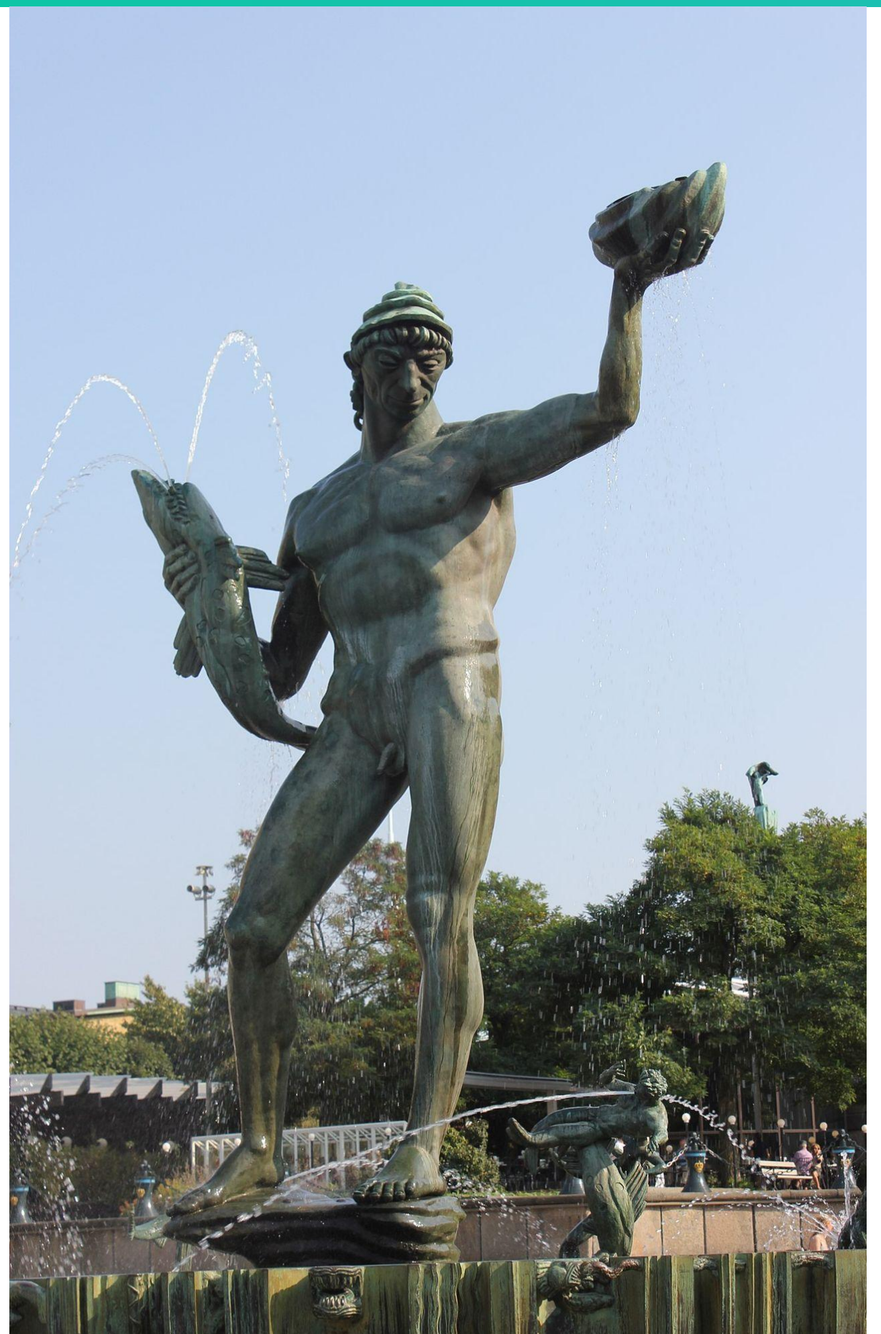
Meltwater Engineering

Swapping the wheels of a running car migrate from Amazon VPC CNI to Cilium in Kubernetes

Platform Engineers Stockholms

Federico Hernandez, Simone Sciarrati

Hallo! Ciao! Hello! Hej! Hola!



Simone Sciarrati

Engineering Team Lead

@dezmodue

@dezmodue

Federico Hernandez

Principal Engineer

@recollir

@recollir

Meltwater

Meltwater Services

- Media monitoring
- Media outreach
- Social listening
- Social publishing and engagement
- Social influencer management

meltwater.com

We offer comprehensive media monitoring and analysis across online news, social media, print, broadcast, and podcasts, capturing more content and conversations than anyone else in the industry.

Kubernetes

70

3000

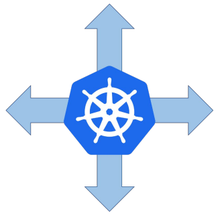
12000

9500

24TB

9 instance types

Meltwater



CA

Observability

AutoScaling

Ingress

DNS

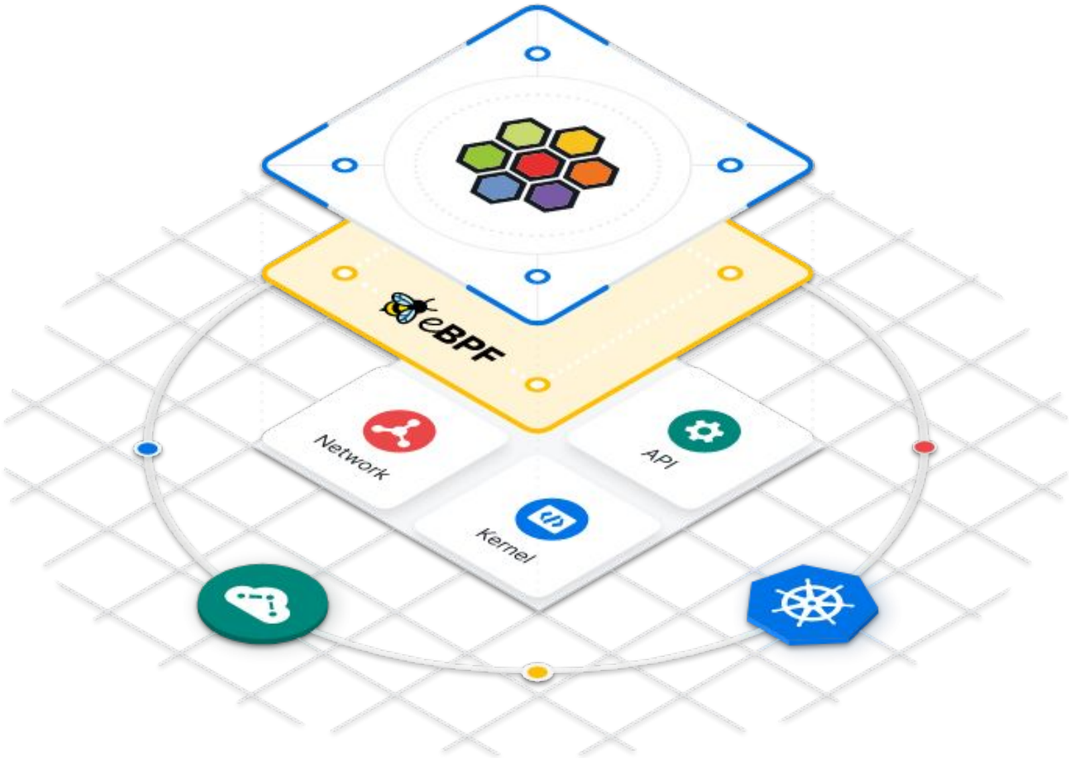
CICD

CNI

kops

Why changing the CNI plugin?

- Additional functionality
- Encryption
- Why Cilium?



Transparent Encryption

CLOUD NATIVE
COMPUTING FOUNDATION

Identity-aware Visibility

Advanced Network Policy

Network Metrics + Policy
Troubleshooting

Multi-cluster Connectivity

Meltwater

Requirements

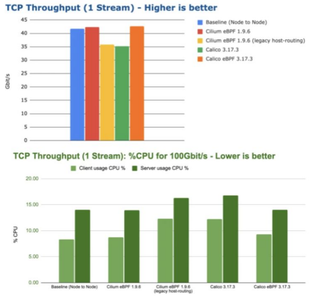
- Seamless
- In place
- Same tooling, and native routing
- No components downtime
- No Kubernetes upgrade

Start small and learn the ropes

cilium GitHub Stars 11.7k Enterprise Learn News and media Documentation

May 11, 2021

CNI Benchmark: Understanding Cilium Network Performance



Understanding Cilium Network Performance

Hello 👋

As more crucial workloads are being migrated to Kubernetes, netv
important selection criteria when deciding what network layer to l
we'll explore the performance characteristics of Cilium based on e
few weeks. Upon popular request, we are also including measurer

However, instead of just presenting the numbers, we are going to
benchmarking a bit deeper and look at:

- The Throughput Benchmark

Search or jump to... Pull requests Issues Marketplace Explore

cilium / cilium-perf-networking Public

Code Issues Pull requests Actions Projects Wiki Security Insights

master 2 branches 0 tags Go to file Add file Code

jibi and kkourt cilium: add support for v1.10 and -rc version 66796a5 on 12 May 2021 30 commits

playbooks	calico: optimize BPF datapath conf	12 months ago
scripts	cilium: add support for v1.10 and -rc version	12 months ago
terraform	cilium performance evaluation scripts	2 years ago
.gitignore	Add zh-lab34-hosts.ini to .gitignore	12 months ago
LICENSE	cilium performance evaluation scripts	2 years ago
README.md	README: notes	2 years ago

README.md

Introduction

https://docs.cilium.io/en/latest/operations/performance/

Additional Details

Installing Cilium

This playbook installs cilium using `kubadm` and `helm`.

Which cilium versions are supported?

We use cilium-install-with-helm.sh script to install cilium. You can pass a version available in the cilium helm repo (e.g., `1.8.3`). You can also pass a version such as `1.8` and the script will try and pick the latest one. Finally, you can also use custom docker images as the example below:

cilium Star 11721

Search docs

GETTING STARTED

- Introduction to Cilium & Hubble
- Getting Started Guides
  - Installation
  - Network Policy Security Tutorials
- Advanced Networking
  - Setting up Cilium in AWS ENI mode
    - Create an AWS cluster
    - Disable VPC CNI (`aws-node` DaemonSet) (EKS only)
    - Deploy Cilium
    - Create a node group
    - Validate the Installation
    - Specify Environment Variables
    - Enable Hubble for Cluster-Wide Visibility
    - Limitations
    - Troubleshooting
  - Using kube-router to run BGP
  - Using BIRD to run BGP
  - Setting up Cluster Mesh
  - Cilium integration with Flannel (beta)
  - IPVLAN based Networking (beta)

Setting up Cilium in AWS ENI mode

Note

The AWS ENI integration is still subject to some limitations. See Limitations for details.

Create an AWS cluster

Setup a Kubernetes on AWS. You can use any method you prefer, but for the simplicity of this tutorial, we are going to use eksctl. For more details on how to set up an EKS cluster using `eksctl`, see the section Installation on AWS EKS.

Copy Line

```
eksctl create cluster --name test-cluster --without-nodegroup
```

Disable VPC CNI (`aws-node` DaemonSet) (EKS only)

If you are running an EKS cluster, you should delete the `aws-node` DaemonSet.

Cilium will manage ENIs instead of VPC CNI, so the `aws-node` DaemonSet has to be deleted to prevent conflict behavior.

Note

Once `aws-node` DaemonSet is deleted, EKS will not try to restore it.

Copy Line

```
kubectl -n kube-system delete daemonset aws-node
```

Deploy Cilium

Note

First, make sure you have Helm 3 installed. Helm 2 is no longer supported.

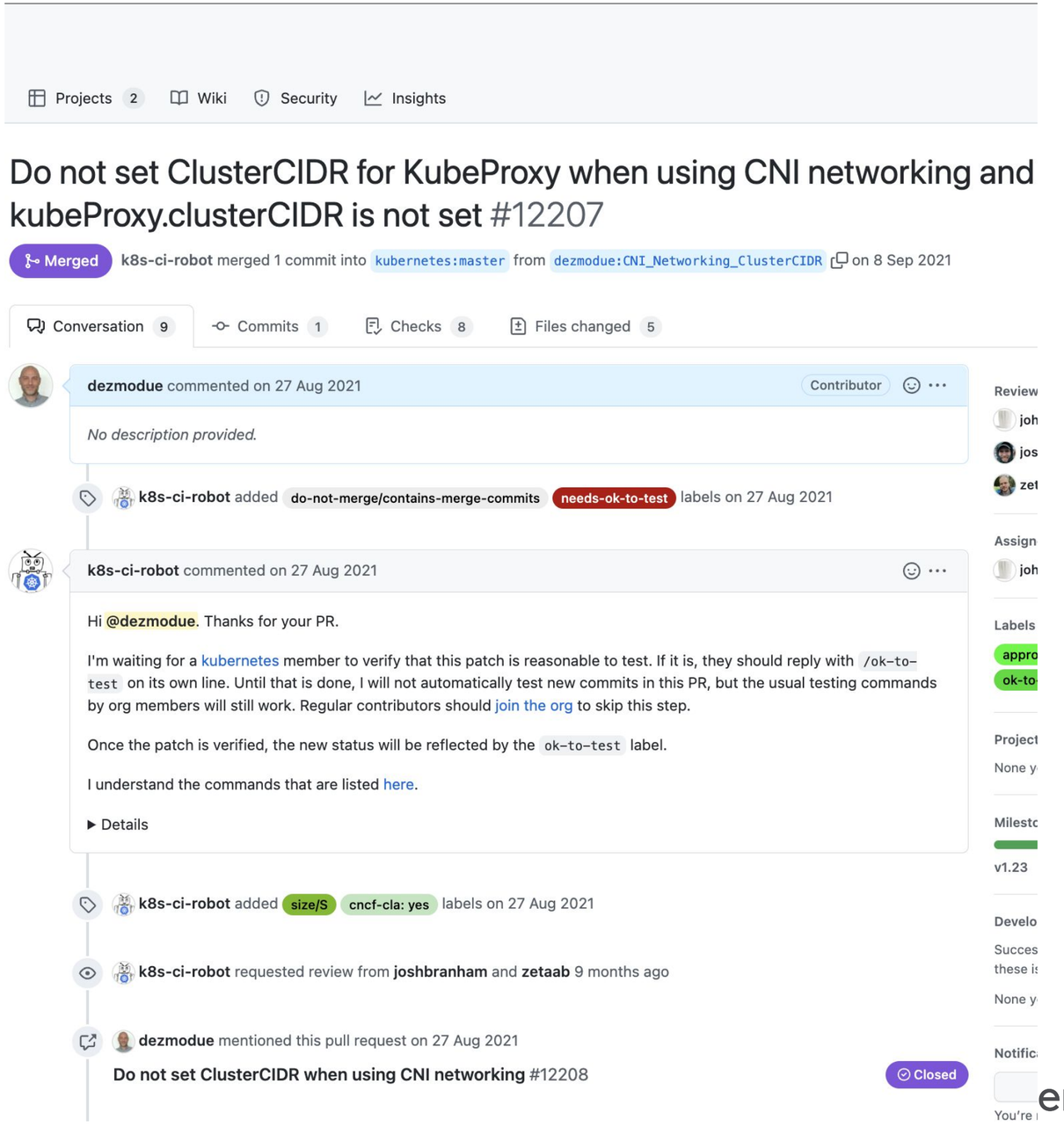
Setup Helm repository:

Copy Line

Meltwater

Adapt tooling

Cilium + kops ?

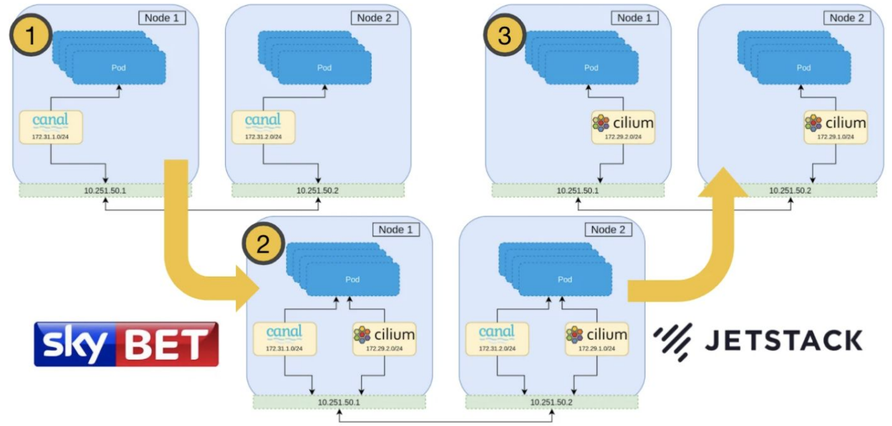


Migration procedure

- Inspiration from blog post
- Easier with native routing
- Intermediate state with both CNI plugins
- Multiple days rollout

Oct 06, 2020

How to perform a CNI Live Migration from Flannel+Calico to Cilium



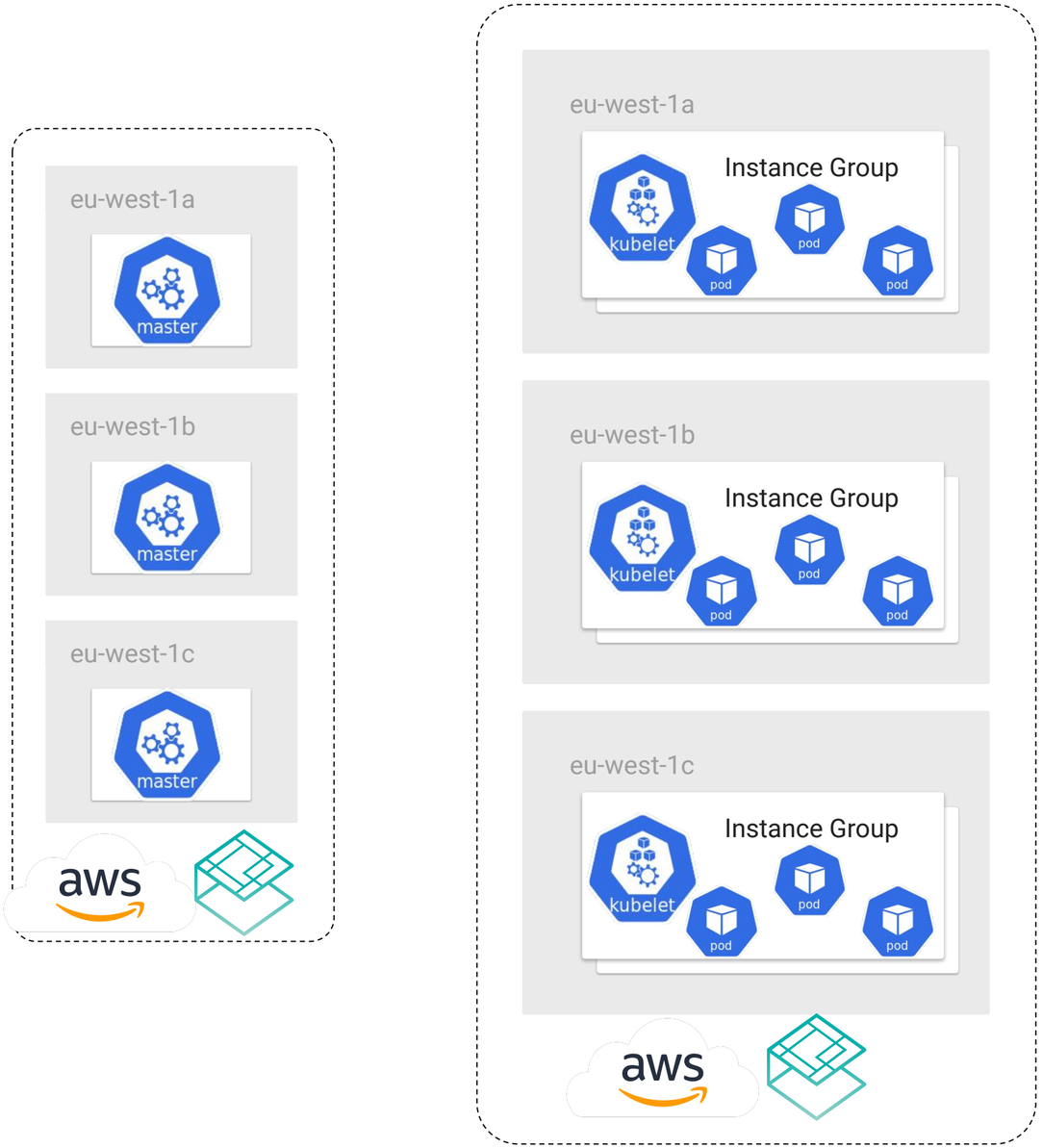
This is a guest blog by Josh Van Leeuwen and covers how Josh implemented a CNI live migration for a customer, Sky Betting and Gaming, to live migrate a cluster running the Canal CNI plugin (Flannel+Calico) to Cilium. Josh is working as a Customer Reliability Engineer for Jetstack. Read on to hear what Josh has to say...

In this post we'll discuss why one might want to change CNIs, what I have learnt developing a solution for live migration, and how it all works.

What is CNI, and why change it?

Container Network Interface (CNI) is a big topic, but in short, CNI is a set of specifications that define an interface used by container orchestrators to set up networking between containers. In the Kubernetes space, the Kubelet is responsible for calling the CNI installed on the cluster so Pods are attached to the Kubernetes cluster network during creation, and its resources are properly released during deletion. CNIs can also be responsible for more advanced features than just setting up routes in the cluster, such as network policy enforcement, encryption, load balancing, etc.

Initial state

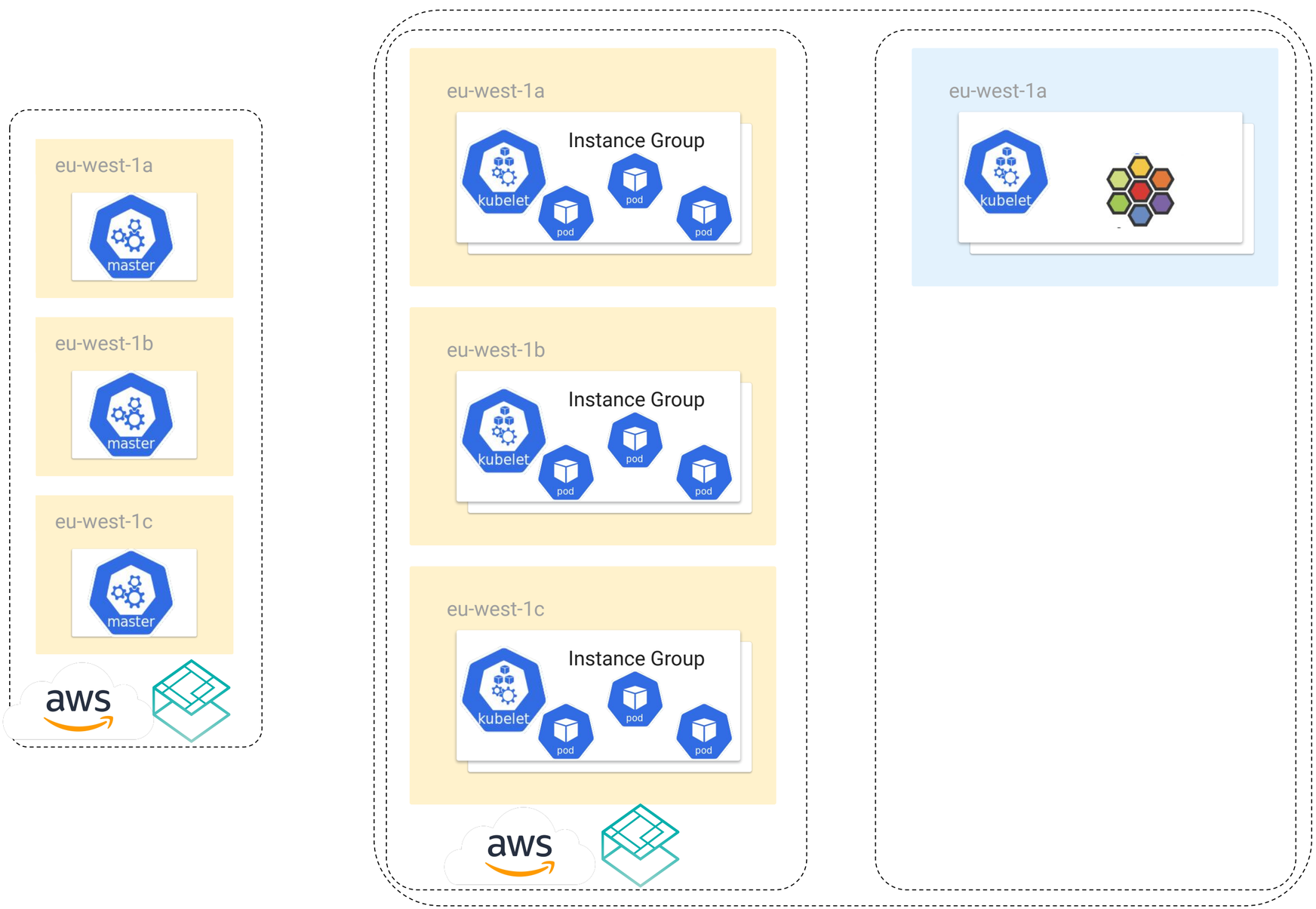


Labelling

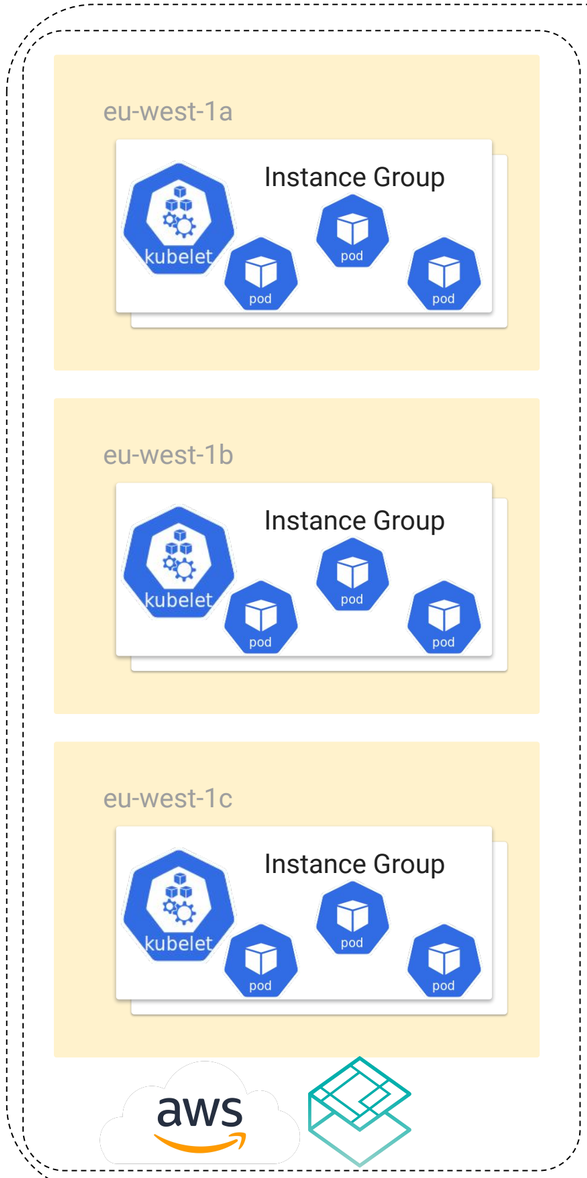
label cniplugin=aws

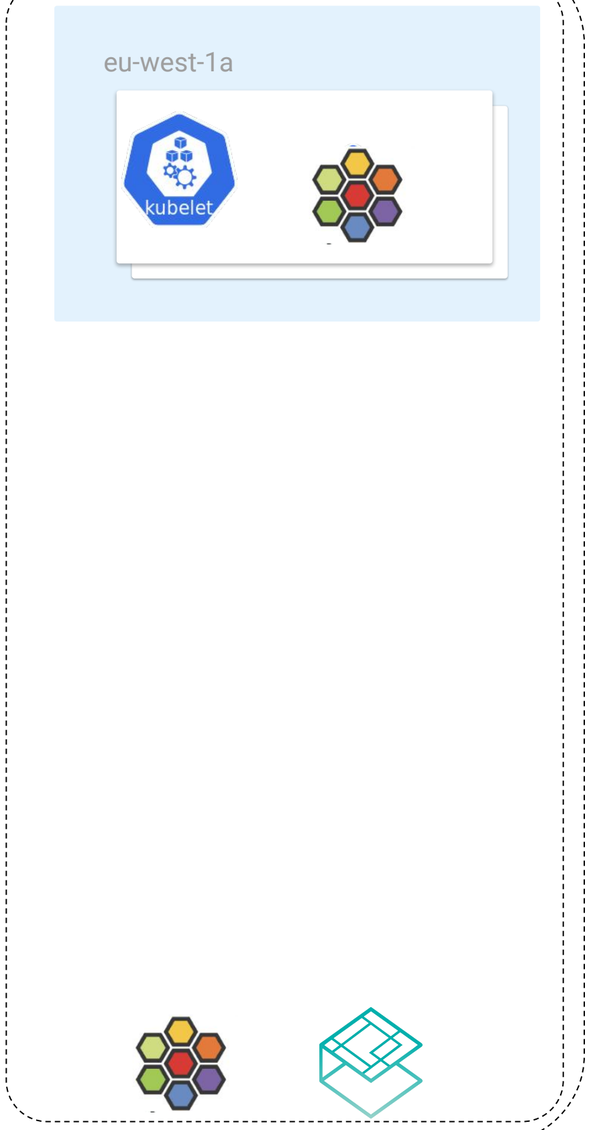
Meltwater

Cilium operator



Cilium daemonset

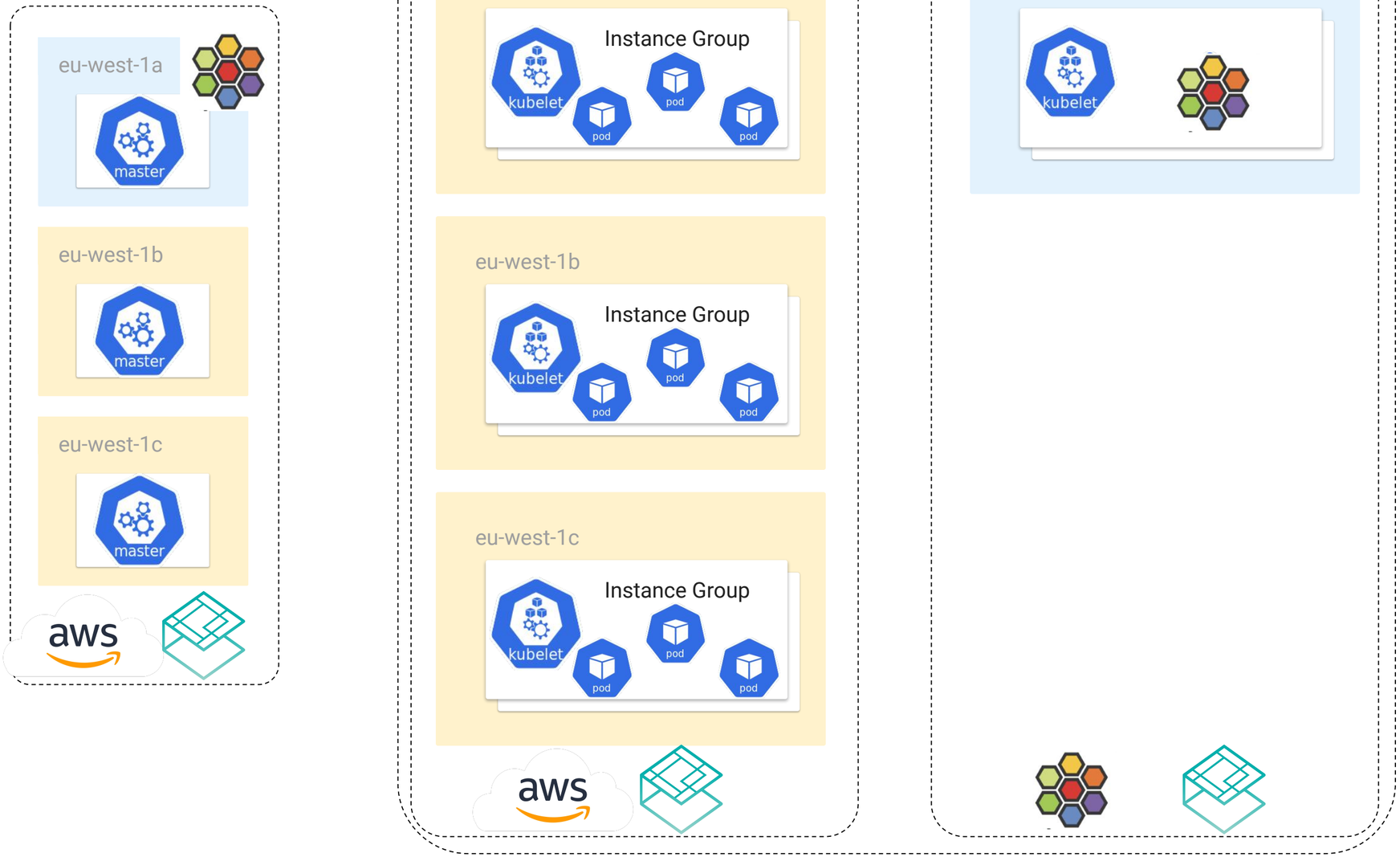




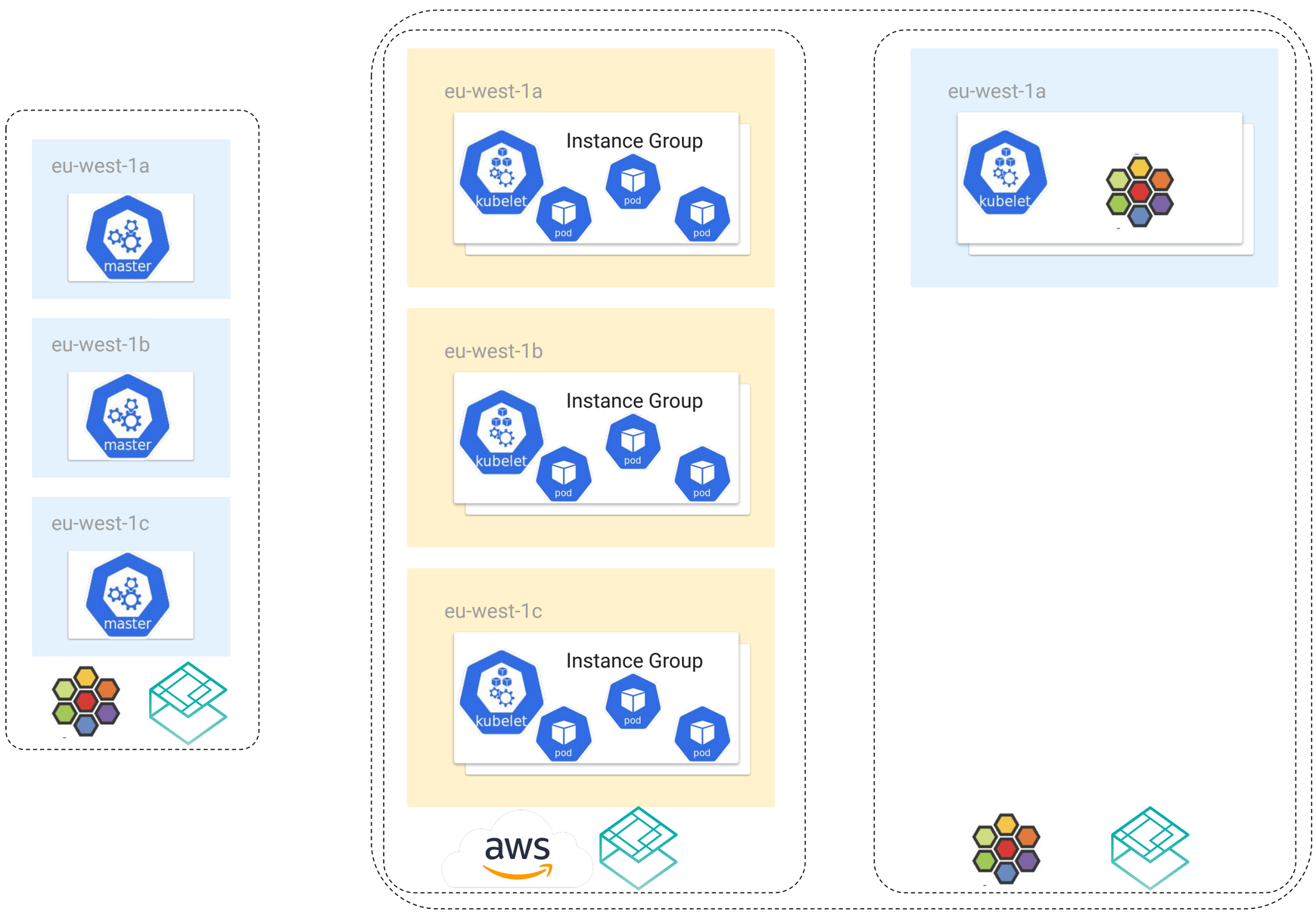
Cilium Agent

Meltwater

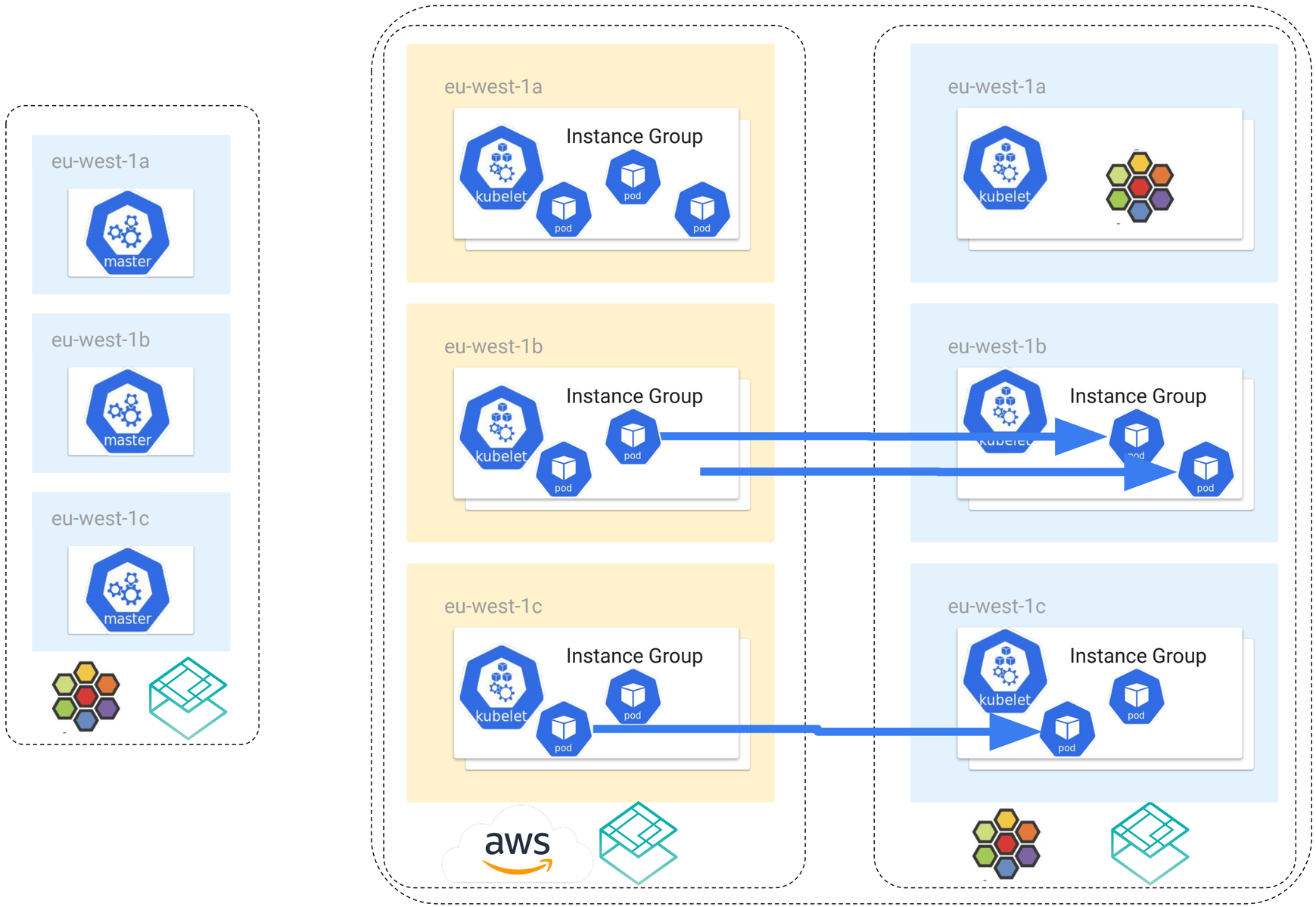
Control plane migration



Control plane migration



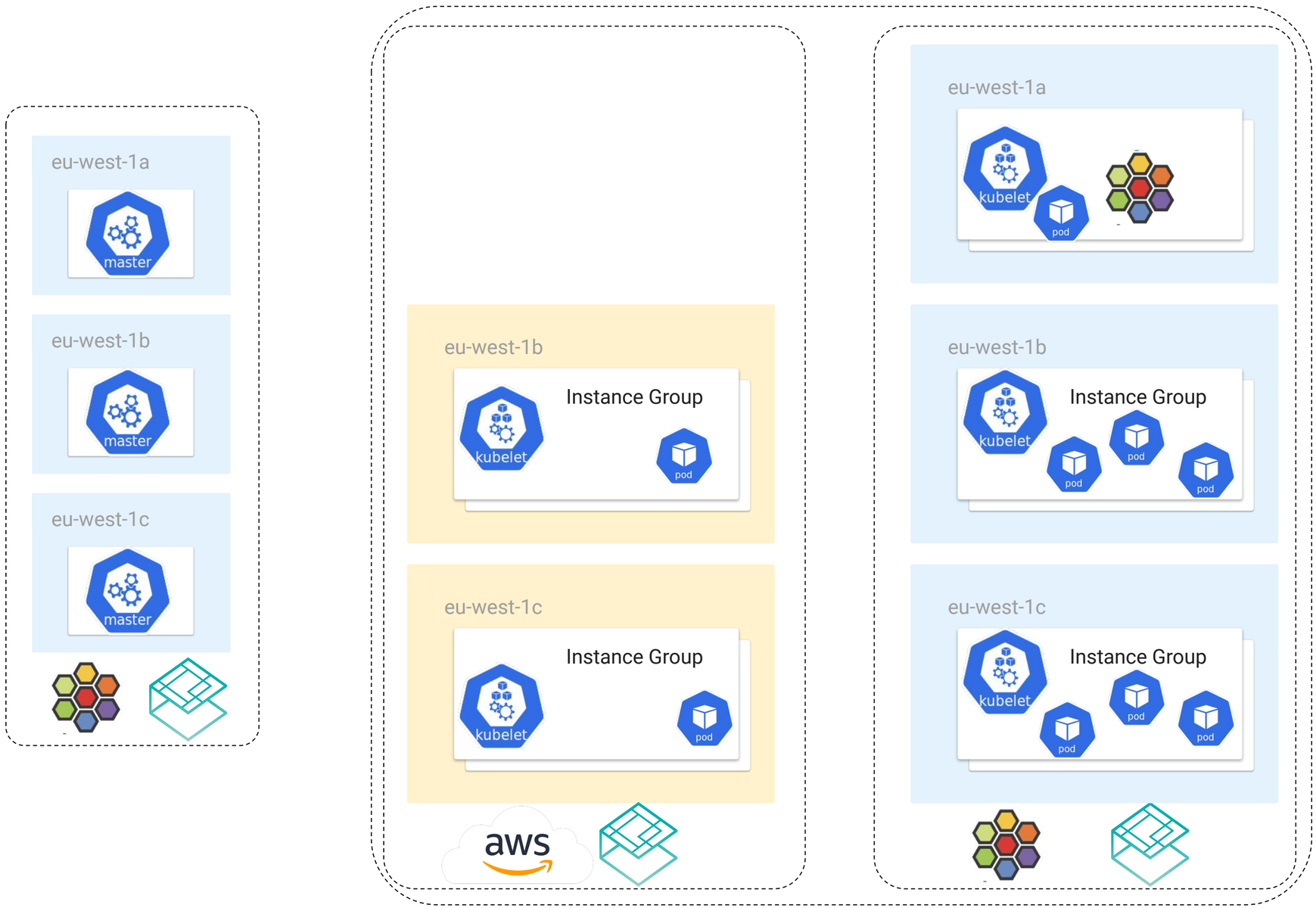
Pods migration



Migrate workloads

Meltwater

Decommissioning



Final state


eu-west-1a
master
eu-west-1b
master
eu-west-1c
master


eu-west-1a
kubelet
pod
eu-west-1b
kubelet
Instance Group
pod
pod
pod
eu-west-1c
kubelet
Instance Group
pod
pod
pod

Prepare for rollout

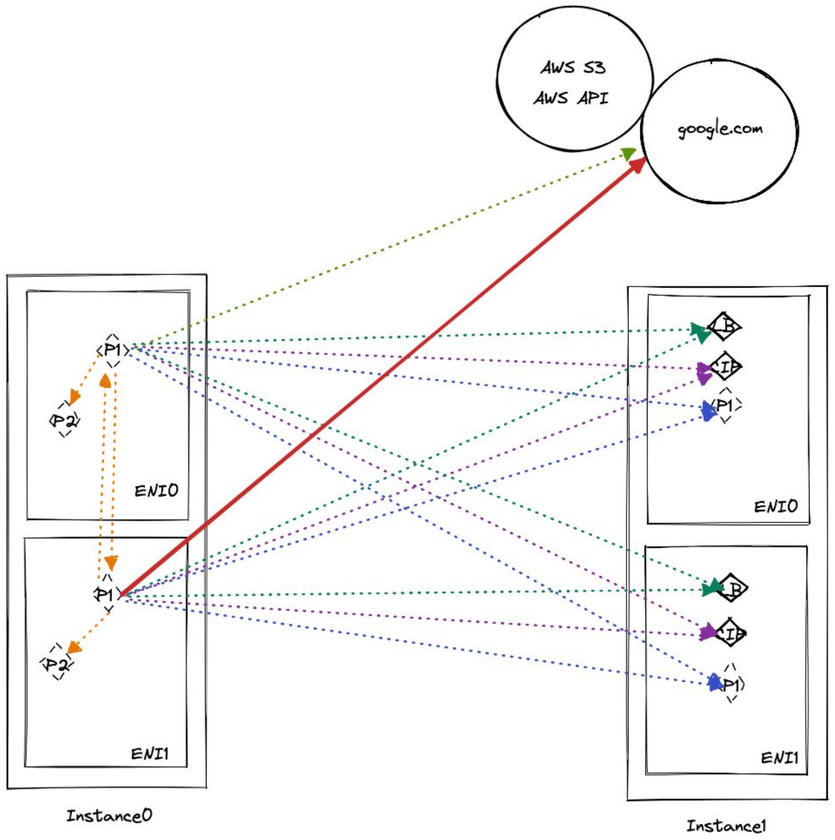
- Codified the procedure
- Testing
- Communication

Dogfooding

- The value of dogfooding
- More controlled environment

Meltwater

More tooling - Bravo pinger




```
egressMasqueradeInterfaces=ens+
```

Show time

Meltwater

Rollout hiccups

- Increased resource requests from daemonsets
- Instances with lower IP limits

Success!!

```
❯ date;cilium status; date
Thu Feb 10 18:24:25 CET 2022
    /¯¯\
 /¯¯\__/¯¯\    Cilium:          OK
 \__/¯¯\__/    Operator:        OK
 /¯¯\__/¯¯\    Hubble:          OK
 \__/¯¯\__/    ClusterMesh:     disabled
    \__/

DaemonSet              cilium                   Desired: 404, Ready: 404/404, Available: 404/404
Deployment             cilium-operator          Desired: 2, Ready: 2/2, Available: 2/2
Deployment             hubble-relay             Desired: 1, Ready: 1/1, Available: 1/1
Deployment             hubble-ui                Desired: 1, Ready: 1/1, Available: 1/1
Containers:            cilium                   Running: 404
                       cilium-operator          Running: 2
                       hubble-relay             Running: 1
                       hubble-ui                Running: 1
Cluster Pods:          7560/7928 managed by Cilium
Image versions         hubble-relay             quay.io/cilium/hubble-relay:v1.10.4@sha256:be17169d2b68a974e9e27bc1
                       hubble-ui                quay.io/cilium/hubble-ui:v0.7.9@sha256:e0e461c680ccd083ac24fe4f9e19
                       hubble-ui                quay.io/cilium/hubble-ui-backend:v0.7.9@sha256:632c938ef6ff30e3a080
                       hubble-ui                docker.io/envoyproxy/envoy:v1.18.2@sha256:e8b37c1d75787dd1e712ff389
                       cilium                   quay.io/cilium/cilium:v1.10.4@sha256:7d354052ccf2a7445101d78cebd144
                       cilium-operator          quay.io/cilium/operator-aws:v1.10.4@sha256:45df7a09f8278a9c2313fa7d
Thu Feb 10 18:28:48 CET 2022
```

Celebrate

Meltwater

Success?!?

It is DNS!

It is always DNS!

It is not DNS...

`getaddrinfo ENOTFOUND`

Success?!?

It is AWS Limits!

It is not AWS limits...

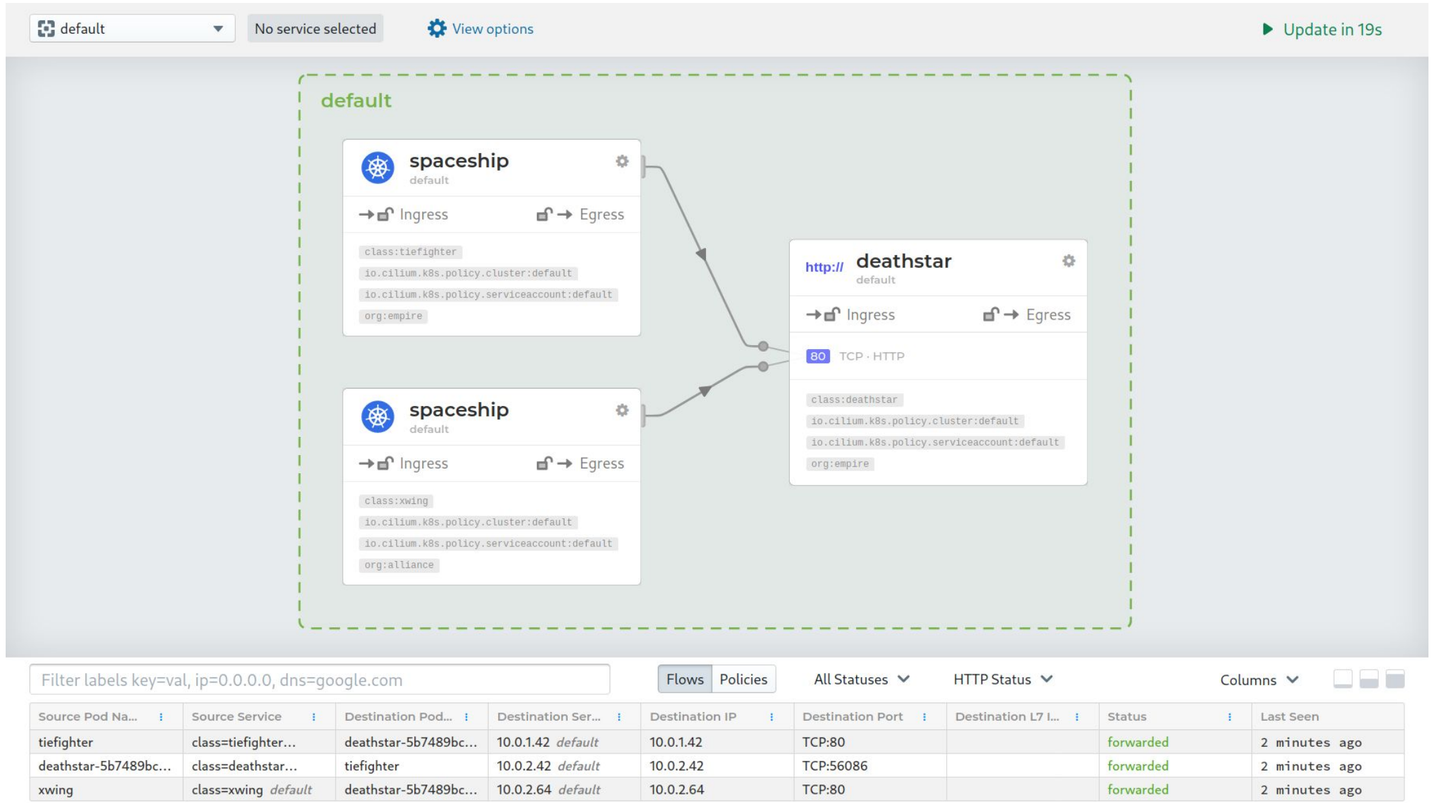
Network Performance Metrics

Additional instance level network performance metrics are available on instances using ENA both for Windows and Linux operating systems. These include:

- ***bw_in_allowance_exceeded*** and ***bw_out_allowance_exceeded*** indicates the number of packets that are shaped and dropped because of instance aggregate bandwidth exceeded BW allowance for the instance.
- ***pps_allowance_exceeded*** indicates number of packets shaped and dropped due to Packet Per Second (PPS) exceeding the allowance for the instance. PPS allowance is enforced separately to the overall bandwidth allowance and while the instance may still be under overall bandwidth allowance the PPS allowance may exceed if the mean packet size is small.
- ***conntrack_allowance_exceeded*** indicates the number of packets shaped and dropped due to exhaustion of tracked session allowance for the instance, new sessions will fail to establish once this allowance is exceeded. Sessions are restored once instance session count drops below the allowance.
- ***linklocal_allowance_exceeded*** indicates the number of packets shaped and dropped due to PPS rate allowance exceeded for local services such as Route 53 DNS Resolver, Instance Metadata Service, Amazon Time Sync Service. This often points to suboptimal design choices or misconfiguration. This allowance is same across instances.

Success?!?

Could it be Cilium?



Routes

```
cilium sysdump --help
```

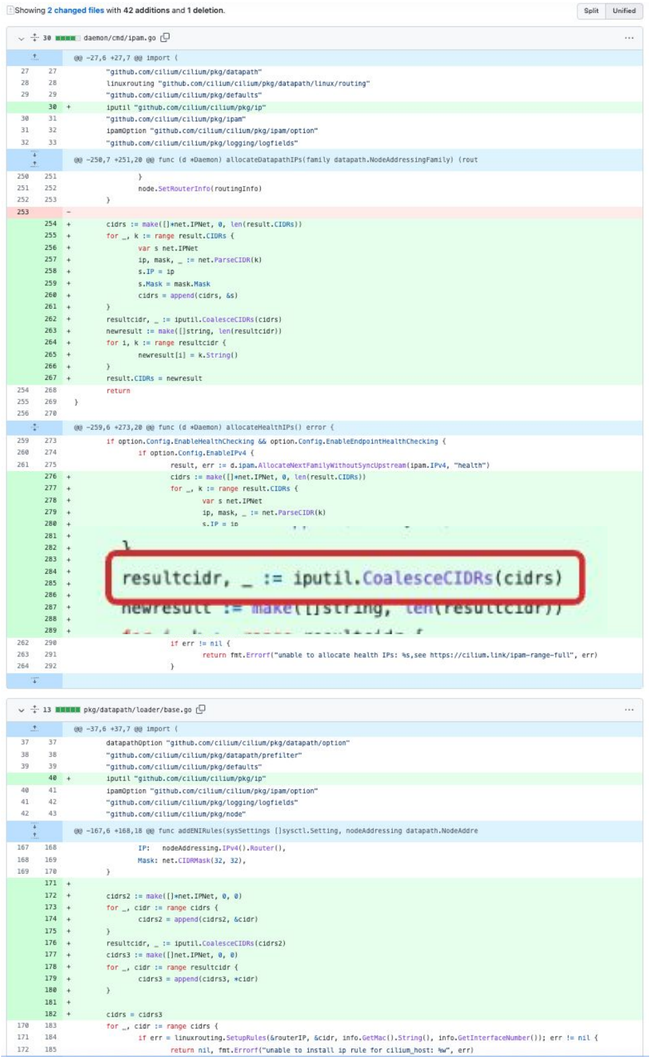
```
root@ip-10-103-220-24:/home/cilium# ip rule | fgrep 10.103.205.47
20:     from all to 10.103.205.47 lookup main
111:    from 10.103.205.47 to 10.103.0.0/16 lookup 10
111:    from 10.103.205.47 to 10.0.0.0/8 lookup 11
```

```
root@ip-10-103-220-24:/home/cilium# ip route show table 10
default via 10.103.192.1 dev ens5
10.103.192.1 dev ens5 scope link
```

```
root@ip-10-103-220-24:/home/cilium# ip route show table 11
default via 10.103.192.1 dev ens6
10.103.192.1 dev ens6 scope link
```

Give back

Contribute



Fin

@meltwaterEng

https://underthehood.meltwater.com

https://github.com/meltwater/MeltwaterEng-public-presentations

Meltwater