

Bringing routes to Kubernetes nodes: Introducing FRR-K8s



About me

Telco Network Team @ Red Hat



Contributed to:

- KubeVirt
- SR-IOV Network Operator
- OPA Gatekeeper
- OVN-Kubernetes
- CNI Plugins
- MetalLB



hachyderm.io/@fedepaol

[@fedepaol](https://twitter.com/fedepaol)

fedepaol@gmail.com

Back to Fosdem 2023

MetalLB and FRR: a match made in heaven

Federico Paolinelli

What is FRR?

FRR

FRRouting (FRR) is a free and open source Internet routing protocol suite for Linux and Unix platforms. It implements BGP, OSPF, RIP, IS-IS, PIM, LDP, BFD, Babel, PBR, OpenFabric and VRRP, with alpha support for EIGRP and NHRP [...] FRR has its roots in the Quagga project.



What is MetalLB?

MetaLB

MetaLB is a load-balancer implementation for bare metal Kubernetes clusters, using standard routing protocols (github.com/metallb/metallb)

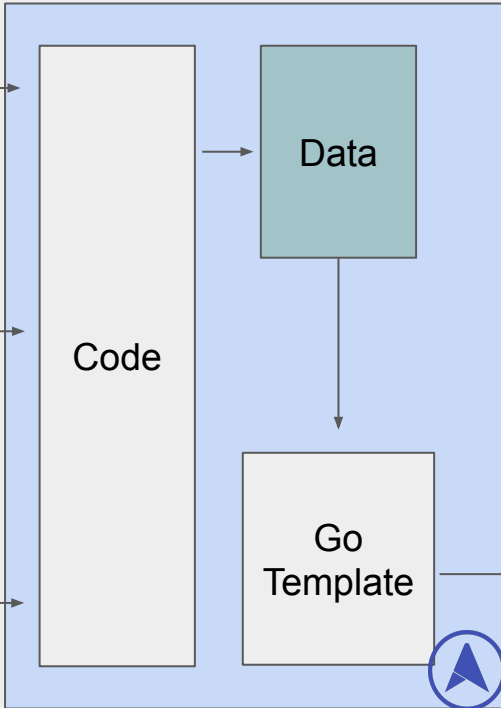


Service

BGP Peer

BGPAdv

Kubernetes API



BGP

BGP

BGP

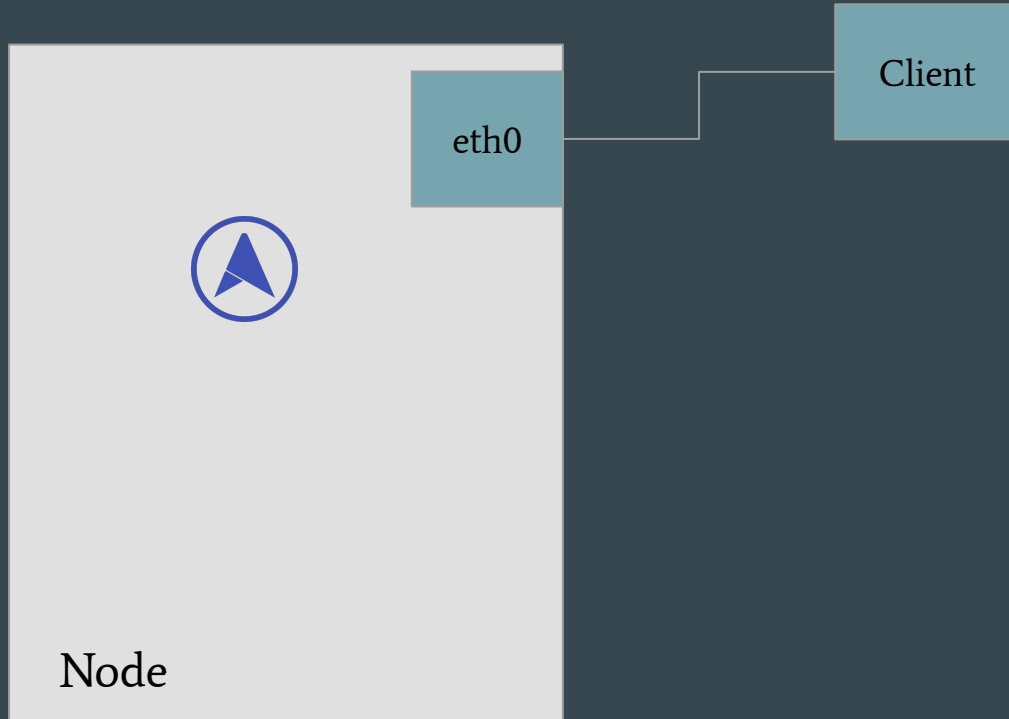
Speaker Pod



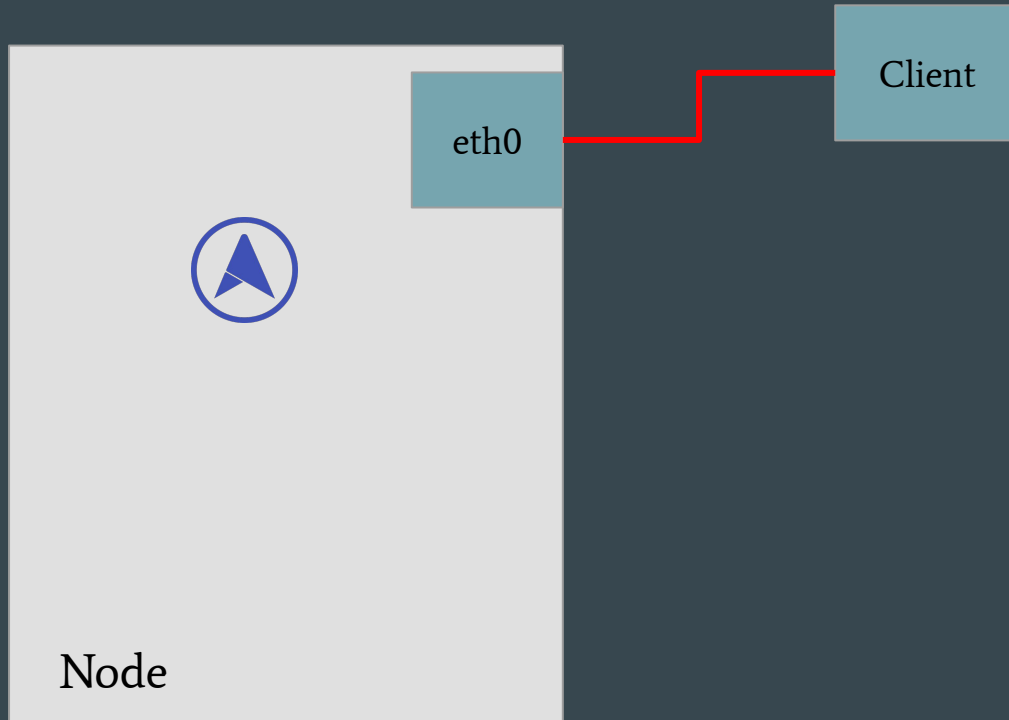
“Can I run MetalLB together with my FRR instance on the cluster nodes?”

A Fosdem 2023 attendee

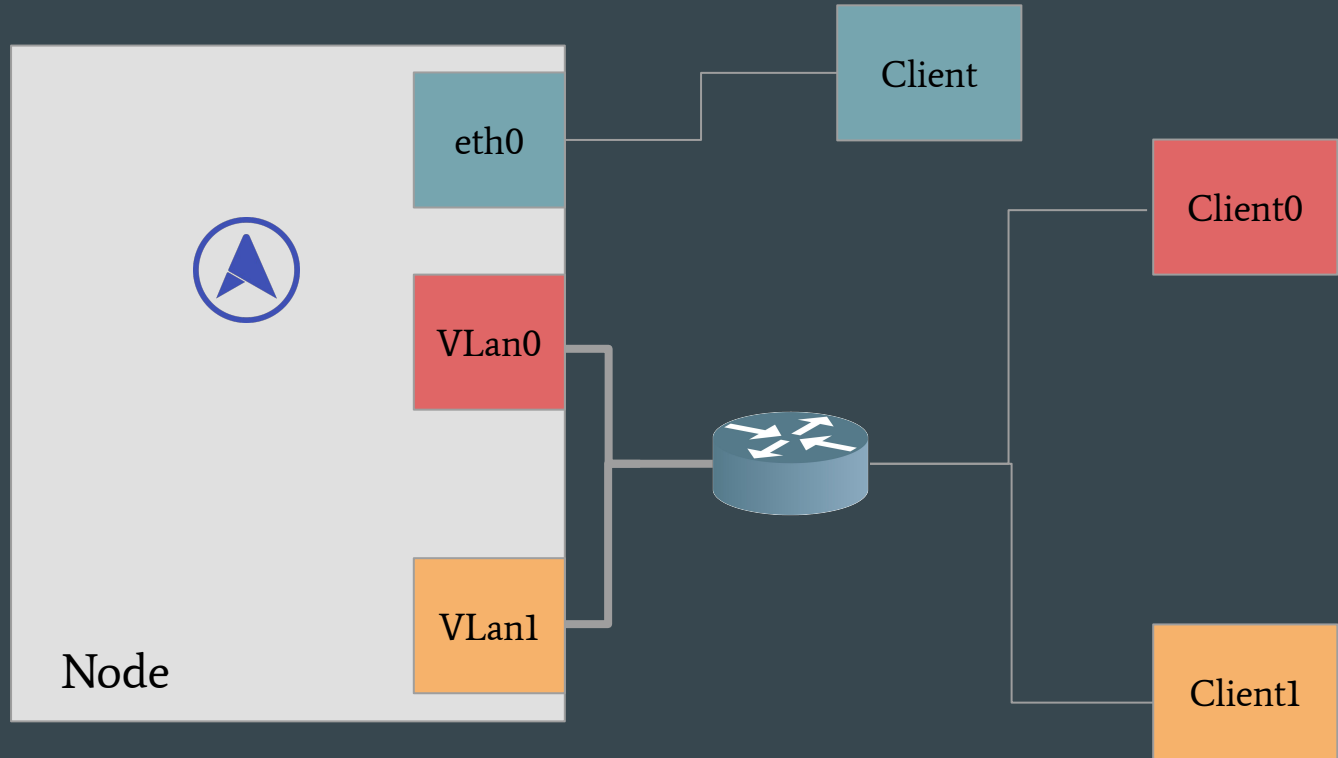
Asymmetric Return Path



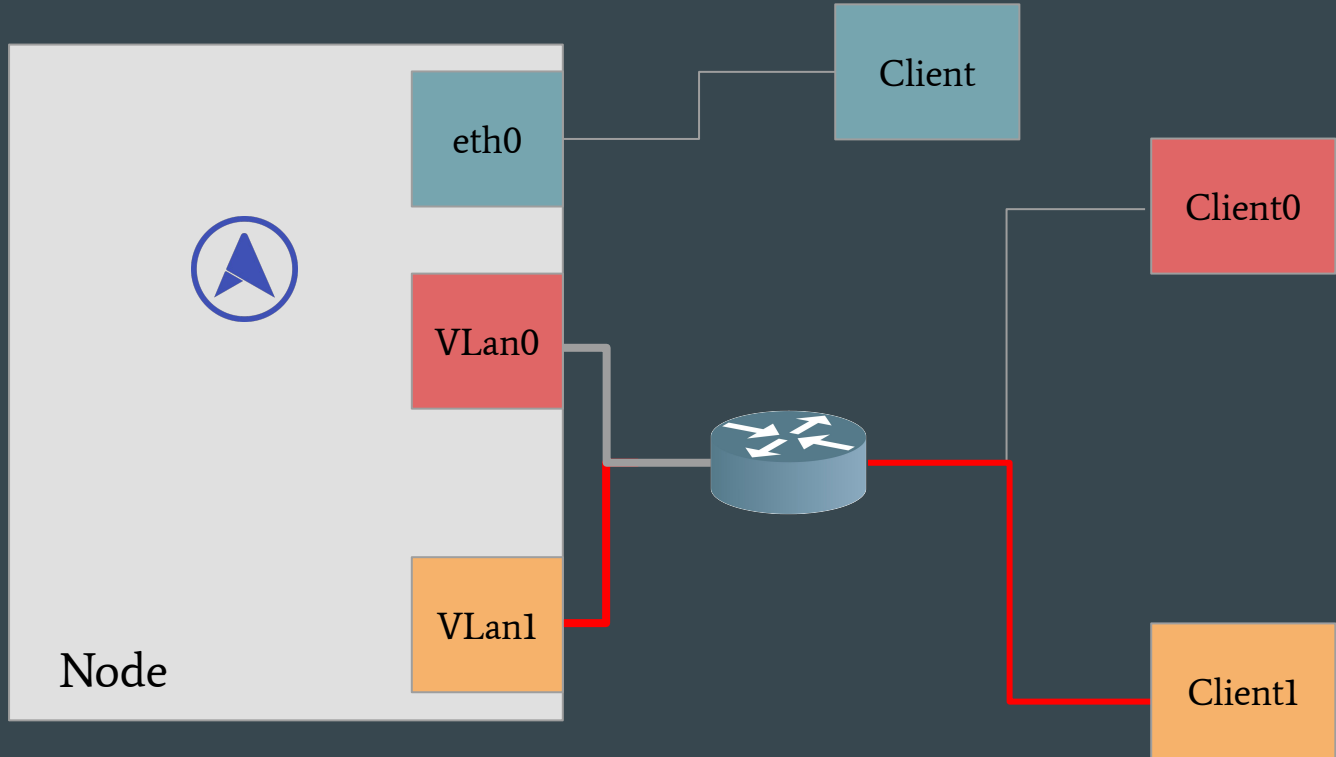
Asymmetric Return Path



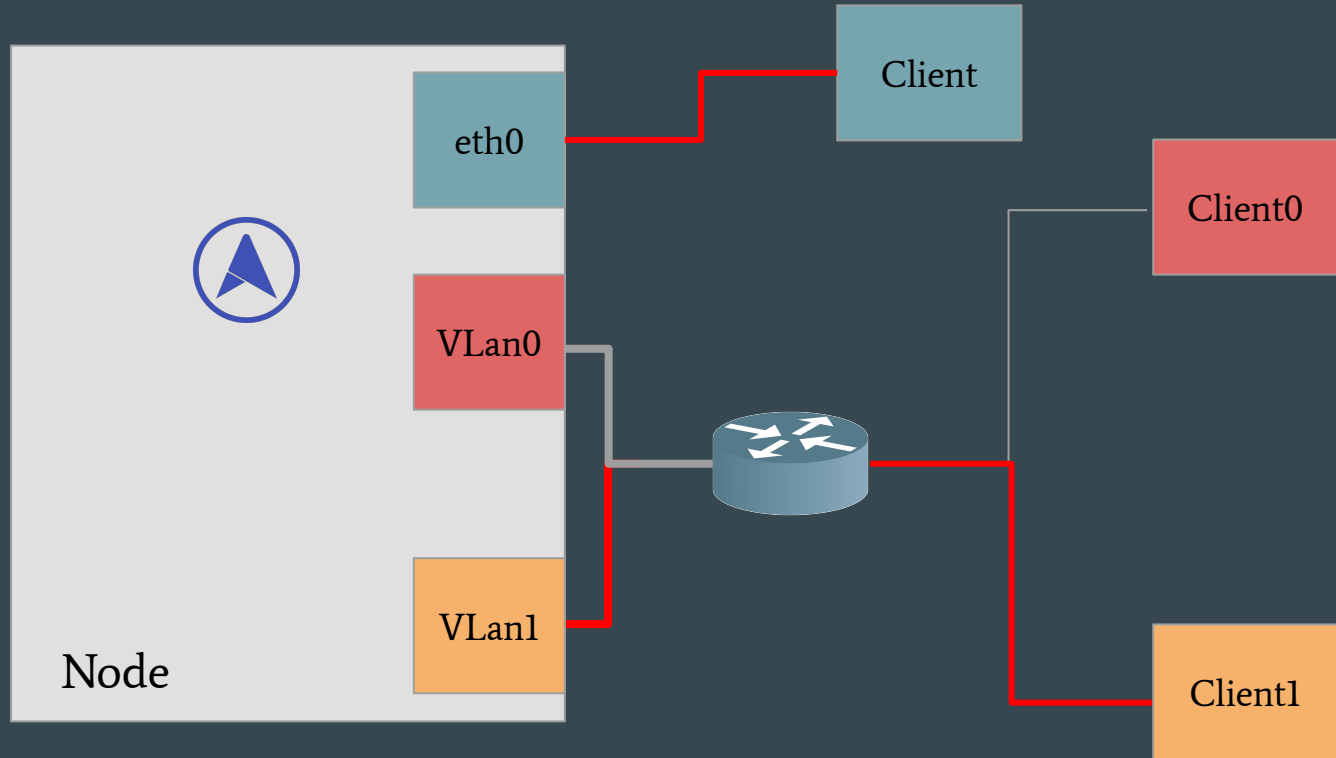
Asymmetric Return Path



Asymmetric Return Path



Asymmetric Return Path





Possible Alternatives

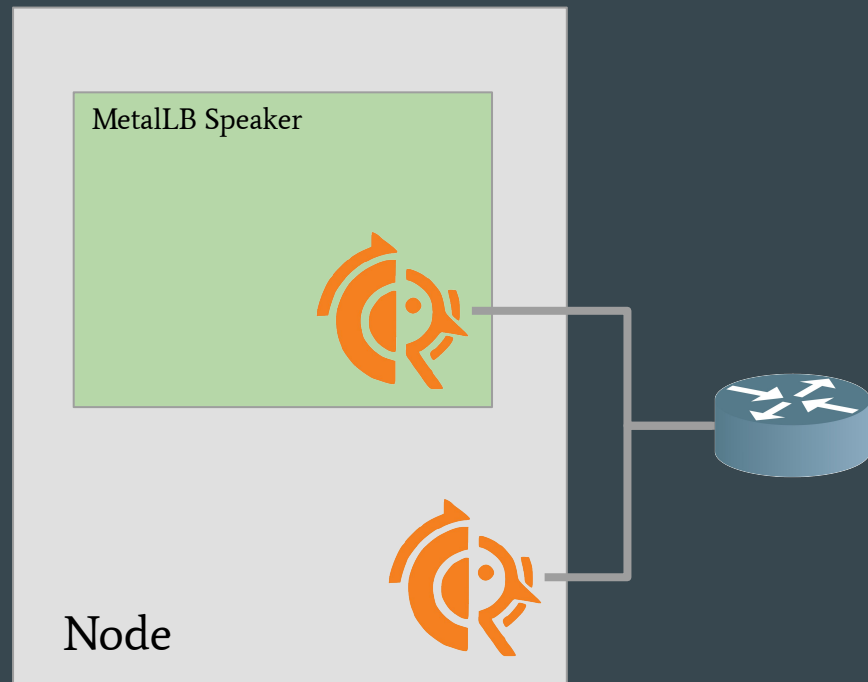
Running two FRR instances

Pros:

- Nothing to do on MetalLB side!

Cons:

- Duplicate number of sessions (and resources)
- Custom ports



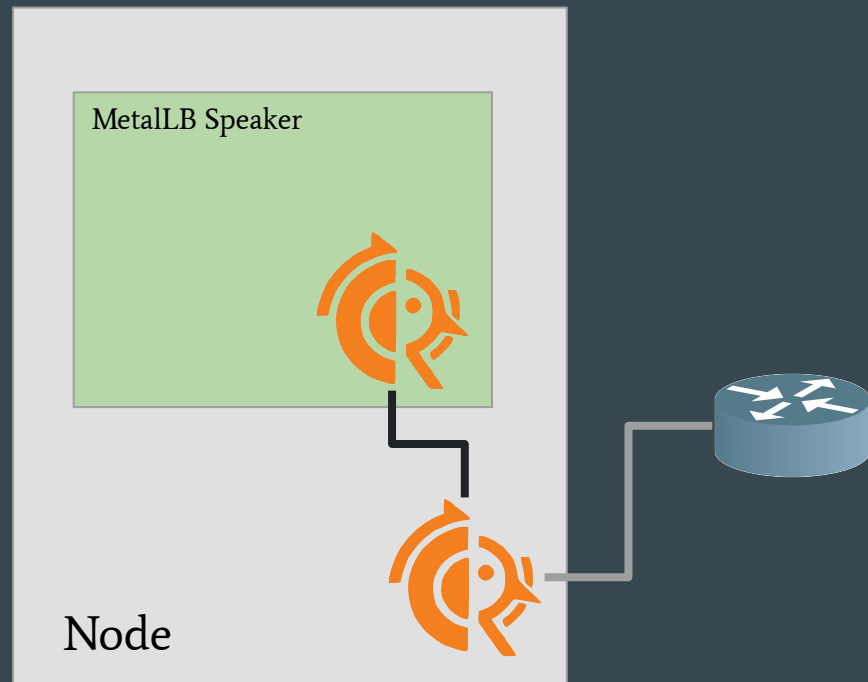
Cascade two FRR instances

Pros:

- The user has control on the FRR config

Cons:

- Limits the flexibility on MetalLB
- FRR was not able to peer with localhost ([Issue 11366](#))
- How about BFD?



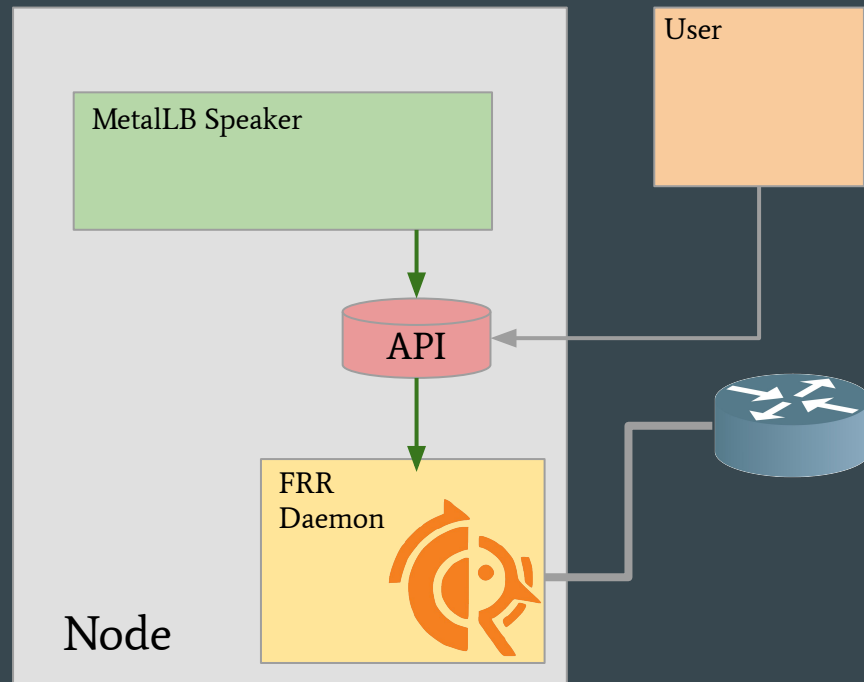
Share the FRR instance with MetalLB

Pros:

- The extra configuration can scale
- No waste of resources / sessions

Cons:

- Getting the right API is tricky
- Not clear how to handle conflicts





Add a proposal to move frr as a separate daemonset #1933

Merged

fedepaol merged 1 commit into `metallb:main` from `fedepaol:design/splitfrr` on Jun 7, 2023

Conversation 38

Commits 1

Checks 0

Files changed 4



fedepaol commented on May 11, 2023

Member ...

Instead of running frr as a side container, here we propose to move it as a separate daemon with its own API, that can be alimented both by the user and by MetalLB.

This would serve all those use cases where the users require to have a FRR-like process running on the cluster's nodes and currently can't because of technical limitations.



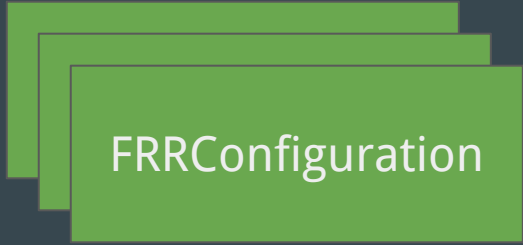
FRR-K8s



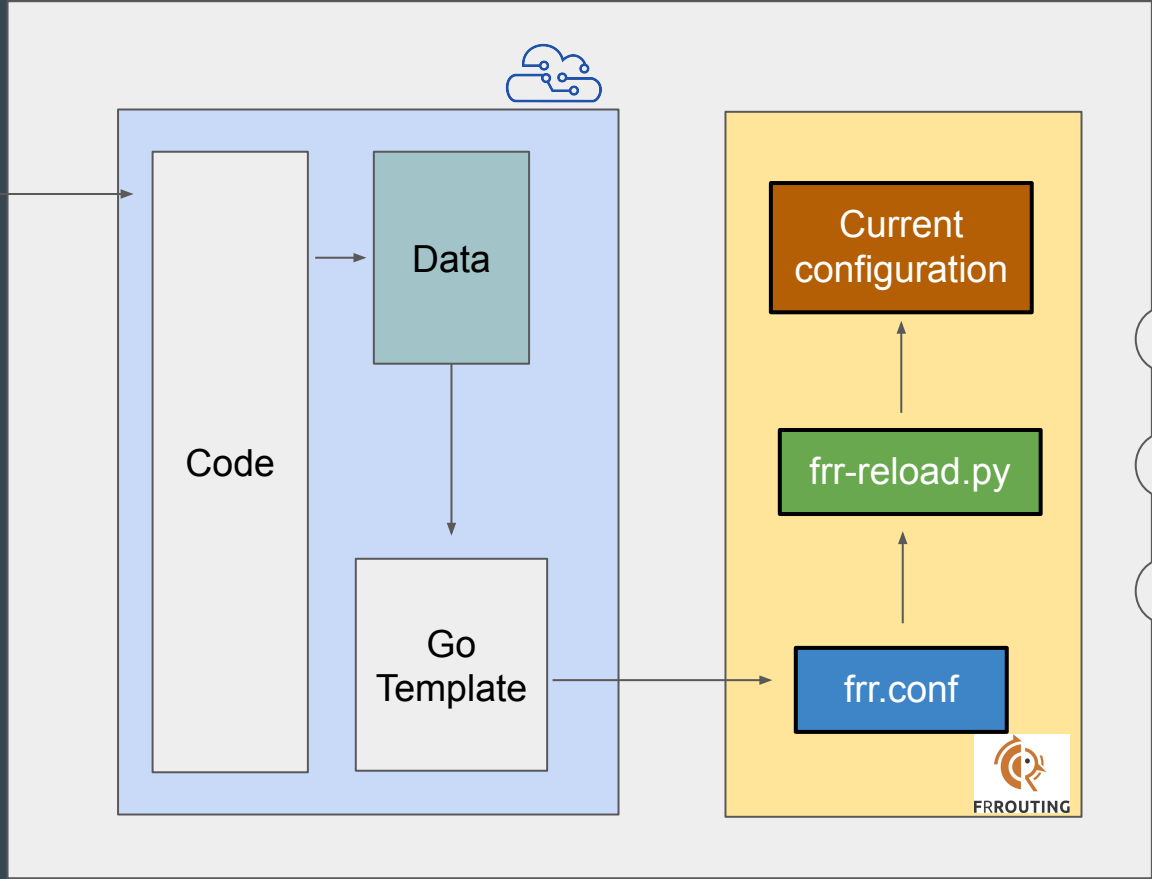
A Kubernetes based daemonset that exposes a subset of the FRR API in a Kubernetes compliant manner.

The rationale behind the creation of this project is to allow multiple actors to share a single FRR instance running on kubernetes nodes.

github.com/metallb/frr-k8s



Kubernetes API



FRR-K8s Pod

How to deploy it

Standalone via:

- Helm
- Kustomize
- All in one manifest

With MetalLB when choosing FRR-K8s mode

github.com/metallb/frr-k8s?tab=readme-ov-file#installing



The API

Challenges

- What granularity of the API should we expose?
- How do we merge configurations?
- Can we intercept configuration conflicts?
- The interaction with MetalLB must be smooth

FRRR Configuration

Configuring BGP

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
metadata:
  name: test
  namespace: frr-k8s-system
spec:
  bgp:
    routers:
      - asn: 64512
        neighbors:
          - address: 172.30.0.3
            asn: 64512
            port: 180
            toAdvertise:
              allowed:
                mode: all
            toReceive:
              allowed:
                mode: all
```

Configuring

```
apiVersion:  
kind: FRRCon  
metadata:  
  name: test  
  namespace:  
spec:  
  bgp:  
    routers:  
      - asn: 64512  
        neighbors:  
          - address: 172.30.0.3  
            asn: 64512  
            port: 180  
            toAdvertise:  
              allowed:  
                mode: all  
            toReceive:  
              allowed:  
                mode: all
```

```
routers:
```

```
- asn: 64512
```

```
  neighbors:
```

```
    - address: 172.30.0.3
```

```
      asn: 64512
```

```
      port: 180
```

```
      toAdvertise:
```

```
        allowed:
```

```
          mode: all
```

```
      toReceive:
```

```
        allowed:
```

```
          mode: all
```

Configuring

```
apiVersion:
```

```
kind: FRRCore
```

```
metadata:
```

```
  name: test
```

```
  namespace:
```

```
spec:
```

```
  bgp:
```

```
    routers:
```

```
      - asn: 64512
```

```
        neighbors:
```

```
          - address: 172.30.0.3
```

```
            asn: 64512
```

```
            port: 180
```

```
            toAdvertise:
```

```
              allowed:
```

```
                mode: all
```

```
            toReceive:
```

```
              allowed:
```

```
                mode: all
```

```
  routers:
```

```
    - asn: 64512
```

```
      neighbors:
```

```
        - address: 172.30.0.3
```

```
          asn: 64512
```

```
          port: 180
```

```
          toAdvertise:
```

```
            allowed:
```

```
              mode: all
```

```
          toReceive:
```

```
            allowed:
```

```
              mode: all
```

Configuring

```
apiVersion:  
kind: FRRCon  
metadata:  
  name: test  
  namespace:  
spec:  
  bgp:  
    routers:  
      - asn: 64512  
        neighb  
          - addr  
            asn:  
            port  
            toAc  
              al  
  
            toRe  
              al
```

```
routers:
```

```
- asn: 64512
```

```
  neighbors:
```

```
    - address: 172.30.0.3
```

```
      asn: 64512
```

```
      port: 180
```

```
      toAdvertise:
```

```
        allowed:
```

```
          mode: all
```

```
      toReceive:
```

```
        allowed:
```

```
          mode: all
```

Configuring

```
apiVersion:  
kind: FRRCore  
metadata:  
  name: test  
  namespace:  
spec:  
  bgp:  
    routers:  
      - asn: 64512  
        neighbors:  
          - address: 172.30.0.3  
            asn: 64512  
            port: 180  
            toAdvertise:  
              allowed:  
                mode: all  
            toReceive:  
              allowed:  
                mode: all
```

```
routers:
```

```
- asn: 64512
```

```
neighbors:
```

```
- address: 172.30.0.3
```

```
asn: 64512
```

```
port: 180
```

```
toAdvertise:
```

```
  allowed:
```

```
    mode: all
```

```
toReceive:
```

```
  allowed:
```

```
    mode: all
```

Configuring

```
apiVersion:  
kind: FRRCore  
metadata:  
  name: test  
  namespace:  
spec:  
  bgp:  
    routers:  
      - asn: 64512  
        neighbors:  
          - address: 172.30.0.3  
            asn: 64512  
            port: 180  
            toAdvertise:  
              allowed:  
                mode: all  
            toReceive:  
              allowed:  
                mode: all
```

```
routers:
```

```
- asn: 64512
```

```
neighbors:
```

```
- address: 172.30.0.3
```

```
asn: 64512
```

```
port: 180
```

```
toAdvertise:
```

```
  allowed:
```

```
    mode: all
```

```
toReceive:
```

```
  allowed:
```

```
    mode: all
```

Advertising Prefixes

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
metadata:
  name: advsome
  namespace: frr-k8s-system
spec:
  bgp:
    routers:
      - asn: 64512
        neighbors:
          - address: 172.18.0.5
            asn: 64512
            toAdvertise:
              allowed:
                prefixes:
                  - 192.168.2.0/24
        prefixes:
          - 192.168.2.0/24
          - 192.169.2.0/24
```


Advertising

```
apiVersion:  
kind: FRRCore  
metadata:  
  name: adv  
  namespace:  
spec:  
  bgp:  
    routers:  
      - asn: 64512  
        neighbors:  
          - address: 172.18.0.5  
            asn: 64512  
            toAdvertise:  
              allowed:  
                prefixes:  
                  - 192.168.2.0/24  
                  - 192.169.2.0/24  
        prefixes:  
          - 192.168.2.0/24  
          - 192.169.2.0/24
```

```
routers:
```

```
- asn: 64512
```

```
neighbors:
```

```
- address: 172.18.0.5
```

```
asn: 64512
```

```
toAdvertise:
```

```
  allowed:
```

```
    prefixes:
```

```
      - 192.168.2.0/24
```

```
prefixes:
```

```
- 192.168.2.0/24
```

```
- 192.169.2.0/24
```

Advertising

```
apiVersion:  
kind: FRRCore  
metadata:  
  name: advs  
  namespace:  
spec:  
  bgp:  
    routers:  
      - asn: 64512  
        neighbors:  
          - address: 172.18.0.5  
            asn: 64512  
            toAdvertise:  
              allowed:  
                prefixes:  
                  - 192.168.2.0/24  
                  - 192.169.2.0/24  
        prefixes:  
          - 192.168.2.0/24  
          - 192.169.2.0/24
```

```
routers:
```

```
- asn: 64512
```

```
  neighbors:
```

```
    - address: 172.18.0.5
```

```
      asn: 64512
```

```
      toAdvertise:
```

```
        allowed:
```

```
          prefixes:
```

```
            - 192.168.2.0/24
```

```
        prefixes:
```

```
          - 192.168.2.0/24
```

```
          - 192.169.2.0/24
```

Advertising

```
apiVersion:  
kind: FRRCon  
metadata:  
  name: adv  
  namespace:  
spec:  
  bgp:  
    routers:  
      - asn: 64512  
        neighb  
          - addr  
            asn:  
            toAc  
            al  
  
    prefix  
      - 192.  
      - 192.
```

```
routers:
```

```
- asn: 64512
```

```
neighbors:
```

```
- address: 172.18.0.5
```

```
asn: 64512
```

```
toAdvertise:
```

```
  allowed:
```

```
    prefixes:
```

```
      - 192.168.2.0/24
```

```
prefixes:
```

```
- 192.168.2.0/24
```

```
- 192.169.2.0/24
```

Receiving Prefixes

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
metadata:
  name: recv
  namespace: frr-k8s-system
spec:
  bgp:
    routers:
      - asn: 64512
        neighbors:
          - address: 172.18.0.5
            asn: 64512
            toReceive:
              allowed:
                prefixes:
                  - prefix: 192.169.1.0/24
                    le: 32g
```

Receiving P

```
apiVersion:  
kind: FRRCon  
metadata:  
  name: recv  
  namespace:  
spec:  
  bgp:  
    routers:  
      - asn: 64512  
        neighb  
          - addr  
            asn:  
            toRe  
            a]
```

```
  neighbors:
```

```
    - address: 172.18.0.5
```

```
      asn: 64512
```

```
      toReceive:
```

```
        allowed:
```

```
          prefixes:
```

```
            - prefix: 192.169.1.0/24
```

```
              le: 32
```

```
    - address: 172.18.0.6
```

```
      asn: 64512
```

```
      toReceive:
```

```
        allowed:
```

```
          mode: all
```

Receiving P

```
apiVersion:  
kind: FRRCon  
metadata:  
  name: recv  
  namespace:  
spec:  
  bgp:  
    routers:  
      - asn: 64512  
        neighb  
          - addr  
            asn:  
            toRe  
            a]
```

```
  neighbors:
```

```
    - address: 172.18.0.5
```

```
      asn: 64512
```

```
      toReceive:
```

```
        allowed:
```

```
          prefixes:
```

```
            - prefix: 192.169.1.0/24
```

```
              le: 32
```

```
    - address: 172.18.0.6
```

```
      asn: 64512
```

```
      toReceive:
```

```
        allowed:
```

```
          mode: all
```

Node Selector

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
metadata:
  name: test
  namespace: frr-k8s-system
spec:
  bgp:
    routers:
      - asn: 64512
        neighbors:
          - address: 172.30.0.3
            asn: 64512
            port: 180
  nodeSelector:
    labelSelector:
      role: "worker"
```

Node Selector

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
metadata:
  name: test
  namespace: frr-k8s-system
spec:
  bgp:
    routers:
      - asn: 64512
        neighbors:
          - address: 172.30.0.3
            asn: 64512
```

```
nodeSelector:
  labelSelector:
    role: "worker"
```


Raw Configuration

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
metadata:
  name: test
  namespace: frr-k8s-system
spec:
  raw:
    rawConfig: |-
      router bgp 64512
        neighbor 172.18.0.5 remote-as 4200000000
        neighbor 172.18.0.5 timers 0 0
        neighbor 172.18.0.5 port 180

        address-family ipv4 unicast
          neighbor 172.18.0.5 activate
          neighbor 172.18.0.5 route-map 172.18.0.5-in in
          neighbor 172.18.0.5 route-map 172.18.0.5-out out
        exit-address-family
```

Raw Configuration

```
apiVersion: frrk8s.metallb.io/v1beta1
```

```
kind: FRRConfiguration
```

```
metadata:
```

```
  name: test
```

```
raw:
```

```
  rawConfig: |-
```

```
    router bgp 64512
```

```
      neighbor 172.18.0.5 remote-as 4200000000
```

```
      neighbor 172.18.0.5 timers 0 0
```

```
      neighbor 172.18.0.5 port 180
```

```
      address-family ipv4 unicast
```

```
        neighbor 172.18.0.5 activate
```

```
        neighbor 172.18.0.5 route-map 172.18.0.5-in in
```

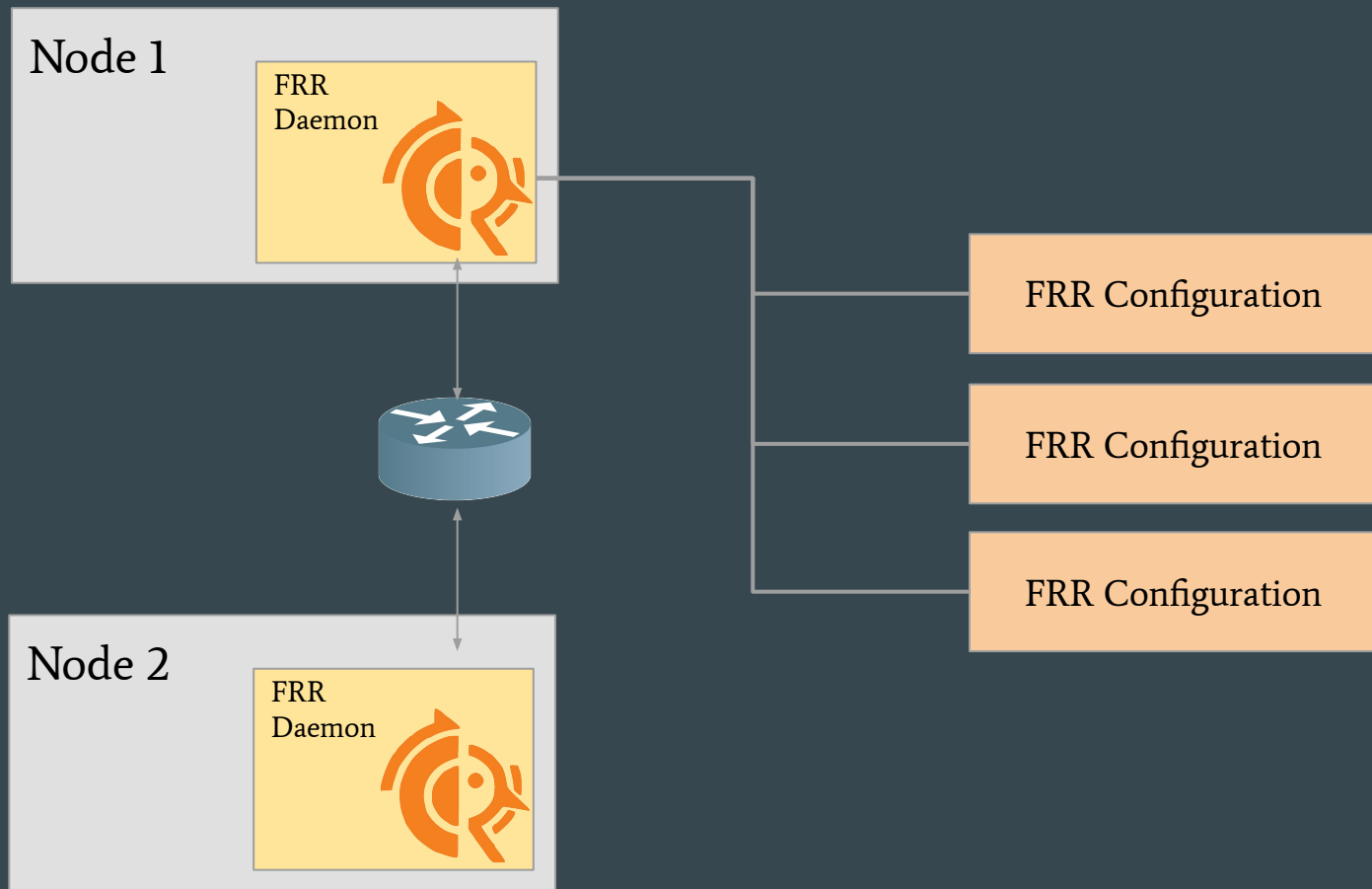
```
        neighbor 172.18.0.5 route-map 172.18.0.5-out out
```

```
      exit-address-family
```



Merging multiple configurations

Merging



Guiding principles

- A configuration must be self-contained
- A configuration can only add to the existing configuration
- A more permissive configuration can override a less permissive one

Merging

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
metadata:
  name: test-0
  namespace: frr-k8s-system
spec:
  bgp:
    routers:
      - asn: 64512
```

```
neighbors:
  - address: 172.18.0.5
    asn: 64512
    toAdvertise:
      allowed:
        prefixes:
          - 192.168.2.0/24
```

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
metadata:
  name: test-1
  namespace: frr-k8s-system
```

```
neighbors:
  - address: 172.18.0.5
    asn: 64512
    toAdvertise:
      allowed:
        mode: all
  - address: 172.30.0.3
    asn: 4200000000
prefixes:
  - 192.168.2.0/24
  - 192.169.2.0/24
```

Merging

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
metadata:
  name: test-0
  namespace: frr-k8s-system
spec:
  bgp:
    routers:
      - asn: 64512
```

```
neighbors:
  - address: 172.18.0.5
    asn: 64512
    toAdvertise:
      allowed:
        prefixes:
          - 192.168.2.0/24
```

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
metadata:
  name: test-1
  namespace: frr-k8s-system
```

```
neighbors:
  - address: 172.18.0.5
    asn: 64512
    toAdvertise:
      allowed:
```

```
        mode: all
  - address: 172.30.0.3
    asn: 4200000000
```

```
prefixes:
  - 192.168.2.0/24
  - 192.169.2.0/24
```

Merging

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
metadata:
  name: test-0
  namespace: frr-k8s-system
spec:
  bgp:
    routers:
      - asn: 64512
```

```
neighbors:
  - address: 172.18.0.5
    asn: 64512
    toAdvertise:
      allowed:
        prefixes:
          - 192.168.2.0/24
```

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
metadata:
  name: test-1
  namespace: frr-k8s-system
```

```
neighbors:
  - address: 172.18.0.5
    asn: 64512
    toAdvertise:
      allowed:
        mode: all
  - address: 172.30.0.3
    asn: 4200000000
prefixes:
  - 192.168.2.0/24
  - 192.169.2.0/24
```


Merging

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
metadata:
  name: test-0
  namespace: frr-k8s-system
spec:
  bgp:
    routers:
      - asn: 64512
```

```
neighbors:
  - address: 172.18.0.5
    asn: 64512
    toAdvertise:
      allowed:
        prefixes:
          - 192.168.2.0/24
```

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
metadata:
  name: test-1
  namespace: frr-k8s-system
```

```
neighbors:
  - address: 172.18.0.5
    asn: 64512
    toAdvertise:
      allowed:
        mode: all
  - address: 172.30.0.3
    asn: 4200000000
prefixes:
  - 192.168.2.0/24
  - 192.169.2.0/24
```



Is it working?

Exposing the Status

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRNodeState
metadata:
  name: frr-k8s-worker
spec: {}
status:
  lastConversionResult: success
  lastReloadResult: success
  runningConfig: |
    Building configuration...

    Current configuration:
    !
    frr version 8.4.2_git
    frr defaults traditional
    hostname frr-k8s-worker
    log file /etc/frr/frr.log informational
    !
    router bgp 64512
      neighbor 172.18.0.5 remote-as 64512
```

Exposing the Status

```
apiVersion: frrk8s.metallb.io/v1beta1  
kind: FRRNodeState
```

```
lastConversionResult: success  
lastReloadResult: success  
runningConfig: |  
  Building configuration...  
  Current configuration:  
  !  
  frr version 8.4.2_git  
  frr defaults traditional  
  hostname frr-k8s-worker  
  log file /etc/frr/frr.log informational  
  !  
  router bgp 64512  
    neighbor 172.18.0.5 remote-as 64512
```

Exposing the Status

```
apiVersion: frrk8s.metallb.io/v1beta1  
kind: FRRNodeState
```

```
lastConversionResult: success
```

```
lastReloadResult: success
```

```
runningConfig: |
```

```
Building configuration...
```

```
Current configuration:
```

```
!
```

```
frr version 8.4.2_git
```

```
frr defaults traditional
```

```
hostname frr-k8s-worker
```

```
log file /etc/frr/frr.log informational
```

```
!
```

```
router bgp 64512
```

```
neighbor 172.18.0.5 remote-as 64512
```

Exposing the Status

```
apiVersion: frrk8s.metallb.io/v1beta1  
kind: FRRNodeState
```

```
lastConversionResult: success  
lastReloadResult: success  
runningConfig: |  
  Building configuration...  
  Current configuration:  
  !  
  frr version 8.4.2_git  
  frr defaults traditional  
  hostname frr-k8s-worker  
  log file /etc/frr/frr.log informational  
  !  
  router bgp 64512  
    neighbor 172.18.0.5 remote-as 64512
```

Exposing the Status

```
apiVersion: frrk8s.metallb.io/v1beta1  
kind: FRRNodeState
```

```
lastConversionResult: success
```

```
lastReloadResult: success
```

```
runningConfig: |
```

```
Building configuration...
```

```
Current configuration:
```

```
!
```

```
frr version 8.4.2_git
```

```
frr defaults traditional
```

```
hostname frr-k8s-worker
```

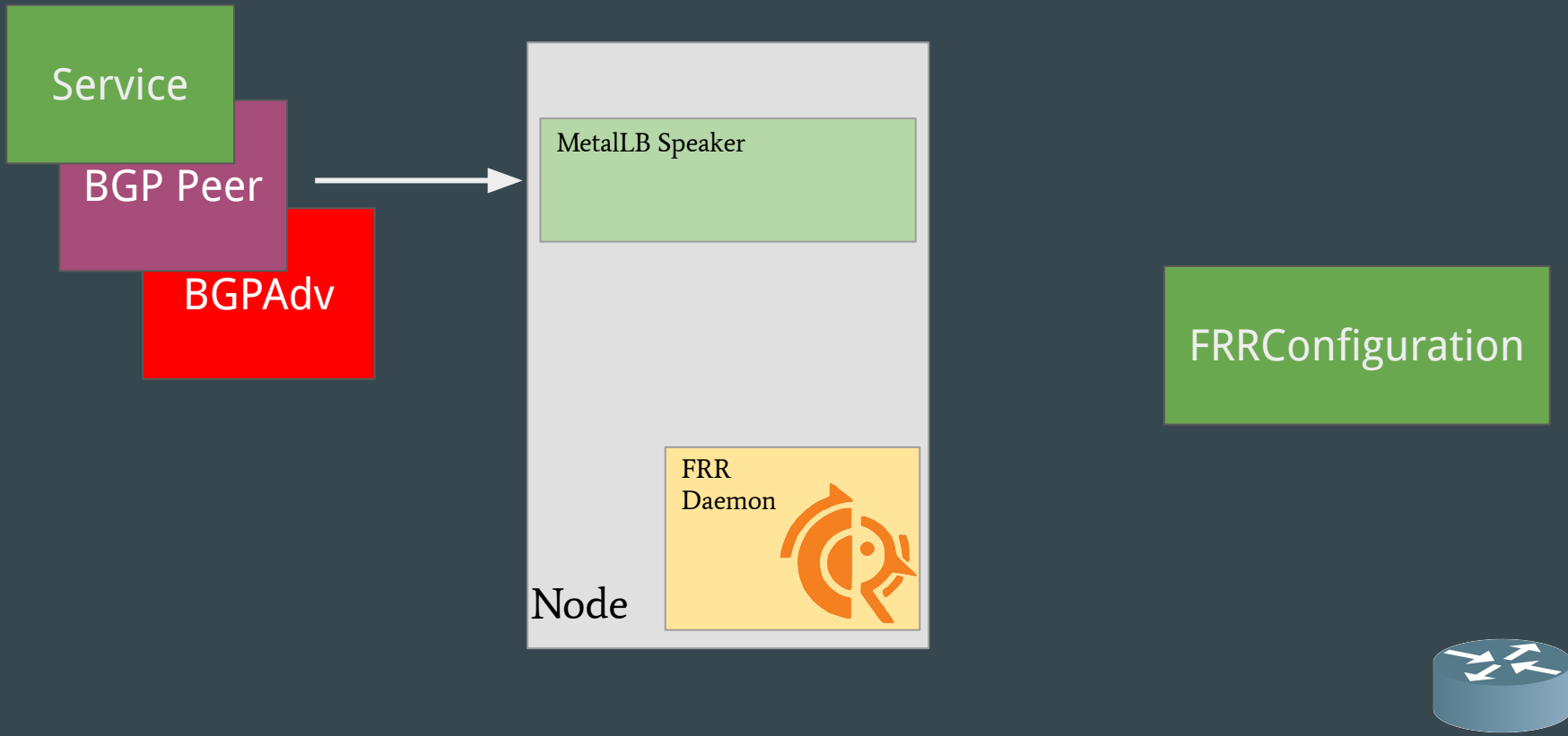
```
log file /etc/frr/frr.log informational
```

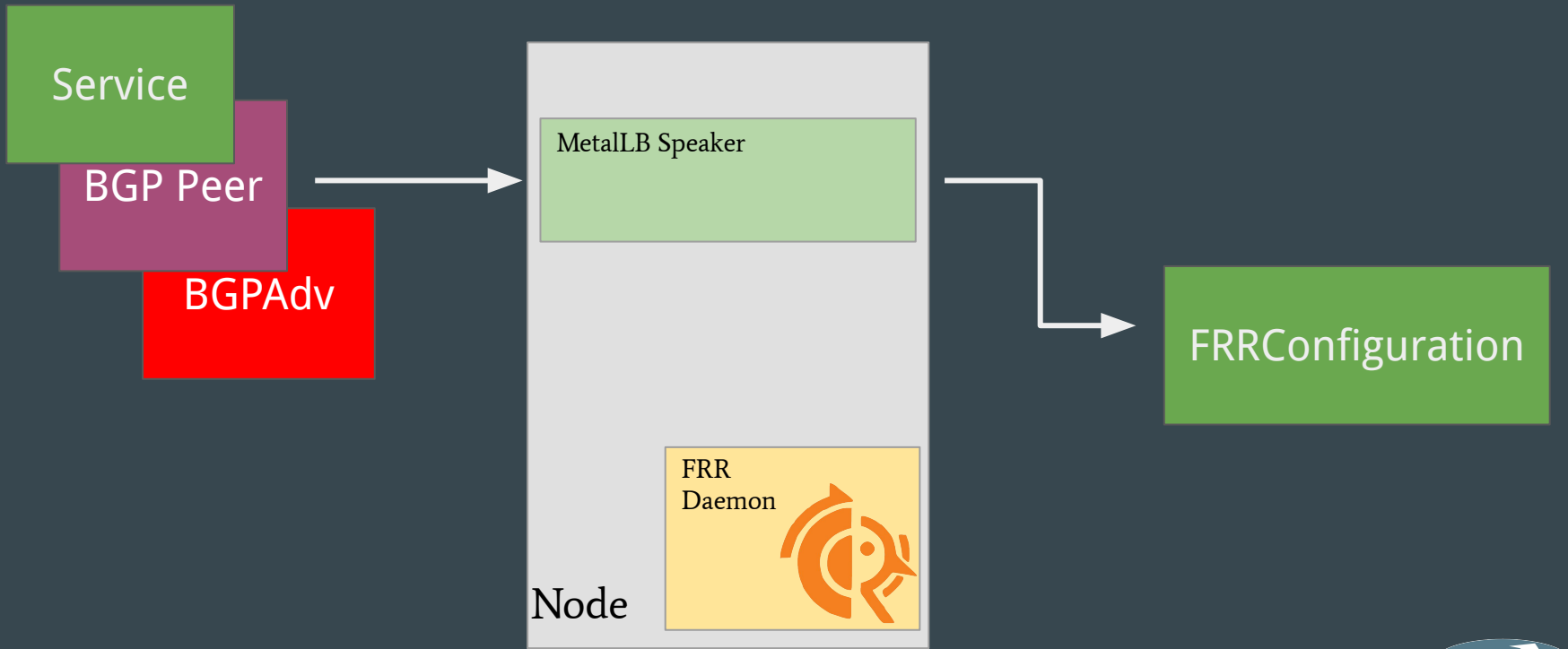
```
!
```

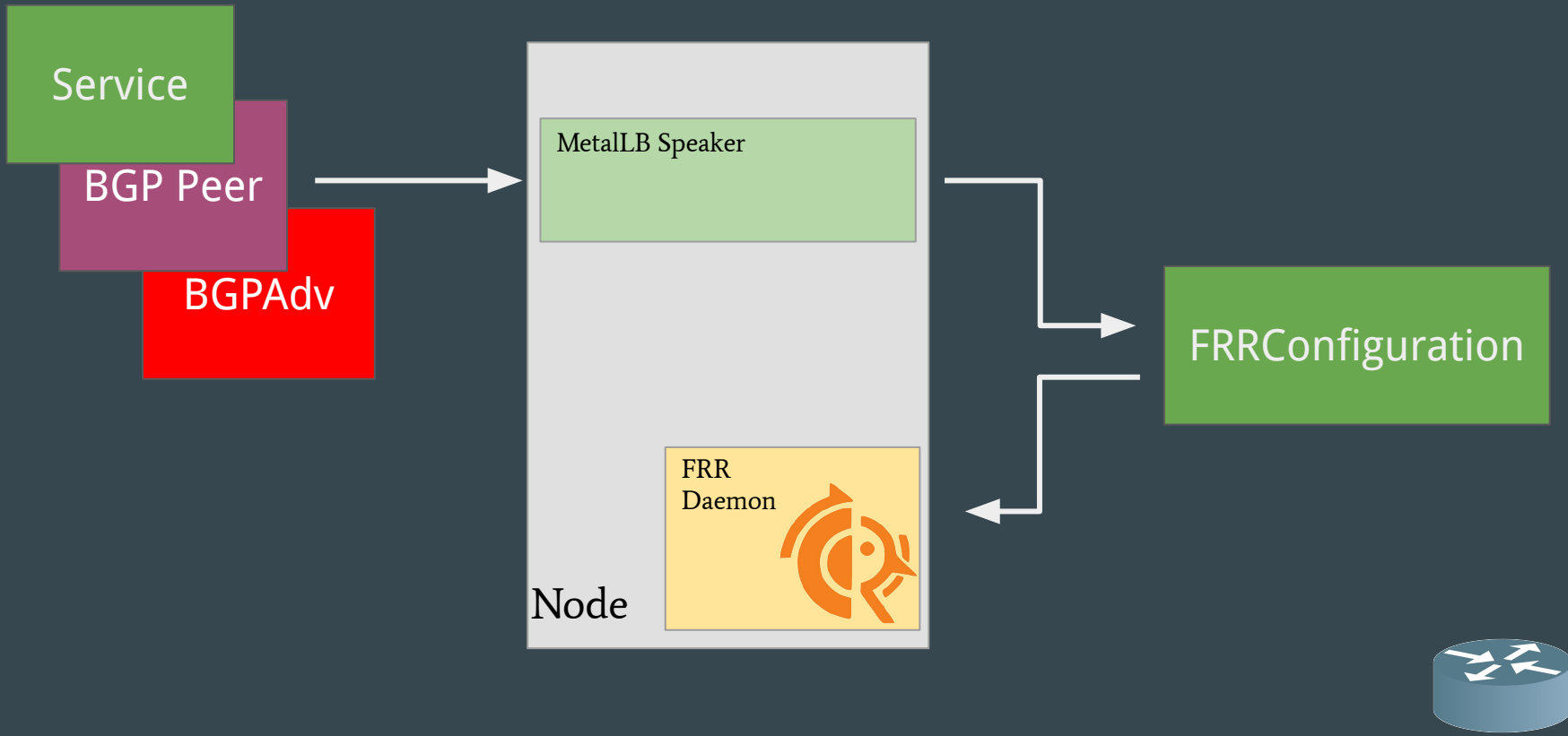
```
router bgp 64512
```

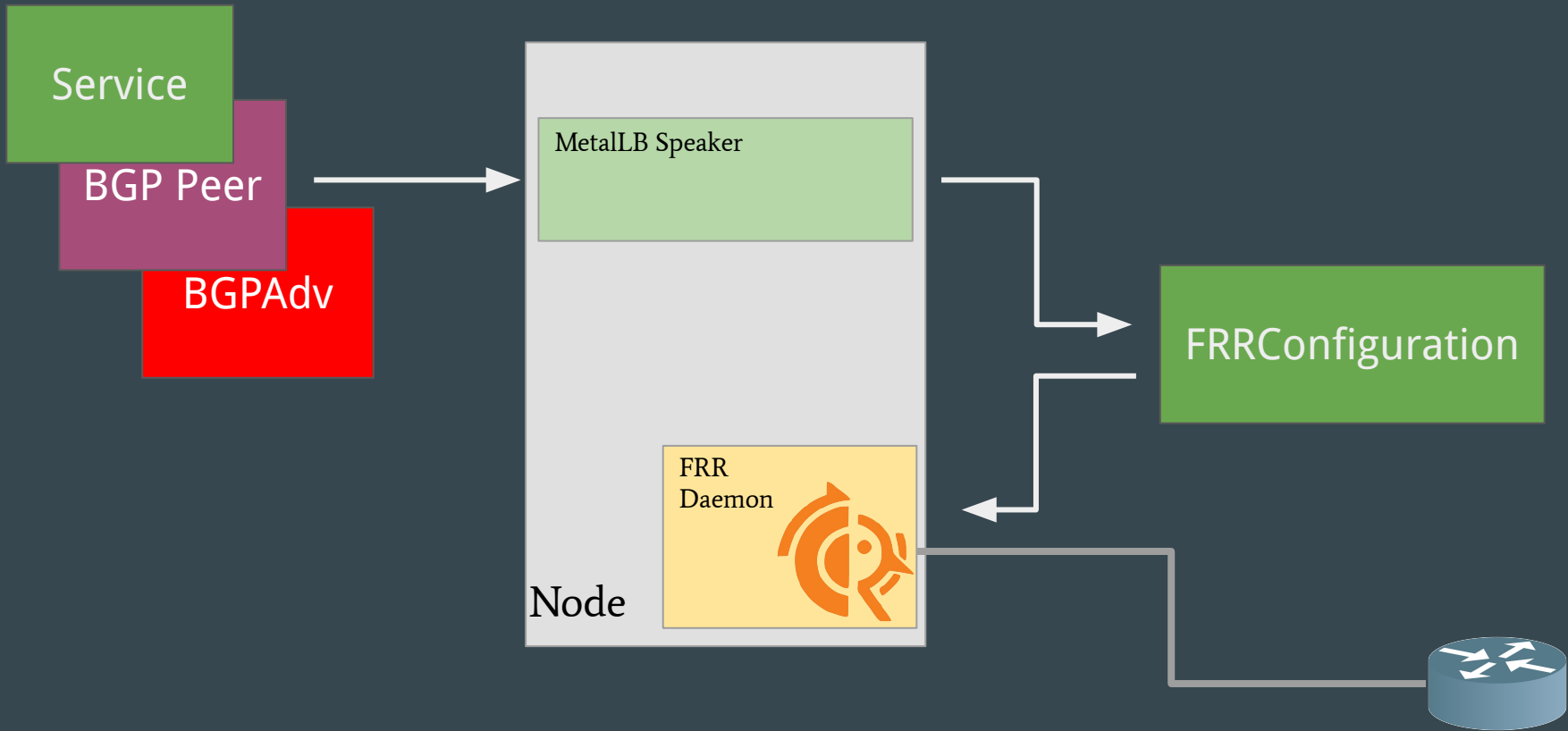
```
neighbor 172.18.0.5 remote-as 64512
```

With MetalLB









With MetalLB

```
apiVersion:
metallb.io/v1beta2
kind: BGPPeer
metadata:
  annotations:
    name: peer
    namespace: metallb-system
spec:
  myASN: 64512
  peerASN: 64512
  peerAddress: 172.18.0.5
```

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
metadata:
  name: metallb-kind-worker
spec:
  bgp:
    routers:
      - asn: 64512
        neighbors:
          - address: 172.18.0.5
            asn: 64512
  nodeSelector:
    matchLabels:
      kubernetes.io/hostname: kind-worker
```

With MetalLB

```
apiVersion:
metallb.io/v1beta2
kind: BGPPeer
metadata:
  annotations:
    name: peer
    namespace: metallb-system
spec:
  myASN: 64512
  peerASN: 64512
  peerAddress: 172.18.0.5
```

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
```

```
  m
  s
  bgp:
    routers:
      - asn: 64512
    neighbors:
      - address: 172.18.0.5
        asn: 64512
  nodeSelector:
    matchLabels:
      hostname: worker0
```

With MetalLB

```
apiVersion:
metallb.io/v1beta2
kind: BGPPeer
metadata:
  annotations:
    name: peer
    namespace: metallb-system
spec:
  myASN: 64512
  peerASN: 64512
  peerAddress: 172.18.0.5
```

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
```

```
  m
  s
  bgp:
    routers:
      - asn: 64512
        neighbors:
          - address: 172.18.0.5
            asn: 64512
    nodeSelector:
      matchLabels:
        hostname: worker0
```

With MetalLB

```
apiVersion:
metallb.io/v1beta2
kind: BGPPeer
metadata:
  annotations:
    name: peer
    namespace: metallb-system
spec:
  myASN: 64512
  peerASN: 64512
  peerAddress: 172.18.0.5
```

```
apiVersion: frrk8s.metallb.io/v1beta1
kind: FRRConfiguration
```

```
  m
  s
  bgp:
    routers:
      - asn: 64512
    neighbors:
      - address: 172.18.0.5
        asn: 64512
  nodeSelector:
    matchLabels:
      hostname: worker0
```


With MetalLB

```
apiVersion:
metallb.io/v1beta2
kind: BGPPeer
metadata:
  annotations:
    name: peer
    namespace: metallb-system
spec:
  myASN: 64512
  peerASN: 64512
  peerAddress: 172.18.0.5
```

```
apiVersion: frrk8s.metallb.io/v1beta1
```

```
neighbors:
- address: 172.18.0.5
  asn: 64512
  toAdvertise:
    allowed:
      mode: filtered
      prefixes:
        - 192.168.10.0/32
prefixes:
- 192.168.10.0/32
```

With MetalLB

```
apiVersion:
metallb.io/v1beta2
kind: BGPPeer
metadata:
  annotations:
    name: peer
    namespace: metallb-system
spec:
  myASN: 64512
  peerASN: 64512
  peerAddress: 172.18.0.5
```

```
apiVersion: frrk8s.metallb.io/v1beta1
```

```
neighbors:
```

```
- address: 172.18.0.5
```

```
asn: 64512
```

```
toAdvertise:
```

```
  allowed:
```

```
    mode: filtered
```

```
    prefixes:
```

```
      - 192.168.10.0/32
```

```
prefixes:
```

```
- 192.168.10.0/32
```

With MetalLB

```
apiVersion:
metallb.io/v1beta2
kind: BGPPeer
metadata:
  annotations:
    name: peer
    namespace: metallb-system
spec:
  myASN: 64512
  peerASN: 64512
  peerAddress: 172.18.0.5
```

```
apiVersion: frrk8s.metallb.io/v1beta1
```

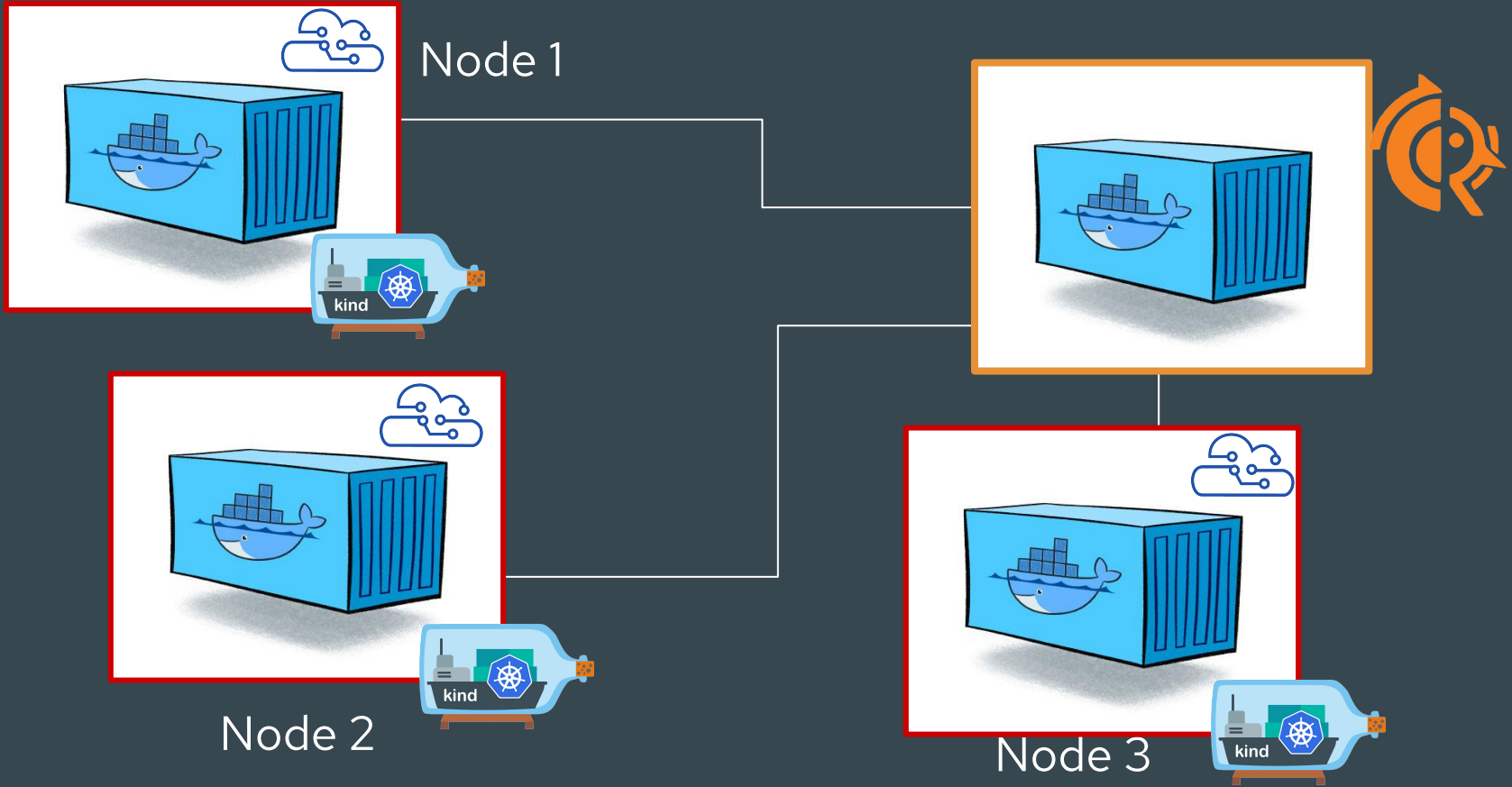
```
neighbors:
- address: 172.18.0.5
  asn: 64512
  toAdvertise:
    allowed:
      mode: filtered
      prefixes:
        - 192.168.10.0/32
prefixes:
- 192.168.10.0/32
```

Demo!



<https://flic.kr/p/dN6fVC>

The demo environment





What's Next?

Resources

- The FRR-K8s [readme](#)
- The MetalLB documentation
- The #metallb channel on the kubernetes slack
- FRR docs [frrouting.org](#)

Thanks!

Any questions?

[@fedepaol](#)

hachyderm.io/@fedepaol

fedepaol@gmail.com

Slides at: speakerdeck.com/fedepaol

fpaoline@redhat.com