



SCHEDULING

Scheduling

kube-scheduler

The larger and more diverse a Kubernetes deployment becomes, the more administration of scheduling can be important. The kube-scheduler determines which nodes will run a Pod, using a topology-aware algorithm.

Users can set the priority of a pod, which will allow preemption of lower priority pods. The eviction of lower priority pods would then allow the higher priority pod to be scheduled.

The scheduler tracks the set of nodes in your cluster, filters them based on a set of predicates, then uses priority functions to score or determine on which node each Pod should be scheduled. The Pod specification as part of a request is sent to the kubelet on the node for creation.

The default scheduling decision can be affected through the use of Labels on nodes or Pods. Labels of podAffinity, taints, and pod bindings allow for configuration from the Pod or the node perspective. Some, like tolerations, allow a Pod to work with a node, even when the node has a taint that would otherwise preclude a Pod being scheduled.

Not all labels are drastic. Affinity settings may encourage a Pod to be deployed on a node, but would deploy the Pod elsewhere if the node was not available. Sometimes, documentation may use the term *require*, but practice shows the setting to be more of a request. As beta features, expect the specifics to change. Some settings will evict Pods from a node should the required condition no longer be true, such as **requiredDuringScheduling**, **RequiredDuringExecution**.

Other options, like a custom scheduler, need to be programmed and deployed into your Kubernetes cluster.