



Envoy's Using 10GB of Memory and It's All My Fault!

Steve Sloka

slokas@vmware.com

@stevesloka



About Me



Sr. Member of Technical Staff

- K8s since 2014
- Kubernetes Contributor
- Contour / Gimbal @ VMware
- CKA / CKAD
- Pittsburgh Kubernetes Meetup Organizer

Observed envoy memory while adding|removing ingress #499



mattalberts opened this issue on Jul 2, 2018 · 47 comments

“Envoy is using a lot of memory!”



Benefits of Contour

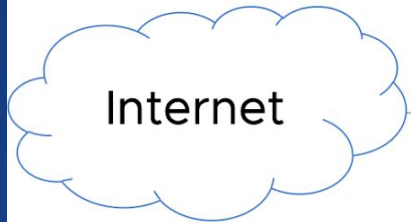


- Traffic consolidation
- Abstract ingress configuration
 - Teams
- TLS management
- Path / Header based routing



kubernetes

GET https://xyz.com/blog



Load Balancer

L7/L4 Proxy



Ingress
Controller



web-1

web-2

app: web-app

blog-1

blog-2

app: blog



kubernetes

Services

Certificates
(Secrets)

Endpoints

"HTTP" Routes

HTTP Listeners
(80/443)



Directed Acyclic Graph (DAG)



CONTOUR



xDS Server (gRPC)



CDS

SDS

EDS

RDS

LDS



envoy

container_memory_usage_bytes{namespace="ingress-system",pod_name=~"contour-ingress-perf.*"}

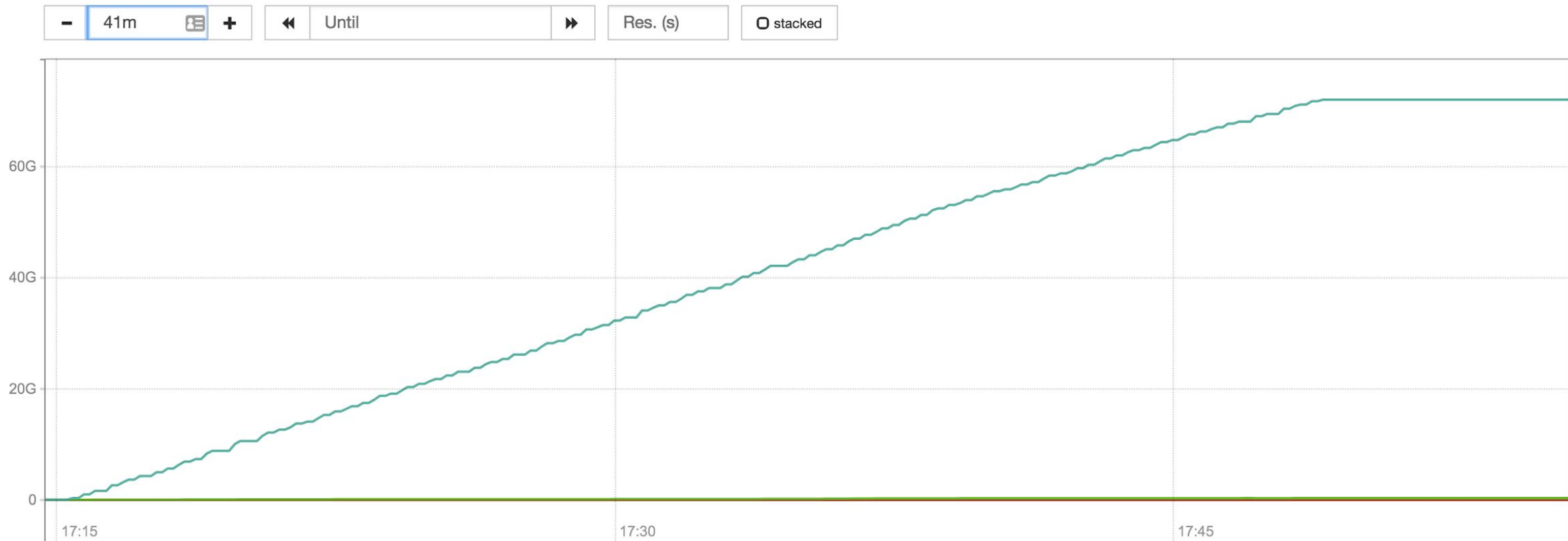
Load time: 53ms
Resolution: 9s
Total time series: 3

Execute

- insert metric at cursor -

Graph

Console



✓ container_memory_usage_bytes{beta_kubernetes_io_arch="amd64",beta_kubernetes_io_os="linux",container_name="envoy",failure_domain_beta_kubernetes_io_region="pdx01",failure_domain_beta_kubernetes_io_zone="225",id="/kubep

✓ container_memory_usage_bytes{beta_kubernetes_io_arch="amd64",beta_kubernetes_io_os="linux",container_name="contour",failure_domain_beta_kubernetes_io_region="pdx01",failure_domain_beta_kubernetes_io_zone="225",id="/kubep

✓ container_memory_usage_bytes{beta_kubernetes_io_arch="amd64",beta_kubernetes_io_os="linux",container_name="POD",failure_domain_beta_kubernetes_io_region="pdx01",failure_domain_beta_kubernetes_io_zone="225",id="/kubepo

Reproducing

- **Start:** No ingress resources
- **Insert:** 5000 total in blocks of 100
- **Stable:** Insert has finished, containers are idle
- **Restart:** Delete the Envoy pod and let it recreate to see the difference in memory vs stable

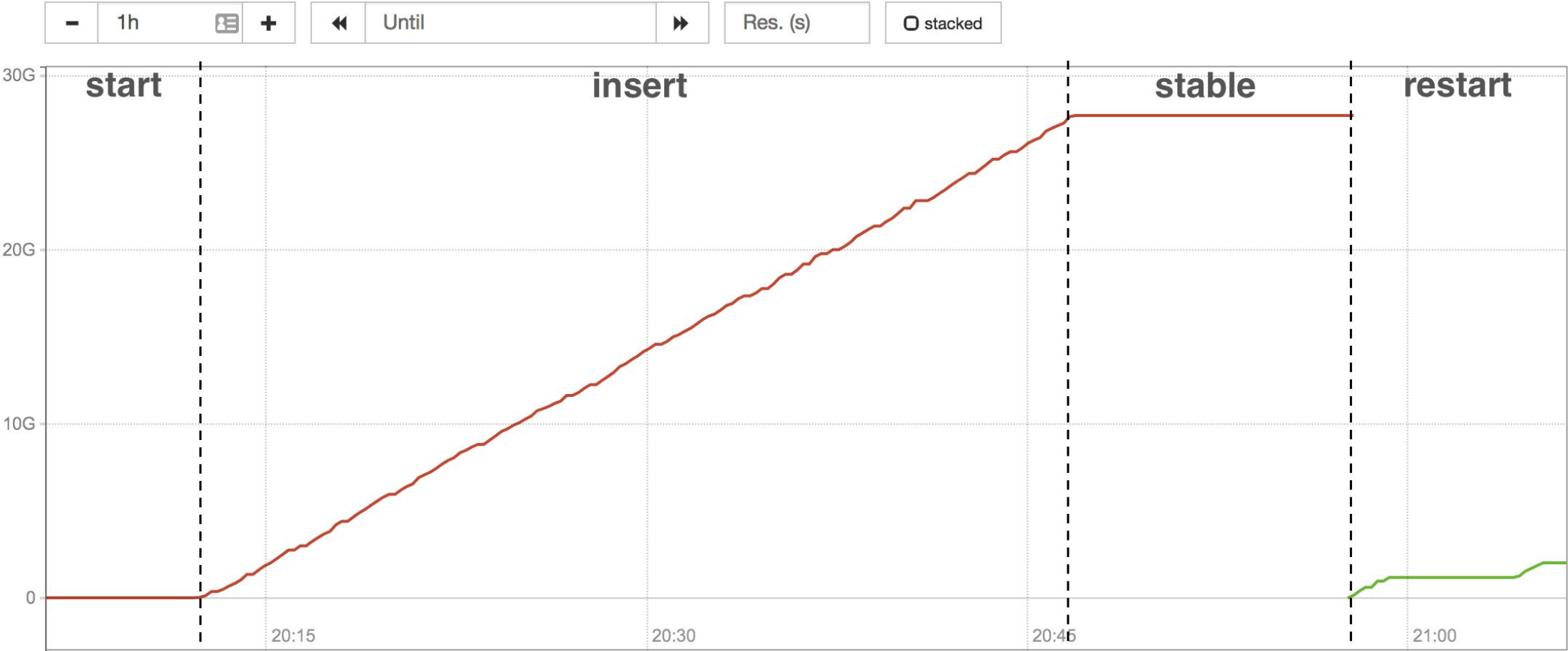
```
container_memory_usage_bytes{namespace="ingress-system",pod_name=~"contour-ingress-debug.*",container_name="envoy"}
```

Load time: 44ms
Resolution: 14s
Total time series: 2

Execute

- insert metric at cursor -

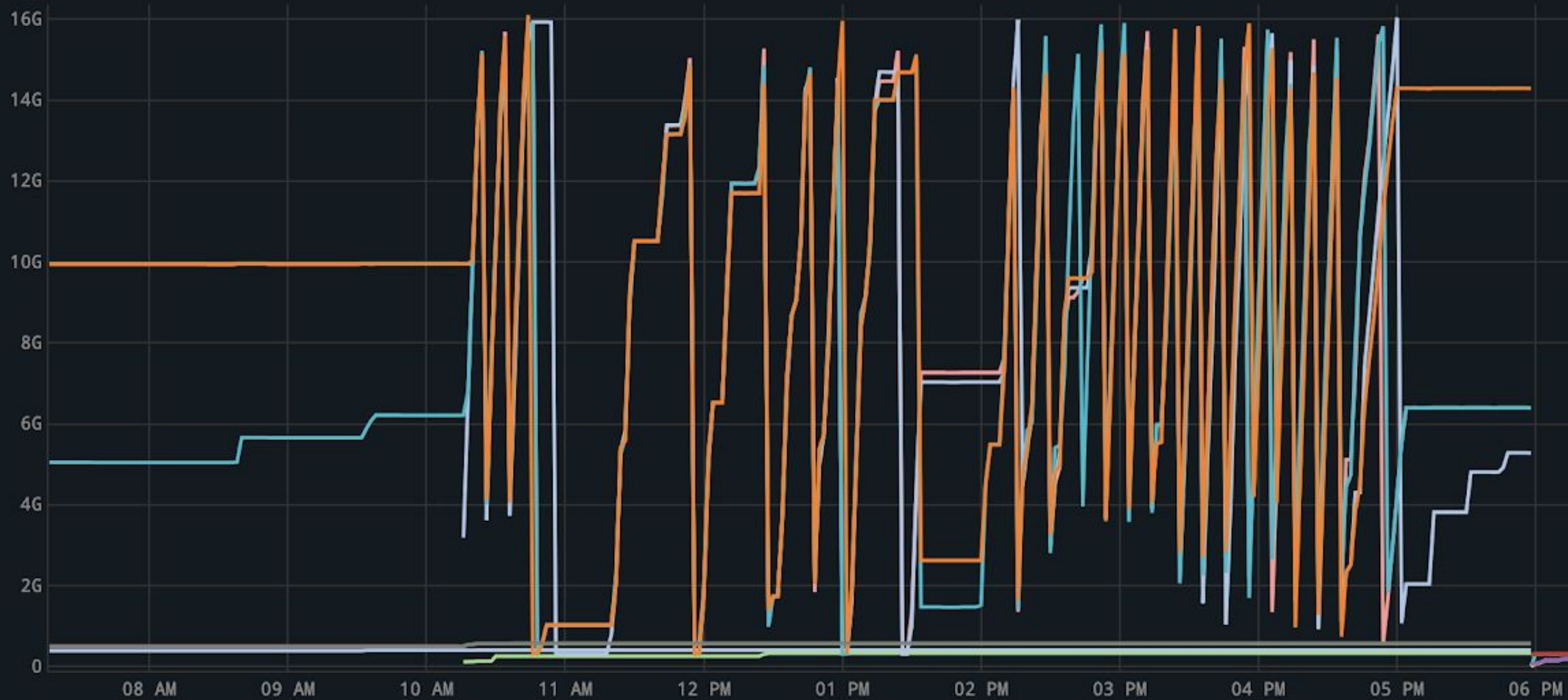
Graph Console



container_memory_usage_bytes{beta_kubernetes_io_arch="amd64",beta_kubernetes_io_os="linux",container_name="envoy",failure_domain_beta_kubernetes_io_region="pdx01",failure_domain_beta_kubernetes_io_zone="us-east-1a",failure_domain_beta_kubernetes_io_availability="a-1",failure_domain_beta_kubernetes_io_hostname="ip-10-0-1-10.ec2.amazonaws.com"}
container_memory_usage_bytes{beta_kubernetes_io_arch="amd64",beta_kubernetes_io_os="linux",container_name="envoy",failure_domain_beta_kubernetes_io_region="pdx01",failure_domain_beta_kubernetes_io_zone="us-east-1a",failure_domain_beta_kubernetes_io_availability="a-1",failure_domain_beta_kubernetes_io_hostname="ip-10-0-1-10.ec2.amazonaws.com"}

Problem Detail

- Changes to Secrets (e.g. Certificates) caused update to LDS
- High rate of change causes lots of old configuration
 - Old listeners go into “draining” (Default 600s)
 - Each draining listener holds onto memory



The “Current” Solution

1. Sort secrets before sending updates to Envoy
2. Don't ignore the ACK/NACK
3. Wait for Envoy to send the next discovery request
4. “Batch” up requests

Tuesday • 11:50am (Room 28ABCDE)

“Take Envoy Beyond a K8s Service Mesh
- to Legacy Bare Metal and VMs + More”

Thank You

@stevesloka

#contour (K8s Slack)

projectcontour.io