

Inhaltsverzeichnis

Vorwort	xv
1 Einführung	1
1.1 Schnelligkeit	2
1.1.1 Der Wert der Immutabilität	3
1.1.2 Deklarative Konfiguration	4
1.1.3 Selbstheilende Systeme	5
1.2 Ihren Service und Ihre Teams skalieren	6
1.2.1 Entkoppeln	6
1.2.2 Einfaches Skalieren für Anwendungen und Cluster	6
1.2.3 Entwicklungs-Teams mit Microservices skalieren	7
1.2.4 Konsistenz und Skalierung durch Separation of Concerns	8
1.3 Abstrahieren Sie Ihre Infrastruktur	10
1.4 Effizienz	11
1.5 Zusammenfassung	12
2 Container erstellen und ausführen	13
2.1 Container-Images	14
2.1.1 Das Docker-Image-Format	15
2.2 Anwendungs-Images mit Docker bauen	16
2.2.1 Dockerfiles	16
2.2.2 Die Image-Größe optimieren	18
2.2.3 Sicherheit von Images	19
2.3 Multistage Image Build	20
2.4 Images in einer Remote-Registry ablegen	22
2.5 Die Docker Container Runtime	23
2.5.1 Container mit Docker ausführen	23
2.5.2 Die kuard-Anwendung erforschen	23
2.5.3 Den Ressourcen-Einsatz begrenzen	24
2.6 Aufräumen	25
2.7 Zusammenfassung	25

3	Ein Kubernetes-Cluster deployen	27
3.1	Kubernetes auf einem öffentlichen Cloud-Provider installieren	28
3.1.1	Google Kubernetes Engine	28
3.1.2	Kubernetes mit dem Azure Kubernetes Service installieren	28
3.1.3	Kubernetes auf den Amazon Web Services installieren	29
3.1.4	Kubernetes mit minikube lokal installieren	29
3.2	Kubernetes in Docker ausführen	30
3.3	Kubernetes auf dem Raspberry Pi ausführen	31
3.4	Der Kubernetes-Client	31
3.4.1	Den Cluster-Status prüfen	31
3.4.2	Worker-Knoten in Kubernetes auflisten	32
3.5	Cluster-Komponenten	34
3.5.1	Kubernetes-Proxy	35
3.5.2	Kubernetes-DNS	35
3.5.3	Kubernetes-UI	36
3.6	Zusammenfassung	36
4	Häufige kubectl-Befehle	37
4.1	Namensräume	37
4.2	Kontexte	37
4.3	Objekte der Kubernetes-API anzeigen	38
4.4	Kubernetes-Objekte erstellen, aktualisieren und löschen	39
4.5	Objekte mit einem Label und Anmerkungen versehen	40
4.6	Debugging-Befehle	40
4.7	Autovervollständigen von Befehlen	42
4.8	Alternative Möglichkeiten zur Kommunikation mit Ihrem Cluster	42
4.9	Zusammenfassung	43
5	Pods	45
5.1	Pods in Kubernetes	46
5.2	In Pods denken	46
5.3	Das Pod-Manifest	47
5.3.1	Einen Pod erstellen	48
5.3.2	Ein Pod-Manifest schreiben	49
5.4	Pods starten	49
5.4.1	Pods auflisten	50
5.4.2	Pod-Details	50
5.4.3	Einen Pod löschen	51

5.5	Auf Ihren Pod zugreifen	52
5.5.1	Port-Forwarding einsetzen	52
5.5.2	Mehr Informationen aus Logs erhalten	53
5.5.3	Befehle in Ihrem Container mit exec ausführen	53
5.5.4	Dateien von und auf Container kopieren	53
5.6	Health-Checks	54
5.6.1	Liveness-Probe	54
5.6.2	Readiness-Probe	56
5.6.3	Arten von Health-Checks	56
5.7	Ressourcen-Management	56
5.7.1	Ressourcen-Anforderungen: Minimal notwendige Ressourcen	57
5.7.2	Den Ressourcen-Einsatz durch Grenzen beschränken	59
5.8	Daten mit Volumes persistieren	60
5.8.1	Volumes in Pods definieren	60
5.8.2	Volumes in Pods nutzen	61
5.8.3	Daten auf Remote-Speicher persistieren	62
5.9	Fügen Sie alles zusammen	62
5.10	Zusammenfassung	64
6	Labels und Anmerkungen	65
6.1	Labels	65
6.1.1	Labels anwenden	67
6.1.2	Labels anpassen	70
6.1.3	Label-Selektoren	71
6.1.4	Label-Selektoren in API-Objekten	73
6.1.5	Labels in der Architektur von Kubernetes	73
6.2	Anmerkungen	74
6.2.1	Anmerkungen definieren	75
6.3	Aufräumen	76
6.4	Zusammenfassung	76
7	Service-Discovery	77
7.1	Was ist Service-Discovery?	77
7.2	Das Service-Objekt	78
7.2.1	Service-DNS	80
7.2.2	Readiness-Checks	81
7.3	Über das Cluster hinausschauen	83
7.4	Cloud-Integration	84

7.5	Weitere Details	85
7.5.1	Endpunkte	85
7.5.2	Manuelle Service-Discovery	86
7.5.3	kube-proxy und Cluster-IPs	87
7.5.4	Umgebungsvariablen zur Cluster-IP	88
7.6	Mit anderen Umgebungen verbinden	89
7.7	Aufräumen	90
7.8	Zusammenfassung	90
8	HTTP Load Balancing mit Ingress	91
8.1	Ingress-Spec versus Ingress-Controller	92
8.2	Contour installieren	93
8.2.1	DNS konfigurieren	94
8.2.2	Eine lokale hosts-Datei konfigurieren	94
8.3	Ingress verwenden	95
8.3.1	Einfachste Anwendung	95
8.3.2	Hostnamen verwenden	96
8.3.3	Pfade verwenden	97
8.3.4	Aufräumen	98
8.4	Fortgeschrittenere Themen und Probleme mit Ingress	98
8.4.1	Mehrere Ingress-Controller laufen lassen	99
8.4.2	Mehrere Ingress-Objekte	99
8.4.3	Ingress und Namensräume	99
8.4.4	Path Rewriting	100
8.4.5	TLS	100
8.5	Alternative Ingress-Implementierungen	102
8.6	Die Zukunft von Ingress	103
8.7	Zusammenfassung	103
9	ReplicaSets	105
9.1	Reconciliation-Schleifen	106
9.2	Die Verbindung zwischen Pods und ReplicaSets	106
9.2.1	Bestehende Container übernehmen	107
9.2.2	Container in Quarantäne stecken	107
9.3	Mit ReplicaSets designen	108
9.4	Spezifikation eines ReplicaSets	108
9.4.1	Pod-Templates	108
9.4.2	Labels	109
9.5	Ein ReplicaSet erstellen	109

9.6	Ein ReplicaSet untersuchen	110
9.6.1	Ein ReplicaSet über einen Pod finden	110
9.6.2	Eine Gruppe von Pods für ein ReplicaSet finden	111
9.7	ReplicaSets skalieren	111
9.7.1	Imperatives Skalieren mit kubectl scale	111
9.7.2	Deklaratives Skalieren mit kubectl apply	112
9.7.3	Ein ReplicaSet automatisch skalieren	113
9.8	ReplicaSets löschen	114
9.9	Zusammenfassung	114
10	Deployments	115
10.1	Ihr erstes Deployment	116
10.1.1	Deployment-Interns	116
10.2	Deployments erstellen	118
10.3	Deployments verwalten	119
10.4	Deployments aktualisieren	120
10.4.1	Ein Deployment skalieren	120
10.4.2	Ein Container-Image aktualisieren	121
10.4.3	Rollout-History	122
10.5	Deployment-Strategien	125
10.5.1	Recreate-Strategie	125
10.5.2	RollingUpdate-Strategie	125
10.5.3	Rollouts verlangsamen, um die Service-Qualität sicherzustellen	129
10.6	Ein Deployment löschen	131
10.7	Ein Deployment überwachen	131
10.8	Zusammenfassung	131
11	DaemonSets	133
11.1	Der DaemonSet-Scheduler	134
11.2	DaemonSets erstellen	135
11.3	DaemonSets auf bestimmte Knoten beschränken	137
11.3.1	Knoten mit Labels versehen	137
11.3.2	Knoten-Selektoren	137
11.4	Ein DaemonSet aktualisieren	139
11.4.1	Rollierendes Update eines DaemonSet	139
11.5	Ein DaemonSet löschen	140
11.6	Zusammenfassung	140

12	Jobs	141
12.1	Das Job-Objekt	141
12.2	Job-Muster	142
12.2.1	Einmalig	142
12.2.2	Parallelism	147
12.2.3	Work-Queues	149
12.3	CronJobs	153
12.4	Zusammenfassung	154
13	ConfigMaps und Secrets	155
13.1	ConfigMaps	155
13.1.1	ConfigMaps erstellen	155
13.1.2	Eine ConfigMap verwenden	156
13.2	Secrets	160
13.2.1	Secrets erstellen	160
13.2.2	Secrets konsumieren	162
13.2.3	Private Docker-Registries	163
13.3	Namensbeschränkungen	164
13.4	ConfigMaps und Secrets managen	165
13.4.1	Ausgabe	165
13.4.2	Erstellen	166
13.4.3	Aktualisieren	166
13.5	Zusammenfassung	168
14	Role-Based Access Control für Kubernetes	169
14.1	Role-Based Access Control	170
14.1.1	Identität in Kubernetes	170
14.1.2	Rollen und Role Bindings verstehen	171
14.1.3	Rollen und Role Bindings in Kubernetes	171
14.2	Techniken zur Arbeit mit RBAC	174
14.2.1	Die Autorisierung mit can-i testen	174
14.2.2	RBAC in der Versionsverwaltung managen	175
14.3	Fortgeschrittene Techniken	175
14.3.1	Cluster-Rollen aggregieren	175
14.3.2	Gruppen für Bindings verwenden	176
14.4	Zusammenfassung	177

15	Storage-Lösungen in Kubernetes integrieren	179
15.1	Externe Services importieren	180
15.1.1	Services ohne Selektoren	181
15.1.2	Grenzen für externe Services: Health-Checking	183
15.2	Zuverlässige Singletons ausführen	183
15.2.1	Ein MySQL-Singleton ausführen	184
15.2.2	Dynamisches Volume-Provisioning	187
15.3	Kubernetes-eigenes Storage mit StatefulSets	189
15.3.1	Eigenschaften von StatefulSets	189
15.3.2	Manuell replizierte MongoDB mit StatefulSets	189
15.3.3	Das MongoDB-Cluster automatisch erstellen	192
15.3.4	Persistente Volumes und StatefulSets	195
15.3.5	Zum Abschluss: Readiness-Proben	196
15.4	Zusammenfassung	196
16	Kubernetes erweitern	197
16.1	Was bedeutet das Erweitern von Kubernetes?	197
16.2	Erweiterungspunkte	198
16.3	Patterns für Custom Resources	206
16.3.1	Just Data	206
16.3.2	Compiler	207
16.3.3	Operator	207
16.3.4	Der Einstieg	208
16.4	Zusammenfassung	208
17	Reale Anwendungen deployen	209
17.1	Jupyter	209
17.2	Parse	211
17.2.1	Voraussetzungen	211
17.2.2	Den Parse-Server bauen	211
17.2.3	Den Parse-Server deployen	212
17.2.4	Parse testen	213
17.3	Ghost	213
17.3.1	Ghost konfigurieren	214
17.4	Redis	217
17.4.1	Redis konfigurieren	217
17.4.2	Einen Redis-Service erstellen	219
17.4.3	Redis deployen	219
17.4.4	Mit unserem Redis-Cluster experimentieren	221
17.5	Zusammenfassung	222

18	Organisieren Sie Ihre Anwendung	223
18.1	Leitprinzipien	223
18.1.1	Dateisysteme als Source of Truth	223
18.1.2	Die Rolle des Code Reviews	224
18.1.3	Feature Gates und Guards	225
18.2	Ihre Anwendung in der Versionsverwaltung managen	226
18.2.1	Struktur im Dateisystem	226
18.2.2	Regelmäßige Versionen managen	227
18.3	Ihre Anwendung für Entwicklung, Testen und Deployment strukturieren	229
18.3.1	Ziele	229
18.3.2	Verlauf eines Releases	230
18.4	Ihre Anwendung durch Templates parametrisieren	231
18.4.1	Mit Helm und Templates parametrisieren	231
18.4.2	Dateisystem-Layout zur Parametrisierung	232
18.5	Ihre Anwendung weltweit deployen	233
18.5.1	Architekturen für ein weltweites Deployment	233
18.5.2	Ein weltweites Deployment implementieren	234
18.5.3	Dashboards und Monitoring für weltweite Deployments	236
18.6	Zusammenfassung	236
	Index	237