进行处理，这也是为什么在虚拟机上虽然使用的是物理机一样的指令，但是虚拟机不能对硬件进行操作的原因，比如重启虚拟机不会引起宿主机的重启。这种解决方案也叫软件全虚拟化方案。

### (3) 改造虚拟机操作系统的方式

通过虚拟化引擎进行指令捕获和翻译的方式虽然可行，但是在虚拟化层要做大量的工作，Xen 项目提供了另外一种思路，就是对虚拟机的操作系统内核进行改造，使虚拟机自己对特殊的指令进行更改，然后和虚拟化层一起配合工作，这也是 Xen 早期一直要使用一个特殊内核的原因，并且不支持 Windows 系统虚拟化。改造的虚拟机虽然使用上有限制，配置比较麻烦，但是这种方式效率非常高，这种方式也被称为半虚拟化方案。

### (4) 对 CPU 指令进行改造

2005 年，Intel 推出了硬件的方案，对 CPU 指令进行改造，即 VT-x，如图 1-5 所示。VT-x 增加了两种操作模式：VMX root operation 和 VMX non-root operation。VMM 运行在 VMX root operation 模式，虚拟机运行在 VMX non-root operation 模式。这两种操作模式都支持 Ring 0 ~ Ring 3 这 4 个特权级。

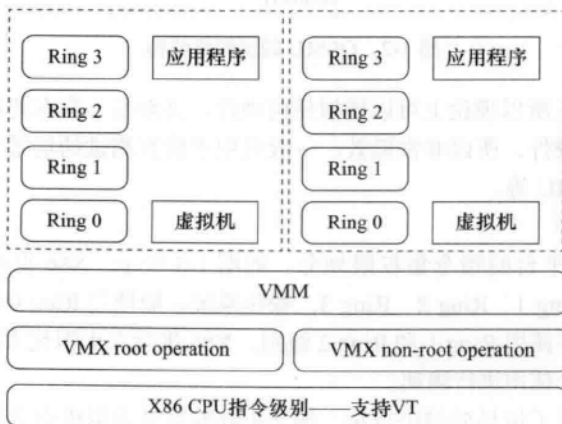


图 1-5 硬件虚拟化 CPU 指令说明

这种方案因为是基于硬件的，所以效率非常高，这种方案也称为硬件支持的全虚拟化方案，如图 1-6 所示。现在的一个发展趋势是不仅 CPU 指令有硬件解决方案，I/O 通信也有硬件解决方案，称为 VT-d；网络通信也有硬件解决方案，称为 VT-c。

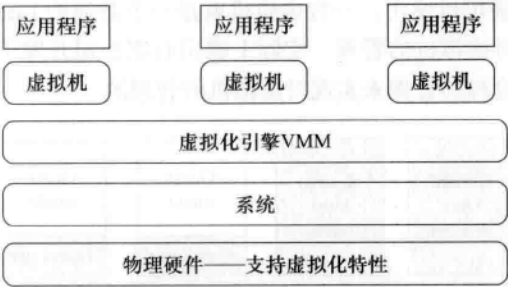


图 1-6 硬件全虚拟化方案架构

**提示** 当前的虚拟化引擎，都是使用硬件支持的虚拟化解决方案。并且最新的操作系统一般都支持一些半虚拟化的特性，所以宿主机和虚拟机使用比较新的版本，性能也会好一些。

(5) 容器虚拟化

容器虚拟化的原理是基于 CGroups、Namespace 等技术将进程隔离，每个进程就像一台单独的虚拟机一样，有自己被隔离出来的资源，也有自己的根目录、独立的进程编号、被隔离的内存空间。基于容器的虚拟化可以实现在单一内核上运行多个实例，因此是一个更高效率的虚拟化方式。目前最热的容器虚拟化技术就是 Docker。Docker 的优势是可以将一个开发环境进行打包，很方便地在另外一个系统上运行起来，并且有版本的概念，可以在前一个版本的基础上累加。但是 Docker 在生产环境的使用还需要一个过程，主要是磁盘、网络性能上还受到很多限制。

2. KVM 的历史

KVM (Kernel-based Virtual Machine) 最初是由以色列的公司 Qumranet 开发的。KVM 在 2007 年 2 月被正式合并到 Linux 2.6.20 核心中，成为内核源代码的一部分。2008 年 9 月 4 日，RedHat 公司收购了 Qumranet，开始在 RHEL 中用 KVM 替换 Xen，第一个包含 KVM 的版本是 RHEL 5.4。从 RHEL 6 开始，KVM 成为默认的虚拟化引擎。KVM 必须在具备 Intel VT 或 AMD-V 功能的 X86 平台上运行。它也被移植到 S/390、PowerPC 与 IA-64 平台上。在 Linux 内核 3.9 版中，加入了对 ARM 架构的支持。

KVM 包含一个为处理器提供底层虚拟化、可加载的核心模块 kvm.ko (kvm-intel.ko 或 kvm-amd.ko)，使用 QEMU (QEMU-KVM) 作为虚拟机上层控制工具。KVM 不需要改变

Linux 或 Windows 系统就能运行。

### 3. KVM 的架构

KVM 的架构非常简单，如图 1-7 所示，KVM 就是内核的一个模块，用户空间通过 QEMU 模拟硬件提供给虚拟机使用，一台虚拟机就是一个普通的 Linux 进程，通过对这个进程的管理，就可以完成对虚拟机的管理。实际上德国有一家公司开发了一个管理平台 Proxmox VE，就是通过对 KVM 进程的管理来实现对虚拟机的管理的。

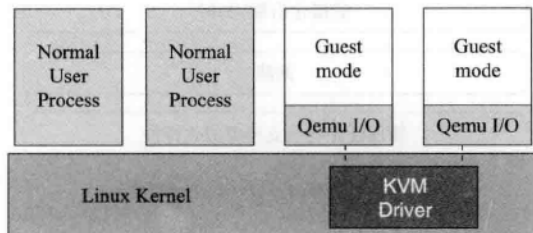


图 1-7 KVM 的架构

因为对进程的管理非常麻烦，RedHat 发布了一个开源项目 Libvirt。Libvirt 有 API，也有一套命令行工具，可以完成对虚拟机的管理，大多数的管理平台都是通过 Libvirt 来完成对 KVM 虚拟机的管理的，比如 OpenStack、CloudStack、OpenNebula 等。

### 4. QEMU 与 KVM

QEMU 是一个开源项目，实际就是一台硬件模拟器，可以模拟许多硬件，包括 X86 架构处理器、AMD64 架构处理器、MIPS R4000、ARM v6/v7 (Cortex-A8, A9, A15)、SPARC sun3 与 PowerPC 架构，还支持其他架构，可以从 QEMU 主页获取完整的列表。

QEMU 可以在其他平台上运行 Linux 的程序，可以存储及还原虚拟机运行状态，还可以虚拟多种设备，包括网卡、多 CPU、IDE 设备、软驱、显卡、声卡、多种并口设备、多种串口设备、多种 USB 设备、PC 喇叭、PS/2 键盘鼠标（默认）和 USB 键盘鼠标、蓝牙设备。

QEMU 还内建 DHCP、DNS、SMB、TFTP 服务器。

比较有意思的是，有人还将 QEMU 编译成 Windows 版本，在 Windows 平台上也可以运行 QEMU。

QEMU 的好处是因为是纯软件模拟，所以可以在支持的平台模拟支持的设备，比如还有人利用 QEMU 在安卓上安装一个 Windows XP 虚拟机出来。

QEMU 的缺点是因为是纯软件模拟，所有非常慢。QEMU 1.0 的时候有一个 QEMU 和 KVM 结合的分支。KVM 只是一个内核的模块，没有用户空间的管理工具，KVM 的虚拟机可以借助 QEMU 的管理工具来管理。QEMU 也可以借助 KVM 来加速，提升虚拟机的性能。QEMU-KVM 的分支版本发布了 3 个正式的版本 1.1、1.2、1.3，随后和 QEMU 的主版本合并，也就是说现在的 QEMU 版本默认支持 KVM，QEMU 和 KVM 已经紧密地结合起来了。



KVM 的最后一个自己的版本是 KVM 83，随后和内核版本一起发布，和内核版本号保持一致，所以要使用 KVM 的最新版本，就要使用最新的内核。

## 5. Libvirt 与 KVM

Libvirt 是一套开源的虚拟化的管理工具，主要由 3 部分组成：

- ❑ 一套 API 的 lib 库，支持主流的编程语言，包括 C、Python、Ruby 等。
- ❑ Libvirtd 服务。
- ❑ 命令行工具 virsh。

Libvirt 的设计目标是通过相同的方式管理不同的虚拟化引擎，比如 KVM、Xen、HyperV、VMware ESX 等。但是目前实际上多数场景使用 Libvirt 的是 KVM，而 Xen、HyperV、VMware ESX 都有各自的管理工具。

Libvirt 可以实现对虚拟机的管理，比如虚拟机的创建、启动、关闭、暂停、恢复、迁移、销毁，以及虚拟机网卡、硬盘、CPU、内存等多种设备的热添加。

Libvirt 还支持远程的宿主机管理，只要在宿主机上启动 Libvirtd 服务并做好配置，就可以通过 Libvirt 进行虚拟机的配置。通道可以是以下方式：

- ❑ SSH。
- ❑ TCP。
- ❑ 基于 TCP 的 TLS。

Libvirt 将虚拟机的管理分为以下几个方面：

第一，存储池资源管理，支持本地文件系统目录、裸设备、lvm、nfs、iscsi 等方式。在虚拟机磁盘格式上支持 qcow2、vmdk、raw 等格式。

第二，网络资源管理，支持 Linux 桥、VLAN、多网卡绑定管理，比较新的版本还支持 Open vSwitch。Libvirt 还支持 nat 和路由方式的网络，Libvirt 可以通过防火墙让虚拟机通过宿主机建立网络通道，和外部的网络进行通信。

## 1.2 KVM 与常用企业级虚拟化产品的 PK

### 1. 常用企业级虚拟化产品的比较

目前常见的企业级的虚拟化产品有 4 款：分别是 VMware、HyperV、Xen、KVM。

#### (1) VMware

VMware 是最早的 X86 平台上的虚拟化引擎，1999 年就发布了第一款产品，经过十几年的发展和市场检验，产品成熟、稳定，兼容性也不错。VMware 的产品线非常全面，不仅有虚拟化的解决方案，在 IaaS、SaaS、PaaS 层都有自己的产品。并且 VMware 在网络、存储方面都有相关的解决方案，VMware 和网络存储厂商在协议层面也有一些私有协议，许多主流的厂商都支持 VMware 一些专用的协议，和 VMware 一起形成了一个生态链。

VMware 目前被 EMC 控股，虚拟化产品线主要有针对个人使用的 VMware Workstation，针对苹果用户的 VMware Fusion，针对企业级用户的 VMware ESXi 服务器。管理工具主要是 VMware vSphere 套件。

VMware 的产品基本上都是非开源产品，并且大部分都是收费产品，一般在传统关键行业使用比较多一些，在中小型企业、互联网行业使用得比较少一点。

## （2）HyperV

HyperV 是微软的虚拟化产品，最近几年发展非常迅速，在 Windows Server 2012 R2 中的 HyperV 支持许多非常新的虚拟化特性。HyperV 必须使用 64 版的 Windows 产品，HyperV 也支持 Linux 系统的虚拟机。

HyperV 也是一款非开源的收费产品，HyperV 的集群管理工具 SCVMM 配置非常复杂，需要配置 Windows 域、Windows Server 集群，然后才能管理多台宿主机。因为 HyperV 的成本相对较低，所以最近几年市场占有率也在提升，主要是一些使用 Windows 系统的企业使用比较多。

## （3）Xen

Xen 是最早的开源虚拟化引擎，由剑桥大学开发，半虚拟化的概念也是 Xen 最早提出的。Xen 后来被思杰收购，推出了一套叫作 XenServer 的管理工具，XenServer 于 2013 年年底宣布免费。Xen 因为推出的时间比较长，兼容性、稳定性都不错，目前使用 Xen 的主要是一些在 Xen 上面技术积累较多的企业。

## （4）KVM

KVM 比较年轻，所以出生的时候就吸取了其他虚拟化技术的优点，一开始就支持硬件虚拟化技术，没有历史兼容包袱。所以 KVM 推出来的时候，性能就非常优异。目前，KVM 是 OpenStack 平台上首选的虚拟化引擎。国内新一代的公有云全部采用 KVM 作为底层的虚拟化引擎。KVM 已经成为开源解决方案的主流选择。

## 2. KVM 优势

KVM 的优势主要体现在以下几点。

### （1）开源

KVM 是一个开源项目，这就决定了 KVM 一直是开放的姿态，许多虚拟化的新技术都是首先在 KVM 上应用，再到其他虚拟化引擎上推广。

虚拟化一般网络和存储都是难点。网络方面，SRIOV 技术就是最先在 KVM 上先有应用，然后再推广到其他虚拟化引擎上。再比如 SDN、Open vSwitch 这些比较新的技术，都是先在 KVM 上得到应用。

磁盘方面，基于 SSD 的分层技术，都是最早在 KVM 上得到应用。

KVM 背靠 Linux 这棵大树，和 Linux 系统紧密结合，在 Linux 上的新技术都可以马上应用到 KVM 上。围绕 KVM 的是一个开源的生态链，从底层的 Linux 系统，到中间层的

Libvirt 管理工具，到云管理平台 OpenStack，莫不是如此。

#### (2) 性能

KVM 吸引许多人使用的一个动因就是性能，在同样的硬件条件下，能提供更好的虚拟机性能，主要是因为 KVM 架构简单，代码只有 2 万行，一开始就支持硬件虚拟化，这些技术特点保证了 KVM 的性能。

#### (3) 免费

KVM 因为是开源项目，绝大部分 KVM 的解决方案都是免费方案，随着 KVM 的发展，KVM 虚拟机越来越稳定，兼容性也越来越好，因而也就得到越来越多的应用。

#### (4) 广泛免费的技术支持

免费并不意味着 KVM 没有技术支持。在 KVM 的开源社区，数量巨大的 KVM 技术支持者都可以提供 KVM 技术支持。另外，如果需要商业级支持，也可以购买红帽公司的服务。

## 1.3 判断企业是否适合使用 KVM 的标准

### 1. 业务类型

目前在互联网行业，KVM 虚拟化技术是使用最广泛的，因为互联网行业有对新技术追求的冲动。另外中小企业也推荐使用 KVM 技术，因为使用 KVM 不需要支付额外的费用。

还有哪些企业、个人适合使用 KVM 呢？可以这样总结：KVM 天生就是为 Linux 而生的，凡是可以使用 Linux 的地方，就可以使用 KVM，KVM 天生和 Linux 在一起。

### 2. 企业对成本的关注度

随着虚拟化这几年的快速发展，在企业的 IT 环境中，虚拟化已经是一个标配的技术。虚拟化带来的好处是巨大的，虚拟化真正实现了资源池化，通过虚拟化可以将服务器资源进行切割，做到资源随取随用，有效节省成本，提高资源利用率。

KVM 虚拟化技术经过几年的发展，已经非常成熟，使用 KVM 技术更不需要在虚拟化方面支付额外的费用，可以进一步节省企业的 IT 成本。

### 3. 企业对快速部署的关注度

虚拟化还有一个重要的功能就是快速部署。在宿主机层面看，虚拟机就是一个镜像文件，要得到另外一台虚拟机，只需要将镜像文件复制一份就可以了，通常只有几分钟。而按照传统方式部署一台物理机，最起码都要一个小时。通过磁盘差量的技术，甚至可以做到秒级生成虚拟机。

通过虚拟化技术，还可以实现在虚拟化层做高可用和在线迁移。虚拟化层的高可用是系统层面的高可用，比基于应用层的高可用配置要简单很多。虚拟机的在线迁移，更是虚拟化技术的独有手段。通过在线迁移，物理机的维护、系统的维护、网络的维护，都可以做到不中断服务，进一步提高了业务应用的可用性。

因为所有的虚拟机都是相同的虚拟硬件，实现了硬件层级的标准化，降低了自动化的难度，很容易搭建私有或者公有的云平台，所以通过虚拟机技术，很容易实现以下功能：

- 资源使用自动化和自助服务。
- 资源池可随时扩展。
- 资源使用标准化。
- 资源使用保持了很好的兼容性。

## 1.4 本章小结

本章介绍了虚拟化及 KVM 技术的演进过程，介绍了 KVM 技术的优势，可以简单地概括：凡是使用 Linux 系统的场景，如果上虚拟机项目，就天然适合使用 KVM 虚拟化技术。

下一章将介绍如何开始自己的第一台虚拟机，在安装第一台虚拟机的时候有哪些地方应该注意。

## 开始自己的第一台虚拟机

通过第 1 章对 KVM 的发展历史、架构及应用场景的介绍，读者对 KVM 技术有了初步的了解和认识。本章将介绍如何创建一台 KVM 虚拟机、操作步骤及要点。对于熟悉或者使用过 KVM 虚拟化的读者，本章内容可以快速阅读或者跳过。

### 2.1 服务器 BIOS 设置

KVM 的使用必须有硬件虚拟化支持，所以需要打开 CPU 的硬件虚拟化特性。对于大多数服务器，如 Dell、HP、IBM、浪潮、联想、华为等，在开机启动的第一个画面都会有一些提示的按键，图 2-1 所示是 Dell R610 开机启动的第一个画面。

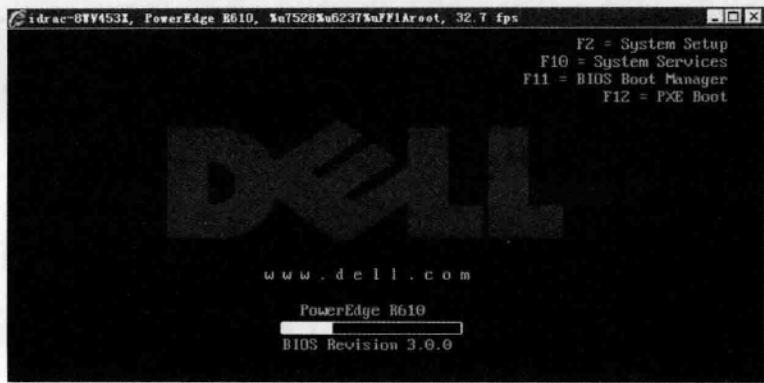


图 2-1 Dell R610 开机启动的第一个画面

该界面提供了如下几个选项：

- ❑ 按 F2 键进入 System Setup，进去之后可以进行一些 BIOS 相关的配置。
- ❑ 按 F10 键进入 System Services，进去之后可以配置一些系统服务，主要用于服务器的配置和检测。
- ❑ 按 F11 键进入 BIOS Boot Manager，进入之后可以选择启动介质，比如从硬盘启动、虚拟 DVD 设备、USB 设备，等等。
- ❑ 按 F12 键进入 PXE，进入之后启动网络引导。

 **提示** 不同的厂商服务器，功能键的定义会有一些差别，可以根据提示来操作。

在安装 KVM 虚拟化之前，需要先确认 CPU 虚拟化支持是否开启。在如图 2-1 所示的服务器开机界面中按 F2 键，然后进入 BIOS 配置，选择 Processor Setting，按 Enter 键进入子菜单，其中一项 Virtualization Technology 必须配置为 Enabled（选择菜单项，按方向键右键可更改配置），如图 2-2 所示。

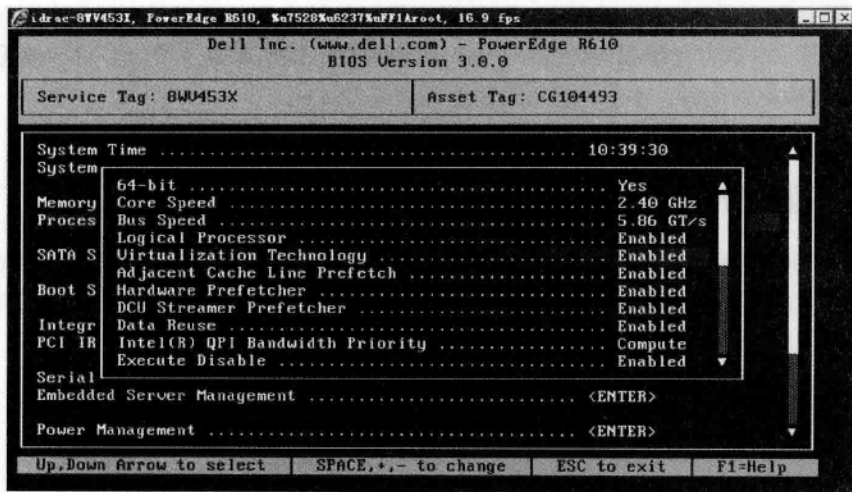



图 2-2 开启 CPU 虚拟化支持

 **提示** 不同厂商的服务器，CPU Virtualization Technology 开关的位置会有些差别，一般都在处理器配置菜单之下。近年来服务器上的 CPU 一般都支持 CPU 虚拟化，在系统中可以执行 `egrep '(vmx|svm)' /proc/CPUinfo` 命令查看，如果有输出内容，说明 CPU 是支持虚拟化的。

## 2.2 宿主机 CentOS 6.5、CentOS 7 系统安装与配置技巧

通过 2.1 节的配置，已经打开了服务器 CPU 对虚拟化的支持，下面开始安装宿主机的操作系统。CentOS 6.x 是目前使用比较多的宿主机操作系统版本，当前最新版本是 CentOS 6.6。CentOS 7 从系统跨度来说，是 CentOS 的一个大版本升级，其中主要包括内核版本的跨越。CentOS 6.x 系列使用的是 2.6.32.x 的内核，CentOS 7.x 则使用 3.10.x 的内核。

### 1. CentOS 6.5 宿主机系统安装及配置

在生产环境中，安装 CentOS 6.5 系统的宿主机，笔者采用 pxe 方式来完成批量的宿主机部署。pxe 配置过程的资料很多，本书就不做详细介绍了，下面分享一个宿主机安装 kickstart 文件的内容，是笔者实际在生产环境中使用的。

```
# Kickstart file automatically generated by anaconda
# 系统安装自动生成的 kickstart 文件
install
url --url=http://10.10.10.1/system/CentOS6564
# 指定安装镜像的目录，可以将不同的系统放在不同的目录，这样可以完成多个系统安装
lang en_US.UTF-8
# 宿主机建议使用英文
keyboard us
network --onboot yes --device eth0 --mtu=1500 --bootproto dhcp
network --onboot yes --device eth1 --noipv4 --noipv6
# 设置网络，笔者装机的时候，一般第一个网卡使用 dhcp 分配 IP，第二个网卡关闭。读者可以根据自己的
  实际情况进行修改
rootpw cEmXc2pkKets # 配置系统密码
text
reboot
# 文本方式安装，安装后重启
firewall --disabled
authconfig --useshadow --passalgo=sha512
# 用户密码加密 selinux --disabled # 关闭 SELinux
#####
# Installation logging level
logging --level=info
# Do not configure the X Window System
skipx
#####
timezone Asia/Shanghai # 配置时区
bootloader --location=mbr --driveorder=sda --append="crashkernel=auto rhgb quiet"
clearpart --all --initlabel # 清除硬盘数据，并创建相应分区
part /boot --fstype ext3 --size=256
part pv.3 --size=40960
part pv.4 --size=100 --grow
volgroup datavg --pesize=32768 pv.4
volgroup KVMvg --pesize=32768 pv.3
logvol /datapool --fstype ext4 --name=datapool --vgname=datavg --size=10240
--grow # 此处创建了一个目录，用于存放虚拟机，笔者习惯的目录是 /datapool
```