

面向互联网 的虚拟计算环境

卢锡城 王怀民 朱 鸿 王 戟 主编



科学出版社

面向互联网的虚拟计算环境

卢锡城 王怀民 朱 鸿 王 戟 主编

科学出版社

北 京

内 容 简 介

本书在互联网计算的背景下系统地研究“虚拟化”问题,阐述“按需聚合,自主协同”的互联网资源管理模型和方法,该模型和方法将计算机时代面向静态统一视图的资源组织模式拓展为互联网时代面向动态相对稳定视图的按需可伸缩的资源组织模式,将传统的基于集中控制的资源共享模式转换为自主协同的资源共享模式,以实现灵活构建支撑高效互联网计算的虚拟计算环境。

本书结合课题组研究实践,详细介绍虚拟计算环境在概念模型与体系结构、资源虚拟化、虚拟资源的分布式管理、虚拟资源的协同机制、编程语言和环境等方面的研究成果和最新进展。

本书可作为高等院校互联网计算、云计算等专业高年级本科生和研究生的教材,也可供相关领域科研人员和工程技术人员参考。

图书在版编目(CIP)数据

面向互联网的虚拟计算环境/卢锡城等主编. —北京:科学出版社, 2016

ISBN 978-7-03-048865-7

I. ①面… II. ①卢… III. ①计算机网络—研究 IV. ①TP393

中国版本图书馆 CIP 数据核字(2016)第 136612 号

责任编辑:王 哲 董素芹 / 责任校对:郭瑞芝

责任印制:张 倩 / 封面设计:迷底书装

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

文林印务有限公司 印刷

科学出版社发行 各地新华书店经销

*

2016 年 6 月第 一 版 开本:720×1 000 1/16

2016 年 6 月第一次印刷 印张:12 1/4

字数:240 000

定价:76.00 元

(如有印装质量问题,我社负责调换)

前 言

互联网正由一般意义下的计算机通信平台逐步演变成为一个覆盖全球的虚拟计算环境。互联网上汇聚的海量计算资源、存储资源、数据资源和应用资源已经成为国家重要战略资源。建立以有效管理和综合利用互联网资源为目标的虚拟计算环境，是全球关注的、具有重大产业价值的课题，已被列入了《国家中长期科学和技术发展规划纲要（2006—2020）》。

课题组从 2005 年开始，在国家 973 计划两轮项目的资助下，历时 10 年，开展了面向互联网的虚拟计算环境（Internet-Based Virtual Computing Environment, iVCE）的研究。iVCE 旨在面向互联网基础设施构建类似计算机操作系统的一体化服务环境，支持用户在开放的网络上有效地共享资源，便捷地协同工作。互联网资源所特有的动态不确定性、自治不可控性和异构多尺度性使得在计算机系统资源管理中的传统方法不再适用，必须研究互联网资源管理的新机理和新的软件环境。本书阐述了 iVCE 的重要成果和最新发展。全书分为 7 章，具体安排如下：

第 1 章是引言部分，介绍互联网的发展历史，分析互联网资源的自然特性，介绍资源虚拟化的基本概念，进而提出面向互联网的虚拟计算环境所面临的科学技术问题。

第 2 章提出 iVCE 的概念模型与体系结构，介绍在解决科学技术问题过程中的设计考虑，并与现有的计算模型进行对比分析，进一步展示 iVCE 的特点。

第 3 章介绍 iVCE 中的资源虚拟化技术，针对计算、内存和网络等资源，分别介绍其技术内涵、研究进展以及在构成 iVCE 自主元素过程中的应用。

第 4 章介绍 iVCE 中的虚拟资源分布式管理技术，包括虚拟资源的分布式组织、虚拟资源的分布式搜索，以及虚拟资源的优化管理等。

第 5 章介绍 iVCE 中的虚拟资源协同技术，在资源管理虚拟共同体模型的基础上，提出基于服务的虚拟执行体协同机制，主要包括事件驱动的协同机制和协同激励机制等。

第 6 章针对 iVCE 自主元素的特点，采用 Agent 的方式对其进行控制和访问，进而提出一种以 Agent 为基本构成单元的程序设计语言 CAOPLE。

第 7 章简要介绍互联网应用的新发展，阐述互联网新型应用发展对互联网规模计算带来的新挑战，讨论 iVCE 为应对这些挑战的最新研究进展，包括面向互联网计算的多尺度弹性体系结构模型、多尺度资源的弹性绑定，以及支持多尺度资源的高效聚合和可信服务。

本书由卢锡城、王怀民、朱鸿、王戟主编，参加编写的人员有李东升、张一鸣、

王意洁、毛新军、褚瑞、李慧霸、符永铨、陈振邦、唐扬斌、朱鸿、周斌、张圣栋等，英国牛津布鲁克斯大学的 Ian Bayley 参加了 CAOPLE 的部分设计，Thomas Fantou 编写了 CAOPLE 的编译程序。由张一鸣对全书进行统稿。由于作者水平有限，再加上时间仓促，书中难免会出现不足之处，恳请读者批评指正。

作 者

2016 年 5 月

目 录

前言

第 1 章	引言	1
1.1	互联网发展进程中的三次跃升	2
1.1.1	第一次跃升 (TCP/IP)	2
1.1.2	第二次跃升 (WWW)	3
1.1.3	第三次跃升 (Cloud)	3
1.2	互联网资源的自然特性	4
1.2.1	成长性	4
1.2.2	自治性	4
1.2.3	多样性	5
1.3	资源虚拟化技术	5
1.4	互联网资源虚拟化的科学技术问题	7
1.4.1	开放环境下的按需聚合	7
1.4.2	分布自治资源的自主协同	8
1.4.3	聚合与协同的计算性质	8
1.5	本章小结	9
	参考文献	9
第 2 章	概念模型与体系结构	11
2.1	概念模型	11
2.1.1	设计原则	11
2.1.2	基本概念	12
2.1.3	概念间的关系	16
2.2	体系结构	17
2.2.1	资源层	17
2.2.2	资源虚拟层	17
2.2.3	聚合层	18
2.2.4	自主协同层	18
2.2.5	可信保障体系	18
2.2.6	编程开发环境	19
2.2.7	应用层	19

2.3 与其他互联网计算模型的联系	19
2.3.1 网格计算	19
2.3.2 服务计算	19
2.3.3 对等计算	20
2.3.4 云计算	20
2.4 本章小结	20
参考文献	21
第3章 资源虚拟化	23
3.1 计算虚拟化	23
3.1.1 基本概念	23
3.1.2 x86 虚拟化技术	25
3.1.3 虚拟机迁移技术	30
3.1.4 面向虚拟计算环境的应用	32
3.2 内存虚拟化	35
3.2.1 基本概念	36
3.2.2 内存超量提交技术	40
3.2.3 内存主动回收技术	44
3.2.4 面向内存虚拟化的服务器整合技术	47
3.3 网络虚拟化	55
3.3.1 单物理网络到多虚拟网络的映射技术	56
3.3.2 多物理网络到单虚拟网络的映射技术	60
3.3.3 多物理网络到多虚拟网络的映射技术	63
3.3.4 硬件实现的网卡虚拟化	63
3.4 本章小结	63
参考文献	63
第4章 虚拟资源的分布式管理	66
4.1 虚拟资源的分布式组织	66
4.1.1 覆盖网技术	66
4.1.2 相关研究	70
4.1.3 面向高效虚拟资源管理的通用覆盖网组织方法	77
4.1.4 基于 Kautz 图覆盖网的高效虚拟资源组织方法	83
4.2 虚拟资源的分布式搜索	88
4.2.1 资源搜索	88
4.2.2 相关研究	89
4.2.3 区间查询	91

4.2.4 Skyline 查询	97
4.2.5 聚合查询	99
4.3 面向资源组织优化的分布式网络延迟探测	99
4.3.1 网络邻近估计相关工作	100
4.3.2 网络邻近估计需求与挑战	103
4.3.3 基于同心环的多尺度邻近节点组织方法	104
4.3.4 分布式最近节点搜索	105
4.3.5 分布式 K 最近节点搜索	106
4.4 本章小结	107
参考文献	107
第 5 章 虚拟资源的协同机制	112
5.1 事件驱动的协同机制	112
5.1.1 基于事件服务的自主元素协同机制	112
5.1.2 相关研究	115
5.1.3 基于属性流行度的事件匹配方法	120
5.1.4 移动环境中高效订阅树重建方法	124
5.2 协同激励机制	129
5.2.1 挑战与问题	129
5.2.2 基本概念	131
5.2.3 相关研究	135
5.2.4 面向 iVCE 的信誉机制	139
5.3 本章小结	146
参考文献	146
第 6 章 程序设计语言	150
6.1 元模型	150
6.1.1 Agent	150
6.1.2 族	151
6.1.3 Agent 的环境	152
6.2 语言设计	153
6.2.1 程序的总体结构	153
6.2.2 环境描述	154
6.2.3 状态与动作说明	155
6.2.4 数据类型说明	155
6.2.5 语句与表达式	157
6.3 实现	159

6.3.1	概述	159
6.3.2	虚拟机 CAVM	159
6.3.3	CAOPLE 编译程序	164
6.4	相关工作	167
6.4.1	面向 Agent 的程序设计语言研究	167
6.4.2	面向 Agent 语言的实现	168
6.5	本章小结	170
	参考文献	170
第 7 章	未来发展	172
7.1	互联网应用发展迅速	172
7.1.1	互联网产业欣欣向荣	172
7.1.2	互联网影响更加广泛	172
7.1.3	互联网计算的规模迅速增长	173
7.1.4	互联网应用的负载变化剧烈	173
7.2	互联网计算面临新挑战	173
7.2.1	互联网新型应用带来的新问题	173
7.2.2	互联网计算面临发展新机遇	174
7.2.3	互联网计算面临的关键技术挑战	175
7.3	高效可信的虚拟计算环境	177
7.3.1	基本概念	178
7.3.2	多尺度资源建模	179
7.3.3	多尺度资源聚合	180
7.3.4	资源的弹性绑定	182
7.3.5	进展简介	183
7.4	本章小结	184
	参考文献	184

第 1 章 引 言

1983 年 1 月 1 日,美国国防部授权在 ARPANET 使用 TCP/IP,标志着互联网技术步入快速发展阶段。互联网功能上经历了全球通信网、全球数据网和全球资源网三次重大跃升,规模和应用的发展速度十分惊人。经历了 30 余年的发展演化,互联网汇聚了海量的计算资源、存储资源、数据资源和应用资源,已成为现代社会的重要信息基础设施。

据统计,2010 年运行在互联网上的服务器数量已超过 5000 万台^[1],且随着信息技术(Information Technology, IT)公司的成长和信息规模的膨胀而不断增长。2013 年亚马逊的服务器数量达到 15.8 万台,比 2009 年增长了 30 倍^[2]。根据国际数据公司(International Data Corporation, IDC)的预测,全球数据中心数量在 2017 年将达到 860 万^[3],每个数据中心的规模按照比较保守的 1000 台服务器来计算(中小型数据中心是指规模小于 3000 个标准机架的数据中心^[4]),届时全球数据中心的服务器总数将达到 86 亿台。

据统计,2010 年全世界一年内产生的数据量超过 1ZB,2014 年全世界一年内产生的数据量达到 7ZB,预计 2020 年全世界一年内产生的数据量将达到 35ZB^[5]。2015 年,Facebook 系统每天要处理 25 亿条消息,数据量超过 500TB,用户单击 Like 按钮的次数达到 27 亿次,上传 3 亿张照片,每半个小时扫描的数据大约为 105TB^[6]。Google 的 Gmail 邮箱的总存储容量理论上达到了 4EB 规模。微软的云存储服务 SkyDrive 的存储容量理论上已经达到 6.25EB 左右。

据统计,2015 年,Android 应用商店中的程序总数突破 140 万个,开发者数量达到 38.8 万^[7]。与此相比,2015 年苹果的 App Store 中的应用程序数量则达到 150 万个,其下载量累计超过 30 亿次,其中 iPhone 应用程序数量已经超过百万。2015 年,微信公众号的数量已经突破 1000 万,每天还在以 1.5 万的速度增长^[8]。

据统计,2014 年,Facebook 的活跃用户数量达到 22 亿,而 2013 年 Facebook 报告其每月活跃用户数为 11 亿人,也就是说,Facebook 公司的每月活跃用户总数在一年之内增加了一倍左右^[9]。腾讯公司 2011 年 1 月 21 日发布第一个微信版本,2013 年微信注册用户数量超过 3 亿,2015 年微信每月活跃用户已达到 5.49 亿^[10]。2013 年,微软的云存储服务 SkyDrive 活跃用户数量突破 2.5 亿^[11]。2015 年,Google 的 Gmail 邮箱用户数量达到 4.25 亿^[12]。2015 年,中国移动 3G 和 4G 用户总数超过 6.7 亿,其中 4G 用户达到 2.25 亿^[13]。

由此可见,互联网已经由一般意义下连接计算机的网络演变成为一个覆盖全球、

提供各类信息服务的计算平台。30 多年前,人们面对的计算环境是独立的计算机;如今,人们面对的计算环境是连接众多类型计算机(包括智能手机)的(移动)互联网。计算机产业得以迅速发展并进入寻常百姓家的一个重要技术基础就是计算机操作系统的成熟,其不仅使计算机硬件资源得到了高效管理和利用,而且极大地方便了普通用户便捷有效地使用计算机。当人们面对互联网的时候,自然期待在互联网上有这样一个类似计算机操作系统的“互联网操作系统”,将互联网管理成为一个统一的计算平台。可以说,这个愿望一直若隐若现地伴随着互联网的发展,每当互联网技术出现一次跃升时,这一愿望之火就会喷发一次。

1.1 互联网发展进程中的三次跃升

1.1.1 第一次跃升(TCP/IP)

互联网的第一次重大技术跃升是 20 世纪 80 年代 TCP/IP 的广泛采用, TCP/IP 被虚拟化为一种特殊的 I/O 模式植入 UNIX 操作系统,能方便实现各类网络的互连,使不同网络的计算机之间可以实现进程间通信,方便地发送邮件、传递文件。这次技术跃升使互联网真正成为全球通信网,20 世纪 80 年代, Sun 公司为提升 TCP/IP 和网络操作系统的重要性,提出了“网络就是计算机”^[14]的口号,“网络就是计算机”是指由许多计算机通过网络连接起来的系统,本身就是一台巨大的计算机,我们称其为“第一次互联网计算”的实践。这一实践的核心是试图在互联网上续写计算机操作系统在资源管理方面的成功,通过“分布式操作系统”管理整个网络资源,形成单一用户视图^[15]。但是,十年之后,这样的分布式操作系统并没有出现。

人们探索把计算机操作系统中经典的资源虚拟化思想和集中控制机制在网络计算时代延拓到企业计算(enterprise computing)^[16,17]中。企业计算是企业计算模型、资源、中间件和应用的统称。企业计算模型的典型代表是客户/服务器(Client/Server, C/S)模型及其变形,如三层 C/S 模型、浏览器/服务器(Browser/Server, B/S)模型等;企业计算资源是指隶属某组织机构的、可依据该组织机构控制力量实施统一管理的资源;企业计算中间件包括支持应用互操作和应用服务器构建的平台软件;企业计算应用是指支撑某机构业务、有相应的服务质量承诺的分布式应用,如金融服务、电信运营、政务管理、军事指挥等方面的信息系统。企业计算也可视为针对资源管理、简化流程等共性问题的业务软件解决方案。20 世纪 90 年代是企业计算大发展的时期^[18]。近年来,虚拟机^[19,20]技术再度兴起的现象表明,在企业计算领域,经典的资源虚拟化在网络化的服务器资源(如集群服务器)管理中仍然成效显著。人们在思考,经典的资源虚拟化技术在网络计算领域还能走多远,在互联网上是否能够构建跨越组织的资源共享应用。

1.1.2 第二次跃升（WWW）

互联网的第二次重大技术跃升是 20 世纪 90 年代环球信息网络（World Wide Web, WWW）技术的发明和广泛应用，使互联网真正成为全球数据网。WWW 成功地向世人展示，在互联网上实现跨越组织的资源共享不仅是可能的，而且其作用和影响是传统企业计算应用不可替代的。这一技术跃升和成功应用在一定意义上激发人们从 20 世纪 90 年代末开始了“Web 服务”^[21]和“网格计算”（grid computing）^[22, 23]的实践，我们称其为“第二次互联网计算”的实践。Web 服务是用以支持网络间不同机器的互动操作的软件系统，试图借鉴 WWW 的形式为用户提供透明、松耦合的第三方服务；网格计算通过汇聚分布的计算资源形成强大的计算能力，其目标既易于理解又十分诱人：使用网络上的计算资源犹如使用水电一样快捷简便。进入 21 世纪，这两方面的研究趋于融合，前者成为后者的技术途径。然而十年过去了，网格计算描绘的景象还没有出现。

两股不同的力量，朝着不同的方向努力，产生了不同的结果。一股是来自计算机领域的力量，按照计算机资源管理的经典准则，让网格计算落地，回归企业计算。例如，英国 e-Science^[24]和我国 CNGrid^[25]作为典型的网格计算应用，从技术角度看实际上是企业计算应用；美国国防部更直截了当地将其全球信息栅格（Global Information Grid, GIG）的平台定位为企业计算平台。另一股是来自互联网应用领域，遵循互联网的自然特性，构造新型的互联网应用，产生了 Web 2.0 和对等计算（Peer-to-Peer, P2P）等令人瞩目的成果^[26, 27]，Foster^[28]在他的著作中也提出网格计算的发展要借鉴 P2P 的思想。

1.1.3 第三次跃升（Cloud）

互联网的第三次重大技术跃升是 21 世纪云计算技术的出现和广泛应用，使互联网成为全球资源网。云计算^[29]是一种通过互联网以服务的方式提供动态可伸缩的虚拟化资源的计算模式，通过采用按使用量付费的模式，为用户提供按需便捷可配置的资源服务。云计算是传统分布计算、并行计算和网格计算的发展，它的出现向世人展示，可以通过互联网实现跨越地域按需可扩展的资源共享，它意味着计算能力、存储空间和各种软件服务可以作为一种商品通过互联网进行流通，有效实现了提高资源利用率和改善用户体验的双赢。

云计算既是互联网的一次技术跃升，更是互联网计算的一次重要实践，我们称其为“第三次互联网计算”的实践。云计算的核心是将计算资源、存储资源、网络资源以虚拟化和自适应的方式通过网络提交给用户使用，实现随用随付（pay-as-you-go）的资源使用方式。云计算从基础设施即服务（Infrastructure as a Service, IaaS）、平台即服务（Platform as a Service, PaaS）和软件即服务（Software as a Service, SaaS）三个层面上为用户提供按需自助的、可伸缩可度量的服务。

1.2 互联网资源的自然特性

实现面向互联网的网络计算所面临的技术挑战主要源自互联网资源的固有特征。与传统计算机环境中的资源相比较,互联网环境下的资源具有“成长性”“自治性”和“多样性”等自然特性^[30]。

1.2.1 成长性

成长性是指互联网资源规模不断膨胀、资源关联关系不断变化的动态特性。互联网是一个不断成长的开放系统,其覆盖地域不断扩大,大量的分布异构资源不断更新与扩展,呈非线性增长。

例如,美国 Lumeta 公司从 1998 年开始启动“互联网地图”项目^[31],旨在对互联网的增长情况进行长期研究。“互联网地图”的每一个节点所代表的可以是计算机群,也有可能是一个大型公司的网络。根据“互联网地图”数据,1998 年 8 月的全球互联网拓扑地图包含 88000 个骨干网路由器;2003 年 4 月的全球互联网拓扑地图包含 159000 个骨干网路由器,其规模是 1998 年的两倍;2008 年 4 月的全球互联网拓扑地图包含 450000 个骨干网路由器,其规模是 1998 年的 5 倍。

随着互联网应用的扩展与深入,资源在快速增长,资源间的关联关系也在不断演化,资源特征信息(即有关资源属性的描述信息)也随之相应地扩充和变化,从而导致互联网的资源视图难以确定。

1.2.2 自治性

自治性是指互联网资源的局部拥有、自主决策的特性。互联网上很多资源的拥有者是特定的组织和个人,资源拥有方可以根据资源需求方的要求提供资源,也可以拒绝提供,不同的拥有者可以有不同的服务策略。

例如,Google 拥有超过 500 个数据中心用于处理不同地域的不同任务,服务器数量超过 100 万台^[32]。其中,一部分服务器用于满足公司的正常运转,一部分服务器(如云计算平台 Azure 和云存储平台 Google Drive)以付费方式提供给外部用户共享使用。Google 拥有的丰富数据资源,既通过免费的网页和图片搜索引擎为普通用户提供服务,也提供了付费的网页和图片搜索服务以限制特定数据资源被访问的用户群体。同样,Google 的软件资源也支持免费和付费两种使用模式,例如,免费的个人邮箱和付费的企业邮箱、免费的较低清晰度的地图和付费的高清晰度的地图等。

显然,传统操作系统的全局化资源控制模式已不再适用,同样地,现有互联网资源的管理模式,也难以有效聚合和使用浩瀚的局部自治的资源,来共同完成大型应用任务。

1.2.3 多样性

多样性是指互联网资源属性存在广泛差异的特性。互联网上的资源多种多样,包括各类信息和数据、各种应用程序和服务等软资源,还包括各种计算设施、仪器等物理资源。

例如, Facebook 作为全球规模最大的社交网络,拥有着浩瀚的资源。在计算资源方面,其 3 个数据中心就拥有 18 万台服务器^[33];在存储资源方面,为管理每天 TB 级增长的数据,不仅拥有数以万计的 SATA (Serial Advanced Technology Attachment) 硬盘用于大容量存储,还拥有大量内存和高容量闪存设备,以提高 MySQL 数据库的 I/O 效率^[34];在数据资源方面,数据中心存放着海量的文本文档、照片、视频、应用程序,且大量是非结构化数据,用户平均每天上传的数据条目超过 120 亿^[9],每月上传照片的容量就高达 7PB;在软件资源方面,2011 年 9 月~2012 年 6 月 Facebook 发布了 4500 款应用,包括网络游戏、个性化搜索、文件共享、社交阅读、在线音乐播放、社交视频、地理位置信息服务、在线视频聊天等^[35]。腾讯作为中国规模最大的社交网络,同样拥有海量资源。在计算资源方面,其服务器总数已超过 30 万台^[36];在存储资源方面,为了应对每天 TB 级增长的数据,腾讯从 2009 年开始自研分布式数据仓库(Tencent Distributed Data Warehouse, TDW)以提升数据存储的可扩展性,仅腾讯云数据中心的服务器数目就达到 20 万台^[37],为每个用户免费提供 10TB 的存储空间;在数据资源方面,腾讯拥有的总数据量超过 100PB^[38],仅其旗下应用 QQ 空间上的照片就超过 1500 亿张^[39];在软件资源方面,截至 2015 年 8 月底,接入腾讯开放平台的应用数量超过 400 万^[40]。

资源不断地快速增长和属性的甚大差异,使得对资源统一建模和管理面临严峻的挑战。

1.3 资源虚拟化技术

虚拟化是指将计算机的各种实体资源予以抽象、转换后以逻辑的形式呈现出来,使用户以更好的方式来管理和应用这些资源。面对互联网资源的成长性、自治性、多样性等自然特性,虚拟化技术正成为人们探索有效管理互联网资源的重要途径。

虚拟化不是一个新概念。事实上,计算机操作系统就是虚拟化技术的代表性成果,它帮助用户摆脱了对复杂计算机实体资源管理的重负,使多个用户能同时和谐地共用一台计算机,有效提高了资源的利用率。

经典的虚拟化技术的重要成果是构建虚拟机。虚拟机是对真实计算环境的抽象和模拟,其技术发展已有半个多世纪的历史。早在 20 世纪 60 年代中,IBM 公司研制的 M44/44X 系统,就利用软件和硬件把一台 IBM 7044 系统(M44)模拟成多台 7044 虚拟机(称为“44X”)。虽然 M44/44X 系统只是用于研究,但它的开拓创新得到业界公

认,被认为是“虚拟机”的开山鼻祖。严格意义上,M44/44X 系统没有完整地模拟基础硬件,还不算完全的虚拟化。20 世纪 70 年代,第一个虚拟机操作系统 VM/370 问世,并随着 IBM System/370 真正投入应用^[41]。

20 世纪 80 年代后,价格低廉的服务器和台式机大量普及,虚拟化技术的发展曾一度放缓。进入 21 世纪,互联网得到广泛深入的应用,虚拟化技术在企业计算、灾难恢复、安全隔离等方面的优势,使之进入快速发展期,并成为学术界的研究热点。近年来,随着多核系统、集群、网格、云计算的广泛部署,虚拟化技术在商业应用上的优势也日益体现,既大幅降低了 IT 成本,又显著增强了系统的安全性和可靠性。据统计,VMware 虚拟机产品已经被 80%以上的全球百强企业所采纳^[42]。

根据向上层应用所提供的接口层次,可以把虚拟机分为四种类型:硬件抽象层虚拟机、操作系统层虚拟机、应用程序编程接口(Application Programming Interface, API)层虚拟机,以及编程语言层虚拟机,如图 1.1 所示。

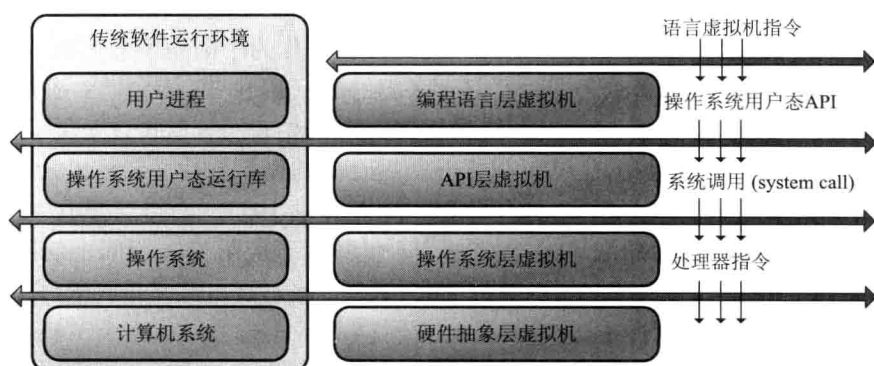


图 1.1 层次化的虚拟机分类

硬件抽象层虚拟机：对上层软件（即客户操作系统）而言，硬件抽象层虚拟机构造了一个完整的计算机硬件系统，与客户操作系统的接口为完整的处理器指令集。人们熟悉的 VMware 虚拟机，就属于此类型。

操作系统层虚拟机：通过动态复制操作系统运行环境，创建多个称为“虚拟容器”的虚拟机，对运行在每个容器之上的软件而言，这种虚拟机均提供了一个完整的操作系统软件环境，而它与上层软件（用户态应用程序或内核模块）的接口就是系统调用。Linux Container 就是比较有名的 Linux 操作系统虚拟机。

API 层虚拟机：这类虚拟机也为上层应用软件（用户态应用程序）提供了特定操作系统运行环境的模拟，但所提供的不是模拟处理器指令或系统调用，而是模拟实现了该操作系统的各类用户态 API。例如，在 Linux 操作系统上支持 Windows 应用程序的 Wine 和在 Windows 操作系统上支持 Linux 应用程序的 cygwin 等就属于此种类型。

编程语言层虚拟机：这类虚拟机利用解释或即时编译（Just-In-Time, JIT）技术来

运行语言虚拟机指令,从而实现软件的跨平台特性。这方面最典型的例子莫过于 Java 语言的虚拟机。

虚拟化技术能支撑计算任务更好地适应互联网资源的自然特性,有力推动了互联网的第三次跃升。例如,互联网资源的动态性给互联网计算带来了重大技术挑战,互联网环境下,计算任务和计算资源的绑定关系随时可能改变,必须实现在保持计算任务不变的情况下,动态切换其承载资源,虚拟机迁移技术就提供了一种很好的技术途径。虚拟机迁移是指在虚拟机不关机的情况下,把虚拟机的所有工作状态暂时冻结,从一台宿主机迁移到另一台宿主机上,并在新的宿主机上恢复虚拟机的工作状态。从原理上讲,实现虚拟机迁移的技术并不困难,难的是提高迁移效率。有两种迁移方式,一种是上面描述的“暂停运行、冻结状态、迁移、恢复执行”方式,即“冷迁移”,该方式暂停运行时间较长;另一种是虚拟机能够一边迁移,一边仍然保持工作,把停机时间缩短到几乎可以忽略不计的程度,即“热迁移”。显然“热迁移”技术的效率高,能更好地适应面向互联网计算的要求,并较为有效地屏蔽了互联网资源的动态性,但实现技术难度大。

把虚拟化技术有效应用到面向互联网的计算中,仍然存在不少关键问题要解决,既有科学层面的问题,更有技术层面的挑战。本章首先概述所面临的科学技术问题,后续章节将针对这些问题开展探讨,介绍相关研究成果。各章汇聚一起完成本书的主要宗旨,阐述面向互联网的虚拟计算环境(Internet-Based Virtual Computing Environment, iVCE)。

1.4 互联网资源虚拟化的科学技术问题

虚拟化技术为互联网资源管理应用开辟了一条途径,但仍面临诸多需要解决的科学技术问题。虚拟化要解决两个重要技术问题:第一是“抽象”,把基础性的实体资源“抽象”为用户易于理解和使用的模型;第二是“绑定”,使用时将抽象的模型高效绑定到实体资源上。传统的操作系统很好地解决了这两个问题,实施了对资源的有效管理。主要技术途径概括起来是两条:一是掌握系统资源全局信息;二是实施资源集中管理和控制。能这样做的基础前提主要有三点:资源范围确定、资源描述统一、资源可以集中管理和控制。

互联网资源所特有的动态不确定性、自治不可控性和异构多尺度性,使得在传统计算机系统中行之有效的“虚拟化”资源管理方法不再适用。为了面对互联网中动态资源的“多对多”使用问题,为了有效地组织、管理、利用跨越成千上万个组织的规模化资源,为了充分满足众多用户的各种各样的需求,必须研究互联网资源管理的新机理,研发新的虚拟化技术,有效解决开放环境下的按需聚合问题、分布自治资源的自主协同问题、聚合与协同的计算性质问题。

1.4.1 开放环境下的按需聚合

在开放的互联网环境下,如何根据任务需求,运用局部信息,实现资源特征信息

的汇聚、组织和综合利用，形成满足任务需求的相对稳定的资源视图，支持任务完成，是互联网计算面临的首个挑战性问题。

互联网是一个不断成长的开放系统，其覆盖地域不断扩大，大量分布异构资源不断动态地更新与扩展。随着互联网应用的扩展与深入，资源及其关联关系的动态演化，会导致相应的资源特征信息（即有关资源属性的描述信息）的扩充和变化。资源的多样性使得资源描述的复杂度加大，必须探索资源特征信息在语义层面上的识别和组织。资源的成长性和自治性使得资源的配置只能在变化中保持视图的动态稳定。总之，在这样开放、动态变化的网络环境中，不可能获得传统意义下全面、时空一致的资源特征信息。因此，按传统方式进行全局资源管理的可能性已不存在，寻找可行途径，针对应用需求，建立相对稳定的资源特征信息计算视图，是实现互联网计算必须要应对的挑战。

1.4.2 分布自治资源的自主协同

互联网资源的分布与自治特性使资源的管理和使用方式发生重大变化，对自治资源实施集中统一的管理方式行不通，协同就成为资源共享和合作工作的一种基本模式。因此，如何支持并实现自治资源的协同，建立可预测、可评估、可调节的协同工作机制和运行环境，达到资源的有效共享和综合利用，完成共同任务，是互联网计算面临的核心问题。

在开放环境下，资源的多样性可能导致不同自治资源在协同模式、协同环境、协同对象和协议等方面具有不确定性，多个自治资源在复杂的协同过程中局部目标与共同目标之间也可能不一致，这些都对自治资源自主协同的能力和计算环境的运行机制提出了更高的要求。实现开放动态环境下资源的灵活按需自主协同，形成互联网计算的核心运行机制，不仅要求资源具有更强的自主决策和适应能力，还必须从运行环境提供协议管理和演化等方面的支持。因此，研究和解决资源的自管理与动态决策、复杂条件下协同的约束满足、开放动态环境下协同模式的互操作、协同协议的管理与演化，以及自治资源的行为激励等问题，都是对传统资源管理理论与技术的挑战。

1.4.3 聚合与协同的计算性质

互联网计算的按需聚合和自主协同的机制复杂，如何建立聚合与协同计算性质的数学描述，相关度量、分析、评价和优化方法，以衡量聚合与协同方案的优劣及效用，指导聚合与协同机制的设计实现与使用，是互联网计算研究的基础性问题。

互联网的开放性，资源的自治性，资源聚合与协同环境存在的不完整性、不一致性和不确定性等问题，需要从理论上（如复杂性和质量）认识聚合与协同的静态性质与动态过程，包括大规模协同所呈现的复杂系统特性（如涌现行为）。传统的计算理论和方法已难以揭示和处理以交互为基本单元的计算的本质规律，也不足以对开放环境下的资源聚合与协同的计算性质进行定性和定量的研究，因此，无论从问题的答案还是从寻找答案的方法上都已经提出了基础性的挑战。