Kubernetes自定义

1. 引言
   1. 编写目的

本文主要是阐释对kubernetes框架做哪些修改，以及使用其内置的接口扩展API。本文主要阐释两个方面的内容，一方面是自定义资源的实现，另一方面是自定义kube-scheduler的优选策略。

1. 自定义资源
   1. 目的

自定义资源的目的是承载规则数据，更好的融入kubernetes框架。

* 1. 实现

这部分包括两个方面，一方面是自定义资源的设计与实现，另一方面是对应controller的实现。

自定义资源的格式大致如下：

apiVersion: apiextensions.k8s.io/v1beta1

kind: CustomResourceDefinition

metadata:

name: foos.samplecontroller.k8s.io

spec:

group: samplecontroller.k8s.io

version: v1alpha1

names:

kind: Foo

plural: foos

scope: Namespaced

参考：<https://blog.csdn.net/cloudvtech/article/details/80277960>

Controller的实现参考如下：

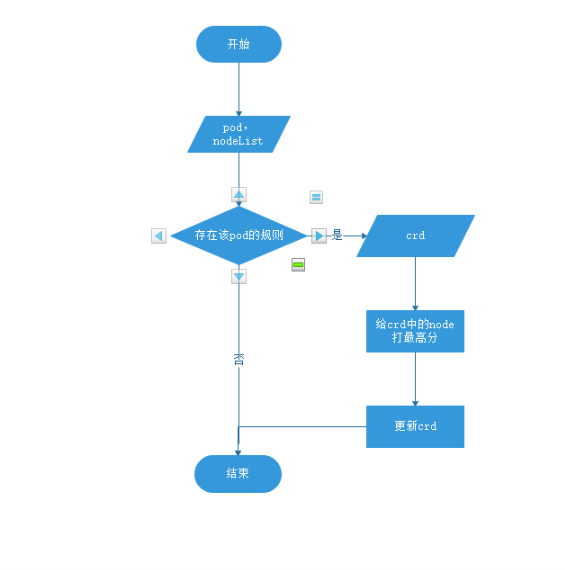
<https://www.cnblogs.com/hindsight/p/9036362.html>

<https://blog.openshift.com/kubernetes-deep-dive-code-generation-customresources/>

1. 优选策略
   1. 目的

添加优选策略是因为kubernetes内置的策略无法满足项目需要。我们希望将一个pod绑定到特定的node上，就kubernetes目前的功能无法实现。所以我们决定自己实现。

* 1. 流程图



说明：

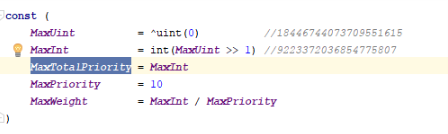
1. pod对应的规则如何判断？

通过namespace，ownType和ownName查询etcd。

1. 可能会查询出多个结果，该怎么消费？

Crd（自定义资源）中包含了timestamp，根据timestamp对查询结果进行排序，优先消费timestamp最小的crd。

1. 当crd的replicas为0时，删除该crd。
2. 如何实现给某个node打最高分？



通过在自定义的优先策略中将node的分打(MaxTotalPriority>>1)分值，既保证后面算总分时不会超过MaxTotalPriority，也可以将node的分值打最高。《kube-scheduler的分值计算》中介绍了默认kube-scheduler如何处理每一个优选策略打的分。

1. 更新CRD，由CRD对应的controller实现。