KubeCon | CloudNativeCon

Europe 2020

Virtual

Kubeadm Deep Dive

Who are we?

Rostislav (Ross) Georgiev

VMware

Rafael Fernández López

SUSE

What is Kubeadm?

Kubeadm is a Kubernetes node bootstrapper

- Someone or something should provide you with the machine
- A container runtime and kubelet must be already installed
- kubeadm does NOT install any CNI
- CNI, container runtime, cloud provider, and machine type agnostic

kubeadm

What is Kubeadm?

Designed to be:

- Easy to use
- Provide sane defaults for 80% of the use cases
- Make the other 20% possible

kubeadm

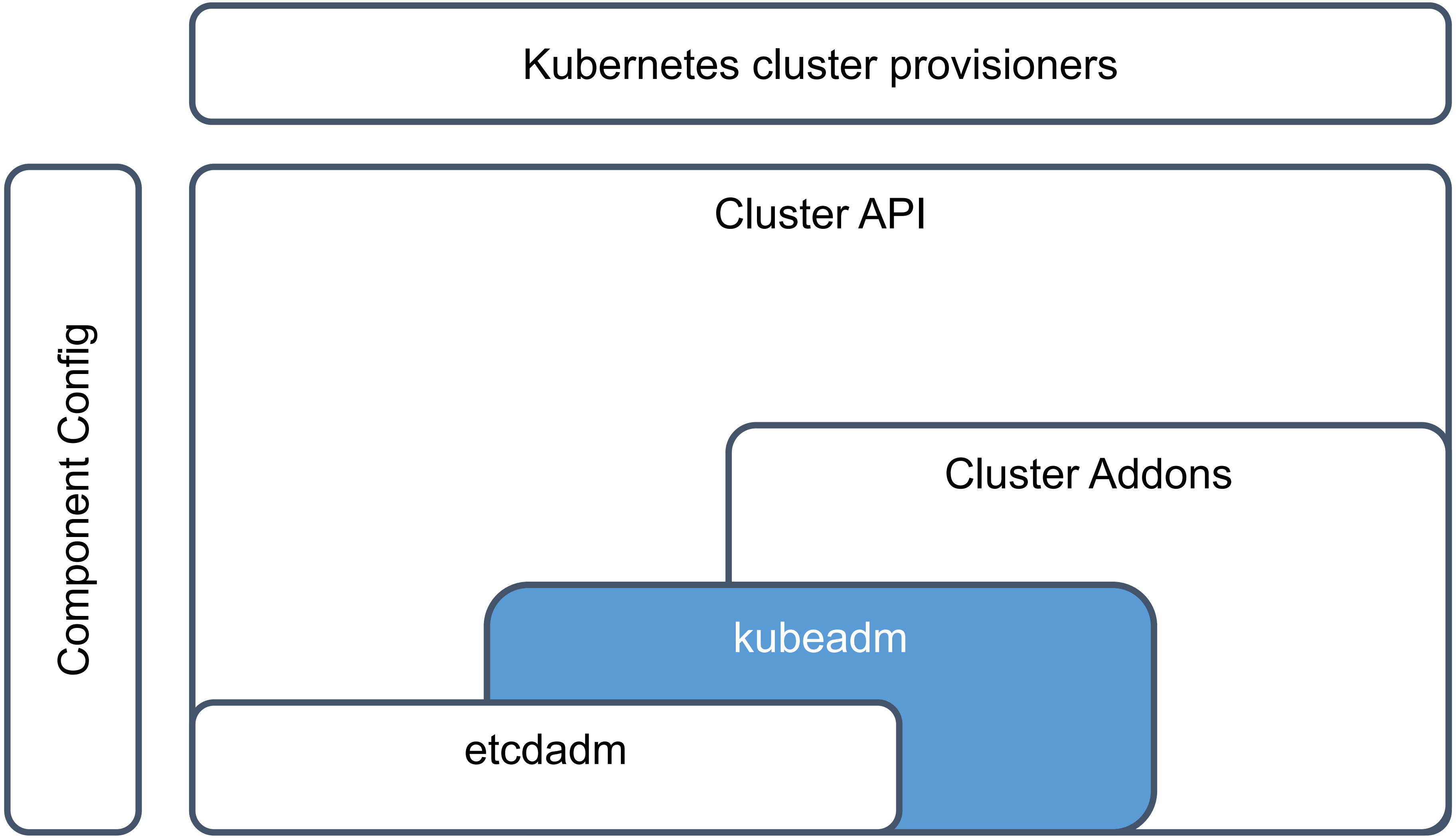
What is Kubeadm?

A project of SIG Cluster Lifecycle

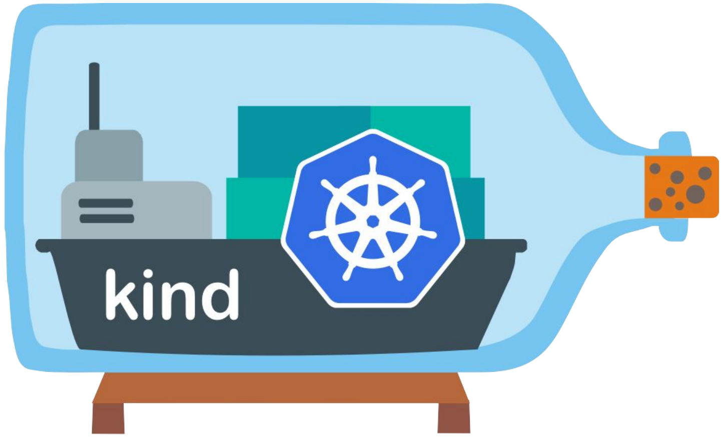
SIG Cluster Lifecycle’s objective is to simplify creation, configuration, upgrade, downgrade, and teardown of Kubernetes clusters and their components.

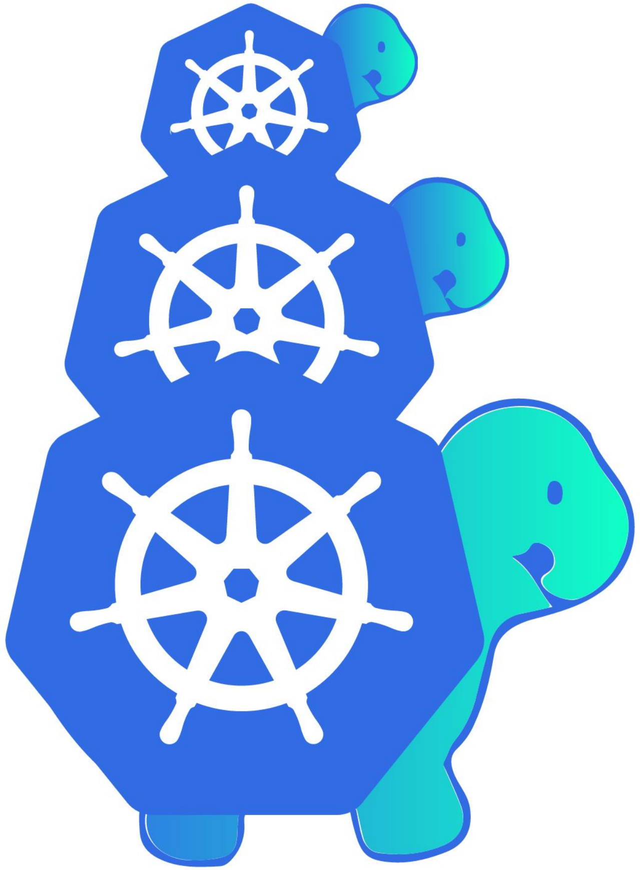
--- The SIG Cluster Lifecycle charter

What is Kubeadm?



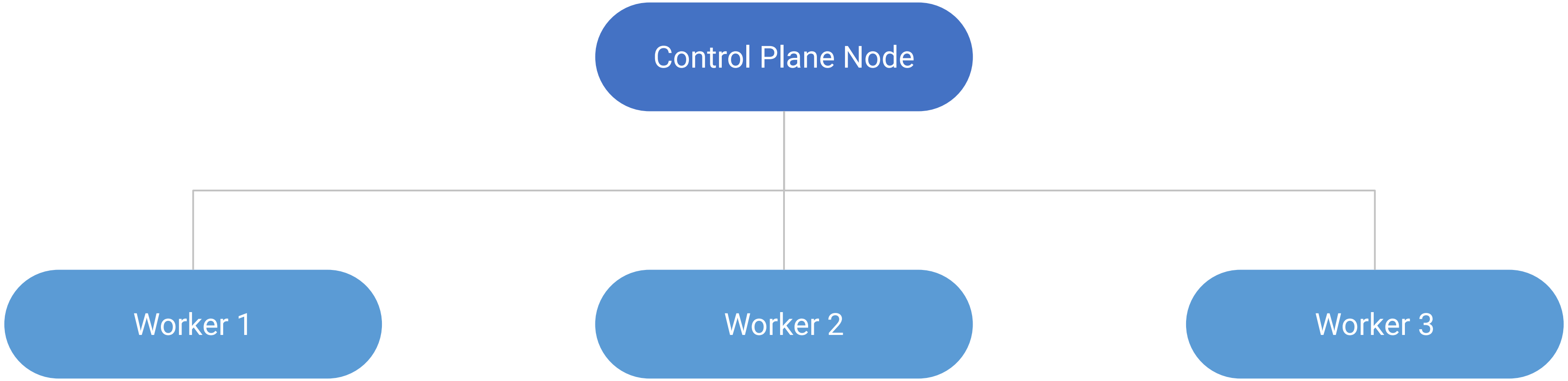
Who uses Kubeadm?





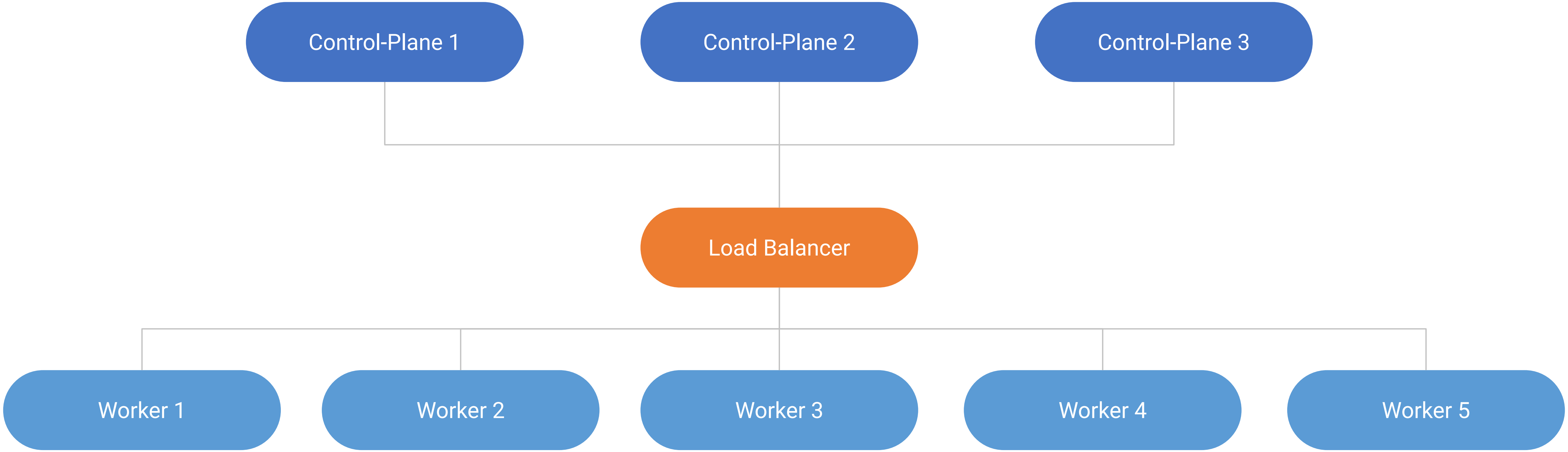
How does it work?

Deployment strategy: Single control-plane node



How does it work?

Deployment strategy: HA (multiple control-plane nodes)



Main Workflow

1. Initialize the cluster and the first control-plane node

```
$ sudo kubeadm init
```

2. Install a POD network addon

```
$ kubectl apply ...
```

3. Join more control-plane nodes

```
$ sudo kubeadm join <control-plane-host>:<control-plane-port> \
  --token <token> --discovery-token-ca-cert-hash sha256:<hash> \
  --control-plane --certificate-key <certificate-decryption-key>
```

4. Join worker nodes

```
$ sudo kubeadm join <control-plane-host>:<control-plane-port> \
  --token <token> --discovery-token-ca-cert-hash sha256:<hash>
```

Upgrade Workflow

1. Check for available upgrades

```
$ sudo kubeadm upgrade plan
```

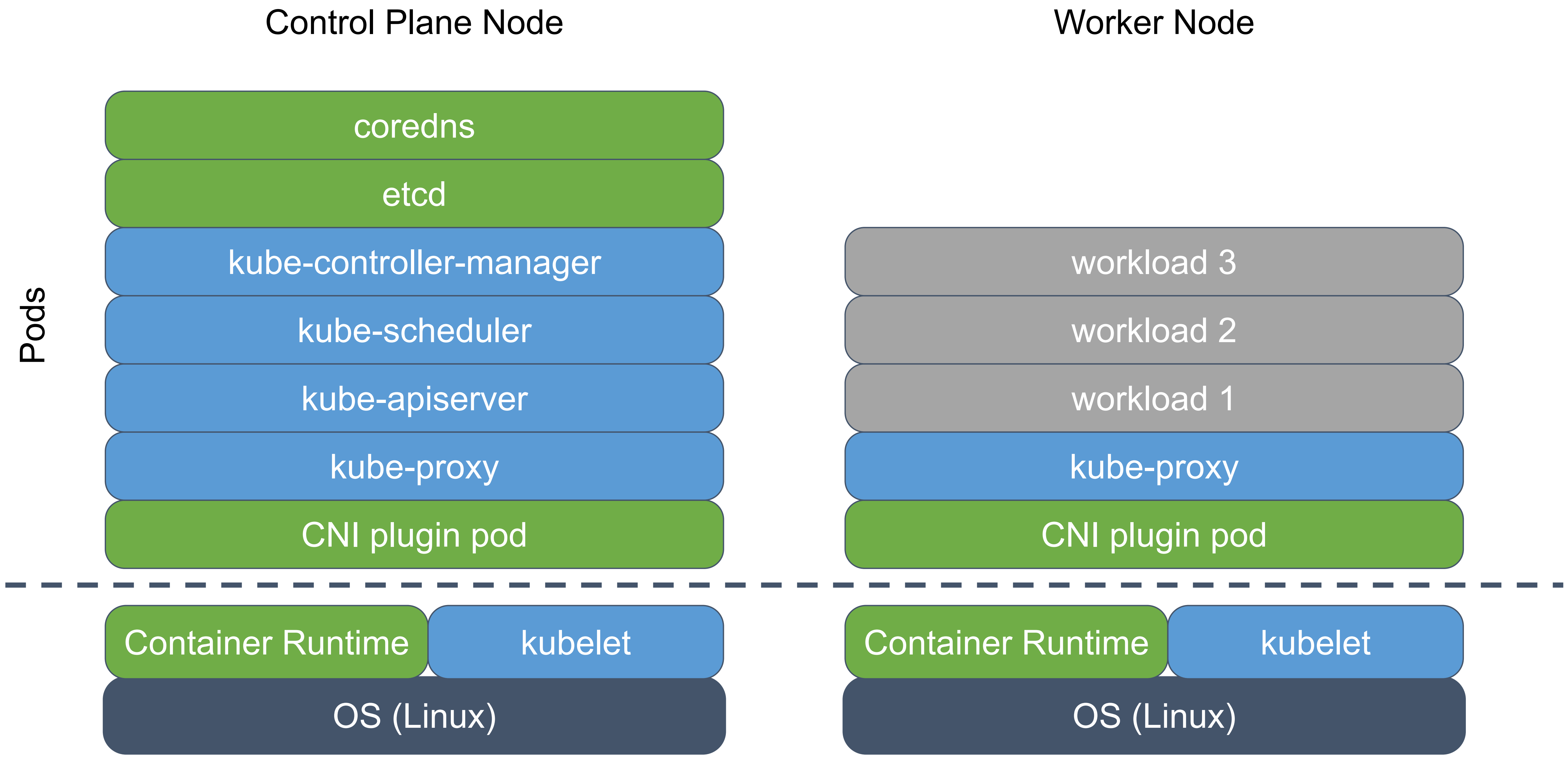
2. Upgrade the first control-plane node on the cluster

```
$ sudo kubeadm upgrade apply v1.19.0
```

3. Upgrade the rest of the nodes

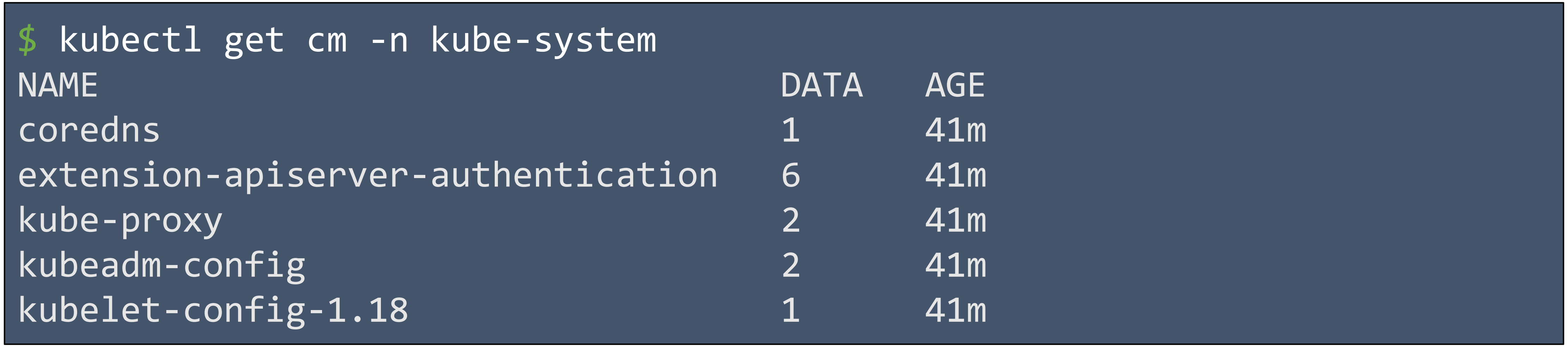
```
$ sudo kubeadm upgrade node
```

What is deployed?



Where’s the config?

KubeCon CloudNativeCon Europe 2020 Virtual



```
$ kubectl get cm -n kube-system
NAME                                  DATA   AGE
coredns                               1      41m
extension-apiserver-authentication    6      41m
kube-proxy                            2      41m
kubeadm-config                        2      41m
kubelet-config-1.18                   1      41m
```

Kubeadm's State

Kubeadm is stable and GA

- HA (multi node control plane)
- Config @ v1beta2
- Dual Stack support
- Customizing static Pods
- Certificate Management
- Kubeadm managed etcd
  - Can be opted out of
- Standard addons (CoreDNS, kube-proxy)
  - Can be opted out of
- Phases support
  - For kubeadm init, join, reset, & upgrade node

New Developments

Kustomize is deprecated and replaced with patches

- Simpler, lighter, more extendable solution
- Enables patching Static Pods on a per-node basis
- Supports “strategic”, “json”, and “merge” patches (see “kubectl patch”)
- Special naming convention:

  `componentname[suffix][+patchtype].{yaml|json}`

  - “`componentname`” is “kube-apiserver”, “kube-controller-manager”, “kube-scheduler”, “etcd”, …
  - “`suffix`” allows you to order the patches (e. g. “01”, “50”, “99”, etc.)
  - “`+patchtype`” allows you to specify the type of the patch (e.g. “+strategic”, “+json”, or “+merge”), if omitted it implies “+strategic”.
  - The patches can be either in JSON or YAML formats.

New Developments

Patches in action

~/patches/kube-controller-manager+strategic.yaml

```
metadata:
  annotations:
    extra-annotation: "Hello!"
```

~/patches/etcd01+merge.json

```
{"metadata":{"annotations":{"extra-annotation":"Hello!"}}}
```

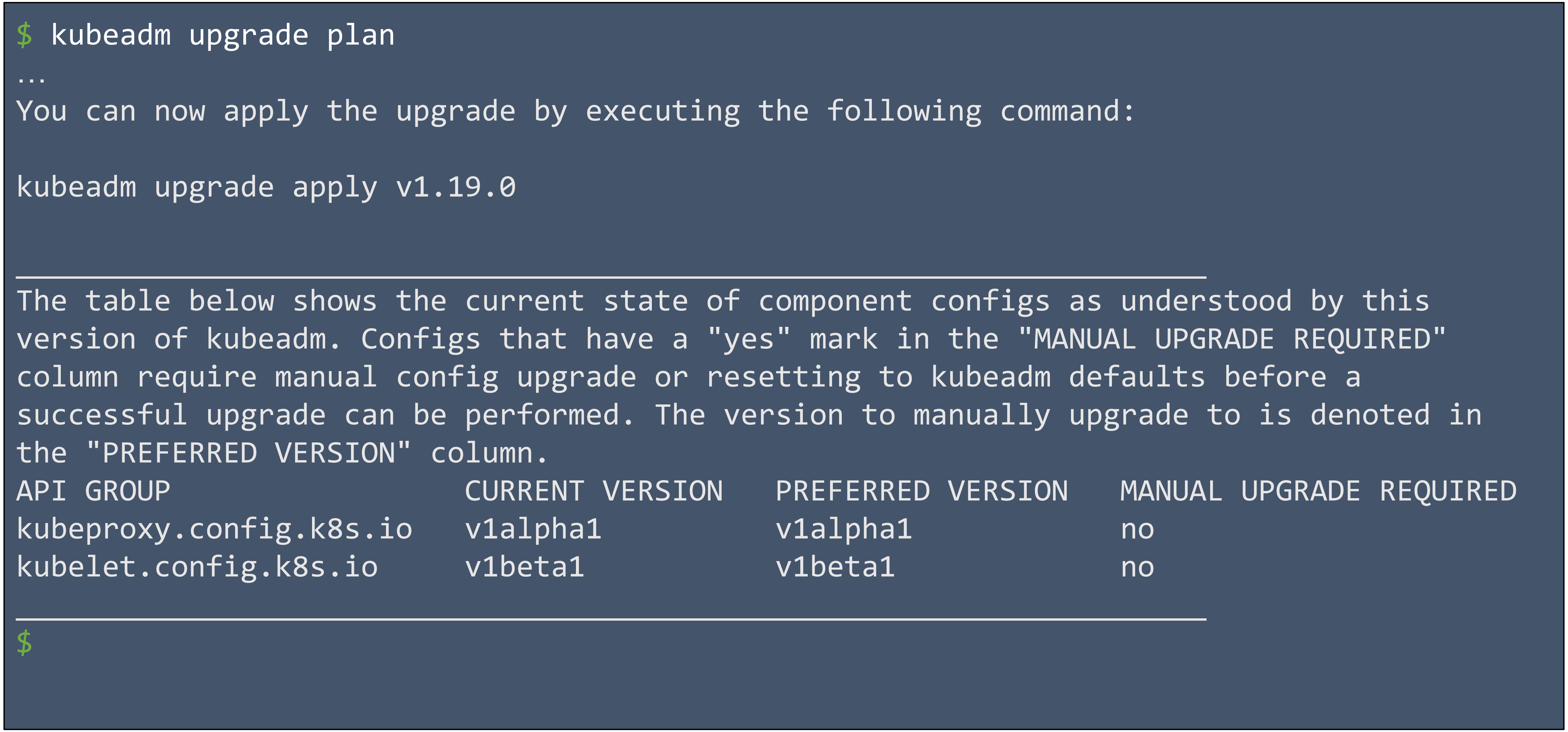
```
$ sudo kubeadm init --experimental-patches ~/patches
```

```
$ sudo kubeadm join --experimental-patches ~/patches ...
```

```
$ sudo kubeadm upgrade apply --experimental-patches ~/patches v1.19.0
```

```
$ sudo kubeadm upgrade node --experimental-patches ~/patches
```

Component Config Updates



```
$ kubeadm upgrade plan
...
You can now apply the upgrade by executing the following command:

kubeadm upgrade apply v1.19.0


_____________________________________________________________________
The table below shows the current state of component configs as understood by this
version of kubeadm. Configs that have a "yes" mark in the "MANUAL UPGRADE REQUIRED"
column require manual config upgrade or resetting to kubeadm defaults before a
successful upgrade can be performed. The version to manually upgrade to is denoted in
the "PREFERRED VERSION" column.
API GROUP                 CURRENT VERSION   PREFERRED VERSION   MANUAL UPGRADE REQUIRED
kubeproxy.config.k8s.io   v1alpha1          v1alpha1            no
kubelet.config.k8s.io     v1beta1           v1beta1             no

_____________________________________________________________________
$
```

What lays ahead?

Moving out of tree

github.com/kubernetes/kubernetes

↓

github.com/kubernetes/kubeadm

What lays ahead?

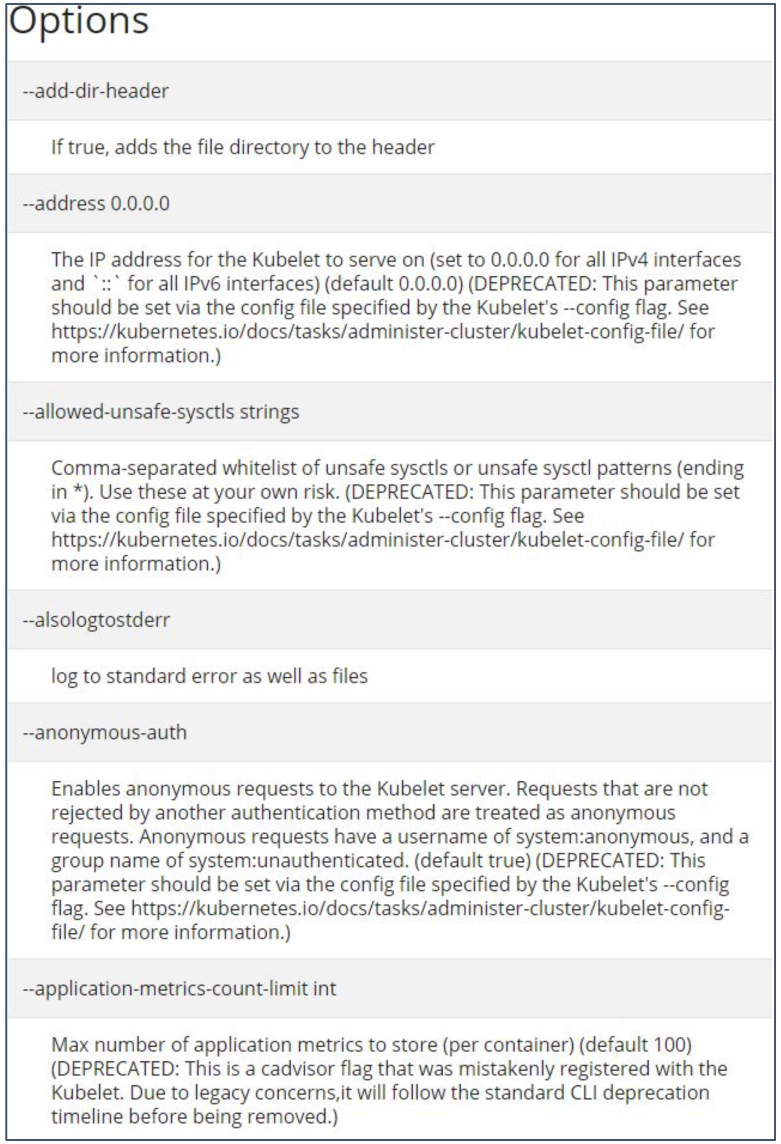
Machine readable output

```
$ sudo kubeadm token list
TOKEN                     TTL         EXPIRES                USAGES                   DESCRIPTION    EXTRA GROUPS
abcdef.0123456789abcdef   23h         2020-07-30T11:13:17Z   authentication,signing   <none>         system:bootstrappers:kubeadm:default-node-token
```

```
$ kubeadm token list -o yaml
apiVersion: output.kubeadm.k8s.io/v1alpha1
expires: "2020-07-30T11:13:17Z"
groups:
- system:bootstrappers:kubeadm:default-node-token
kind: BootstrapToken
token: abcdef.0123456789abcdef
usages:
- authentication
- signing
```

```
$ kubeadm token list -o json
{
    "kind": "BootstrapToken",
    "apiVersion": "output.kubeadm.k8s.io/v1alpha1",
    "token": "abcdef.0123456789abcdef",
    "expires": "2020-07-30T11:13:17Z",
    "usages": [
        "authentication",
        "signing"
    ],
    "groups": [
        "system:bootstrappers:kubeadm:default-node-token"
    ]
}
```

Component configs



Options

--add-dir-header

If true, adds the file directory to the header

--address 0.0.0.0

The IP address for the Kubelet to serve on (set to 0.0.0.0 for all IPv4 interfaces and `::` for all IPv6 interfaces) (default 0.0.0.0) (DEPRECATED: This parameter should be set via the config file specified by the Kubelet's --config flag. See https://kubernetes.io/docs/tasks/administer-cluster/kubelet-config-file/ for more information.)

--allowed-unsafe-sysctls strings

Comma-separated whitelist of unsafe sysctls or unsafe sysctl patterns (ending in *). Use these at your own risk. (DEPRECATED: This parameter should be set via the config file specified by the Kubelet's --config flag. See https://kubernetes.io/docs/tasks/administer-cluster/kubelet-config-file/ for more information.)

--alsologtostderr

log to standard error as well as files

--anonymous-auth

Enables anonymous requests to the Kubelet server. Requests that are not rejected by another authentication method are treated as anonymous requests. Anonymous requests have a username of system:anonymous, and a group name of system:unauthenticated. (default true) (DEPRECATED: This parameter should be set via the config file specified by the Kubelet's --config flag. See https://kubernetes.io/docs/tasks/administer-cluster/kubelet-config-file/ for more information.)

--application-metrics-count-limit int

Max number of application metrics to store (per container) (default 100) (DEPRECATED: This is a cadvisor flag that was mistakenly registered with the Kubelet. Due to legacy concerns,it will follow the standard CLI deprecation timeline before being removed.)

```
apiVersion: kubelet.config.k8s.io/v1beta1
kind: KubeletConfiguration
authentication:
  anonymous:
    enabled: false
  webhook:
    cacheTTL: 0s
    enabled: true
  x509:
    clientCAFile: /etc/kubernetes/pki/ca.crt
authorization:
  mode: Webhook
  webhook:
    cacheAuthorizedTTL: 0s
    cacheUnauthorizedTTL: 0s
clusterDNS:
- 10.96.0.10
clusterDomain: cluster.local
healthzBindAddress: 127.0.0.1
healthzPort: 10248
imageGCHighThresholdPercent: 100
rotateCertificates: true
staticPodPath: /etc/kubernetes/manifests
```

What lays ahead?

Kubeadm Configuration Changes

v1beta1 → ✗

v1beta2 → ✓

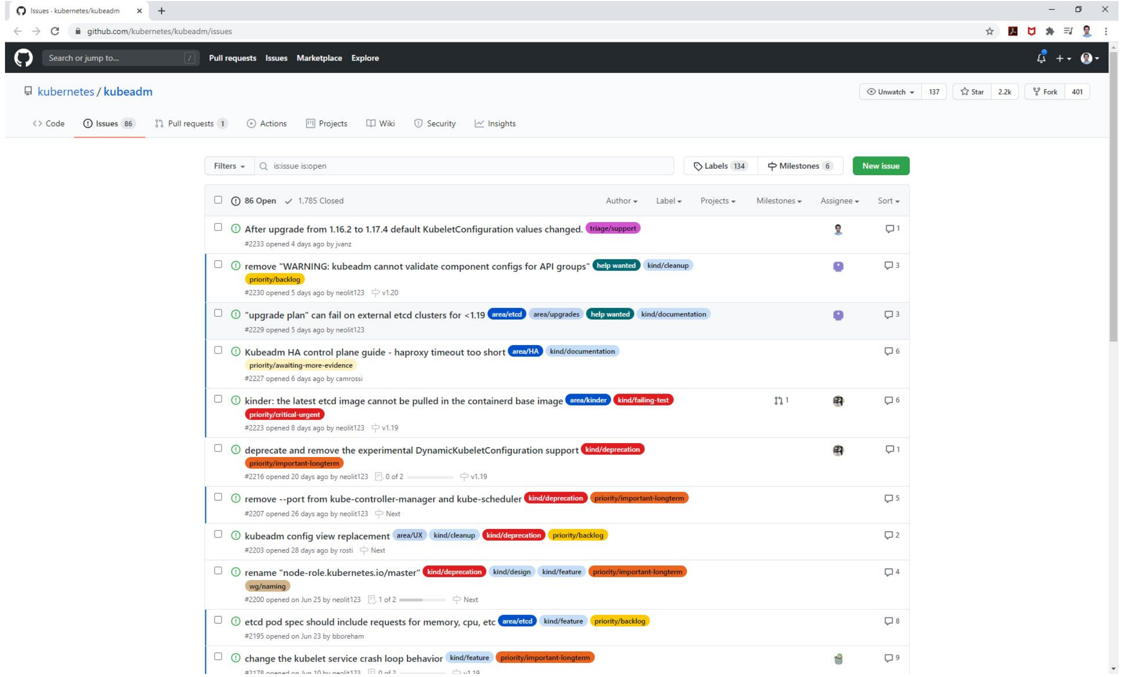
v1beta3 → ✓

What lays ahead?

Also...

- Kubeadm operator and existing clusters modification
- Cluster addons project integration
- Trimming the backlog

The Issue Tracker

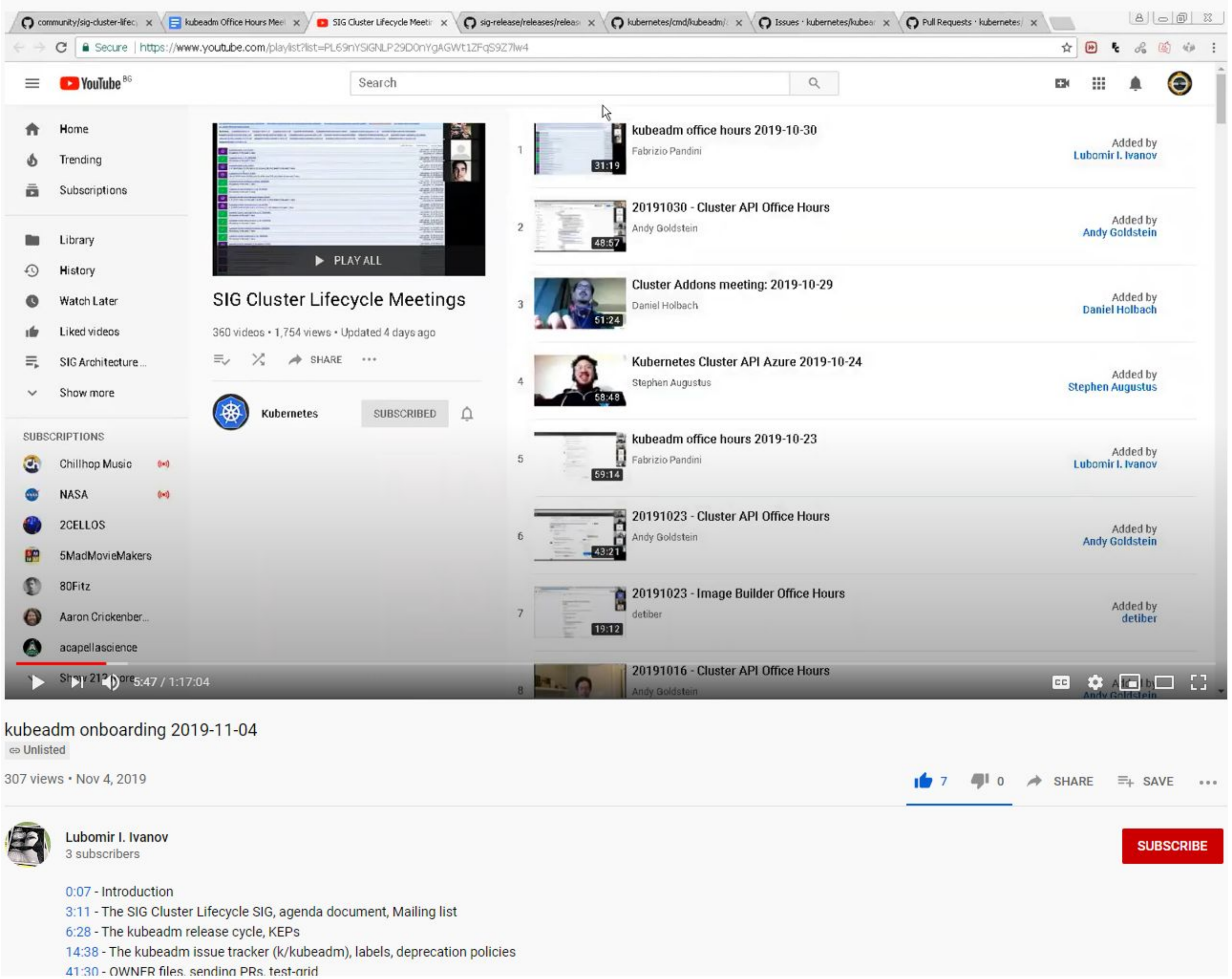


86 Open ✓ 1,785 Closed

- After upgrade from 1.16.2 to 1.17.4 default KubeletConfiguration values changed. triage/support — 1
  #2233 opened 4 days ago by jvanz
- remove "WARNING: kubeadm cannot validate component configs for API groups" help wanted kind/cleanup priority/backlog — 3
  #2230 opened 5 days ago by neolit123 v1.20
- "upgrade plan" can fail on external etcd clusters for <1.19 area/etcd area/upgrades help wanted kind/documentation — 3
  #2229 opened 5 days ago by neolit123
- Kubeadm HA control plane guide - haproxy timeout too short area/HA kind/documentation priority/awaiting-more-evidence — 6
  #2227 opened 6 days ago by camrossi
- kinder: the latest etcd image cannot be pulled in the containerd base image area/kinder kind/failing-test priority/critical-urgent — 1 — 6
  #2223 opened 8 days ago by neolit123 v1.19
- deprecate and remove the experimental DynamicKubeletConfiguration support kind/deprecation priority/important-longterm — 1
  #2216 opened 20 days ago by neolit123 0 of 2 v1.19
- remove --port from kube-controller-manager and kube-scheduler kind/deprecation priority/important-longterm — 5
  #2207 opened 26 days ago by neolit123 Next
- kubeadm config view replacement area/UX kind/cleanup kind/deprecation priority/backlog — 2
  #2203 opened 28 days ago by rosti Next
- rename "node-role.kubernetes.io/master" kind/deprecation kind/design kind/feature priority/important-longterm wg/naming — 4
  #2200 opened on Jun 25 by neolit123 1 of 2 Next
- etcd pod spec should include requests for memory, cpu, etc area/etcd kind/feature priority/backlog — 8
  #2195 opened on Jun 23 by bboreham
- change the kubelet service crash loop behavior kind/feature priority/important-longterm — 9
  #2178 opened on Jun 10 by neolit123 0 of 2 v1.19

github.com/kubernetes/kubeadm

Onboarding Video



SIG Cluster Lifecycle Meetings

360 videos • 1,754 views • Updated 4 days ago

1	kubeadm office hours 2019-10-30	Fabrizio Pandini	Added by Lubomir I. Ivanov
2	20191030 - Cluster API Office Hours	Andy Goldstein	Added by Andy Goldstein
3	Cluster Addons meeting: 2019-10-29	Daniel Holbach	Added by Daniel Holbach
4	Kubernetes Cluster API Azure 2019-10-24	Stephen Augustus	Added by Stephen Augustus
5	kubeadm office hours 2019-10-23	Fabrizio Pandini	Added by Lubomir I. Ivanov
6	20191023 - Cluster API Office Hours	Andy Goldstein	Added by Andy Goldstein
7	20191023 - Image Builder Office Hours	detiber	Added by detiber
8	20191016 - Cluster API Office Hours	Andy Goldstein	Added by Andy Goldstein

kubeadm onboarding 2019-11-04

Unlisted

307 views • Nov 4, 2019

Lubomir I. Ivanov
3 subscribers

0:07 - Introduction
3:11 - The SIG Cluster Lifecycle SIG, agenda document, Mailing list
6:28 - The kubeadm release cycle, KEPs
14:38 - The kubeadm issue tracker (k/kubeadm), labels, deprecation policies
41:30 - OWNER files, sending PRs, test-grid

Talk to us!

Slack

#kubeadm
#sig-cluster-lifecycle

on Kubernetes Slack

Meetings

Wednesdays
9:00 AM US PT

Google Group

kubernetes-sig-cluster-lifecycle

The kubeadm team

Lucas, Tim, Fabrizio, Lubomir, Ross, Rafael,
Alexander, Jason, Di Xu, Yago, SataQiu,
Yassine, Marek, Ed, Liz, Chuck, Leigh, ...

Reminder

SIG Cluster Lifecycle Survey

Kubeadm Deep Dive

KubeCon | CloudNativeCon
Europe 2020
Virtual

Q&A