

# Kubernetes

## 权威指南



龚正 吴治辉 王伟 等编著  
崔秀龙 闫健勇

从Docker到Kubernetes  
实践全接触



下一代分布式架构的王者

Kubernetes: The Definitive Guide



中国工信出版集团



电子工业出版社  
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY  
<http://www.phei.com.cn>

## 作者介绍



### 龚正 / HPE高级顾问

拥有十多年的IT从业经验，具备丰富的云计算、大数据分析和大型企业级应用的架构设计和实施经验，是电信、金融、互联网等领域的资深专家。



### 吴治辉 / HPE资深架构师

拥有超过15年的软件研发经验，专注于电信软件和云计算方面的软件研发，拥有丰富的大型项目架构设计经验，是业界少有的具备很强Coding能力的S级资深架构师，也是《ZeroC Ice权威指南》《架构解密：从分布式到微服务》的作者。



### 王伟 / HPE资深系统架构师、大数据和云计算技术专家

拥有多年IT行业从业经验，参与过多个大型应用的架构设计、系统开发和实施落地，精通大数据、云计算及大型系统架构和开发的相关技术，对互联网和电信行业的热点技术有着深刻的理解，是云计算和大数据方面的技术专家。



### 崔秀龙 / HPE资深架构师

开源软件、自动化爱好者，拥有十多年从业经验，对软件生命周期的各个环节均有深刻的理解。



### 闫健勇 / HPE高级项目经理、总架构师

拥有超过15年的电信行业系统建设经验，主导了多项电信大型系统的架构设计和管理，对于云计算和大数据在电信行业中的应用拥有丰富的经验。



### 崔晓宁 / HPE高级顾问

拥有超过7年的测试咨询和质量管理经验，在云计算、大数据和分布式运算架构下的业务质量控制方面有非常丰富的项目实践和心得，并对推动组织架构优化有丰富的经验。帮助多个超过百人的大型项目建立软件产品管理规范 and 体系，并对其运营提供指导。



### 刘晓红 / HPE高级咨询顾问

拥有超过10年的电信行业从业经验，亲历中国移动BSS/OSS领域核心系统的建设发展历程，具备丰富的咨询规划、需求分析、产品设计、项目管理、测试管理经验，专注于云计算、大数据等前沿技术的研究。

**纪念并不是为了结束，而是为了新的写作思路的展开**

近年来，Kubernetes的版本和社区发生了翻天覆地的变化

而本书的作者们紧紧抓住了这些变化中的机遇

适时出版了

《Kubernetes权威指南：从Docker到Kubernetes实践全接触》的第1版、第2版和纪念版

这3本书均记录了Kubernetes发展历程中各里程碑版本的核心思想

之后

你会发现作者们的新思路和新视角

这些将为你开启全新的Kubernetes实践之旅

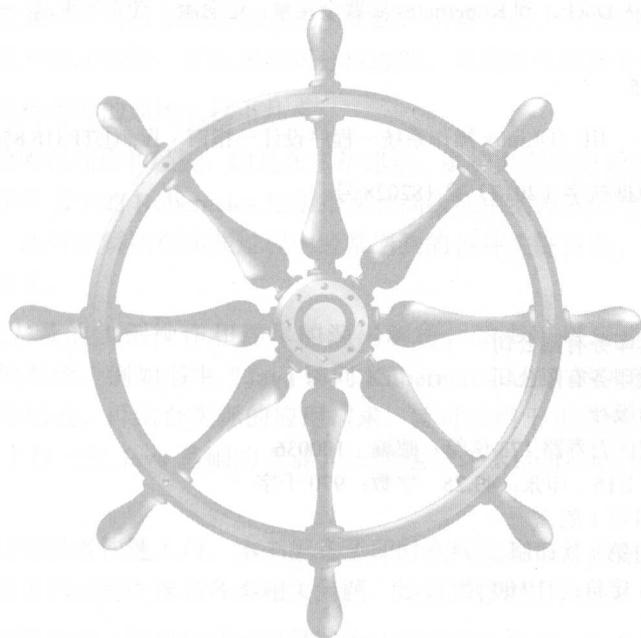
# Kubernetes

## 权威指南



从Docker到Kubernetes  
实践全接触

龚正 吴治辉 王伟 / 等编著  
崔秀龙 闫健勇



电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

## 内 容 简 介

Kubernetes 是由谷歌开源的 Docker 容器集群管理系统，为容器化的应用提供了资源调度、部署运行、服务发现、扩容及缩容等一整套功能。《Kubernetes 权威指南：从 Docker 到 Kubernetes 实践全接触（纪念版）》从架构师、开发人员和运维人员的角度，阐述了 Kubernetes 的基本概念、实践指南、核心原理、开发指导、运维指南及源码分析等内容，图文并茂、内容丰富、由浅入深、讲解全面；围绕着生产环境中可能出现的问题，给出了大量的典型案例，比如安全配置、网络方案、共享存储方案、高可用性方案及 Trouble Shooting 技巧等，有很强的实战指导意义。本书随着 Kubernetes 版本更新不断完善，目前涵盖了 Kubernetes 从 v1.0 到 v1.6 版本的全部特性，尽力为 Kubernetes 用户提供全方位的指南。

无论是对于软件工程师、测试工程师、运维工程师、软件架构师、技术经理，还是对于资深 IT 人士来说，本书都极具参考价值。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有，侵权必究。

### 图书在版编目（CIP）数据

Kubernetes 权威指南：从 Docker 到 Kubernetes 实践全接触：纪念版 / 龚正等编著. —北京：电子工业出版社，2017.9

ISBN 978-7-121-32351-5

I. ①K… II. ①龚… III. ①Linux 操作系统—程序设计—指南 IV. ①TP316.85-62

中国版本图书馆 CIP 数据核字（2017）第 182028 号

策划编辑：张国霞

责任编辑：徐津平

印 刷：三河市良远印务有限公司

装 订：三河市良远印务有限公司

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：43.25 字数：970 千字

版 次：2017 年 9 月第 1 版

印 次：2017 年 9 月第 1 次印刷

印 数：2500 册 定价：119.00 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）88254888，88258888。

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。

本书咨询联系方式：010-51260888-819，faq@phei.com.cn。

# 推荐序

经过作者们多年的实践经验积累及近一年的精心准备，本书终于与我们大家见面了。我有幸作为首批读者，提前见证和学习了在云时代引领业界技术方向的 Kubernetes 和 Docker 的最新动态。

从内容上讲，本书从一个开发者的角度去理解、分析和解决问题：从基础入门到架构原理，从运行机制到开发源码，再从系统运维到应用实践，讲解全面。本书图文并茂，内容丰富，由浅入深，对基本原理阐述清晰，对程序源码分析透彻，对实践经验体会深刻。

我认为本书值得推荐的原因有以下几点。

首先，作者的所有观点和经验，均是在多年建设、维护大型应用系统的过程中积累形成的。例如，读者通过学习书中的 Kubernetes 运维指南和高级应用实践案例章节的内容，不仅可以直接提高开发技能，还可以解决在实践中经常遇到的各种关键问题。书中的这些内容具有很高的借鉴和推广意义。

其次，通过大量的实例操作和详尽的源码解析，本书可以帮助读者进一步深刻理解 Kubernetes 的各种概念。例如书中“Java 访问 Kubernetes API”的几种方法，读者参照其中的案例，只要稍做修改，再结合实际的应用需求，就可以用于正在开发的项目中，达到事半功倍的效果，对于有一定 Java 基础的专业人士快速学习 Kubernetes 的各种细节和实践操作十分有利。

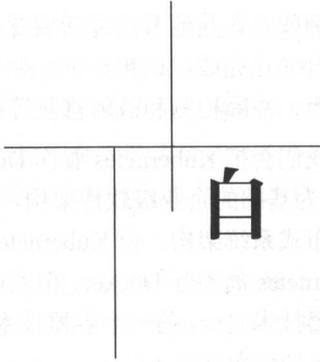
再次，为了让初学者快速入门，本书配备了即时在线交流工具和专业后台技术支持团队。如果你在开发和应用的过程中遇到各类相关问题，均可直接联系该团队的开发支持专家。

最后，我们可以看到，容器化技术已经成为计算模型演化的一个开端，Kubernetes 作为谷歌开源的 Docker 容器集群管理技术，在这场新的技术革命中扮演着重要的角色。Kubernetes 正在被众多知名企业所采用，例如 RedHat、VMware、CoreOS 及腾讯等，因此，Kubernetes 站在了容器新技术变革的浪潮之巅，将具有不可预估的发展前景和商业价值。

如果你是初级程序员，那么你有必要好好学习本书；如果你正在 IT 领域进行高级进阶修炼，那你也有必要阅读本书。无论是架构师、开发者、运维人员，还是对容器技术比较好奇的读者，本书都是一本不可多得的带你从入门向高级进阶的精品书，值得大家选择！

初瑞

中国移动业务支撑中心高级经理



# 自序

我不知道你是如何获得这本书的，可能是在百度头条、网络广告、朋友圈中听说本书后购买的，也可能是某一天逛书店时，这本书恰好神奇地出现在你面前的书架上，让你想起一千多年前那个意外得到《太公兵法》的传奇少年，你觉得这是冥冥之中上天的恩赐，于是果断带走。不管怎样，我相信多年以后，这本书仍然值得你回忆。

Kubernetes 这个名字起源于古希腊，是舵手的意思，所以它的 Logo 既像一张渔网，又像一个罗盘。谷歌采用这个名字的一层深意就是：既然 Docker 把自己定位为驮着集装箱在大海上自在遨游的鲸鱼，那么谷歌就要以 Kubernetes 掌舵大航海时代的话语权，“捕获”和“指引”这条鲸鱼按照“主人”设定的路线巡游，确保谷歌倾力打造的新一代容器世界的宏伟蓝图顺利实现。

虽然 Kubernetes 自诞生至今才 1 年多，其第一个正式版本 Kubernetes 1.0 于 2015 年 7 月才发布，完全是个新生事物，但其影响力巨大，已经吸引了包括 IBM、惠普、微软、红帽、Intel、VMware、CoreOS、Docker、Mesosphere、Mirantis 等在内的众多业界巨头纷纷加入。红帽这个软件虚拟化领域的领导者之一，在容器技术方面已经完全“跟从”谷歌了，不仅把自家的第三代 OpenShift 产品的架构底层换成了 Docker+Kubernetes，还直接在其新一代容器操作系统 Atomic 内原生集成了 Kubernetes。

Kubernetes 是第一个将“一切以服务（Service）为中心，一切围绕服务运转”作为指导思想创新型产品，它的功能和架构设计自始至终地遵循了这一指导思想，构建在 Kubernetes 上的系统不仅可以独立运行在物理机、虚拟机集群或者企业私有云上，也可以被托管在公有云中。Kubernetes 方案的另一个亮点是自动化，在 Kubernetes 的解决方案中，一个服务可以自我扩展、自我诊断，并且容易升级，在收到服务扩容的请求后，Kubernetes 会触发调度流程，最终在选定的目标节点上启动相应数量的服务实例副本，这些副本在启动成功后会自动加入负载均衡器中并生效，整个过程无须额外的人工操作。另外，Kubernetes 会定时巡查每个服务的所有实例的可用性，确保服务实例的数量始终保持为预期的数量，当它发现某个实例不可用时，会自动

重启该实例或者在其他节点重新调度、运行一个新实例，这样，一个复杂的过程无须人工干预即可全部自动化完成。试想一下，如果一个包括几十个节点且运行着几万个容器的复杂系统，其负载均衡、故障检测和故障修复等都需要人工介入进行处理，那将是多么的难以想象。

通常我们会把 Kubernetes 看作 Docker 的上层架构，就好像 Java 与 J2EE 的关系一样：J2EE 是以 Java 为基础的企业级软件架构，而 Kubernetes 则以 Docker 为基础打造了一个云计算时代的全新分布式系统架构。但 Kubernetes 与 Docker 之间还存在着更为复杂的关系，从表面上看，似乎 Kubernetes 离不开 Docker，但实际上在 Kubernetes 的架构里，Docker 只是其目前支持的两种底层容器技术之一，另一个容器技术则是 Rocket，后者来源于 CoreOS 这个 Docker 昔日的“恋人”所推出的竞争产品。

Kubernetes 同时支持这两种互相竞争的容器技术，这是有深刻的历史原因的。快速发展的 Docker 打败了谷歌曾经名噪一时的开源容器技术 lxc，并迅速风靡世界。但是，作为一个已经对全球 IT 公司产生重要影响的技术，Docker 背后的容器标准的制定注定不可能被任何一个公司私有控制，于是就有了后来引发危机的 CoreOS 与 Docker 分手事件，其导火索是 CoreOS 撇开了 Docker，推出了与 Docker 相对抗的开源容器项目——Rocket，并动员一些知名 IT 公司成立委员会来试图主导容器技术的标准化，该分手事件愈演愈烈，最终导致 CoreOS “傍上”谷歌一起宣布“叛逃”Docker 阵营，共同发起了基于 CoreOS+Rocket+Kubernetes 的新项目 Tectonic。这让当时的 Docker 阵营和 Docker 粉丝们无比担心 Docker 的命运，不管最终鹿死谁手，容器技术分裂态势的加剧对所有牵涉其中的人来说都没有好处，于是 Linux 基金会出面调和矛盾，双方都退让一步，最终的结果是 Linux 基金会于 2015 年 6 月宣布成立开放容器技术项目（Open Container Project），谷歌、CoreOS 及 Docker 都加入了 OCP 项目。但通过查看 OCP 项目的成员名单，你会发现 Docker 在这个名单中只能算一个小角色了。OCP 的成立最终结束了这场让无数人揪心的“战争”，Docker 公司被迫放弃了自己的独家控制权。作为回报，Docker 的容器格式被 OCP 采纳为新标准的基础，并且由 Docker 负责起草 OCP 草案规范的初稿文档，当然这个“标准起草者”的角色也不是那么容易担当的，Docker 要提交自己的容器执行引擎的源码作为 OCP 项目的启动资源。

事到如今，我们再来回顾当初 CoreOS 与谷歌的叛逃事件，从表面上看，谷歌貌似是被诱拐“出柜”的，但局里人都明白，谷歌才是这一系列事件背后的主谋，其不仅为当年失败的 lxc 报了一箭之仇，还重新掌控了容器技术的未来。容器标准之战大捷之后，谷歌进一步扩大了联盟并提高了自身影响力。2015 年 7 月，谷歌正式宣布加入 OpenStack 阵营，其目标是确保 Linux 容器及关联的容器管理技术 Kubernetes 能够被 OpenStack 生态圈所容纳，并且成为 OpenStack 平台上与 KVM 虚拟机一样的一等公民。谷歌加入 OpenStack 意味着对数据中心控制平面的争夺已经结束，以容器为代表的形态与以虚拟化为代表的系统形态将会完美融合于 OpenStack 之上，并与软件定义网络和软件定义存储一起统治下一代数据中心。

谷歌凭借着几十年大规模容器使用的丰富经验，步步为营，先是祭出 Kubernetes 这个神器，然后又掌控了容器技术的制定标准，最后又入驻 OpenStack 阵营全力将 Kubernetes 扶上位，谷歌这个 IT 界的领导者和创新者再次王者归来。我们都明白，在 IT 世界里只有那些被大公司掌控和推广的，同时被业界众多巨头都认可和支撑的新技术才能生存和壮大下去。Kubernetes 就是当今 IT 界里符合要求且为数不多的热门技术之一，它的影响力可能长达十年，所以，我们每个 IT 人都有理由重视这门新技术。

谁能比别人领先一步掌握新技术，谁就在竞争中赢得了先机。惠普中国电信解决方案领域的资深专家团一起分工协作，并行研究，废寝忘食地合力撰写，完成了这部近 700 页的 Kubernetes 权威指南。经过两年的高速发展，Kubernetes 先后发布了 v1.0~v1.6 这 6 个大版本，每个版本都带来了大量的新特性，能够处理的应用场景也越来越丰富。本书遵循从入门到精通的学习路线，全书共分为六大章节，涵盖了入门、实践指南、架构原理、开发指南、高级案例、运维指南和源码分析等内容，内容详实、图文并茂，几乎囊括了 Kubernetes 到 v1.6 版本的方方面面，无论是对于软件工程师、测试工程师、运维工程师、软件架构师、技术经理，还是对于资深 IT 人士来说，本书都极具参考价值。

吴治辉

惠普公司系统架构师

轻松注册成为博文视点社区用户（[www.broadview.com.cn](http://www.broadview.com.cn)），您即可享受以下服务。

- **提交勘误：**您对书中内容的修改意见可在【提交勘误】处提交，若被采纳，将获赠博文视点社区积分（在您购买电子书时，积分可用来抵扣相应金额）。
- **交流互动：**在页面下方【读者评论】处留下您的疑问或观点，与我们和其他读者一同学习交流。

页面入口：<http://www.broadview.com.cn/32351>

二维码：



# 目 录

## 第 1 章 Kubernetes 入门 1

1.1	Kubernetes 是什么	1
1.2	为什么要用 Kubernetes	4
1.3	从一个简单的例子开始	5
1.3.1	环境准备	6
1.3.2	启动 MySQL 服务	6
1.3.3	启动 Tomcat 应用	9
1.3.4	通过浏览器访问网页	10
1.4	Kubernetes 基本概念和术语	12
1.4.1	Master	12
1.4.2	Node	12
1.4.3	Pod	15
1.4.4	Label (标签)	18
1.4.5	Replication Controller	22
1.4.6	Deployment	26
1.4.7	Horizontal Pod Autoscaler	28
1.4.8	StatefulSet	29
1.4.9	Service (服务)	30

1.4.10	Volume (存储卷)	37
1.4.11	Persistent Volume	41
1.4.12	Namespace (命名空间)	42
1.4.13	Annotation (注解)	43
1.4.14	小结	44
<b>第 2 章 Kubernetes 实践指南</b>		<b>45</b>
2.1	Kubernetes 安装与配置	45
2.1.1	系统要求	45
2.1.2	使用 kubeadm 工具快速安装 Kubernetes 集群	46
2.1.3	以二进制文件方式安装 Kubernetes 集群	51
2.1.4	Kubernetes 集群的安全设置	59
2.1.5	Kubernetes 集群的网络配置	64
2.1.6	内网中的 Kubernetes 相关配置	64
2.1.7	Kubernetes 的版本升级	65
2.1.8	Kubernetes 核心服务配置详解	66
2.2	kubectl 命令行工具用法详解	86
2.2.1	kubectl 用法概述	86
2.2.2	kubectl 子命令详解	88
2.2.3	kubectl 参数列表	90
2.2.4	kubectl 输出格式	90
2.2.5	kubectl 操作示例	92
2.3	深入掌握 Pod	93
2.3.1	Pod 定义详解	93
2.3.2	Pod 的基本用法	98
2.3.3	静态 Pod	103

2.3.4	Pod 容器共享 Volume	104
2.3.5	Pod 的配置管理	106
2.3.6	在容器内获取 Pod 信息（Downward API）	119
2.3.7	Pod 生命周期和重启策略	124
2.3.8	Pod 健康检查	125
2.3.9	玩转 Pod 调度	127
2.3.10	Init Container（初始化容器）	149
2.3.11	Pod 的升级和回滚	152
2.3.12	Pod 的扩容和缩容	166
2.3.13	使用 StatefulSet 搭建 MongoDB 集群	171
2.4	深入掌握 Service	180
2.4.1	Service 定义详解	181
2.4.2	Service 基本用法	182
2.4.3	Headless Service	187
2.4.4	集群外部访问 Pod 或 Service	192
2.4.5	DNS 服务搭建指南	196
2.4.6	自定义 DNS 和上游 DNS 服务器	204
2.4.7	Ingress：HTTP 7 层路由机制	208

## 第 3 章 Kubernetes 核心原理 226

---

3.1	Kubernetes API Server 原理分析	226
3.1.1	Kubernetes API Server 概述	226
3.1.2	独特的 Kubernetes Proxy API 接口	229
3.1.3	集群功能模块之间的通信	230
3.2	Controller Manager 原理分析	231
3.2.1	Replication Controller	232

3.2.2	Node Controller	234
3.2.3	ResourceQuota Controller	235
3.2.4	Namespace Controller	237
3.2.5	Service Controller 与 Endpoint Controller	237
3.3	Scheduler 原理分析	238
3.4	kubelet 运行机制分析	242
3.4.1	节点管理	242
3.4.2	Pod 管理	243
3.4.3	容器健康检查	244
3.4.4	cAdvisor 资源监控	245
3.5	kube-proxy 运行机制分析	247
3.6	深入分析集群安全机制	251
3.6.1	API Server 认证管理 (Authentication)	251
3.6.2	API Server 授权管理 (Authorization)	253
3.6.3	Admission Control (准入控制)	272
3.6.4	Service Account	274
3.6.5	Secret 私密凭据	279
3.7	网络原理	282
3.7.1	Kubernetes 网络模型	282
3.7.2	Docker 的网络基础	284
3.7.3	Docker 的网络实现	296
3.7.4	Kubernetes 的网络实现	304
3.7.5	Pod 和 Service 网络实战	308
3.7.6	CNI 网络模型	321
3.7.7	Kubernetes 网络策略	331
3.7.8	开源的网络组件	333

3.8 共享存储原理	363
3.8.1 共享存储机制概述	363
3.8.2 PV 详解	364
3.8.3 PVC 详解	368
3.8.4 PV 和 PVC 的生命周期	370
3.8.5 StorageClass 详解	373
3.8.6 动态存储管理实战：GlusterFS	376

## 第 4 章 Kubernetes 开发指南 388

---

4.1 REST 简述	388
4.2 Kubernetes API 详解	390
4.2.1 Kubernetes API 概述	390
4.2.2 API 版本	395
4.2.3 API Groups（API 组）	395
4.2.4 API 方法说明	397
4.2.5 API 响应说明	398
4.3 使用 Java 程序访问 Kubernetes API	400
4.3.1 Jersey	401
4.3.2 Fabric8	412
4.3.3 使用说明	413

## 第 5 章 Kubernetes 运维指南 434

---

5.1 Kubernetes 集群管理指南	434
5.1.1 Node 的管理	434
5.1.2 更新资源对象的 Label	436
5.1.3 Namespace：集群环境共享与隔离	437

5.1.4	Kubernetes 资源管理	441
5.1.5	资源紧缺时的 Pod 驱逐机制	475
5.1.6	Pod Disruption Budget (主动驱逐保护)	483
5.1.7	Kubernetes 集群的高可用部署方案	485
5.1.8	Kubernetes 集群监控	496
5.1.9	集群统一日志管理	513
5.1.10	Kubernetes 审计日志 (Audit Log)	522
5.1.11	使用 Web UI (Dashboard) 管理集群	523
5.1.12	Helm: Kubernetes 应用包管理工具	527
5.2	Trouble Shooting 指导	538
5.2.1	查看系统 Event 事件	538
5.2.2	查看容器日志	540
5.2.3	查看 Kubernetes 服务日志	541
5.2.4	常见问题	542
5.2.5	寻求帮助	546
5.3	Kubernetes 开发中的新功能	546
5.3.1	Pod Preset (运行时参数注入策略)	546
5.3.2	Cluster Federation (集群联邦)	553
5.3.3	容器运行时接口 (Container Runtime Interface-CRI)	557
5.3.4	对 GPU 的支持	561
5.3.5	Kubernetes 的演进路线 (Roadmap) 和开发模式	565

## 第 6 章 Kubernetes 源码导读

568

6.1	Kubernetes 源码结构和编译步骤	568
6.2	kube-apiserver 进程源码分析	572
6.2.1	进程启动过程	572

6.2.2	关键代码分析	574
6.2.3	设计总结	589
6.3	kube-controller-manager 进程源码分析	592
6.3.1	进程启动过程	592
6.3.2	关键代码分析	595
6.3.3	设计总结	603
6.4	kube-scheduler 进程源码分析	605
6.4.1	进程启动过程	605
6.4.2	关键代码分析	610
6.4.3	设计总结	617
6.5	kubelet 进程源码分析	619
6.5.1	进程启动过程	619
6.5.2	关键代码分析	624
6.5.3	设计总结	647
6.6	kube-proxy 进程源码分析	648
6.6.1	进程启动过程	648
6.6.2	关键代码分析	650
6.6.3	设计总结	665
6.7	kubectl 进程源码分析	666
6.7.1	kubectl create 命令	667
6.7.2	rolling-update 命令	671