



IstioCon

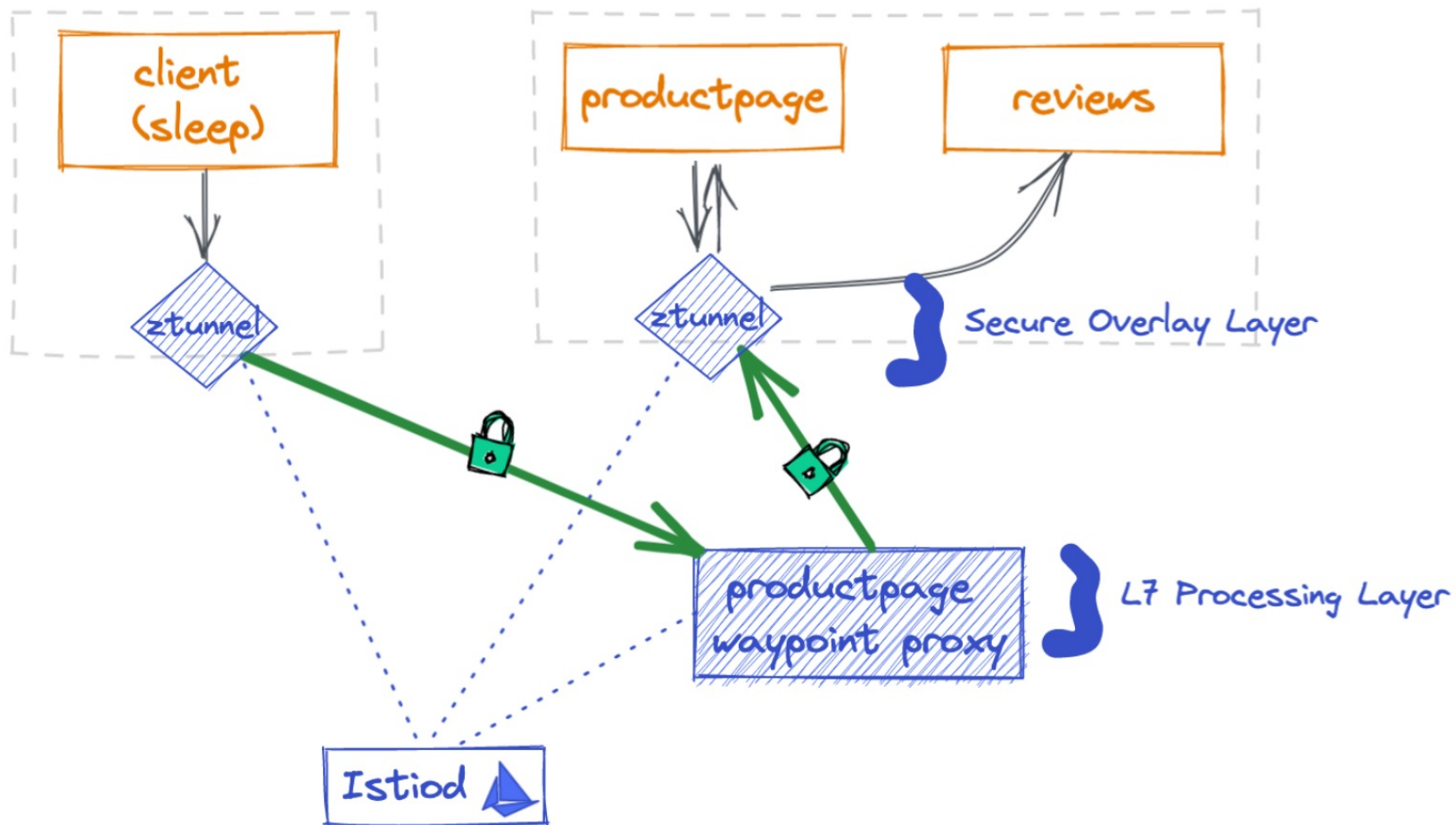


# Building an Efficient Service Mesh: Merbridge's Innovations in eBPF Implementation and Istio Ambient

*Kebe Liu*

# What is Ambient Mesh?

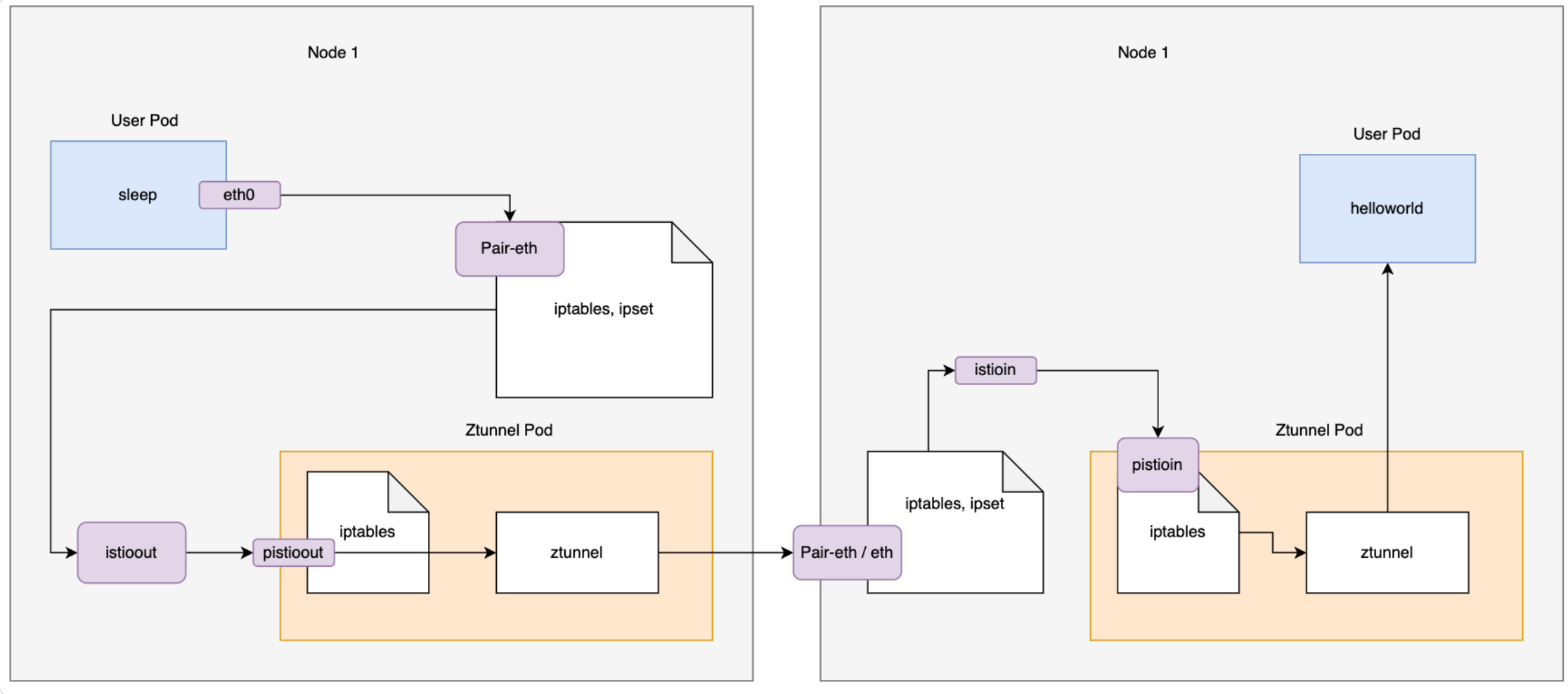
Ambient is a new data plane proxy mode that, through node proxy + Waypoint mode, can reduce resource usage and lower the difficulty of use compared to the Sidecar mode.



From: <https://istio.io/latest/blog/2022/introducing-ambient-mesh/>

# Ambient mode traffic forwarding path

kubectl exec deploy/sleep – curl helloworld:5000/hello Traffic path



# Ambient mode challenges

Because the Ambient mode uses iptables by default to control and forward traffic on the host, and many CNIs also manage container networks through policy routing or iptables, it can easily cause policy conflicts and become unavailable.

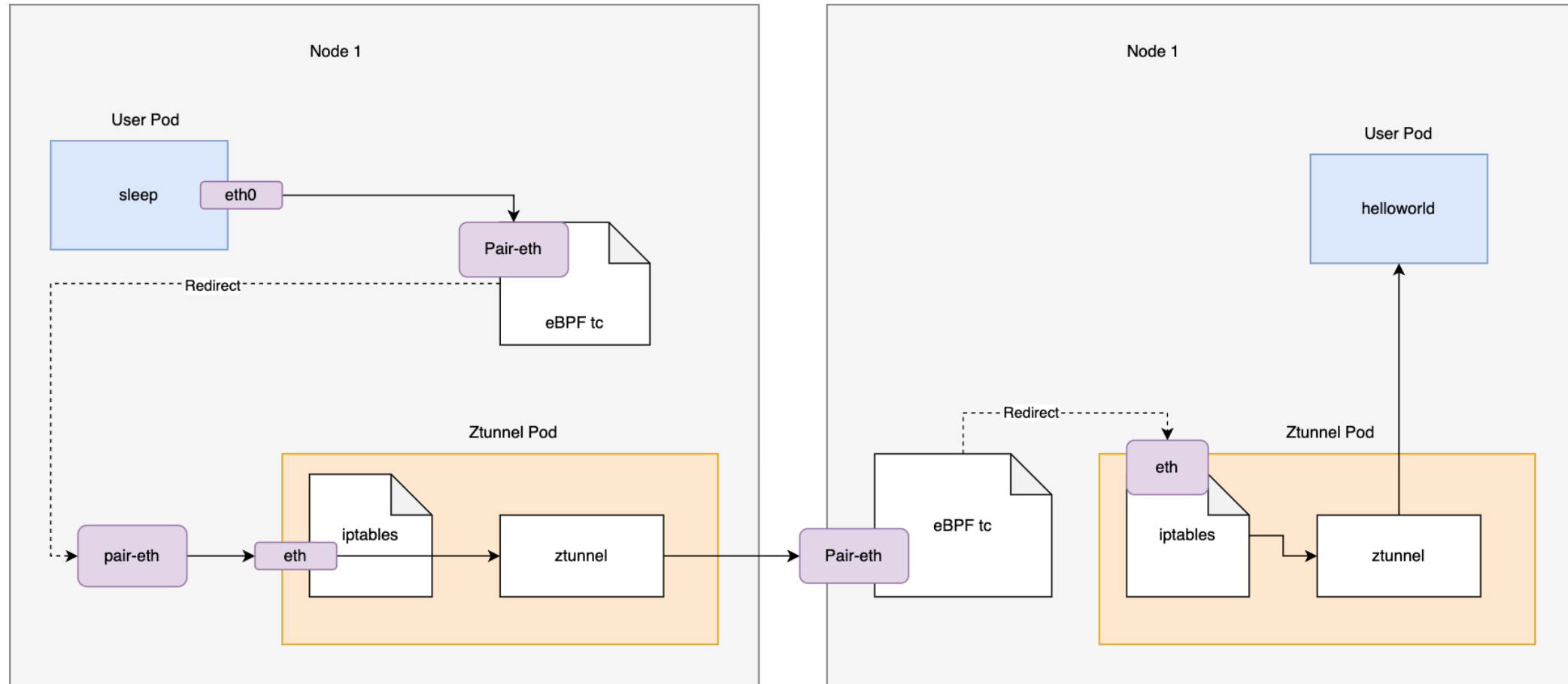
Therefore, the Ambient mode cannot currently adapt to all CNIs.

Based on this, the community has added support for eBPF to bypass iptables for traffic forwarding.

<input type="checkbox"/>	<input checked="" type="radio"/> 19 Open	<input checked="" type="radio"/> 34 Closed	Author ▾	Label ▾	Projects ▾
<input type="checkbox"/>	<input checked="" type="radio"/>		ambient: consider blocking usage in bad installation patterns	area/ambient	
			#46524 opened on Aug 15 by howardjohn		
<input type="checkbox"/>	<input checked="" type="radio"/>		ambient does not work on minikube	area/ambient	
			#46163 opened on Jul 25 by jmazzitelli	2 tasks done	
<input type="checkbox"/>	<input checked="" type="radio"/>		L4 AuthorizationPolicy not work properly	Ambient Beta	area/ambient
			#46057 opened on Jul 18 by fyuan1316	3 of 16 tasks	
<input type="checkbox"/>	<input checked="" type="radio"/>		Install Istio ambient failed - kubernetes1.27.3 + Calico3.25	area/ambient	area/environments
			#46054 opened on Jul 18 by Snuger	4 of 16 tasks	kind/need more info
<input type="checkbox"/>	<input checked="" type="radio"/>		Loose CNI for ambient	area/ambient	
			#46032 opened on Jul 15 by costinm	2 of 16 tasks	
<input type="checkbox"/>	<input checked="" type="radio"/>		Istio-CNI double pod get, ebpf handling	area/ambient	
			#45932 opened on Jul 11 by costinm	3 of 16 tasks	
<input type="checkbox"/>	<input checked="" type="radio"/>		ambient: app pod's readinessProbe cannot work when ambient is installed via manifests	area/ambient	area/environments
			#45781 opened on Jul 3 by nak3	2 tasks done	Ambient Beta
<input type="checkbox"/>	<input checked="" type="radio"/>		Ambient does not work with Azure Networking Policy	area/ambient	area/networking
			#45760 opened on Jul 1 by Stevenjin8	4 of 16 tasks	
<input type="checkbox"/>	<input checked="" type="radio"/>		Pod with sidecar not going Ready when Ambient is running	area/ambient	area/environments
			#45393 opened on Jun 10 by myshkin5	4 of 16 tasks	
<input type="checkbox"/>	<input checked="" type="radio"/>		Add integration test for ambient eBPF redirection	area/ambient	
			#44466 opened on Apr 21 by PlatformLC		

# Using eBPF and TC to Optimize Traffic Paths

kubectl exec deploy/sleep – curl helloworld:5000/hello Traffic path with eBPF / TC

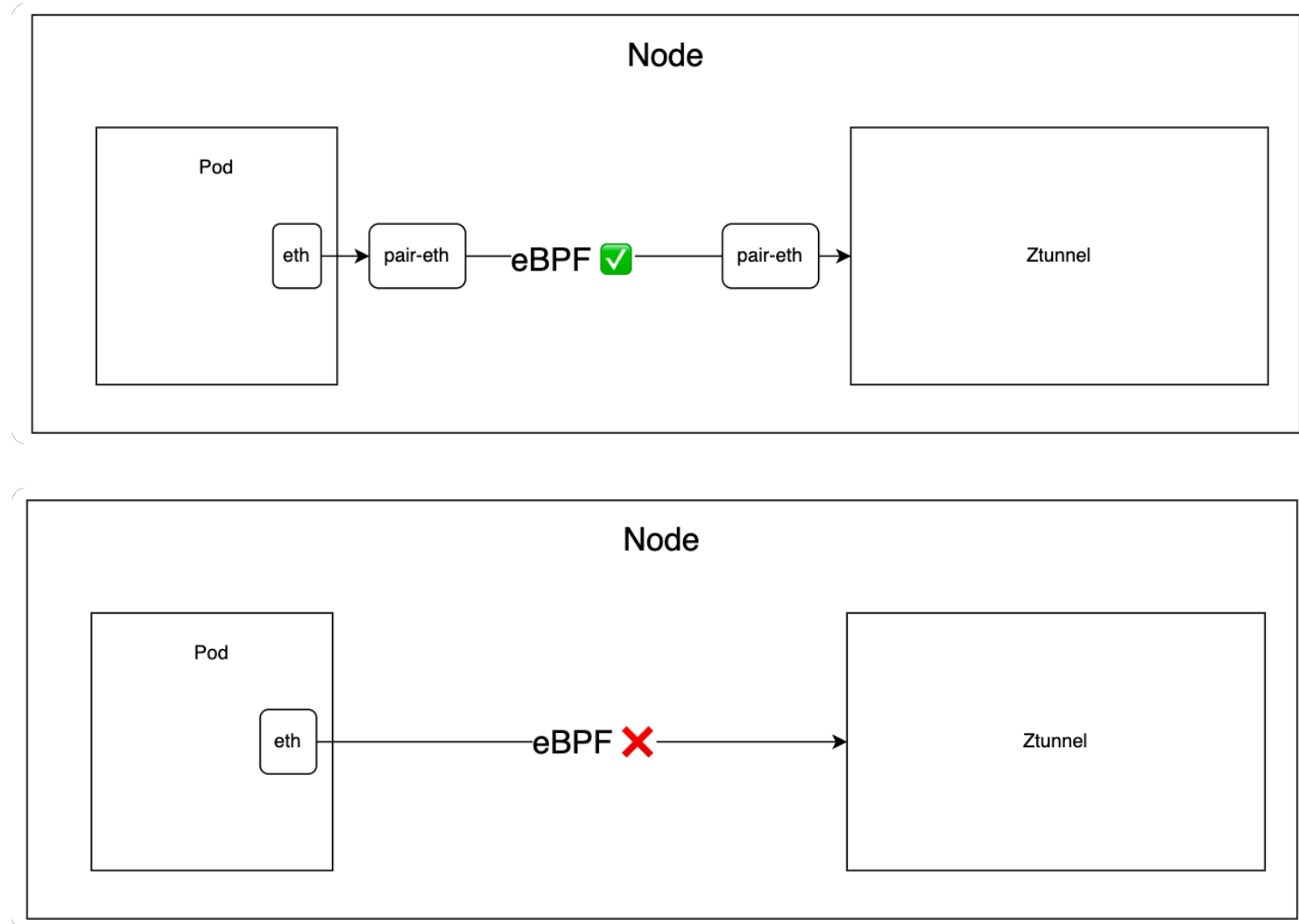


# Problems with Ambient eBPF Mode

Ambient's eBPF mode, by applying eBPF programs on the pair network interface of the Pod, forwards traffic to the pair network interface of ztunnel, thereby allowing traffic to enter ztunnel and replacing the functionality of iptables. This approach can alleviate the difficulty of supporting CNI.

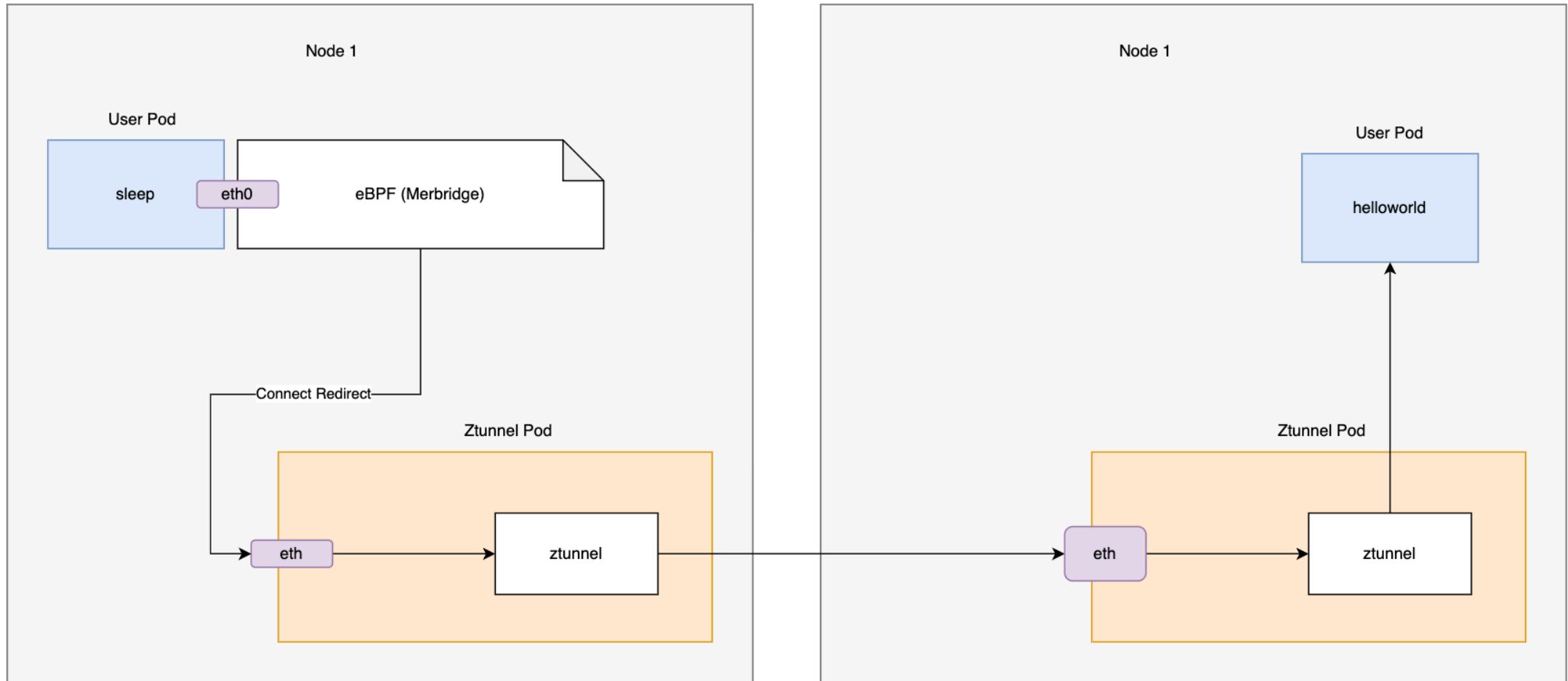
However, this mode relies on CNIs that use the pair network interface mode. For some CNIs implemented in macvlan mode, since there is no pair network interface on the host, this mode cannot be used.

(PS: The core reason is that tc redirect cannot cross network namespaces.)



# Using Merbridge to Enhance Ambient

kubectl exec deploy/sleep – curl helloworld:5000/hello Traffic path with Merbridge

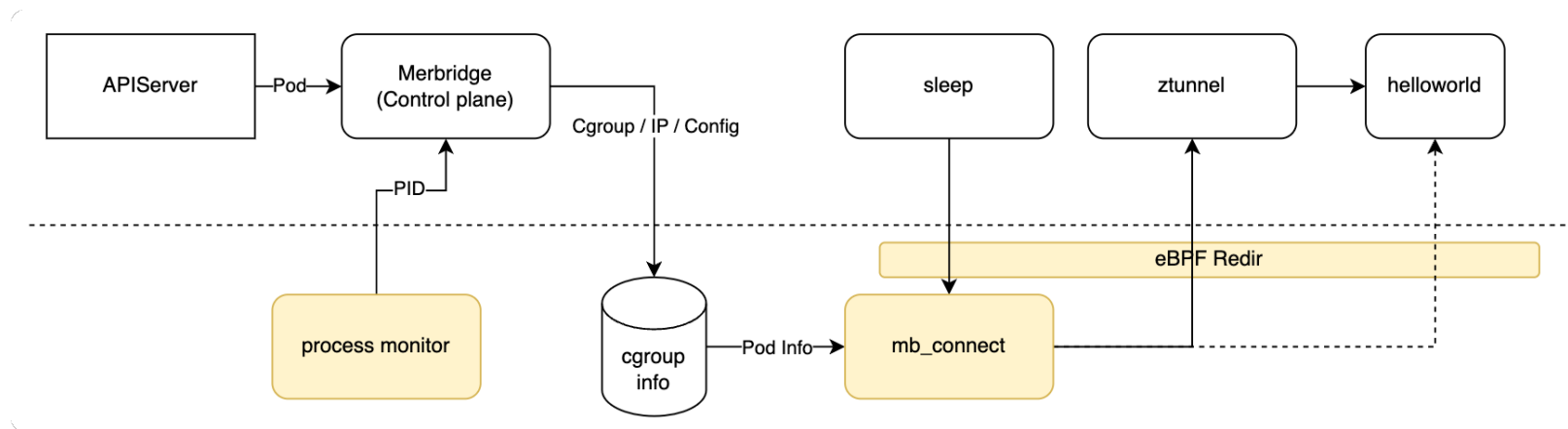




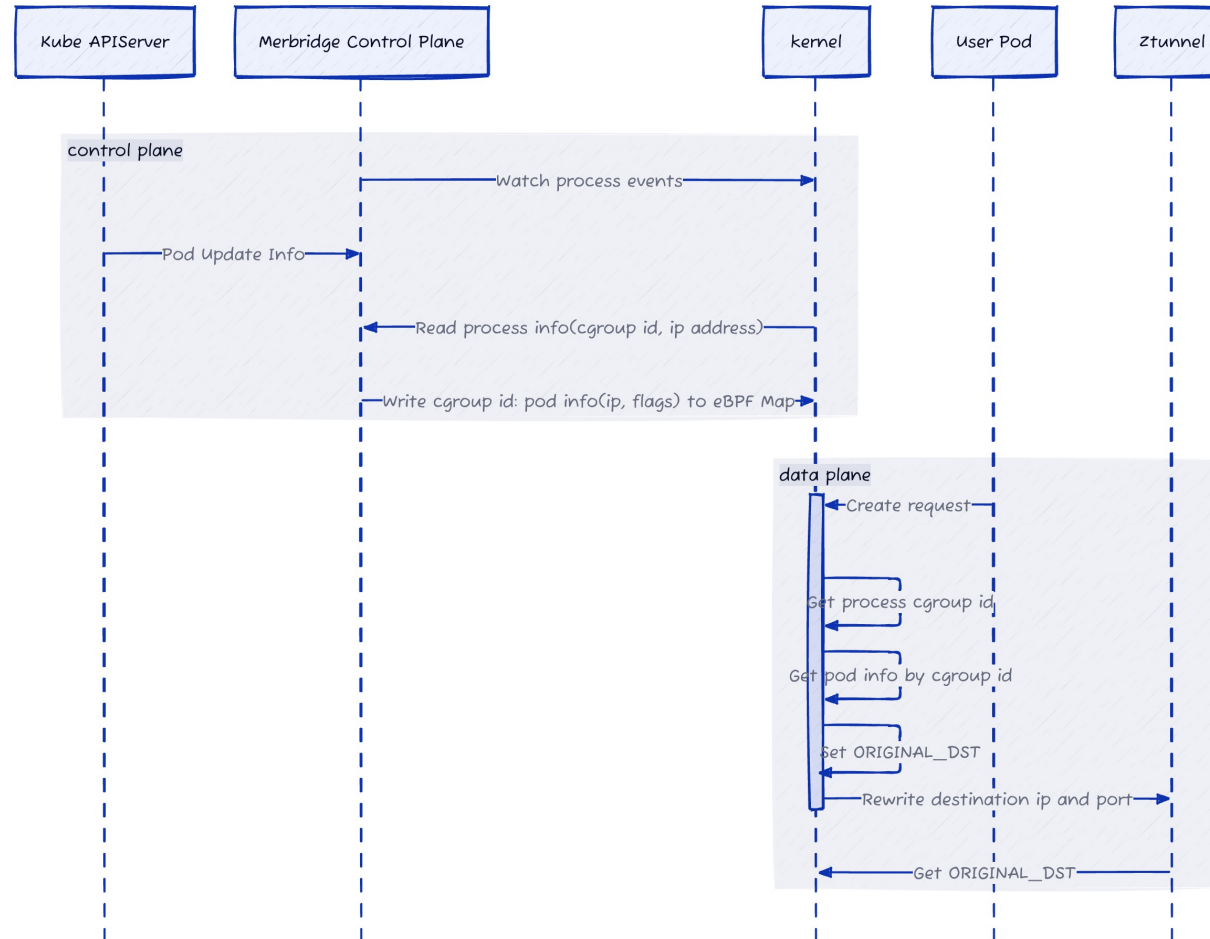
# How to handle traffic forwarding?

In Ambient mode, by matching whether the source IP of the traffic is in Ambient mode on the host, it is determined whether to forward the traffic to Ztunnel.

However, the eBPF program of Merbridge will directly determine whether the traffic needs to be forwarded to ztunnel when the process in the Pod initiates a request (before it is processed by the kernel protocol stack and the source IP is not bound yet).



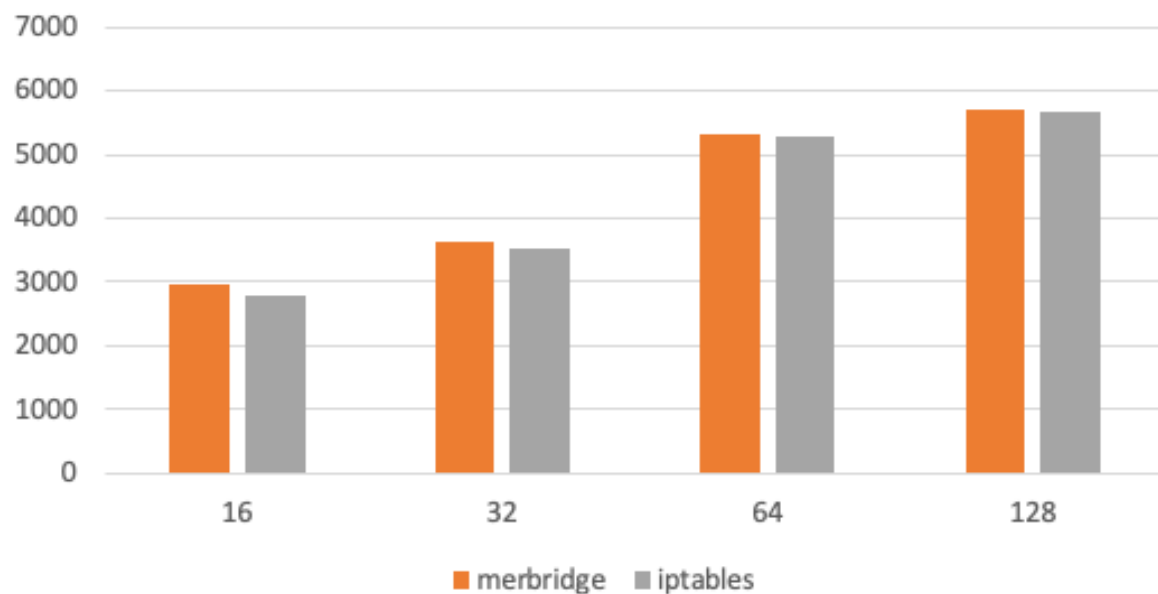
# How to handle traffic forwarding?



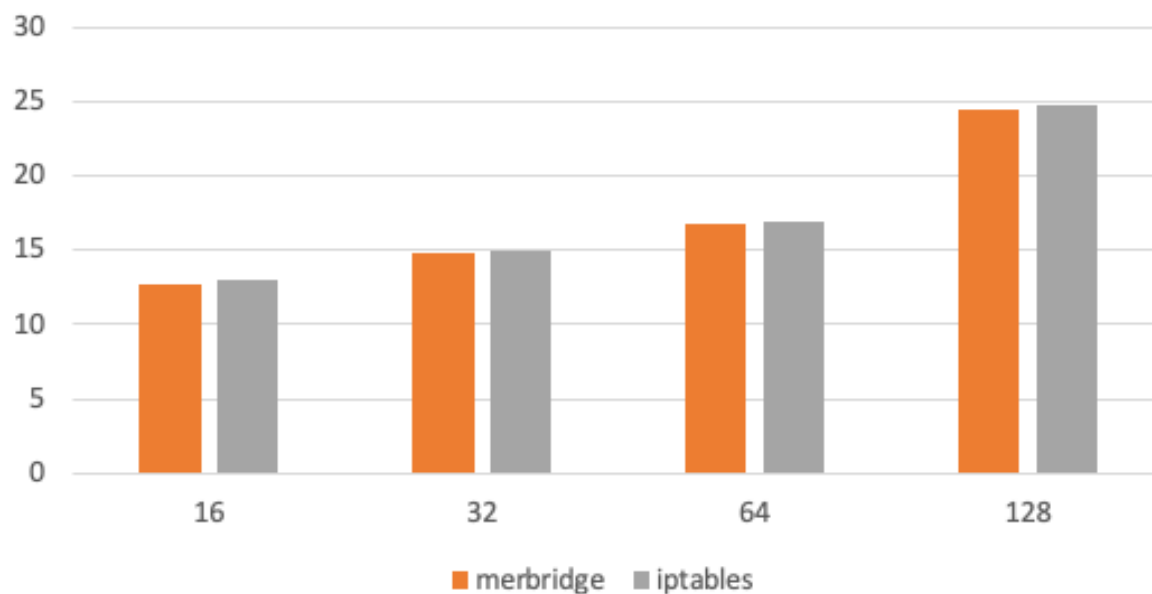
- ORIGINAL\_SRC  
Ztunnel will bind the source address of its traffic sent to other Pods to the IP address of the requesting Pod (rather than ztunnel itself). However, Merbridge does not rely on iptables, so it will be modified when ztunnel binds the source address.
- msg\_redirect  
Because ztunnel uses zero copy between downstream and upstream to accelerate the transmission of layer 4 data, in this situation, bpf\_msg\_redirect cannot be used directly. This will result in the acceleration of bpf\_msg\_redirect being unavailable in some paths.

```
if ((cg_info.detected_flags & ZTUNNEL_FLAG) &&
    (cg_info.flags & ZTUNNEL_FLAG)) {
    // ztunnel
    __u32 *ztunnel_ip = get_ztunnel_ip();
    if (!ztunnel_ip) {
        debugf("can not get ztunnel pod ip in bind");
        return 1;
    }
    // ztunnel will bind the source pod ip to upstream,
    // we will rollback this operation because we not support TPROXY mode.
    ctx->user_ip4 = ztunnel_ip[3];
    debugf("successfully rewrite ztunnel bind");
}
return 1;
```

QPS



Latency



Github: <https://github.com/merbridge/merbridge>

Slack: [https://join.slack.com/t/merbridge/shared\\_invite/zt-11uc3z0w7-DMyv42eQ6s5YUxO5mZ5hwQ](https://join.slack.com/t/merbridge/shared_invite/zt-11uc3z0w7-DMyv42eQ6s5YUxO5mZ5hwQ)