python × PYCHINA Python 中国社区

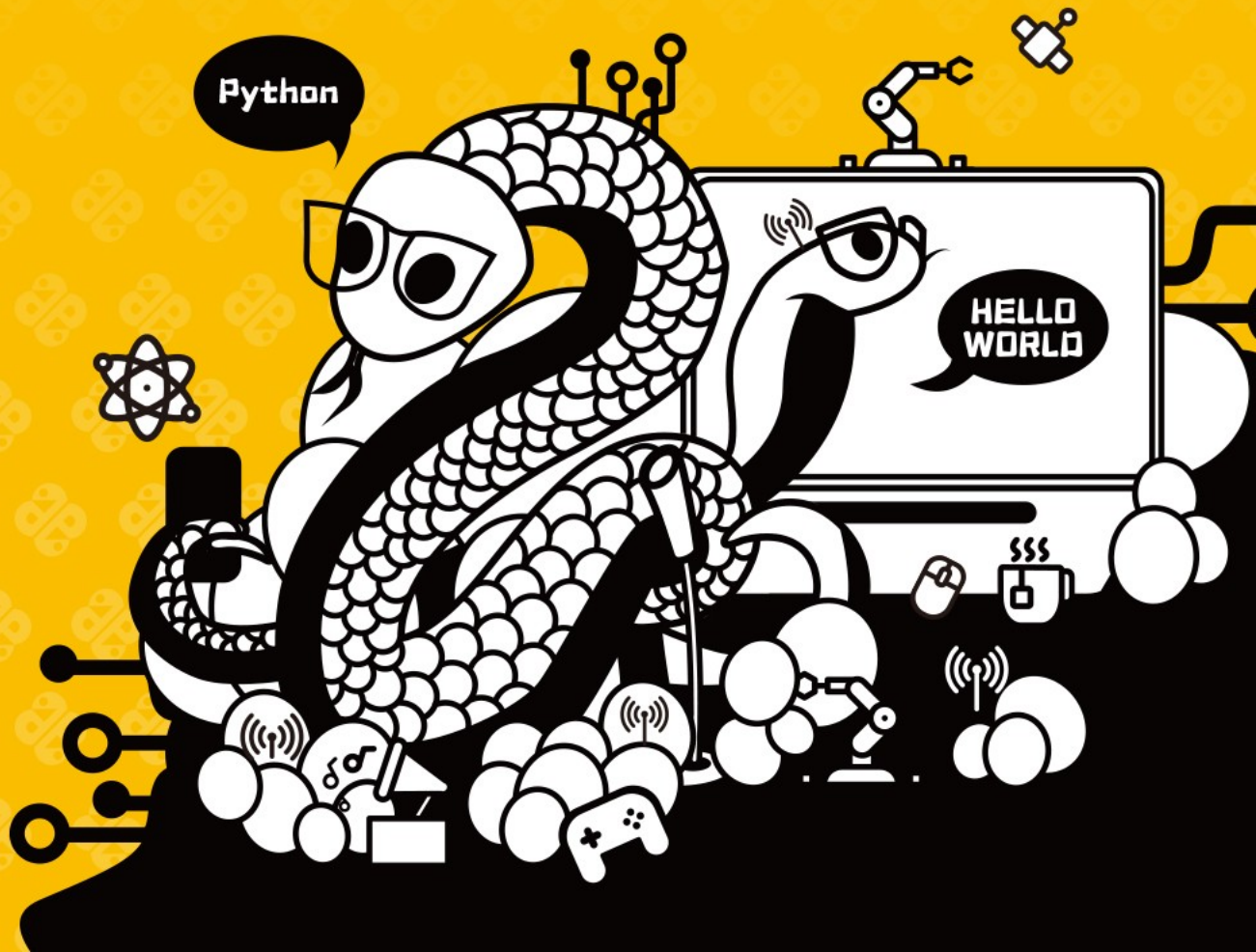
Python for Good

PyCon China 2022

Microsoft® MVP Most Valuable Professional

用 Python 给 Kubernetes 写个控制器

主讲人：张晋涛



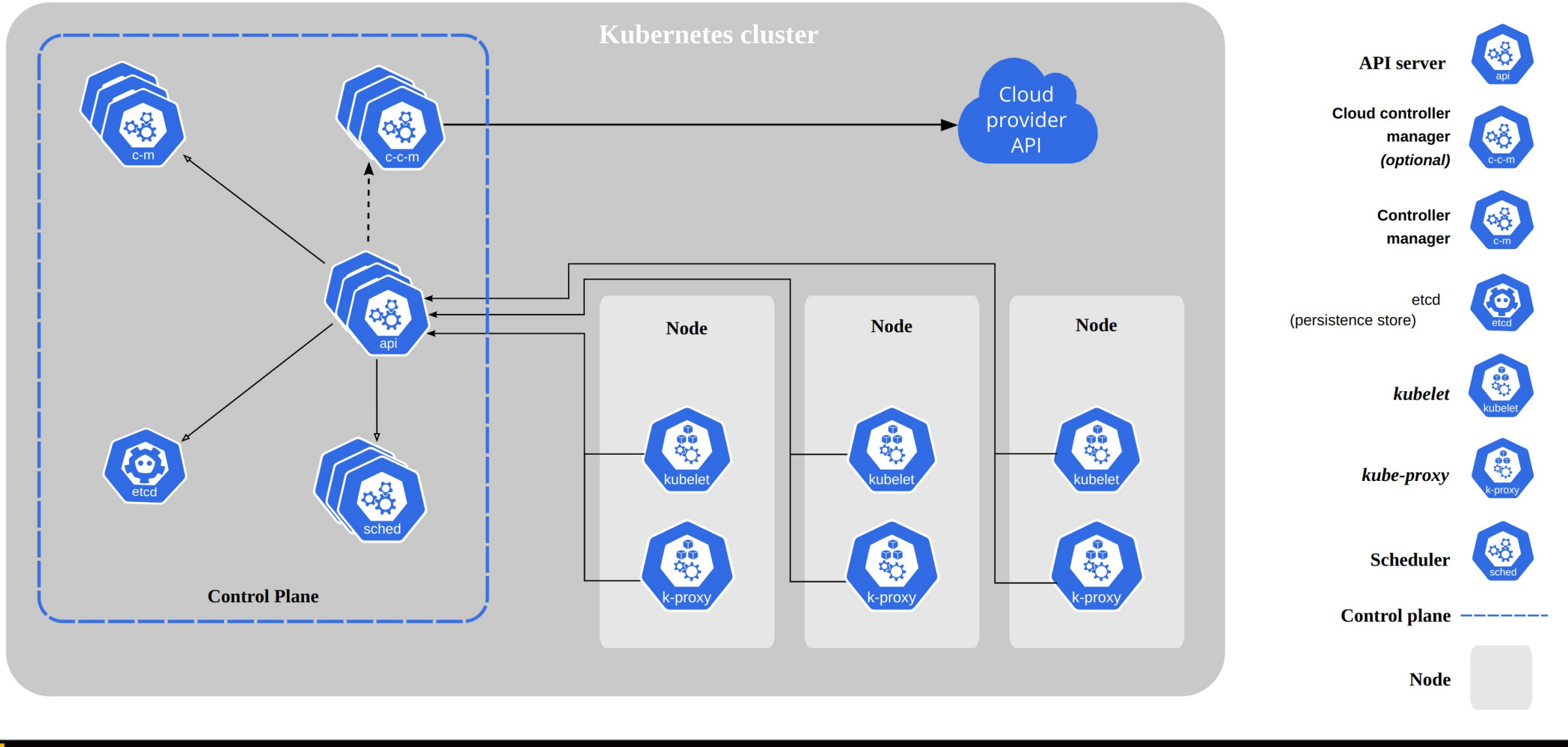
个人介绍

- Apache APISIX PMC
- Kubernetes Ingress NGINX maintainer
- Microsoft MVP
- 『K8S 生态周报』发起人和维护者
- GitHub：tao12345666333
- Mail: zhangjintao@apache.org

Agenda

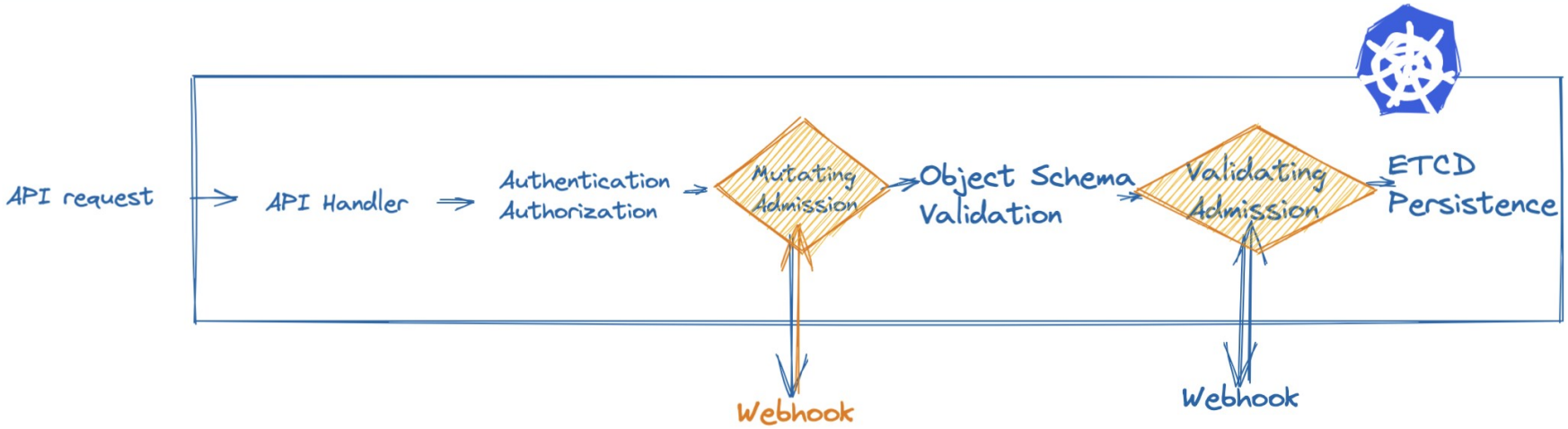
- Kubernetes 中请求处理流程
- 什么是准入控制器
- 用 Python 实现准入控制器
- 与其他方案对比

Kubernetes 架构



- Kubernetes 集群的核心组件
- 处理集群内外的所有请求

Kubernetes 请求处理流程



- API Handler 匹配处理链路（/apis）
- 认证 / 授权
- Mutating Admission： 可进行变更操作
- Object schema validation： schema 校验
- Validating Admission： 可进行验证操作
- etcd： 持久化

什么是准入控制器

- 在 Mutating Admission 或 Validating Admission 执行相关操作的代码逻辑或者组件
- （静态）准入控制器： Kubernetes 代码中携带，不可动态调整的
- 动态准入控制器：利用 Kubernetes 提供的 MutatingAdmissionWebhook 和 ValidatingAdmissionWebhook 扩展点，由用户自行开发的组件，接收 HTTP 回调。

为什么需要准入控制器

- Kubernetes 中一系列复杂的校验 / 事务逻辑
- 用户场景中会有各种需求
- 安全合规：镜像源检查 / 启动用户等；
- 应用治理：资源配额 /label 标识等；

如何实现动态准入控制器

- 建议 Kubernetes v1.16 以上用 v1 API；
- 构建 web server 接收请求并作出响应
- 在 Kubernetes 中创建 AdmissionConfiguration 或 ValidatingWebhookConfiguration 资源进行配置

示例实现

- 请求：以 POST 发送来的 AdmissionReview 对象

https://github.com/tao12345666333/py-admission-controller


```
{
  "kind": "AdmissionReview",
  "apiVersion": "admission.k8s.io/v1beta1",
  "request": {
    "uid": "66666666-5f5f-11e8-bc74-36e6bb280816",
    "kind": {
      "group": "",
      "version": "v1",
      "kind": "Pod"
    },
    "resource": {
      "group": "",
      "version": "v1",
      "resource": "pods"
    },
    "namespace": "dummy",
    "operation": "CREATE",
    "userInfo": {
        ...
    },
    "object": {
      "metadata": {
      },
      "spec": {
          ...
      },
      "status": {}
    },
    "oldObject": null
  }
}
```

示例实现

- 响应： 200 状态码的 AdmissionReview 对象
- uid 从 request 复制过来
- allowed: 决定是否允许
- 可包含 status 或者 patch 等响应

```
{
  "apiVersion": "admission.k8s.io/v1",
  "kind": "AdmissionReview",
  "response": {
    "uid": "<value from request.uid>",
    "allowed": true,
    "patchType": "JSONPatch",
    "patch":
"W3sib3AiOiAiYWRkIiwgInBhdGgiOiAiL3NwZWMvcmVwbGljYXMiLCAidmFsdWUiOiAzfV0
="
  }
}
```

https://github.com/tao12345666333/py-admission-controller

其他方案

- OPA/Gatekeeper
- Kyverno
- Kubernetes v1.26 ValidatingAdmissionPolicy 新特性

OPA/Gatekeeper

限制副本范围

```yaml
apiVersion: templates.gatekeeper.sh/v1beta1
kind: ConstraintTemplate
metadata:
  name: k8sreplicalimits
  annotations:
    description: Requires a number of replicas to be set for a deployment between a min and max value.
spec:
  crd:
    spec:
      names:
        kind: k8sreplicalimits
      validation:
        # Schema for the `parameters` field
        openAPIV3Schema:
          type: object
          properties:
            ranges:
              type: array
              items:
                type: object
                properties:
                  min_replicas:
                    type: integer
                  max_replicas:
                    type: integer
  targets:
    - target: admission.k8s.gatekeeper.sh
      rego: |
        package k8sreplicalimits
        deployment_name = input.review.object.metadata.name
        violation[{"msg": msg}] {
          spec := input.review.object.spec
          not input_replica_limit(spec)
          msg := sprintf("The provided number of replicas is not allowed for deployment: %v. Allowed ranges: %v", [deployment_name, input.parameters])
        }
        input_replica_limit(spec) {
          provided := input.review.object.spec.replicas
          count(input.parameters.ranges) > 0
          range := input.parameters.ranges[_]
          value_within_range(range, provided)
        }
        value_within_range(range, value) {
          range.min_replicas ≤ value
          range.max_replicas ≥ value
        }

```

```yaml
apiVersion: constraints.gatekeeper.sh/v1beta1
kind: k8sreplicalimits
metadata:
  name: replica-limits
spec:
  match:
    kinds:
      - apiGroups: ["apps"]
        kinds: ["Deployment"]
  parameters:
    ranges:
    - min_replicas: 3
      max_replicas: 10
```

Kyverno

- 需要 3 个副本

```yaml
apiVersion: kyverno.io/v1
kind: ClusterPolicy
metadata:
  name: deployment-has-multiple-replicas
  annotations:
    policies.kyverno.io/title: Require Multiple Replicas
    policies.kyverno.io/category: Sample
    policies.kyverno.io/severity: medium
    policies.kyverno.io/subject: Deployment
spec:
  validationFailureAction: audit
  background: true
  rules:
    - name: deployment-has-multiple-replicas
      match:
        resources:
          kinds:
          - Deployment
      validate:
        message: "Deployments should have three replicas to ensure availability."
        pattern:
          spec:
            replicas: ">2"
```

ValidatingAdmissionPolicy

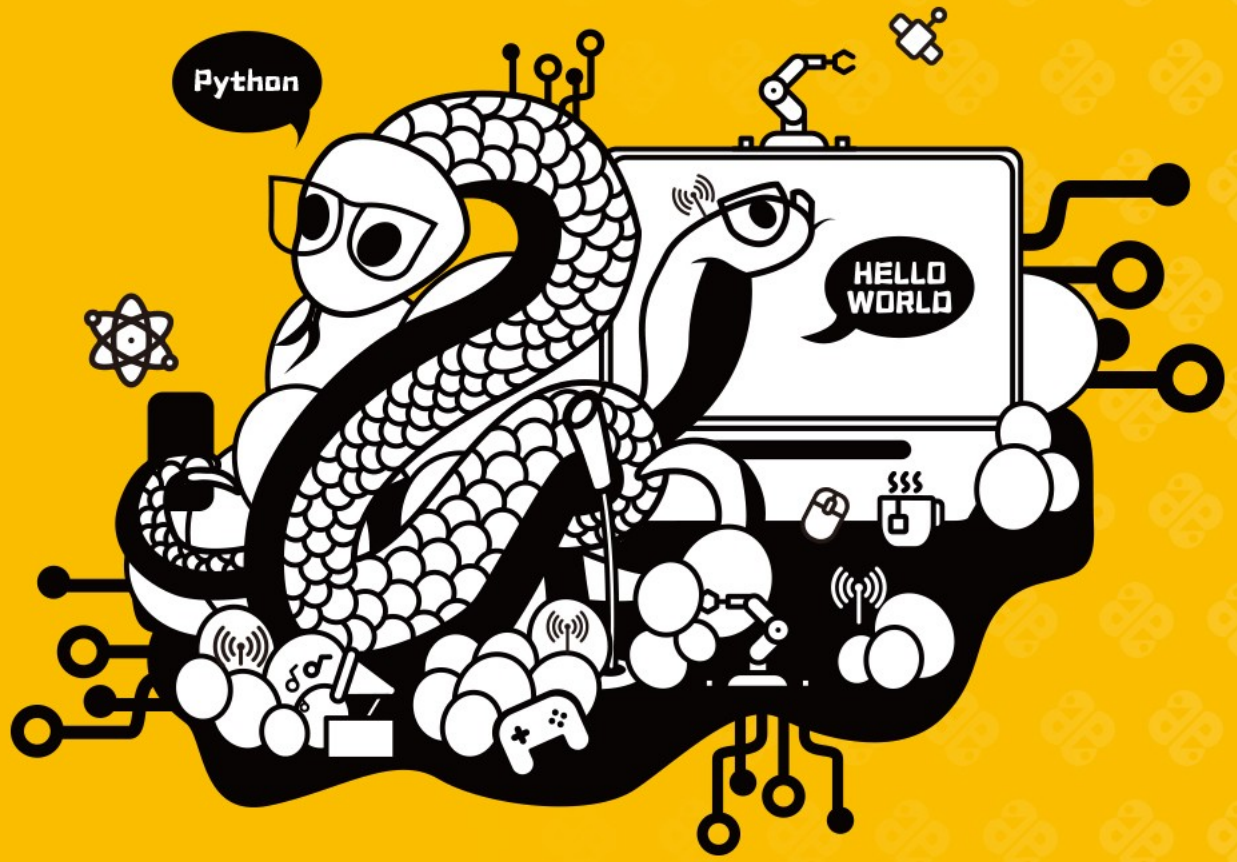
- Deploy 副本数小于等于 2 失败

```yaml
apiVersion: admissionregistration.k8s.io/v1alpha1
kind: ValidatingAdmissionPolicy
metadata:
  name: "demo-policy.moelove.info"
Spec:
  failurePolicy: Fail
  matchConstraints:
    resourceRules:
    - apiGroups:   ["apps"]
      apiVersions: ["v1"]
      operations:  ["CREATE", "UPDATE"]
      resources:   ["deployments"]
  validations:
    - expression: "object.spec.replicas ≤ 2"

```

对比

- 自研：更灵活，与一些内部系统集成。但需要开发和维护成本；
- OPA/Gatekeeper：简单，需要学习 Rego；
- Kyverno：简单，通过 YAML 即可使用；
- Kubernetes v1.26 ValidatingAdmissionPolicy 新特性：默认未开启，尚不稳定，仅能进行 Validating。但属于原生特性，无需其他组件；



Microsoft® MVP Most Valuable Professional

Thanks!

感谢观看