On The Edge of Small Data

Shannon Weyrick
VP Research/Fellow • NS1
sweyrick@ns1.com

NS1.

Who is processing flow or other visibility data from their infrastructure?

NS1.

Who thinks they should be getting more out of their solution?

NS1.

preface: the case for small data

NS1 Case Study

- Managed Authoritative DNS with 26 Global Anycasted POPs
- >100 billion DNS queries per average day
- >70 million flows/day
- 3.5 TB storage for only 30 days of flow history

The Data Conundrum

What we think we want:
All The Data

...because we think we *may* use it all *someday*

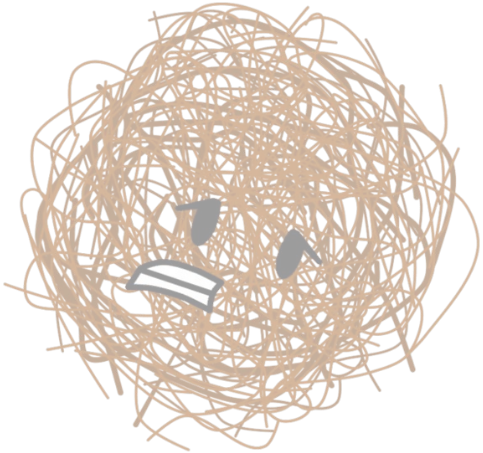
What we actually want:
Targeted Insights

...to help us operate, debug, scale and protect our networks *today*

There is a price to pay for streaming raw data to a central solution

The Costs of Raw Data

- Complicated data pipelines for centralized collection
- Batch processing costs to make it actionable
- Inability to make sense of or take advantage of all the data
- Slow dashboards, short retention times
- Slow reaction times to critical events
- Ingestion costs (esp. SaaS)



Paradigm Shift: Small Data

- Push the conversion of raw → actionable out to the edge
  - Distribute as close to the source as possible
- React quicker
  - Make those insights available at the edge *and* centrally
- Collect, process and store less
- Don’t find the needles in the haystack: just collect the needles
- Dynamically decide what your team needs at any time

Shannon Weyrick

Orb Founder, VP Research @ NS1

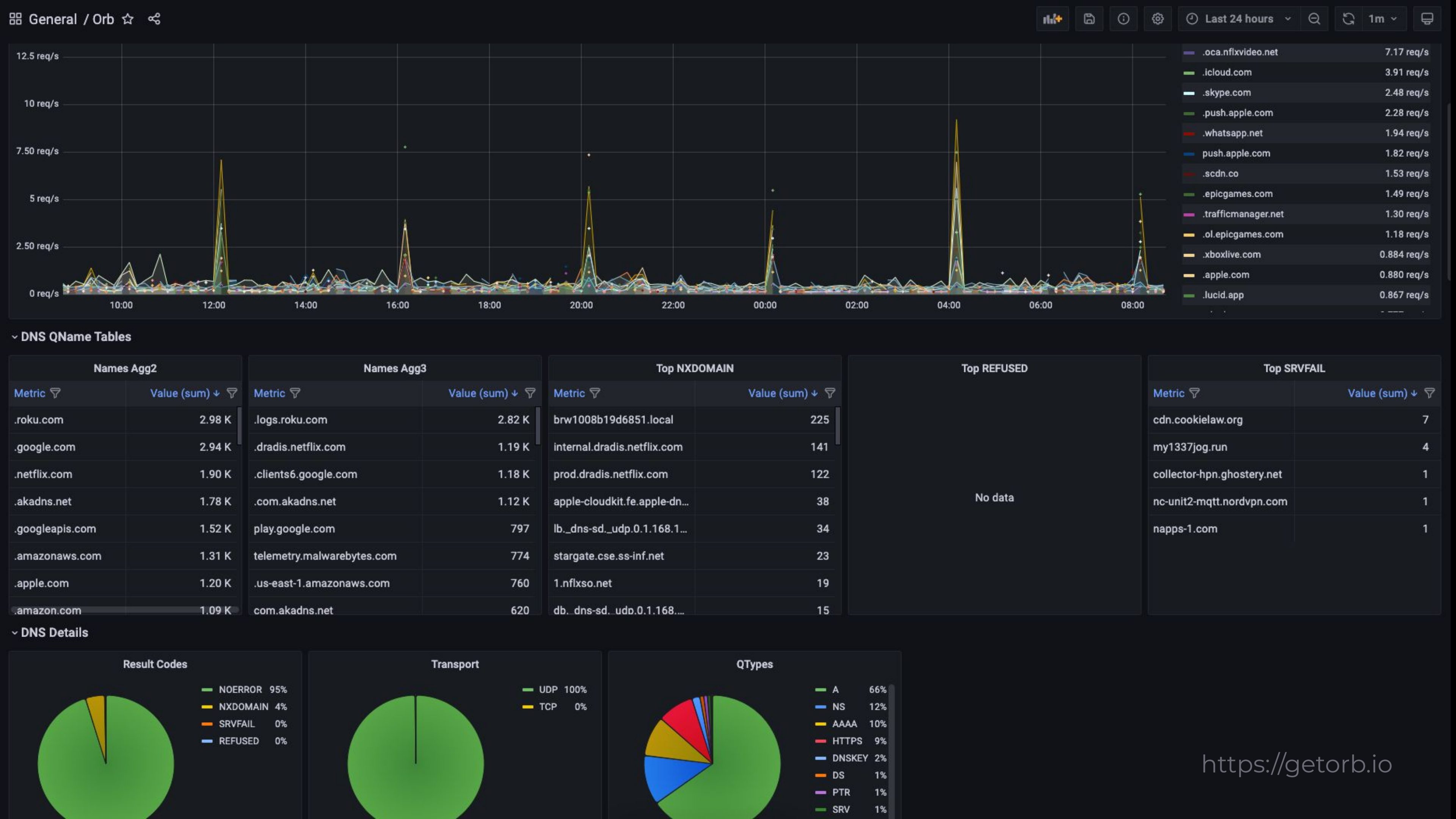
- 26 years in industry, 8 years at NS1
- NS1 engineering leadership
- Since start of 2021 focused on Orb open source innovation @ NS1 Labs
- sweyrick@ns1.com

If you remember just one thing from this talk...

NS1.

Orb is Open Source Edge Observability

- **Observability tool** designed for **distributed edge networks**
- Uses **small data** paradigm with **dynamic policy orchestration**
- Real-time **insights** into **data flow** on the **distributed edge**
- Integrates with **modern observability stacks**
- **Free** and **open source**, backed by NS1



.oca.nflxvideo.net	7.17 req/s
.icloud.com	3.91 req/s
.skype.com	2.48 req/s
.push.apple.com	2.28 req/s
.whatsapp.net	1.94 req/s
push.apple.com	1.82 req/s
.scdn.co	1.53 req/s
.epicgames.com	1.49 req/s
.trafficmanager.net	1.30 req/s
.ol.epicgames.com	1.18 req/s
.xboxlive.com	0.884 req/s
.apple.com	0.880 req/s
.lucid.app	0.867 req/s

DNS QName Tables

Names Agg2

Metric	Value (sum)
.roku.com	2.98 K
.google.com	2.94 K
.netflix.com	1.90 K
.akadns.net	1.78 K
.googleapis.com	1.52 K
.amazonaws.com	1.31 K
.apple.com	1.20 K
.amazon.com	1.09 K

Names Agg3

Metric	Value (sum)
.logs.roku.com	2.82 K
.dradis.netflix.com	1.19 K
.clients6.google.com	1.18 K
.com.akadns.net	1.12 K
play.google.com	797
telemetry.malwarebytes.com	774
.us-east-1.amazonaws.com	760
com.akadns.net	620

Top NXDOMAIN

Metric	Value (sum)
brw1008b19d6851.local	225
internal.dradis.netflix.com	141
prod.dradis.netflix.com	122
apple-cloudkit.fe.apple-dn...	38
lb._dns-sd._udp.0.1.168.1...	34
stargate.cse.ss-inf.net	23
1.nflxso.net	19
db._dns-sd._udp.0.1.168....	15

Top REFUSED

No data

Top SRVFAIL

Metric	Value (sum)
cdn.cookielaw.org	7
my1337jog.run	4
collector-hpn.ghostery.net	1
nc-unit2-mqtt.nordvpn.com	1
napps-1.com	1

DNS Details

Result Codes

NOERROR	95%
NXDOMAIN	4%
SRVFAIL	0%
REFUSED	0%

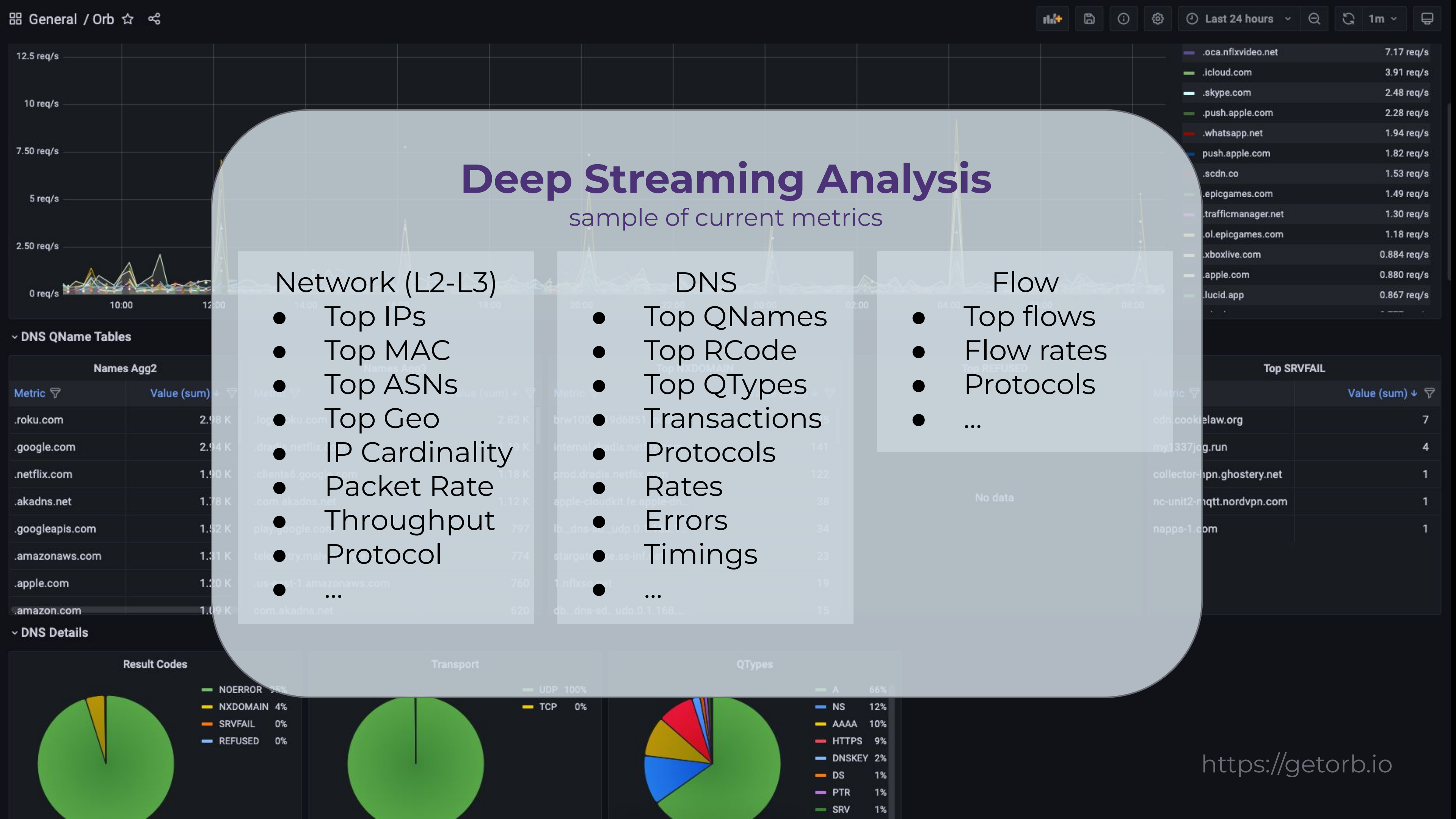
Transport

UDP	100%
TCP	0%

QTypes

A	66%
NS	12%
AAAA	10%
HTTPS	9%
DNSKEY	2%
DS	1%
PTR	1%
SRV	1%

https://getorb.io



Deep Streaming Analysis

sample of current metrics

Network (L2-L3)

- Top IPs
- Top MAC
- Top ASNs
- Top Geo
- IP Cardinality
- Packet Rate
- Throughput
- Protocol
- ...

DNS

- Top QNames
- Top RCode
- Top QTypes
- Transactions
- Protocols
- Rates
- Errors
- Timings
- ...

Flow

- Top flows
- Flow rates
- Protocols
- ...

.oca.nflxvideo.net	7.17 req/s
.icloud.com	3.91 req/s
.skype.com	2.48 req/s
.push.apple.com	2.28 req/s
.whatsapp.net	1.94 req/s
push.apple.com	1.82 req/s
.scdn.co	1.53 req/s
.epicgames.com	1.49 req/s
.trafficmanager.net	1.30 req/s
.ol.epicgames.com	1.18 req/s
.xboxlive.com	0.884 req/s
.apple.com	0.880 req/s
.lucid.app	0.867 req/s

Names Agg2

Metric	Value (sum)
.roku.com	2.98 K
.google.com	2.94 K
.netflix.com	1.90 K
.akadns.net	1.78 K
.googleapis.com	1.52 K
.amazonaws.com	1.31 K
.apple.com	1.20 K
.amazon.com	1.09 K

Top SRVFAIL

Metric	Value (sum)
cdn.cookielaw.org	7
my1337jog.run	4
collector-hpn.ghostery.net	1
nc-unit2-mqtt.nordvpn.com	1
napps-1.com	1

Result Codes

NOERROR	95%
NXDOMAIN	4%
SRVFAIL	0%
REFUSED	0%

Transport

UDP	100%
TCP	0%

QTypes

A	66%
NS	12%
AAAA	10%
HTTPS	9%
DNSKEY	2%
DS	1%
PTR	1%
SRV	1%

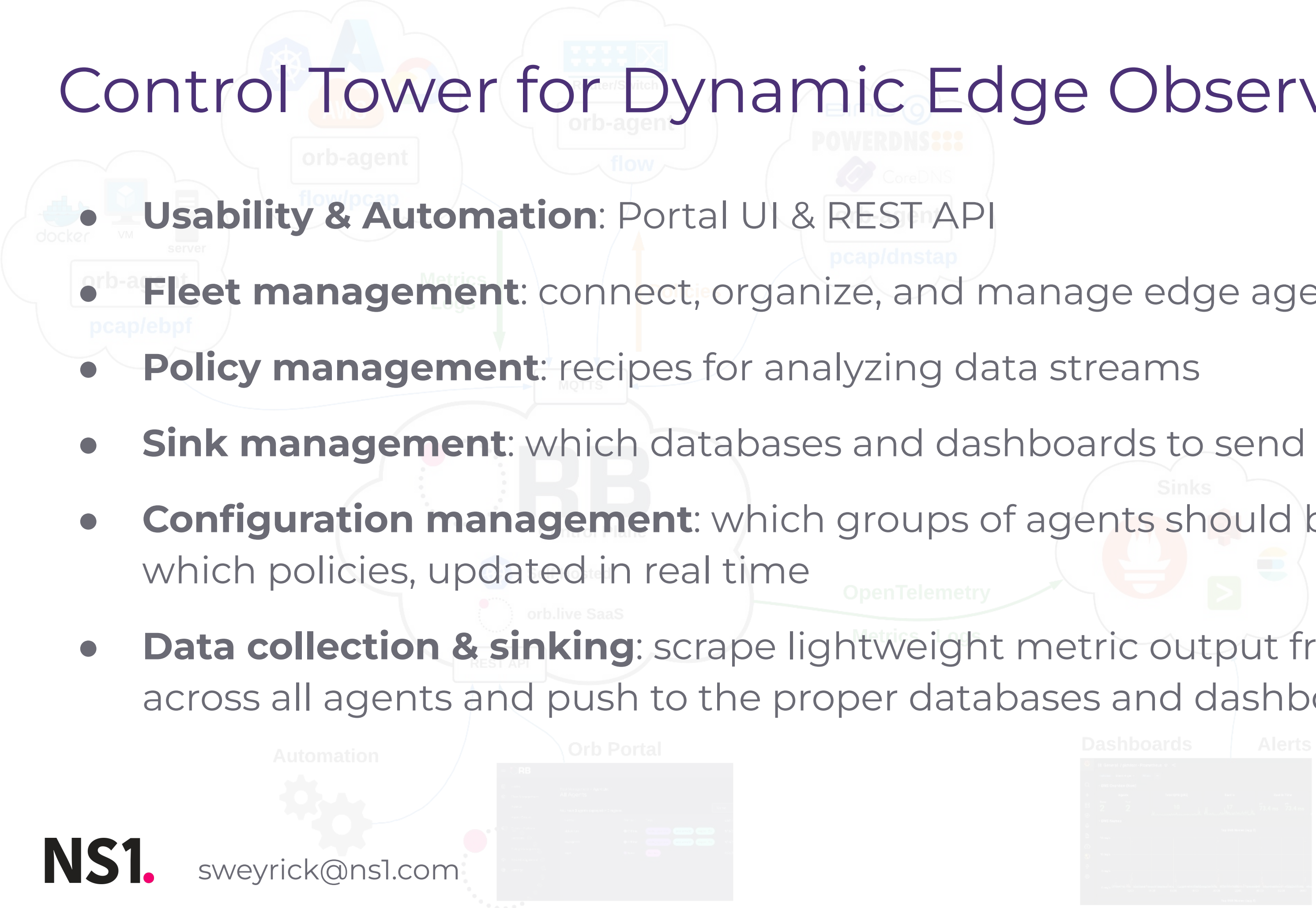
https://getorb.io

control tower for the edge

Orb control plane: cloud native application

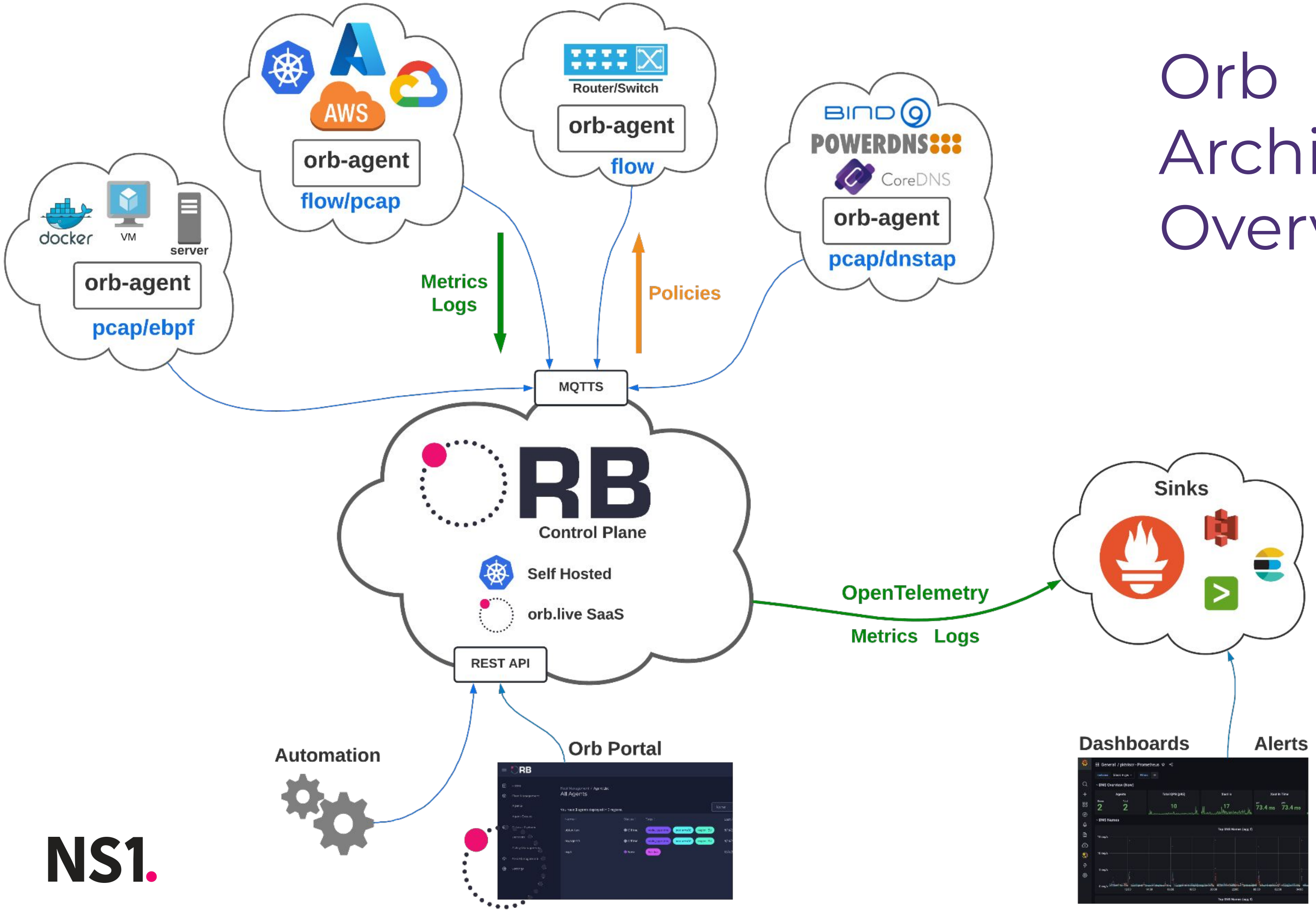
NS1.

Control Tower for Dynamic Edge Observability



- **Usability & Automation**: Portal UI & REST API
- **Fleet management**: connect, organize, and manage edge agents
- **Policy management**: recipes for analyzing data streams
- **Sink management**: which databases and dashboards to send metrics to
- **Configuration management**: which groups of agents should be running which policies, updated in real time
- **Data collection & sinking**: scrape lightweight metric output from all policies across all agents and push to the proper databases and dashboards

Orb Architecture Overview

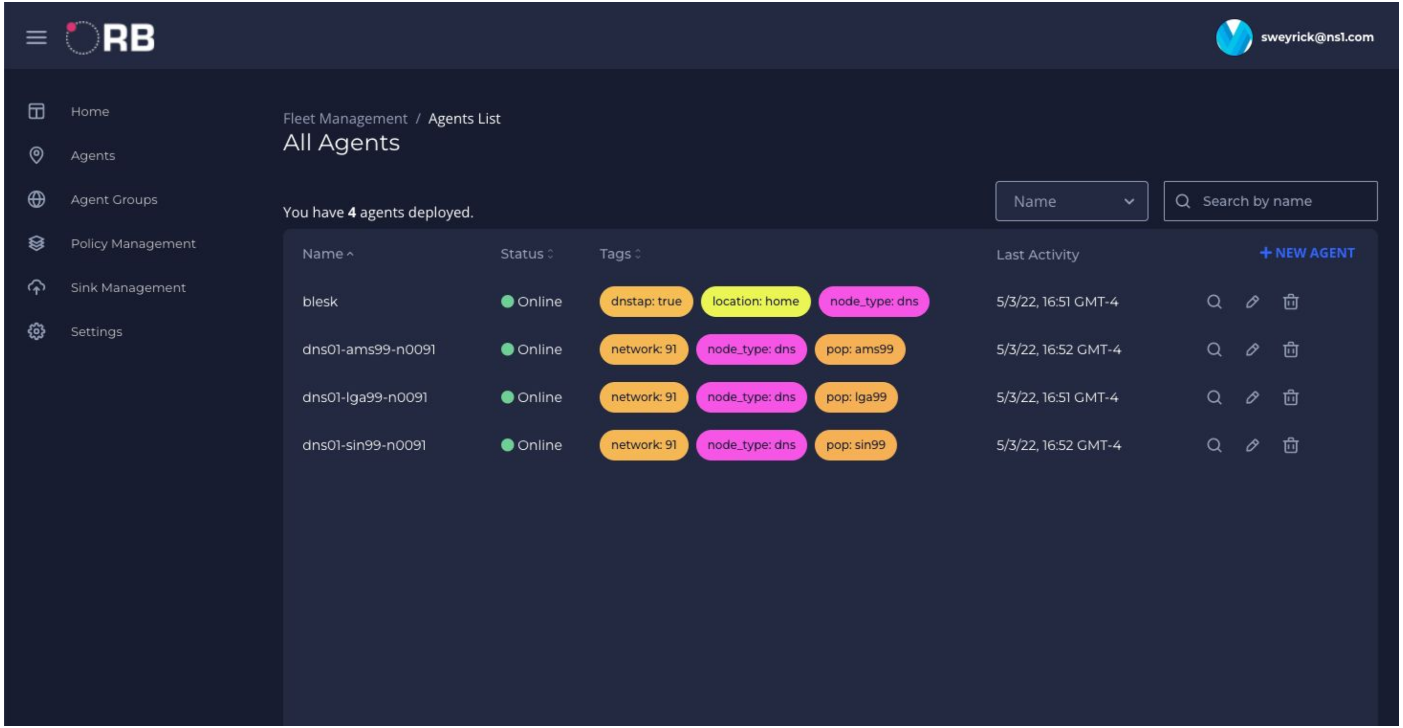


NS1.

https://getorb.io

Fleet Management

Connect, organize, and manage edge agents



Fleet Management / Agents List

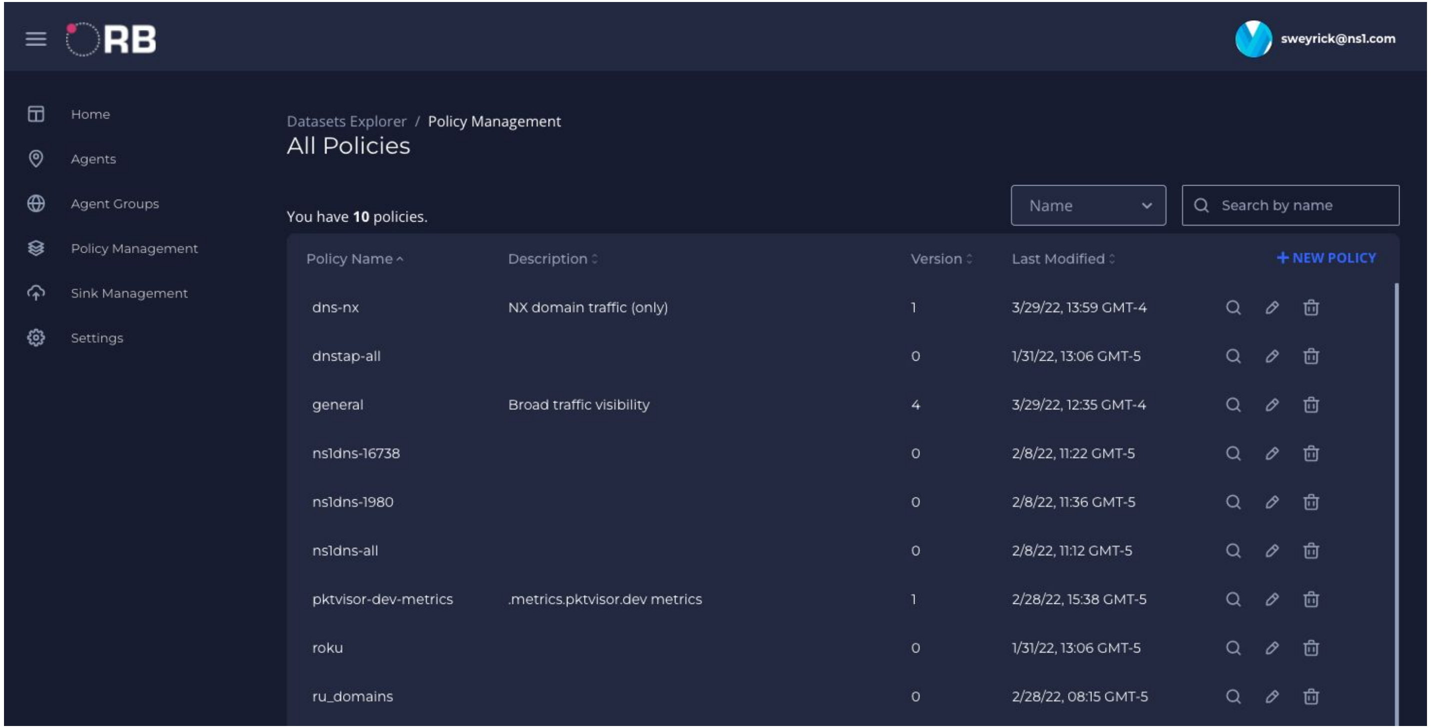
All Agents

You have **4** agents deployed.

Name	Status	Tags	Last Activity
blesk	● Online	dnstap: true location: home node_type: dns	5/3/22, 16:51 GMT-4
dns01-ams99-n0091	● Online	network: 91 node_type: dns pop: ams99	5/3/22, 16:52 GMT-4
dns01-lga99-n0091	● Online	network: 91 node_type: dns pop: lga99	5/3/22, 16:51 GMT-4
dns01-sin99-n0091	● Online	network: 91 node_type: dns pop: sin99	5/3/22, 16:52 GMT-4

Policy Management

Recipes for analyzing data streams



Datasets Explorer / Policy Management

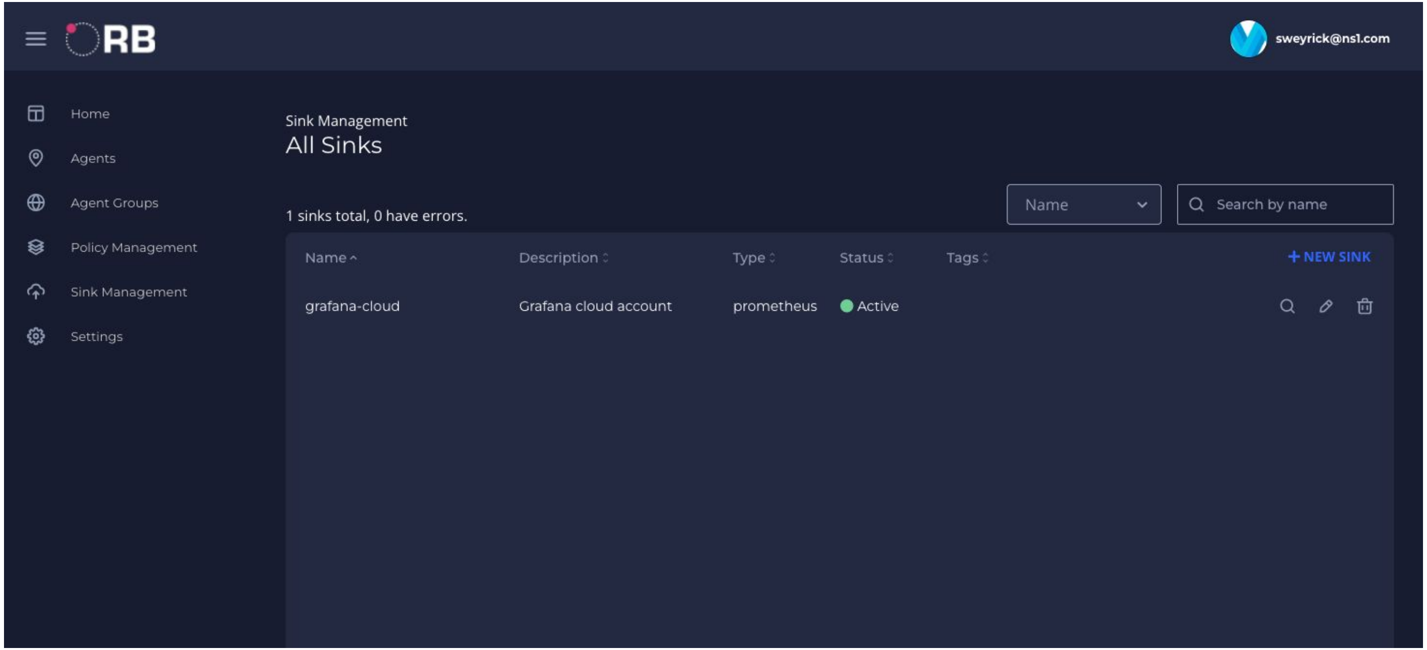
All Policies

You have **10** policies.

Policy Name	Description	Version	Last Modified
dns-nx	NX domain traffic (only)	1	3/29/22, 13:59 GMT-4
dnstap-all		0	1/31/22, 13:06 GMT-5
general	Broad traffic visibility	4	3/29/22, 12:35 GMT-4
ns1dns-16738		0	2/8/22, 11:22 GMT-5
ns1dns-1980		0	2/8/22, 11:36 GMT-5
ns1dns-all		0	2/8/22, 11:12 GMT-5
pktvisor-dev-metrics	.metrics.pktvisor.dev metrics	1	2/28/22, 15:38 GMT-5
roku		0	1/31/22, 13:06 GMT-5
ru_domains		0	2/28/22, 08:15 GMT-5

Sink Management

Which databases and dashboards to send metrics to



Sink Management

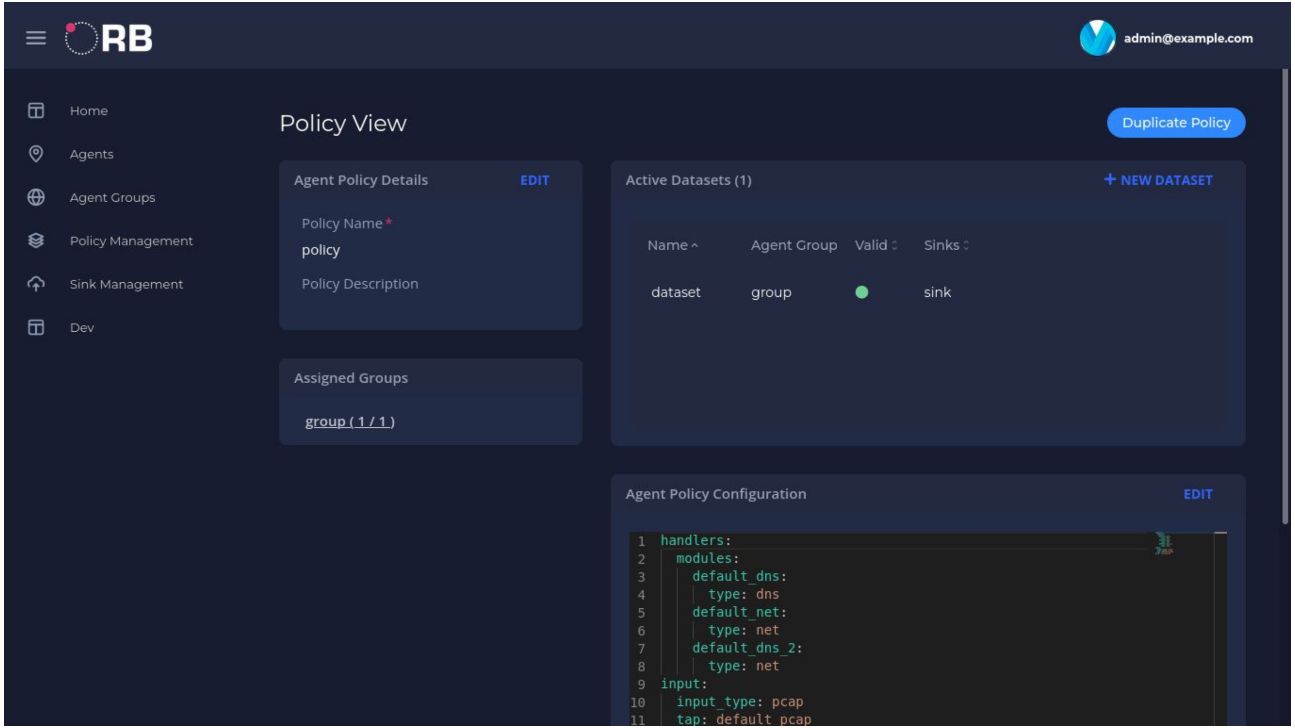
All Sinks

1 sinks total, 0 have errors.

Name	Description	Type	Status	Tags
grafana-cloud	Grafana cloud account	prometheus	● Active	

Configuration Management

Which agents should run which policies, update in real time



Policy View

Agent Policy Details

Policy Name *
policy

Policy Description

Assigned Groups

group (1 / 1)

Active Datasets (1)

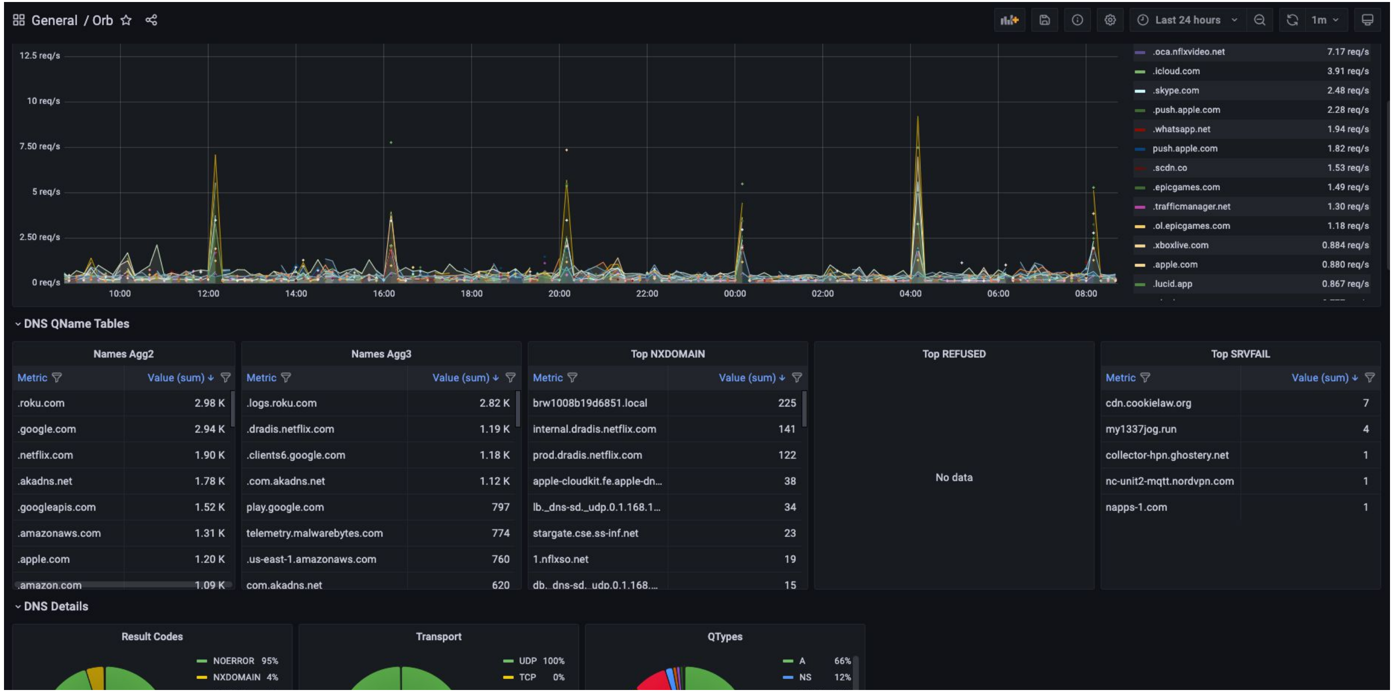
Name	Agent Group	Valid	Sinks
dataset	group	●	sink

Agent Policy Configuration

```
handlers:
  modules:
    default_dns:
      type: dns
    default_net:
      type: net
    default_dns_2:
      type: net
input:
  input_type: pcap
  tap: default_pcap
```

Data Collection & Sinking

Scrape lightweight metric output from all policies across all agents and push to the proper databases and dashboards



.oca.nflxvideo.net	7.17 req/s
.icloud.com	3.91 req/s
.skype.com	2.48 req/s
.push.apple.com	2.28 req/s
.whatsapp.net	1.94 req/s
push.apple.com	1.82 req/s
.scdn.co	1.53 req/s
.epicgames.com	1.49 req/s
.trafficmanager.net	1.30 req/s
.ol.epicgames.com	1.18 req/s
.xboxlive.com	0.884 req/s
.apple.com	0.880 req/s
.lucid.app	0.867 req/s

˅ DNS QName Tables

Names Agg2

Metric	Value (sum) ↓
.roku.com	2.98 K
.google.com	2.94 K
.netflix.com	1.90 K
.akadns.net	1.78 K
.googleapis.com	1.52 K
.amazonaws.com	1.31 K
.apple.com	1.20 K
.amazon.com	1.09 K

Names Agg3

Metric	Value (sum) ↓
.logs.roku.com	2.82 K
.dradis.netflix.com	1.19 K
.clients6.google.com	1.18 K
.com.akadns.net	1.12 K
play.google.com	797
telemetry.malwarebytes.com	774
.us-east-1.amazonaws.com	760
com.akadns.net	620

Top NXDOMAIN

Metric	Value (sum) ↓
brw1008b19d6851.local	225
internal.dradis.netflix.com	141
prod.dradis.netflix.com	122
apple-cloudkit.fe.apple-dn...	38
lb._dns-sd._udp.0.1.168.1...	34
stargate.cse.ss-inf.net	23
1.nflxso.net	19
db._dns-sd._udp.0.1.168....	15

Top REFUSED

No data

Top SRVFAIL

Metric	Value (sum) ↓
cdn.cookielaw.org	7
my1337jog.run	4
collector-hpn.ghostery.net	1
nc-unit2-mqtt.nordvpn.com	1
napps-1.com	1

˅ DNS Details

Result Codes

NOERROR	95%
NXDOMAIN	4%

Transport

UDP	100%
TCP	0%

QTypes

A	66%
NS	12%

edge agent for streaming analysis

orb-agent

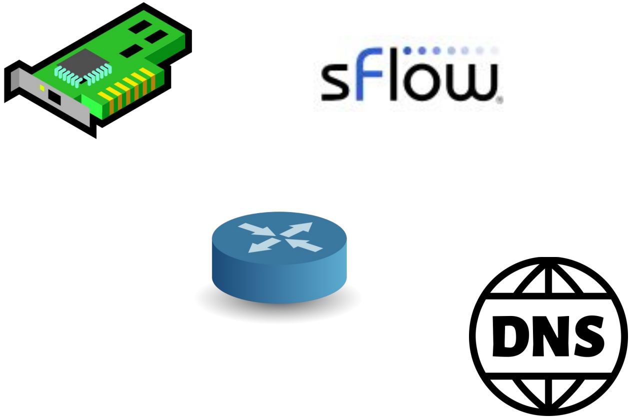
NS1.

What Is The Orb Edge Agent?

- **Taps into** multiple, concurrent data streams at the edge
- Uses **fast streaming algorithms** to **analyze deeply** in real time
- **Efficiently summarizes** important insights, generate metrics
- Can be **reprogrammed in real time** with dynamic policies
- Can **scale up** and **scale down**

What Can It Tap Into?

- Packet capture
- dnstap
- Network flow (sFlow, Netflow/IPFIX)
- SNMP (soon)
- envoy taps (soon)
- eBPF (soon)
- Expandable via custom loadable modules



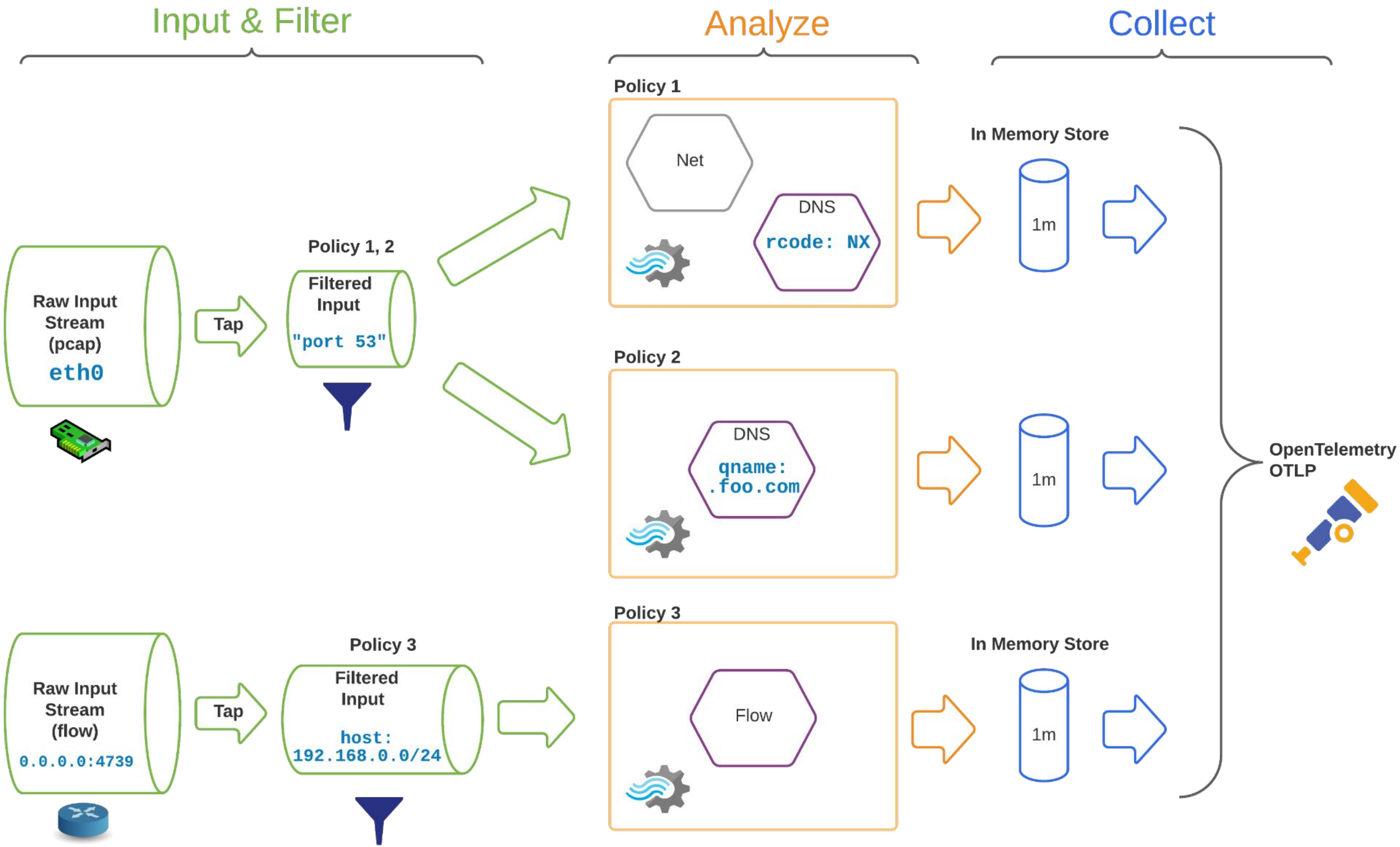
What Can It Generate Metrics For?

- L2-L3 Network
- DNS
- DHCP
- Flows
- Policy resource usage
- Expandable via custom loadable modules

Orb Edge Agent

Embedded Stream Processing

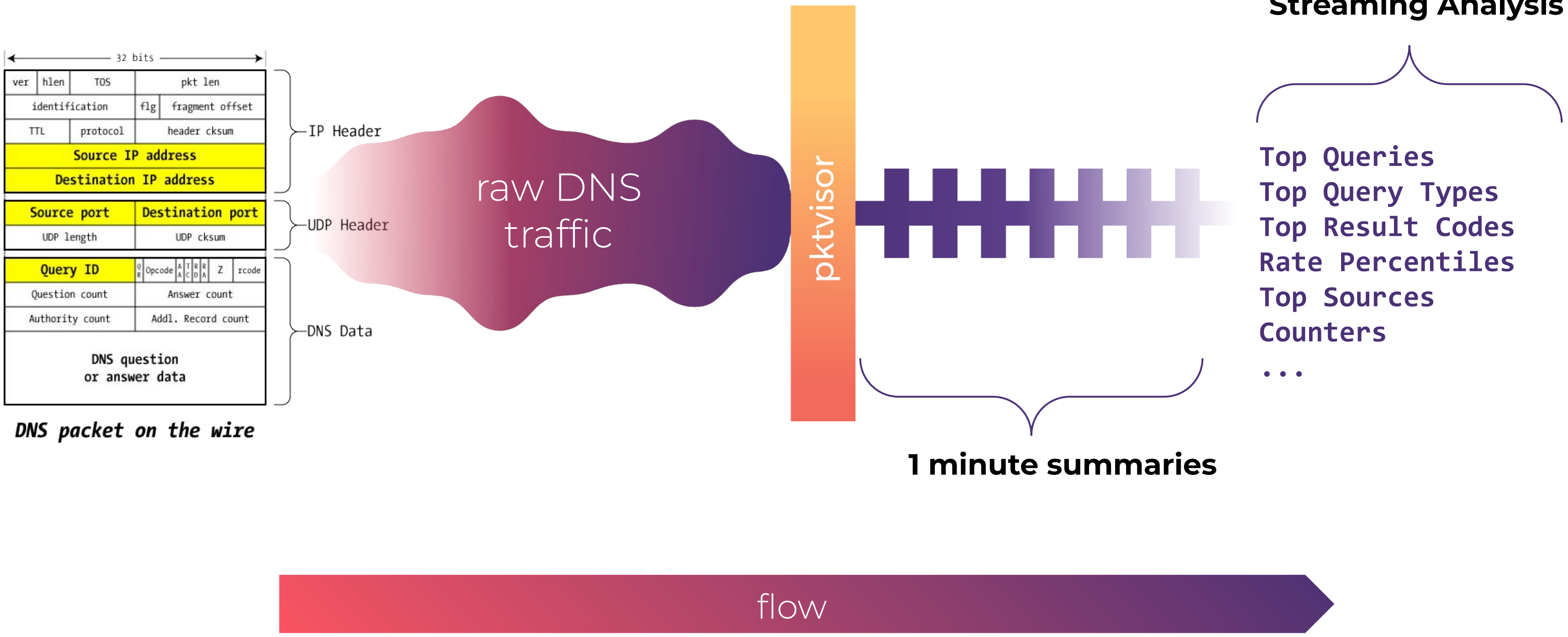
powered by pktvisor



NS1.   sweyrick@ns1.com

https://getorb.io

Use Case: DNS Analysis



32 bits

ver | hlen | TOS | pkt len
identification | flg | fragment offset
TTL | protocol | header cksum
Source IP address
Destination IP address
Source port | Destination port
UDP length | UDP cksum
Query ID | QR | Opcode | AA | TC | RD | RA | Z | rcode
Question count | Answer count
Authority count | Addl. Record count
DNS question or answer data

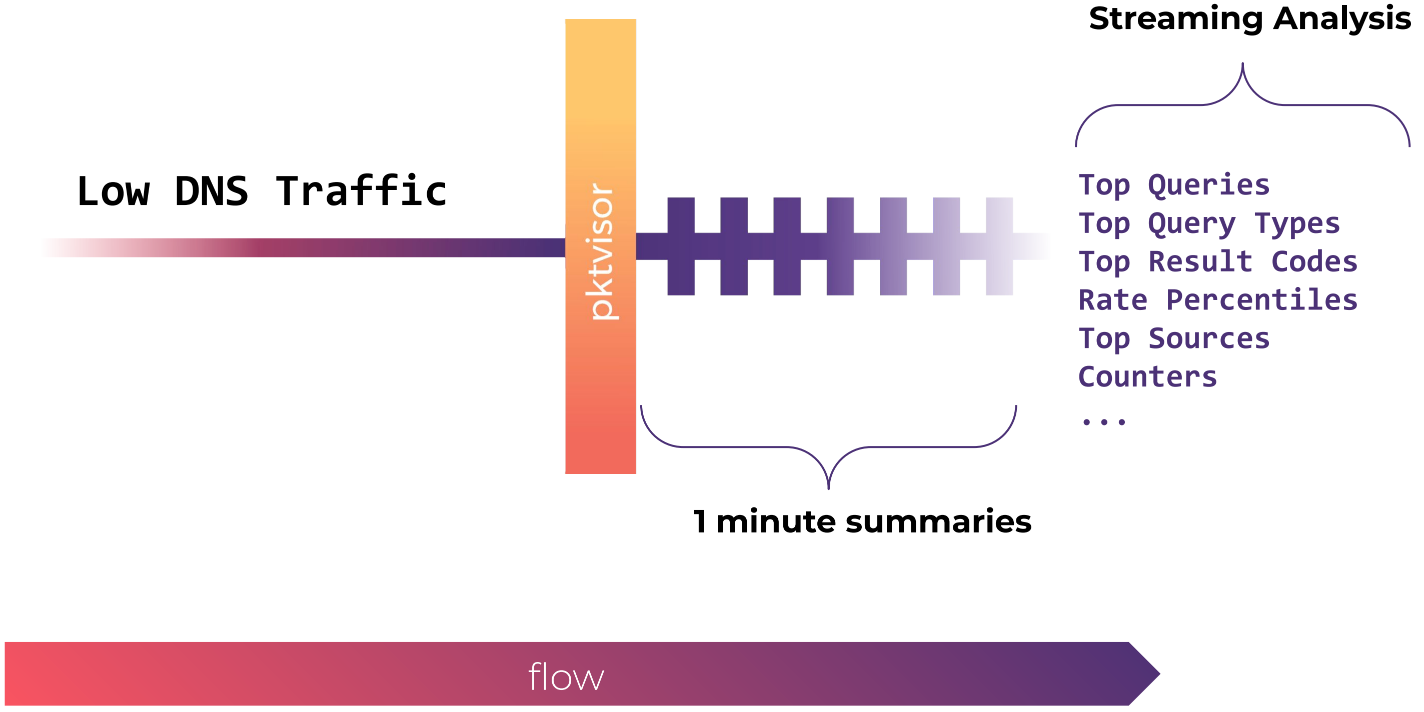
DNS packet on the wire

1 minute summaries

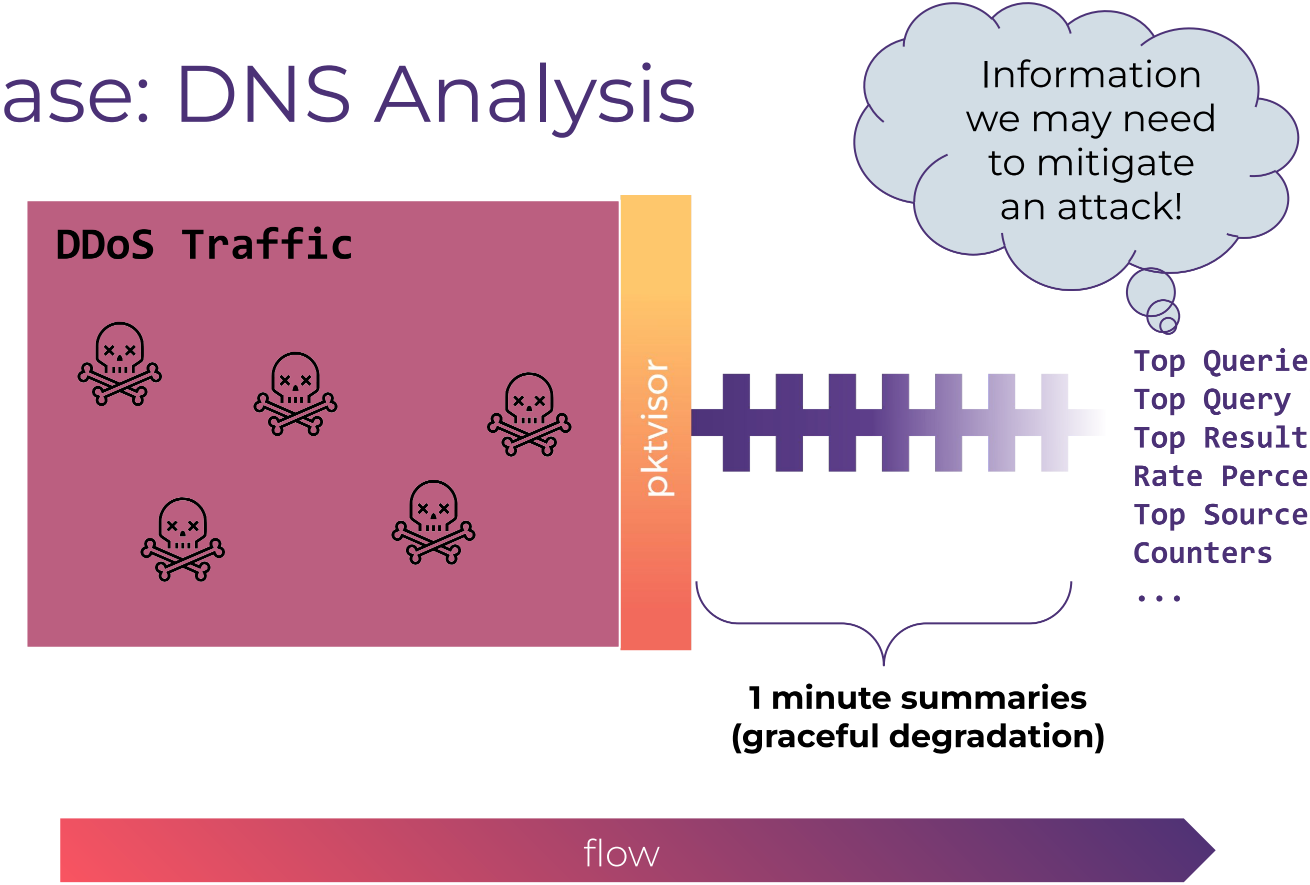
Streaming Analysis

Top Queries
Top Query Types
Top Result Codes
Rate Percentiles
Top Sources
Counters
...

Use Case: DNS Analysis



Use Case: DNS Analysis



Top Queries
Top Query Types
Top Result Codes
Rate Percentiles
Top Sources
Counters
...

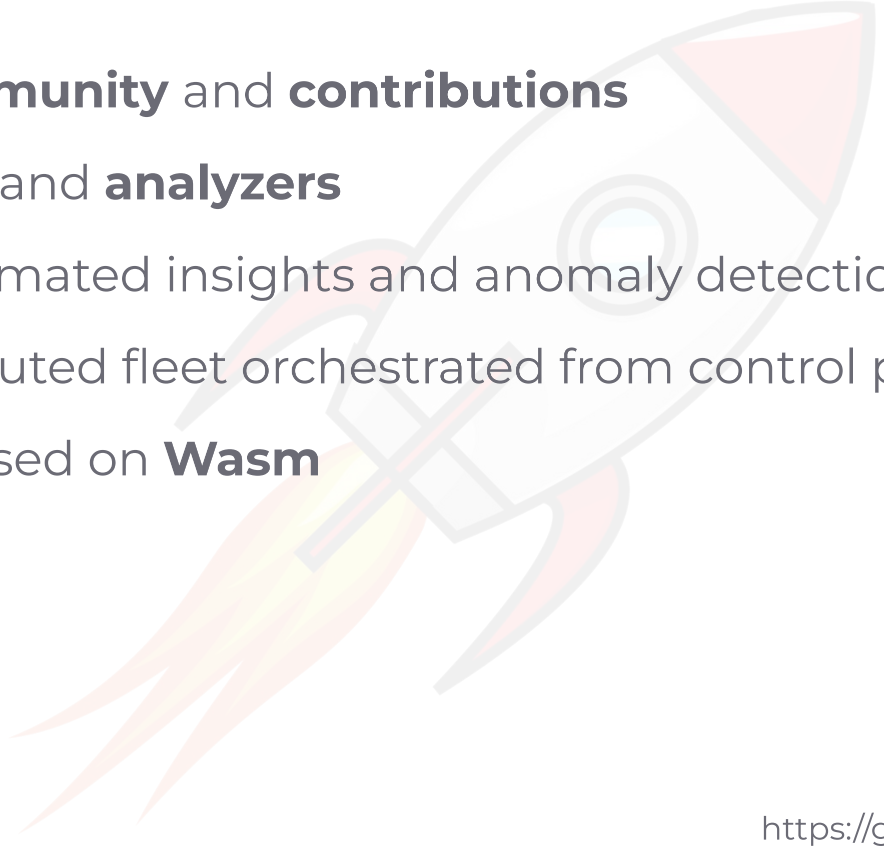
1 minute summaries (graceful degradation)

Tech Notes

- Orb edge agent runs on Linux x86_64 and ARM
  - Available as Docker containers or statically linked binaries
  - Connect to Orb control plane over MQTT over TLS
- Orb control plane runs in Kubernetes or Docker Compose
  - Helm chart available
- Today Orb sinks metrics to Prometheus compatible TSDB
  - remote_write is compatible with several TSDBs and cloud services
  - Wholesale replacement with OpenTelemetry nearly complete

Exciting Future

- Expanding our active **community** and **contributions**
- New input stream sources and **analyzers**
- **Machine learning** for automated insights and anomaly detection
- **pcap samples** from distributed fleet orchestrated from control plane
- Custom edge analyzers based on **Wasm**
- What are your ideas?



conclusion

Remember This

- **Observability tool** designed for **distributed edge networks**
- Uses **small data** paradigm with **dynamic policy orchestration**
- Real-time **insights** into **data flow** on the **distributed edge**
- Integrates with **modern observability stacks**
- **Free** and **open source**, backed by NS1

Do This

- Join the community: https://getorb.io
- Try Orb SaaS for free: https://orb.live
- Star the project: github.com/ns1labs/orb
- Give us your feedback! We’d love to understand your use case

thank you

NS1.