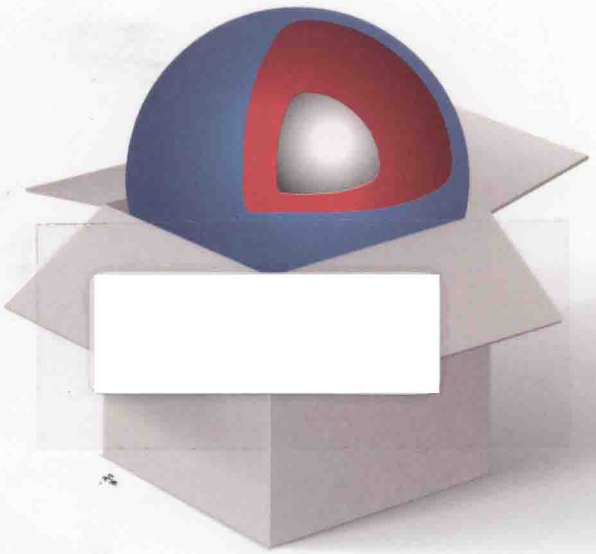
Broadview®
www.broadview.com.cn

CoreOS 实践之路

林帆◎著



中国工信出版集团
電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

CoreOS 实践之路

林帆◎著

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

内 容 简 介

本书是一本介绍 CoreOS 操作系统使用和周边技术的入门实践类书籍。本书内容分为三个主要部分。第一部分（第 1 章）主要介绍 CoreOS 的基本概念和系统的安装，为后续各个组件的使用做好铺垫工作；第二部分（第 2~6 章）主要介绍 CoreOS 中最核心的内置组件，通过这些组件，使用者能够完成大部分 CoreOS 的日常操作和开发任务；第三部分（第 7~9 章）主要针对 CoreOS 中一些比较进阶的话题以及组件进行更具体的讲解，并介绍一些 CoreOS 使用技巧。

在通读了这些内容后，相信读者会对 CoreOS 系统有一个比较全面的认识。

图书在版编目（CIP）数据

CoreOS 实践之路 / 林帆著. —北京：电子工业出版社，2015.12
ISBN 978-7-121-27509-8

Ⅰ. ①C… Ⅱ. ①林… Ⅲ. ①操作系统 Ⅳ. ①TP316

中国版本图书馆 CIP 数据核字（2015）第 263671 号

策划编辑：张春雨
责任编辑：葛 娜
印　　刷：三河市双峰印刷装订有限公司
装　　订：三河市双峰印刷装订有限公司
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：787×980　1/16　　印张：21　　字数：482千字
版　　次：2015年12月第1版
印　　次：2015年12月第1次印刷
印　　数：3000册　　定价：79.00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。

序

CoreOS 是一个非常年轻而且充满想象力的项目。 CoreOS 专注于集群操作系统的定制、更新以及维护等，实现了自动将集群内的操作系统、基础软件包持续更新至最新稳定版本，并缩小更新系统所需时间到分钟以内的级别。

自 2014 年登陆中国以来，CoreOS 在国内掀起了不少风波。CoreOS 的 AppC 项目制定 Application Container 标准部署规范，在容器部署规范上迈出了一大步； Etcd 逐渐也被大多数云产品引入作为在配置管理、服务发现方面取代 Zookeeper 的不二选择。总之，CoreOS 作为一个整体的集群操作系统，可以极大的解放 App 部署、弹性调度以及服务发现的复杂实施工作，开启了微服务新世界的大门，从研发流程和运维方式上面都产生了极大的改变。

但是国内一直缺少坚持在这方面潜心研究、并将其推广到生产环境的团队和气氛，一方面在于 CoreOS 本身的复杂性和超前的技术路线；另一方面，国内的 CaaS 云计算市场，特别是编排和部署更新方面，都处于刚兴起阶段。但是作者绝对算一个在 CoreOS 领域保持坚持不懈，努力钻研的勇士。认识作者是在 CoreOS 第一个国内 Meetup。当时本身也怀着对 CoreOS 的万分崇拜，跑了将近半个北京城去参加。作者当时大胆的提出了 CoreOS 不少的痛点，让我影响颇深，我也是从这个时候开始，不间断的跟作者保持在 CoreOS 方面技术的探讨和学习。随后作者发表了一系列关于 CoreOS 的文章：《漫步云端:CoreOS 实践指南》，是国内学习 CoreOS 的必备资料。

本书也是作者积累了 2 年多学习和实践经验的总结和沉淀。本书以 CoreOS 为基础，详细的介绍了 CoreOS 基础组件 Fleet、Systemd、Etcd、Flannel、Rkt、Cloudinit 等的配置以及对应的使用方式。理论离不开实践，特别是集群化的分布式系统，实践部分更加具有挑战性，因此作者在每一章都备有详细的实践，以引导各位读者更加深层次的掌握和领悟 CoreOS，然后通过一个综合案例，再次让大家得以融汇贯通。不仅如此，作者在最后详细介绍了 CoreOS 生态圈

的发展，重点介绍了开源版本的 Borg：Kubernetets，如同主菜之后的一道美味甜点，读了之后收益颇丰。

如果你是一个刚接触云计算，特别是 CaaS(Container as a Service)的朋友，这本书可以帮助你迅速掌握目前业界最火的几款系统的使用；

如果你是一个耕耘于云计算的朋友，你可以通过本书更深的了解 CoreOS 的前生，现在，以及未来，开阔自己的视野！

段兵　百度资深工程师

2015 年 11 月 2 日于北京

前　言

关于 CoreOS 系统

一直以来，服务器操作系统的升级都是运维人员感到棘手的事情。目前市面上的各种服务器操作系统普遍存在版本壁垒，无法保证安全的系统升级和回滚，这使得许多服务器不得不长时间运行在已经过时的内核和系统组件上，然后手工安装紧急的安全补丁或者索性完全不在意系统的安全问题。CoreOS 系统并不是第一个尝试解决这种现状的服务器系统，但它却是被最先设计出的能够安全可靠地用于生产环境中系统持续升级解决方案的操作系统。

出于这样的初衷，CoreOS 采用了基于双系统分区、容器技术和集群架构的设计思路，克服了由于用户修改系统内容、用户服务对系统组件依赖，以及系统重启时服务中断等种种导致升级过程不可靠的因素，最终以一种轻量级、平台定制化的操作系统呈现出来。它尽可能地适应各种不同的基础设施环境，使得系统具备十分便捷的集群组建能力，并鼓励用户通过容器技术隔离服务运行环境。

熟悉 CoreOS 系统的操作，除了理解它的只读系统分区和双系统分区等特殊性，更多的要求还在于熟悉 CoreOS 内置的容器和集群工具，例如 Docker、Rkt、Systemd、Fleet、Etcd、Locksmith，以及与 Confd、Flannel 和 Kubernetes 等非内置服务的集成使用。这些内容都会在本书的相应章节中逐一介绍。

本书的内容

本书是一本介绍 CoreOS 操作系统使用和周边技术的入门实践类书籍。本书内容分为三个主要部分。

第一部分，包括第 1 章的内容。主要介绍 CoreOS 的基本概念和系统的安装，为后续各个组件的使用做好铺垫工作。

第二部分，包括第 2~6 章的内容。主要介绍 CoreOS 中最核心的内置组件，通过这些组件，使用者能够完成大部分 CoreOS 的日常操作和开发任务。

第三部分，包括第 7~9 章的内容。主要针对 CoreOS 中一些比较进阶的话题以及组件进行更具体的讲解，并介绍一些 CoreOS 使用技巧。

在通读了这些内容后，相信读者会对 CoreOS 系统有一个比较全面的认识。

关于本书

本书的诞生源于我在 CSDN 发表的《CoreOS 实践指南》系列文章，在许多章节中都依然可以看见该系列文章的影子。但由于 CoreOS 周边的技术发展迅速，当时刊在网络上的许多内容都已经逐渐过时，本书针对这部分内容进行了修改，并扩充了大量在网络文章中由于篇幅原因没有详细介绍说明的技巧和信息，同时增加了如 Flannel、Kubernetes 等周边内容，从字数上看来，其容量大约是原系列文章的 5 倍。本书在编写过程中持续更新了系统的最新特性，直至定稿前 CoreOS 的 v835.0.0 版本和 Kubernetes 的 v1.0.6 版本，从新特性变化的频率来看，目前这些技术都已经进入相对稳定的阶段，其内容会在未来较长时间内适用。

由于作者时间与水平的局限，尽管在后期已经对书的内容进行过校检，但书中难免依然存在一些纰漏和错误。如果读者发现了问题，请发送至作者的邮箱：linfan.china@gmail.com。同时，在本书的网站 http://coreos.space 中也会及时发布相应的后续内容更新和勘误。

致谢

本书的出版，首先要感谢鼓励我发表《CoreOS 实践指南》系列文章的前 CSDN 编辑周小璐女士，她是我在技术写作路上的一位伯乐，没有她的帮助，我肯定无法完成这个原本只是写在自己默默无闻的博客上的技术分享。

其次，非常感谢电子工业出版社的张春雨先生。在本书写作的近 10 个月时间里，张先生的敦促和指导使得本书得以最终成型。特别是他耐心而友好地对待我一次次的拖稿，以及在我将稿件提交排版后，还继续对内容进行多次较大幅度更新和修改给予了支持。

最后，感谢我的父母，你们的关心和鼓励一直是我持续向前的动力；以及我的女友杨斌清，在我每天晚上挑灯写作时，包容和陪伴我，愿我们的辛勤付出能收获幸福的果实。我爱你们。

目　录

1 CoreOS 简介和安装

在互联网规模越来越庞大的今天，服务器集群对许多互联网软件公司都已经不再是一个陌生的词汇。与此同时，面对不可预期的业务扩展速度和层出不穷的网络安全漏洞，大型服务器集群的快速组建和安全管理也成为很多企业在进行基础设施架构时需要特别考虑的问题。

目前，网络上大多数服务器集群仍然使用着 Web 1.0 时代延续下来的传统服务器系统。这些操作系统虽然能够在集群中正常地支撑业务运作，但随着业务量和服务器节点数量的增加，许多弊端也逐渐凸显出来。例如，软件对特定基础服务的依赖，以及操作系统本身的升级路线导致系统版本无法持续更新，机房中分布着大量不同内核版本和类型操作系统的差异使得系统安全补丁难以高效地应用，特定主机安装了具有缺陷的软件版本而没有被及时发现等。

这些集群规模化造成的基础设施安全和效率问题给运维带来了新的挑战，单纯地通过制度、工具等外部手段来进行管控并不能够真正解决操作系统版本老化、应用补丁复杂混乱的状况。从根本上消除这些潜在的隐患，需要从最底层的操作系统角度进行改变。

CoreOS 是一个基于 Linux 内核的轻量级操作系统，为计算机集群的基础设施建设而生，专注于自动化、轻松部署、安全、可靠、规模化。作为一个操作系统，CoreOS 提供了在应用容器内部署应用所需要的基础功能环境，以及一系列用于服务发现和配置共享的内建工具。

在本章中，将从 CoreOS 系统的起源、组成和安装部署等方面介绍这个系统的独特之处，为之后章节的进一步深入做好基础准备工作。

1.1 CoreOS 简介

1.1.1 CoreOS 是什么

概括地说，CoreOS 是专为集群环境设计的轻量级 Linux 发行版。

作为新一代的服务器集群专用操作系统，CoreOS 提供了比传统服务器系统更快的启动速度、更小的系统资源消耗、更安全稳定的运行环境，以及更高效的集群组建体验。

CoreOS 的轻盈，得益于它的血统。CoreOS 是基于 Google 的 ChromeOS 系统二次定制的发行版，而 ChromeOS 本身就是一款内核充分优化、只需要几秒即可完成启动的轻量系统，两者的 Logo 如图 1-1 所示。站在巨人的肩膀上，自然能够飞得更高。在这样的基础上，CoreOS 进一步剔除了与图形界面相关的软件包、Perl 运行环境、Python 运行环境等非系统必需的组件。管理 CoreOS 系统的一般方法是通过 SSH Shell 登录或基于 Web 的工具（例如 cAdvisor）。对于有管理 Linux 服务器经验的运维人员来说，这些都是最常用的管理服务器系统的工具。

图 1-1 ChromeOS 系统和 CoreOS 系统的 Logo

1.1.2 CoreOS 的诞生和发展

如果要为美国当下 IT 界的传奇公司合写一本传记，就会发现其中的许多故事都开始于某个不起眼的车库。亚马逊、谷歌、Youtube 和苹果公司等皆不例外。

2013 年 3 月，CoreOS 的三位创始人 Alex Polvi、Michael Marineau 和 Brandon Philips，在美国加利福尼亚州帕洛阿尔托（Palo Alto）市的一个车库里，开始了他们的改变服务端领域的梦想：让遍布在全球的服务器能够像现代的浏览器软件一样自动和及时地获得系统的安全补丁和更新。

这是项目发起者 Alex 的第二次创业，第一次创业他将自己的产品卖给了云计算领域的巨头 RackSpace 公司[1]。CoreOS 公司成立后，打出了产品的第一个旗号——“以全新的方式思考服务器（a new way to think about servers）”。由 Alex 出任公司 CEO，而 Brandon 出任 CTO。

2013 年 4 月，首个 CoreOS 内部测试版本发布，只提供了 RackSpace 云的安装脚本。这个

[1] http://www.csdn.net/article/2013-08-22/2816672-coreos-the-new-linux

版本涉及了操作系统后台自动更新功能的雏形，也就是后来为人们所知的双系统分区更新方案。值得注意的是，CoreOS 采用了首先发布云端平台的系统版本，并且在之后很长一段时间里，通过云端平台运行 CoreOS 一直是优先推荐的部署方案。

2013 年 5 月，CoreOS 启动了两个十分重要的子项目：Etcd 和 Fleet，并分别于同年 6 月和 10 月发布到 Github。这两个项目实现了当系统自动升级时，对外服务不中断的能力，也标志着 CoreOS 真正成为了“集群化的操作系统”。

2013 年 8 月，经过几个月的紧张开发，此时的 CoreOS 已经在 Linux 技术社区里有了一些前期用户和测试反馈。同时，CoreOS 新增了用于 OpenStack、VMware 和 KVM 的官方镜像，并开始以 Vagrant Box 的方式发布测试系统。这一阶段在研发方向的探索和实际进展上都有不少突破，系统基本已经成型，CoreOS 用户群持续增长。值得指出的是，这一阶段的新增用户有很大一部分来自于 Docker 的爱好者。

2014 年 5 月，第一个正式的 Beta 版本上线，CoreOS 提供了一个新的服务：Locksmith，用来解决系统更新时由于过多节点同时重启导致的集群不稳定问题。

2014 年 7 月，CoreOS 迎来一个重要的里程碑时刻，第一个稳定版本发布。这时 CoreOS 已经形成了完整的 Alpha、Beta、Stable 三个官方升级通道的更新模式。

2014 年 8 月，CoreOS 第一次并购，收购了企业级 Docker 镜像服务提供商 Quay.io，开发规模进一步扩大。

2014 年 9 月，美国的著名云服务提供商 DigitalOcean 与 CoreOS 达成战略合作，将 CoreOS 提升为继 Ubuntu、CentOS、Fedora 和 Debian 后，该平台的第 5 个默认支持系统。许多互联网企业用户开始接纳 CoreOS 作为其首选的操作系统发行版。

2014 年 12 月，CoreOS 由于与 Docker 对于容器的发展方向见解不合，自立门户建立了 AppC 容器标准和 Rkt 容器项目。这个项目为后来的容器统一标准 OCI 的诞生奠定了基础。同月，国内的第一次 CoreOS Meetup 活动在北京 ThoughtWorks 办公室举办。

2015 年 4 月，CoreOS 接受来自 Google 的 1200 万美元投资，成为 Google Kubernetes 项目的战略合作平台，并共同推出了 CoreOS 的企业版产品 Tectonic。

2015 年 5 月，第一届全球 CoreOS 开发者大会 CoreOS Fest 在美国硅谷所在地旧金山市举办，这次大会向人们展现了 CoreOS 已经颇具规模的国际社区。

2015 年 8 月，CoreOS 开始内置集成 Kubelet 组件，这是 CoreOS 与 Kubernetes 生态圈融合的重要一步，明确了两者相互促进、共同发展的前景。

至今，CoreOS 已经发展成为了包括多地区的独立组织（CoreOS Meetup）和全球规模的技

术大会（CoreOS Fest）活动组成的技术社区型组织，促进了应用容器化和服务器集群领域的技术发展。

1.1.3 CoreOS 的用户体验

1.1.3.1 来自现代浏览器的启发

与传统的服务器操作系统集大成的设计理念不同的是，CoreOS 采用了高度精简的系统内核及外围定制，剔除了对于服务端非必需的功能，甚至没有提供 GUI 界面和包管理器。然而，CoreOS 内置了许多额外的服务组件，这些组件主要包括三类：应用容器和辅助服务（Docker/Rkt/Flannel/Kubelet）、云部署和分布式调度服务（Cloudinit/Ignition/Etcd/Fleet）、系统安全和认证增强服务（SELinux/Dex），它们构成了 CoreOS 分布式服务集群的关键元素。

CoreOS 的核心思想来自于现代浏览器（特别是 Chrome 浏览器）的用户体验：快速启动、后台更新、跨版本无缝升级、每个 Tab 页采用独立沙盒、单个 Tab 页崩溃能快速修复，以及整个浏览器也不会因为单个沙盒进程的崩溃而崩溃。引申到服务器上，试想将一个应用托管在应用容器中的服务从一台服务器转移到另一台服务器上，就像用鼠标将 Tab 页从一个浏览器拖拽到另一个浏览器界面上那样简单。而这些，正是 CoreOS 希望带给每一个用户的体验。

1.1.3.2 更快的启动速度

因为轻，所以快。CoreOS 团队对操作系统内核和外围做了大量的精简，不仅减少了硬件资源的占用，更获得了极大的启动速度提升。

根据官方宣传，其系统运行时内存使用量只有 114MB（过去的主页内容，现在已经看不到了）。在 Vagrant 环境下实测目前的最新版本（v801.0.0），启动后加上文件和磁盘缓存的总内存使用大约为 115MB，扣除文件和磁盘缓存部分后的净内存使用大约为 32MB，均远远低于其他主流服务器系统。对比数据如表 1-1 所示。

表 1-1 几种主流服务器操作系统启动后的内存和磁盘占用对比

系统	净内存消耗（扣除 Cache/Buffer）	磁盘占用
CoreOS 801.0	32MB	416MB
CentOS 7.0	111MB	930MB
Ubuntu 14.04	112MB	1.1GB
Debian 8.1	42MB	1.1GB