k8s ipv6 双栈集群 公众组件适配方案

作者：bingoct
原文链接：https://ld246.com/article/1658889099228
来源网站：链滴

背景

k8s: ipv4/ipv6 双栈

1. 服务规约中没有显式设定 .spec.ipFamilyPolicy 。当你创建此服务时，Kubernetes 从所配的第一个 service-cluster-ip-range 种为服务分配一个集群IP，并设置 .spec.ipFamilyPolicy 为 SingleStack 。

2. .spec.ipFamilyPolicy 设置为 PreferDualStack 。

- 当你在双栈集群上创建此服务时，Kubernetes 会为该服务分配 IPv4 和 IPv6 地址。 控制平更新服务的 .spec 以记录 IP 地址分配。 字段 .spec.ClusterIPs 是主要字段，包含两个分配的 IP 址； .spec.ClusterIP 是次要字段， 其取值从 .spec.ClusterIPs 计算而来。
- 对于单协议栈的集群， .spec.ClusterIPs 和 .spec.ClusterIP 字段都 仅仅列出一个地址。

通用方案

1. charts中service相关的模板需要暴露 .spec.ipFamilyPolicy字段，在values.yaml中显示指定。用kyverno去注入这个字段，而不是批量修改charts。

```
# 查找Service模板文件
grep -irw "kind: Service" -R
```

2. 确保公众组件能同时监听 0.0.0.0和 ::地址。内部服务可以直接通过ipv6通信。（可以关闭coredn的ipv4解析来测试）

在验证服务是从ipv6的域名解析访问的，coredns需要修改配置，将A的请求，重定向到AAAA，这样以保证只通过访问ipv6服务。

```
.:53 {
        rewrite stop type A AAAA
}
```

3. 外部应用能访问公众组件ipv6服务。

4. 对于公众组件的集群模式，集群内通信端口需要能支持ipv6通信。（这个不大好验证）

kyverno 字段注入

```
helm repo add kyverno https://kyverno.github.io/kyverno/
helm repo update
helm install kyverno kyverno/kyverno --version v2.5.1 -n kyverno --create-namespace
```

创建集群策略 ipv6-policy.yaml，kubectl apply -f ipv6-policy.yaml

```
apiVersion: kyverno.io/v1
kind: ClusterPolicy
metadata:
  name: add-ipfamily
  annotations:
    policies.kyverno.io/title: Add IP Family
    policies.kyverno.io/category: Sample
    policies.kyverno.io/severity: medium
    policies.kyverno.io/subject: ipFamily
```

```
spec:
  rules:
  - name: add-ipfamily
    match:
      resources:
        kinds:
        - Service
    mutate:
      patchStrategicMerge:
        spec:
          ipFamilyPolicy: PreferDualStack
```

kyverno 策略验证

创建python服务，不显示指定 ipFamilyPolicy。

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: python-deployment
  labels:
    app: python-http
spec:
  replicas: 1
  selector:
    matchLabels:
      app: python-http
  template:
    metadata:
      labels:
        app: python-http
    spec:
      containers:
      - name: python-http
        image: python:3.8.13-slim
        ports:
        - containerPort: 8080
        command: ["python"]
        args: ["-m","http.server","8080","--bind","::"]
---
apiVersion: v1
kind: Service
metadata:
  name: python-service
spec:
  selector:
    app: python-http
  ports:
  - protocol: TCP
    port: 8080
    targetPort: 8080
```

查看python-service的.spec.ipFamilyPolicy字段，注入成功，验证通过！

```
 kubectl get svc python-service -o yaml
apiVersion: v1
kind: Service
metadata:
  annotations:
    kubectl.kubernetes.io/last-applied-configuration: |
      {"apiVersion":"v1","kind":"Service","metadata":{"annotations":{},"name":"python-service","
amespace":"default"},"spec":{"ports":[{"port":8080,"protocol":"TCP","targetPort":8080}],"select
r":{"app":"python-http"}}}
    policies.kyverno.io/last-applied-patches: |
      add-ipfamily.add-ipfamily.kyverno.io: added /spec/ipFamilyPolicy
  creationTimestamp: "2022-07-18T12:07:52Z"
  name: python-service
  namespace: default
  resourceVersion: "4662253"
  uid: e1ab7dd7-688c-4db1-b033-034651537a27
spec:
  clusterIP: 10.110.61.121
  clusterIPs:
  - 10.110.61.121
  - fd00::1234:5678:1:c603
  internalTrafficPolicy: Cluster
  ipFamilies:
  - IPv4
  - IPv6
  ipFamilyPolicy: PreferDualStack
  ports:
  - port: 8080
    protocol: TCP
    targetPort: 8080
  selector:
    app: python-http
  sessionAffinity: None
  type: ClusterIP
status:
  loadBalancer: {}
```

ingress-controller

适配的版本见# Ingress NGINX Controller，选择charts 4.1.2，ingress 为1.2.0

```
helmfile -f /data/bkhelmfile/blueking/ingress-nginx-1.2.0.gotmpl sync
```

values配置 controller.service.ipFamilies=["ipv6"]才能设置为ipv6单栈模式。

mysql

chart：bitnami/mysql-4.5.5

覆盖values.yaml 中 master.config 和 slave.config 中的bind-address，绑定至全端口

```
master:
  config: |-
```

```
    [mysqld]
    bind-address=*
slave:
  config: |-
    [mysqld]
    bind-address=*
```

rabbitmq

chart： bitnami/rabbitmq-8.24.12

https://www.rabbitmq.com/management.html#single-listener-port，rabbitmq是支持双栈监的

empd 默认监听ipv6.

rabbitmq: networking

By default, RabbitMQ will listen on port 5672 on **all available interfaces**. It is possible to limit client connections to a subset of the interfaces or even just one, for example, IPv6-only interfces. The following few sections demonstrate how to do it.

rbmq的HTTP API客户端15672，没有监听ipv6

https://rabbitmq.com/rabbitmq-server.8.html

RABBITMQ_NODE_IP_ADDRESSBy default RabbitMQ will bind to all IPv6 and IPv4 interfaces available. This variable limits the node to one network interface or address family. To learn more see the RabbitMQ Networking guide

解决方法：

```
extraConfiguration: |-
  management.tcp.port = 15672
  management.tcp.ip   = ::
```

rabbitmq集群内部通过ipv6通信（端口25672）需要配置开启，见rabbitmq-networking-distribution-ipv6，**注意集群通信端口只能选择ipv4、ipv6中的一种方式通信**!

在ipv6 **单栈**集群，如果需要开启rbmq的集群模式，需要通过：github #issues/6845解决，如下

```
initContainers:
- name: rabbitmq-ipv6-init
  image: "docker.io/busybox:1.33.1"
  imagePullPolicy: IfNotPresent
  volumeMounts:
  - name: rabbitmq-ipv6-cfg
    mountPath: /ipv6
  command: ['sh', '-c', 'echo "{inet6, true}." > /ipv6/erl_inetrc']
extraVolumes:
- name: rabbitmq-ipv6-cfg
  emptyDir: {}
extraVolumeMounts:
- name: rabbitmq-ipv6-cfg
  mountPath: /ipv6
extraEnvVars:
- name: RABBITMQ_SERVER_ADDITIONAL_ERL_ARGS
```

```
  value: "-kernel inetrc '/ipv6/erl_inetrc' -proto_dist inet6_tcp"
- name: RABBITMQ_CTL_ERL_ARGS
  value: "-proto_dist inet6_tcp"
```

redis

chart bitnami/redis-15.3.3

redis: config 中，可以看到在redis6.2 中 bind 127.0.0.1 -::1，默认同时监听ipv4/ipv6地址。

redis-cluster

chart： bitnami/redis-cluster-7.4.6

和redis一致

mongodb

chart：bitnami/mongodb-10.30.6

mongodb: 参考 > 安全 >
络和配置强化 > IP绑定 ，需要手动绑定ipv6。

从MongoDB 3.6，MongoDB二进制文件mongod和开始 mongos，默认情况下绑定到localhost。果 为二进制net.ipv6文件设置了 配置文件设置或 --ipv6 命令行选项，则二进制文件还会绑定到本地IP6地址。

在values.yaml中，提供了开启ipv6的values，无需通过Values.configuration修改配置文件。

```
# values.yaml
enableIPv6: true
```

elastic

chart：bitnami/elasticsearch-17.5.4

Indexing IPv6 addresses in Elasticsearch，elastic已支持ipv6。

Starting with Elasticsearch 5.0, the ip field will support indexing IPv6 addresses.

elastic7.16: IP field type。 An ip field can index/store either IPv4 or IPv6 addresses.

zookeeper

chart：bitnami/zookeeper-9.0.4

zookeeperAdmin Guide: A Guide to Deployment and Administration，zk已支持ipv6。

clientPortAddress**New in 3.3.0:** the address (ipv4, ipv6 or hostname) to listen for client connections; that is, the address that clients attempt to connect to. This is optional, by default we bind in such a way that any connection to the **clientPort** for any address/interface/nic on the server will be accepted.

因为未开启集群模式， leader监听端ipv6口2888和集群内通信3888端口是关闭的

etcd

chart bitnami/etcd-6.2.11 （bcs-system依赖）

无需额外配置，默认支持ipv6。