Docker in Production

Docker生产环境

实践指南

[美] Joe Johnston [西] Antoni Batchelli
[英] Justin Cormack [美] John Fiedler 著
[英] Milos Gajdos

DockOne.io Community of Container 吴佳兴 梁晓勇 译

中国工信出版集团 人民邮电出版社 POSTS & TELECOM PRESS

Docker in Production

Docker生产环境

实践指南

[美] Joe Johnston [西] Antoni Batchelli
[英] Justin Cormack [美] John Fiedler [英] Milos Gajdos 著

DockOne.io Community of Container 吴佳兴 梁晓勇 译

人民邮电出版社
北京

图书在版编目（CIP）数据

Docker生产环境实践指南 / (美) 约翰斯顿(Johnston, J.) 等著 ; 吴佳兴, 梁晓勇译. -- 北京 : 人民邮电出版社, 2016.7
书名原文: Docker in Production
ISBN 978-7-115-42225-5

Ⅰ. ①D… Ⅱ. ①约… ②吴… ③梁… Ⅲ. ①Linux操作系统－程序设计－指南 Ⅳ. ①TP316.89-62

中国版本图书馆CIP数据核字(2016)第098850号

◆ 著　[美] Joe Johnston　[西] Antoni Batchelli
[英] Justin Cormack　[美] John Fiedler
[英] Milos Gajdos
译　吴佳兴　梁晓勇
责任编辑　杨海玲
责任印制　焦志炜

◆ 人民邮电出版社出版发行　北京市丰台区成寿寺路 11 号
邮编　100164　电子邮件　315@ptpress.com.cn
网址　http://www.ptpress.com.cn
北京天宇星印刷厂印刷

◆ 开本：800×1000　1/16
印张：13.5
字数：274 千字　2016 年 7 月第 1 版
印数：1－3 000 册　2016 年 7 月北京第 1 次印刷
著作权合同登记号　图字：01-2015-6664 号

定价：49.00 元

读者服务热线：(010)81055410　印装质量热线：(010)81055316
反盗版热线：(010)81055315

内容提要

本书围绕“Docker 该如何应用到生产环境”这一核心问题展开。在本书中，读者将接触到多个 IT 企业应用 Docker 到生产环境的成功案例，了解 Docker 实际投产时将会面临的问题，以及它与现有基础设施存在的矛盾与冲突，了解构建 Docker 生态系统所需的配套设施，包括安全、构建镜像、持续集成/持续交付、镜像存储、配置管理、网络实现、服务发现、持久化存储以及日志监控等模块的具体选型方案及利弊所在。本书编写时一些案例参考的 Docker 版本是 Docker 1.6 或 Docker 1.7。

本书要求读者具备一定的容器管理和运维的基础知识，适合在生产环境中使用 Docker 的相关技术人员阅读，尤其适合具有中高级 DevOps 和运维背景的读者阅读。

对本书的赞誉

经过 2015 年 Docker 项目的飞速发展，国内的互联网企业开始关注容器技术在生产环境下的使用案例。但是，由于与 Docker 相关的技术资料还没有得到系统性地整理，国内企业一直在各种实际场景中进行实践，所以本书的出现正好填补了当前生产环境实践的实际需要，让国内容器技术的推广迈出坚实的一步。

——肖德时，数人科技 CTO

2015 年，不管是传统 IT 企业、互联网巨头，还是初创公司，大大小小的公司都在实践如何在生产环境中应用 Docker 来解决其实际问题。在这个过程中会遇到各种各样的问题，如网络和存储驱动如何选型，资源限制不够彻底怎么办，镜像仓库如何做权限控制，容器内健康状况如何监测，容器有哪些安全隐患，以及在容器集群化过程中会遇到哪些问题，如服务发现、弹性伸缩等。这些问题都没有标准答案，这也激发了工程师们的探索乐趣，于是就有了各种各样的实战方案。

本书恰好就是对一些实战方案的归纳和总结。相对于市面上其他 Docker 图书，我认为这本书价值更大：对工程师们来说，将技术应用到“生产环境”中才是最大的价值和乐趣所在！

——陈轶飞，原百度 PaaS 平台负责人，国内最早大规模应用 Docker 的实践者

我在 Google 从事基于容器的基础设施和集群管理研发多年，许多关于容器使用最佳实践的知识都是通过“代代相传、口口相传”的方式获得的。在 Docker 迅速流行的同时，在开源社区里却缺少如 Google（或其他公司）内部“老工程师对新人倾囊相授实践+真理”的这种奢侈。

在众多讲述 Docker 自身原理、使用方法的书中，这本书从生产角度出发，将作者在实战中积累的一线经验系统地汇总成了基于 Docker 搭建生产系统的经验，值得我们借鉴。

——张鑫，才云科技 CEO

在深入学习 Docker 时将面对这样的问题：Docker 的流行催生出大量与之相关的新技术，

这些新技术有哪些，它们是什么关系，又该如何正确选择？Docker 也有适用场景，到底哪些场景适合用？把 Docker 用在生产环境，是否有成功案例或最佳实践，又该如何操作？如果你正面对这些问题，本书就非常适合你。另外，本书不仅以可以直接应用的真实生产环境示例贯穿始终，还讲解了技术的权衡之道，是 Docker 进阶的难得好书。

——刘凡，好雨云创始人

Docker 进入大家视野已经有段时间了，也历经了多个重要的版本升级。有报告表明，在一线互联网公司中，已经有 27%的企业在生产环节中使用到了 Docker，先行者已开始从中获益。但目前 Docker 资料仍停留在概念或者实验级别，关于真正在生产环境中使用 Docker 的内容则少之又少，最佳实践方面的信息与资料也十分稀缺。

Docker 技术对于 IT 架构乃至软件架构的改变是巨大的，对于想在生产环境中使用 Docker 的企业和团队来讲，只掌握概念和基本原理是不够的，如何能使用 Docker 解决自身问题，获得由此带来的收益，需要更多的生产实践方面的内容。

本书以生产部署为背景，讲述 Docker 在真实环境中的使用，能够给读者一个很好的参考，进而达到让读者“举一反三”的效果，使其能让自身的 IT 架构提升到一个新的技术高度。

——周东波，首都在线总工程师

Docker 的出现不仅是单一技术的诞生，而是整个开发、测试、运维体系的一次变革，是从传统的底层主机化真正向资源服务化转变。Docker 同时改变了广大技术人员的思维模式，让原先枯燥重复的工作变得有趣。

本书非常系统地讲解了 Docker 在生产环境的搭建和应用，同时详细地讲解了其镜像、网络、存储、安全等方面的知识，是有志于研究 Docker 的读者必备的书籍。

——孙斌，去哪儿网运维总监

Docker 是互联网领域最近两年最火热的技术词语，没有之一。以 Docker 为代表的容器技术，通过务实的工程创新，正在影响无数企业的 IT 基础设施以及主流的公有云服务。国内越来越多的互联网企业开始将 Docker 应用到生产环境中。可以预见，掌握 Docker 相关知识即将成为软件工程师或运维工程师的必备技能。DockOne.io 社区组织翻译的这本书恰逢其时，值得每名 IT 从业者阅读。

——刘海锋，京东云平台总架构师

容器技术已经是2015年最热的技术。容器技术已经存在数年，为什么Docker作为容器技术的布道者能够有翻天覆地的影响？为什么容器技术能够给传统应用带来好处？如何把传统应用容器化并通过DevOps方式简洁化？如何通过具体参数和方法解决现实问题？如何使用Docker以及Docker实践？

其实还有很多问题需要解决，如应用的复杂性和难维护性、开发与运维的脱离性，以及成本不断攀升的现实性。本书从简入繁，通过具体案例解释具体问题，通过实践帮助读者理解具体问题，对正在或者即将使用基于容器的DevOps的读者有很多益处。

——陈冉，Linker Networks CTO & VP

Docker的出现使得IaaS和PaaS的界限进一步模糊化，引领了云计算技术变革。时下诸多公司正如火如荼地探索在生产环境中如何玩转 Docker。本书的特点在于不仅介绍“是什么”，更进一步探讨“如何用”和“为什么用”；不限于某种特定的框架，而是对比各种不同的方案。书中涵盖容器监控、配置管理、安全、调度、持续交付等生产实践经验，想在生产环境中玩转Docker的技术人员定能从中获益。

——吴毅挺，携程CIS-系统研发部总监

译者介绍

吴佳兴，毕业于华东理工大学计算机系，目前是携程网系统研发团队的一名 DevOps 工程师，主要研究方向有 Python 开发、运维自动化、配置管理及 PaaS 平台的构建等。2014 年年底有幸加入 DockOne 社区，作为译者，利用闲暇时间为社区贡献一些微薄的力量。个人博客 devopstarter.info。欢迎邮件联系（wjx_colstu@hotmail.com）。

梁晓勇，毕业于厦门大学，现任某互联网金融公司架构师，DockOne 社区编外人员，非著名互联网从业者。长期奋战在技术研发第一线，在网络管理、技术开发、架构设计等方面略有心得。热爱互联网技术，积极投身开源社区，对 Docker 等容器技术具有浓厚兴趣。欢迎邮件联系（sunlxy@yahoo.com）。

前言

Docker 是基础设施的新成员。很少有新兴技术能像它这样，在 DevOps 和基础设施领域中快速风靡起来。在不到两年的时间内，Google、亚马逊、微软、IBM 以及几乎所有云供应商都宣布支持运行 Docker 容器。大量与 Docker 相关的创业公司在 2014 年和 2015 年年初都获得了风险资本的投资。Docker 开源技术背后的同名公司——Docker 公司，在 2015 年第一季度的 D 轮融资中估值为 10 亿美元左右。

大大小小的公司都在转换其应用，使之运行于容器内，以此实现面向服务架构（SOA）和微服务。不论是参加从旧金山到柏林的任何 DevOps 聚会，还是阅读最热门的公司工程博客，都可以看出全世界的运维领导者们如今都在云上运行 Docker。

毫无疑问，容器已经成为应用程序打包和基础设施自动化的重要组成部分。但有一个棘手的问题，促使本书作者和同僚们创作了另一本 Docker 图书。

本书面向的读者

具有中高级 DevOps 和运维背景的读者将从本书获益最多。因而，强烈建议读者应具备在生产环境中运行服务器以及创建和管理容器这两方面的基本经验。

很多图书和博客文章已经涵盖了与 Docker 安装及运行相关的话题，但能把在生产环境中运行 Docker 时产生的大量甚至是令人挠头的关注点结合在一起的材料则少之又少。不用担心，如果你很喜欢《盗梦空间》（Inception）这部电影，在云服务器的虚拟机中运行容器会让你感觉很自然。

本书将带读者深入理解生产环境中架构的组成部分、关注点，以及如何运行基于 Docker 的基础设施。

谁真的在生产环境中使用 Docker

换个更深刻的说法，对于在真实生产环境中使用 Docker 遇到的问题，如何找到解决之

道？本书综合了访谈、真实公司端到端的生产环境实例，以及来自 DevOps 杰出专家的参考文献，以此来解答这些问题。虽然本书包含了一些有用的示例，但它并不是一本复制粘贴的“教程式”参考书。相反，本书侧重于生产环境中对前沿技术进行评估、风险抵御及运维所需的实践理论和经验。

作为作者，我们希望这本书所包含的内容能够为那些正在评估如何及何时将 Docker 相关技术引入其 DevOps 栈的团队提供一个可靠的决策指南，这远比代码片段要来得长久。

生产环境中运行的 Docker 为企业提供了多个新的运行和管理服务器端软件的方式。很多现成的用例讲解了如何使用 Docker，但很少有公司公开分享过他们的全栈生产环境经验。本书汇集了作者在生产环境中运行 Docker 的多个实例和一组选定的友好公司分享的使用经验。

为什么使用 Docker

Docker 所使用的底层容器技术已经存在了很多年，甚至早于 dotCloud[①]这家平台即服务（PaaS）创业公司，即后来我们所熟知的 Docker。在 dotCloud 之前，许多知名的公司（如 Heroku[②]和 Iron.io[③]）已经在生产环境中运行大型容器集群，以获取额外的超越虚拟机的性能优势。与虚拟机相比，在容器中运行软件赋予了这些公司秒级而非分钟级的实例启动与停止的能力，同时能使用更少的机器运行更多实例。

既然这项技术并不新鲜，为什么 Docker 能获得如此巨大的成功呢？主要是因为它的易用性。Docker 创造了一种统一的方式，通过简便的命令行及 HTTP API 工具来打包、运行和维护容器。这种简化降低了将应用程序及其运行时环境打包成一个自包含镜像的入门门槛，使之变得可行且有趣，而不需要类似 Chef、Puppet 及 Capistrano 之类的配置管理和发布系统。

Docker 提供了一种统一手段，将应用程序及其运行时环境打包到一个简单的 Dockerfile 里，这从根本上改变了开发人员与 DevOps 团队之间的交互界面。从而极大简化了开发团队与 DevOps 之间的沟通需求与责任边界。

在 Docker 出现之前，各个公司的开发与运维团队之间经常会爆发史诗般的战争。开发

① https://www.dotcloud.com/
② https://www.heroku.com/
③ http://www.iron.io/

团队想要快速前进，整合最新版的软件及依赖，以及持续部署。运维团队则以保证稳定为己任，他们负责把关可以运行于生产环境中的内容。如果运维团队对新的依赖或需求感到不适，他们通常会站在保守的立场上，要求开发人员使用旧版软件以确保糟糕的代码不会搞垮整台服务器。

Docker 一下子改变了 DevOps 的决策思维，从“基本上说不”变成了“好的，只要运行在 Docker 中就可以”，因为糟糕的代码只会让容器崩溃，而不会影响到同一服务器上的其他服务。在这种泛型中，DevOps 有效地负责为开发人员提供 PaaS，而开发人员负责保证其代码能正常运行。如今，很多团队将开发人员加入到 PagerDuty 中，以监控他们在生产环境中的代码，让 DevOps 和运维人员专注于平台的稳定运行及安全。

开发环境与生产环境

对大多数团队而言，采用 Docker 是受开发人员更快的迭代和发布周期需求推动的。这对于开发环境是非常有益的，但对于生产环境，在单台宿主机上运行多个 Docker 容器可能会导致安全漏洞，这一点我们将在第 6 章“安全”中讲述。事实上，几乎所有关于在生产环境中运行 Docker 的话题都是围绕着将开发环境与生产环境区分开的两个关注点进行的：一是编配，二是安全。

有些团队试图让开发环境和生产环境尽可能保持一致。这种方法看起来很好，但是限于开发环境这样做所需定制工具的数量又或者说模拟云服务（如 AWS）的复杂度，这种方法并不实际。

为了简化这本书的范畴，我们将介绍一些部署代码的用例，但判定最佳开发环境设置的实践机会将留给读者。作为基本原则之一，尽量保持生产环境和开发环境的相似性，并使用一个持续集成/持续交付（CI/CD）系统以获取最佳结果。

我们所说的“生产环境”

对于不同的团队，生产环境意味着不同的东西。在本书中，我们所说的生产环境是指真实客户用于运行代码的环境。这是相对于开发环境、预演环境及测试环境而言的，后者的停机时间不会被客户感知到。

在生产环境中，Docker 有时是用于接收公共网络流量的容器，有时则是用于处理来自队列负荷的异步的后台作业。不管哪种用途，在生产环境中运行 Docker 与在其他环境中运行相比，最主要的差异都是需要在其安全性与稳定性上投入较多的注意力。

编写本书的动力之一是，与 Docker 相关的文档和博客文章中缺乏对实际生产环境与其他环境的明确区分。我们认为，80%的 Docker 博客文章中的建议在尝试在生产环境中运行 6 个月之后会被放弃（或至少修改）。为什么？因为大多数博客文章中举的都是理想化的例子，使用了最新、最好用的工具，一旦某个极端的情况变成了致命缺陷，这些工具将被遗弃（或延期），被更简单的方法所取代。这是 Docker 技术生态系统现状的一个反映，而非技术博客的缺陷。

总的来说，生产环境很难管理。Docker 简化了从开发到生产的工作流程，但同时增加了安全和编配的复杂度（更多关于编配的内容参见第 4 章）。

为了节省时间，下面给出本书的重点综述。

所有在生产环境中运行 Docker 的团队，都会在传统的安全最佳实践上做出一项或多项妥协。如果无法完全信任容器内运行的代码，那么就只得选用容器与虚拟机一对一的拓扑方式。对于很多团队而言，在生产环境中运行 Docker 的优势远远大于其带来的安全与编配问题。如果遇到工具方面的问题，请等待一到两个月，以便 Docker 社区对其进行修复，不要浪费时间去修补其他人的工具。保持 Docker 设置最小化。让一切自动化。最后，对成熟的编配工具（如 Mesos、Kubernates 等）的需求远比想象的要少得多。

功能内置与组合工具

Docker 社区一个常见的口头禅是“电池内置但可移除”，指的是将很多功能捆绑在一起的单体二进制文件，这有别于传统 Unix 哲学下相对较小、功能单一、管道化的二进制文件。

这种单体式的做法是由两个主要因素决定的：（1）使 Docker 易于开箱即用；（2）Golang 缺少动态链接。Docker 及多数相关工具都是用 Google 的 Go 编程语言[①]编写的，该语言可以简化高并发代码的编写与部署。虽然 Go 是一门出色的编程语言，但用它来构建的 Docker

① https://golang.org/

生态系统中也因此迟迟无法实现一个可插拔的架构，在这种架构中可以很容易用替代品对工具进行更换。

如果读者有 Unix 系统背景，最好是编译自己的精简版 Docker 守护进程，以符合生产环境的需求。如果读者有开发背景，预计到 2015 年下半年，Docker 插件将成为现实[①]。在此期间，估计 Docker 生态系统中的工具将会出现明显的重叠现象，某些情况下甚至是相互排斥的。

换句话说，要让 Docker 运行于生产环境中，用户的一半工作将是决定哪些工具对自己的技术栈最有意义。与 DevOps 所有事情一样，先从最简单的解决方案入手，然后在必要时增加其复杂性。

2015 年 5 月，Docker 公司发布了 Compose[②]、Machine[③]及 Swarm[④]，与 Docker 生态系统内的同类工具进行竞争。所有这些工具都是可选的，请根据实际情况对其进行评估，而不要认为 Docker 公司提供的工具就一定是最佳解决方案。

探索 Docker 生态系统时的另一项关键建议是：评估每个开源工具的资金来源及其商业目标。目前，Docker 公司和 CoreOS 经常发布工具，以争夺关注度和市场份额。一个新工具发布后，最好等上几个月，看看社区的反应，不要因为它看起来很酷就切换到最新、最好用的工具上。

哪些东西不要 Docker 化

最后一个关键点是，不要期望能在 Docker 容器中运行所有东西。Heroku 风格的“十二要素”（12 factor[⑤]）应用是最容易 Docker 化的，因为它们不维护状态。在理想的微服务环境中，容器能在几毫秒内启动、停止而不影响集群的健康或应用程序的状态。

类似 ClusterHQ[⑥]这样的创业公司正着手实现 Docker 化数据库和有状态的应用程序，但眼下，由于编配和性能方面的原因，可能需要继续直接在虚拟机或裸机上运行数据库。

Docker 还不适用于任何需要动态调整 CPU 和内存要求的应用[⑦]。允许动态调整的代码

① Docker 1.7 版中正式引入了插件系统。——译者注
② https://docs.docker.com/compose/
③ https://docs.docker.com/machine/
④ https://docs.docker.com/swarm/
⑤ http://12factor.net/
⑥ https://clusterhq.com/
① Docker 1.10 版中新增的 docker update 命令可实现 CPU 和内存的动态调整。——译者注

已经完成，但尚不清楚何时才能在一般的生产环境中投入使用。目前，若对容器的 CPU 和内存的限制进行调整，需要停止并重新启动容器。

另外，对网络吞吐量有高要求的应用进行最佳优化时不要使用 Docker，因为 Docker 使用 iptables 来完成宿主机 IP 到容器 IP 的 NAT 转换。通过禁用 Docker 的 NAT 来提升网络性能是可行的，但这是一个高级的使用场景，很少有团队会在生产环境中这么做。

技术审稿人

衷心感谢以下技术审稿人提供的早期反馈及细致的评论：Mika Turunen、Xavier Bruhiere 和 Felix Rabe。

作者团队介绍

作为作者，我们的首要目标是尽可能方便地组织和传播我们的知识，使之对社区产生价值。容器和 Docker 基础设施场景发展如此之快，以致几乎没有时间去完成传统纸质书[①]。

本书是由一个 5 人团队在几个月时间内撰写完成的，这 5 位作者在生产环境基础设施和 DevOps 方面都具有丰富的经验。书中的内容与时俱进，但我们还是特别谨慎，确保概念能经得起时间的考验。

Joe Johnston 是一名全栈开发人员、企业家及服务于旧金山创业公司的顾问。他是 Airstack（一家微服务基础设施创业公司）、California Labs 和 Connect.Me 的联合创始人。@joejohnston[②]

Antoni Batchelli 是 PeerSpace 公司[③]的工程副总裁和 PalletOps 公司[④]（一家基础设施自动化咨询公司）的联合创始人。他的主要工作是将函数式编程语言与基础设施结合，以及帮助工程团队打造杰出的软件。@tbatchelli[⑤]

Justin Cormack 是一名顾问，他对开源软件、云计算及分布式系统方面的创新机会尤其感兴趣。他目前就职于 unikernels。读者可以在 GitHub[⑥]上找到他。@justincormack[⑦]

---

① 本书英文版仅以电子书形式发布。——译者注
② https://twitter.com/joejohnston
③ https://www.peerspace.com/
④ http://palletops.com/
⑤ https://twitter.com/tbatchelli
⑥ https://github.com/justincormack
⑦ https://twitter.com/justincormack

John Fiedler 是 RelateIQ 公司的工程运营总监。他的团队专注于基于 Docker 的解决方案，为其 SaaS 基础设施及 DevOps 提供动力。@johnfielder[①]

Milos Gajdos 不仅是一名独立咨询师，还是 Infrahackers 有限公司的基础设施的负责人。他协助企业更好地理解 Linux 容器技术，并实现基于容器的基础设施。他偶尔会撰写一些有关容器[②]的博客文章。@milosgajdos[③]

① https://twitter.com/johnfielder
② http://containerops.org/
③ https://twitter.com/milosgajdos

目录