eBPF at LINE's Private Cloud
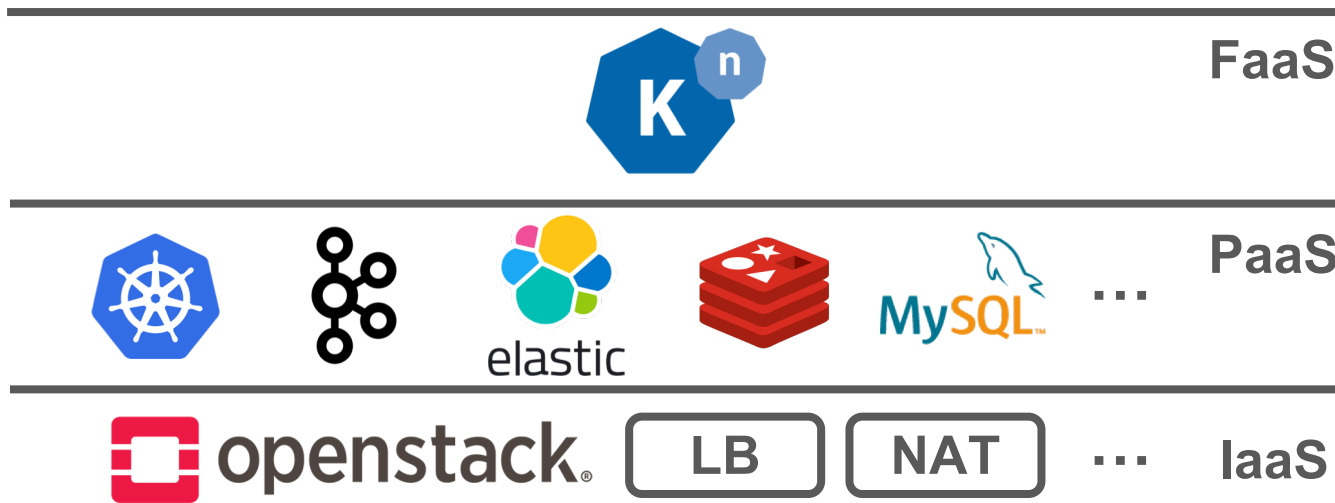
Yutaro Hayakawa

October 28, 2020

LINE

- **Messaging & many family services**

- **185 million global MAU**

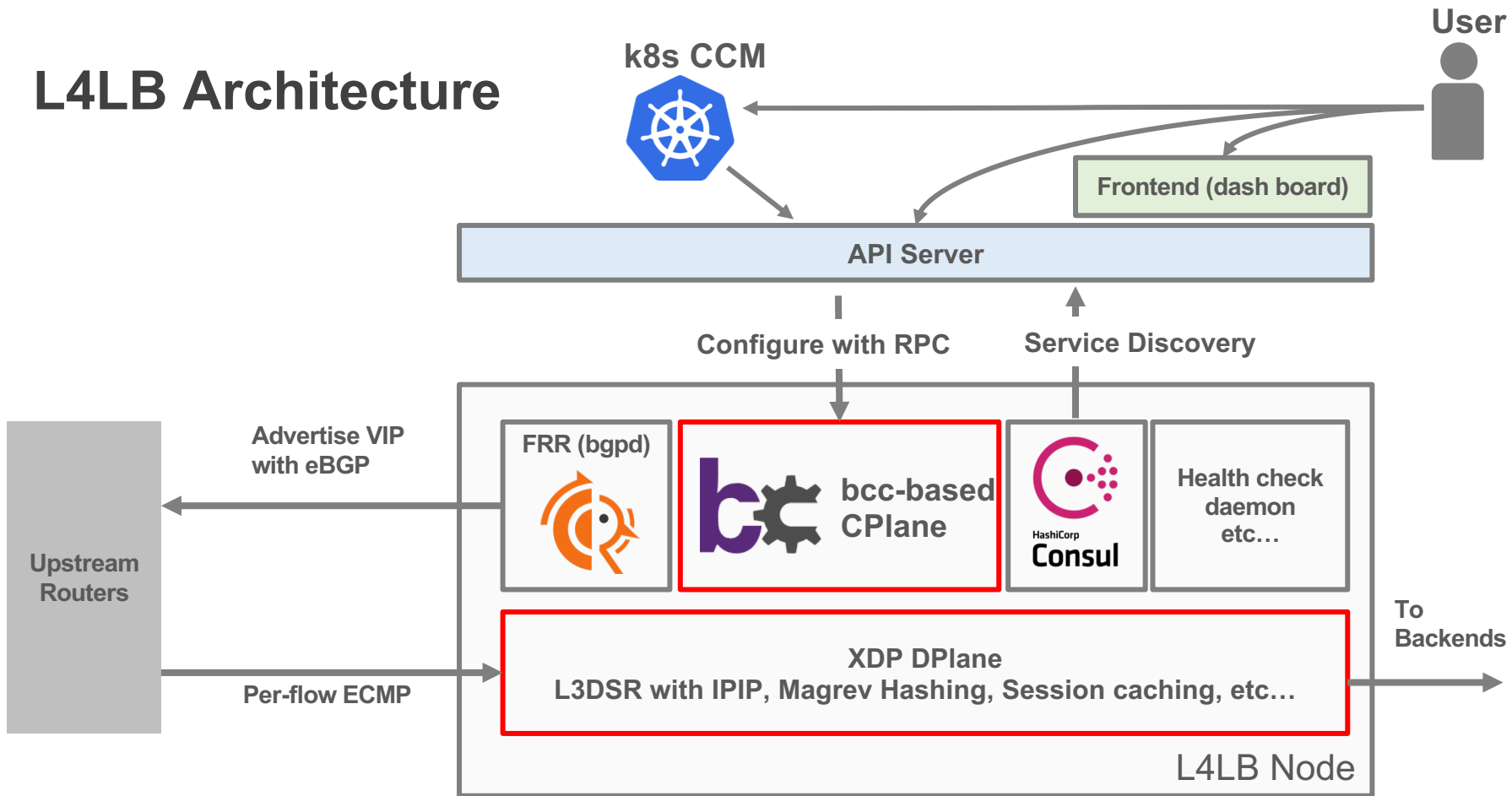- **3Tbps+ network traffic in total**

Verda: LINE's Private Cloud Service

Verda and XDP Based L4 Load Balancer Service

- **Part of our private cloud service since 2017**

- **5100 private, 760 public VIPs**

- **k8s CCM integration (Type: LoadBalancer)**

L4LB Architecture

For More Information

- **Our motivation, detailed architecture, etc… (en)**
  - **https://www.youtube.com/watch?v=UE6rPA1Js2s&feature=emb_title**
  - **https://speakerdeck.com/line_devday2019/software-engineering-that-supports-line-original-lbaas**

ipftrace

- **Network domain specific function call tracer**

- **Trace "which packets have gone through which functions"**

```
// Trace the TCP packets with destination 10.0.0.10

# iptables -t raw -A OUTPUT -p tcp -d 10.0.0.10 -j MARK --set-mark 0xdeadbeef

# ipft -m 0xdeadbeef
```

Output

Attaching program (total 1803, succeeded 1001, failed 0 filtered: 0 untraceable: 802)
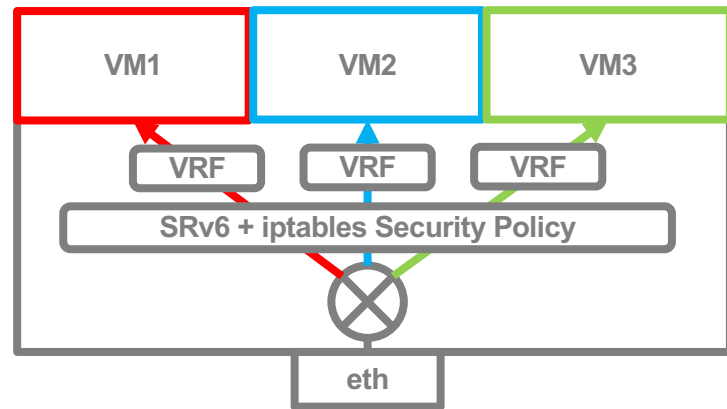Trace ready!
Samples: 246 Lost: 0^C
Trace done!
===

Time Stamp	CPU ID	Functions the packets have gone through	User defined tracing data (with Lua script)
3347634373462	0000	selinux_ipv4_output	(len: 5764 gso_type: tcpv4)
3347634379670	0000	ip_output	(len: 5764 gso_type: tcpv4)
3347634382597	0000	nf_hook_slow	(len: 5764 gso_type: tcpv4)
3347634385879	0000	selinux_ipv4_postroute	(len: 5764 gso_type: tcpv4)
3347634388958	0000	selinux_ip_postroute	(len: 5764 gso_type: tcpv4)
3347634391979	0000	ip_finish_output	(len: 5764 gso_type: tcpv4)
3347634394932	0000	__cgroup_bpf_run_filter_skb	(len: 5764 gso_type: tcpv4)
3347634398196	0000	ip_finish_output2	(len: 5764 gso_type: tcpv4)
3347634401431	0000	neigh_direct_output	(len: 5764 gso_type: tcpv4)
3347634404503	0000	dev_queue_xmit	(len: 5764 gso_type: tcpv4)
3347634407363	0000	__dev_queue_xmit	(len: 5764 gso_type: tcpv4)
3347634410290	0000	netdev_pick_tx	(len: 5764 gso_type: tcpv4)
3347634413287	0000	validate_xmit_skb	(len: 5764 gso_type: tcpv4)
3347634416425	0000	netif_skb_features	(len: 5764 gso_type: tcpv4)
3347634419602	0000	skb_network_protocol	(len: 5764 gso_type: tcpv4)
3347634422951	0000	skb_csum_hwoffload_help	(len: 5764 gso_type: tcpv4)

Use case

- **Multi tenant HV networking using SRv6 + VRF**

- **Contributed to find the bug in SRv6 GSO handling**

- **Upstream change**
  - **https://github.com/torvalds/linux/commit/62ebaeaedee7591c257543d040677a60e35c7aec**

For More Information

- **Our SRv6 DC network architecture (en)**

  - **https://speakerdeck.com/line_developers/line-data-center-networking-with-srv6**

- **Detailed investigation of SRv6 TSO/GSO issue (jp)**

  - **https://engineering.linecorp.com/ja/blog/tso-problems-srv6-based-multi-tenancy-environment/**

- **ipftrace source**

  - **https://github.com/YutaroHayakawa/ipftrace2**

And more…

- **SRv6 acceleration using XDP (jp)**
  - **https://engineering.linecorp.com/ja/blog/intern2019-report-infra/**
  - **https://www.janog.gr.jp/meeting/janog45/application/files/3815/7952/0335/009_srv6xdp_saito.pdf**

- **UDP and PMTUD support for our load balancer (jp)**
  - **https://engineering.linecorp.com/ja/blog/network-development-in-verda/**

Thank you for listening!

Twitter/Slack: @YutaroHayakawa