Accelerate Istio-CNI with ebpf

Xu Yizhou & Guo Ruijing

Agenda

- Istio-CNI
- tcp/ip stack overhead between sidecar and service
- Background knowledge of ebpf
- Acceleration for Inbound/Outbound/Envoy to Envoy

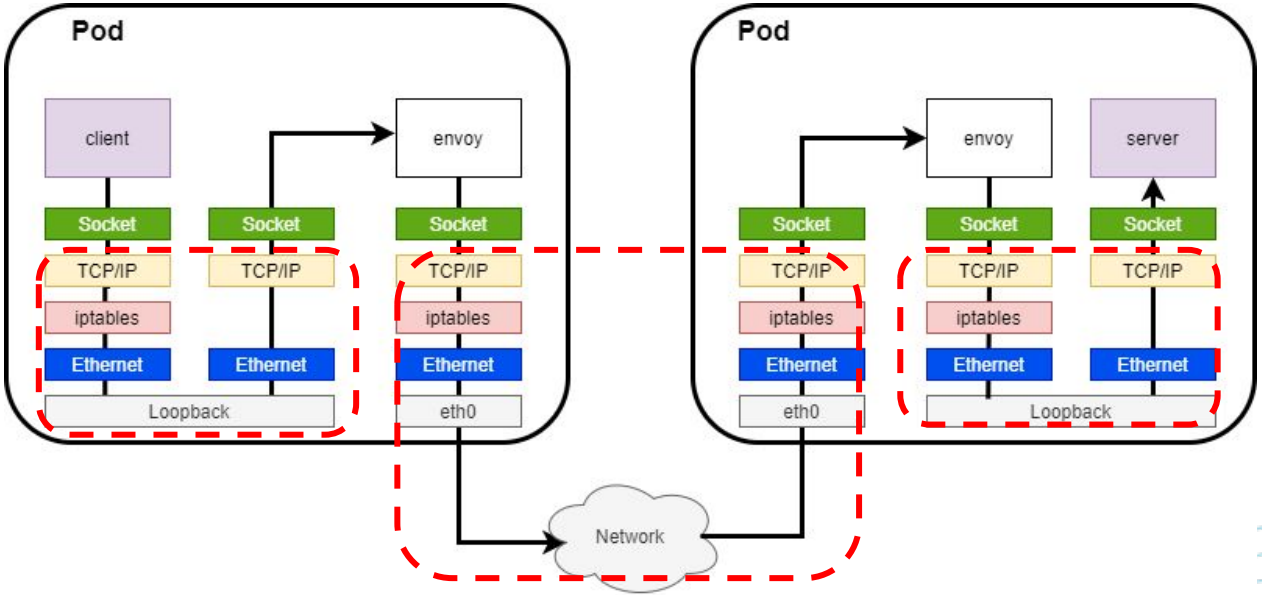
Istio-CNI

- The Istio CNI plugin performs the Istio mesh pod traffic redirection in the Kubernetes pod life-cycle's network setup phase,
- Removing the requirement for the NET_ADMIN and NET_RAW capabilities for users deploying pods into the Istio mesh.
- The Istio CNI plugin replaces the functionality provided by the istio-init container.

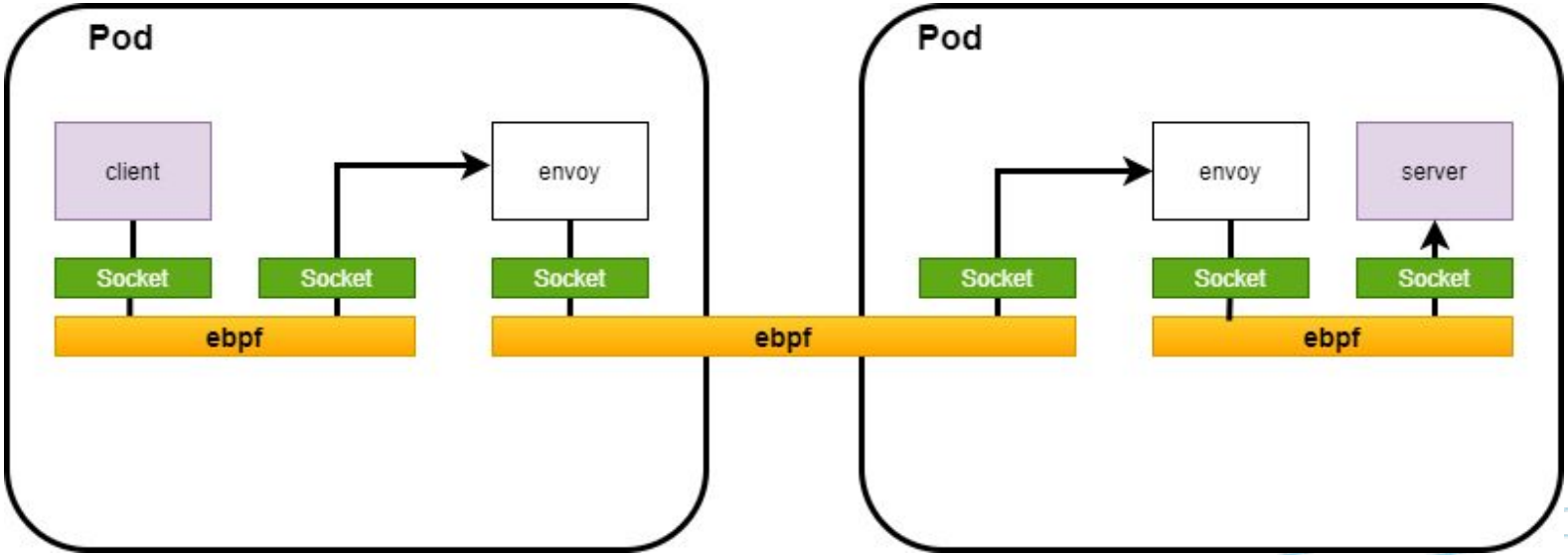
Tcp/ip stack overhead between sidecar and service

Overhead sidecar traffic from 3 scopes

- Inbound
- Outbound
- Envoy to Envoy(same host)



Dataflow After Acceleration(same host)



ebpf Background Knowledge

Prog type

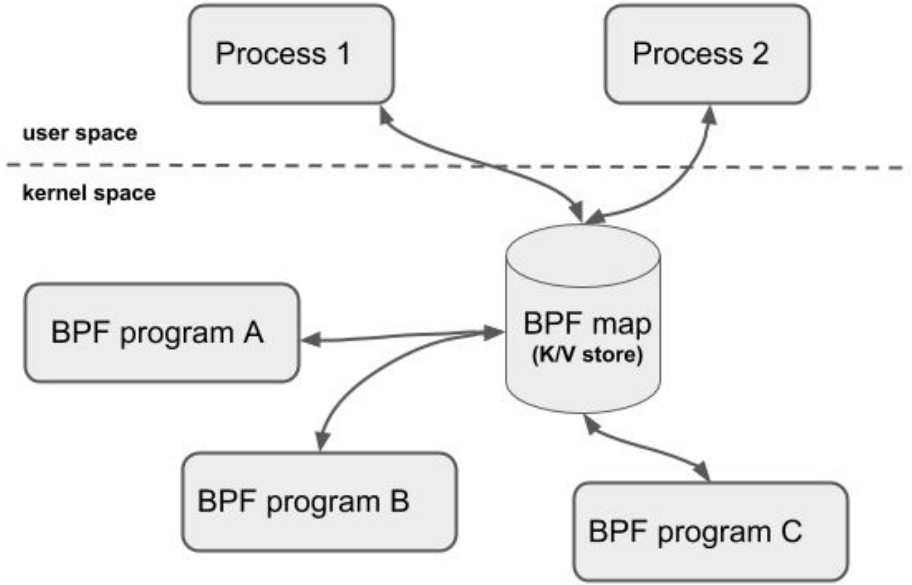
- ebpf provide various programs type for different purpose
- We choose SOCK_OPS & SK_SKB to implement function

```
enum bpf_prog_type {
    BPF_PROG_TYPE_UNSPEC,
    BPF_PROG_TYPE_SOCKET_FILTER,
    BPF_PROG_TYPE_KPROBE,
    BPF_PROG_TYPE_SCHED_CLS,
    BPF_PROG_TYPE_SCHED_ACT,
    BPF_PROG_TYPE_TRACEPOINT,
    BPF_PROG_TYPE_XDP,
    BPF_PROG_TYPE_PERF_EVENT,
    BPF_PROG_TYPE_CGROUP_SKB,
    BPF_PROG_TYPE_CGROUP_SOCK,
    BPF_PROG_TYPE_LWT_IN,
    BPF_PROG_TYPE_LWT_OUT,
    BPF_PROG_TYPE_LWT_XMIT,
    BPF_PROG_TYPE_SOCK_OPS,
    BPF_PROG_TYPE_SK_SKB,
};
```

ebpf Background Knowledge

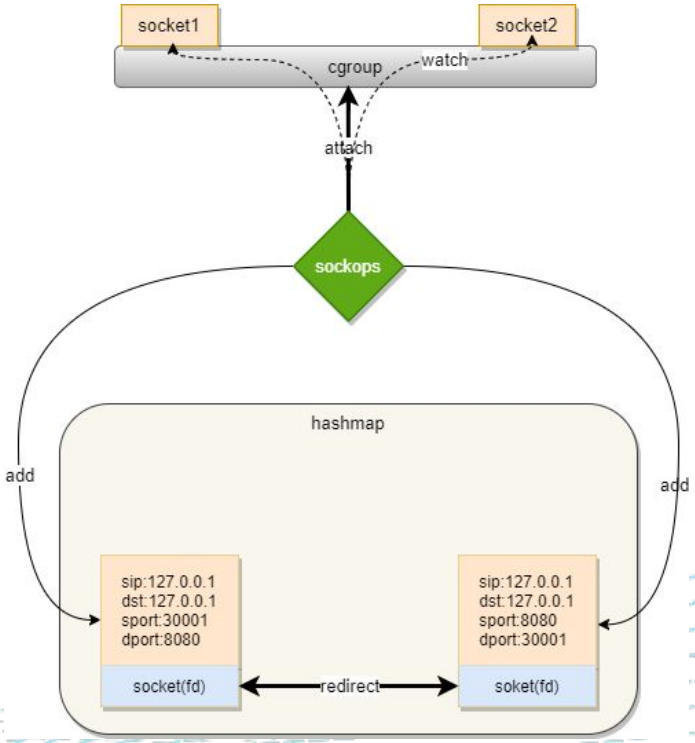
map

- Share collected information and to store state
- Accessed from eBPF programs as well as from applications in user space

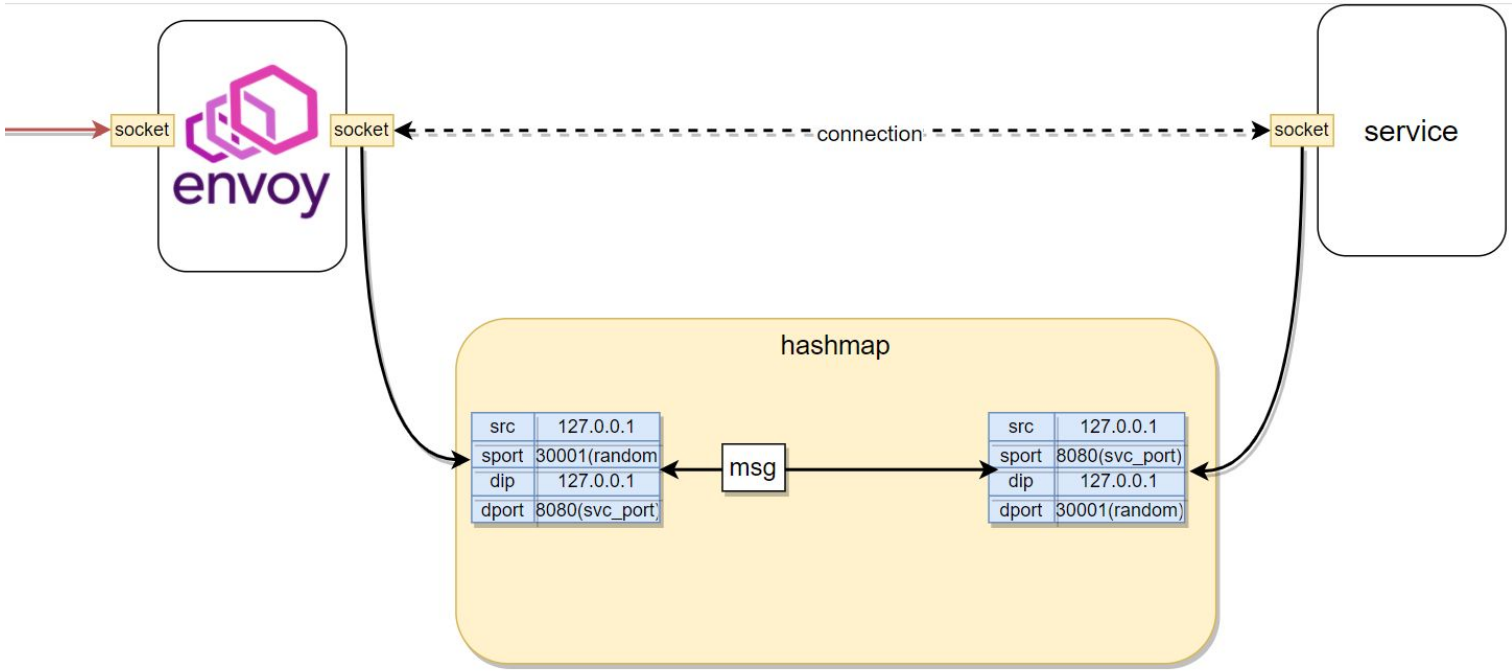


Work Flow of Acceleration

- Attach SOCK_OPS program to global cgroup
- Capture socket in established state and add to hashmap
- Attach sk_skb program to hashmap
- When socket send a msg, lookup peer socket in sockmap
- Redirect

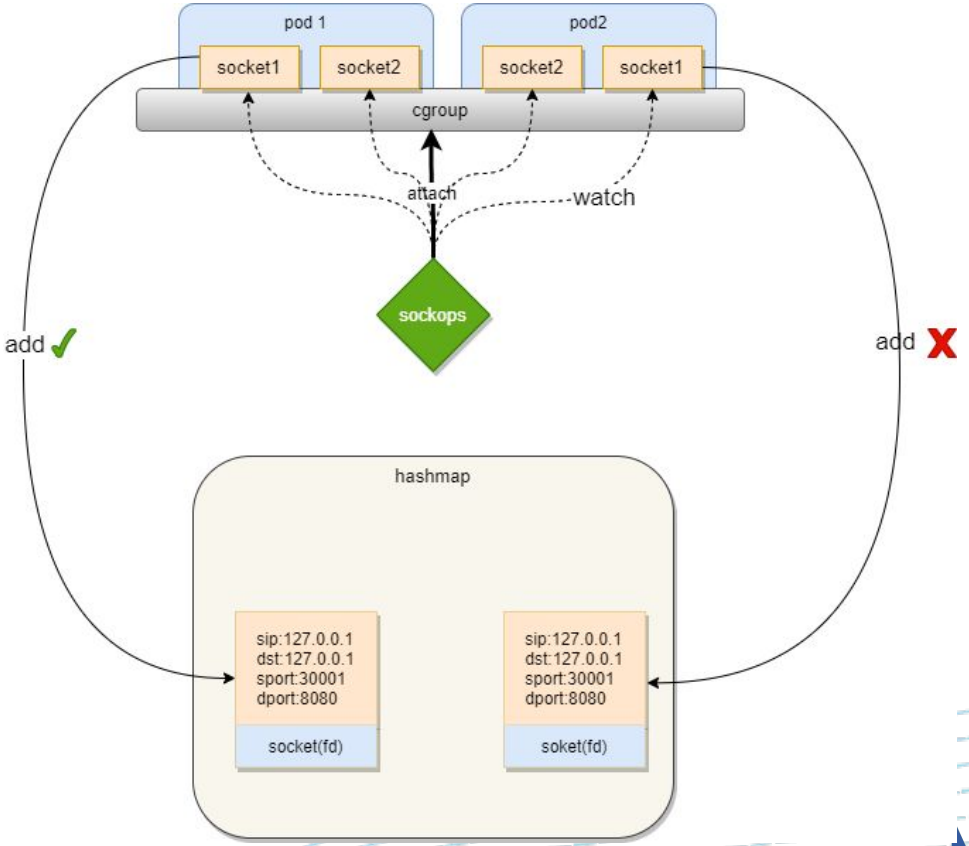


Inbound Acceleration



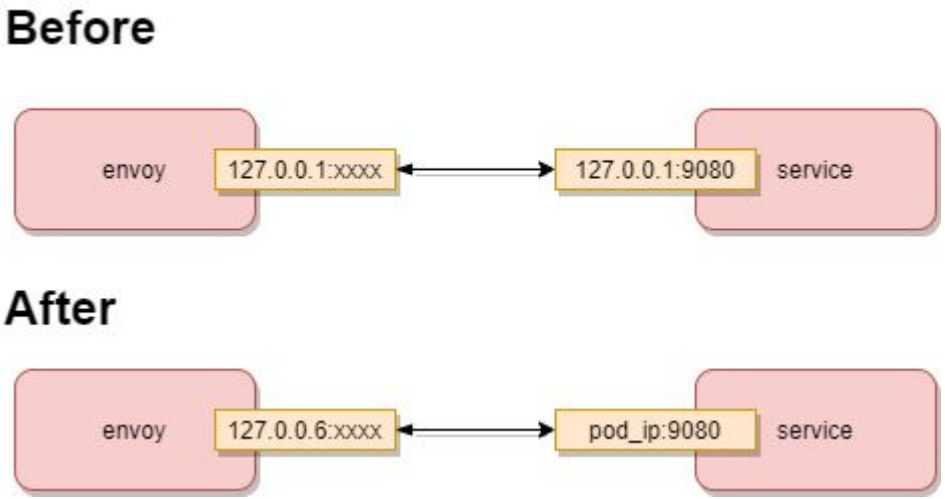
Problem

In the case of Inbound, 4-tuple key may conflict due to same src/dst ip address

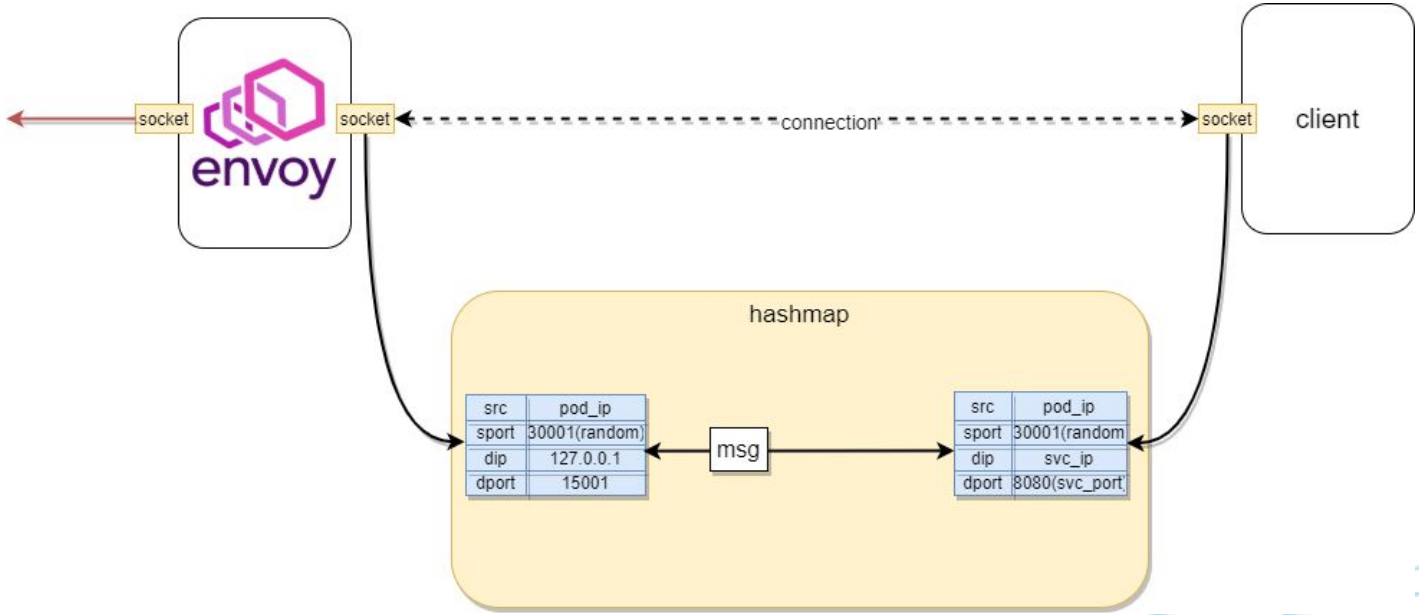


Use pod ip as hash key

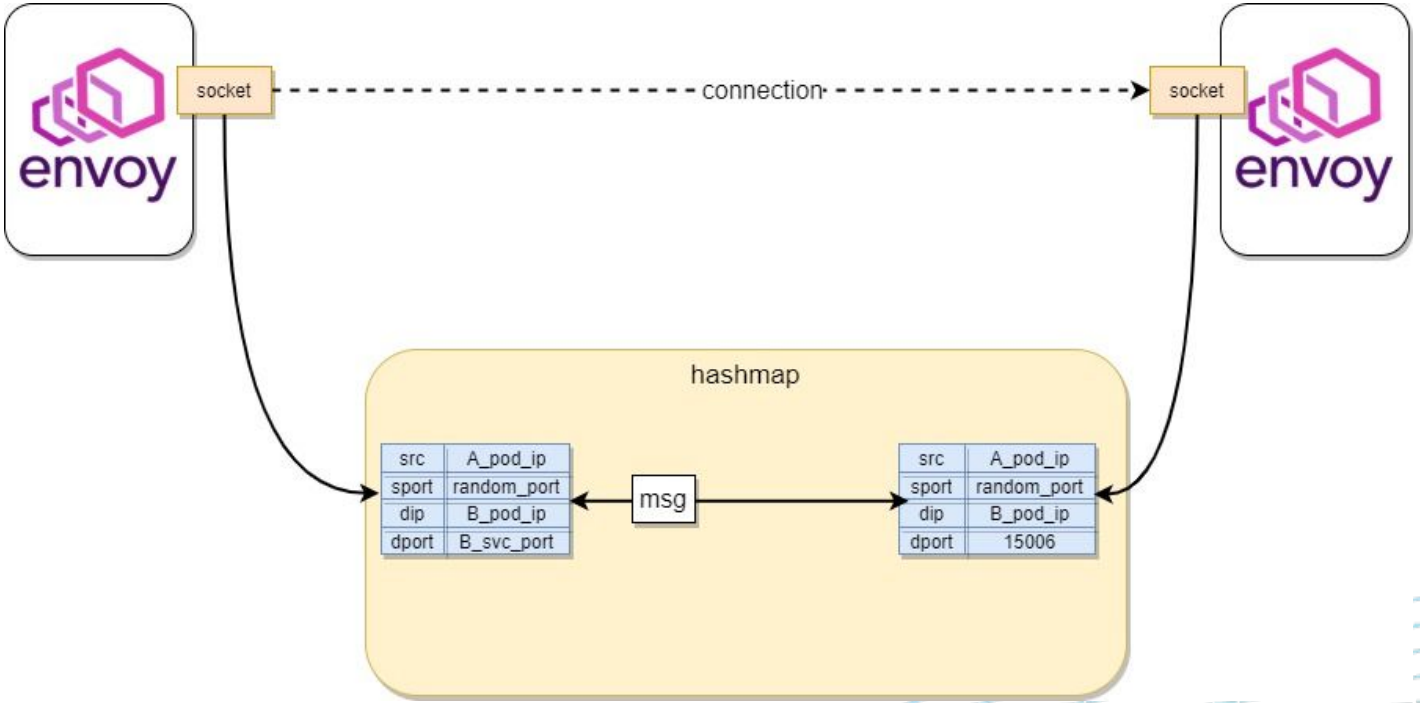
Use pod_ip to generate a unique key is a way to distinguish socket from different network namespace



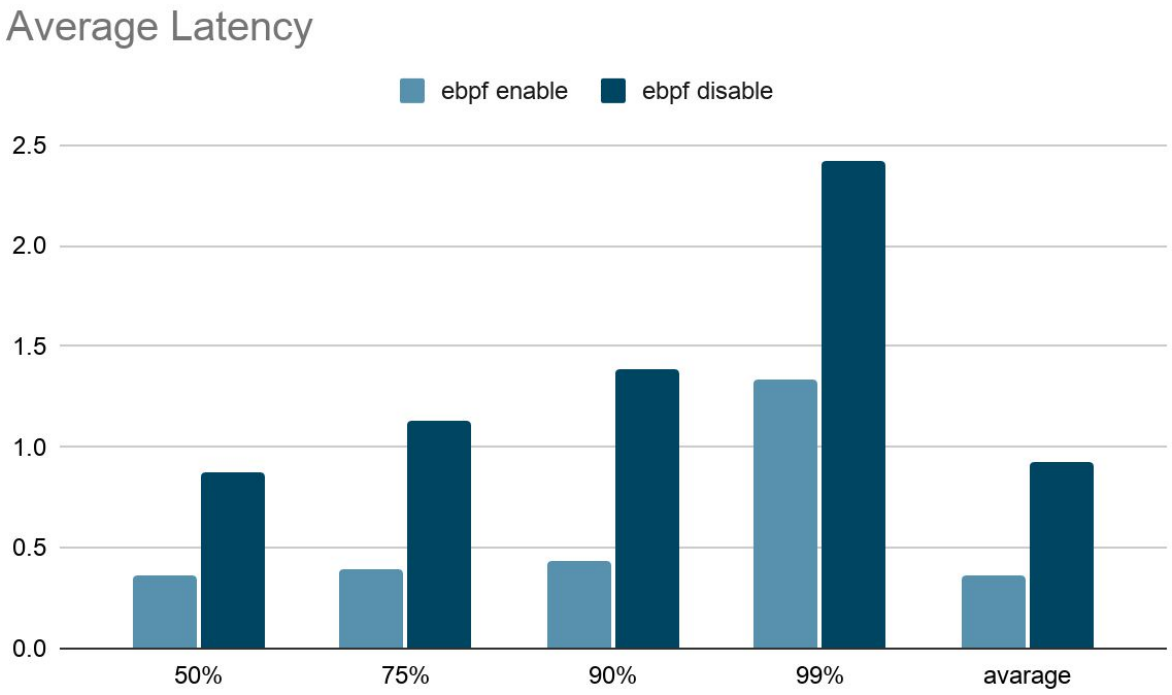
Outbound Acceleration



Envoy to Envoy Acceleration(same host)



Performance Comparison



Thank you!

#IstioCon