



Deep Dive TCP/IP Bypass with eBPF in Service Mesh

Luyao Zhong @ Intel
2022 April

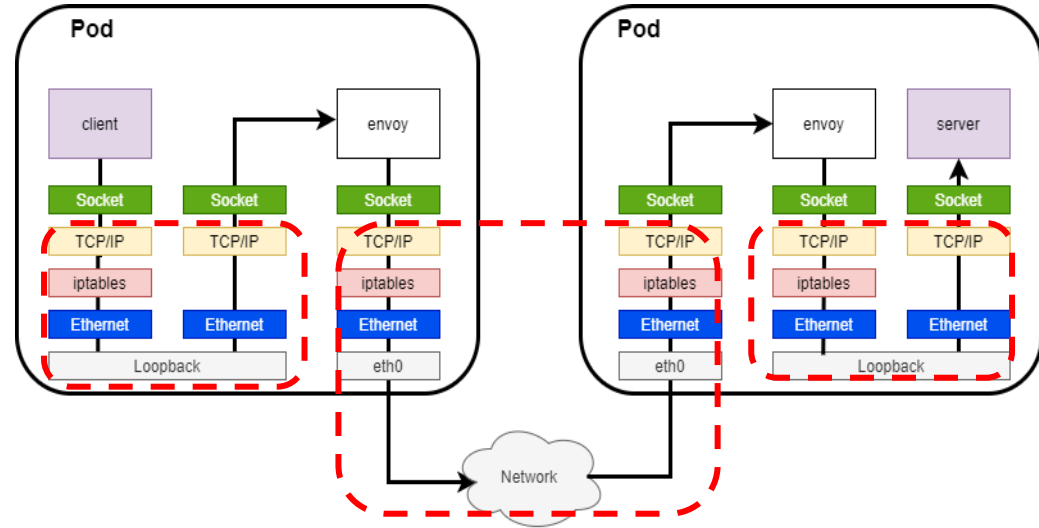
Agenda

- TCP/IP stack overhead
- Background knowledge of eBPF
- How to bypass TCP/IP stack
- Performance Comparison
- Project links

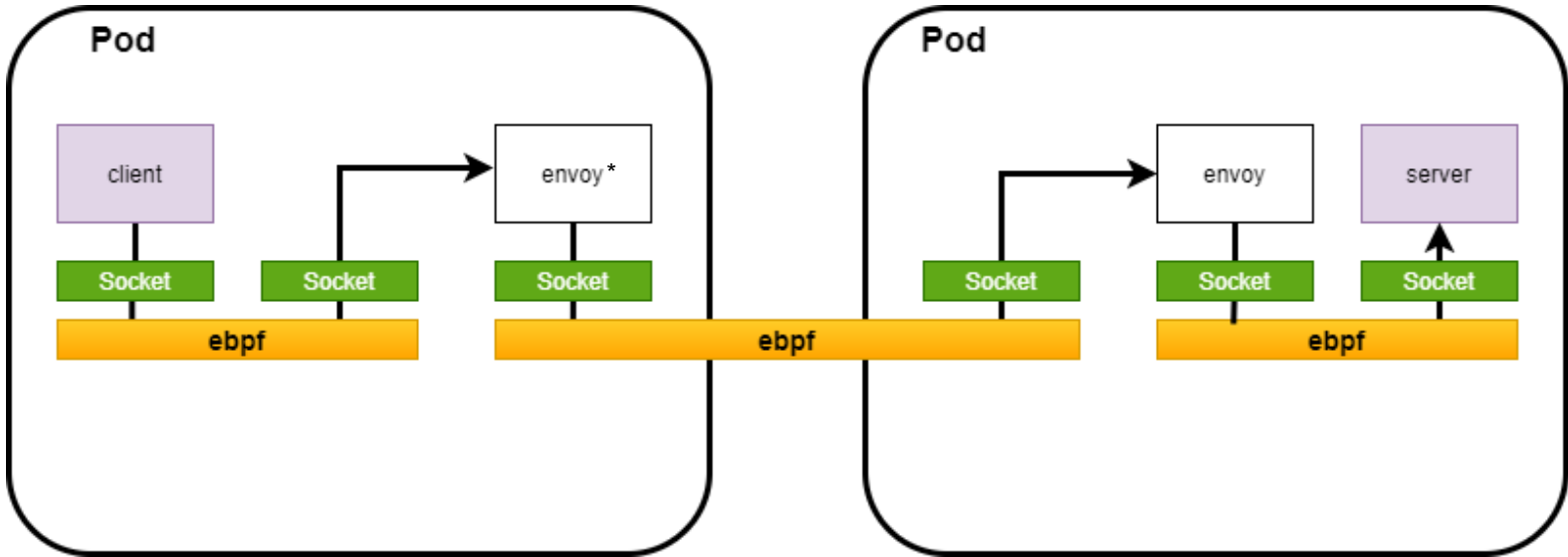


TCP/IP stack overhead

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

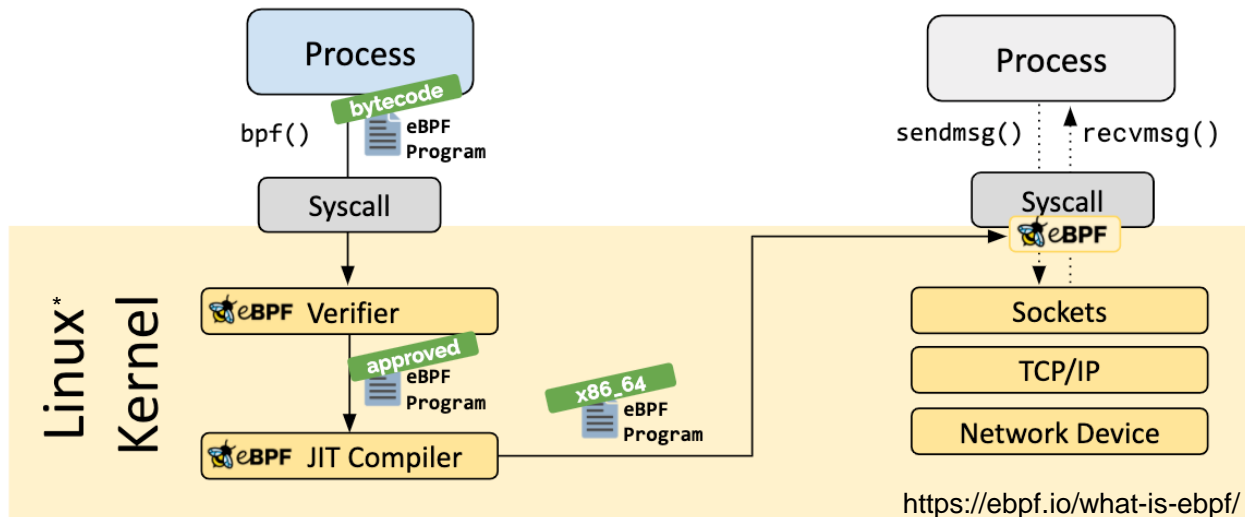


Dataflow after acceleration (same host)



ebpf background knowledge

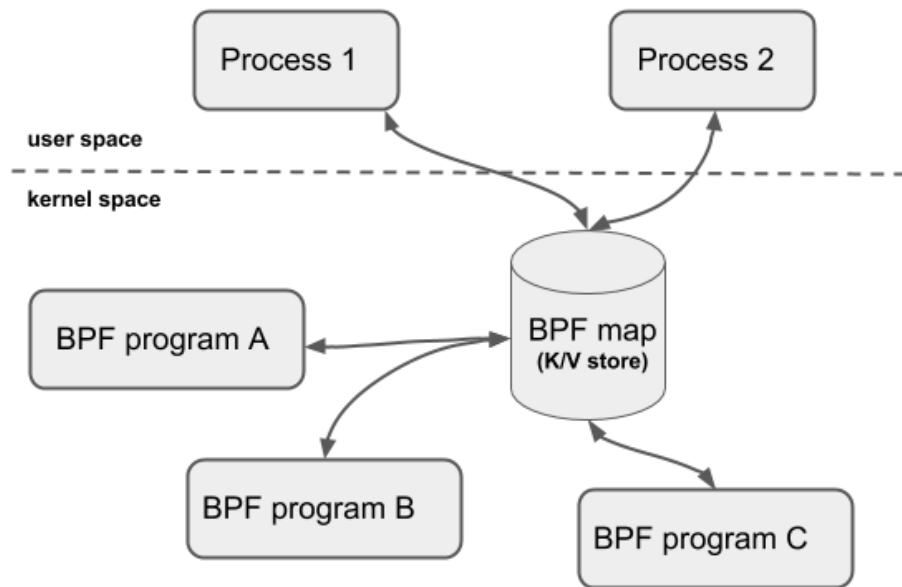
Loader and verification architecture



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

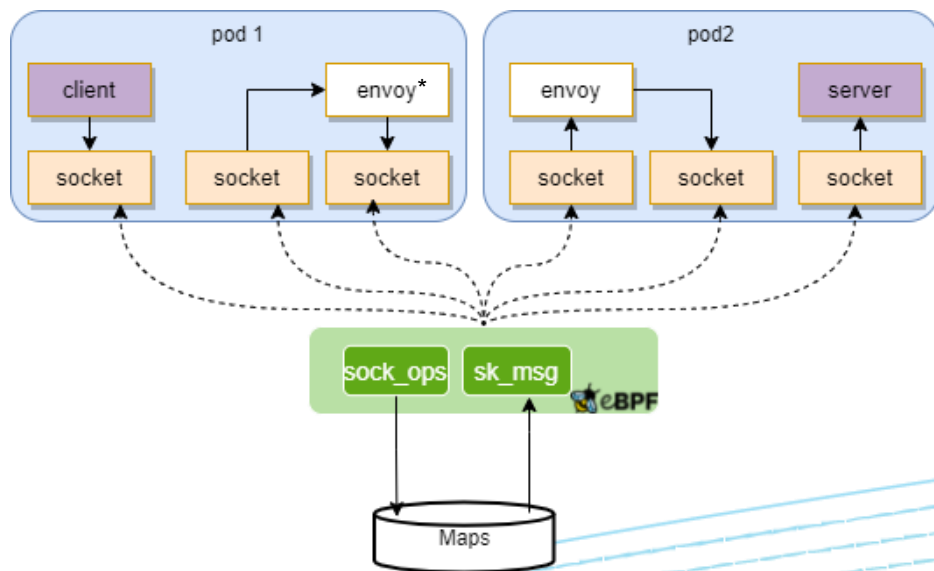
- SK_MSG

- 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



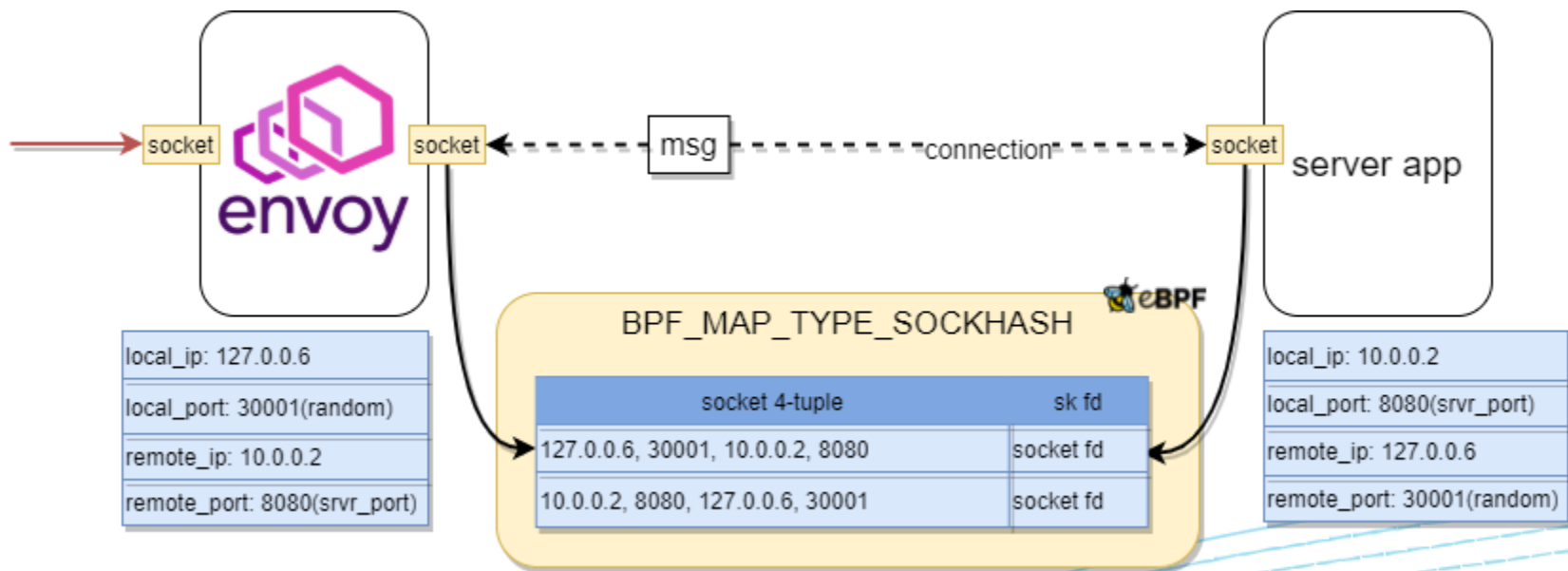
Work flow of acceleration

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

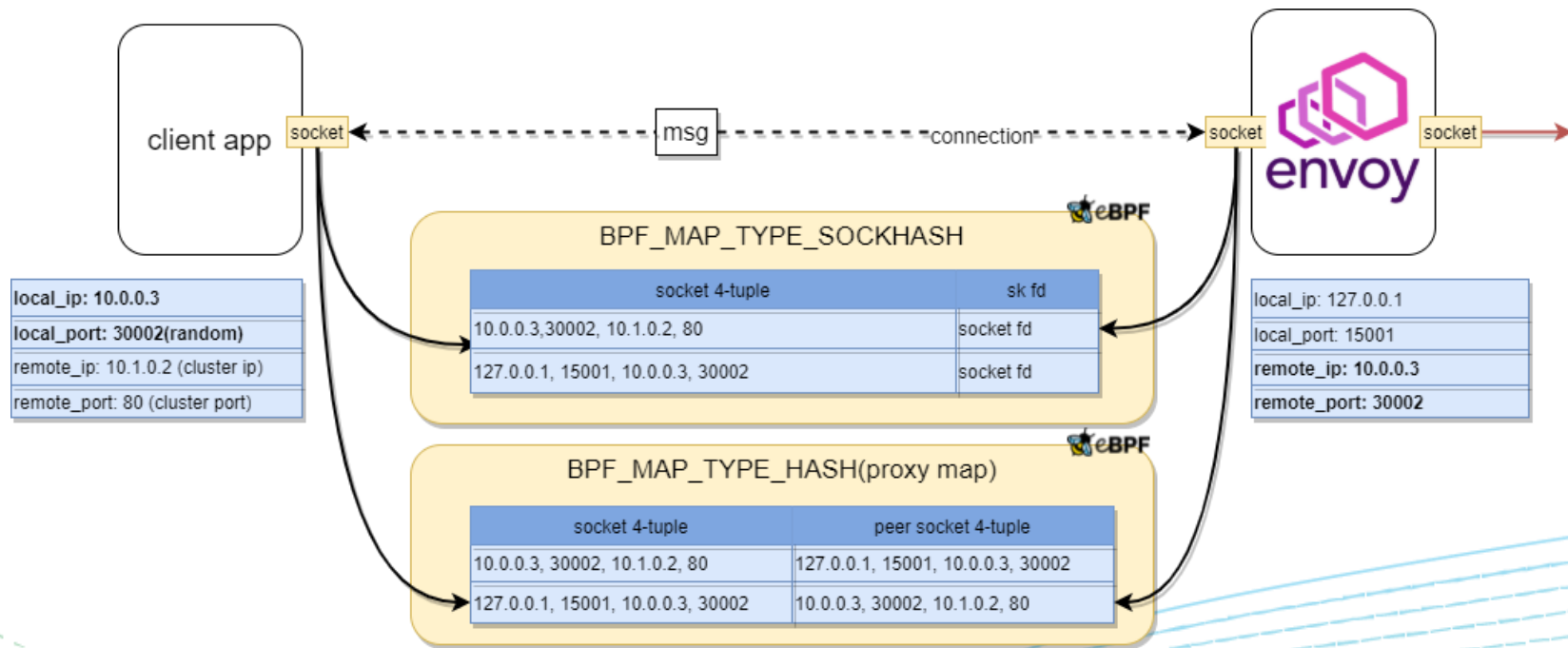


Inbound acceleration

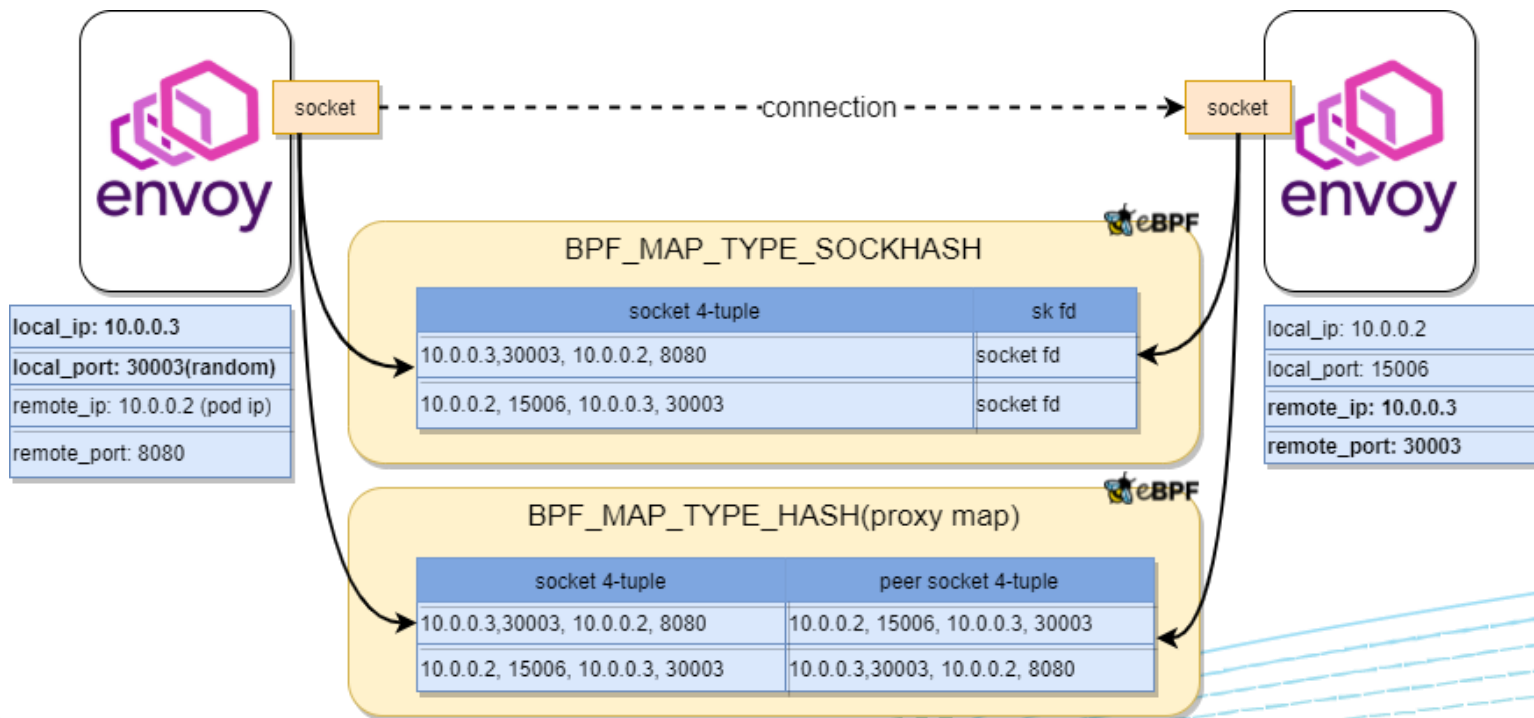
Istio 1.10: PILOT_ENABLE_INBOUND_PASSTHROUGH



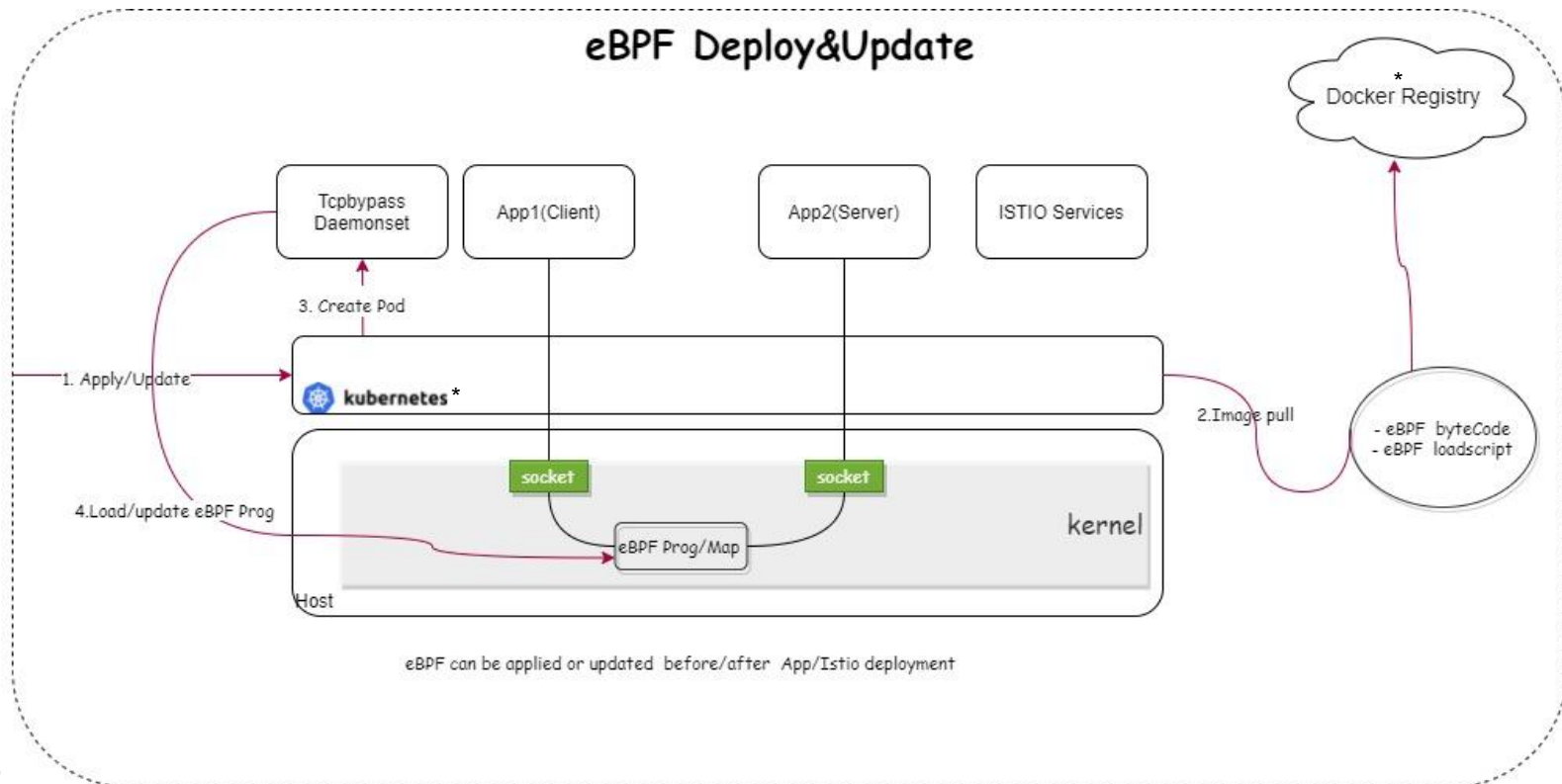
Outbound acceleration



Envoy* to Envoy Acceleration (same host)

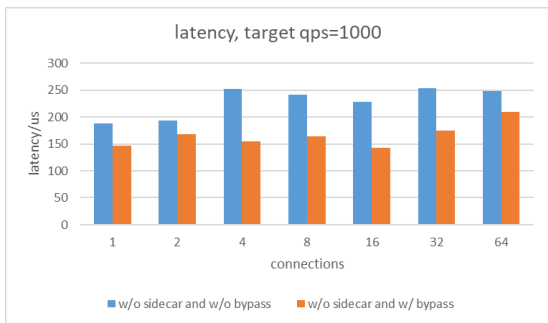


Deploy eBPF

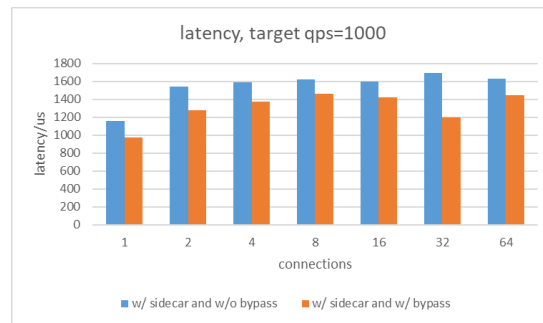


Performance Comparison

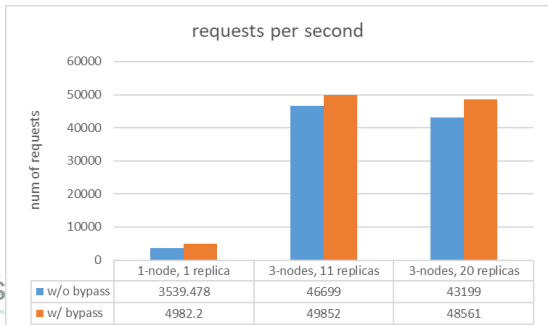
Nighthawk w/o sidecar injected



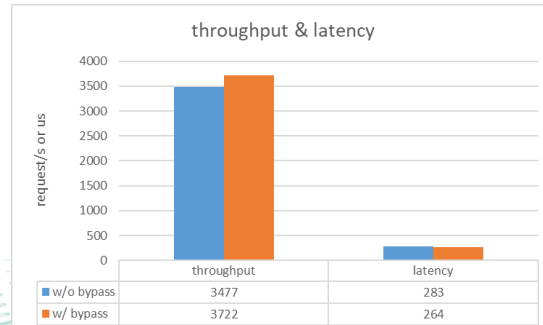
Nighthawk w/ sidecar injected



DSB-hotel w/o sidecar injected



GMS w/ sidecar injected



Project is public NOW!

- Docker* image
 - <https://hub.docker.com/r/intel/istio-tcpip-bypass/tags>
- Project GitHub* repo
 - <https://github.com/intel/istio-tcpip-bypass>
- Tips for Debug
 - ss command: Inspect the socket status
 - bpf_trace_printk() helper function: kernel debug tracepipe
 - /sys/kernel/debug/tracing/trace_pipe



Thank you!

luyao.Zhong@intel.com

#IstioCon



Notices and Disclaimers

Intel technologies may require enabled hardware, software or service activation.

No product or component can be absolutely secure.

Your costs and results may vary.

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

