Building High-Availability Clusters

Each of master replicas will run the following components in the following mode:

- etcd instance: all instances will be clustered together

- API server: each server will talk to local etcd - all API servers in the cluster will be available;

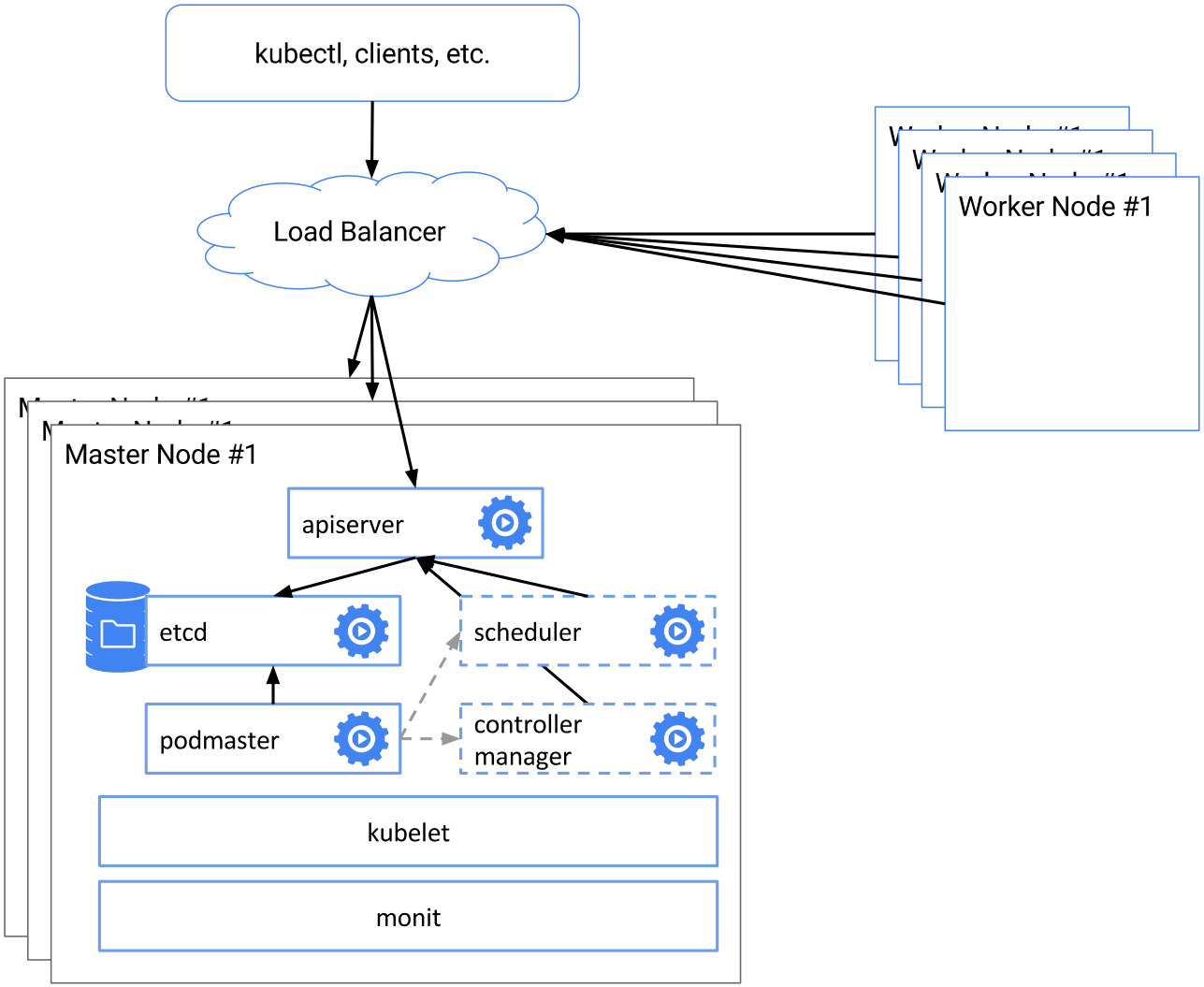
- controllers, scheduler: will use lease mechanism - only one instance of each of them will be active in the cluster;

- add-on manager: each manager will work independently trying to keep add-ons in sync.

- kubelet: run a number of processes that implement the Kubernetes API. Works as a process watcher.

- monit: monitor the kubelet

Building High-Availability Clusters



Building High-Availability Clusters

Master elected components

Use a lease-lock in the API to perform master election. Use the --leader-elect flag for each scheduler and controller-manager, using a lease in the API will ensure that only 1 instance of the scheduler and controller-manager are running at once.

Building High-Availability Clusters

Best practices for replicating masters for HA clusters

- Try to place masters replicas in different zones. During a zone failure, all master placed inside the zone will fail. To survive zone failure, also place nodes in multiple zones

Building High-Availability Clusters

Best practices for replicating masters for HA clusters

- Do not use a cluster with two master replicas. Consensus on a two replica cluster requires both replicas running when changing persistent state. the autofailover is based on all but one node being able to communicate. In a two node scenario, the cluster can't guarantee that. It can with three.

Building High-Availability Clusters

Best practices for replicating masters for HA clusters

- When you add a master replica, cluster state (etcd) is copied to a new instance. If the cluster is large, it may take a long time to duplicate its state. This operation may be sped up by migrating etcd data directory, as described here (we are considering adding support for etcd data dir migration in future)