

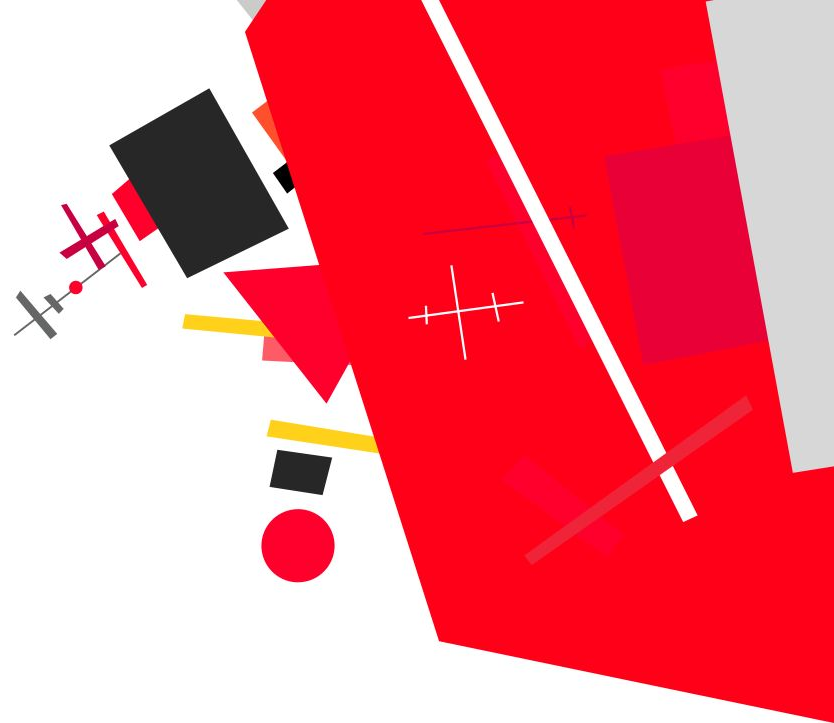
HIGH SPEED LOAD BALANCING FROM THE LINUX KERNEL

Laura García
ZEVENET



HighLoad⁺⁺

Профессиональная конференция
разработчиков высоконагруженных
систем



Capabilities



Multilayer



Multiplatform



REST API



Web GUI

Focused on



Security



High Availability



Scalability



Performance

Research



+ Concurrent users



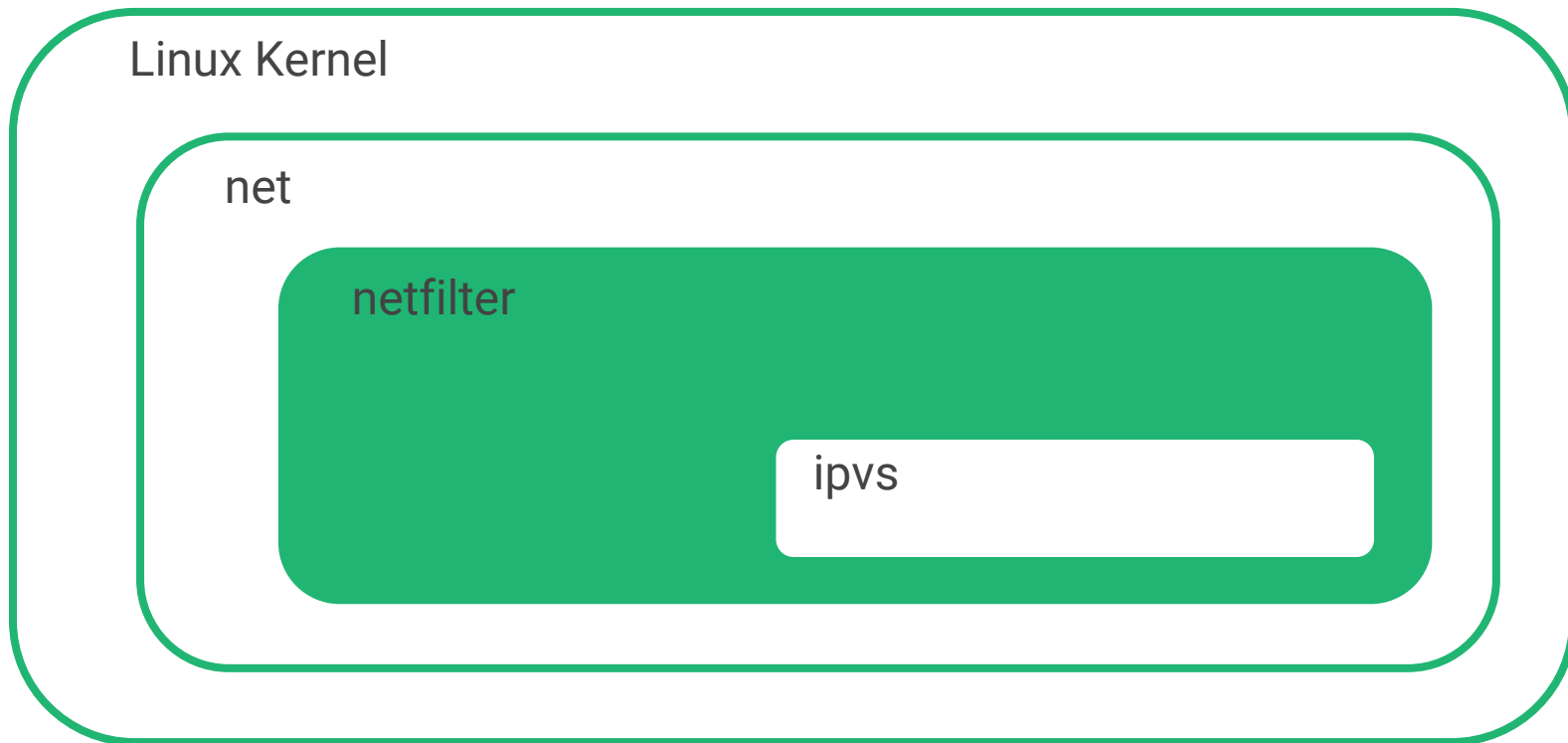
- CPU cycles



Research



Research



Development with iptables

{ sNAT
dNAT }

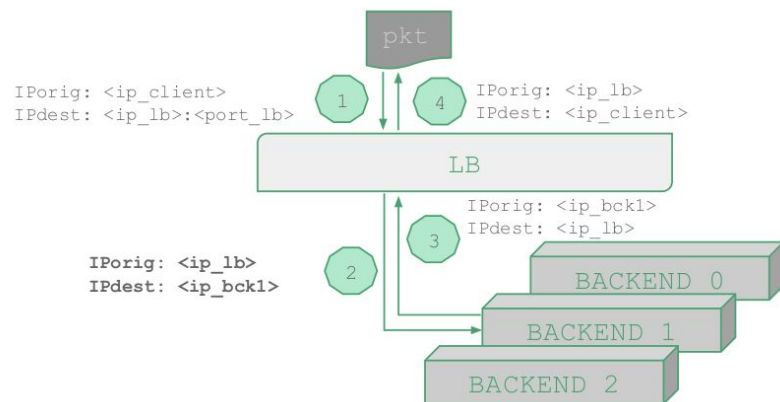
{ conntrack
helpers
conntrackd
syncd
xtables }

{ sip
(t)ftp
sctp
...

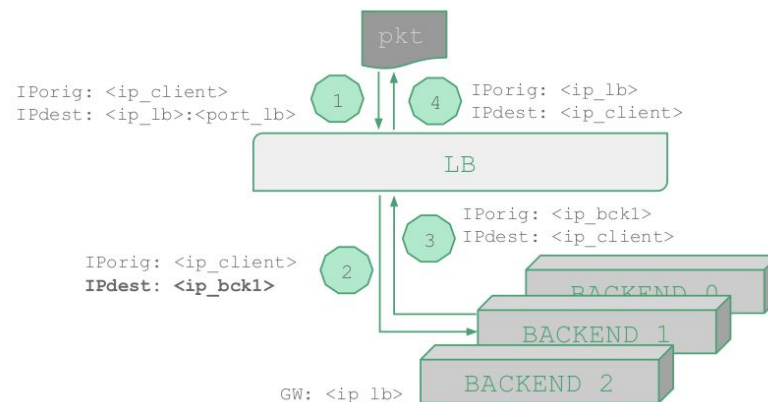
{ multiport
limits
recent
statistic }

Development with iptables

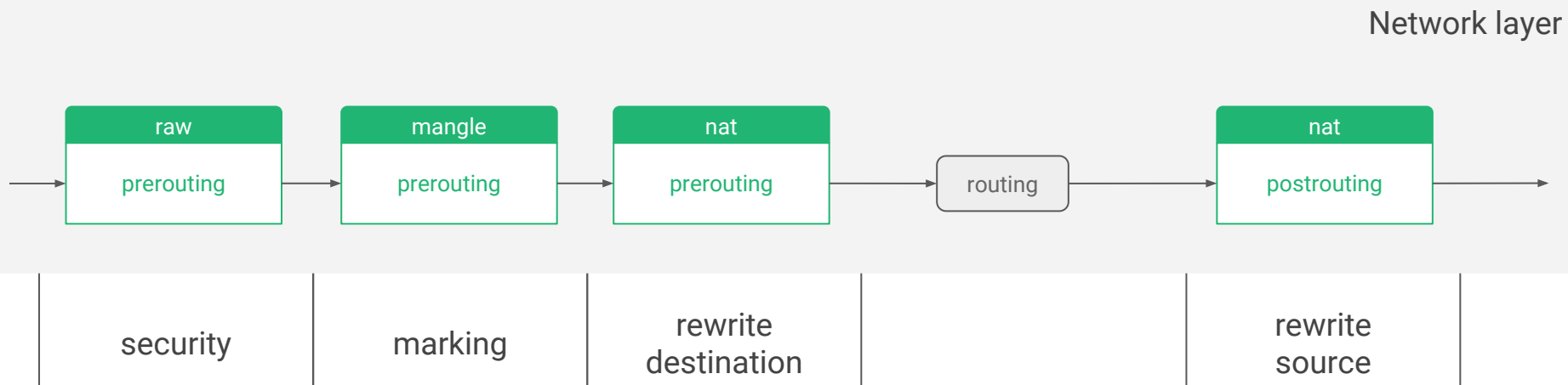
source NAT



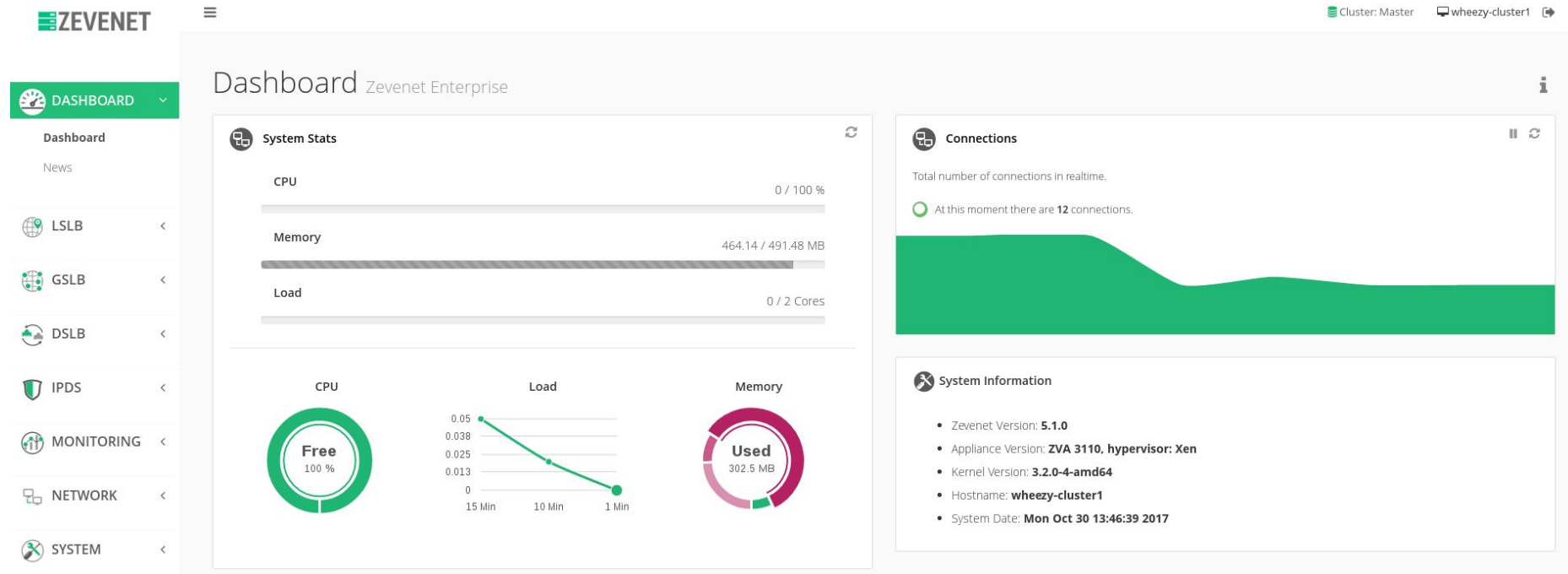
destination NAT



Development with iptables



Web GUI



API

```
curl -k -H "ZAPI_KEY: MyK3y...X" https://192.168.100.204:444/zapi/v3.1/zapi.cgi/farms
```

```
{
  "description" : "List farms",
  "params" : [
    {
      "farmname" : "http-profile-farm1",
      "profile" : "http",
      "status" : "up",
      "vip" : "192.168.100.208",
      "vport" : "80"
    },
    {
      "farmname" : "LSLB-farm1",
      "profile" : "l4xnat",
      "status" : "up",
      "vip" : "192.168.100.207",
      "vport" : "222"
    },
    {
      "farmname" : "NewGSLB-farm1",
      "profile" : "gslb",
      "status" : "up",
      "vip" : "192.168.100.207",
      "vport" : "53"
    }
  ]
}
```

API

```
curl -k -X POST -H 'Content-Type: application/json' -H "ZAPI_KEY: MyK3y...X" -d '{"ip":"192.168.100.254","port":80}' \
https://192.168.100.204:444/zapi/v3.1/zapi.cgi/farms/webfrontend/backends

{
  "description" : "New farm backend",
  "message" : "Backend added",
  "params" : {
    "id" : 1,
    "ip" : "192.168.100.254",
    "max_conns" : "0",
    "port" : 80,
    "priority" : null,
    "weight" : null
  },
  "status" : "up"
}
```

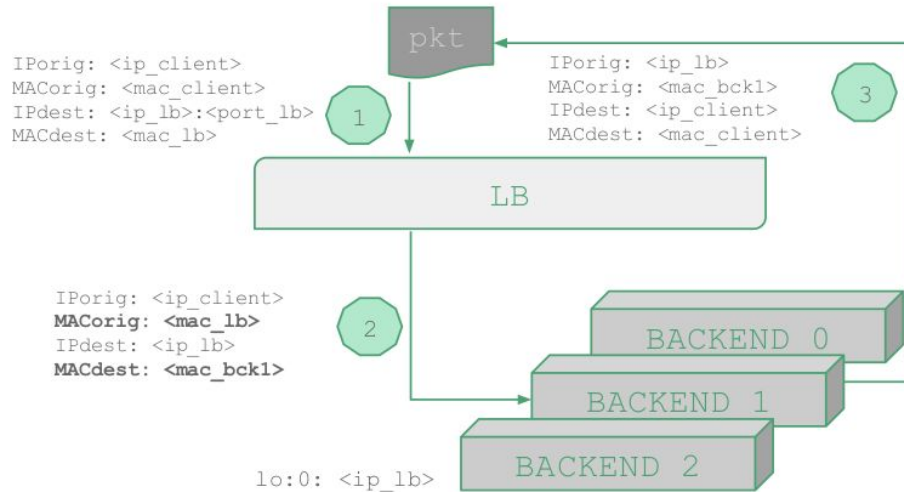
Development with nftables

{ expressive
native expressions
Ingress, egress hook
conntrack, helpers, etc.
stateless NAT }

{ Expressions: nth, random, hash, etc.
Models for dNAT, sNAT and DSR }

Development with nftables

Direct Server Return



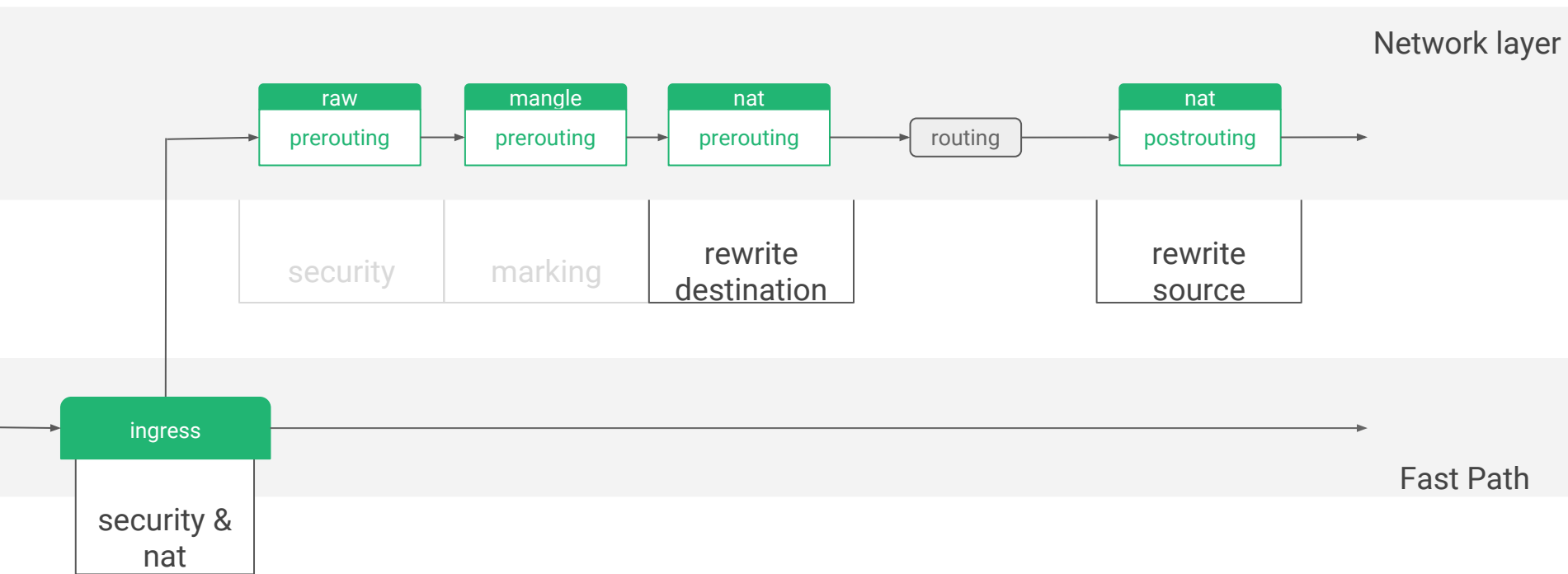
Development with nftables

```
table netdev filter {
    chain ingress {

        type filter hook ingress device <if_lb> priority 0; policy accept;
        ip daddr <ip_lb> udp dport <port_lb> ether saddr set <mac_lb> \
            ether daddr set numgen inc mod 3 \
            map { \
                0: <mac_bck0>, \
                1: <mac_bck1>, \
                2: <mac_bck2> } \
            fwd to <if_lb>

    }
}
```

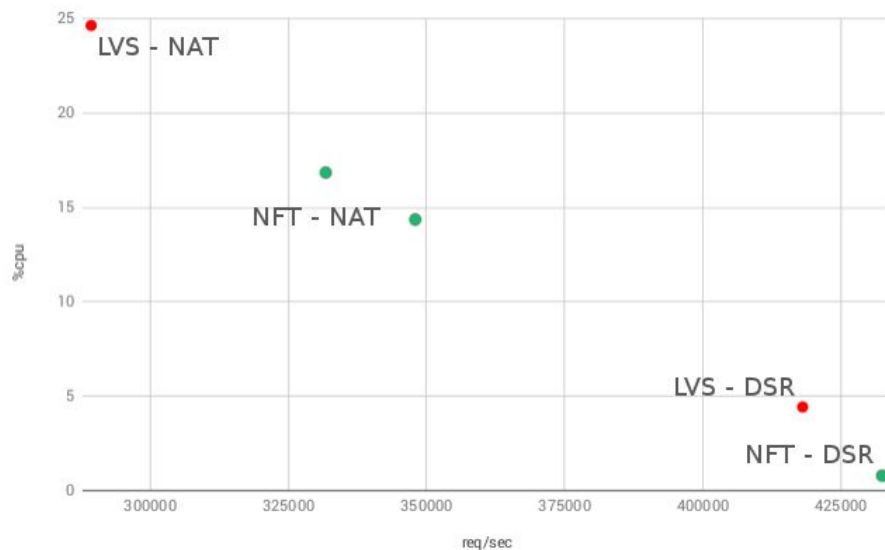
Development with nftables



Benchmarks

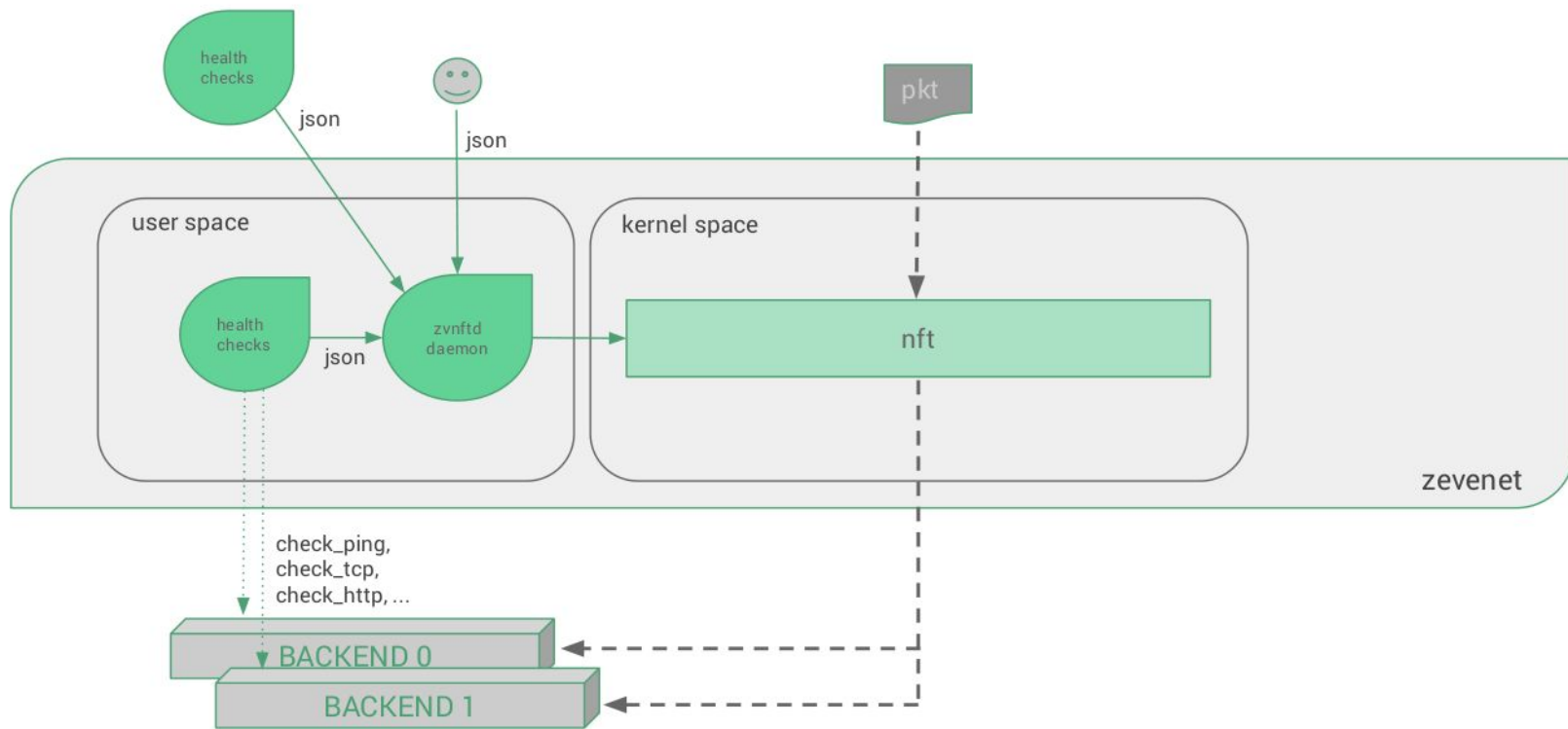


Benchmarks - %cpu vs. req/sec



~5x-6x
10x

New L4 core zvnftd



New challenges

- ★ libnftables
- ★ Layer 7 preprocessor
- ★ kTLS
- ★ HW offload
- ★ programmability

Thank you!



laura.garcia@zevenet.com



<https://github.com/zevenet>



<https://hub.docker.com/r/zevenet/>