Segment routing in container networks

Ben de Graaff

Supervisor: Marijke Kaat (SURFnet/UvA)
Best path
Arbitrary paths

A > 1 > 4 > B

Background
Pure IPv6 (SRv6)
Container networks
Platform independent

Background

Internet

docker

LXC

cilium
Multi-tenancy

Background

Internet
Example: load balancer

Background

Transit policy
Research

State of segment routing in IPv6

Proof of concept:
Container networking
Network functions
Network programming

src > 1 > 4 > dst
Segment routing header

- Header
- Segments left
- Segment 0
- ...
- Segment N
- Extensions
Proof of concept

Validate policy

Apply policy
Container discovery/mobility

Routing opaque addresses

Results

1: Inject SRH
2: Compute final hop
3: Deliver
Implementation

Hardware/software

Results
Linux kernel 4.10+

Basic routing/policy

Limited extension support

Implementation quirks...
SRv6 availability

Hardware

Software

Vector Packet Processing

Results

NCS 5500

The Fast Data Project
Technical implementation
Virtual topology

VM

Container

VPP

Container

VPP

Container

Network function

VM
eBPF

Latest & greatest

Process directly in kernel

Fast, powerful

myprog.c + LLVM = bytecode

\[\downarrow\]

bpf() syscall

\[\downarrow\]

Kernel network stack
eBPF

- Ingress eBPF
- Validate policy
- eth0
- Control app
- Egress eBPF
- Apply policy
- veth
- Container

Results

✗
✗
Linux do-it-yourself

Tun/tap
AF_PACKET
ip rule
iptables fwmark
PF_RING
...

Results
Linux do-it-yourself

Validate policy

Ingress app

AF_PACKET

eth0

veth

Container

ip rule

Apply policy

Egress app

tun
Summary

Validate policy

Apply policy

Results
Ingress policy enforcement?

Discussion

Internet

FW

eth0

Ingress

veth

Container

Egress
Future work

Ingress path control

Linux segment routing Netlink API

Develop useful extensions
Conclusion

Proof of concept: works

SDN easy, at cost of overhead

Hardware not strictly required
Related work

Cisco, Bell Canada, Comcast, et al, technical workshops @ [www.segment-routing.net](http://www.segment-routing.net)

NFV with SRv6, with SRH unaware hosts (NetSoft 2017, presented *today*)
Segment routing in container networks

Segment routing is effective at enabling **SDN** and **network functions** between **containers**

However, it is **not yet widely supported** in hardware, software

Check out the report for a full list of references

Backup slides
Security/RH0

Enforce policy at network edges

SIDs must be explicitly enabled

HMAC: check at ingress
Simplify the network

https://xkcd.com/927/
MPLS

No LDP, RSVP required
Multi-tenancy

Layer 2 and 3 cross-connects

Multi-tenancy: Segment ID or extension?