SACC

2022 中国系统架构师大会

SYSTEM ARCHITECT CONFERENCE CHINA 2022

激发架构性能 点亮业务活力

云上会议 网络直播 | 2022年10月20-22日

ITT168.com ChinaUnix.net ITPUB

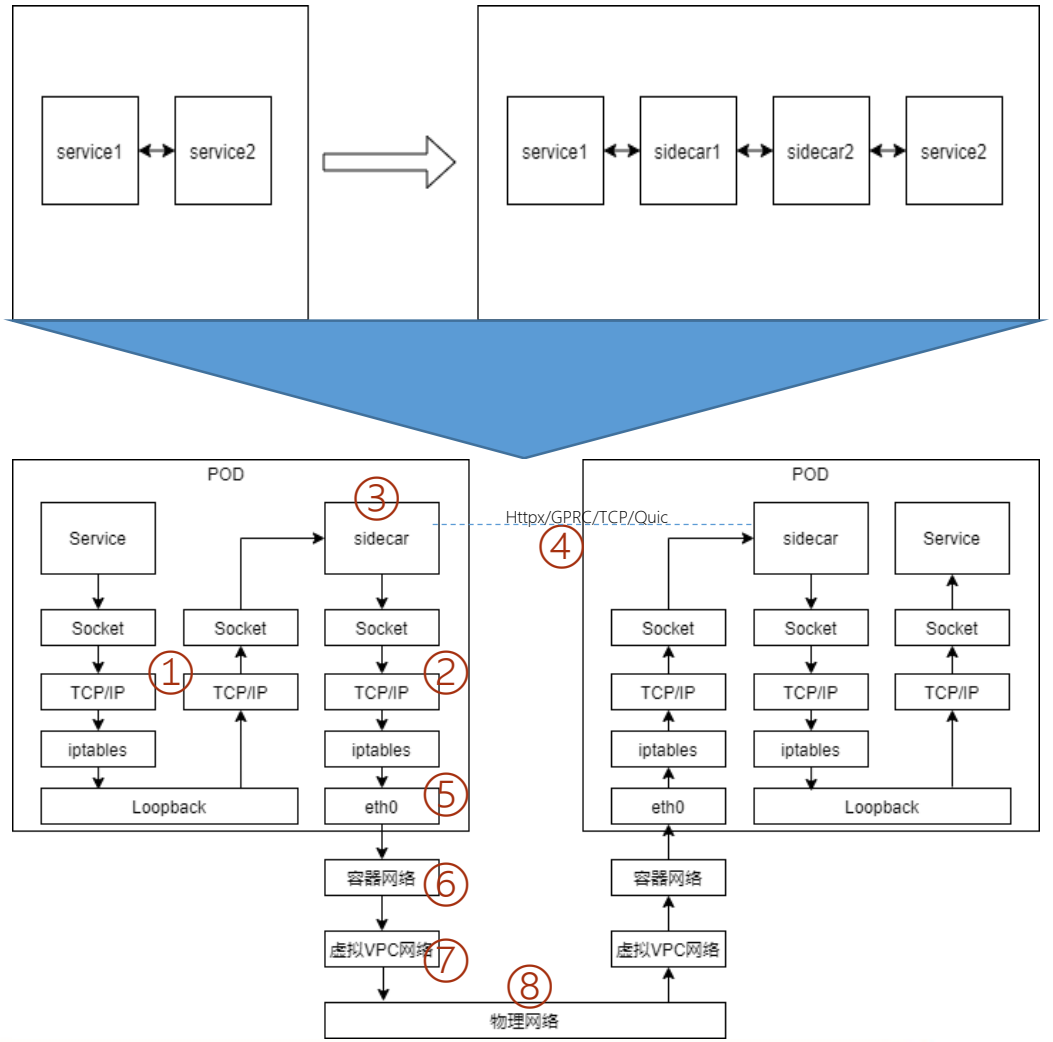
延迟降低50%，深入服务网格数据面性能调优

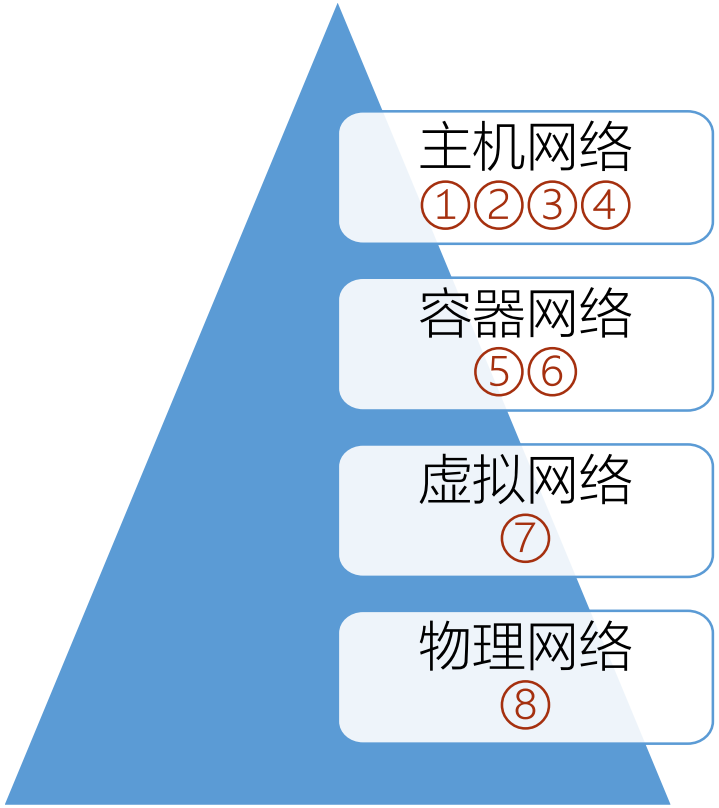
网易高级技术专家 汪翰林

目录

- [x] 服务网格网络数据面介绍
- [ ] 服务网格网络数据面优化实践
- [ ] 服务网格网络数据面后续演进

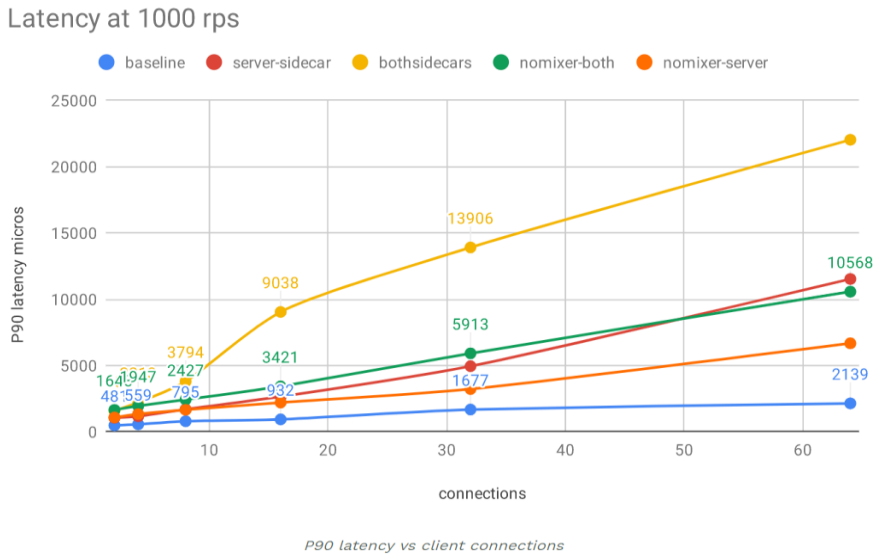
服务网格网络数据面

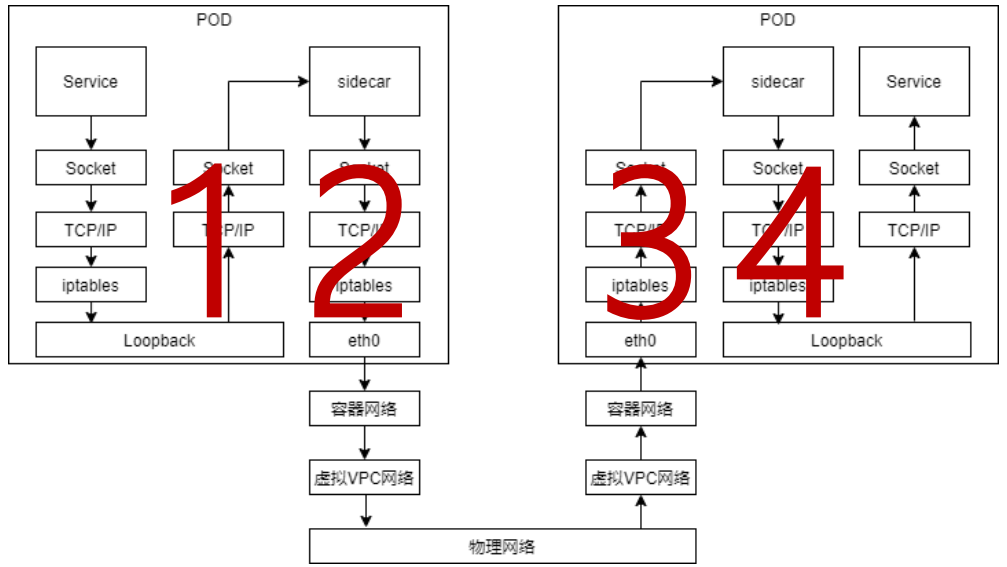




服务网格底层数据面链路要经过物理网络、虚拟网络、容器网络、主机网络

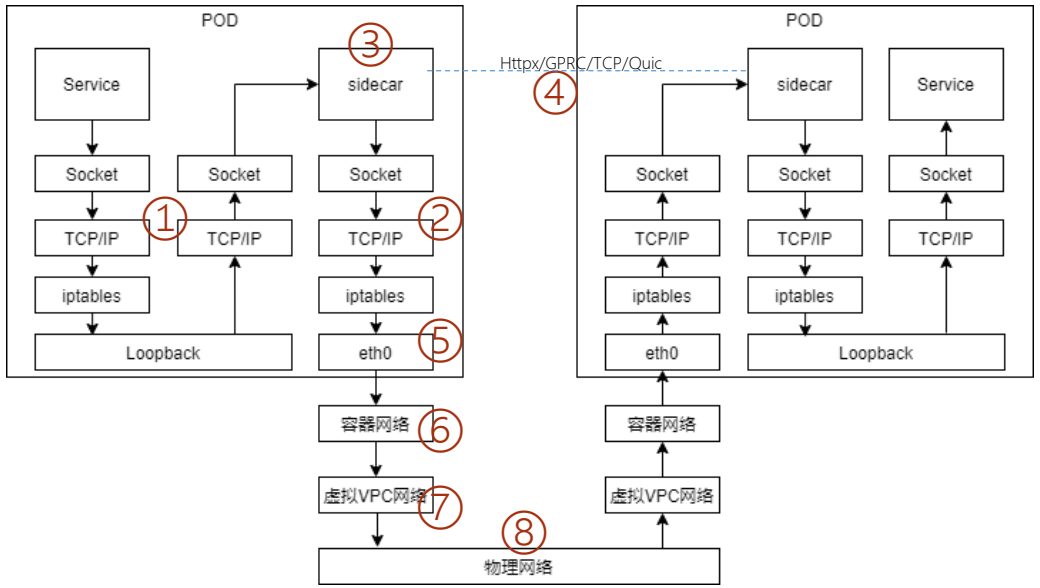
服务网格性能劣化





服务网格额外引入的四次内核协议栈调用，导致链路变长，时延增加！

服务网格性能优化方向

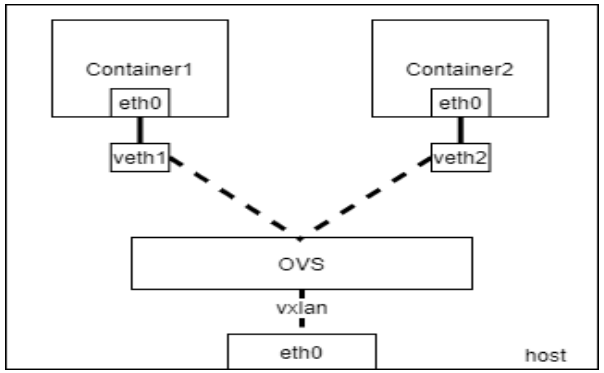


物理网络、虚拟网络、容器网络的时延，并不是服务网格化引入，但是性能优化也可以让服务网格收益。

网络	编号	优化方向
物理网络	⑧	不作讨论
虚拟网络	⑦	虚拟化会带来10%开销，可以支持使用裸机容器
容器网络	⑤⑥	改造，如用户态OVS替代内核态OVS
主机网络	①	Bypass内核协议栈，如eBPF加速/用户态协议栈
	②	Bypass内核协议栈，如用户态协议栈
	③④	代理逻辑优化，如mixer 通信协议优化，如GRPC/Quic

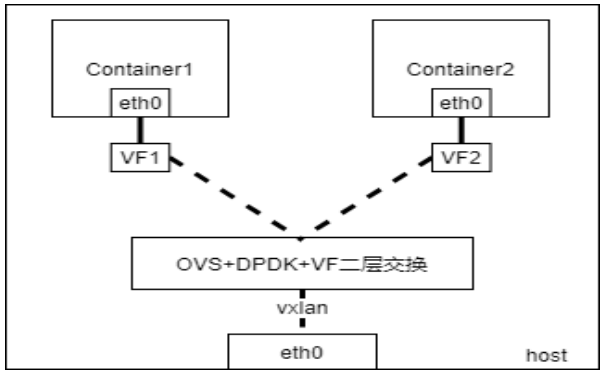
目录

容器网络 – 用户态OVS+SRIOV



OVS+VETH

- ✓ VETH pair接入容器
- ✓ 内核态OVS实现流量分发和VXLAN封装/解封

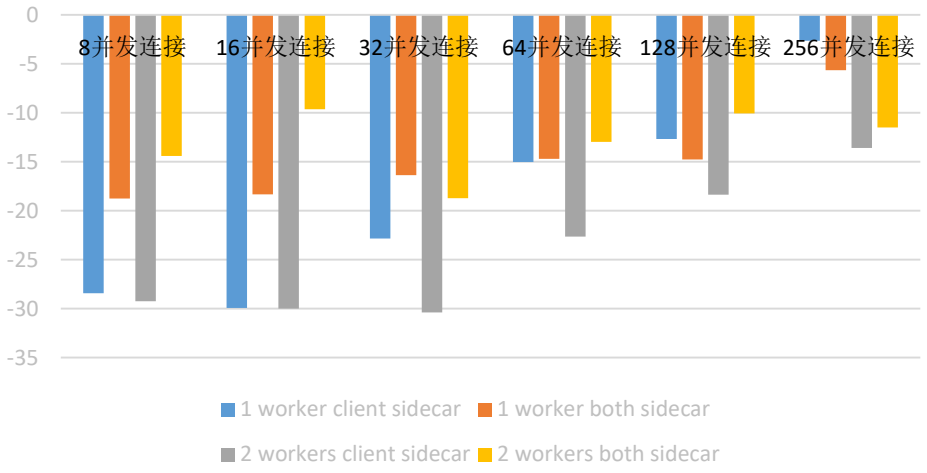


OVS+DPDK+SRIOV

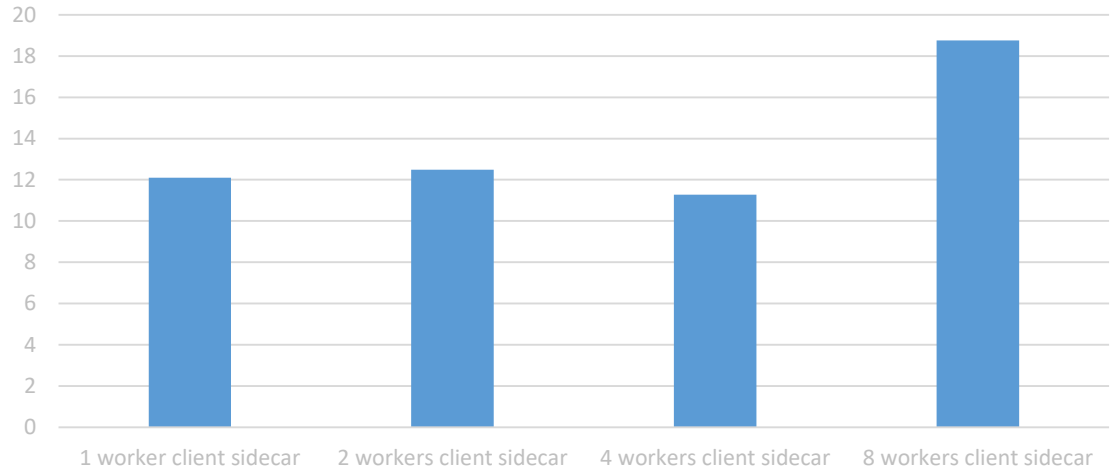
- ✓ VF接入容器和OVS
- ✓ 用户态OVS实现流量分发和VXLAN封装/解封
- ✓ VF多队列
- ✓ DPDK高效分发

容器网络 – 用户态OVS+SRIOV

时延降低比例

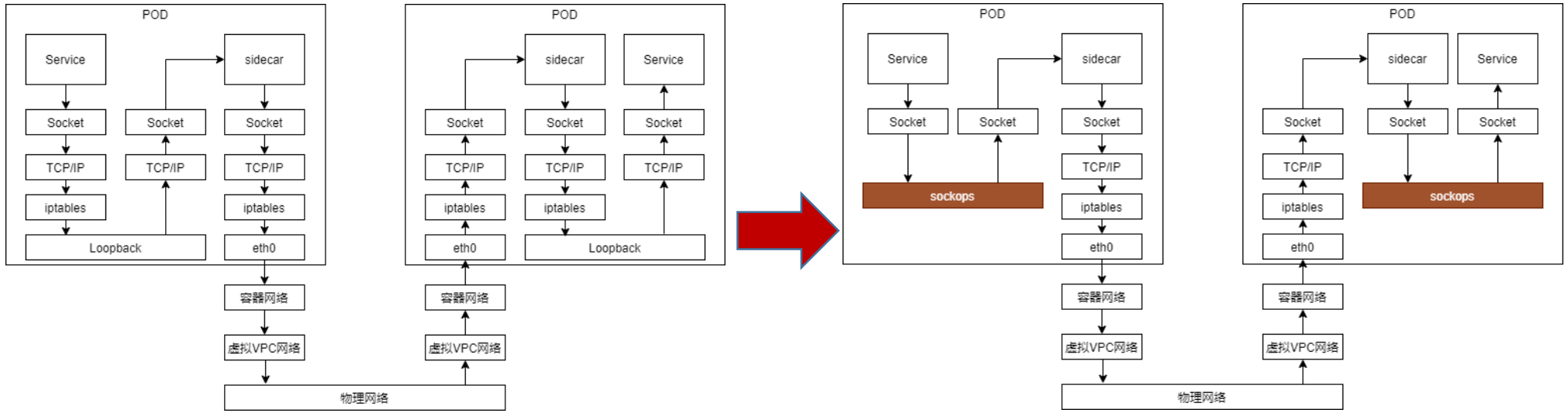


QPS提升比例



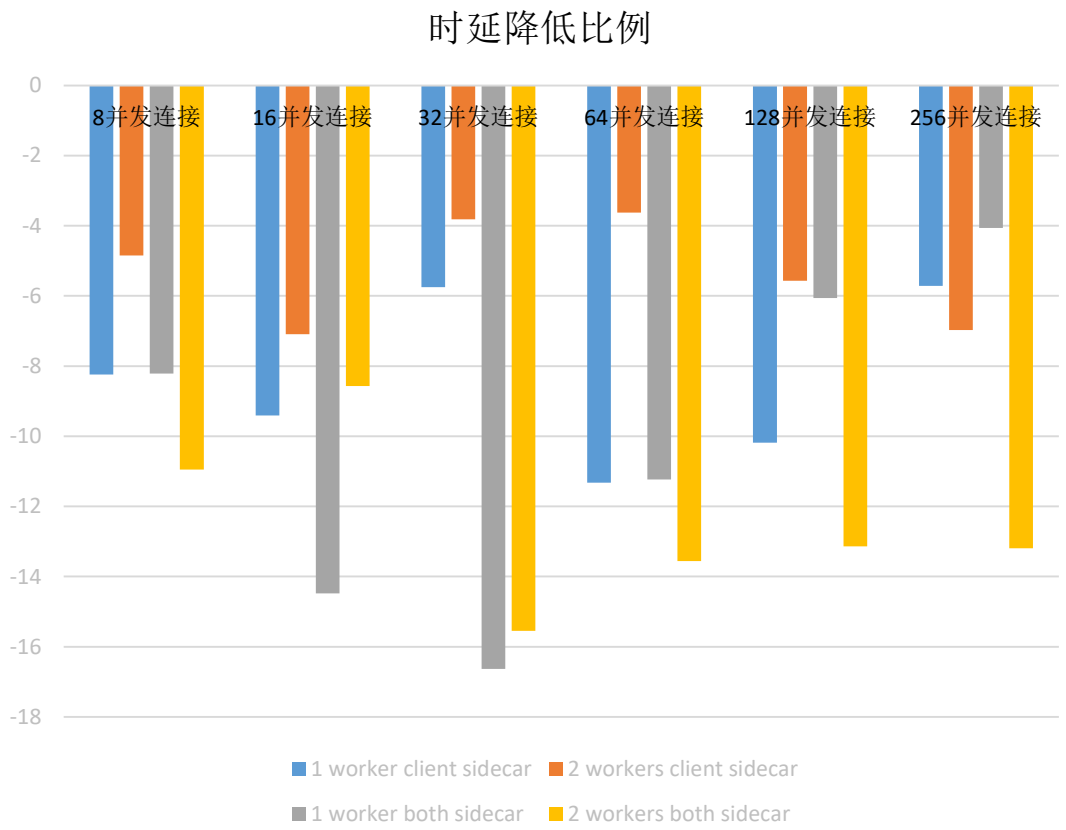
- ✓ 受益于用户态OVS高效的基于DPDK的PMD转发，以及VF多队列
- ✓ 时延降低幅度在10%-30%之间
- ✓ QPS提升在10%-20%之间

主机网络 – eBPF Sockops加速



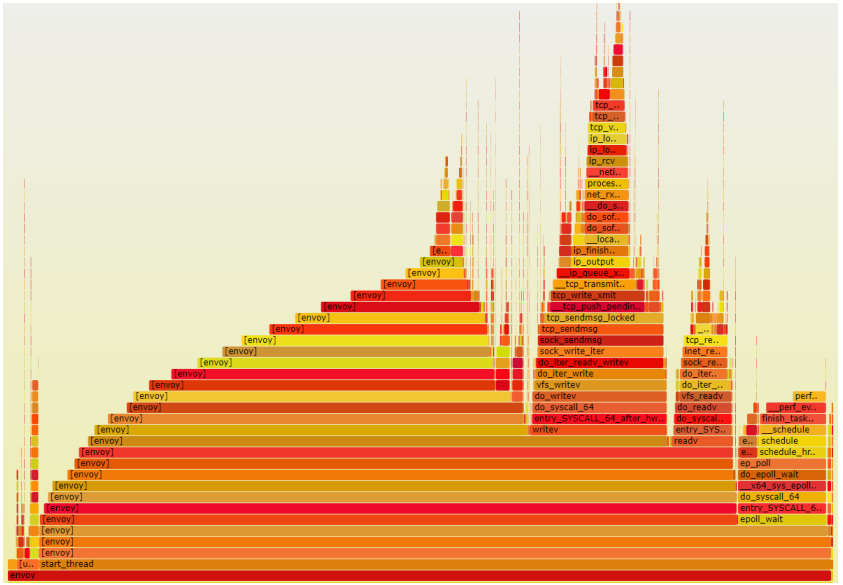
- ✓ 加速Service和Sidecar之间的通信路径
- ✓ 基于Sockmap和sk redirect技术，Bypass内核协议栈
- ✓ 不适用于跨节点加速

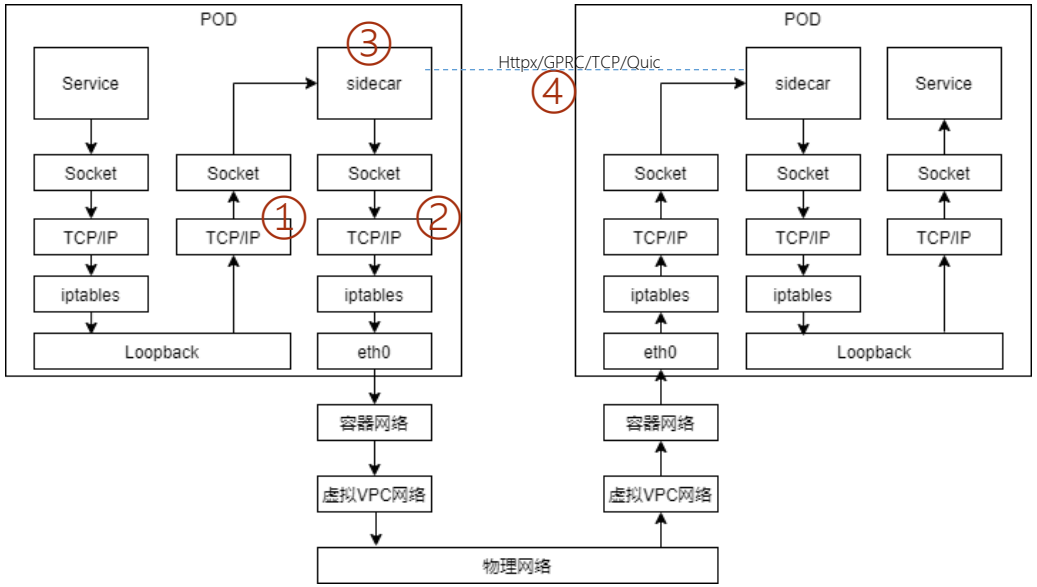
主机网络 – eBPF Sockops加速



Bypass 内核协议栈，
5-15%的时延降低

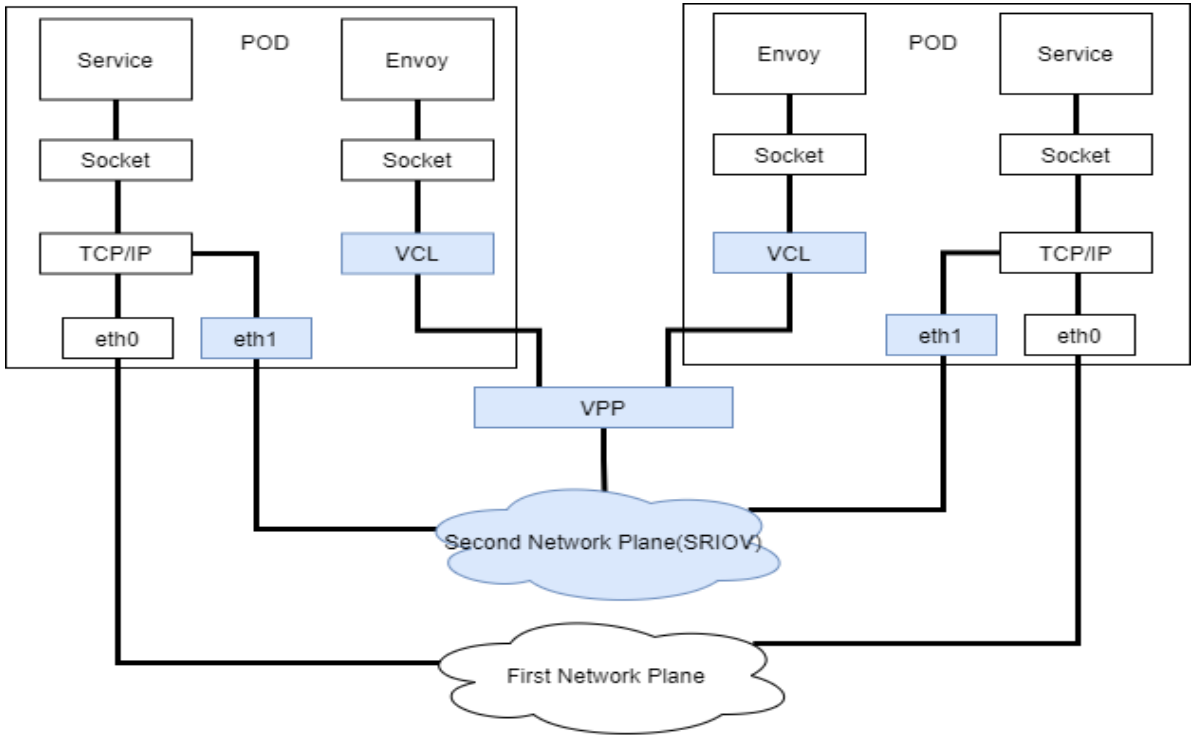
主机网络 – 用户态协议栈





- ✓ 内核协议栈（①+②）占比近50%
- ✓ ①②都可以替换成用户态协议栈
- ✓ ③④社区会持续优化，①②占比提升，则用户态协议栈优化效果越明显

主机网络 – 用户态协议栈性能优先模式



- ✓ 用户态协议栈加速基于SRIOV容器网络方案，此容器网络作为第二平面存在，避免对现有网络影响
- ✓ 用户态协议栈独立部署到VPP进程，解放envoy的CPU
- ✓ VCL库通过LD_PRELOAD劫持socket调用做到针对Envoy的无侵入加速
- ✓ Service容器基于原有内核协议栈无改动

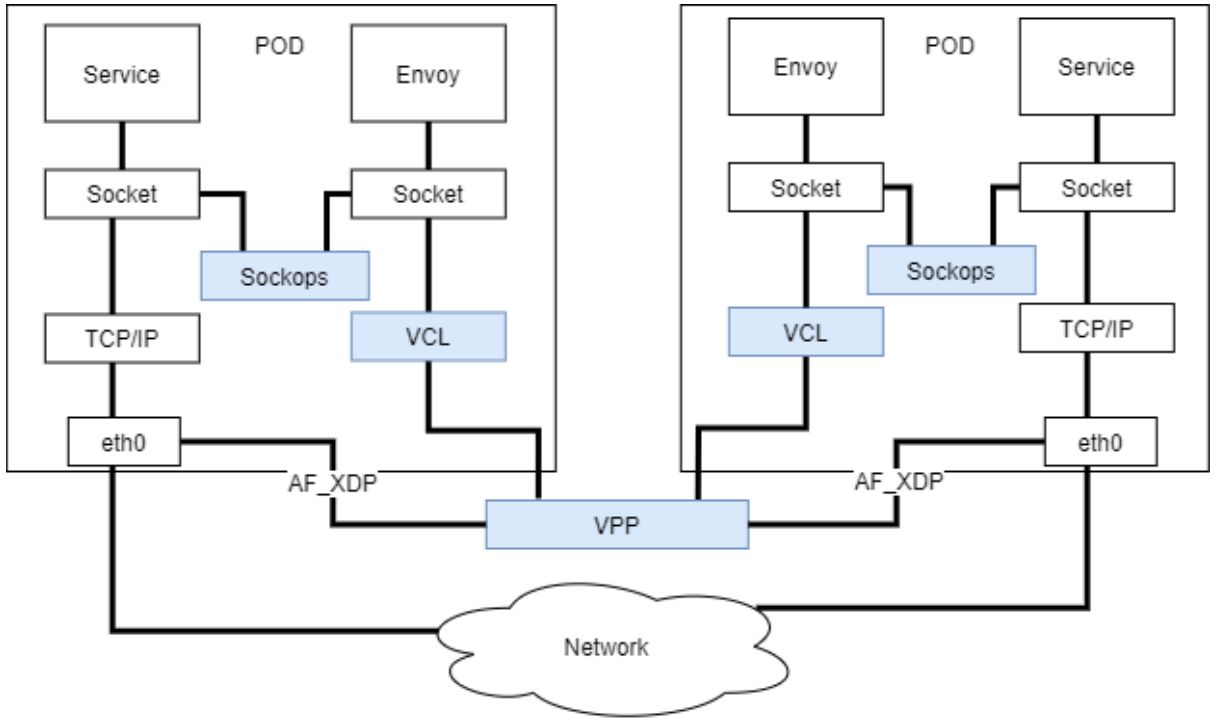
优点：

- ✓ 叠加SRIOV容器网络以及用户态协议栈，加速性能最好

缺点：

- ✓ 额外引入第二网络平面，组网较复杂
- ✓ 用户态协议栈需要独占1个物理核和20GB内存

主机网络 – 用户态协议栈兼容优先模式



- ✓ 借助AF_XDP对报文分流
- ✓ 用户态协议栈独立部署到VPP进程，解放envoy的CPU
- ✓ VCL库通过LD_PRELOAD劫持socket调用做到针对Envoy的无侵入加速
- ✓ Service和Envoy基于Sockops进行加速

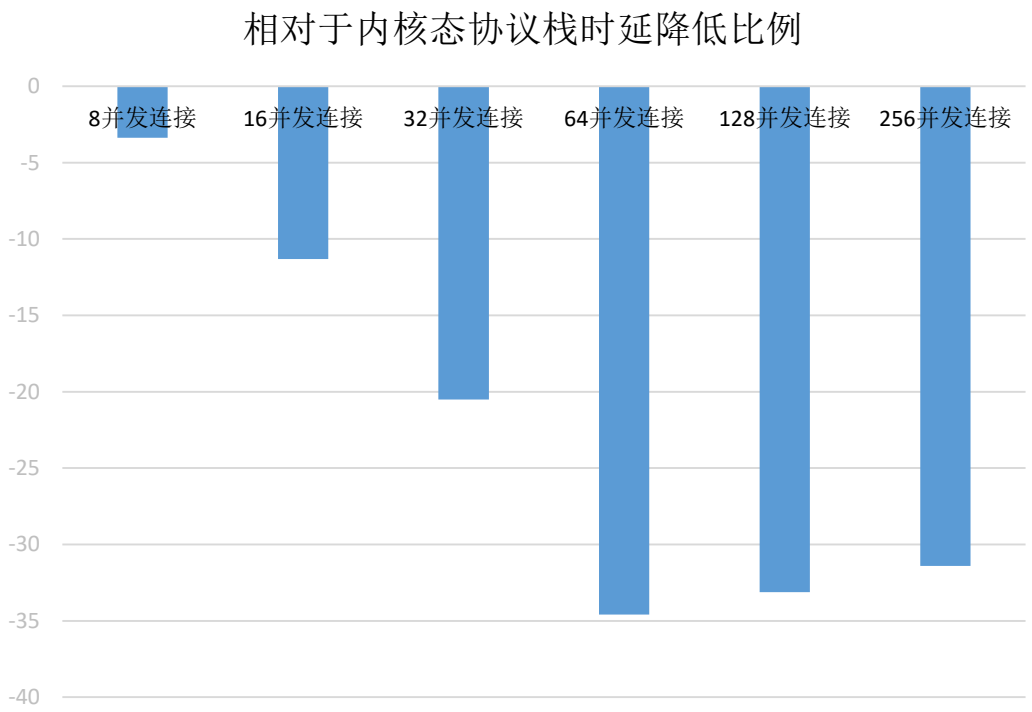
优点：

- ✓ 叠加eBPF Sockops以及用户态协议栈，加速性能好
- ✓ 即插即用，对现有网络无改动

缺点：

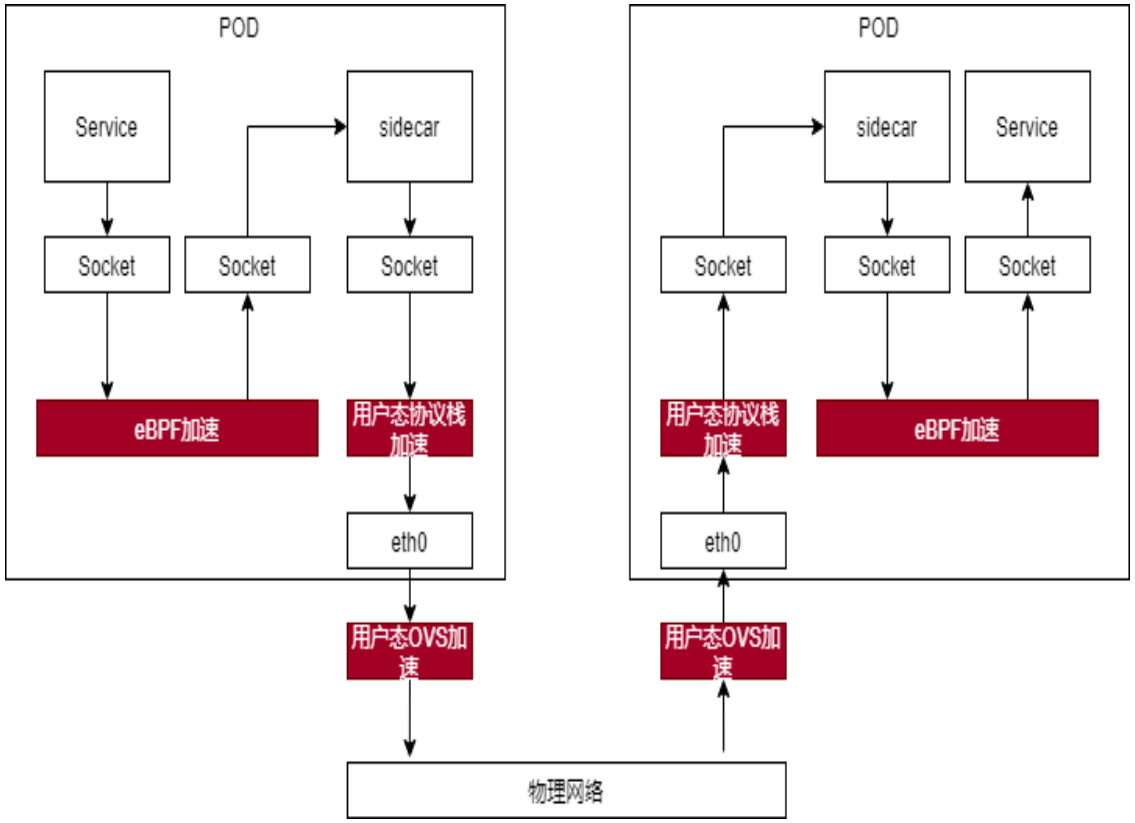
- ✓ 性能稍差于性能优先模式
- ✓ 用户态协议栈需要独占1个物理核和20GB内存

主机网络 – 用户态协议栈



- ✓ 压测场景下，如使用nighthawk加压背景流量，时延可以进一步降低10-20%

总结

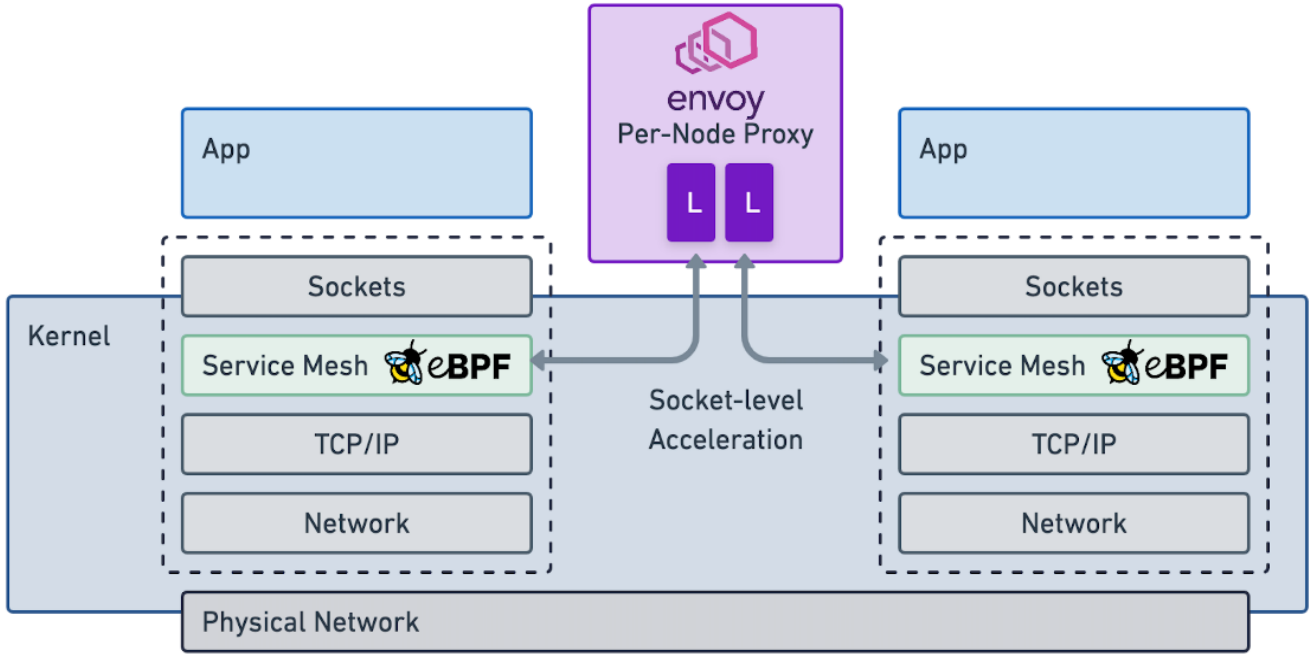


- ✓ 去除虚拟网络直接基于裸机，*10%*
- ✓ 用户态OVS加速容器网络，*20%*
- ✓ eBPF加速主机网络，*10%*
- ✓ 用户态协议栈加速主机网络，*30%*

可选的加速方式	适用场景	优势
用户态协议栈+eBPF	1.客户新建容器网络 2.客户已有容器网络，改造意愿不强	1.即插即用，现有网络无改造； 2.加速性能好
用户态协议栈+用户态OVS	1.客户已有容器网络，可适当改造 2.追求极致性能	1.加速性能最好
eBPF	1.客户新建容器网络 2.客户已有容器网络，改造意愿不强	1.即插即用，现有网络无改造，且无需额外占用CPU和内存资源； 2.加速性能一般

目录

Cilium service mesh



资源消耗

- ✓ PerPod->PerNode

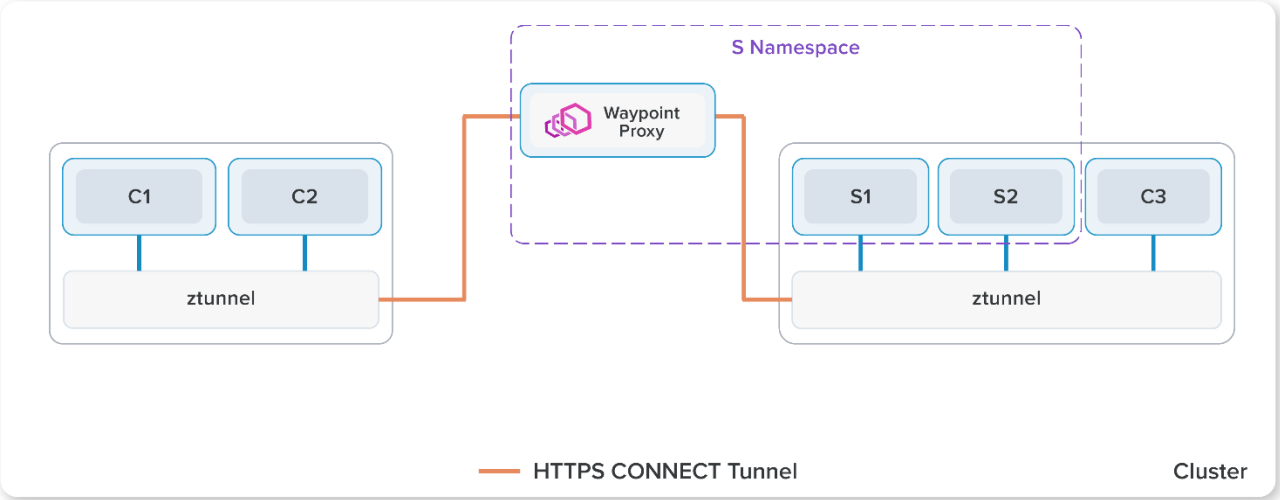
性能

- ✓ Cilium容器网络+Sockops
- ✓ L3/4不经过envoy，同节点少一级envoy
- ✓ 可结合用户态协议栈，实现envoy加速

限制

- ✓ 依赖于Cilium容器网络
- ✓ 部分envoy扩展功能不支持，如限速、故障注入
- ✓ 配置依赖于CNP/CEC，较复杂，无法和Istio兼容

Istio ambient mesh



资源消耗

✓ PerPod->PerServiceAccount

性能

✓ ztunnel 和 Waypoint 依 赖 于 envoy 和 iptables，相比sidecar无优势

✓ 可结合用户态协议栈优化Waypoint，Sockops优化ztunnel

限制

✓ beta版本，部分功能支持不全

THANKS