
Table of Contents

Foreword to the Second Edition.....	xvii
Foreword to the First Edition.....	xix
Preface.....	xxi
1. Revolution in the Cloud.....	1
The Creation of the Cloud	2
Buying Time	3
Infrastructure as a Service	3
The Dawn of DevOps	3
Improving Feedback Loops	4
What Does DevOps Mean?	4
Infrastructure as Code	6
Learning Together	6
The Coming of Containers	7
The State of the Art	7
Thinking Inside the Box	7
Putting Software in Containers	8
Plug and Play Applications	9
Conducting the Container Orchestra	10
Kubernetes	11
From Borg to Kubernetes	11
Why Kubernetes?	11
Will Kubernetes Disappear?	13
Kubernetes Is Not a Panacea	14
Cloud Native	15
The Future of Operations	17

Distributed DevOps	18
Some Things Will Remain Centralized	18
Developer Productivity Engineering	18
You Are the Future	19
Summary	19
2. First Steps with Kubernetes.....	21
Running Your First Container	21
Installing Docker Desktop	22
What Is Docker?	22
Running a Container Image	23
The Demo Application	23
Looking at the Source Code	24
Introducing Go	24
How the Demo App Works	25
Building a Container	25
Understanding Dockerfiles	26
Minimal Container Images	26
Running Docker Image Build	27
Naming Your Images	27
Port Forwarding	28
Container Registries	28
Authenticating to the Registry	29
Naming and Pushing Your Image	29
Running Your Image	29
Hello, Kubernetes	29
Running the Demo App	30
If the Container Doesn't Start	31
Minikube	31
Summary	31
3. Getting Kubernetes.....	33
Cluster Architecture	33
The Control Plane	34
Node Components	35
High Availability	35
The Costs of Self-Hosting Kubernetes	37
It's More Work Than You Think	37
It's Not Just About the Initial Setup	38
Tools Don't Do All the Work for You	39
Kubernetes the Hard Way	39
Kubernetes Is Hard	39

Administration Overhead	39
Start with Managed Services	40
Managed Kubernetes Services	41
Google Kubernetes Engine (GKE)	41
Cluster Autoscaling	41
Autopilot	41
Amazon Elastic Kubernetes Service (EKS)	42
Azure Kubernetes Service (AKS)	42
IBM Cloud Kubernetes Service	43
DigitalOcean Kubernetes	43
Kubernetes Installers	43
kops	43
Kubespray	44
kubeadm	44
Rancher Kubernetes Engine (RKE)	44
Puppet Kubernetes Module	44
Buy or Build: Our Recommendations	45
Run Less Software	45
Use Managed Kubernetes if You Can	45
But What About Vendor Lock-in?	46
Bare-Metal and On-Prem	46
Multicloud Kubernetes Clusters	47
OpenShift	48
Anthos	48
Use Standard Kubernetes Self-Hosting Tools if You Must	48
Clusterless Container Services	49
AWS Fargate	49
Azure Container Instances (ACI)	49
Google Cloud Run	50
Summary	50
4. Working with Kubernetes Objects.....	53
Deployments	53
Supervising and Scheduling	54
Restarting Containers	54
Creating Deployments	55
Pods	55
ReplicaSets	56
Maintaining Desired State	57
The Kubernetes Scheduler	58
Resource Manifests in YAML Format	59
Resources Are Data	59

Deployment Manifests	59
Using kubectl apply	60
Service Resources	61
Querying the Cluster with kubectl	63
Taking Resources to the Next Level	64
Helm: A Kubernetes Package Manager	64
Installing Helm	65
Installing a Helm Chart	65
Charts, Repositories, and Releases	66
Listing Helm Releases	67
Summary	67
5. Managing Resources.....	69
Understanding Resources	69
Resource Units	70
Resource Requests	70
Resource Limits	71
Quality of Service	71
Managing the Container Life Cycle	72
Liveness Probes	72
Probe Delay and Frequency	73
Other Types of Probes	73
Readiness Probes	74
Startup Probes	75
gRPC Probes	75
File-Based Readiness Probes	76
minReadySeconds	76
Pod Disruption Budgets	77
Using Namespaces	78
Working with Namespaces	79
What Namespaces Should I Use?	79
Service Addresses	80
Resource Quotas	80
Default Resource Requests and Limits	82
Optimizing Cluster Costs	83
Kubecost	83
Optimizing Deployments	83
Optimizing Pods	84
Vertical Pod Autoscaler	85
Optimizing Nodes	85
Optimizing Storage	87
Cleaning Up Unused Resources	87

Checking Spare Capacity	90
Using Reserved Instances	90
Using Preemptible (Spot) Instances	91
Keeping Your Workloads Balanced	93
Summary	94
6. Operating Clusters.....	97
Cluster Sizing and Scaling	97
Capacity Planning	98
Nodes and Instances	101
Scaling the Cluster	103
Conformance Checking	104
CNCF Certification	105
Conformance Testing with Sonobuoy	106
Kubernetes Audit Logging	108
Chaos Testing	108
Only Production Is Production	109
chaoskube	109
kube-monkey	110
PowerfulSeal	110
Summary	111
7. Kubernetes Power Tools.....	113
Mastering kubectl	113
Shell Aliases	113
Using Short Flags	114
Abbreviating Resource Types	114
Auto-Completing kubectl Commands	115
Getting Help	115
Getting Help on Kubernetes Resources	116
Showing More Detailed Output	116
Working with JSON Data and jq	116
Watching Objects	117
Describing Objects	118
Working with Resources	118
Imperative kubectl Commands	118
When Not to Use Imperative Commands	119
Generating Resource Manifests	120
Exporting Resources	120
Diffing Resources	120
Working with Containers	121
Viewing a Container's Logs	121

Attaching to a Container	122
Watching Kubernetes Resources with kubespy	123
Forwarding a Container Port	123
Executing Commands on Containers	123
Running Containers for Troubleshooting	124
Using BusyBox Commands	125
Adding BusyBox to Your Containers	126
Installing Programs on a Container	127
Contexts and Namespaces	127
kubeconfig files	128
kubectl and kubens	129
kubectl-ps1	130
Kubernetes Shells and Tools	130
kubectl-shell	130
Click	130
kubed-sh	131
Stern	131
Kubernetes IDEs	131
Lens	132
VS Code Kubernetes Extension	132
Building Your Own Kubernetes Tools	132
Summary	133
8. Running Containers.....	135
Containers and Pods	135
What Is a Container?	136
Container Runtimes in Kubernetes	137
What Belongs in a Container?	137
What Belongs in a Pod?	138
Container Manifests	139
Image Identifiers	139
The latest Tag	140
Container Digests	141
Base Image Tags	141
Ports	142
Resource Requests and Limits	142
Image Pull Policy	143
Environment Variables	143
Container Security	144
Running Containers as a Non-Root User	144
Blocking Root Containers	145
Setting a Read-Only Filesystem	146

Disabling Privilege Escalation	146
Capabilities	146
Pod Security Contexts	148
Pod Service Accounts	148
Volumes	148
emptyDir Volumes	149
Persistent Volumes	150
Restart Policies	151
Image Pull Secrets	151
Init Containers	152
Summary	152
9. Managing Pods.....	155
Labels	155
What Are Labels?	155
Selectors	156
More Advanced Selectors	157
Other Uses for Labels	158
Labels and Annotations	159
Node Affinities	159
Hard Affinities	160
Soft Affinities	160
Pod Affinities and Anti-Affinities	161
Keeping Pods Together	161
Keeping Pods Apart	162
Soft Anti-Affinities	163
When to Use Pod Affinities	163
Taints and Tolerations	164
Pod Controllers	165
DaemonSets	166
StatefulSets	167
Jobs	168
CronJobs	169
Horizontal Pod Autoscalers	170
Operators and Custom Resource Definitions (CRDs)	172
Ingress	173
Ingress Controllers	174
Ingress Rules	175
Terminating TLS with Ingress	176
Service Mesh	176
Istio	177
Linkerd	177

Consul Connect	177
NGINX Service Mesh	177
Summary	178
10. Configuration and Secrets.....	181
ConfigMaps	181
Creating ConfigMaps	182
Setting Environment Variables from ConfigMaps	183
Setting the Whole Environment from a ConfigMap	185
Using Environment Variables in Command Arguments	186
Creating Config Files from ConfigMaps	187
Updating Pods on a Config Change	188
Kubernetes Secrets	189
Using Secrets as Environment Variables	189
Writing Secrets to Files	190
Reading Secrets	191
Access to Secrets	192
Encryption at Rest	192
Keeping Secrets and ConfigMaps	192
Secrets Management Strategies	193
Encrypt Secrets in Version Control	193
Use a Dedicated Secrets Management Tool	194
Encrypting Secrets with Sops	195
Encrypting a File with Sops	195
Using a KMS Backend	197
Sealed Secrets	197
Summary	198
11. Security, Backups, and Cluster Health.....	199
Access Control and Permissions	199
Managing Access by Cluster	199
Introducing Role-Based Access Control (RBAC)	200
Understanding Roles	201
Binding Roles to Users	201
What Roles Do I Need?	202
Guard Access to cluster-admin	202
Applications and Deployment	203
RBAC Troubleshooting	204
Cluster Security Scanning	205
Gatekeeper/OPA	205
kube-bench	205
Kubescape	206

Container Security Scanning	206
Clair	206
Aqua	207
Anchore Engine	207
Synk	207
Backups	208
Do I Need to Back Up Kubernetes?	209
Backing Up etcd	209
Backing Up Resource State	210
Backing Up Cluster State	210
Large and Small Disasters	211
Velero	211
Monitoring Cluster Status	214
kubectl	214
CPU and Memory Utilization	216
Cloud Provider Console	216
Kubernetes Dashboard	217
Weave Scope	218
kube-ops-view	218
node-problem-detector	218
Further Reading	218
Summary	219
12. Deploying Kubernetes Applications.....	221
Building Manifests with Helm	221
What's Inside a Helm Chart?	222
Helm Templates	223
Interpolating Variables	224
Quoting Values in Templates	225
Specifying Dependencies	225
Deploying Helm Charts	226
Setting Variables	226
Specifying Values in a Helm Release	227
Updating an App with Helm	227
Rolling Back to Previous Versions	228
Creating a Helm Chart Repo	228
Managing Helm Chart Secrets with Sops	229
Managing Multiple Charts with Helmfile	231
What's in a Helmfile?	231
Chart Metadata	232
Applying the Helmfile	233
Advanced Manifest Management Tools	234

kustomize	234
Tanka	235
Kapitan	236
kompose	236
Ansible	236
kubeval	237
Summary	238
13. Development Workflow.....	239
Development Tools	239
Skaffold	239
Telepresence	241
Waypoint	241
Knative	242
OpenFaaS	242
Crossplane	242
Deployment Strategies	244
Rolling Updates	244
Recreate	245
maxSurge and maxUnavailable	245
Blue/Green Deployments	246
Rainbow Deployments	247
Canary Deployments	247
Handling Migrations with Helm	248
Helm Hooks	248
Handling Failed Hooks	249
Other Hooks	249
Chaining Hooks	250
Summary	250
14. Continuous Deployment in Kubernetes.....	253
What Is Continuous Deployment?	253
Which CD Tool Should I Use?	254
Hosted CI/CD Tools	254
Azure Pipelines	255
Google Cloud Build	255
Codefresh	255
GitHub Actions	255
GitLab CI	256
Self-Hosted CI/CD Tools	256
Jenkins	256
Drone	256

Tekton	256
Concourse	256
Spinnaker	257
Argo	257
Keel	257
A CI/CD Pipeline with Cloud Build	257
Setting Up Google Cloud and GKE	258
Forking the Demo Repository	258
Create Artifact Registry Container Repository	258
Configuring Cloud Build	259
Building the Test Container	259
Running the Tests	260
Building the Application Container	260
Substitution Variables	261
Git SHA Tags	261
Validating the Kubernetes Manifests	261
Publishing the Image	262
Creating the First Build Trigger	262
Testing the Trigger	263
Deploying from a CI/CD Pipeline	263
Creating a Deploy Trigger	265
Adapting the Example Pipeline	266
GitOps	266
Flux	266
Summary	269
15. Observability and Monitoring.....	271
What Is Observability?	271
What Is Monitoring?	271
Closed-Box Monitoring	271
What Does “Up” Mean?	273
Logging	274
Introducing Metrics	276
Tracing	277
Observability	278
The Observability Pipeline	279
Monitoring in Kubernetes	280
External Closed-Box Checks	280
Internal Health Checks	282
Summary	284

16. Metrics in Kubernetes.....	285
What Are Metrics, Really?	285
Time-Series Data	285
Counters and Gauges	286
What Can Metrics Tell Us?	287
Choosing Good Metrics	287
Services: The RED Pattern	288
Resources: The USE Pattern	289
Business Metrics	290
Kubernetes Metrics	291
Analyzing Metrics	294
What’s Wrong with a Simple Average?	295
Means, Medians, and Outliers	295
Discovering Percentiles	296
Applying Percentiles to Metrics Data	296
We Usually Want to Know the Worst	298
Beyond Percentiles	298
Graphing Metrics with Dashboards	299
Use a Standard Layout for All Services	299
Build an Information Radiator with Primary Dashboards	300
Dashboard Things That Break	301
Alerting on Metrics	303
What’s Wrong with Alerts?	303
On-Call Should Not Be Hell	304
Urgent, Important, and Actionable Alerts	304
Track Your Alerts, Out-of-Hours Pages, and Wake-Ups	305
Metrics Tools and Services	306
Prometheus	306
Google Operations Suite	308
AWS CloudWatch	309
Azure Monitor	309
Datadog	309
New Relic	310
Summary	311
Afterword.....	313
Index.....	317