Network Monitoring with eBPF, Vector and ClickHouse

Who Are We?

Matt Franklin
Senior Production Engineering Manager, Observability Team

Who Are We?

Sebastian Rabenhorst
Senior Production Engineer, Observability Team

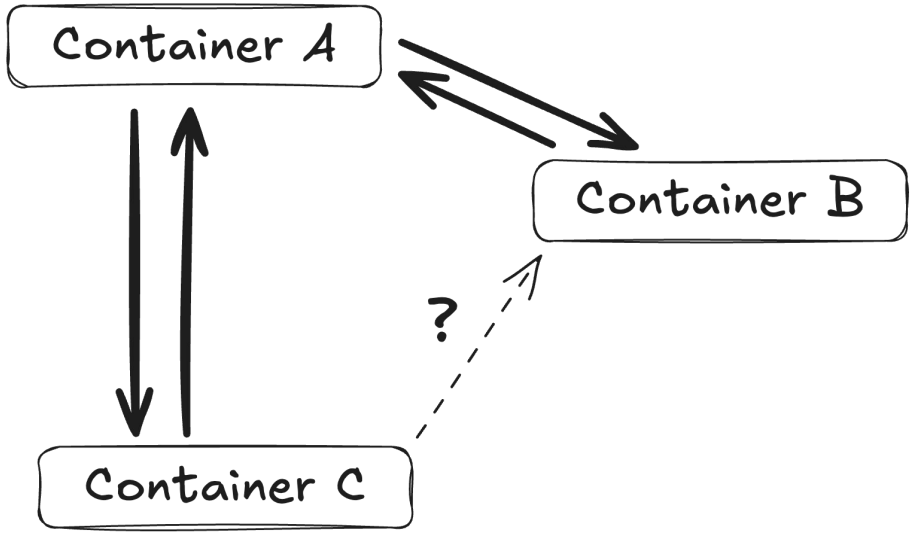
Shopify Observability

- **Observe**: Shopify's Observability platform
- **Investigate**: Search & Analytics on event data within Observe

Context

Why build a network monitoring system?

- Restore a capability lost in the migration from our previous metrics vendor
- Every incident starts with “Is it the network?”



Hundreds

Clusters

Millions

Containers

Millions

DNS Queries/s

Tens of Millions

Connections/s

Requirements

01

Replace previous network monitoring solution

02

Debug network flows and DNS queries (filter/group by k8s metadata)

03

Minimal impact on production nodes

04

Ensure scalability and extensibility

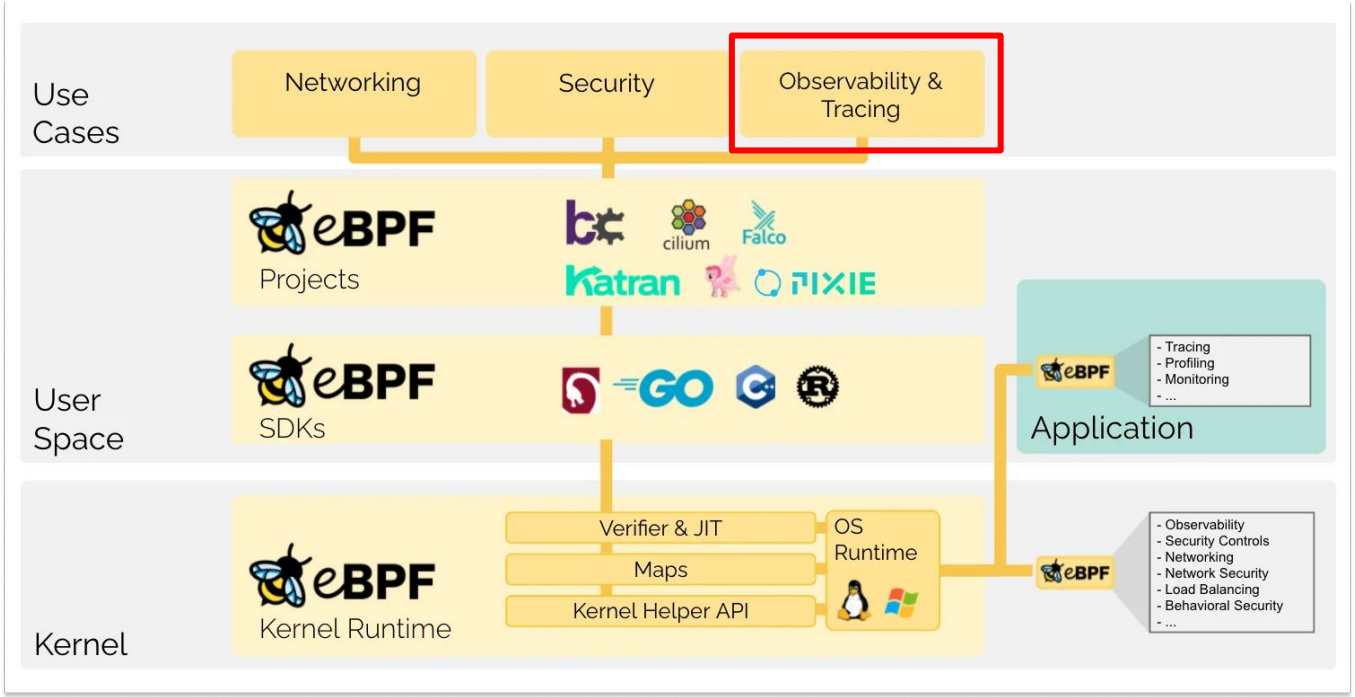
05

Handle high volumes of network traffic (TiB/min)

Pipeline Architecture

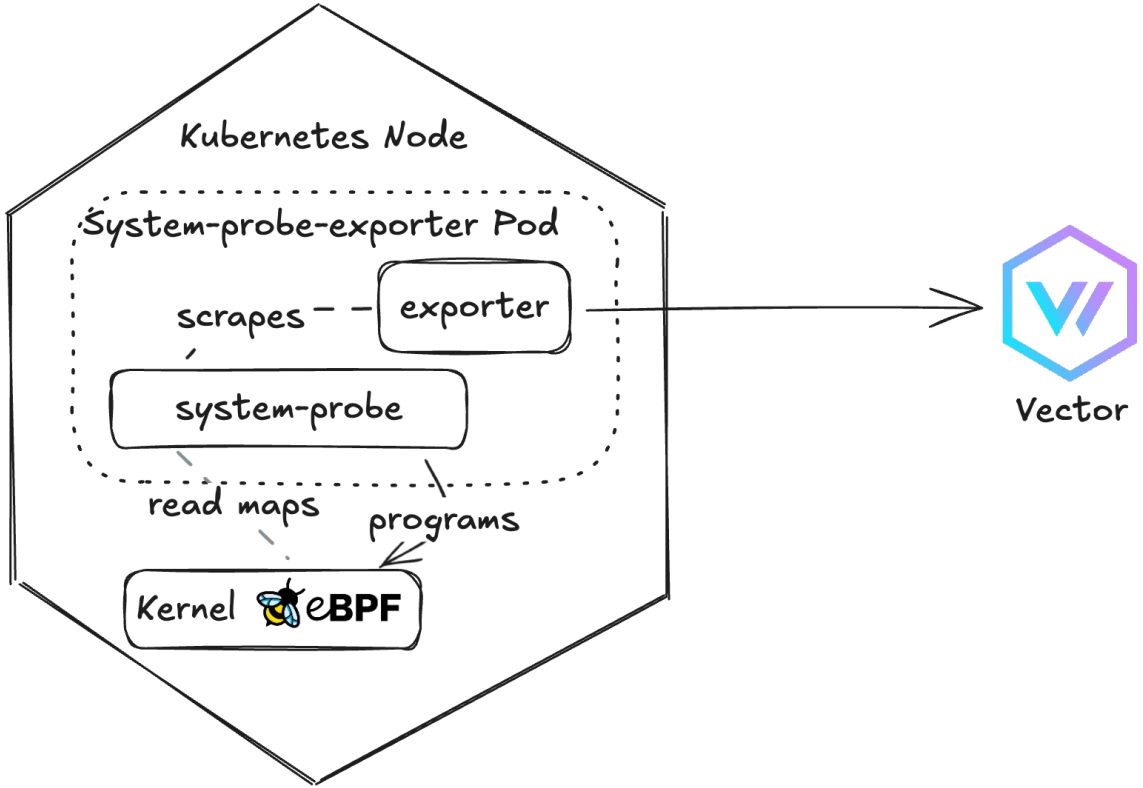
- Capture network traffic statistics from our k8s fleet.
- Low resource consumption.

eBPF



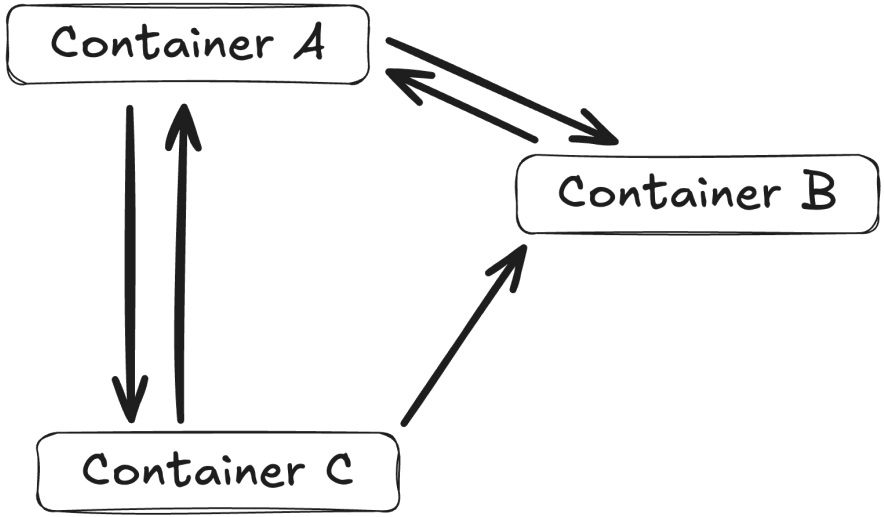
From https://ebpf.io/what-is-ebpf/

eBPF Exporter and Exporting Network/DNS Events

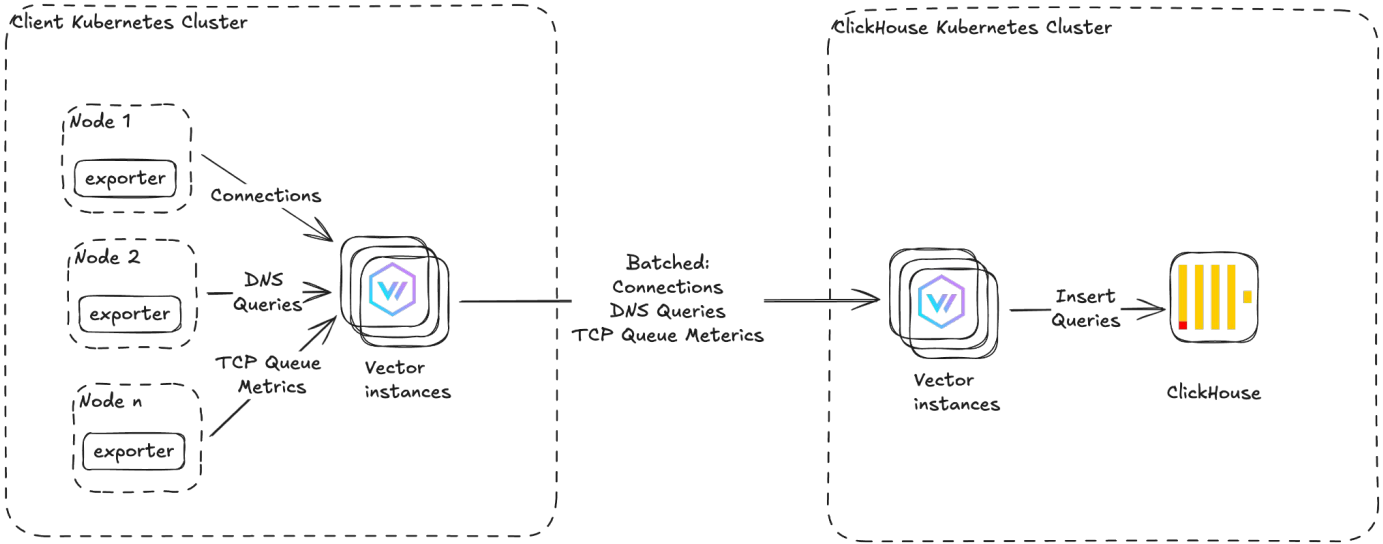


Connections

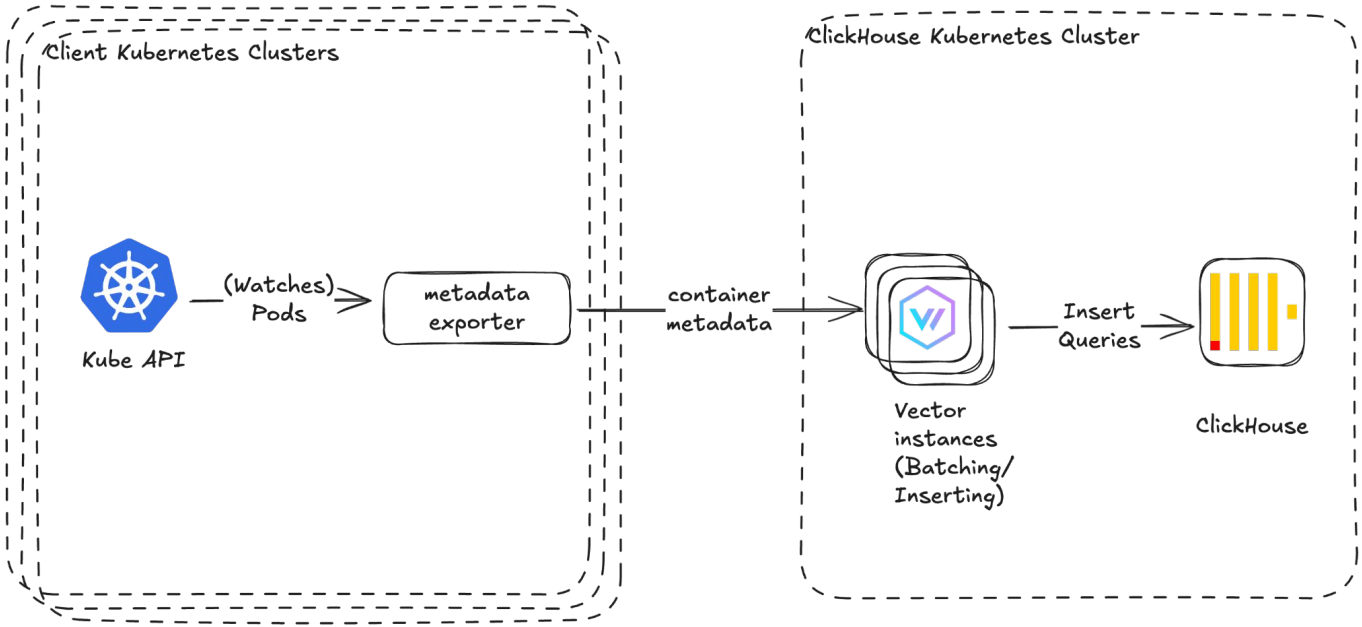
- PID
- **Local Address** (IP, Port, Container ID)
- **Remote Address** (IP, Port, Container ID)
- Type
- Bytes TX/RX
- Packets TX/RX
- Direction



Vector for Forwarding and Aggregation



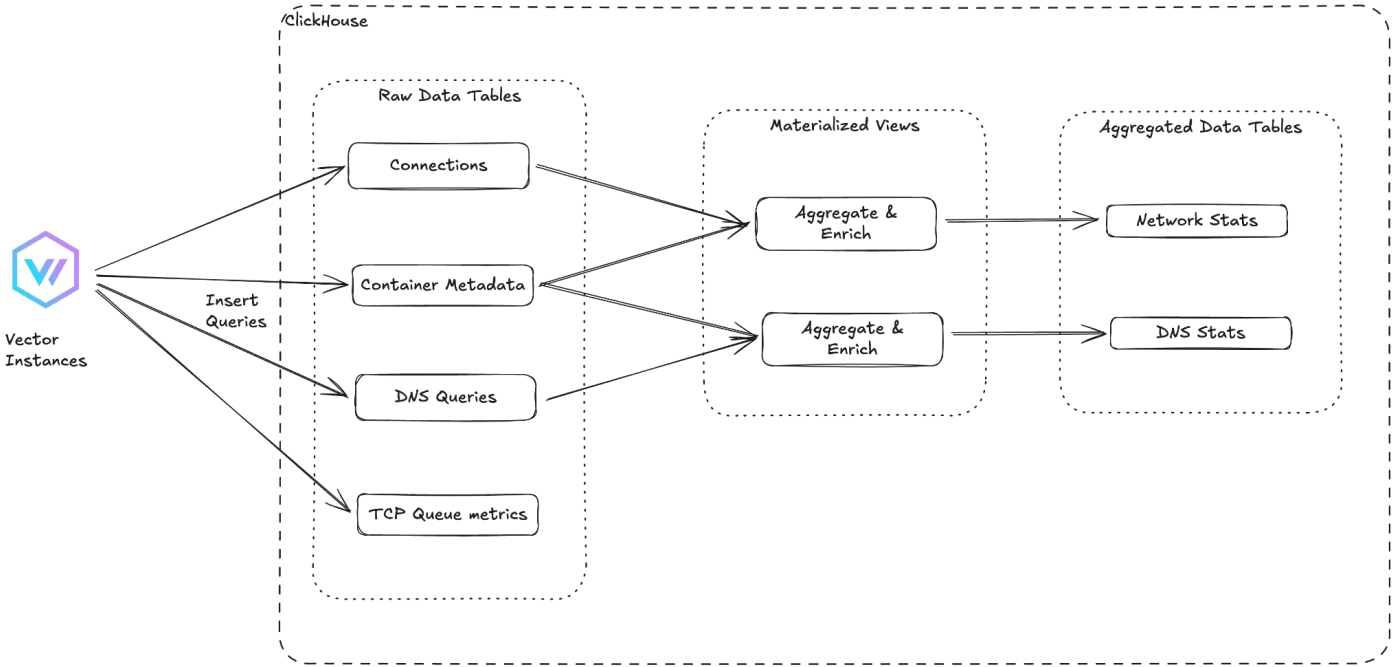
Metadata Pipeline

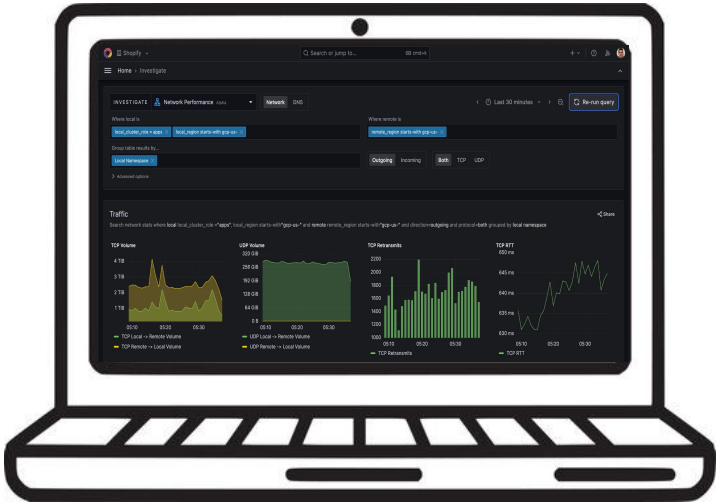


Metadata

```
message ContainerMetadata {
  string timestamp = 1;
  string network = 2;
  string project = 3;
  string region = 4;
  string zone = 5;
  string cluster = 6;
  string cluster_role = 7;
  string environment = 8;
  string namespace = 9;
  string node = 10;
  string deployment = 11;
  string pod_id = 12;
  string pod = 13;
  string container_id = 14;
  string container = 15;
  string hostname = 16;
  string ip = 17;
  bool host_network = 18;
  string pod_uid = 19;
}
```

ClickHouse

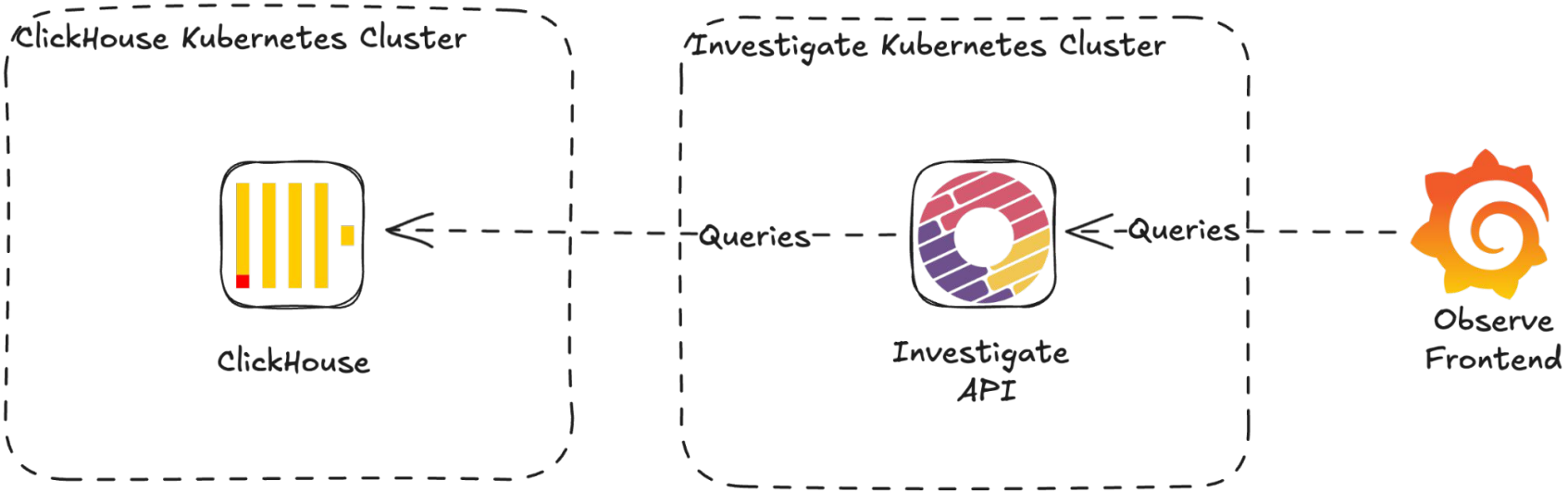

ClickHouse
Raw Data Tables
Materialized Views
Aggregated Data Tables
Connections
Container Metadata
DNS Queries
TCP Queue metrics
Aggregate & Enrich
Aggregate & Enrich
Network Stats
DNS Stats
Insert Queries
Vector Instances



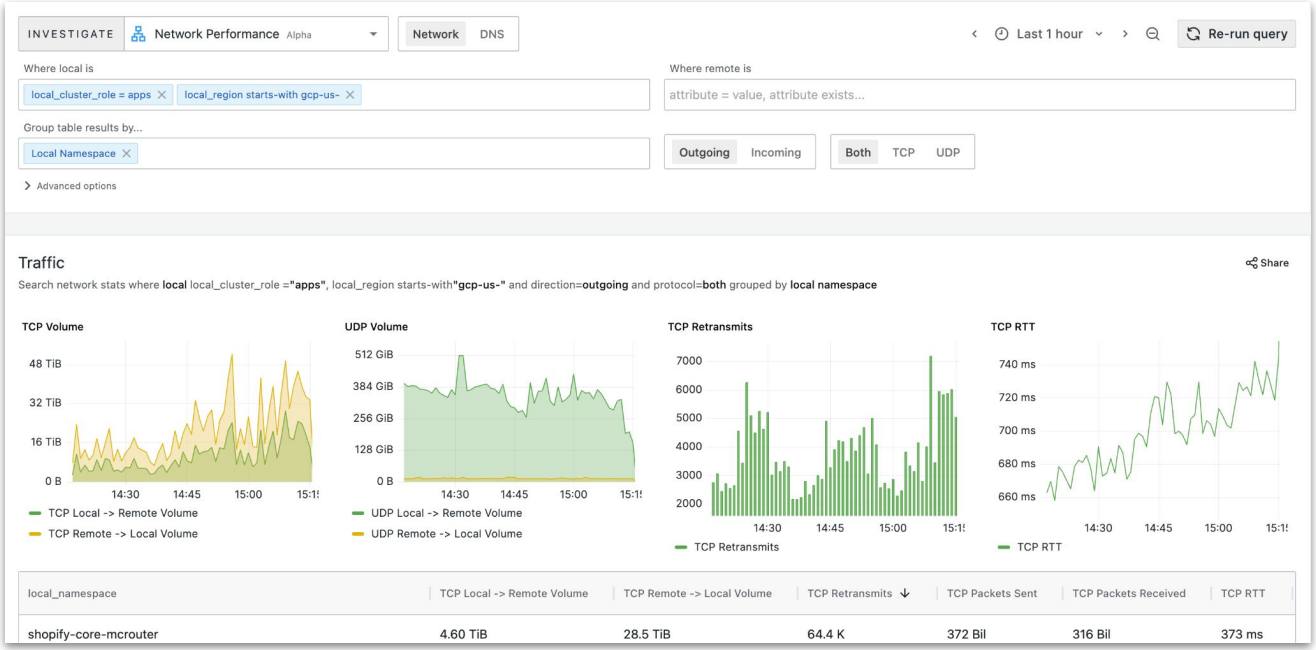
UI

- NPM is primarily used for troubleshooting to understand network state in relation to Shopify operations.
- It overlaps with Observe logs and traces investigation, but is more constrained.
- Using Investigate as a platform lets us leverage its existing capabilities while tailoring the user experience.

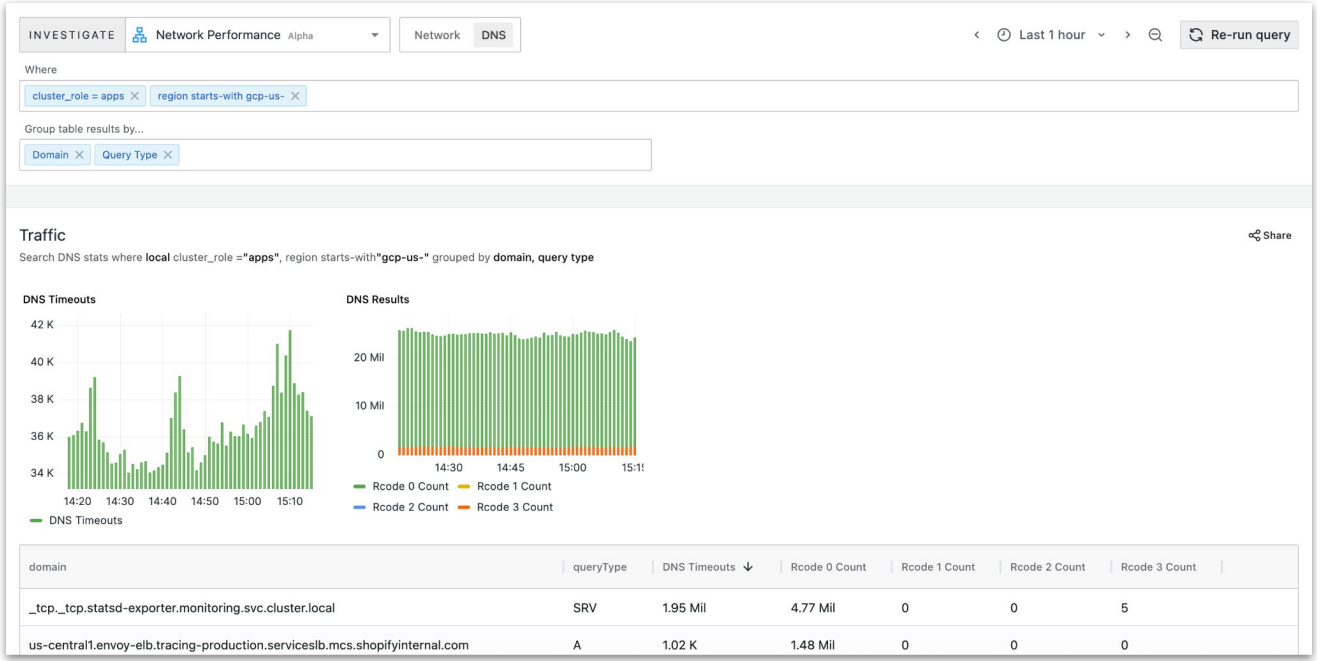
Querying


ClickHouse Kubernetes Cluster
ClickHouse
Queries
Investigate Kubernetes Cluster
Investigate API
Queries
Observe Frontend

Network View (Connections)



DNS View


INVESTIGATE
Network Performance Alpha
Network
DNS
Last 1 hour
Re-run query
Where
cluster_role = apps
region starts-with gcp-us-
Group table results by...
Domain
Query Type
Traffic
Share
Search DNS stats where local cluster_role ="apps", region starts-with"gcp-us-" grouped by domain, query type
DNS Timeouts
42 K
40 K
38 K
36 K
34 K
14:20
14:30
14:40
14:50
15:00
15:10
DNS Timeouts
DNS Results
20 Mil
10 Mil
0
14:30
14:45
15:00
Rcode 0 Count
Rcode 1 Count
Rcode 2 Count
Rcode 3 Count
domain
queryType
DNS Timeouts
Rcode 0 Count
Rcode 1 Count
Rcode 2 Count
Rcode 3 Count
_tcp._tcp.statsd-exporter.monitoring.svc.cluster.local
SRV
1.95 Mil
4.77 Mil
0
0
5
us-central1.envoy-elb.tracing-production.serviceslb.mcs.shopifyinternal.com
A
1.02 K
1.48 Mil
0
0
0

Demo

Limitations

- No support for cluster IPs and host network.
- Filtering and grouping remote currently limited to containers within Shopify Kubernetes clusters.
- No data from non-Kubernetes workloads.

Outlook

- Add support for node and cluster IPs, non-Kubernetes workloads and external services.
- Enhance UI and iterate on existing features based on user feedback.
- Leverage the underlying platform to surface new insights into network activity.

Questions?

Feedback