

6045 // X00

Włodzisławski

May 2023

Ondrej Blazek
Infrastructure Engineer
Seznam.cz

- BPF is super flexible (networking, tracing, security)
- BPF since 1992 (cBPF, eBPF, kernel 3.18)
- programs verified, "safe to run" (privileged)
- efficient - compiled into native instructions (1M limit)
- event driven (network events, syscalls, function entry/exit, kprobes, uprobes)

- XDP - earliest possible point in the NIC driver
- Traffic control ingress/egress (tc) - after initial packet processing
- Socket operations - specific cgroup (e.g. redirecting to another socket)

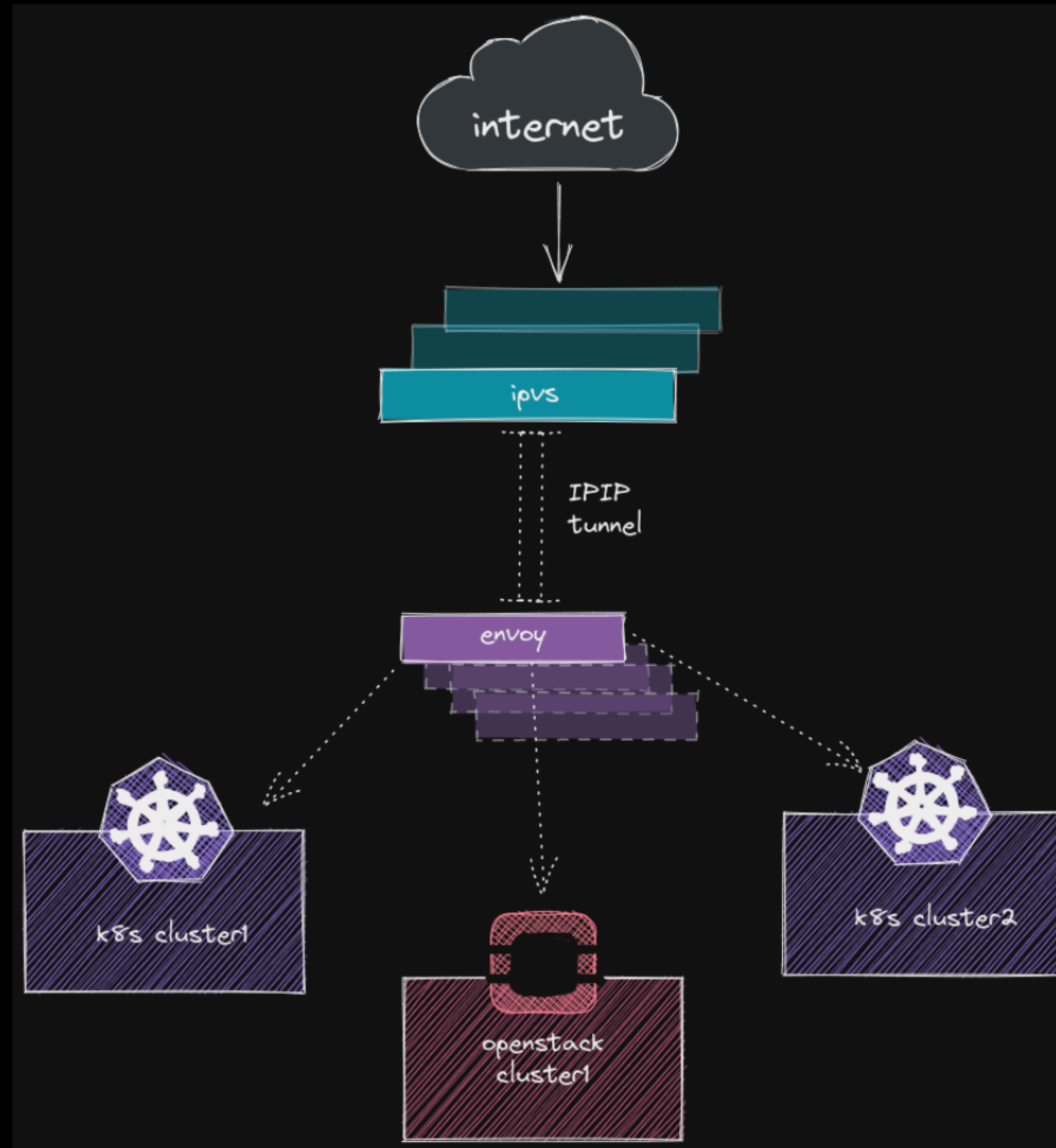
- Packet resides in kernel space (no user/kernel space crossing)
- Compatible with kernel's TCP/IP stack (no bypass)
- No need for dedicated CPUs
- For a while now core part of kernel
- Compatible with most 10G+ NICs
- Operates directly on packet data (via registers)
- Ideal for:
 - DDoS protection (extremely low per-packet cost)
 - Loadbalancing XDP_TX, XDP_REDIRECT

```
1 #include <linux/bpf.h>
2
3 #ifndef __section
4 # define __section(NAME)          \
5     __attribute__((section(NAME), used))
6 #endif
7
8 __section("prog")
9 int xdp_drop(struct xdp_md *ctx)
10 {
11     return XDP_DROP;
12 }
13
14 char __license[] __section("license") = "GPL";
```

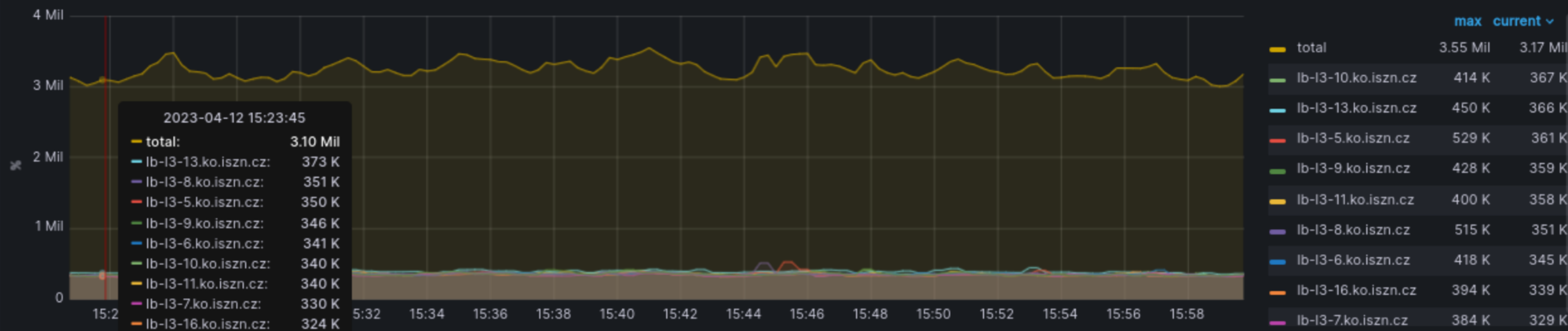
```
$ clang -O2 -Wall -target bpf -c xdp-example.c -o xdp-example.o  
# ip link set dev eth0 xdp obj xdp-example.o
```



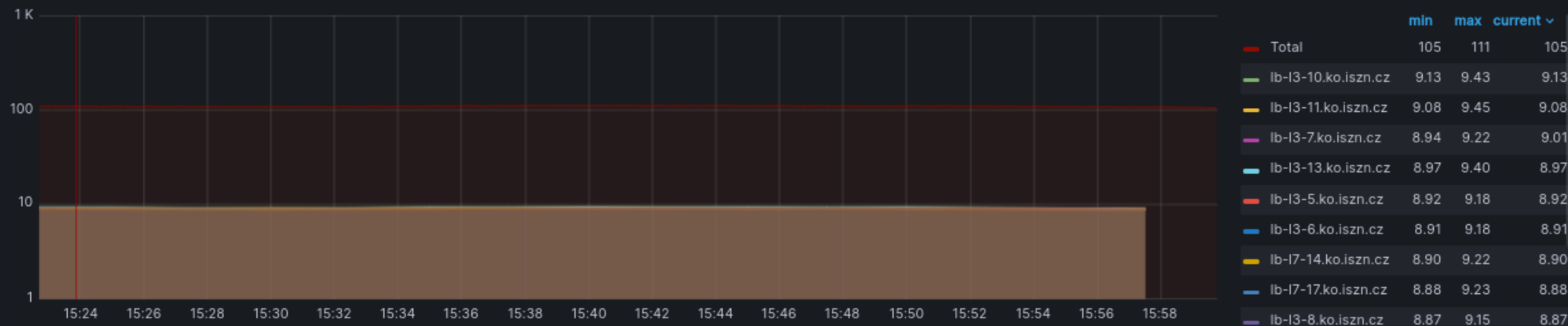
```
# ip l sh eth0
2: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 xdp qdisc mq state UP mode
  DEFAULT group default qlen 1000
    link/ether b0:26:28:7a:f9:40 brd ff:ff:ff:ff:ff:ff
    prog/xdp id 485 tag b2933a93f4e20b92 jited
```

Packets per second L7 nodes

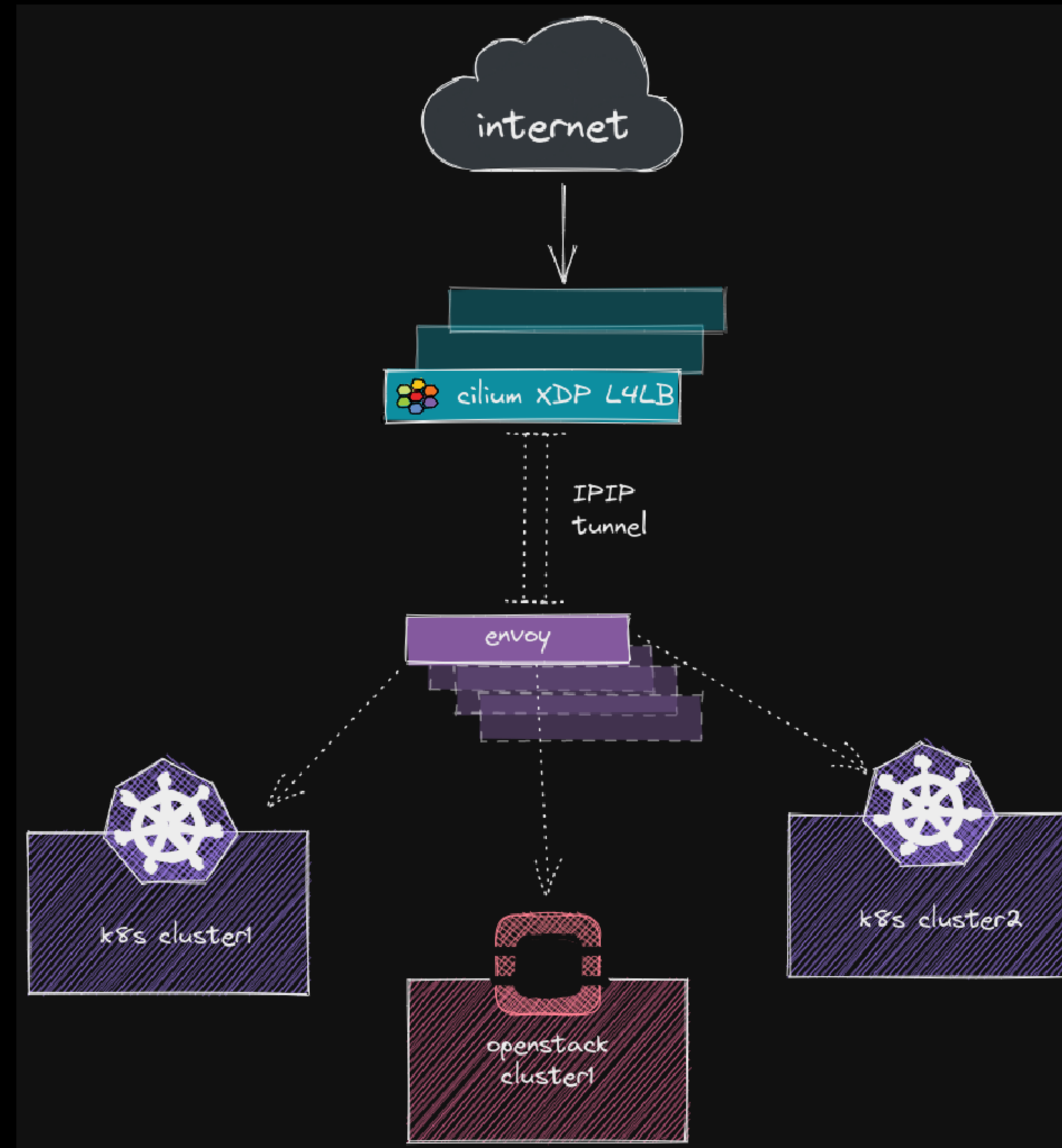


CPU load (all CPUs)



IPVS L4LB

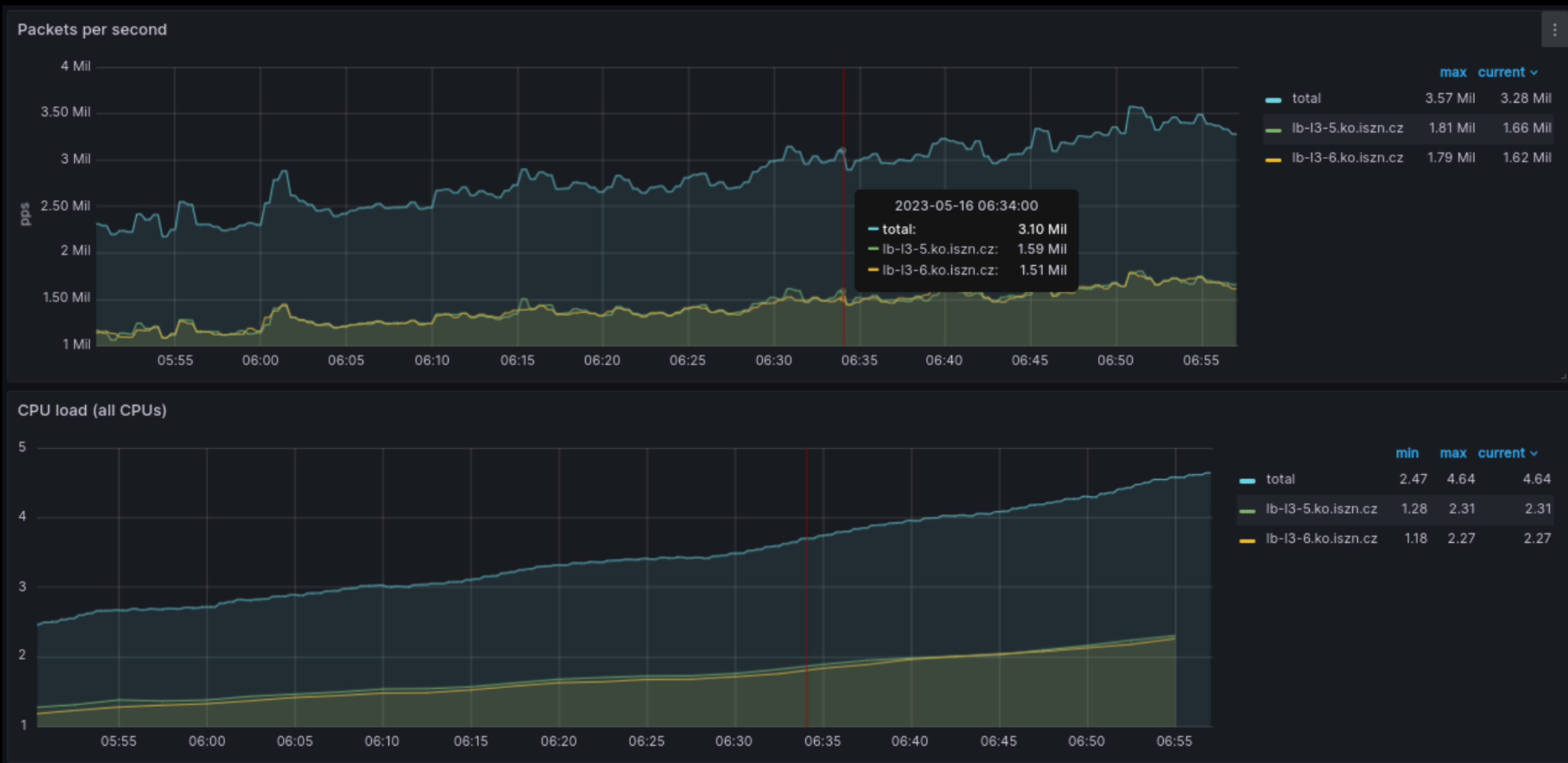
- highly scalable software loadbalancer
- in linux kernel since version 2.4
- ipip, DSR, hairpin mode
- used also by FB before Katran announcement



Cilium XDP

L4LB

- generically programmable L4LB based on eBPF
- same features as IPVS (and more)
- part of cilium 1.10 (today 1.13+)



#####

THANK YOU

#