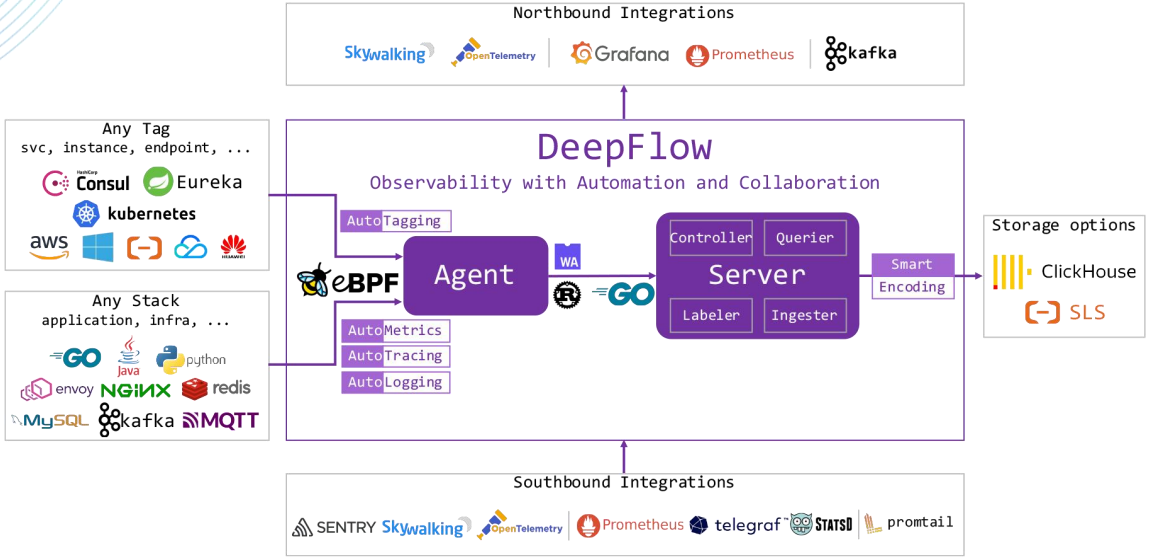
DeepFlow在Kube-OVN环境的可观测实践

宋建昌 / 云杉网络 云原生工程师

可观测性技术实践系列 第10期直播 · 2022-09-21

内容目录

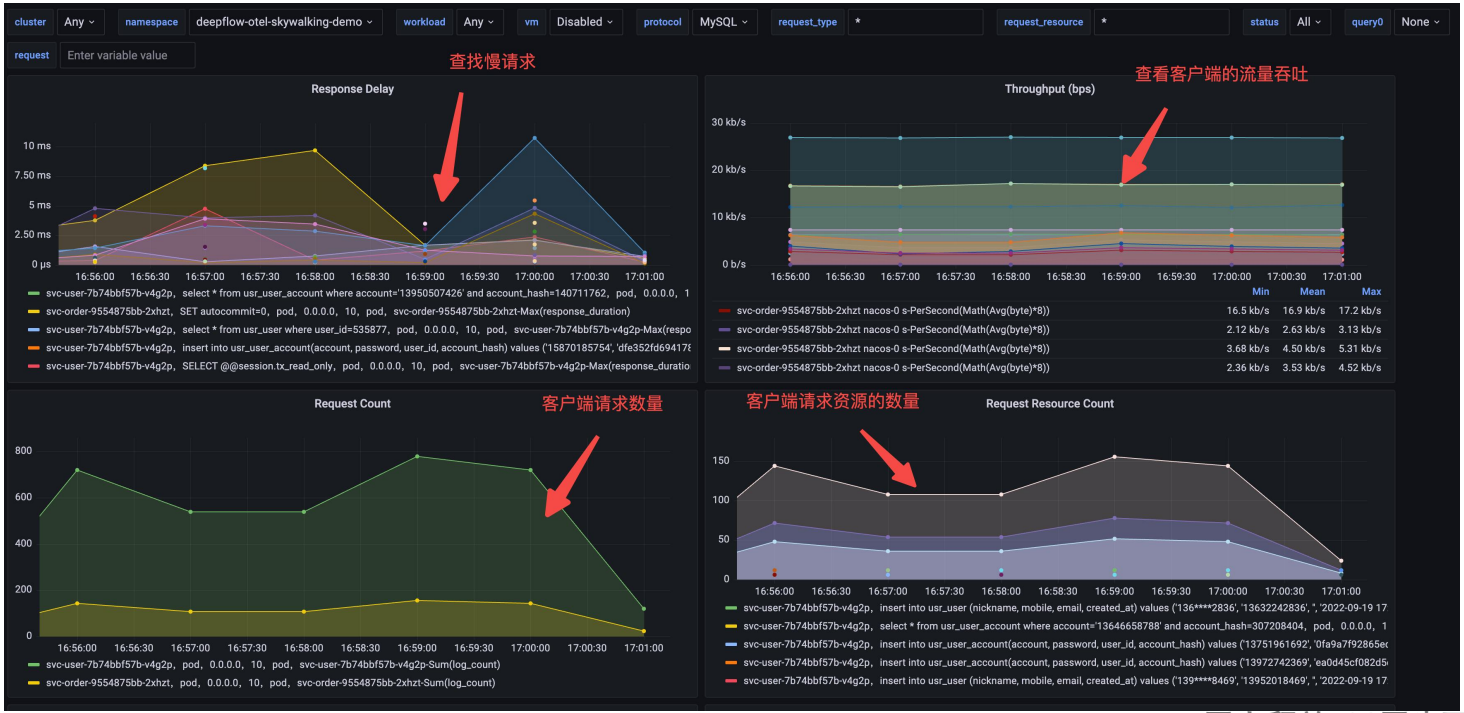
DeepFlow 高度自动化的可观测性能力



- DB流量突增：如何判断哪些服务的哪些SQL请求量变大?
- 响应变慢：如何判断延迟在应用、网络、还是数据库?
- 业务异常：如何快速判断服务端口/接口异常?

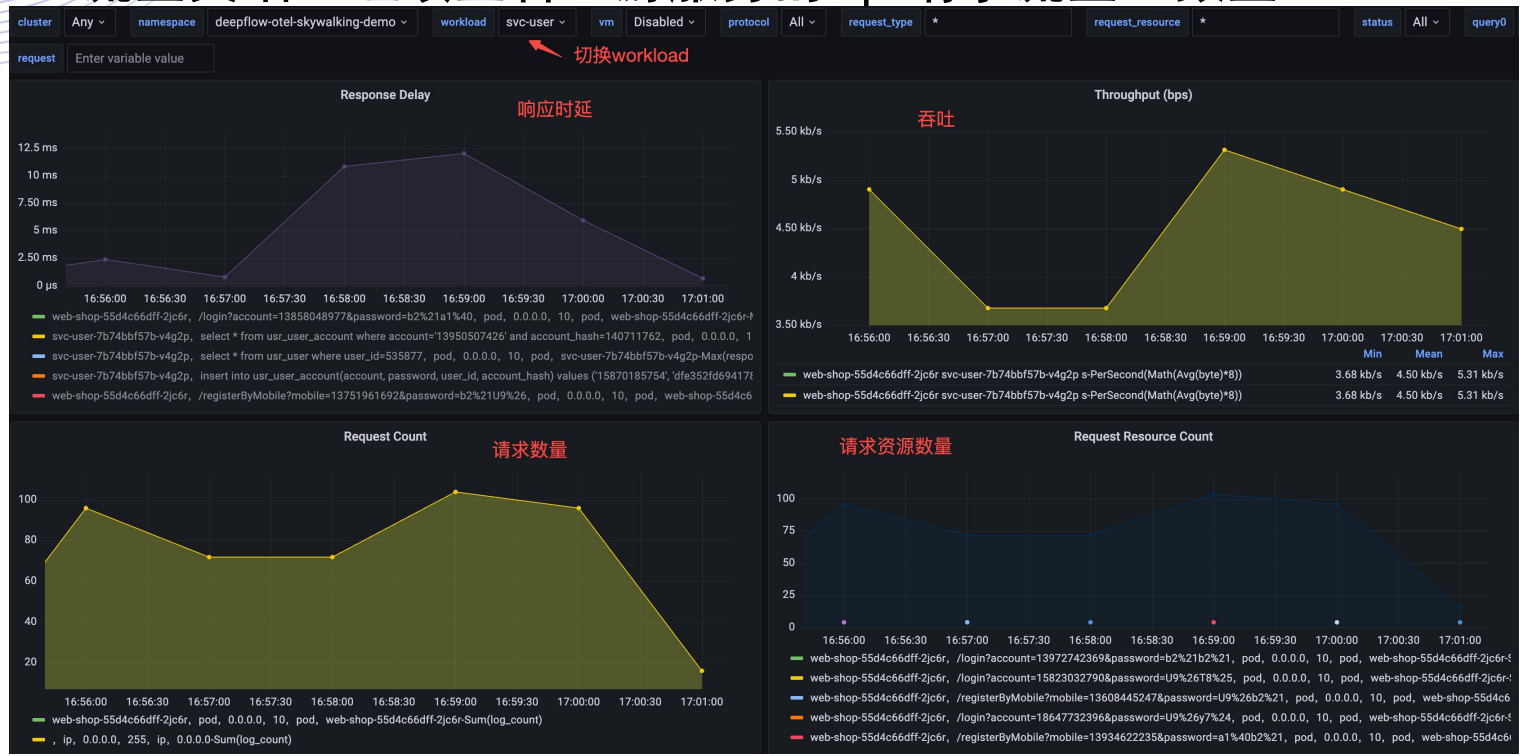
DeepFlow 在云原生场景下如何快速定位问题

DB 流量突增：如何判断哪些服务的哪些SQL请求量变大?



DeepFlow 在云原生场景下如何快速定位问题

DB 流量突增：继续查看上游服务的api请求流量、数量？

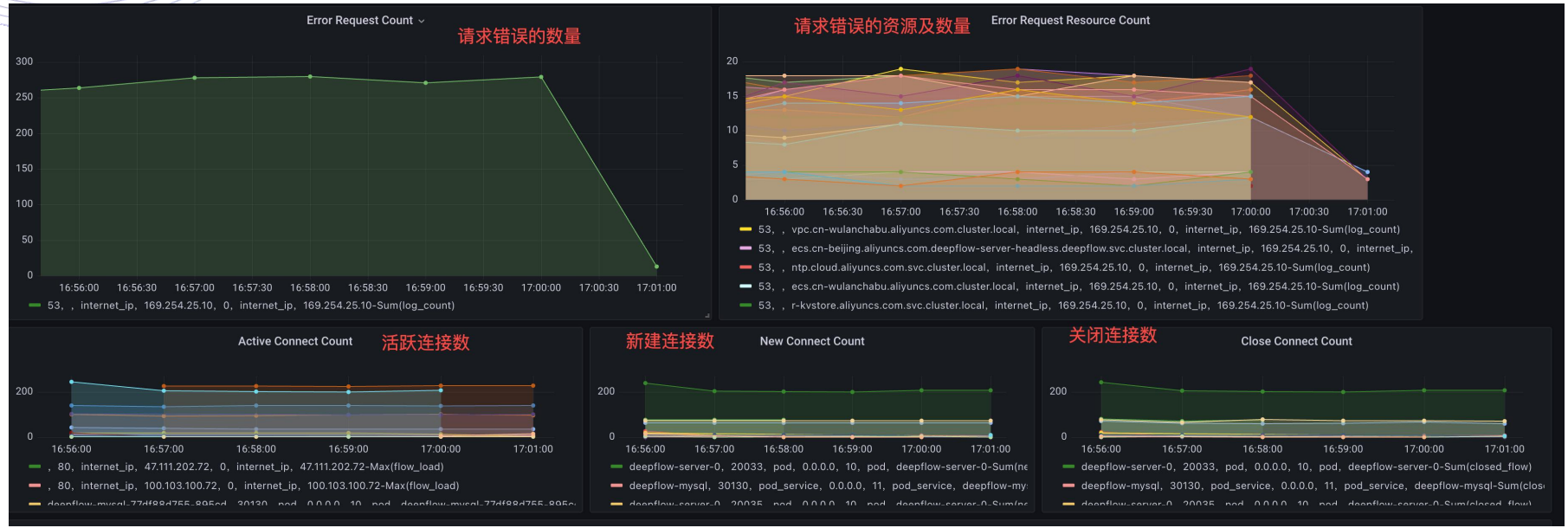


DeepFlow在云原生场景下如何快速定位问题

响应变慢：如何判断延迟在应用、网络、还是数据库？

DeepFlow在云原生场景下如何快速定位问题

业务异常：如何快速找到服务端口/接口异常?



内容目录

DeepFlow 一键开启 Kube-OVN 的可观测性

```
helm repo add deepflow https://deepflowys.github.io/deepflow
helm repo update deepflow

helm install deepflow -n deepflow deepflow/deepflow \
    --create-namespace
```

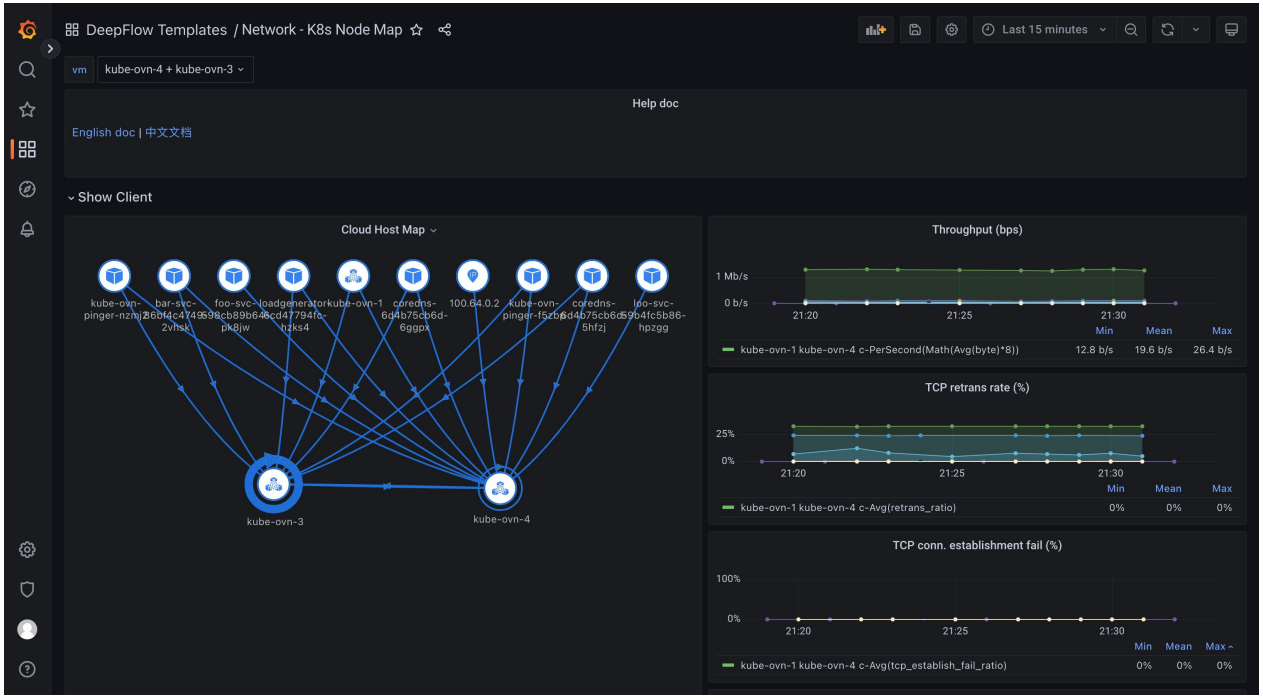
了解 DeepFlow dashboard

- **网络可观测**
  - Node/Pod的流量拓扑
  - Node/Pod流日志

- **应用可观测**
  - 服务性能总览
  - 微服务调用拓扑
  - 服务调用日志
  - 调用链追踪(tracing)

DeepFlow 一键开启 Kube-OVN 的可观测性

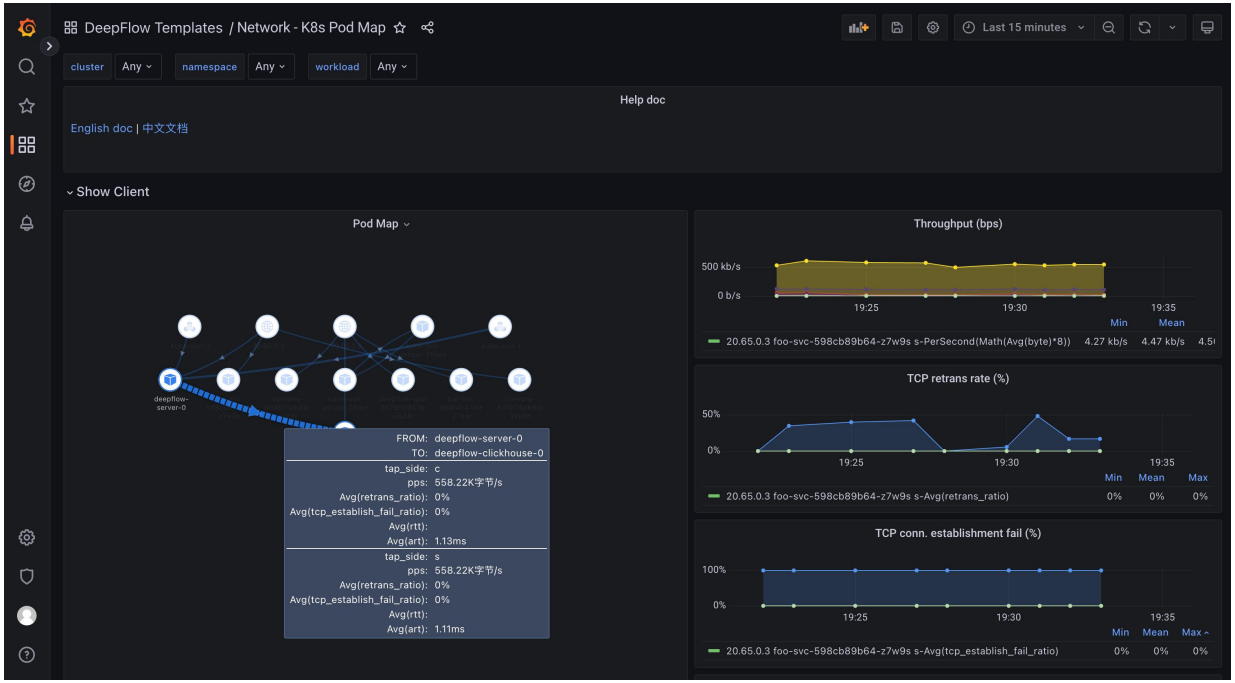
Node流量拓扑



Node/Pod流量拓扑 | Node/Pod流日志 | 微服务性能总览 | 微服务调用拓扑 | 微服务调用日志 | 调用链追踪(tracing)

DeepFlow 一键开启 Kube-OVN 的可观测性

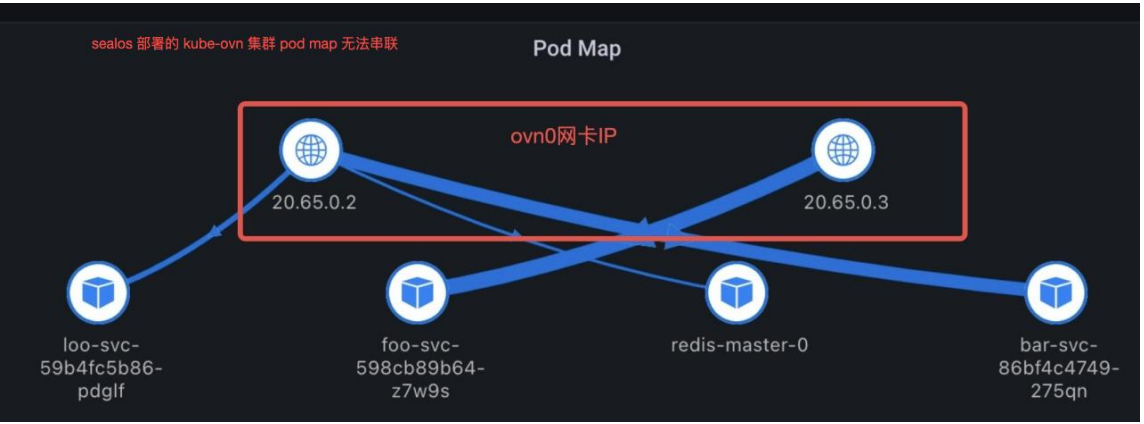
Pod流量拓扑

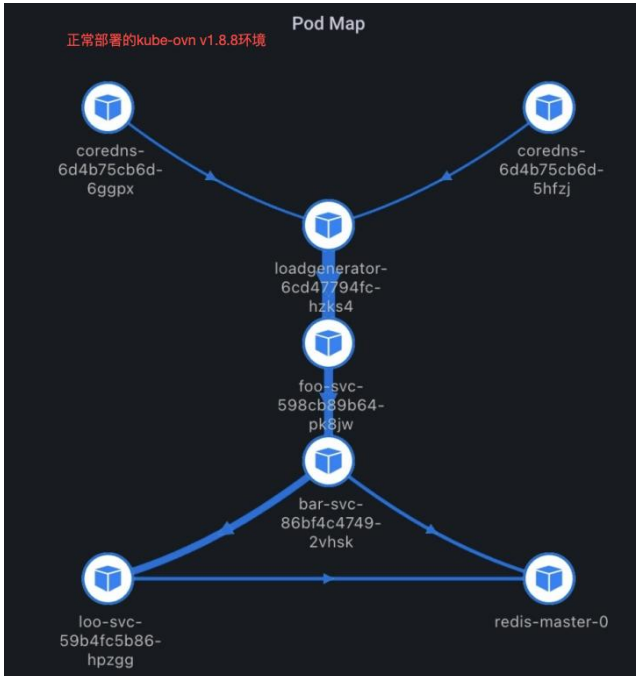


Node/Pod流量拓扑 | Node/Pod流日志 | 微服务性能总览 | 微服务调用拓扑 | 微服务调用日志 | 调用链追踪(tracing)

DeepFlow 一键开启 Kube-OVN 的可观测性

DeepFlow 在 sealos 安装的 kube-ovn 上看到的map不对？原来是 sealos 里默认把 ovn-lb 给关了，打开就 ok 了

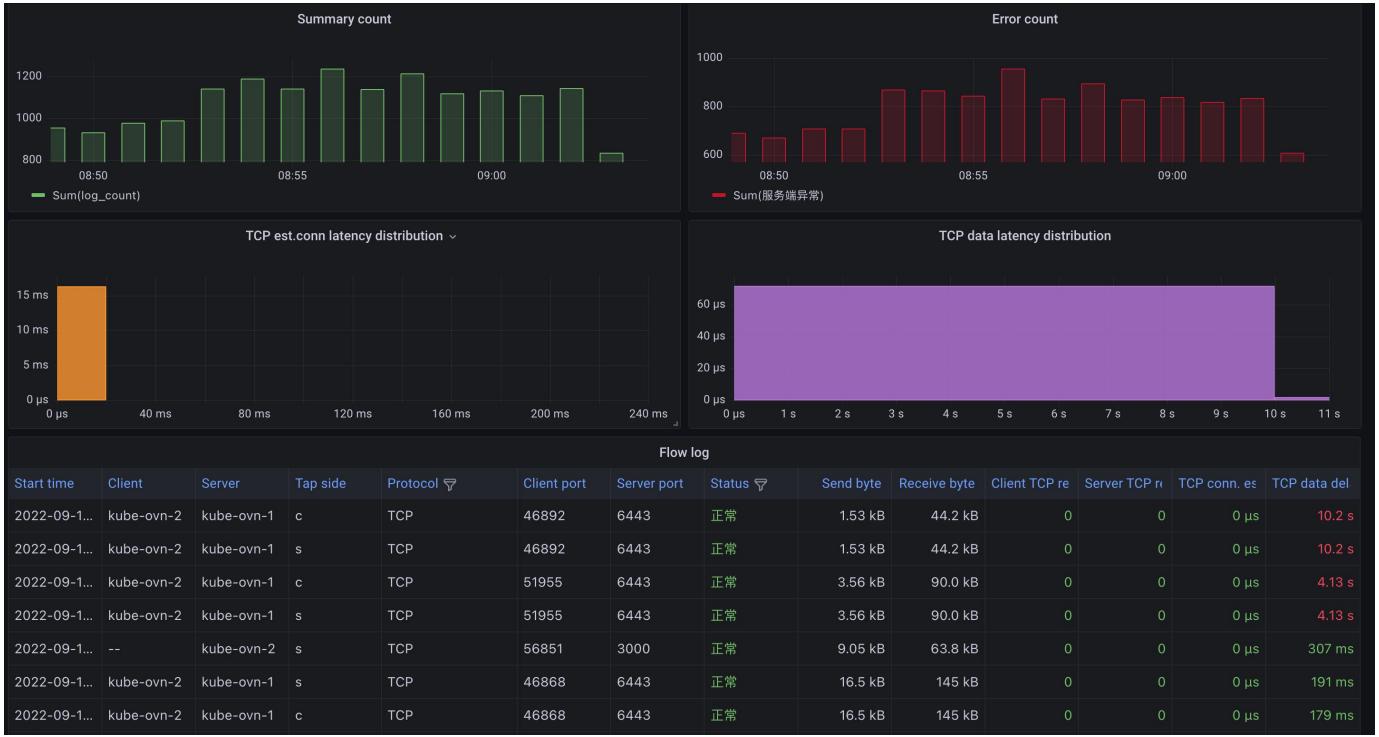




Node/Pod流量拓扑 | Node/Pod流日志 | 微服务性能总览 | 微服务调用拓扑 | 微服务调用日志 | 调用链追踪(tracing)

DeepFlow 一键开启 Kube-OVN 的可观测性

Node流日志

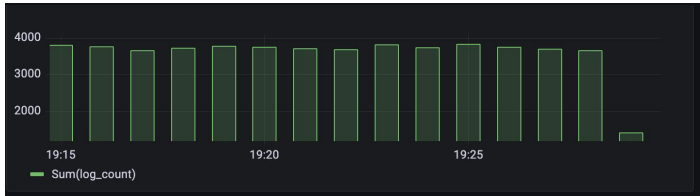


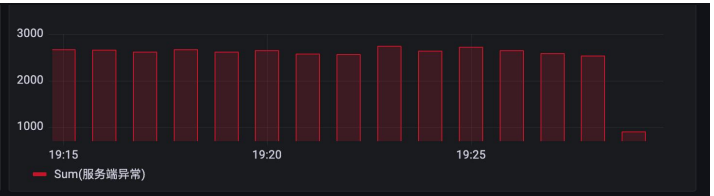
Start time	Client	Server	Tap side	Protocol	Client port	Server port	Status	Send byte	Receive byte	Client TCP re	Server TCP re	TCP conn. es	TCP data del
2022-09-1...	kube-ovn-2	kube-ovn-1	c	TCP	46892	6443	正常	1.53 kB	44.2 kB	0	0	0 µs	10.2 s
2022-09-1...	kube-ovn-2	kube-ovn-1	s	TCP	46892	6443	正常	1.53 kB	44.2 kB	0	0	0 µs	10.2 s
2022-09-1...	kube-ovn-2	kube-ovn-1	c	TCP	51955	6443	正常	3.56 kB	90.0 kB	0	0	0 µs	4.13 s
2022-09-1...	kube-ovn-2	kube-ovn-1	s	TCP	51955	6443	正常	3.56 kB	90.0 kB	0	0	0 µs	4.13 s
2022-09-1...	--	kube-ovn-2	s	TCP	56851	3000	正常	9.05 kB	63.8 kB	0	0	0 µs	307 ms
2022-09-1...	kube-ovn-2	kube-ovn-1	s	TCP	46868	6443	正常	16.5 kB	145 kB	0	0	0 µs	191 ms
2022-09-1...	kube-ovn-2	kube-ovn-1	c	TCP	46868	6443	正常	16.5 kB	145 kB	0	0	0 µs	179 ms

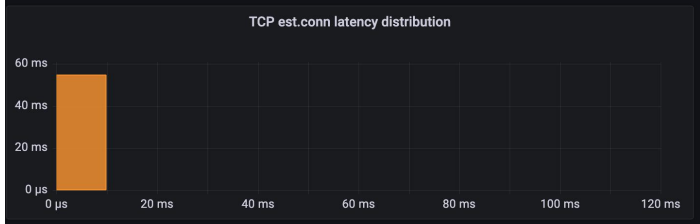
Node/Pod流量拓扑 | Node/Pod流日志 | 微服务性能总览 | 微服务调用拓扑 | 微服务调用日志 | 调用链追踪(tracing)

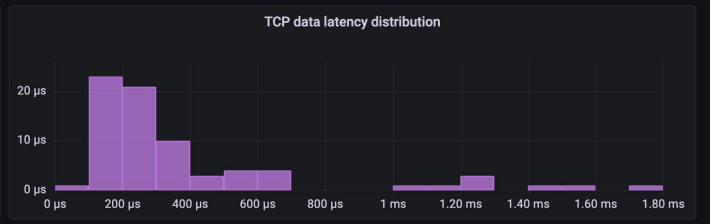
DeepFlow 一键开启 Kube-OVN 的可观测性

Pod流日志







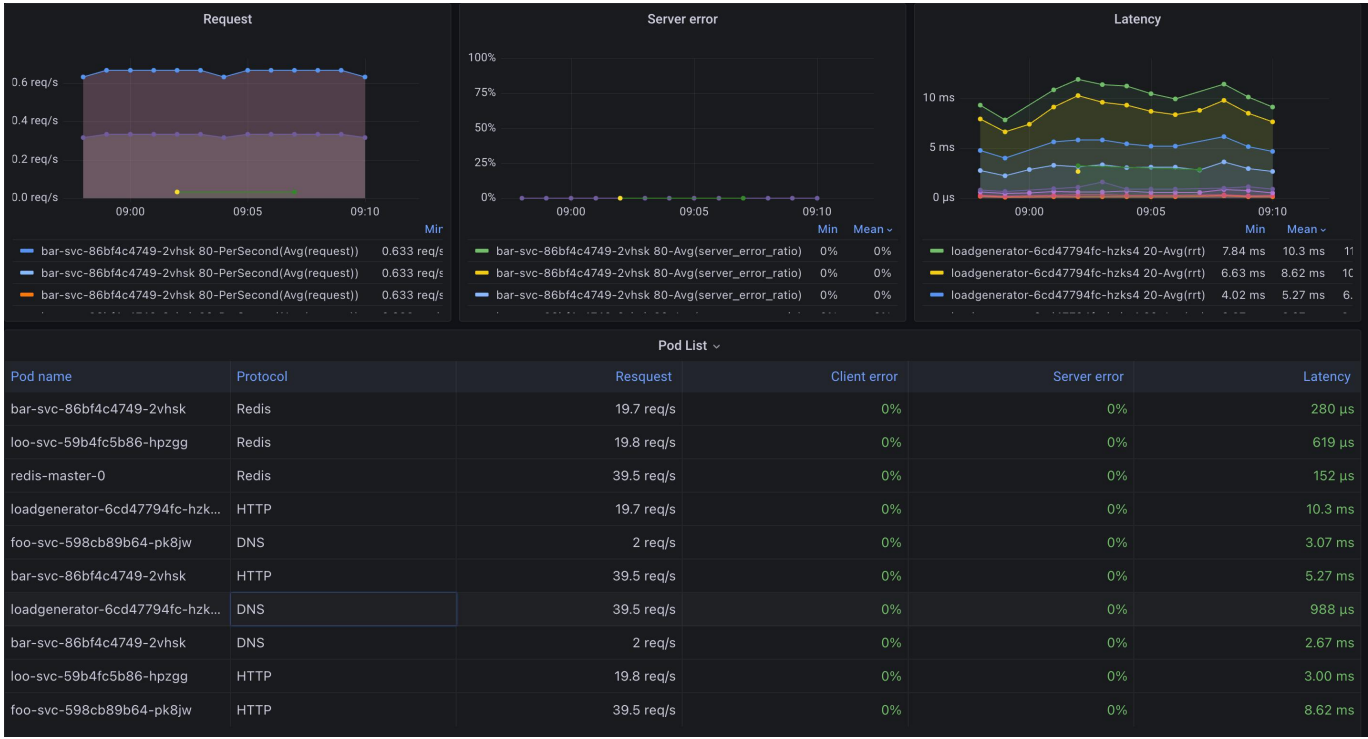


Flow log

2022-09-19 ...	loadgenerat...	productpage	c	TCP	46704	9080	正常	499 B	4.85 kB	0	0	1.17 ms	27.3 ms
2022-09-19 ...	loadgenerat...	productpag...	s-nd	TCP	46704	9080	正常	499 B	4.99 kB	0	0	1.49 ms	26.4 ms
2022-09-19 ...	loadgenerat...	productpag...	s	TCP	46704	9080	正常	499 B	4.85 kB	0	0	1.52 ms	26.2 ms
2022-09-19 ...	loadgenerat...	coredns-6d...	s	UDP	48877	53	正常	727 B	5.25 kB	0	0	1.05 ms	14.5 ms
2022-09-19 ...	loadgenerat...	--	c	TCP	59508	80	正常	479 B	1.21 kB	0	0	741 µs	9.43 ms
2022-09-19 ...	loadgenerat...	foo-svc-598...	s-nd	TCP	59508	8080	正常	479 B	1.21 kB	0	0	553 µs	9.12 ms
2022-09-19 ...	loadgenerat...	foo-svc-598...	s	TCP	59508	8080	正常	479 B	1.21 kB	0	0	552 µs	8.90 ms
2022-09-19 ...	loadgenerat...	--	c	TCP	57392	80	正常	479 B	1.21 kB	0	0	760 µs	8.46 ms

DeepFlow 一键开启 Kube-OVN 的可观测性

服务性能总览



Pod List

Pod name	Protocol	Resquest	Client error	Server error	Latency
bar-svc-86bf4c4749-2vhsk	Redis	19.7 req/s	0%	0%	280 µs
loo-svc-59b4fc5b86-hpzgg	Redis	19.8 req/s	0%	0%	619 µs
redis-master-0	Redis	39.5 req/s	0%	0%	152 µs
loadgenerator-6cd47794fc-hzk...	HTTP	19.7 req/s	0%	0%	10.3 ms
foo-svc-598cb89b64-pk8jw	DNS	2 req/s	0%	0%	3.07 ms
bar-svc-86bf4c4749-2vhsk	HTTP	39.5 req/s	0%	0%	5.27 ms
loadgenerator-6cd47794fc-hzk...	DNS	39.5 req/s	0%	0%	988 µs
bar-svc-86bf4c4749-2vhsk	DNS	2 req/s	0%	0%	2.67 ms
loo-svc-59b4fc5b86-hpzgg	HTTP	19.8 req/s	0%	0%	3.00 ms
foo-svc-598cb89b64-pk8jw	HTTP	39.5 req/s	0%	0%	8.62 ms

Node/Pod流量拓扑 | Node/Pod流日志 | 微服务性能总览 | 微服务调用拓扑 | 微服务调用日志 | 调用链追踪(tracing)

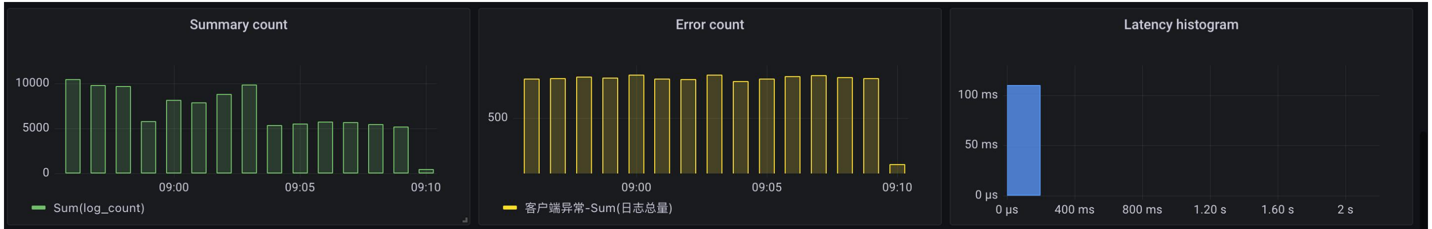
DeepFlow 一键开启 Kube-OVN 的可观测性

微服务调用拓扑

Node/Pod流量拓扑 | Node/Pod流日志 | 微服务性能总览 | 微服务调用拓扑 | 微服务调用日志 | 调用链追踪(tracing)

DeepFlow 一键开启 Kube-OVN 的可观测性

服务调用日志



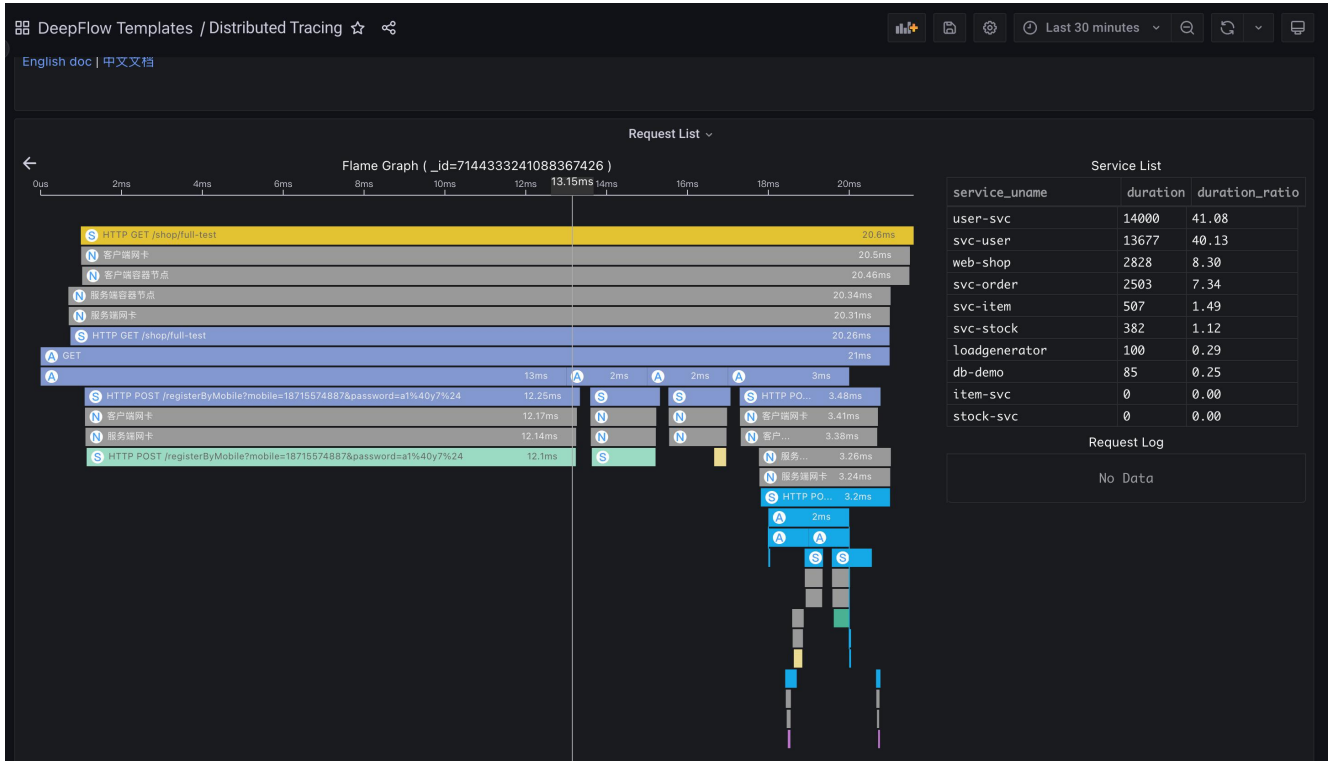
Request log

Start time	Client	Server	Protocol	Request type	Request domain	Request resource	Response Status	Resonse Code
2022-09-18 08:59...	loo-svc-59b4fc5b...	redis-master-0	Redis	GET	--	GET loo	未知	--
2022-09-18 08:59...	loo-svc-59b4fc5b...	redis-master-0	Redis	GET	--	GET loo	未知	--
2022-09-18 08:59...	loo-svc-59b4fc5b...	redis-master-0	Redis	GET	--	GET loo	未知	--
2022-09-18 08:59...	loo-svc-59b4fc5b...	redis-master-0	Redis	GET	--	GET loo	未知	--
2022-09-18 08:59...	loo-svc-59b4fc5b...	redis-master-0	Redis	GET	--	GET loo	未知	--
2022-09-18 08:59...	loo-svc-59b4fc5b...	redis-master-0	Redis	GET	--	GET loo	未知	--
2022-09-18 08:58...	loo-svc-59b4fc5b...	redis-master-0	Redis	GET	--	GET loo	未知	--
2022-09-18 08:58...	loo-svc-59b4fc5b...	redis-master-0	Redis	GET	--	GET loo	未知	--
2022-09-18 08:58...	loo-svc-59b4fc5b...	redis-master-0	Redis	GET	--	GET loo	未知	--
2022-09-18 08:58...	loo-svc-59b4fc5b...	redis-master-0	Redis	GET	--	GET loo	未知	--
2022-09-18 08:59...	--	redis-master-0	Redis	GET	--	GET bar	未知	--
2022-09-18 08:59...	--	redis-master-0	Redis	GET	--	GET bar	未知	--
2022-09-18 08:59...	--	redis-master-0	Redis	GET	--	GET bar	未知	--

Node/Pod流量拓扑 | Node/Pod流日志 | 微服务性能总览 | 微服务调用拓扑 | **微服务调用日志** | 调用链追踪(tracing)

DeepFlow 一键开启 Kube-OVN 的可观测性

调用链追踪(tracing)

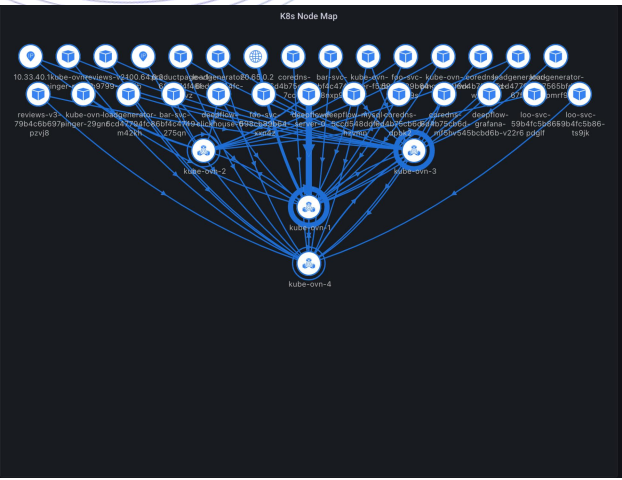


Node/Pod流量拓扑 | Node/Pod流日志 | 微服务性能总览 | 微服务调用拓扑 | 微服务调用日志 | 调用链追踪(tracing)

内容目录

网络通信场景追踪

DeepFlow增强Kube-OVN diagnose工具的观测能力




```
7] kube-ovn-3: exporter connect successfully
s-vswitchd and ovsdb are up
n_controller is up
art to check port binding
hassis id is 6e71c8d2-69b0-47d3-9118-168ca3ad1b50
rt in sb is [reviews-v2-c44db9799-qqp5b.deepflow-ebpf-istio-demo coredns-6d4b75cb6
rnal kube-ovn-pinger-nzmj2.kube-system loo-svc-59b4fc5b86-ts9jk.deepflow-ebpf-spri
w-ebpf-istio-demo productpage-v1-684874f46f-746vz.deepflow-ebpf-istio-demo node-ku
ebpf-spring-demo bar-svc-86bf4c4749-8nxp9.deepflow-ebpf-spring-demo]
s and ovn-sb binding check passed
start to check apiserver connectivity
connect to apiserver success in 3.04ms
I0919 21:01:48.541837    3405 ping.go:129] start to check pod connectivity
I0919 21:01:48.727805    3405 ping.go:159] ping pod: kube-ovn-pinger-f5zbp 10.16.0.6, count: 3, loss count 0, average rtt 22.
56ms
I0919 21:01:48.831262    3405 ping.go:159] ping pod: kube-ovn-pinger-nzmj2 10.16.0.7, count: 3, loss count 0, average rtt 0.1
9ms
I0919 21:01:48.831412    3405 ping.go:83] start to check node connectivity
I0919 21:01:49.140021    3405 ping.go:108] ping node: kube-ovn-3 10.33.40.21, count: 3, loss count 0, average rtt 0.35ms
I0919 21:01:49.442895    3405 ping.go:108] ping node: kube-ovn-4 10.33.40.22, count: 3, loss count 0, average rtt 0.86ms
I0919 21:01:49.443068    3405 ping.go:223] start to check dns connectivity
I0919 21:01:49.526744    3405 ping.go:236] resolve dns kubernetes.default to [10.96.0.1] in 83.58ms
```

网络通信场景追踪

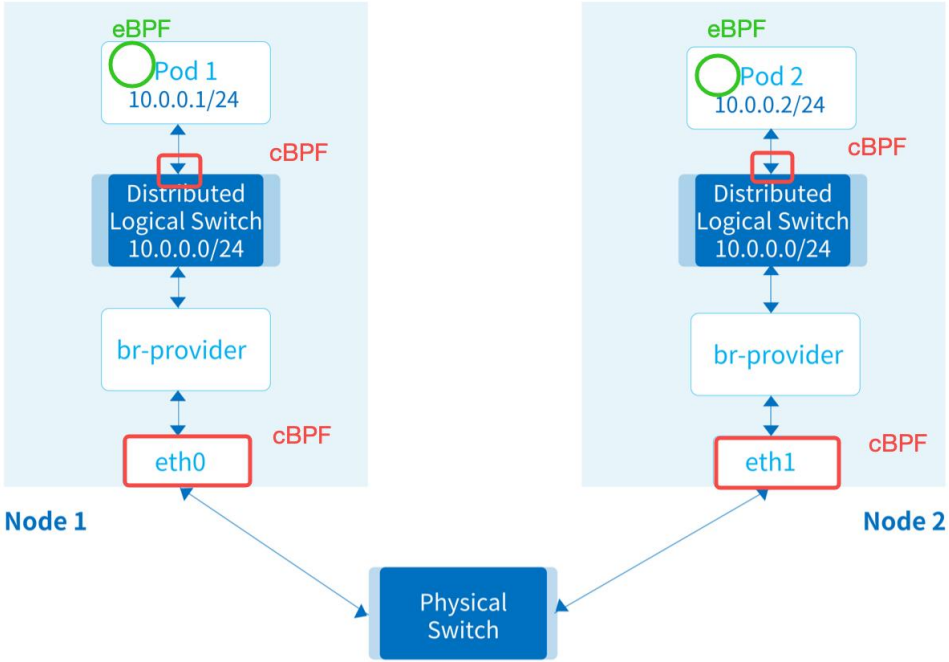
DeepFlow 增强 Kube-OVN 的观测能力

	DeepFlow
可观测范围	客户端应用（智能汽车）
	服务端应用 （CT 服务：5G 核心网）
	服务端应用 (IT 服务)
	网络中间件(L4/L7网关、服务网格)
	存储中间件（数据库、消息队列）
	容器 (K8s、 OpenShift、Serverless)
	虚拟化 (KVM、ESXi、Hyperv、Xen)
	物理网络
告警	Grafana告警（6.1.4支持） 企业版本告警
数据采集、存储、查询	支持

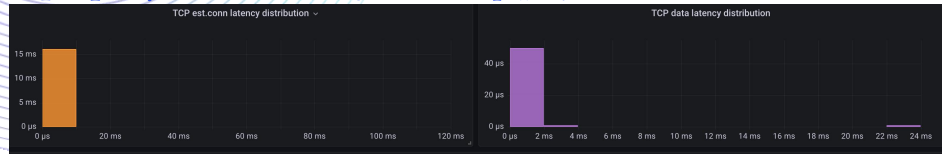
网络通信场景追踪

DeepFlow在kube-ovn的同子网、跨子网、跨VPC、双栈等场景的实际应用

- 以同子网跨节点场景为例
  - 流量采集位置
  - 隧道解封装
  - 跨Node的流量关联
  - 流量关联Pod、Node等资源



自动生成网络流日志



Flow log

Start time	Client	Server	Tap side	Protocol	Client port	Server port	Status	Send byte	Receive byte	Client TCP re	Server TCP re	TCP conn. es	TCP data del
2022-09-1...	kube-ovn-2	kube-ovn-1	c	TCP	46892	6443	正常	1.82 kB	37.6 kB	0	0	0 µs	5.86 s
2022-09-1...	kube-ovn-2	kube-ovn-1	s	TCP	46892	6443	正常	1.82 kB	37.6 kB	0	0	0 µs	5.86 s
2022-09-1...	kube-ovn-2	kube-ovn-1	c	TCP	51955	6443	正常	3.10 kB	82.6 kB	0	0	0 µs	3.90 s
2022-09-1...	kube-ovn-2	kube-ovn-1	s	TCP	51955	6443	正常	3.10 kB	82.6 kB	0	0	0 µs	3.90 s
2022-09-1...	kube-ovn-2	kube-ovn-1	c	TCP	44066	6443	正常	40.6 kB	186 kB	1	0	0 µs	73.1 ms
2022-09-1...	kube-ovn-2	kube-ovn-1	s	TCP	44066	6443	正常	40.6 kB	186 kB	1	0	0 µs	73.0 ms
2022-09-1...	kube-ovn-2	kube-ovn-1	c	TCP	52552	6443	正常	3.17 kB	6.23 kB	0	0	0 µs	25.2 ms
2022-09-1...	kube-ovn-2	kube-ovn-1	s	TCP	52552	6443	正常	3.17 kB	6.23 kB	0	0	0 µs	25.1 ms
2022-09-1...	--	kube-ovn-2	s	UDP	56237	53	服务端异常	0 B	0 B	0	0	0 µs	23.5 ms
2022-09-1...	kube-ovn-2	--	c	UDP	53219	53	服务端异常	0 B	0 B	0	0	0 µs	5.38 ms
2022-09-1...	kube-ovn-2	--	c	UDP	53219	53	服务端异常	0 B	0 B	0	0	0 µs	5.34 ms
2022-09-1...	kube-ovn-2	kube-ovn-1	c	TCP	33008	6443	正常	38.6 kB	68.7 kB	0	0	0 µs	3.30 ms
2022-09-1...	kube-ovn-2	kube-ovn-1	s	TCP	33008	6443	正常	38.6 kB	68.7 kB	0	0	0 µs	3.18 ms

```
53Z|00004|acl_log(ovn_pinctrl0)|INFO|name="<unnamed>", verdict=drop, severity=warning: icmp,vl
00:00:00:a5:30:25,dl_dst=00:00:00:8e:41:c5,nw_src=10.16.0.6,nw_dst=10.16.0.7,nw_tos=0,nw_ecn=0
8,icmp_code=0
53Z|00005|acl_log(ovn_pinctrl0)|INFO|name="<unnamed>", verdict=drop, severity=warning: icmp,vl
00:00:00:a5:30:25,dl_dst=00:00:00:8e:41:c5,nw_src=10.16.0.6,nw_dst=10.16.0.7,nw_tos=0,nw_ecn=0
8,icmp_code=0
53Z|00006|acl_log(ovn_pinctrl0)|INFO|name="<unnamed>", verdict=drop, severity=warning: icmp,vl
00:00:00:a5:30:25,dl_dst=00:00:00:8e:41:c5,nw_src=10.16.0.6,nw_dst=10.16.0.7,nw_tos=0,nw_ecn=0
8,icmp_code=0
74Z|00007|acl_log(ovn_pinctrl0)|INFO|name="<unnamed>", verdict=drop, severity=warning: udp,vla
0:00:00:8e:41:c5,dl_dst=00:00:00:cb:a2:e2,nw_src=10.16.0.7,nw_dst=10.16.0.4,nw_tos=0,nw_ecn=0,
nw_ttl=64,tp_src=35880,tp_dst=53
2022-09-18T01:19:55.475Z|00008|acl_log(ovn_pinctrl0)|INFO|name="<unnamed>", verdict=drop, severity=warning: udp,vla
n_tci=0x0000,dl_src=00:00:00:8e:41:c5,dl_dst=00:00:00:cb:a2:e2,nw_src=10.16.0.7,nw_dst=10.16.0.5,nw_tos=0,nw_ecn=0,
nw_ttl=64,tp_src=40441,tp_dst=53
2022-09-18T01:20:00.475Z|00009|acl_log(ovn_pinctrl0)|INFO|name="<unnamed>", verdict=drop, severity=warning: udp,vla
n_tci=0x0000,dl_src=00:00:00:8e:41:c5,dl_dst=00:00:00:cb:a2:e2,nw_src=10.16.0.7,nw_dst=10.16.0.4,nw_tos=0,nw_ecn=0,
nw_ttl=64,tp_src=45968,tp_dst=53
2022-09-18T01:20:00.475Z|00010|acl_log(ovn_pinctrl0)|INFO|name="<unnamed>", verdict=drop, severity=warning: udp,vla
n_tci=0x0000,dl_src=00:00:00:8e:41:c5,dl_dst=00:00:00:cb:a2:e2,nw_src=10.16.0.7,nw_dst=10.16.0.5,nw_tos=0,nw_ecn=0,
nw_ttl=64,tp_src=50392,tp_dst=53
2022-09-18T01:20:05.476Z|00011|acl_log(ovn_pinctrl0)|INFO|name="<unnamed>", verdict=drop, severity=warning: udp,vla
n_tci=0x0000,dl_src=00:00:00:8e:41:c5,dl_dst=00:00:00:cb:a2:e2,nw_src=10.16.0.7,nw_dst=10.16.0.4,nw_tos=0,nw_ecn=0,
nw_ttl=64,tp_src=42886,tp_dst=53
2022-09-18T01:20:05.476Z|00012|acl_log(ovn_pinctrl0)|INFO|name="<unnamed>", verdict=drop, severity=warning: udp,vla
n_tci=0x0000,dl_src=00:00:00:8e:41:c5,dl_dst=00:00:00:cb:a2:e2,nw_src=10.16.0.7,nw_dst=10.16.0.4,nw_tos=0,nw_ecn=0,
nw_ttl=64,tp_src=46931,tp_dst=53
```

自动生成网络流日志

差异点	DeepFlow	kube-ovn networkpolicy 日志
额外cpu开销	无	有
捕获位置	VPC网络、容器网络、物理网络、系统接口	未涉及
隧道-TAG	支持（类型、外层IP/MAC等）	未涉及
VPC网络-TAG	支持（VPC/VM/RDS/LB/NAT约20种）	未涉及
容器网络-TAG	支持（cluster/ns/workload/service/pod约10种）	未涉及
采集位置-TAG	支持（采集点/采集器/网卡/MAC等）	未涉及
指标量-吞吐	支持（packet/byte/网络层载荷/传输层载荷等）	未涉及
指标量-性能	支持（零窗/重传等）	未涉及
指标量-时延	支持（建连/传输/协议栈等）	未涉及
流状态	支持(排期中)	支持

流量分发

DeepFlow 对 Kube-OVN 流量分发能力的增强

全局流量镜像配置

流量镜像功能默认为关闭状态，如果需要开启请修改 `kube-ovn-cni` DaemonSet 的启动参数：

- `--enable-mirror=true`： 是否开启流量镜像。
- `--mirror-iface=mirror0`：流量镜像所复制到的网卡名。该网卡可为主机上已存在的一块物理网卡， 此时该网卡会被桥接进 br-int 网桥，镜像流量会直接接入底层交换机。若网卡名不存在，Kube-OVN 会自动 创建一块同名的虚拟网卡，管理员或开发者可以在宿主机上通过该网卡获取当前节点所有流量。默认为 `mirror0` 。

接下来可以用 tcpdump 或其他流量分析工具监听 `mirror0` 上的流量：

```
tcpdump -ni mirror0
```

Pod 级别流量镜像配置

如果只需对部分 Pod 流量进行镜像，则需要关闭全局的流量镜像功能，然后在特定 Pod 上增加 `ovn.kubernetes.io/mirror` annotation 来开启 Pod 级别流量镜像。

```yaml
apiVersion: v1
kind: Pod
metadata:
  name: mirror-pod
  namespace: ls1
  annotations:
    ovn.kubernetes.io/mirror: "true"
spec:
  containers:
  - name: mirror-pod
    image: nginx:alpine
```

享会

流量分发

DeepFlow 对 Kube-OVN 流量分发能力的增强

- 面向业务的流量过滤
- 源端Payload截断
- 多层流量标签
- 流量全局去重
- 流量多路分发
- 资源变更感知

Cloud

试用 **SaaS** 服务

Community

访问 **GitHub** 仓库

Enterprise

DeepFlow 咨询

https://github.com/deepflowys/deepflow

我们在招聘，HR 微信：holidayxd

QA互动问答
扫码提问

可观测性技术实践直播互...

扫一扫二维码打开或分享给好友