Accelerate Service Mesh Network with ebpf
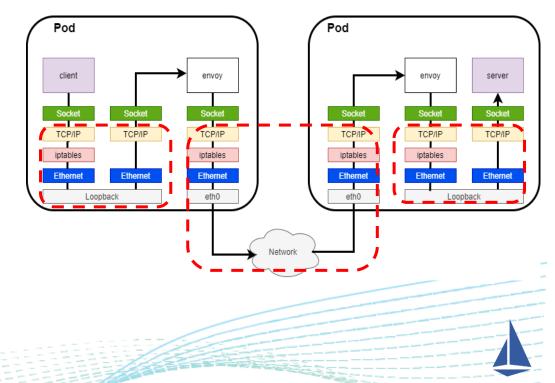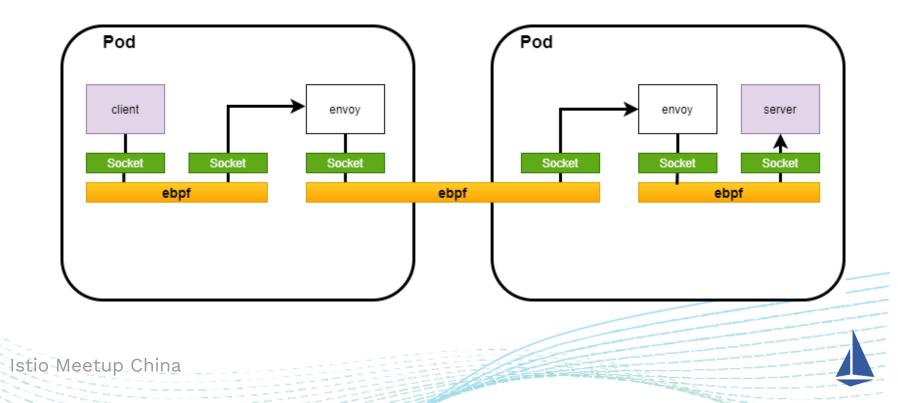
Luyao Zhong

Agenda

- TCP/IP stack overhead in service mesh
- Background knowledge of eBPF
- Independent solution to bypass TCP/IP stack
- Performance Comparision

TCP/IP stack overhead

- All the application data goes via sidecar (envoy)
- All the data passes TCP/IP stack 3 times
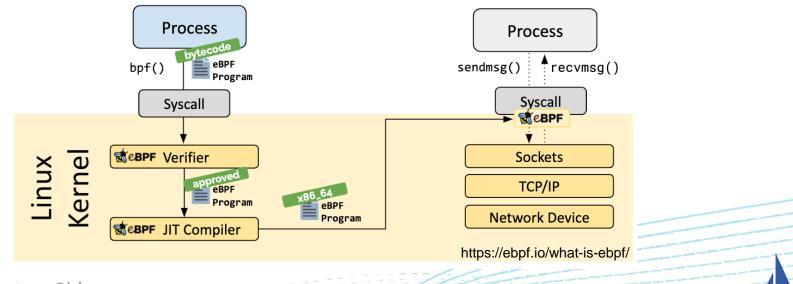 - Inbound
  - Outbound
  - Envoy to Envoy(same host)

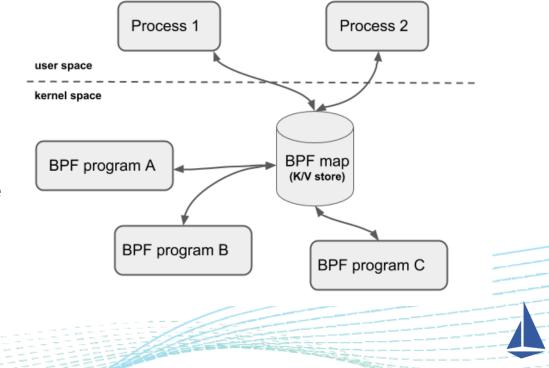# Dataflow After Acceleration(same host)

ebpf Background Knowledge

Loader & Verification Architecture

https://ebpf.io/what-is-ebpf/

ebpf Background Knowledge

map

- Share collected information

- Accessed from eBPF programs as well as from applications in user space

- Map type
  - HASHMAP
  - SOCKHASH: Hold socket as value

ebpf Background Knowledge

Prog type

- SOCK_OPS

  ➢ Set callbacks for TCP state changing

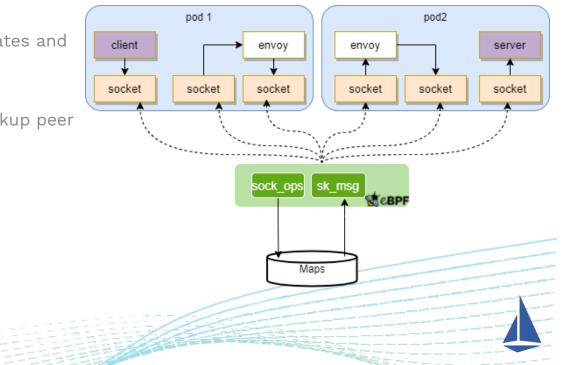 ➢ Help functions: BPF_MAP_UPDATE_ELEM, BPF_SOCK_HASH_UPDATE

- SK_MSG

  ➢ Attach to a SOCKHASH map, capture the packets sent by a socket in SOCKHASH map and determine its destination socket

  ➢ Help functions: BPF_MSG_REDIRECT_HASH
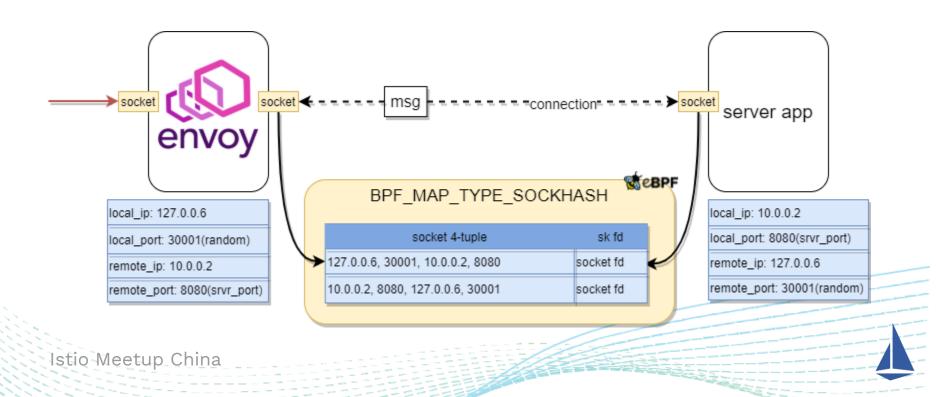
Work Flow of Acceleration

- sock_ops
  - Capture socket in specific states and populate the maps
- sk_msg
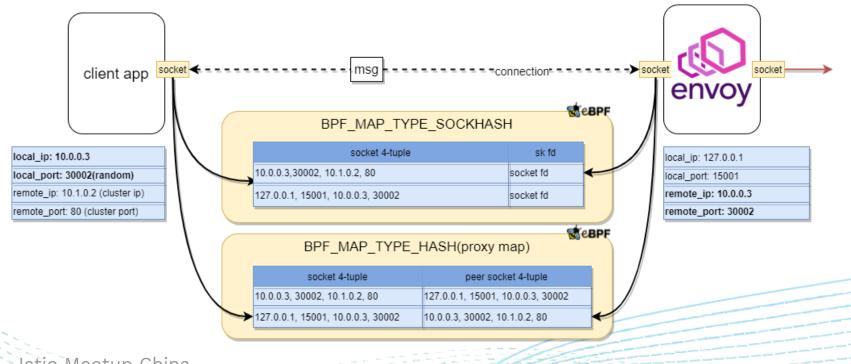 - When socket send a msg, lookup peer socket
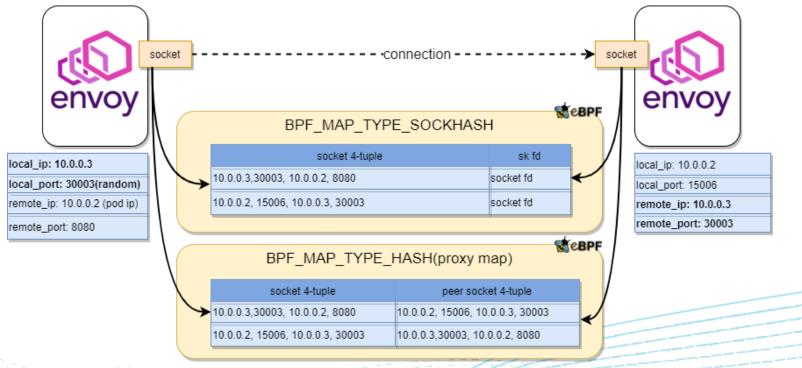 - Redirect

Inbound Acceleration
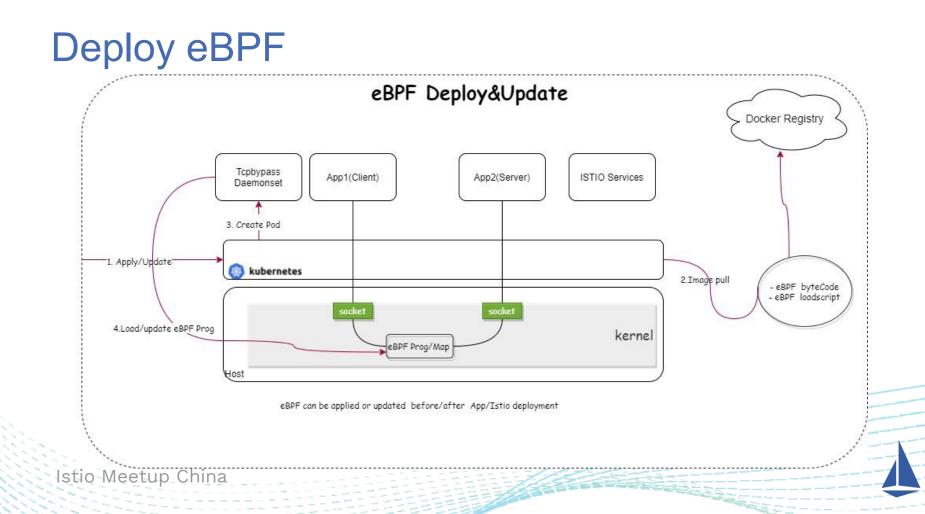## Istio 1.10: PILOT_ENABLE_INBOUND_PASSTHROUGH

Outbound Acceleration

Envoy to Envoy Acceleration(same host)

Deploy eBPF

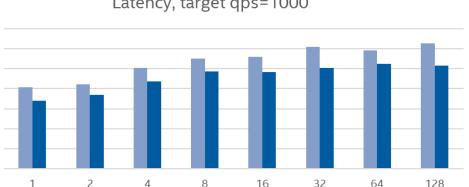# Performance Comparison

Refactored istio benchmarking tool
◦ Two pods run on the same node

Configurations
◦ mTLS enabled
◦ Number of Envoy workers: 2
◦ Response payload size: 1KB

Latency
◦ 11-17% improvement

Latency, target qps=1000

disable bypass(baseline) ■ enable bypass

Summary

- eBPF functionality **enabled with a DaemonSet** pod
  - eBPF program tracks connections from client to redirected Envoy (127.0.0.1) and back (**outbound**)
  - eBPF program also tracks connections from Envoy (127.0.0.6) to Pod IP address and back (**inbound**)
  - eBPF program also tracks connections from Envoy to Envoy(in the same node) and back (**envoy to envoy**)
- Works with **Istio >= 1.10**
- **CNI agnostic** and should work with all CNIs (wo/ eBPF)
- Latency：11~17% improvement

Thank you!

luyao.Zhong@intel.com