

软件定义数据中心

SOFTWARE DEFINED DATA CENTER **技术与实践**

TECHNOLOGY AND APPLICATION 陈熹 Ricky Sun 主编



- 国内首部系统介绍软件定义数据中心的专业书籍
- 众多业界专家倾力奉献，揭秘如何实现软件定义数据中心
- 理论与企业案例完美融合，呈现云计算时代的数据中心最佳解决方案

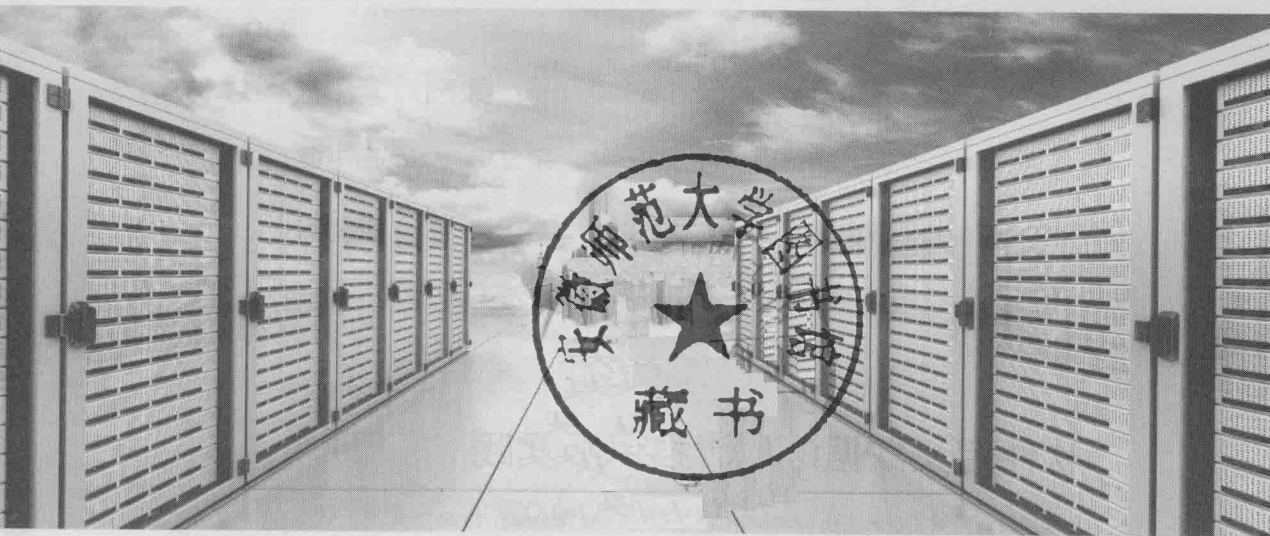


机械工业出版社
China Machine Press

软件定义数据中心

SOFTWARE DEFINED DATA CENTER 技术与实践

TECHNOLOGY AND APPLICATION 陈熹 Ricky Sun 主编



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

软件定义数据中心：技术与实践 / 陈熹等主编. —北京：机械工业出版社，2014.12

ISBN 978-7-111-48317-5

I. 软… II. 陈… III. 数据库系统 IV. TP311.13

中国版本图书馆 CIP 数据核字 (2014) 第 241862 号

本书从与软件定义数据中心有关的基本概念入手，通过实例介绍软件定义数据中心涉及的技术、应用、前景。在此基础上，深入介绍构建软件定义数据中心的计算、网络、存储、安全、自动化管理和高可用性等基本技术，并辅以解决方案和大型实例，力求使读者全面了解当前软件定义数据中心的技术动态和发展趋势，为实际构建软件定义数据中心提供必要的技术指导。

本书适作为数据中心分析、设计、研发、管理工程师的技术普及读物，亦可作为高等学校相关专业课程的教材或参考书。

软件定义数据中心：技术与实践

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：朱 劼 余 洁

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 1 月第 1 版第 1 次印刷

开 本：186mm×240mm 1/16

印 张：21.5（含 2 面彩插）

书 号：ISBN 978-7-111-48317-5

定 价：69.00 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) 88378991 88361066

投稿热线：(010) 88379604

购书热线：(010) 68326294 88379649 68995259

读者信箱：hzjsj@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

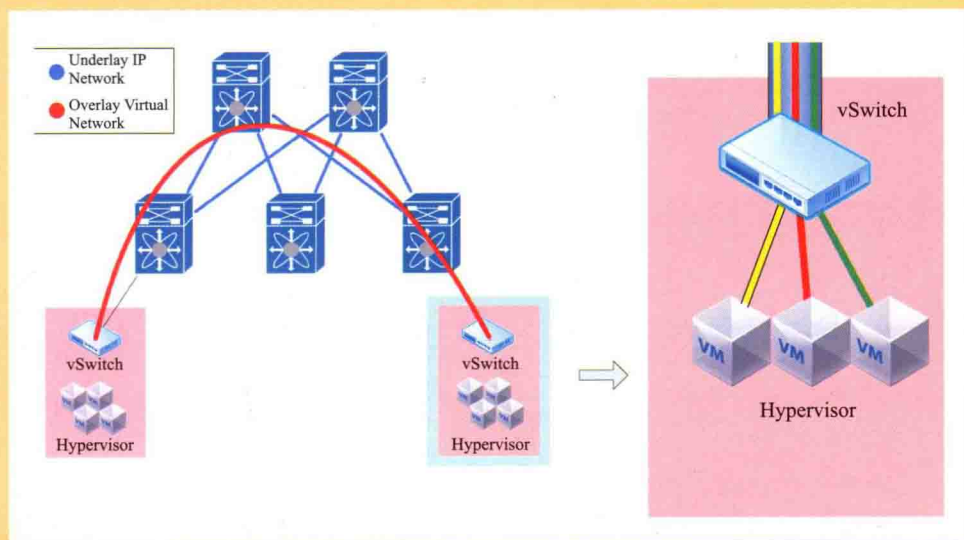


图 4-16 在共享物理网络的基础上实现隔离的逻辑网络

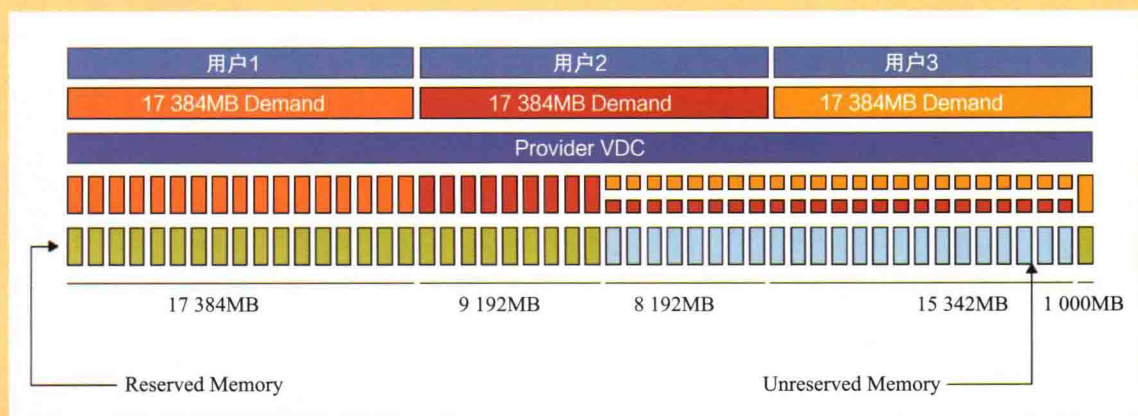


图 5-14 Provider VDC 内存分配

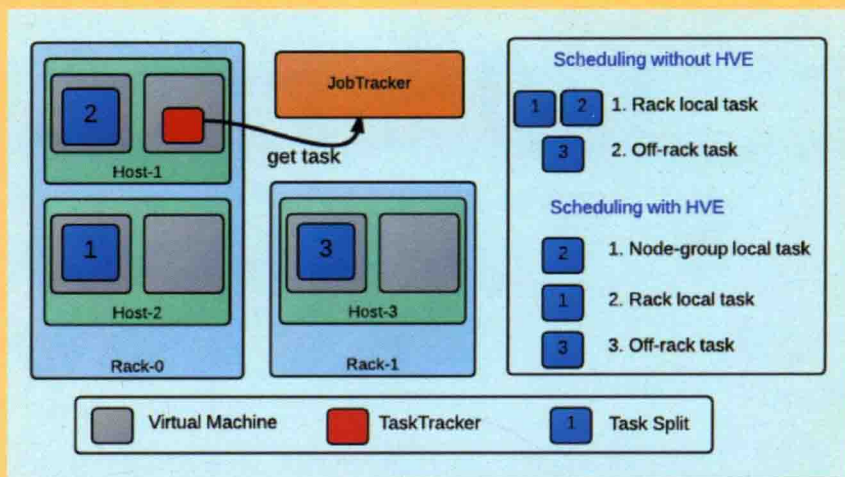


图 11-10 HVE 任务调度策略

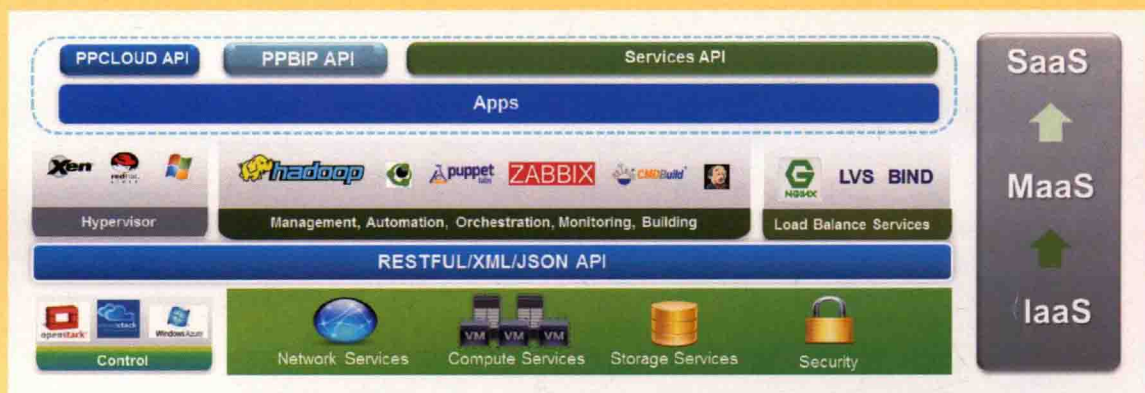


图 13-3 系统架构



图 13-7 基于云的应用交付流程

编委会

主编

陈 熹

EMC 中国研发集团高级经理

Ricky Sun (孙宇熙)

EMC 首席技术官办公室技术总监

编写组 (按姓氏拼音排序): EMC 中国研究院

曹 逾

陈 平

董 哲

范晨辉

郭小燕

李三平

刘 伟

陶 隽

王俊元

杨子夜

赵丽媛

周宝曜

特约编委

黄彦林

金 昀

陈文春

Ray Feng

Ariel Duan

联合策划

朱 捷 (EMC 中国研究院)

序

说到软件定义数据中心，我们要先从互联网说起。互联网在过去二十年里，更新了我们的沟通模式，加速了我们的信息获取，变革了我们的购物习惯，彻底地改变了我们的生活。不仅如此，互联网还颠覆了许多行业，而此颠覆仍是进行时。在企业级 IT 行业，云计算就是这场颠覆的名字。

由于互联网的远程商业服务模式与超大规模技术架构的推广与成熟，使得各行各业看到一个新的 IT 云服务模式和一个新的 IT 云基础架构。这个服务模式就是“as a Service”（即服务）的模式，包含了 IaaS、PaaS、SaaS，及总称 ITaaS。而这个新的基础架构就是这本书的主角——SDDC，即软件定义数据中心。

如果拜访任何一个成功的云服务商或任何一个大规模的互联网服务商的数据中心，我们会发现它们的创新与技术几乎完全是由软件来完成的。这些数据中心的硬件往往会非常统一，而不同的计算、存储、网络以及管理功能则由软件来实现。由于这个特性，使得它们的硬件使用率提高，伸缩规模相对简单，部署应用尤其迅捷。

而回头看现在的企业级数据中心，还不完全是这样。同时，大多数企业并不能摒弃自己的数据中心而完全依赖公有云的服务。但是，把云基础架构运用到企业级数据中心的时机已经成熟。企业可以在自己的数据中心搭起软件定义的私有云，并与公有云相通，形成混合云。

有了以软件定义数据中心为基础的混合云，企业就可以进退有度，游刃有余。加上成功管理新的移动终端技术，可轻松进入“云移动”时代！这也是为什么软件定义数据中心最近获得大家关注的根本原因。EMC 中国研究院编著的这本《软件定义数据中心：技术与实践》恰逢其时，它会向读者详细解说怎么实现软件定义数据中心。

从 2006 年开始，我与 EMC 中国研究院（ELC）的同事一起工作、合作，深深地被 ELC 的

研究员们的踏实、专业与聪颖所打动。这八年是云计算和软件定义数据中心从无到有、从稚嫩到完善的一个激动人心的过程，也是 EMC 中国研究院从无到有、不断发展、日益成熟的过程。

我相信对所有前瞻性的正在迈入云移动时代的企业 IT 部门来说，这是一本有“干货”的好书；对业界与学界研究云计算与数据中心架构的专家、学者、学生来说，也是一本有参考价值的好书。

希望你会喜欢它。

Charles Fan

VMware 高级副总裁，EMC 中国卓越研发集团创始人

前言

对于 IT 设备和服务厂商来说，“这是一个最好的时代，也是一个最坏的时代”。一方面，Amazon、Google 和层出不穷的初创公司能轻而易举地从资本市场拿到源源不断的投资，似乎根本不用担心成本压力和盈利预期；另一方面，IBM、HP 等传统 IT 巨头不得不焦头烂额地应对业绩的压力、对发展前景的质疑，希望通过转型继续在企业 IT 市场生存下来。外行人看来，还不是做着一样的生意吗？怎么前两年风生水起的大公司，这么快就裁的裁、撤的撤，纷纷转型自救了？俗话说得好：“形势比人强”。任由你是业界的“巨无霸”，也抵挡不住时代的大潮。这一波拍过来的浪潮无疑就是第三平台。在第三平台的大浪中，移动应用、社交应用、大数据应用是冒头的浪尖，而提供动力的是云计算。

作为第三平台的支柱，云计算吸引了 IT 厂商最多的资源。因为大家都对第二平台时代 Windows 操作系统 + Intel CPU（简称 Wintel）的联盟记忆犹新，谁占据了基础架构平台，谁就能主宰一个时代。在这个问题上，传统 IT 厂商和互联网背景的 IT 服务商是有重大分歧的。互联网公司没有老本可吃，也没有历史包袱，他们希望能让用户一步到位，直接上公有云；而传统 IT 厂商则多立足于现有的产品线，希望能让用户实现从 On-premise（IT 的本地运营）到 off-premise（异地运营，如 IDC）到私有云、混合云的过渡。然而无论是哪种云，都会碰到一系列共同的问题：硬件资源利用率、扩展性、自动化管理等。硬件的更新换代需要经年累月的时间，通常很难满足快速发展的业务需求，软件定义才是现实可行的出路。这也是为什么软件定义数据中心迅速成为 IT 产业的热门关键词的原因。

本书呈现了 EMC 中国研究院在这一领域多年的研究成果，并在此基础上总结梳理出软件定义数据中心的发展历程和未来方向。在内容组织上，全书分为四个部分：第一部分是总体介绍，试图回答一些关于软件定义数据中心的基本问题，告诉读者“什么是软件定义数据中心”、“为什么需要软件定义”；第二部分深入介绍了软件定义数据中心的关键技术，涵盖了计算、存储、网络、资源管理和调度、安全和高可用性；第三部分在了解了关键技术的基础上，向读者

展示了软件定义数据中心可以提供的一些应用场景；第四部分选取了两个软件定义数据中心的实例，一个是面向公有云的 AWS，另一个是面向流媒体的 PPTV，希望能让读者更有临场感，能看一看现在业界的“大拿们”是怎么“玩”的。

在本书的编写过程中，我们得到了 Pivotal 公司技术总监 Ray Feng 先生、我们的前 EMC 同事的大力支持，Ray 对第三平台的精辟论述也出现在了本书第 1 章。同时感谢 PPTV 的特约编委们、Ricky 的前微软同事们、以 PPTV 研发副总裁 Bill Huang 为领导的研发团队的同事们，他们提供了大量第一手资料，并共同编写了一章大型实例。此外，还要感谢 EMC 中国研发集团总经理 Wei Liu 和 EMC 高级总监 Xiaoye Jiang 在本书编写的过程中给予的大力支持。最后，我们衷心感谢在本书的撰写和出版过程中对我们给予巨大帮助的机械工业出版社华章公司的编辑们，没有她们的辛勤工作和耐心配合，这本书不会成为现实。

由于时间有限，本书的内容难免存在错漏之处，还请各位读者和专家不吝赐教。

陈熹、Ricky Sun (孙宇熙)

目 录

编委会
序
前 言

第一部分 总体介绍

第 1 章 基本概念	2
1.1 数据中心的历史	3
1.2 继续发展的推动力	6
1.3 软件定义的必要性	8
1.4 架构分析	13
1.4.1 基本功能模块	14
1.4.2 层次细分	14
1.4.3 接口与标准	17
1.5 现状与发展	18
1.6 第三平台: SDDC 上的 IT 新浪潮	20

第二部分 关键技术

第 2 章 软件定义的计算	26
2.1 虚拟化的定义与基本概念	26
2.1.1 虚拟化定义	26
2.1.2 虚拟化产生背景	26

2.1.3 计算虚拟化	27
2.1.4 VMM 的要求与基本特征	30
2.1.5 虚拟化平台的不同架构	30
2.2 虚拟化技术分类	31
2.2.1 x86 平台虚拟化面临的问题 与挑战	31
2.2.2 全虚拟化	32
2.2.3 半虚拟化	32
2.2.4 硬件虚拟化	33
2.2.5 小结	33
2.3 CPU 虚拟化	34
2.3.1 二进制翻译	34
2.3.2 硬件解决方案	35
2.4 内存虚拟化	36
2.4.1 软件解决方案	36
2.4.2 硬件解决方案	37
2.4.3 内存虚拟化管理面临的 挑战	38
2.5 I/O 虚拟化	38
2.5.1 背景介绍	38
2.5.2 基于软件的 I/O 虚拟化	39
2.5.3 基于硬件的 I/O 虚拟化	42
2.6 实例剖析	46
2.6.1 VMware ESX	46
2.6.2 Xen	53

第3章 软件定义存储	59	4.3.2 Neutron 在 OpenStack 中的 工作机制	134
3.1 新的存储管理模式	59	4.3.3 Nicira NVP 插件	134
3.1.1 传统存储面临的挑战	59	4.3.4 小结	137
3.1.2 新的管理模式：软件 定义存储	61	第5章 自动化资源管理	138
3.2 与存储虚拟化的比较	63	5.1 资源管理定义	138
3.3 架构、功能与特性	66	5.2 资源管理对象	140
3.3.1 数据模块与控制模块	67	5.3 资源管理策略	143
3.3.2 系统架构设计	68	5.3.1 资源管理一般性评价 指标和标准	143
3.4 解决方案：分类与比较	72	5.3.2 资源管理的主要策略	144
3.4.1 分类方法	72	5.4 多租户管理	147
3.4.2 现有产品简介	74	5.4.1 网络多租户管理	148
3.4.3 分类映射	77	5.4.2 计算多租户管理	149
3.5 市场现状与分析	78	5.4.3 存储多租户管理	149
3.5.1 技术影响	79	5.5 性能管理	150
3.5.2 软件定义存储的商业价值	80	5.6 对外服务接口	153
3.5.3 市场展望	81	5.7 资源管理典型实现	156
3.6 典型实现	81	第6章 流程控制	163
3.6.1 基于传统外置存储：ViPR	81	6.1 概述	163
3.6.2 基于服务器内置存储： ScaleIO	85	6.2 架构和功能	168
第4章 软件定义网络	91	6.2.1 数据中心 Orchestrator 的架构	168
4.1 概述	91	6.2.2 数据中心 Orchestrator 的功能	170
4.1.1 什么是 SDN	92	6.3 实现数据中心自动化	171
4.1.2 SDN 的架构和特征	94	6.3.1 数据中心的自动化势在 必行	171
4.1.3 SDN 相关组织介绍	97	6.3.2 自动化的好处	171
4.1.4 各大厂商对 SDN 的态度 和应用	101	6.3.3 自动化实施的对象	173
4.2 SDN 的技术实现	108	6.3.4 如何实现自动化	173
4.2.1 以网络为中心的实现	109	6.4 实例分析	175
4.2.2 以主机为中心的实现	120	6.4.1 VMware vCloud Orchestrator	175
4.3 SDN 的典型实现：OpenStack 中 的网络组件 Neutron	131	6.4.2 System Center Orchestrator	178
4.3.1 Neutron 在 OpenStack 中 的架构	132		

第7章 软件定义数据中心的安全

- 7.1 数据中心安全设计原则 184
- 7.2 物理基础设施的安全 186
- 7.3 软件定义层的安全 187
 - 7.3.1 安全的计算 187
 - 7.3.2 安全的存储 193
 - 7.3.3 安全的网络 197
- 7.4 软件资源协调层的安全 203
 - 7.4.1 统一的身份与访问授权管理 203
 - 7.4.2 安全技术的统一运用 205
- 7.5 小结 207

第8章 软件定义的高可用性

- 8.1 高可用性系统设计 209
 - 8.1.1 不可用的常见原因 209
 - 8.1.2 冗余的组件部署 210
 - 8.1.3 高可用性集群 210
 - 8.1.4 典型的冗余配置 211
- 8.2 软件定义之路——计算的高可用性 213
 - 8.2.1 高可用性对应用的需求 213
 - 8.2.2 高可用性集群——VMware HA 214
 - 8.2.3 零停机保障——VMware FT 218
- 8.3 软件定义之路——存储的高可用性 219
 - 8.3.1 基于 VPLEX 的高可用性 219
 - 8.3.2 ScaleIO 的高可用性 222
 - 8.3.3 ViPR 的控制器集群与 HA 数据服务 223
- 8.4 软件定义之路——网络的高可用性 224
 - 8.4.1 网络虚拟化 225
 - 8.4.2 逻辑端口镜像 226
 - 8.4.3 网络控制器集群 226

- 8.4.4 网关服务的高可用性 227

- 8.5 软件定义数据中心的高可用性 228
 - 8.5.1 整合的解决方案 229
 - 8.5.2 持续可用性 230
 - 8.5.3 分布式快速数据恢复 231
- 8.6 典型实现 232
 - 8.6.1 VMware SDDC 的高可用性 232
 - 8.6.2 OpenStack 的高可用性设计 234

第三部分 解决方案与应用

第9章 总体解决方案

- 9.1 SDDC 的基本要素 238
- 9.2 SDDC 实例：VMware 解决方案 239
 - 9.2.1 VMware SDDC 的计算 242
 - 9.2.2 VMware SDDC 的存储 243
 - 9.2.3 VMware SDDC 的网络 246
 - 9.2.4 VMware SDDC 的高可用性和容错 249
 - 9.2.5 VMware SDDC 的自动化 250
 - 9.2.6 VMware SDDC 的安全机制 251
 - 9.2.7 VMware SDDC 的管理 252
 - 9.2.8 VMware SDDC 实现小结 253
- 9.3 SDDC 实例：OpenStack 解决方案 254
 - 9.3.1 Horizon 控制面板 256
 - 9.3.2 Nova 计算组件 256
 - 9.3.3 Swift 对象存储 257
 - 9.3.4 Glance 镜像存储 257
 - 9.3.5 KeyStone 身份控制 257
 - 9.3.6 Quantum 网络 257
 - 9.3.7 Cinder 块存储 258
- 9.4 小结 258

第 10 章 云存储应用 259

10.1 云存储案例 259

10.2 云存储实现 260

10.2.1 可管理性 262

10.2.2 云存储系统的类型 262

10.2.3 访问方法 263

10.2.4 性能 263

10.2.5 多租户 264

10.2.6 可扩展性 264

10.2.7 可用性 264

10.2.8 可控性 265

10.2.9 效率 265

10.2.10 成本 266

10.3 云存储模式 266

10.3.1 公有云存储 266

10.3.2 私有云存储 267

10.3.3 混合云存储 267

10.3.4 三种云存储模式比较 267

10.4 主要云存储服务提供商 268

10.4.1 企业级云存储 268

10.4.2 个人云存储 269

第 11 章 虚拟化大数据平台 270

11.1 概述 270

11.2 VMware Serengeti 272

11.3 AWS EMR 280

11.4 小结 283

第四部分 大型实例分析

第 12 章 AWS 数据中心实例 286

12.1 AWS 概述 286

12.2 EC2 管理计算能力 287

12.2.1 EC2 概述 287

12.2.2 EC2 架构 287

12.2.3 EC2 存储 288

12.2.4 自动缩放 289

12.2.5 网络路由 289

12.2.6 EC2 实例 289

12.3 可扩展的存储 291

12.3.1 块存储 291

12.3.2 对象存储 292

12.3.3 冷数据归档 293

12.3.4 云存储网关 294

12.4 弹性十足的网络 295

12.4.1 亚马逊的 VPC 296

12.4.2 VPC 的特性 296

12.4.3 VPC 的应用场景 298

12.4.4 VPC 对 SDN 的践行 301

12.5 自动化的管理和部署 302

12.6 效益分析与未来发展 305

第 13 章 PPTV 基础平台管理

体系 306

13.1 系统概述 306

13.1.1 云部署模型 306

13.1.2 自建 IDC 部署概述 307

13.1.3 系统架构和组成 307

13.2 IaaS 部署和管理实践 308

13.2.1 基于 CloudStack 的 IaaS

管理平台 308

13.2.2 存储服务 310

13.2.3 基于 CloudStack 的私有云

平台最佳实践 310

13.3 MaaS 管理和基础服务体系 311

13.3.1 MaaS 管理架构 311

13.3.2 自动化基础设施管理

架构概述 312

13.3.3 开源工具链 312

参考文献 319

• SDDC 是虚拟化的基础设施。

• SDDC 部署如左图所示。

接下来,我们将学习什么是网络虚拟化,并将在 SDDC 出现之前,在物理网络中部署什么设备。

1.1 数据中心的概述

第一部分

总体介绍

■ 第1章 基本概念

软件定义数据中心 (Softwares Defined Data Center, SDDC) 是个新概念。新到什么程度呢? 2012 年以前还没有人系统阐述它。随着软件定义计算、软件定义存储、软件定义网络等一系列“软件定义”新技术的蓬勃发展, 已经有几十年发展历史的数据中心眼看着将要迎来另一场深刻的变革。原有的设备还可以继续运转, 但是管理员不再需要频繁出入轰鸣的机房去照看它们; 网络不需要重新连线也可以被划分成完全隔离的区域, 并且不用担心 IP 地址之间会发生冲突; 在数据中心部署负载均衡、备份恢复、数据库不再需要变动硬件, 也不再需要动辄几天的部署测试, 管理员只需点几下鼠标, 几秒钟就能完成; 资源是按需分配的, 再也不会有机房长年累月全速运转, 而没有人知道上面运行的是什么业务; 软件导致的系统崩溃几乎总是不可避免的, 但是在系统管理员甚至还没有发现这些问题的时候, 它们已经被自动修复了, 当然, 所有的过程都被记录了下来……

在机房里汗流浹背地摆弄过服务器的网线、光纤线、串口线和各种按钮的系统管理员看到这种情景会是什么心情? 回忆起往日给上百台服务器装系统、打补丁时手忙脚乱的画面, 如今都已经成了过眼云烟, 不免有些悲喜交加。不管是悲是喜, 这些事情都正在发生。也许你所接触到的一些计算环境已经开始大规模应用计算虚拟化, 但是还在使用传统的以太网和基于 IP 的网络划分; 也许有人已经将存储资源全部抽象成了块存储、文件存储和对象存储, 但是还需要大量的手工配置去设置一个备份服务……这不是一场风暴, 原有的技术和架构不会在一夜之间被摧毁; 这也不是海底火山喷发, 信息孤岛不会转眼间就消失。SDDC 所涉及的概念、技术、架构、规范都在迅速发展, 但又并不同步。我们要展示给大家的是一个日新月异的领域。要想用一两句话为 SDDC 下一个准确的定义本身就不够严谨。

要了解什么是 SDDC, 至少要回答以下几个基本的问题:

- SDDC 是在什么基础上发展而来的?
- 是什么驱动了 SDDC 的演化? (解决了什么问题?)