Services: Control

Setup

Installation with `kubeadm init` would set up certificates, configs and initial static **kublet** manifests to start control-plane using **kublet** configured container runtime. Here we will use certificates and configs generated by **kubeadm** but will set up services directly on the host using **runit**.

Files

- Values and Path

Kubelet TLS certificate issuing

- bootstrap-tokens - generate shared secret (token), it is then used by kubelet to get actual certs from API server

Kubernetes API server

- Implementation details - API Server

```
exec chpst -u kube:kube kube-apiserver \
  --enable-bootstrap-token-auth \
  --service-account-issuer https://localhost:6443 \
  --service-cluster-ip-range 172.19.1.0/24 \
  --etcd-servers http://127.0.0.1:2379 \
  --enable-admission-plugins
NamespaceLifecycle,LimitRanger,ServiceAccount,DefaultStorageClass,DefaultTolerationSeconds,Pri
ority,ResourceQuota \
  --client-ca-file /etc/kubernetes/pki/ca.crt \
```

```
--tls-cert-file /etc/kubernetes/pki/apiserver.crt \
--tls-private-key-file /etc/kubernetes/pki/apiserver.key \
--kubelet-client-certificate /etc/kubernetes/pki/apiserver-kubelet-client.crt \
--kubelet-client-key /etc/kubernetes/pki/apiserver-kubelet-client.key \
--service-account-key-file /etc/kubernetes/pki/sa.pub \
--service-account-signing-key-file /etc/kubernetes/pki/sa.key \
--requestheader-client-ca-file /etc/kubernetes/pki/front-proxy-ca.crt \
--proxy-client-cert-file /etc/kubernetes/pki/front-proxy-client.crt \
--proxy-client-key-file /etc/kubernetes/pki/front-proxy-client.key
```

Flags generated by kuebadm init

```
kube-apiserver \
--advertise-address=192.168.50.247 \
--allow-privileged=true \
--authorization-mode=Node,RBAC \
--client-ca-file=/etc/kubernetes/pki/ca.crt \
--enable-admission-plugins=NodeRestriction \
--enable-bootstrap-token-auth=true \
--etcd-cafile=/etc/kubernetes/pki/etcd/ca.crt \
--etcd-certfile=/etc/kubernetes/pki/apiserver-etcd-client.crt \
--etcd-keyfile=/etc/kubernetes/pki/apiserver-etcd-client.key \
--etcd-servers=https://127.0.0.1:2379 \
--kubelet-client-certificate=/etc/kubernetes/pki/apiserver-kubelet-client.crt \
--kubelet-client-key=/etc/kubernetes/pki/apiserver-kubelet-client.key \
--kubelet-preferred-address-types=InternalIP,ExternalIP,Hostname \
--proxy-client-cert-file=/etc/kubernetes/pki/front-proxy-client.crt \
--proxy-client-key-file=/etc/kubernetes/pki/front-proxy-client.key \
--requestheader-allowed-names=front-proxy-client \
--requestheader-client-ca-file=/etc/kubernetes/pki/front-proxy-ca.crt \
--requestheader-extra-headers-prefix=X-Remote-Extra- \
--requestheader-group-headers=X-Remote-Group \
--requestheader-username-headers=X-Remote-User \
--secure-port=6443 \
--service-account-issuer=https://kubernetes.default.svc.cluster.local \
--service-account-key-file=/etc/kubernetes/pki/sa.pub \
--service-account-signing-key-file=/etc/kubernetes/pki/sa.key \
--service-cluster-ip-range=10.96.0.0/12 \
--tls-cert-file=/etc/kubernetes/pki/apiserver.crt \
```

```
--tls-private-key-file=/etc/kubernetes/pki/apiserver.key
```

Flag	Value	Info
`--enable-bootstrap-token-auth`		• Bootstrap tokens
`--enable-admission-plugins`	`NamespaceLifecycle,LimitRanger,ServiceAccount,DefaultStorageClass,DefaultTolerationSeconds,Priority,ResourceQuota`	• A Guide to Kubernetes Admission Controllers • Admission Controllers Reference
`--service-account-issuer`	`https://localhost:6443` - URL to the API server (load balancer in multi-server setup)	• kube-apiserver fails init. receive "--service-account-signing-key-file and --service-account-issuer are required flag" • No `ServiceAccountIssuerDiscovery` feature gate - so probably was in dev but now is always on
`--client-ca-file`	Enabled client authentication using certificates signed by this CA.	• X509 Client Certs

Admission plugins

Defaults

Symbol	Description
CertificateApproval	This admission controller observes requests to approve CertificateSigningRequest resources and performs additional authorization checks to ensure the approving user has permission to **approve** certificate requests with the `spec.signerName` requested on the CertificateSigningRequest resource. See Certificate Signing Requests for more information on the permissions required to perform different actions on CertificateSigningRequest resources.

CertificateSigning	This admission controller observes updates to the `status.certificate` field of CertificateSigningRequest resources and performs an additional authorization checks to ensure the signing user has permission to **sign** certificate requests with the `spec.signerName` requested on the CertificateSigningRequest resource. See Certificate Signing Requests for more information on the permissions required to perform different actions on CertificateSigningRequest resources.
CertificateSubjectRestriction	This admission controller observes creation of CertificateSigningRequest resources that have a `spec.signerName` of `kubernetes.io/kube-apiserver-client`. It rejects any request that specifies a 'group' (or 'organization attribute') of `system:masters`.
DefaultIngressClass	This admission controller observes creation of `Ingress` objects that do not request any specific ingress class and automatically adds a default ingress class to them. This way, users that do not request any special ingress class do not need to care about them at all and they will get the default one. This admission controller does not do anything when no default ingress class is configured. When more than one ingress class is marked as default, it rejects any creation of `Ingress` with an error and an administrator must revisit their `IngressClass` objects and mark only one as default (with the annotation "ingressclass.kubernetes.io/is-default-class"). This admission controller ignores any `Ingress` updates; it acts only on creation. See the Ingress documentation for more about ingress classes and how to mark one as default.
DefaultStorageClass	This admission controller observes creation of `PersistentVolumeClaim` objects that do not request any specific storage class and automatically adds a default storage class to them. This way, users that do not request any special storage class do not need to care about them at all and they will get the default one. This admission controller does not do anything when no default storage class is configured. When more than one storage class is marked as default, it rejects any creation of `PersistentVolumeClaim` with an error and an administrator must revisit their `StorageClass` objects and mark only one as default. This admission controller ignores any `PersistentVolumeClaim` updates; it acts only on creation. See persistent volume documentation about persistent volume claims and storage classes and how to mark a storage class as default.

DefaultTolerationSeconds	This admission controller sets the default forgiveness toleration for pods to tolerate the taints `notready:NoExecute` and `unreachable:NoExecute` based on the k8s-apiserver input parameters `default-not-ready-toleration-seconds` and `default-unreachable-toleration-seconds` if the pods don't already have toleration for taints `node.kubernetes.io/not-ready:NoExecute` or `node.kubernetes.io/unreachable:NoExecute`. The default value for `default-not-ready-toleration-seconds` and `default-unreachable-toleration-seconds` is 5 minutes.
LimitRanger	This admission controller will observe the incoming request and ensure that it does not violate any of the constraints enumerated in the `LimitRange` object in a `Namespace`. If you are using `LimitRange` objects in your Kubernetes deployment, you MUST use this admission controller to enforce those constraints. LimitRanger can also be used to apply default resource requests to Pods that don't specify any; currently, the default LimitRanger applies a 0.1 CPU requirement to all Pods in the `default` namespace. See the LimitRange API reference and the example of LimitRange for more details.
MutatingAdmissionWebhook	This admission controller calls any mutating webhooks which match the request. Matching webhooks are called in serial; each one may modify the object if it desires.
NamespaceLifecycle	This admission controller enforces that a `Namespace` that is undergoing termination cannot have new objects created in it, and ensures that requests in a non-existent `Namespace` are rejected. This admission controller also prevents deletion of three system reserved namespaces `default`, `kube-system`, `kube-public`.
PersistentVolumeClaimResize	This admission controller implements additional validations for checking incoming `PersistentVolumeClaim` resize requests. Enabling the `PersistentVolumeClaimResize` admission controller is recommended. This admission controller prevents resizing of all claims by default unless a claim's `StorageClass` explicitly enables resizing by setting `allowVolumeExpansion` to `true`.
PodSecurity	The PodSecurity admission controller checks new Pods before they are admitted, determines if it should be admitted based on the requested security context and the restrictions on permitted Pod Security Standards for the namespace that the Pod would be in. See the Pod Security Admission documentation for more information.
Priority	The priority admission controller uses the `priorityClassName` field and populates the integer value of the priority. If the priority class is not found, the Pod is rejected.

ResourceQuota	This admission controller will observe the incoming request and ensure that it does not violate any of the constraints enumerated in the `ResourceQuota` object in a `Namespace`. If you are using `ResourceQuota` objects in your Kubernetes deployment, you MUST use this admission controller to enforce quota constraints. See the ResourceQuota API reference and the example of Resource Quota for more details.
RuntimeClass	If you define a RuntimeClass with Pod overhead configured, this admission controller checks incoming Pods. When enabled, this admission controller rejects any Pod create requests that have the overhead already set. For Pods that have a RuntimeClass configured and selected in their `.spec`, this admission controller sets `.spec.overhead` in the Pod based on the value defined in the corresponding RuntimeClass. See also Pod Overhead for more information.
ServiceAccount	This admission controller implements automation for serviceAccounts. The Kubernetes project strongly recommends enabling this admission controller. You should enable this admission controller if you intend to make any use of Kubernetes `ServiceAccount` objects.
StorageObjectInUseProtection	The `StorageObjectInUseProtection` plugin adds the `kubernetes.io/pvc-protection` or `kubernetes.io/pv-protection` finalizers to newly created Persistent Volume Claims (PVCs) or Persistent Volumes (PV). In case a user deletes a PVC or PV the PVC or PV is not removed until the finalizer is removed from the PVC or PV by PVC or PV Protection Controller. Refer to the Storage Object in Use Protection for more detailed information.
TaintNodesByCondition	This admission controller taints newly created Nodes as `NotReady` and `NoSchedule`. That tainting avoids a race condition that could cause Pods to be scheduled on new Nodes before their taints were updated to accurately reflect their reported conditions.
ValidatingAdmissionPolicy	This admission controller implements the CEL validation for incoming matched requests. It is enabled when both feature gate `validatingadmissionpolicy` and `admissionregistration.k8s.io/v1alpha1` group/version are enabled. If any of the ValidatingAdmissionPolicy fails, the request fails.
ValidatingAdmissionWebhook	This admission controller calls any validating webhooks which match the request. Matching webhooks are called in parallel; if any of them rejects the request, the request fails. This admission controller only runs in the validation phase; the webhooks it calls may not mutate the object, as opposed to the webhooks called by the `MutatingAdmissionWebhook` admission controller.

Appendix

Usage: kube-apiserver

```
The Kubernetes API server validates and configures data
for the api objects which include pods, services, replicationcontrollers, and
others. The API Server services REST operations and provides the frontend to the
cluster's shared state through which all other components interact.

Usage:
  kube-apiserver [flags]

Generic flags:

      --advertise-address ip
                The IP address on which to advertise the apiserver to members of the
cluster. This
                address must be reachable by the rest of the cluster. If blank, the --
bind-address will
                be used. If --bind-address is unspecified, the host's default interface
will be used.
      --cloud-provider-gce-l7lb-src-cidrs cidrs
                CIDRs opened in GCE firewall for L7 LB traffic proxy & health checks
(default
                130.211.0.0/22,35.191.0.0/16)
      --cors-allowed-origins strings
                List of allowed origins for CORS, comma separated.  An allowed origin can
be a regular
                expression to support subdomain matching. If this list is empty CORS will
not be enabled.
      --default-not-ready-toleration-seconds int
                Indicates the tolerationSeconds of the toleration for notReady:NoExecute
that is added
                by default to every pod that does not already have such a toleration.
(default 300)
      --default-unreachable-toleration-seconds int
                Indicates the tolerationSeconds of the toleration for
```

```
unreachable:NoExecute that is
                added by default to every pod that does not already have such a
toleration. (default 300)
      --enable-priority-and-fairness
                If true and the APIPriorityAndFairness feature gate is enabled, replace
the
                max-in-flight handler with an enhanced one that queues and dispatches with
priority and
                fairness (default true)
      --external-hostname string
                The hostname to use when generating externalized URLs for this master
(e.g. Swagger API
                Docs or OpenID Discovery).
      --feature-gates mapStringBool
                A set of key=value pairs that describe feature gates for
alpha/experimental features.
                Options are:
                APIListChunking=true|false (BETA - default=true)
                APIPriorityAndFairness=true|false (BETA - default=true)
                APIResponseCompression=true|false (BETA - default=true)
                APISelfSubjectReview=true|false (ALPHA - default=false)
                APIServerIdentity=true|false (BETA - default=true)
                APIServerTracing=true|false (ALPHA - default=false)
                AggregatedDiscoveryEndpoint=true|false (ALPHA - default=false)
                AllAlpha=true|false (ALPHA - default=false)
                AllBeta=true|false (BETA - default=false)
                AnyVolumeDataSource=true|false (BETA - default=true)
                AppArmor=true|false (BETA - default=true)
                CPUManagerPolicyAlphaOptions=true|false (ALPHA - default=false)
                CPUManagerPolicyBetaOptions=true|false (BETA - default=true)
                CPUManagerPolicyOptions=true|false (BETA - default=true)
                CSIMigrationPortworx=true|false (BETA - default=false)
                CSIMigrationRBD=true|false (ALPHA - default=false)
                CSINodeExpandSecret=true|false (ALPHA - default=false)
                CSIVolumeHealth=true|false (ALPHA - default=false)
                ComponentSLIs=true|false (ALPHA - default=false)
                ContainerCheckpoint=true|false (ALPHA - default=false)
                ContextualLogging=true|false (ALPHA - default=false)
```

```
                CronJobTimeZone=true|false (BETA - default=true)
                CrossNamespaceVolumeDataSource=true|false (ALPHA - default=false)
                CustomCPUCFSQuotaPeriod=true|false (ALPHA - default=false)
                CustomResourceValidationExpressions=true|false (BETA - default=true)
                DisableCloudProviders=true|false (ALPHA - default=false)
                DisableKubeletCloudCredentialProviders=true|false (ALPHA - default=false)
                DownwardAPIHugePages=true|false (BETA - default=true)
                DynamicResourceAllocation=true|false (ALPHA - default=false)
                EventedPLEG=true|false (ALPHA - default=false)
                ExpandedDNSConfig=true|false (BETA - default=true)
                ExperimentalHostUserNamespaceDefaulting=true|false (BETA - default=false)
                GRPCContainerProbe=true|false (BETA - default=true)
                GracefulNodeShutdown=true|false (BETA - default=true)
                GracefulNodeShutdownBasedOnPodPriority=true|false (BETA - default=true)
                HPAContainerMetrics=true|false (ALPHA - default=false)
                HPAScaleToZero=true|false (ALPHA - default=false)
                HonorPVReclaimPolicy=true|false (ALPHA - default=false)
                IPTablesOwnershipCleanup=true|false (ALPHA - default=false)
                InTreePluginAWSUnregister=true|false (ALPHA - default=false)
                InTreePluginAzureDiskUnregister=true|false (ALPHA - default=false)
                InTreePluginAzureFileUnregister=true|false (ALPHA - default=false)
                InTreePluginGCEUnregister=true|false (ALPHA - default=false)
                InTreePluginOpenStackUnregister=true|false (ALPHA - default=false)
                InTreePluginPortworxUnregister=true|false (ALPHA - default=false)
                InTreePluginRBDUnregister=true|false (ALPHA - default=false)
                InTreePluginvSphereUnregister=true|false (ALPHA - default=false)
                JobMutableNodeSchedulingDirectives=true|false (BETA - default=true)
                JobPodFailurePolicy=true|false (BETA - default=true)
                JobReadyPods=true|false (BETA - default=true)
                KMSv2=true|false (ALPHA - default=false)
                KubeletInUserNamespace=true|false (ALPHA - default=false)
                KubeletPodResources=true|false (BETA - default=true)
                KubeletPodResourcesGetAllocatable=true|false (BETA - default=true)
                KubeletTracing=true|false (ALPHA - default=false)
                LegacyServiceAccountTokenTracking=true|false (ALPHA - default=false)
                LocalStorageCapacityIsolationFSQuotaMonitoring=true|false (ALPHA -
default=false)
                LogarithmicScaleDown=true|false (BETA - default=true)
```

```
LoggingAlphaOptions=true|false (ALPHA - default=false)
LoggingBetaOptions=true|false (BETA - default=true)
MatchLabelKeysInPodTopologySpread=true|false (ALPHA - default=false)
MaxUnavailableStatefulSet=true|false (ALPHA - default=false)
MemoryManager=true|false (BETA - default=true)
MemoryQoS=true|false (ALPHA - default=false)
MinDomainsInPodTopologySpread=true|false (BETA - default=false)
MinimizeIPTablesRestore=true|false (ALPHA - default=false)
MultiCIDRRangeAllocator=true|false (ALPHA - default=false)
NetworkPolicyStatus=true|false (ALPHA - default=false)
NodeInclusionPolicyInPodTopologySpread=true|false (BETA - default=true)
NodeOutOfServiceVolumeDetach=true|false (BETA - default=true)
NodeSwap=true|false (ALPHA - default=false)
OpenAPIEnums=true|false (BETA - default=true)
OpenAPIV3=true|false (BETA - default=true)
PDBUnhealthyPodEvictionPolicy=true|false (ALPHA - default=false)
PodAndContainerStatsFromCRI=true|false (ALPHA - default=false)
PodDeletionCost=true|false (BETA - default=true)
PodDisruptionConditions=true|false (BETA - default=true)
PodHasNetworkCondition=true|false (ALPHA - default=false)
PodSchedulingReadiness=true|false (ALPHA - default=false)
ProbeTerminationGracePeriod=true|false (BETA - default=true)
ProcMountType=true|false (ALPHA - default=false)
ProxyTerminatingEndpoints=true|false (BETA - default=true)
QOSReserved=true|false (ALPHA - default=false)
ReadWriteOncePod=true|false (ALPHA - default=false)
RecoverVolumeExpansionFailure=true|false (ALPHA - default=false)
RemainingItemCount=true|false (BETA - default=true)
RetroactiveDefaultStorageClass=true|false (BETA - default=true)
RotateKubeletServerCertificate=true|false (BETA - default=true)
SELinuxMountReadWriteOncePod=true|false (ALPHA - default=false)
SeccompDefault=true|false (BETA - default=true)
ServerSideFieldValidation=true|false (BETA - default=true)
SizeMemoryBackedVolumes=true|false (BETA - default=true)
StatefulSetAutoDeletePVC=true|false (ALPHA - default=false)
StatefulSetStartOrdinal=true|false (ALPHA - default=false)
StorageVersionAPI=true|false (ALPHA - default=false)
StorageVersionHash=true|false (BETA - default=true)
```

```
                TopologyAwareHints=true|false (BETA - default=true)
                TopologyManager=true|false (BETA - default=true)
                TopologyManagerPolicyAlphaOptions=true|false (ALPHA - default=false)
                TopologyManagerPolicyBetaOptions=true|false (BETA - default=false)
                TopologyManagerPolicyOptions=true|false (ALPHA - default=false)
                UserNamespacesStatelessPodsSupport=true|false (ALPHA - default=false)
                ValidatingAdmissionPolicy=true|false (ALPHA - default=false)
                VolumeCapacityPriority=true|false (ALPHA - default=false)
                WinDSR=true|false (ALPHA - default=false)
                WinOverlay=true|false (BETA - default=true)
                WindowsHostNetwork=true|false (ALPHA - default=true)
      --goaway-chance float
                To prevent HTTP/2 clients from getting stuck on a single apiserver,
randomly close a
                connection (GOAWAY). The client's other in-flight requests won't be
affected, and the
                client will reconnect, likely landing on a different apiserver after going
through the
                load balancer again. This argument sets the fraction of requests that will
be sent a
                GOAWAY. Clusters with single apiservers, or which don't use a load
balancer, should NOT
                enable this. Min is 0 (off), Max is .02 (1/50 requests); .001 (1/1000) is
a recommended
                starting point.
      --livez-grace-period duration
                This option represents the maximum amount of time it should take for
apiserver to
                complete its startup sequence and become live. From apiserver's start time
to when this
                amount of time has elapsed, /livez will assume that unfinished post-start
hooks will
                complete successfully and therefore return true.
      --max-mutating-requests-inflight int
                This and --max-requests-inflight are summed to determine the server's
total concurrency
                limit (which must be positive) if --enable-priority-and-fairness is true.
Otherwise,
```

```
                this flag limits the maximum number of mutating requests in flight, or a
zero value
                disables the limit completely. (default 200)
      --max-requests-inflight int
                This and --max-mutating-requests-inflight are summed to determine the
server's total
                concurrency limit (which must be positive) if --enable-priority-and-
fairness is true.
                Otherwise, this flag limits the maximum number of non-mutating requests in
flight, or a
                zero value disables the limit completely. (default 400)
      --min-request-timeout int
                An optional field indicating the minimum number of seconds a handler must
keep a request
                open before timing it out. Currently only honored by the watch request
handler, which
                picks a randomized value above this number as the connection timeout, to
spread out
                load. (default 1800)
      --request-timeout duration
                An optional field indicating the duration a handler must keep a request
open before
                timing it out. This is the default request timeout for requests but may be
overridden by
                flags such as --min-request-timeout for specific types of requests.
(default 1m0s)
      --shutdown-delay-duration duration
                Time to delay the termination. During that time the server keeps serving
requests
                normally. The endpoints /healthz and /livez will return success, but
/readyz immediately
                returns failure. Graceful termination starts after this delay has elapsed.
This can be
                used to allow load balancer to stop sending traffic to this server.
      --shutdown-send-retry-after
                If true the HTTP Server will continue listening until all non long running
request(s) in
                flight have been drained, during this window all incoming requests will be
```

```
rejected with
                a status code 429 and a 'Retry-After' response header, in addition
'Connection: close'
                response header is set in order to tear down the TCP connection when idle.
      --strict-transport-security-directives strings
                List of directives for HSTS, comma separated. If this list is empty, then
HSTS
                directives will not be added. Example: 'max-
age=31536000,includeSubDomains,preload'

Etcd flags:

      --delete-collection-workers int
                Number of workers spawned for DeleteCollection call. These are used to
speed up
                namespace cleanup. (default 1)
      --enable-garbage-collector
                Enables the generic garbage collector. MUST be synced with the
corresponding flag of the
                kube-controller-manager. (default true)
      --encryption-provider-config string
                The file containing configuration for encryption providers to be used for
storing
                secrets in etcd
      --encryption-provider-config-automatic-reload
                Determines if the file set by --encryption-provider-config should be
automatically
                reloaded if the disk contents change. Setting this to true disables the
ability to
                uniquely identify distinct KMS plugins via the API server healthz
endpoints.
      --etcd-cafile string
                SSL Certificate Authority file used to secure etcd communication.
      --etcd-certfile string
                SSL certification file used to secure etcd communication.
      --etcd-compaction-interval duration
                The interval of compaction requests. If 0, the compaction request from
apiserver is
```

```
                disabled. (default 5m0s)
      --etcd-count-metric-poll-period duration
                Frequency of polling etcd for number of resources per type. 0 disables the
metric
                collection. (default 1m0s)
      --etcd-db-metric-poll-interval duration
                The interval of requests to poll etcd and update metric. 0 disables the
metric
                collection (default 30s)
      --etcd-healthcheck-timeout duration
                The timeout to use when checking etcd health. (default 2s)
      --etcd-keyfile string
                SSL key file used to secure etcd communication.
      --etcd-prefix string
                The prefix to prepend to all resource paths in etcd. (default "/registry")
      --etcd-readycheck-timeout duration
                The timeout to use when checking etcd readiness (default 2s)
      --etcd-servers strings
                List of etcd servers to connect with (scheme://ip:port), comma separated.
      --etcd-servers-overrides strings
                Per-resource etcd servers overrides, comma separated. The individual
override format:
                group/resource#servers, where servers are URLs, semicolon separated. Note
that this
                applies only to resources compiled into this server binary.
      --lease-reuse-duration-seconds int
                The time in seconds that each lease is reused. A lower value could avoid
large number of
                objects reusing the same lease. Notice that a too small value may cause
performance
                problems at storage layer. (default 60)
      --storage-backend string
                The storage backend for persistence. Options: 'etcd3' (default).
      --storage-media-type string
                The media type to use to store objects in storage. Some resources or
storage backends
                may only support a specific media type and will ignore this setting.
Supported media
```

```
                types: [application/json, application/yaml,
application/vnd.kubernetes.protobuf]
                (default "application/vnd.kubernetes.protobuf")
      --watch-cache
                Enable watch caching in the apiserver (default true)
      --watch-cache-sizes strings
                Watch cache size settings for some resources (pods, nodes, etc.), comma
separated. The
                individual setting format: resource[.group]#size, where resource is
lowercase plural (no
                version), group is omitted for resources of apiVersion v1 (the legacy core
API) and
                included for others, and size is a number. This option is only meaningful
for resources
                built into the apiserver, not ones defined by CRDs or aggregated from
external servers,
                and is only consulted if the watch-cache is enabled. The only meaningful
size setting to
                supply here is zero, which means to disable watch caching for the
associated resource;
                all non-zero values are equivalent and mean to not disable watch caching
for that resource

Secure serving flags:

      --bind-address ip
                The IP address on which to listen for the --secure-port port. The
associated
                interface(s) must be reachable by the rest of the cluster, and by CLI/web
clients. If
                blank or an unspecified address (0.0.0.0 or ::), all interfaces will be
used. (default
                0.0.0.0)
      --cert-dir string
                The directory where the TLS certs are located. If --tls-cert-file and
                --tls-private-key-file are provided, this flag will be ignored. (default
                "/var/run/kubernetes")
      --http2-max-streams-per-connection int
```

```
                The limit that the server gives to clients for the maximum number of
streams in an
                HTTP/2 connection. Zero means to use golang's default.
      --permit-address-sharing
                If true, SO_REUSEADDR will be used when binding the port. This allows
binding to
                wildcard IPs like 0.0.0.0 and specific IPs in parallel, and it avoids
waiting for the
                kernel to release sockets in TIME_WAIT state. [default=false]
      --permit-port-sharing
                If true, SO_REUSEPORT will be used when binding the port, which allows
more than one
                instance to bind on the same address and port. [default=false]
      --secure-port int
                The port on which to serve HTTPS with authentication and authorization. It
cannot be
                switched off with 0. (default 6443)
      --tls-cert-file string
                File containing the default x509 Certificate for HTTPS. (CA cert, if any,
concatenated
                after server cert). If HTTPS serving is enabled, and --tls-cert-file and
                --tls-private-key-file are not provided, a self-signed certificate and key
are generated
                for the public address and saved to the directory specified by --cert-dir.
      --tls-cipher-suites strings
                Comma-separated list of cipher suites for the server. If omitted, the
default Go cipher
                suites will be used.
                Preferred values: TLS_AES_128_GCM_SHA256, TLS_AES_256_GCM_SHA384,
                TLS_CHACHA20_POLY1305_SHA256, TLS_ECDHE_ECDSA_WITH_AES_128_CBC_SHA,
                TLS_ECDHE_ECDSA_WITH_AES_128_GCM_SHA256,
TLS_ECDHE_ECDSA_WITH_AES_256_CBC_SHA,
                TLS_ECDHE_ECDSA_WITH_AES_256_GCM_SHA384,
TLS_ECDHE_ECDSA_WITH_CHACHA20_POLY1305,
                TLS_ECDHE_ECDSA_WITH_CHACHA20_POLY1305_SHA256,
TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA,
                TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256, TLS_ECDHE_RSA_WITH_AES_256_CBC_SHA,
                TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384,
```

```
TLS_ECDHE_RSA_WITH_CHACHA20_POLY1305,
                TLS_ECDHE_RSA_WITH_CHACHA20_POLY1305_SHA256, TLS_RSA_WITH_AES_128_CBC_SHA,
                TLS_RSA_WITH_AES_128_GCM_SHA256, TLS_RSA_WITH_AES_256_CBC_SHA,
                TLS_RSA_WITH_AES_256_GCM_SHA384.
                Insecure values: TLS_ECDHE_ECDSA_WITH_AES_128_CBC_SHA256,
                TLS_ECDHE_ECDSA_WITH_RC4_128_SHA, TLS_ECDHE_RSA_WITH_3DES_EDE_CBC_SHA,
                TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA256, TLS_ECDHE_RSA_WITH_RC4_128_SHA,
                TLS_RSA_WITH_3DES_EDE_CBC_SHA, TLS_RSA_WITH_AES_128_CBC_SHA256,
TLS_RSA_WITH_RC4_128_SHA.
      --tls-min-version string
                Minimum TLS version supported. Possible values: VersionTLS10,
VersionTLS11,
                VersionTLS12, VersionTLS13
      --tls-private-key-file string
                File containing the default x509 private key matching --tls-cert-file.
      --tls-sni-cert-key namedCertKey
                A pair of x509 certificate and private key file paths, optionally suffixed
with a list
                of domain patterns which are fully qualified domain names, possibly with
prefixed
                wildcard segments. The domain patterns also allow IP addresses, but IPs
should only be
                used if the apiserver has visibility to the IP address requested by a
client. If no
                domain patterns are provided, the names of the certificate are extracted.
Non-wildcard
                matches trump over wildcard matches, explicit domain patterns trump over
extracted
                names. For multiple key/certificate pairs, use the --tls-sni-cert-key
multiple times.
                Examples: "example.crt,example.key" or
"foo.crt,foo.key:*.foo.com,foo.com". (default [])

Auditing flags:

      --audit-log-batch-buffer-size int
                The size of the buffer to store events before batching and writing. Only
used in batch
```

```
                mode. (default 10000)
      --audit-log-batch-max-size int
                The maximum size of a batch. Only used in batch mode. (default 1)
      --audit-log-batch-max-wait duration
                The amount of time to wait before force writing the batch that hadn't
reached the max
                size. Only used in batch mode.
      --audit-log-batch-throttle-burst int
                Maximum number of requests sent at the same moment if ThrottleQPS was not
utilized
                before. Only used in batch mode.
      --audit-log-batch-throttle-enable
                Whether batching throttling is enabled. Only used in batch mode.
      --audit-log-batch-throttle-qps float32
                Maximum average number of batches per second. Only used in batch mode.
      --audit-log-compress
                If set, the rotated log files will be compressed using gzip.
      --audit-log-format string
                Format of saved audits. "legacy" indicates 1-line text format for each
event. "json"
                indicates structured json format. Known formats are legacy,json. (default
"json")
      --audit-log-maxage int
                The maximum number of days to retain old audit log files based on the
timestamp encoded
                in their filename.
      --audit-log-maxbackup int
                The maximum number of old audit log files to retain. Setting a value of 0
will mean
                there's no restriction on the number of files.
      --audit-log-maxsize int
                The maximum size in megabytes of the audit log file before it gets
rotated.
      --audit-log-mode string
                Strategy for sending audit events. Blocking indicates sending events
should block server
                responses. Batch causes the backend to buffer and write events
asynchronously. Known
```

```
                modes are batch,blocking,blocking-strict. (default "blocking")
      --audit-log-path string
                If set, all requests coming to the apiserver will be logged to this
file.  '-' means
                standard out.
      --audit-log-truncate-enabled
                Whether event and batch truncating is enabled.
      --audit-log-truncate-max-batch-size int
                Maximum size of the batch sent to the underlying backend. Actual
serialized size can be
                several hundreds of bytes greater. If a batch exceeds this limit, it is
split into
                several batches of smaller size. (default 10485760)
      --audit-log-truncate-max-event-size int
                Maximum size of the audit event sent to the underlying backend. If the
size of an event
                is greater than this number, first request and response are removed, and
if this doesn't
                reduce the size enough, event is discarded. (default 102400)
      --audit-log-version string
                API group and version used for serializing audit events written to log.
(default
                "audit.k8s.io/v1")
      --audit-policy-file string
                Path to the file that defines the audit policy configuration.
      --audit-webhook-batch-buffer-size int
                The size of the buffer to store events before batching and writing. Only
used in batch
                mode. (default 10000)
      --audit-webhook-batch-max-size int
                The maximum size of a batch. Only used in batch mode. (default 400)
      --audit-webhook-batch-max-wait duration
                The amount of time to wait before force writing the batch that hadn't
reached the max
                size. Only used in batch mode. (default 30s)
      --audit-webhook-batch-throttle-burst int
                Maximum number of requests sent at the same moment if ThrottleQPS was not
utilized
```

```
                before. Only used in batch mode. (default 15)
      --audit-webhook-batch-throttle-enable
                Whether batching throttling is enabled. Only used in batch mode. (default
true)
      --audit-webhook-batch-throttle-qps float32
                Maximum average number of batches per second. Only used in batch mode.
(default 10)
      --audit-webhook-config-file string
                Path to a kubeconfig formatted file that defines the audit webhook
configuration.
      --audit-webhook-initial-backoff duration
                The amount of time to wait before retrying the first failed request.
(default 10s)
      --audit-webhook-mode string
                Strategy for sending audit events. Blocking indicates sending events
should block server
                responses. Batch causes the backend to buffer and write events
asynchronously. Known
                modes are batch,blocking,blocking-strict. (default "batch")
      --audit-webhook-truncate-enabled
                Whether event and batch truncating is enabled.
      --audit-webhook-truncate-max-batch-size int
                Maximum size of the batch sent to the underlying backend. Actual
serialized size can be
                several hundreds of bytes greater. If a batch exceeds this limit, it is
split into
                several batches of smaller size. (default 10485760)
      --audit-webhook-truncate-max-event-size int
                Maximum size of the audit event sent to the underlying backend. If the
size of an event
                is greater than this number, first request and response are removed, and
if this doesn't
                reduce the size enough, event is discarded. (default 102400)
      --audit-webhook-version string
                API group and version used for serializing audit events written to
webhook. (default
                "audit.k8s.io/v1")
```

```
Features flags:

      --contention-profiling
                Enable lock contention profiling, if profiling is enabled
      --profiling
                Enable profiling via web interface host:port/debug/pprof/ (default true)

Authentication flags:

      --anonymous-auth
                Enables anonymous requests to the secure port of the API 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)
      --api-audiences strings
                Identifiers of the API. The service account token authenticator will
validate that
                tokens used against the API are bound to at least one of these audiences.
If the
                --service-account-issuer flag is configured and this flag is not, this
field defaults to
                a single element list containing the issuer URL.
      --authentication-token-webhook-cache-ttl duration
                The duration to cache responses from the webhook token authenticator.
(default 2m0s)
      --authentication-token-webhook-config-file string
                File with webhook configuration for token authentication in kubeconfig
format. The API
                server will query the remote service to determine authentication for
bearer tokens.
      --authentication-token-webhook-version string
                The API version of the authentication.k8s.io TokenReview to send to and
expect from the
                webhook. (default "v1beta1")
      --client-ca-file string
                If set, any request presenting a client certificate signed by one of the
```

```
authorities in
                the client-ca-file is authenticated with an identity corresponding to the
CommonName of
                the client certificate.
      --enable-bootstrap-token-auth
                Enable to allow secrets of type 'bootstrap.kubernetes.io/token' in the
'kube-system'
                namespace to be used for TLS bootstrapping authentication.
      --oidc-ca-file string
                If set, the OpenID server's certificate will be verified by one of the
authorities in
                the oidc-ca-file, otherwise the host's root CA set will be used.
      --oidc-client-id string
                The client ID for the OpenID Connect client, must be set if oidc-issuer-
url is set.
      --oidc-groups-claim string
                If provided, the name of a custom OpenID Connect claim for specifying user
groups. The
                claim value is expected to be a string or array of strings. This flag is
experimental,
                please see the authentication documentation for further details.
      --oidc-groups-prefix string
                If provided, all groups will be prefixed with this value to prevent
conflicts with other
                authentication strategies.
      --oidc-issuer-url string
                The URL of the OpenID issuer, only HTTPS scheme will be accepted. If set,
it will be
                used to verify the OIDC JSON Web Token (JWT).
      --oidc-required-claim mapStringString
                A key=value pair that describes a required claim in the ID Token. If set,
the claim is
                verified to be present in the ID Token with a matching value. Repeat this
flag to
                specify multiple claims.
      --oidc-signing-algs strings
                Comma-separated list of allowed JOSE asymmetric signing algorithms. JWTs
with a
```

```
                supported 'alg' header values are: RS256, RS384, RS512, ES256, ES384,
ES512, PS256,
                PS384, PS512. Values are defined by RFC 7518
                https://tools.ietf.org/html/rfc7518#section-3.1. (default [RS256])
      --oidc-username-claim string
                The OpenID claim to use as the user name. Note that claims other than the
default
                ('sub') is not guaranteed to be unique and immutable. This flag is
experimental, please
                see the authentication documentation for further details. (default "sub")
      --oidc-username-prefix string
                If provided, all usernames will be prefixed with this value. If not
provided, username
                claims other than 'email' are prefixed by the issuer URL to avoid clashes.
To skip any
                prefixing, provide the value '-'.
      --requestheader-allowed-names strings
                List of client certificate common names to allow to provide usernames in
headers
                specified by --requestheader-username-headers. If empty, any client
certificate
                validated by the authorities in --requestheader-client-ca-file is allowed.
      --requestheader-client-ca-file string
                Root certificate bundle to use to verify client certificates on incoming
requests before
                trusting usernames in headers specified by --requestheader-username-
headers. WARNING:
                generally do not depend on authorization being already done for incoming
requests.
      --requestheader-extra-headers-prefix strings
                List of request header prefixes to inspect. X-Remote-Extra- is suggested.
      --requestheader-group-headers strings
                List of request headers to inspect for groups. X-Remote-Group is
suggested.
      --requestheader-username-headers strings
                List of request headers to inspect for usernames. X-Remote-User is common.
      --service-account-extend-token-expiration
                Turns on projected service account expiration extension during token
```

```
generation, which
                helps safe transition from legacy token to bound service account token
feature. If this
                flag is enabled, admission injected tokens would be extended up to 1 year
to prevent
                unexpected failure during transition, ignoring value of
                service-account-max-token-expiration. (default true)
      --service-account-issuer stringArray
                Identifier of the service account token issuer. The issuer will assert
this identifier
                in "iss" claim of issued tokens. This value is a string or URI. If this
option is not a
                valid URI per the OpenID Discovery 1.0 spec, the
ServiceAccountIssuerDiscovery feature
                will remain disabled, even if the feature gate is set to true. It is
highly recommended
                that this value comply with the OpenID spec:
                https://openid.net/specs/openid-connect-discovery-1_0.html. In practice,
this means that
                service-account-issuer must be an https URL. It is also highly recommended
that this URL
                be capable of serving OpenID discovery documents at
                {service-account-issuer}/.well-known/openid-configuration. When this flag
is specified
                multiple times, the first is used to generate tokens and all are used to
determine which
                issuers are accepted.
      --service-account-jwks-uri string
                Overrides the URI for the JSON Web Key Set in the discovery doc served at
                /.well-known/openid-configuration. This flag is useful if the discovery
docand key set
                are served to relying parties from a URL other than the API server's
external (as
                auto-detected or overridden with external-hostname).
      --service-account-key-file stringArray
                File containing PEM-encoded x509 RSA or ECDSA private or public keys, used
to verify
                ServiceAccount tokens. The specified file can contain multiple keys, and
```

```
the flag can be
                specified multiple times with different files. If unspecified, --tls-
private-key-file is
                used. Must be specified when --service-account-signing-key-file is
provided
      --service-account-lookup
                If true, validate ServiceAccount tokens exist in etcd as part of
authentication.
                (default true)
      --service-account-max-token-expiration duration
                The maximum validity duration of a token created by the service account
token issuer. If
                an otherwise valid TokenRequest with a validity duration larger than this
value is
                requested, a token will be issued with a validity duration of this value.
      --token-auth-file string
                If set, the file that will be used to secure the secure port of the API
server via token
                authentication.

Authorization flags:

      --authorization-mode strings
                Ordered list of plug-ins to do authorization on secure port. Comma-
delimited list of:
                AlwaysAllow,AlwaysDeny,ABAC,Webhook,RBAC,Node. (default [AlwaysAllow])
      --authorization-policy-file string
                File with authorization policy in json line by line format, used with
                --authorization-mode=ABAC, on the secure port.
      --authorization-webhook-cache-authorized-ttl duration
                The duration to cache 'authorized' responses from the webhook authorizer.
(default 5m0s)
      --authorization-webhook-cache-unauthorized-ttl duration
                The duration to cache 'unauthorized' responses from the webhook
authorizer. (default 30s)
      --authorization-webhook-config-file string
                File with webhook configuration in kubeconfig format, used with
                --authorization-mode=Webhook. The API server will query the remote service
```

```
to determine
                access on the API server's secure port.
      --authorization-webhook-version string
                The API version of the authorization.k8s.io SubjectAccessReview to send to
and expect
                from the webhook. (default "v1beta1")

Cloud provider flags:

      --cloud-config string
                The path to the cloud provider configuration file. Empty string for no
configuration file.
      --cloud-provider string
                The provider for cloud services. Empty string for no provider.

API enablement flags:

      --runtime-config mapStringString
                A set of key=value pairs that enable or disable built-in APIs. Supported
options are:
                v1=true|false for the core API group
                <group>/<version>=true|false for a specific API group and version (e.g.
apps/v1=true)
                api/all=true|false controls all API versions
                api/ga=true|false controls all API versions of the form v[0-9]+
                api/beta=true|false controls all API versions of the form v[0-9]+beta[0-
9]+
                api/alpha=true|false controls all API versions of the form v[0-9]+alpha[0-
9]+
                api/legacy is deprecated, and will be removed in a future version

Egress selector flags:

      --egress-selector-config-file string
                File with apiserver egress selector configuration.

Admission flags:
```

```
      --admission-control strings
                Admission is divided into two phases. In the first phase, only mutating
admission
                plugins run. In the second phase, only validating admission plugins run.
The names in
                the below list may represent a validating plugin, a mutating plugin, or
both. The order
                of plugins in which they are passed to this flag does not matter. Comma-
delimited list
                of: AlwaysAdmit, AlwaysDeny, AlwaysPullImages, CertificateApproval,
CertificateSigning,
                CertificateSubjectRestriction, DefaultIngressClass, DefaultStorageClass,
                DefaultTolerationSeconds, DenyServiceExternalIPs, EventRateLimit,
                ExtendedResourceToleration, ImagePolicyWebhook,
LimitPodHardAntiAffinityTopology,
                LimitRanger, MutatingAdmissionWebhook, NamespaceAutoProvision,
NamespaceExists,
                NamespaceLifecycle, NodeRestriction, OwnerReferencesPermissionEnforcement,
                PersistentVolumeClaimResize, PersistentVolumeLabel, PodNodeSelector,
PodSecurity,
                PodTolerationRestriction, Priority, ResourceQuota, RuntimeClass,
SecurityContextDeny,
                ServiceAccount, StorageObjectInUseProtection, TaintNodesByCondition,
                ValidatingAdmissionPolicy, ValidatingAdmissionWebhook. (DEPRECATED: Use
                --enable-admission-plugins or --disable-admission-plugins instead. Will be
removed in a
                future version.)
      --admission-control-config-file string
                File with admission control configuration.
      --disable-admission-plugins strings
                admission plugins that should be disabled although they are in the default
enabled
                plugins list (NamespaceLifecycle, LimitRanger, ServiceAccount,
TaintNodesByCondition,
                PodSecurity, Priority, DefaultTolerationSeconds, DefaultStorageClass,
                StorageObjectInUseProtection, PersistentVolumeClaimResize, RuntimeClass,
                CertificateApproval, CertificateSigning, CertificateSubjectRestriction,
                DefaultIngressClass, MutatingAdmissionWebhook, ValidatingAdmissionPolicy,
```

```
                ValidatingAdmissionWebhook, ResourceQuota). Comma-delimited list of
admission plugins:
                AlwaysAdmit, AlwaysDeny, AlwaysPullImages, CertificateApproval,
CertificateSigning,
                CertificateSubjectRestriction, DefaultIngressClass, DefaultStorageClass,
                DefaultTolerationSeconds, DenyServiceExternalIPs, EventRateLimit,
                ExtendedResourceToleration, ImagePolicyWebhook,
LimitPodHardAntiAffinityTopology,
                LimitRanger, MutatingAdmissionWebhook, NamespaceAutoProvision,
NamespaceExists,
                NamespaceLifecycle, NodeRestriction, OwnerReferencesPermissionEnforcement,
                PersistentVolumeClaimResize, PersistentVolumeLabel, PodNodeSelector,
PodSecurity,
                PodTolerationRestriction, Priority, ResourceQuota, RuntimeClass,
SecurityContextDeny,
                ServiceAccount, StorageObjectInUseProtection, TaintNodesByCondition,
                ValidatingAdmissionPolicy, ValidatingAdmissionWebhook. The order of
plugins in this flag
                does not matter.
      --enable-admission-plugins strings
                admission plugins that should be enabled in addition to default enabled
ones
                (NamespaceLifecycle, LimitRanger, ServiceAccount, TaintNodesByCondition,
PodSecurity,
                Priority, DefaultTolerationSeconds, DefaultStorageClass,
StorageObjectInUseProtection,
                PersistentVolumeClaimResize, RuntimeClass, CertificateApproval,
CertificateSigning,
                CertificateSubjectRestriction, DefaultIngressClass,
MutatingAdmissionWebhook,
                ValidatingAdmissionPolicy, ValidatingAdmissionWebhook, ResourceQuota).
Comma-delimited
                list of admission plugins: AlwaysAdmit, AlwaysDeny, AlwaysPullImages,
                CertificateApproval, CertificateSigning, CertificateSubjectRestriction,
                DefaultIngressClass, DefaultStorageClass, DefaultTolerationSeconds,
                DenyServiceExternalIPs, EventRateLimit, ExtendedResourceToleration,
ImagePolicyWebhook,
                LimitPodHardAntiAffinityTopology, LimitRanger, MutatingAdmissionWebhook,
```

```
                NamespaceAutoProvision, NamespaceExists, NamespaceLifecycle,
NodeRestriction,
                OwnerReferencesPermissionEnforcement, PersistentVolumeClaimResize,
                PersistentVolumeLabel, PodNodeSelector, PodSecurity,
PodTolerationRestriction, Priority,
                ResourceQuota, RuntimeClass, SecurityContextDeny, ServiceAccount,
                StorageObjectInUseProtection, TaintNodesByCondition,
ValidatingAdmissionPolicy,
                ValidatingAdmissionWebhook. The order of plugins in this flag does not
matter.

Metrics flags:

      --allow-metric-labels stringToString
                The map from metric-label to value allow-list of this label. The key's
format is
                <MetricName>,<LabelName>. The value's format is
<allowed_value>,<allowed_value>...e.g.
                metric1,label1='v1,v2,v3', metric1,label2='v1,v2,v3'
metric2,label1='v1,v2,v3'. (default [])
      --disabled-metrics strings
                This flag provides an escape hatch for misbehaving metrics. You must
provide the fully
                qualified metric name in order to disable it. Disclaimer: disabling
metrics is higher in
                precedence than showing hidden metrics.
      --show-hidden-metrics-for-version string
                The previous version for which you want to show hidden metrics. Only the
previous minor
                version is meaningful, other values will not be allowed. The format is
<major>.<minor>,
                e.g.: '1.16'. The purpose of this format is make sure you have the
opportunity to notice
                if the next release hides additional metrics, rather than being surprised
when they are
                permanently removed in the release after that.

Logs flags:
```

```
      --log-flush-frequency duration
                Maximum number of seconds between log flushes (default 5s)
      --log-json-info-buffer-size quantity
                [Alpha] In JSON format with split output streams, the info messages can be
buffered for
                a while to increase performance. The default value of zero bytes disables
buffering. The
                size can be specified as number of bytes (512), multiples of 1000 (1K),
multiples of
                1024 (2Ki), or powers of those (3M, 4G, 5Mi, 6Gi). Enable the
LoggingAlphaOptions
                feature gate to use this.
      --log-json-split-stream
                [Alpha] In JSON format, write error messages to stderr and info messages
to stdout. The
                default is to write a single stream to stdout. Enable the
LoggingAlphaOptions feature
                gate to use this.
      --logging-format string
                Sets the log format. Permitted formats: "json" (gated by
LoggingBetaOptions), "text".
                (default "text")
  -v, --v Level
                number for the log level verbosity
      --vmodule pattern=N,...
                comma-separated list of pattern=N settings for file-filtered logging (only
works for
                text log format)

Traces flags:

      --tracing-config-file string
                File with apiserver tracing configuration.

Misc flags:

      --aggregator-reject-forwarding-redirect
```

```
                Aggregator reject forwarding redirect response back to client. (default
true)
      --allow-privileged
                If true, allow privileged containers. [default=false]
      --enable-aggregator-routing
                Turns on aggregator routing requests to endpoints IP rather than cluster
IP.
      --endpoint-reconciler-type string
                Use an endpoint reconciler (master-count, lease, none) master-count is
deprecated, and
                will be removed in a future version. (default "lease")
      --event-ttl duration
                Amount of time to retain events. (default 1h0m0s)
      --kubelet-certificate-authority string
                Path to a cert file for the certificate authority.
      --kubelet-client-certificate string
                Path to a client cert file for TLS.
      --kubelet-client-key string
                Path to a client key file for TLS.
      --kubelet-preferred-address-types strings
                List of the preferred NodeAddressTypes to use for kubelet connections.
(default
                [Hostname,InternalDNS,InternalIP,ExternalDNS,ExternalIP])
      --kubelet-timeout duration
                Timeout for kubelet operations. (default 5s)
      --kubernetes-service-node-port int
                If non-zero, the Kubernetes master service (which apiserver
creates/maintains) will be
                of type NodePort, using this as the value of the port. If zero, the
Kubernetes master
                service will be of type ClusterIP.
      --max-connection-bytes-per-sec int
                If non-zero, throttle each user connection to this number of bytes/sec.
Currently only
                applies to long-running requests.
      --proxy-client-cert-file string
                Client certificate used to prove the identity of the aggregator or kube-
apiserver when
```

```
                it must call out during a request. This includes proxying requests to a
user api-server
                and calling out to webhook admission plugins. It is expected that this
cert includes a
                signature from the CA in the --requestheader-client-ca-file flag. That CA
is published
                in the 'extension-apiserver-authentication' configmap in the kube-system
namespace.
                Components receiving calls from kube-aggregator should use that CA to
perform their half
                of the mutual TLS verification.
      --proxy-client-key-file string
                Private key for the client certificate used to prove the identity of the
aggregator or
                kube-apiserver when it must call out during a request. This includes
proxying requests
                to a user api-server and calling out to webhook admission plugins.
      --service-account-signing-key-file string
                Path to the file that contains the current private key of the service
account token
                issuer. The issuer will sign issued ID tokens with this private key.
      --service-cluster-ip-range string
                A CIDR notation IP range from which to assign service cluster IPs. This
must not overlap
                with any IP ranges assigned to nodes or pods. Max of two dual-stack CIDRs
is allowed.
      --service-node-port-range portRange
                A port range to reserve for services with NodePort visibility.  This must
not overlap
                with the ephemeral port range on nodes.  Example: '30000-32767'. Inclusive
at both ends
                of the range. (default 30000-32767)

Global flags:

  -h, --help
                help for kube-apiserver
      --version version[=true]
```

```
                Print version information and quit
```

Kubernetes controller

```
exec chpst -u kube:kube kube-controller-manager \
--kubeconfig /etc/kubernetes/controller-manager.conf \
--client-ca-file /etc/kubernetes/pki/ca.crt \
--tls-cert-file /etc/kubernetes/pki/apiserver.crt \
--tls-private-key-file /etc/kubernetes/pki/apiserver.key \
--requestheader-client-ca-file /etc/kubernetes/pki/front-proxy-ca.crt \
--cluster-signing-cert-file /etc/kubernetes/pki/ca.crt \
--cluster-signing-key-file /etc/kubernetes/pki/ca.key
```

Flag	Value	Info
`--cluster-signing-cert-file` and `--cluster-signing-key-file`	`/etc/kubernetes/pki/ca.crt` `/etc/kubernetes/pki/ca.key`	• Configuring your cluster to provide signing

Flags generated by kubeadm init

```
kube-controller-manager \
--authentication-kubeconfig=/etc/kubernetes/controller-manager.conf \
--authorization-kubeconfig=/etc/kubernetes/controller-manager.conf \
--bind-address=127.0.0.1 \
--client-ca-file=/etc/kubernetes/pki/ca.crt \
--cluster-name=kubernetes \
--cluster-signing-cert-file=/etc/kubernetes/pki/ca.crt \
--cluster-signing-key-file=/etc/kubernetes/pki/ca.key \
--controllers=*,bootstrapsigner,tokencleaner \
--kubeconfig=/etc/kubernetes/controller-manager.conf \
--leader-elect=true \
--requestheader-client-ca-file=/etc/kubernetes/pki/front-proxy-ca.crt \
--root-ca-file=/etc/kubernetes/pki/ca.crt \
--service-account-private-key-file=/etc/kubernetes/pki/sa.key \
--use-service-account-credentials=true
```

Usage: kube-controller-manager

```
The Kubernetes controller manager is a daemon that embeds
the core control loops shipped with Kubernetes. In applications of robotics and
automation, a control loop is a non-terminating loop that regulates the state of
the system. In Kubernetes, a controller is a control loop that watches the shared
state of the cluster through the apiserver and makes changes attempting to move the
current state towards the desired state. Examples of controllers that ship with
Kubernetes today are the replication controller, endpoints controller, namespace
controller, and serviceaccounts controller.

Usage:
  kube-controller-manager [flags]

Debugging flags:

      --contention-profiling
                Enable lock contention profiling, if profiling is enabled
      --profiling
                Enable profiling via web interface host:port/debug/pprof/ (default true)

Leader-migration flags:

      --enable-leader-migration
                Whether to enable controller leader migration.
      --leader-migration-config string
                Path to the config file for controller leader migration, or empty to use
the value that reflects default configuration of the controller manager. The config file
should be of type LeaderMigrationConfiguration, group controllermanager.config.k8s.io,
                version v1alpha1.

Generic flags:

      --allocate-node-cidrs
                Should CIDRs for Pods be allocated and set on the cloud provider.
      --cidr-allocator-type string
                Type of CIDR allocator to use (default "RangeAllocator")
      --cloud-config string
                The path to the cloud provider configuration file. Empty string for no
```

```
configuration file.
      --cloud-provider string
                The provider for cloud services. Empty string for no provider.
      --cluster-cidr string
                CIDR Range for Pods in cluster. Requires --allocate-node-cidrs to be true
      --cluster-name string
                The instance prefix for the cluster. (default "kubernetes")
      --configure-cloud-routes
                Should CIDRs allocated by allocate-node-cidrs be configured on the cloud
provider. (default true)
      --controller-start-interval duration
                Interval between starting controller managers.
      --controllers strings
                A list of controllers to enable. '*' enables all on-by-default
controllers, 'foo' enables the controller named 'foo', '-foo' disables the controller
named 'foo'.
                All controllers: attachdetach, bootstrapsigner, cloud-node-lifecycle,
clusterrole-aggregation, cronjob, csrapproving, csrcleaner, csrsigning, daemonset,
deployment, disruption, endpoint, endpointslice, endpointslicemirroring, ephemeral-volume,
                garbagecollector, horizontalpodautoscaling, job, namespace, nodeipam,
nodelifecycle, persistentvolume-binder, persistentvolume-expander, podgc, pv-protection,
pvc-protection, replicaset, replicationcontroller, resourcequota, root-ca-cert-publisher,
                route, service, serviceaccount, serviceaccount-token, statefulset,
tokencleaner, ttl, ttl-after-finished
                Disabled-by-default controllers: bootstrapsigner, tokencleaner (default
[*])
      --external-cloud-volume-plugin string
                The plugin to use when cloud provider is set to external. Can be empty,
should only be set when cloud-provider is external. Currently used to allow node and
volume controllers to work for in tree cloud providers.
      --feature-gates mapStringBool
                A set of key=value pairs that describe feature gates for
alpha/experimental features. Options are:
                APIListChunking=true|false (BETA - default=true)
                APIPriorityAndFairness=true|false (BETA - default=true)
                APIResponseCompression=true|false (BETA - default=true)
                APISelfSubjectReview=true|false (ALPHA - default=false)
                APIServerIdentity=true|false (BETA - default=true)
                APIServerTracing=true|false (ALPHA - default=false)
```

```
AggregatedDiscoveryEndpoint=true|false (ALPHA - default=false)
AllAlpha=true|false (ALPHA - default=false)
AllBeta=true|false (BETA - default=false)
AnyVolumeDataSource=true|false (BETA - default=true)
AppArmor=true|false (BETA - default=true)
CPUManagerPolicyAlphaOptions=true|false (ALPHA - default=false)
CPUManagerPolicyBetaOptions=true|false (BETA - default=true)
CPUManagerPolicyOptions=true|false (BETA - default=true)
CSIMigrationPortworx=true|false (BETA - default=false)
CSIMigrationRBD=true|false (ALPHA - default=false)
CSINodeExpandSecret=true|false (ALPHA - default=false)
CSIVolumeHealth=true|false (ALPHA - default=false)
ComponentSLIs=true|false (ALPHA - default=false)
ContainerCheckpoint=true|false (ALPHA - default=false)
ContextualLogging=true|false (ALPHA - default=false)
CronJobTimeZone=true|false (BETA - default=true)
CrossNamespaceVolumeDataSource=true|false (ALPHA - default=false)
CustomCPUCFSQuotaPeriod=true|false (ALPHA - default=false)
CustomResourceValidationExpressions=true|false (BETA - default=true)
DisableCloudProviders=true|false (ALPHA - default=false)
DisableKubeletCloudCredentialProviders=true|false (ALPHA - default=false)
DownwardAPIHugePages=true|false (BETA - default=true)
DynamicResourceAllocation=true|false (ALPHA - default=false)
EventedPLEG=true|false (ALPHA - default=false)
ExpandedDNSConfig=true|false (BETA - default=true)
ExperimentalHostUserNamespaceDefaulting=true|false (BETA - default=false)
GRPCContainerProbe=true|false (BETA - default=true)
GracefulNodeShutdown=true|false (BETA - default=true)
GracefulNodeShutdownBasedOnPodPriority=true|false (BETA - default=true)
HPAContainerMetrics=true|false (ALPHA - default=false)
HPAScaleToZero=true|false (ALPHA - default=false)
HonorPVReclaimPolicy=true|false (ALPHA - default=false)
IPTablesOwnershipCleanup=true|false (ALPHA - default=false)
InTreePluginAWSUnregister=true|false (ALPHA - default=false)
InTreePluginAzureDiskUnregister=true|false (ALPHA - default=false)
InTreePluginAzureFileUnregister=true|false (ALPHA - default=false)
InTreePluginGCEUnregister=true|false (ALPHA - default=false)
InTreePluginOpenStackUnregister=true|false (ALPHA - default=false)
InTreePluginPortworxUnregister=true|false (ALPHA - default=false)
```

```
                InTreePluginRBDUnregister=true|false (ALPHA - default=false)
                InTreePluginvSphereUnregister=true|false (ALPHA - default=false)
                JobMutableNodeSchedulingDirectives=true|false (BETA - default=true)
                JobPodFailurePolicy=true|false (BETA - default=true)
                JobReadyPods=true|false (BETA - default=true)
                KMSv2=true|false (ALPHA - default=false)
                KubeletInUserNamespace=true|false (ALPHA - default=false)
                KubeletPodResources=true|false (BETA - default=true)
                KubeletPodResourcesGetAllocatable=true|false (BETA - default=true)
                KubeletTracing=true|false (ALPHA - default=false)
                LegacyServiceAccountTokenTracking=true|false (ALPHA - default=false)
                LocalStorageCapacityIsolationFSQuotaMonitoring=true|false (ALPHA -
default=false)
                LogarithmicScaleDown=true|false (BETA - default=true)
                LoggingAlphaOptions=true|false (ALPHA - default=false)
                LoggingBetaOptions=true|false (BETA - default=true)
                MatchLabelKeysInPodTopologySpread=true|false (ALPHA - default=false)
                MaxUnavailableStatefulSet=true|false (ALPHA - default=false)
                MemoryManager=true|false (BETA - default=true)
                MemoryQoS=true|false (ALPHA - default=false)
                MinDomainsInPodTopologySpread=true|false (BETA - default=false)
                MinimizeIPTablesRestore=true|false (ALPHA - default=false)
                MultiCIDRRangeAllocator=true|false (ALPHA - default=false)
                NetworkPolicyStatus=true|false (ALPHA - default=false)
                NodeInclusionPolicyInPodTopologySpread=true|false (BETA - default=true)
                NodeOutOfServiceVolumeDetach=true|false (BETA - default=true)
                NodeSwap=true|false (ALPHA - default=false)
                OpenAPIEnums=true|false (BETA - default=true)
                OpenAPIV3=true|false (BETA - default=true)
                PDBUnhealthyPodEvictionPolicy=true|false (ALPHA - default=false)
                PodAndContainerStatsFromCRI=true|false (ALPHA - default=false)
                PodDeletionCost=true|false (BETA - default=true)
                PodDisruptionConditions=true|false (BETA - default=true)
                PodHasNetworkCondition=true|false (ALPHA - default=false)
                PodSchedulingReadiness=true|false (ALPHA - default=false)
                ProbeTerminationGracePeriod=true|false (BETA - default=true)
                ProcMountType=true|false (ALPHA - default=false)
                ProxyTerminatingEndpoints=true|false (BETA - default=true)
                QOSReserved=true|false (ALPHA - default=false)
```

```
                ReadWriteOncePod=true|false (ALPHA - default=false)
                RecoverVolumeExpansionFailure=true|false (ALPHA - default=false)
                RemainingItemCount=true|false (BETA - default=true)
                RetroactiveDefaultStorageClass=true|false (BETA - default=true)
                RotateKubeletServerCertificate=true|false (BETA - default=true)
                SELinuxMountReadWriteOncePod=true|false (ALPHA - default=false)
                SeccompDefault=true|false (BETA - default=true)
                ServerSideFieldValidation=true|false (BETA - default=true)
                SizeMemoryBackedVolumes=true|false (BETA - default=true)
                StatefulSetAutoDeletePVC=true|false (ALPHA - default=false)
                StatefulSetStartOrdinal=true|false (ALPHA - default=false)
                StorageVersionAPI=true|false (ALPHA - default=false)
                StorageVersionHash=true|false (BETA - default=true)
                TopologyAwareHints=true|false (BETA - default=true)
                TopologyManager=true|false (BETA - default=true)
                TopologyManagerPolicyAlphaOptions=true|false (ALPHA - default=false)
                TopologyManagerPolicyBetaOptions=true|false (BETA - default=false)
                TopologyManagerPolicyOptions=true|false (ALPHA - default=false)
                UserNamespacesStatelessPodsSupport=true|false (ALPHA - default=false)
                ValidatingAdmissionPolicy=true|false (ALPHA - default=false)
                VolumeCapacityPriority=true|false (ALPHA - default=false)
                WinDSR=true|false (ALPHA - default=false)
                WinOverlay=true|false (BETA - default=true)
                WindowsHostNetwork=true|false (ALPHA - default=true)
      --kube-api-burst int32
                Burst to use while talking with kubernetes apiserver. (default 30)
      --kube-api-content-type string
                Content type of requests sent to apiserver. (default
"application/vnd.kubernetes.protobuf")
      --kube-api-qps float32
                QPS to use while talking with kubernetes apiserver. (default 20)
      --leader-elect
                Start a leader election client and gain leadership before executing the
main loop. Enable this when running replicated components for high availability. (default
true)
      --leader-elect-lease-duration duration
                The duration that non-leader candidates will wait after observing a
leadership renewal until attempting to acquire leadership of a led but unrenewed leader
slot. This is effectively the maximum duration that a leader can be stopped before it is
```

replaced by another candidate. This is only applicable if leader election is enabled. (default 15s)

--leader-elect-renew-deadline duration

The interval between attempts by the acting master to renew a leadership slot before it stops leading. This must be less than the lease duration. This is only applicable if leader election is enabled. (default 10s)

--leader-elect-resource-lock string

The type of resource object that is used for locking during leader election. Supported options are 'leases', 'endpointsleases' and 'configmapsleases'. (default "leases")

--leader-elect-resource-name string

The name of resource object that is used for locking during leader election. (default "kube-controller-manager")

--leader-elect-resource-namespace string

The namespace of resource object that is used for locking during leader election. (default "kube-system")

--leader-elect-retry-period duration

The duration the clients should wait between attempting acquisition and renewal of a leadership. This is only applicable if leader election is enabled. (default 2s)

--min-resync-period duration

The resync period in reflectors will be random between MinResyncPeriod and 2*MinResyncPeriod. (default 12h0m0s)

--node-monitor-period duration

The period for syncing NodeStatus in NodeController. (default 5s)

--route-reconciliation-period duration

The period for reconciling routes created for Nodes by cloud provider. (default 10s)

--use-service-account-credentials

If true, use individual service account credentials for each controller.

Service controller flags:

--concurrent-service-syncs int32

The number of services that are allowed to sync concurrently. Larger number = more responsive service management, but more CPU (and network) load (default 1)

Secure serving flags:

```
      --bind-address ip
                The IP address on which to listen for the --secure-port port. The
associated interface(s) must be reachable by the rest of the cluster, and by CLI/web
clients. If blank or an unspecified address (0.0.0.0 or ::), all interfaces will be used.
(default
                0.0.0.0)
      --cert-dir string
                The directory where the TLS certs are located. If --tls-cert-file and --
tls-private-key-file are provided, this flag will be ignored.
      --http2-max-streams-per-connection int
                The limit that the server gives to clients for the maximum number of
streams in an HTTP/2 connection. Zero means to use golang's default.
      --permit-address-sharing
                If true, SO_REUSEADDR will be used when binding the port. This allows
binding to wildcard IPs like 0.0.0.0 and specific IPs in parallel, and it avoids waiting
for the kernel to release sockets in TIME_WAIT state. [default=false]
      --permit-port-sharing
                If true, SO_REUSEPORT will be used when binding the port, which allows
more than one instance to bind on the same address and port. [default=false]
      --secure-port int
                The port on which to serve HTTPS with authentication and authorization. If
0, don't serve HTTPS at all. (default 10257)
      --tls-cert-file string
                File containing the default x509 Certificate for HTTPS. (CA cert, if any,
concatenated after server cert). If HTTPS serving is enabled, and --tls-cert-file and --
tls-private-key-file are not provided, a self-signed certificate and key are generated
                for the public address and saved to the directory specified by --cert-dir.
      --tls-cipher-suites strings
                Comma-separated list of cipher suites for the server. If omitted, the
default Go cipher suites will be used.
                Preferred values: TLS_AES_128_GCM_SHA256, TLS_AES_256_GCM_SHA384,
TLS_CHACHA20_POLY1305_SHA256, TLS_ECDHE_ECDSA_WITH_AES_128_CBC_SHA,
TLS_ECDHE_ECDSA_WITH_AES_128_GCM_SHA256, TLS_ECDHE_ECDSA_WITH_AES_256_CBC_SHA,
                TLS_ECDHE_ECDSA_WITH_AES_256_GCM_SHA384,
TLS_ECDHE_ECDSA_WITH_CHACHA20_POLY1305, TLS_ECDHE_ECDSA_WITH_CHACHA20_POLY1305_SHA256,
TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA, TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256,
TLS_ECDHE_RSA_WITH_AES_256_CBC_SHA,
                TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384,
TLS_ECDHE_RSA_WITH_CHACHA20_POLY1305, TLS_ECDHE_RSA_WITH_CHACHA20_POLY1305_SHA256,
```

```
TLS_RSA_WITH_AES_128_CBC_SHA, TLS_RSA_WITH_AES_128_GCM_SHA256,
TLS_RSA_WITH_AES_256_CBC_SHA, TLS_RSA_WITH_AES_256_GCM_SHA384.
                Insecure values: TLS_ECDHE_ECDSA_WITH_AES_128_CBC_SHA256,
TLS_ECDHE_ECDSA_WITH_RC4_128_SHA, TLS_ECDHE_RSA_WITH_3DES_EDE_CBC_SHA,
TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA256, TLS_ECDHE_RSA_WITH_RC4_128_SHA,
TLS_RSA_WITH_3DES_EDE_CBC_SHA,
                TLS_RSA_WITH_AES_128_CBC_SHA256, TLS_RSA_WITH_RC4_128_SHA.
      --tls-min-version string
                Minimum TLS version supported. Possible values: VersionTLS10,
VersionTLS11, VersionTLS12, VersionTLS13
      --tls-private-key-file string
                File containing the default x509 private key matching --tls-cert-file.
      --tls-sni-cert-key namedCertKey
                A pair of x509 certificate and private key file paths, optionally suffixed
with a list of domain patterns which are fully qualified domain names, possibly with
prefixed wildcard segments. The domain patterns also allow IP addresses, but IPs should
                only be used if the apiserver has visibility to the IP address requested
by a client. If no domain patterns are provided, the names of the certificate are
extracted. Non-wildcard matches trump over wildcard matches, explicit domain patterns
trump over
                extracted names. For multiple key/certificate pairs, use the --tls-sni-
cert-key multiple times. Examples: "example.crt,example.key" or
"foo.crt,foo.key:*.foo.com,foo.com". (default [])

Authentication flags:

      --authentication-kubeconfig string
                kubeconfig file pointing at the 'core' kubernetes server with enough
rights to create tokenreviews.authentication.k8s.io. This is optional. If empty, all token
requests are considered to be anonymous and no client CA is looked up in the cluster.
      --authentication-skip-lookup
                If false, the authentication-kubeconfig will be used to lookup missing
authentication configuration from the cluster.
      --authentication-token-webhook-cache-ttl duration
                The duration to cache responses from the webhook token authenticator.
(default 10s)
      --authentication-tolerate-lookup-failure
                If true, failures to look up missing authentication configuration from the
cluster are not considered fatal. Note that this can result in authentication that treats
```

```
all requests as anonymous.
      --client-ca-file string
                If set, any request presenting a client certificate signed by one of the
authorities in the client-ca-file is authenticated with an identity corresponding to the
CommonName of the client certificate.
      --requestheader-allowed-names strings
                List of client certificate common names to allow to provide usernames in
headers specified by --requestheader-username-headers. If empty, any client certificate
validated by the authorities in --requestheader-client-ca-file is allowed.
      --requestheader-client-ca-file string
                Root certificate bundle to use to verify client certificates on incoming
requests before trusting usernames in headers specified by --requestheader-username-
headers. WARNING: generally do not depend on authorization being already done for incoming
                requests.
      --requestheader-extra-headers-prefix strings
                List of request header prefixes to inspect. X-Remote-Extra- is suggested.
(default [x-remote-extra-])
      --requestheader-group-headers strings
                List of request headers to inspect for groups. X-Remote-Group is
suggested. (default [x-remote-group])
      --requestheader-username-headers strings
                List of request headers to inspect for usernames. X-Remote-User is common.
(default [x-remote-user])

Authorization flags:

      --authorization-always-allow-paths strings
                A list of HTTP paths to skip during authorization, i.e. these are
authorized without contacting the 'core' kubernetes server. (default
[/healthz,/readyz,/livez])
      --authorization-kubeconfig string
                kubeconfig file pointing at the 'core' kubernetes server with enough
rights to create subjectaccessreviews.authorization.k8s.io. This is optional. If empty,
all requests not skipped by authorization are forbidden.
      --authorization-webhook-cache-authorized-ttl duration
                The duration to cache 'authorized' responses from the webhook authorizer.
(default 10s)
      --authorization-webhook-cache-unauthorized-ttl duration
                The duration to cache 'unauthorized' responses from the webhook
```

```
authorizer. (default 10s)

Attachdetach controller flags:

      --attach-detach-reconcile-sync-period duration
                The reconciler sync wait time between volume attach detach. This duration
must be larger than one second, and increasing this value from the default may allow for
volumes to be mismatched with pods. (default 1m0s)
      --disable-attach-detach-reconcile-sync
                Disable volume attach detach reconciler sync. Disabling this may cause
volumes to be mismatched with pods. Use wisely.

Csrsigning controller flags:

      --cluster-signing-cert-file string
                Filename containing a PEM-encoded X509 CA certificate used to issue
cluster-scoped certificates.  If specified, no more specific --cluster-signing-* flag may
be specified.
      --cluster-signing-duration duration
                The max length of duration signed certificates will be given.  Individual
CSRs may request shorter certs by setting spec.expirationSeconds. (default 8760h0m0s)
      --cluster-signing-key-file string
                Filename containing a PEM-encoded RSA or ECDSA private key used to sign
cluster-scoped certificates.  If specified, no more specific --cluster-signing-* flag may
be specified.
      --cluster-signing-kube-apiserver-client-cert-file string
                Filename containing a PEM-encoded X509 CA certificate used to issue
certificates for the kubernetes.io/kube-apiserver-client signer.  If specified, --cluster-
signing-{cert,key}-file must not be set.
      --cluster-signing-kube-apiserver-client-key-file string
                Filename containing a PEM-encoded RSA or ECDSA private key used to sign
certificates for the kubernetes.io/kube-apiserver-client signer.  If specified, --cluster-
signing-{cert,key}-file must not be set.
      --cluster-signing-kubelet-client-cert-file string
                Filename containing a PEM-encoded X509 CA certificate used to issue
certificates for the kubernetes.io/kube-apiserver-client-kubelet signer.  If specified, --
cluster-signing-{cert,key}-file must not be set.
      --cluster-signing-kubelet-client-key-file string
                Filename containing a PEM-encoded RSA or ECDSA private key used to sign
```

```
certificates for the kubernetes.io/kube-apiserver-client-kubelet signer.  If specified, --
cluster-signing-{cert,key}-file must not be set.
      --cluster-signing-kubelet-serving-cert-file string
                Filename containing a PEM-encoded X509 CA certificate used to issue
certificates for the kubernetes.io/kubelet-serving signer.  If specified, --cluster-
signing-{cert,key}-file must not be set.
      --cluster-signing-kubelet-serving-key-file string
                Filename containing a PEM-encoded RSA or ECDSA private key used to sign
certificates for the kubernetes.io/kubelet-serving signer.  If specified, --cluster-
signing-{cert,key}-file must not be set.
      --cluster-signing-legacy-unknown-cert-file string
                Filename containing a PEM-encoded X509 CA certificate used to issue
certificates for the kubernetes.io/legacy-unknown signer.  If specified, --cluster-
signing-{cert,key}-file must not be set.
      --cluster-signing-legacy-unknown-key-file string
                Filename containing a PEM-encoded RSA or ECDSA private key used to sign
certificates for the kubernetes.io/legacy-unknown signer.  If specified, --cluster-
signing-{cert,key}-file must not be set.

Deployment controller flags:

      --concurrent-deployment-syncs int32
                The number of deployment objects that are allowed to sync concurrently.
Larger number = more responsive deployments, but more CPU (and network) load (default 5)

Statefulset controller flags:

      --concurrent-statefulset-syncs int32
                The number of statefulset objects that are allowed to sync concurrently.
Larger number = more responsive statefulsets, but more CPU (and network) load (default 5)

Endpoint controller flags:

      --concurrent-endpoint-syncs int32
                The number of endpoint syncing operations that will be done concurrently.
Larger number = faster endpoint updating, but more CPU (and network) load (default 5)
      --endpoint-updates-batch-period duration
                The length of endpoint updates batching period. Processing of pod changes
will be delayed by this duration to join them with potential upcoming updates and reduce
```

```
the overall number of endpoints updates. Larger number = higher endpoint programming
                latency, but lower number of endpoints revision generated

Endpointslice controller flags:

      --concurrent-service-endpoint-syncs int32
                The number of service endpoint syncing operations that will be done
concurrently. Larger number = faster endpoint slice updating, but more CPU (and network)
load. Defaults to 5. (default 5)
      --endpointslice-updates-batch-period duration
                The length of endpoint slice updates batching period. Processing of pod
changes will be delayed by this duration to join them with potential upcoming updates and
reduce the overall number of endpoints updates. Larger number = higher endpoint
                programming latency, but lower number of endpoints revision generated
      --max-endpoints-per-slice int32
                The maximum number of endpoints that will be added to an EndpointSlice.
More endpoints per slice will result in less endpoint slices, but larger resources.
Defaults to 100. (default 100)

Endpointslicemirroring controller flags:

      --mirroring-concurrent-service-endpoint-syncs int32
                The number of service endpoint syncing operations that will be done
concurrently by the EndpointSliceMirroring controller. Larger number = faster endpoint
slice updating, but more CPU (and network) load. Defaults to 5. (default 5)
      --mirroring-endpointslice-updates-batch-period duration
                The length of EndpointSlice updates batching period for
EndpointSliceMirroring controller. Processing of EndpointSlice changes will be delayed by
this duration to join them with potential upcoming updates and reduce the overall number
of EndpointSlice
                updates. Larger number = higher endpoint programming latency, but lower
number of endpoints revision generated
      --mirroring-max-endpoints-per-subset int32
                The maximum number of endpoints that will be added to an EndpointSlice by
the EndpointSliceMirroring controller. More endpoints per slice will result in less
endpoint slices, but larger resources. Defaults to 100. (default 1000)

Ephemeralvolume controller flags:
```

```
      --concurrent-ephemeralvolume-syncs int32
                The number of ephemeral volume syncing operations that will be done
concurrently. Larger number = faster ephemeral volume updating, but more CPU (and network)
load (default 5)

Garbagecollector controller flags:

      --concurrent-gc-syncs int32
                The number of garbage collector workers that are allowed to sync
concurrently. (default 20)
      --enable-garbage-collector
                Enables the generic garbage collector. MUST be synced with the
corresponding flag of the kube-apiserver. (default true)

Horizontalpodautoscaling controller flags:

      --concurrent-horizontal-pod-autoscaler-syncs int32
                The number of horizontal pod autoscaler objects that are allowed to sync
concurrently. Larger number = more responsive horizontal pod autoscaler objects
processing, but more CPU (and network) load. (default 5)
      --horizontal-pod-autoscaler-cpu-initialization-period duration
                The period after pod start when CPU samples might be skipped. (default
5m0s)
      --horizontal-pod-autoscaler-downscale-stabilization duration
                The period for which autoscaler will look backwards and not scale down
below any recommendation it made during that period. (default 5m0s)
      --horizontal-pod-autoscaler-initial-readiness-delay duration
                The period after pod start during which readiness changes will be treated
as initial readiness. (default 30s)
      --horizontal-pod-autoscaler-sync-period duration
                The period for syncing the number of pods in horizontal pod autoscaler.
(default 15s)
      --horizontal-pod-autoscaler-tolerance float
                The minimum change (from 1.0) in the desired-to-actual metrics ratio for
the horizontal pod autoscaler to consider scaling. (default 0.1)

Namespace controller flags:

      --concurrent-namespace-syncs int32
```

The number of namespace objects that are allowed to sync concurrently. Larger number = more responsive namespace termination, but more CPU (and network) load (default 10)

--namespace-sync-period duration

The period for syncing namespace life-cycle updates (default 5m0s)

Nodeipam controller flags:

--node-cidr-mask-size int32

Mask size for node cidr in cluster. Default is 24 for IPv4 and 64 for IPv6.

--node-cidr-mask-size-ipv4 int32

Mask size for IPv4 node cidr in dual-stack cluster. Default is 24.

--node-cidr-mask-size-ipv6 int32

Mask size for IPv6 node cidr in dual-stack cluster. Default is 64.

--service-cluster-ip-range string

CIDR Range for Services in cluster. Requires --allocate-node-cidrs to be true

Nodelifecycle controller flags:

--large-cluster-size-threshold int32

Number of nodes from which NodeController treats the cluster as large for the eviction logic purposes. --secondary-node-eviction-rate is implicitly overridden to 0 for clusters this size or smaller. (default 50)

--node-eviction-rate float32

Number of nodes per second on which pods are deleted in case of node failure when a zone is healthy (see --unhealthy-zone-threshold for definition of healthy/unhealthy). Zone refers to entire cluster in non-multizone clusters. (default 0.1)

--node-monitor-grace-period duration

Amount of time which we allow running Node to be unresponsive before marking it unhealthy. Must be N times more than kubelet's nodeStatusUpdateFrequency, where N means number of retries allowed for kubelet to post node status. (default 40s)

--node-startup-grace-period duration

Amount of time which we allow starting Node to be unresponsive before marking it unhealthy. (default 1m0s)

--secondary-node-eviction-rate float32

Number of nodes per second on which pods are deleted in case of node failure when a zone is unhealthy (see --unhealthy-zone-threshold for definition of

```
healthy/unhealthy). Zone refers to entire cluster in non-multizone clusters. This value is
                implicitly overridden to 0 if the cluster size is smaller than --large-
cluster-size-threshold. (default 0.01)
      --unhealthy-zone-threshold float32
                Fraction of Nodes in a zone which needs to be not Ready (minimum 3) for
zone to be treated as unhealthy.  (default 0.55)

Persistentvolume-binder controller flags:

      --enable-dynamic-provisioning
                Enable dynamic provisioning for environments that support it. (default
true)
      --enable-hostpath-provisioner
                Enable HostPath PV provisioning when running without a cloud provider.
This allows testing and development of provisioning features.  HostPath provisioning is
not supported in any way, won't work in a multi-node cluster, and should not be used for
                anything other than testing or development.
      --flex-volume-plugin-dir string
                Full path of the directory in which the flex volume plugin should search
for additional third party volume plugins. (default "/usr/libexec/kubernetes/kubelet-
plugins/volume/exec/")
      --pv-recycler-increment-timeout-nfs int32
                the increment of time added per Gi to ActiveDeadlineSeconds for an NFS
scrubber pod (default 30)
      --pv-recycler-minimum-timeout-hostpath int32
                The minimum ActiveDeadlineSeconds to use for a HostPath Recycler pod.
 This is for development and testing only and will not work in a multi-node cluster.
(default 60)
      --pv-recycler-minimum-timeout-nfs int32
                The minimum ActiveDeadlineSeconds to use for an NFS Recycler pod (default
300)
      --pv-recycler-pod-template-filepath-hostpath string
                The file path to a pod definition used as a template for HostPath
persistent volume recycling. This is for development and testing only and will not work in
a multi-node cluster.
      --pv-recycler-pod-template-filepath-nfs string
                The file path to a pod definition used as a template for NFS persistent
volume recycling
      --pv-recycler-timeout-increment-hostpath int32
```

the increment of time added per Gi to ActiveDeadlineSeconds for a HostPath scrubber pod. This is for development and testing only and will not work in a multi-node cluster. (default 30)

--pvclaimbinder-sync-period duration

The period for syncing persistent volumes and persistent volume claims (default 15s)

--volume-host-allow-local-loopback

If false, deny local loopback IPs in addition to any CIDR ranges in --volume-host-cidr-denylist (default true)

--volume-host-cidr-denylist strings

A comma-separated list of CIDR ranges to avoid from volume plugins.

Podgc controller flags:

--terminated-pod-gc-threshold int32

Number of terminated pods that can exist before the terminated pod garbage collector starts deleting terminated pods. If <= 0, the terminated pod garbage collector is disabled. (default 12500)

Replicaset controller flags:

--concurrent-replicaset-syncs int32

The number of replica sets that are allowed to sync concurrently. Larger number = more responsive replica management, but more CPU (and network) load (default 5)

Replicationcontroller flags:

--concurrent-rc-syncs int32

The number of replication controllers that are allowed to sync concurrently. Larger number = more responsive replica management, but more CPU (and network) load (default 5)

Resourcequota controller flags:

--concurrent-resource-quota-syncs int32

The number of resource quotas that are allowed to sync concurrently. Larger number = more responsive quota management, but more CPU (and network) load (default 5)

--resource-quota-sync-period duration

```
                The period for syncing quota usage status in the system (default 5m0s)

Serviceaccount controller flags:

      --concurrent-serviceaccount-token-syncs int32
                The number of service account token objects that are allowed to sync
concurrently. Larger number = more responsive token generation, but more CPU (and network)
load (default 5)
      --root-ca-file string
                If set, this root certificate authority will be included in service
account's token secret. This must be a valid PEM-encoded CA bundle.
      --service-account-private-key-file string
                Filename containing a PEM-encoded private RSA or ECDSA key used to sign
service account tokens.

Ttl-after-finished controller flags:

      --concurrent-ttl-after-finished-syncs int32
                The number of TTL-after-finished controller workers that are allowed to
sync concurrently. (default 5)

Metrics flags:

      --allow-metric-labels stringToString
                The map from metric-label to value allow-list of this label. The key's
format is <MetricName>,<LabelName>. The value's format is
<allowed_value>,<allowed_value>...e.g. metric1,label1='v1,v2,v3',
metric1,label2='v1,v2,v3' metric2,label1='v1,v2,v3'.
                (default [])
      --disabled-metrics strings
                This flag provides an escape hatch for misbehaving metrics. You must
provide the fully qualified metric name in order to disable it. Disclaimer: disabling
metrics is higher in precedence than showing hidden metrics.
      --show-hidden-metrics-for-version string
                The previous version for which you want to show hidden metrics. Only the
previous minor version is meaningful, other values will not be allowed. The format is
<major>.<minor>, e.g.: '1.16'. The purpose of this format is make sure you have the
                opportunity to notice if the next release hides additional metrics, rather
than being surprised when they are permanently removed in the release after that.
```

```
Logs flags:

      --log-flush-frequency duration
                Maximum number of seconds between log flushes (default 5s)
      --log-json-info-buffer-size quantity
                [Alpha] In JSON format with split output streams, the info messages can be
buffered for a while to increase performance. The default value of zero bytes disables
buffering. The size can be specified as number of bytes (512), multiples of 1000 (1K),
                multiples of 1024 (2Ki), or powers of those (3M, 4G, 5Mi, 6Gi). Enable the
LoggingAlphaOptions feature gate to use this.
      --log-json-split-stream
                [Alpha] In JSON format, write error messages to stderr and info messages
to stdout. The default is to write a single stream to stdout. Enable the
LoggingAlphaOptions feature gate to use this.
      --logging-format string
                Sets the log format. Permitted formats: "json" (gated by
LoggingBetaOptions), "text". (default "text")
  -v, --v Level
                number for the log level verbosity
      --vmodule pattern=N,...
                comma-separated list of pattern=N settings for file-filtered logging (only
works for text log format)

Misc flags:

      --kubeconfig string
                Path to kubeconfig file with authorization and master location
information.
      --master string
                The address of the Kubernetes API server (overrides any value in
kubeconfig).

Global flags:

  -h, --help
                help for kube-controller-manager
      --version version[=true]
                Print version information and quit
```

Revision #26
Created 2023-04-20 17:38:46 IST by hxd
Updated 2023-11-11 11:45:20 GMT by hxd