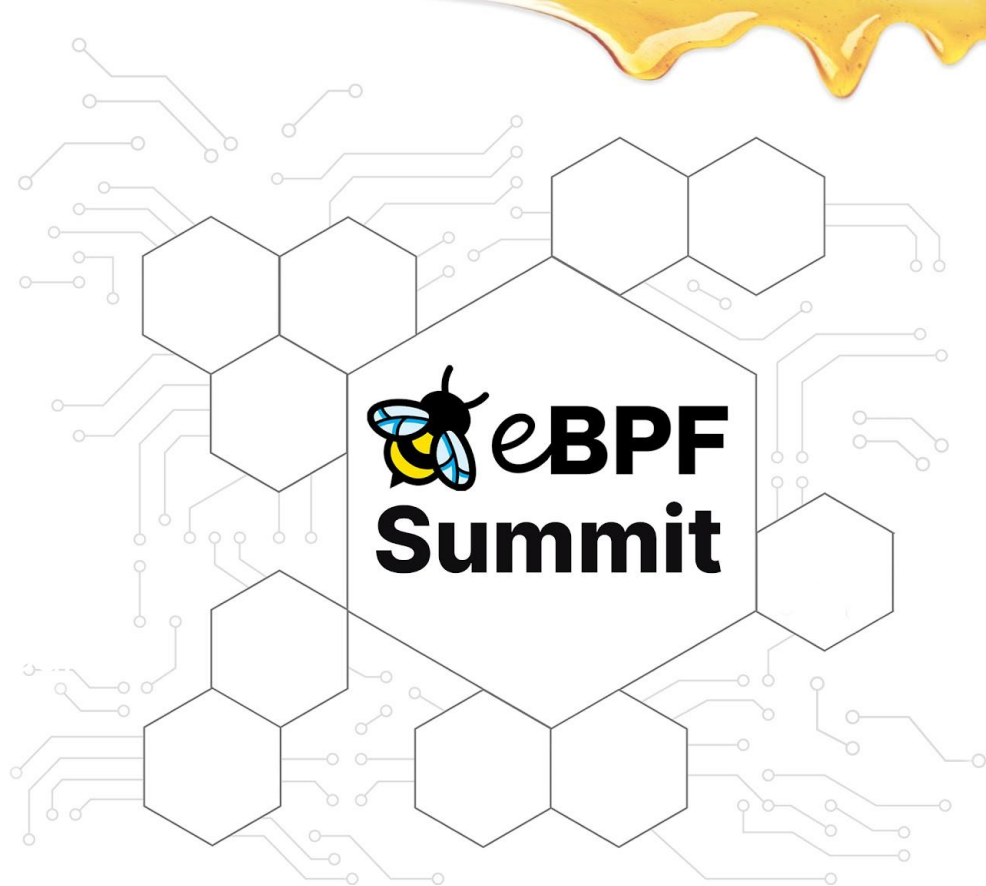


# Improving Cilium's eBPF-based DSR performance by adding support for IP-in-IP



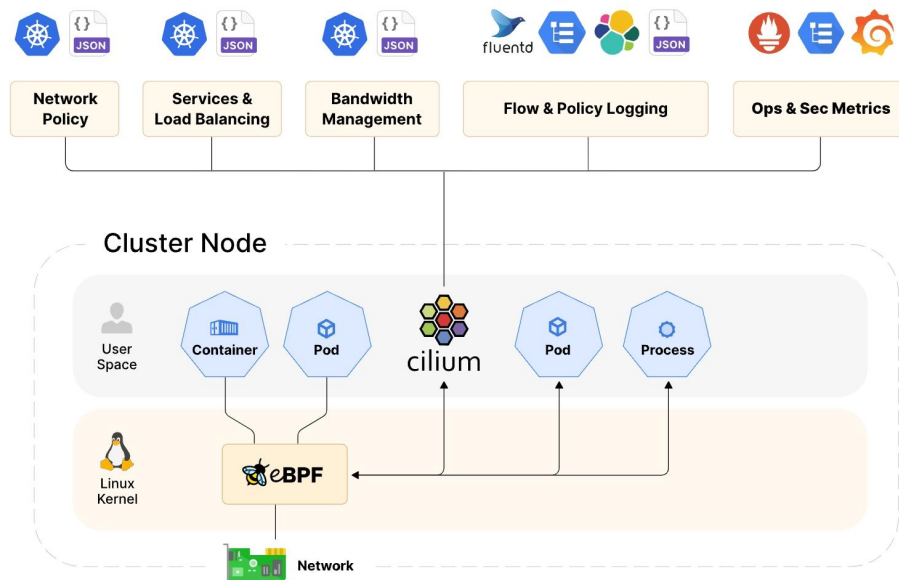
**Jiang Wang**

@System Technology and Engineering, ByteDance

ByteDance 字节跳动

# Cilium

- Cilium
  - open source software
  - provide, secure and observe network connectivity between container workloads
- ByteDance deployment
  - Deployed on a few Edge network sites
  - Used as k8s CNI to provide networking
  - Replace kube-routers
  - Less CPU utilization.



# SNAT mode

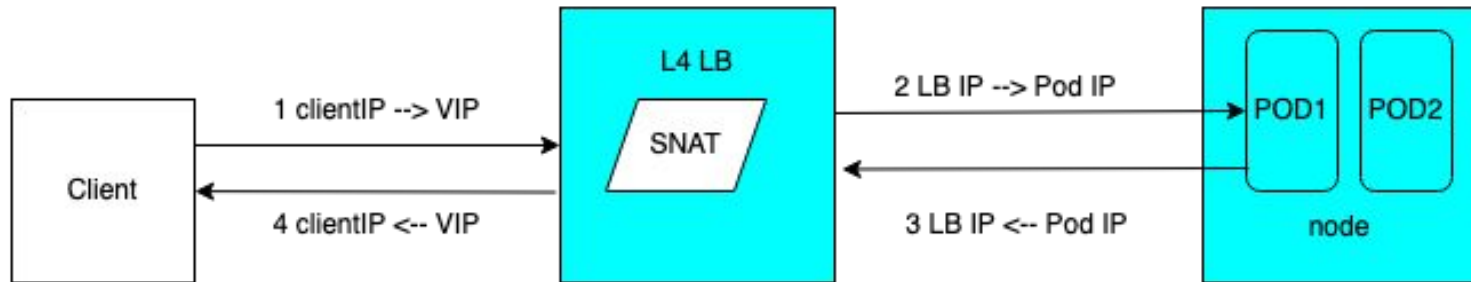


Figure1: SNAT packet flow

# Direct Server Return (DSR) mode

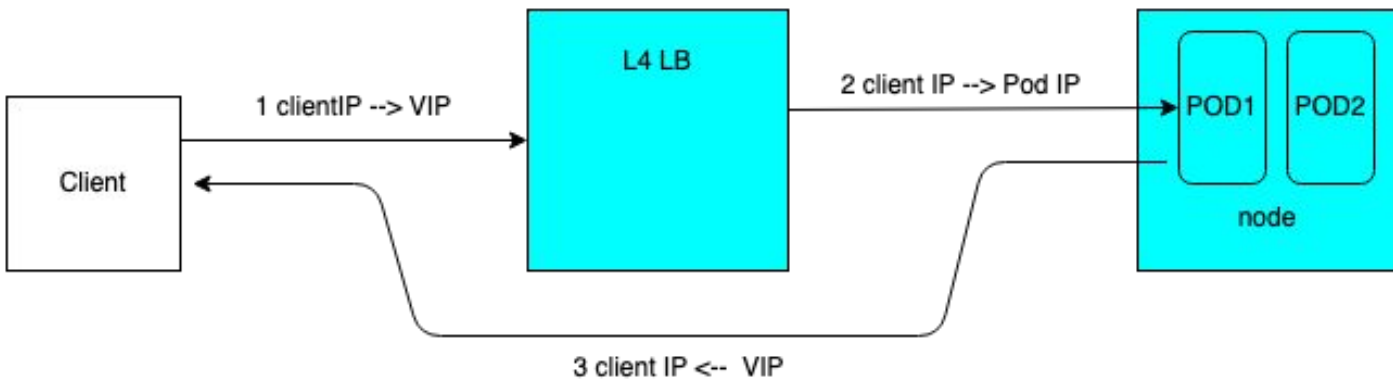


Figure1: DSR packet flow

- Benefit: less processing on LB.
- Where to store VIP?

# Current Cilium DSR support for IPv4

- Use IPv4 options

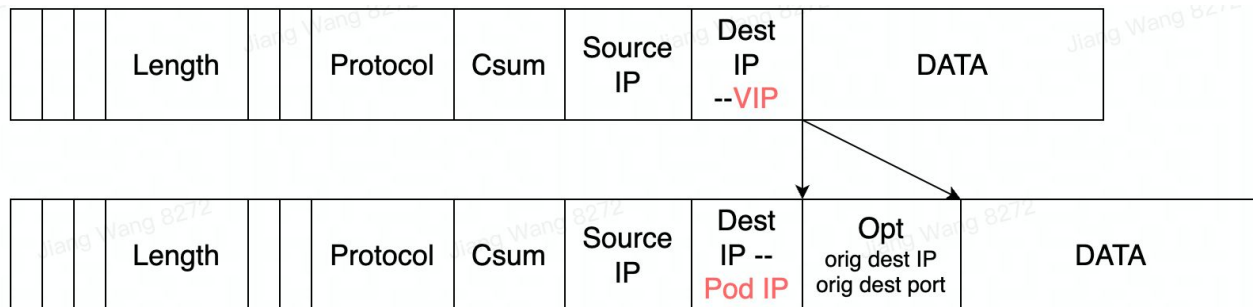


Figure 2: IP packet format in DSR with IP option



# Problem with Current DSR support for IPv4

- Problem: will go to **slow** path on switches
  - Network switch has a fast path and slow path
  - IP packets with option will go via slow path on many switches
  - Switches CPU usage arrived 100% for some cores. **Bottleneck!**
- Could we do better?

# Proposal: Use IP-in-IP for DSR on IPv4

- Use IP-in-IP

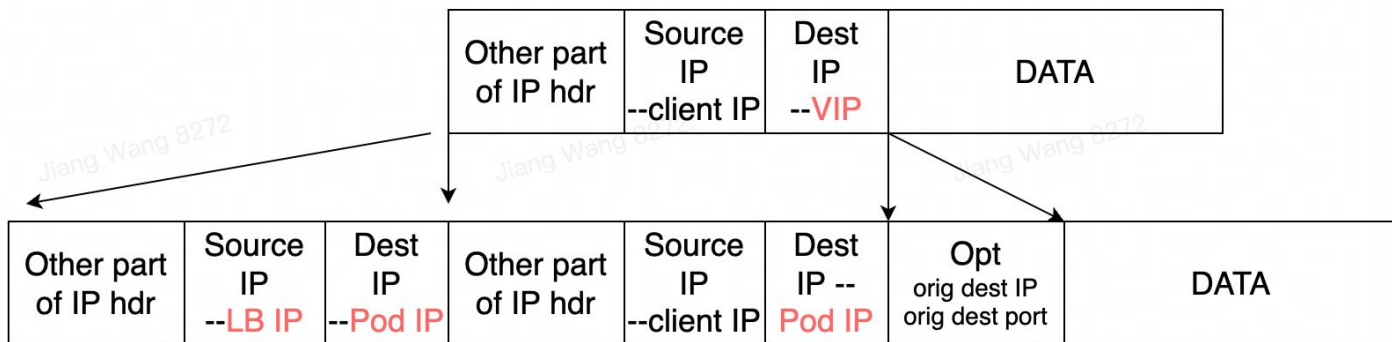


Figure 3: IP packet format in DSR with IP-in-IP

- Will go to **fast** path on switches
- Drawbacks: smaller MTU for UDP and TCP Syn pkt



# Thank you!