A new traffic capture and visualisation tool for IMRS

Jim Hague jim@sinodun.com
https://sinodun.com
@SinodunCom

Traffic capture and visualisation for IMRS

- **What is IMRS?** ICANN Managed Root Server (the root server formally known as L-root)

- **How is traffic captured?** Using RFC8618 C-DNS (Compressed-DNS) a CBOR based DNS specific file format for traffic capture.
    - Pairs query/responses and indexes common data
    - Why use it? Much smaller than PCAP with most of same info

- **How to visualise the data?** Import into ClickHouse and display in Grafana. Aggregation of data is important factor here!

Project Background

- Sinodun contract for ICANN DNS Eng Team who manage IMRS.
- Open source code developed via DNS-STATS (dns-stats.org).
- Historically used a combination DSC XML + Hedgehog (+PCAP)
- Have now migrated to C-DNS + ClickHouse + Grafana solution
  - Presented C-DNS at OARC29, C-DNS RFC8618 published Sept 2019
  - We will describe this full solution today

C-DNS targets limited use case

IMRS is ~280 (mainly) hosted servers in challenging environments
Managed as ~170 ‘instances’ in different locations

Total traffic is **~17 billion queries per day**

- Data collection on **same hardware** as nameserver
- Minimise server resources conflict: **1 RU server**
- Collected data **stored on same hardware**
- **Upload** will use the same interface as DNS traffic

C-DNS File sizes

Format	PCAP	C-DNS
File size (Mb)	660	75
Compressed with ‘xz -9’ (Mb)	49	18
User time for compression (s)	161	39

C-DNS File sizes

Format	PCAP	C-DNS
File size (Mb)	660	75
Compressed with ‘xz -9’ (Mb)	49	18
User time for compression (s)	161	39

***COMPRESSED SIZE**: C-DNS is 30-40% size of PCAP*

C-DNS File sizes

Format	PCAP	C-DNS
File size (Mb)	660	75
Compressed with ‘xz -9’ (Mb)	49	18
User time for compression (s)	161	39

***COMPRESSED SIZE**: C-DNS is 30-40% size of PCAP*

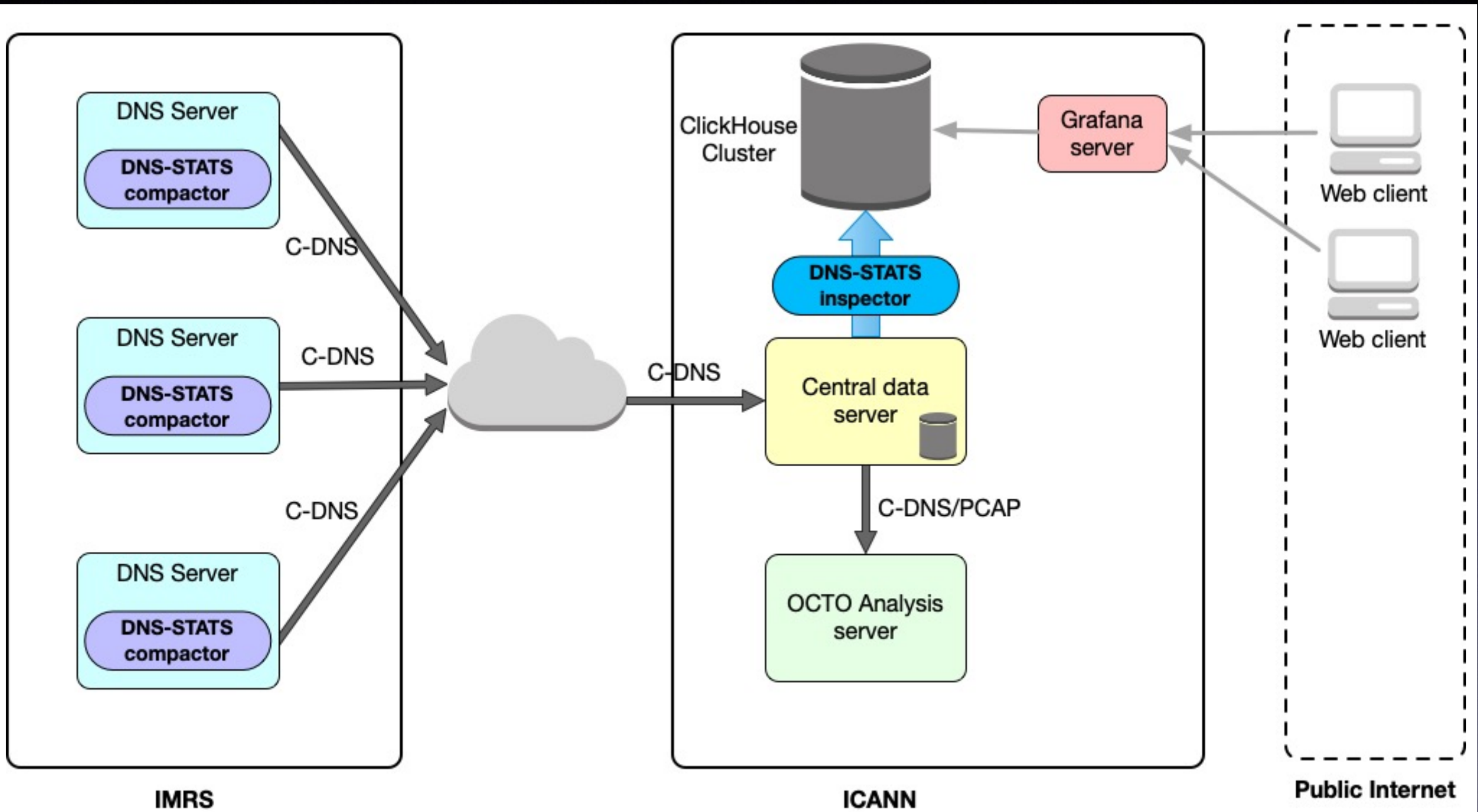
***COMPRESSION CPU**: C-DNS uses ~25% of PCAP*

C-DNS Implementation Status

- dns-stats github: https://github.com/dns-stats/compactor
- Software has two components:
  - **compactor**: Captures & compresses traffic in C-DNS format
  - **inspector**: Reads C-DNS and has 2 output formats
    - Templated text output for import to database
    - Lossy reconstruction of PCAP (files used by OCTO)

- **NOTE**: v0.9 uses -04 C-DNS draft format but v1.0 supports RFC8618 format (writes RFC format, reads both)

Architectural Overview

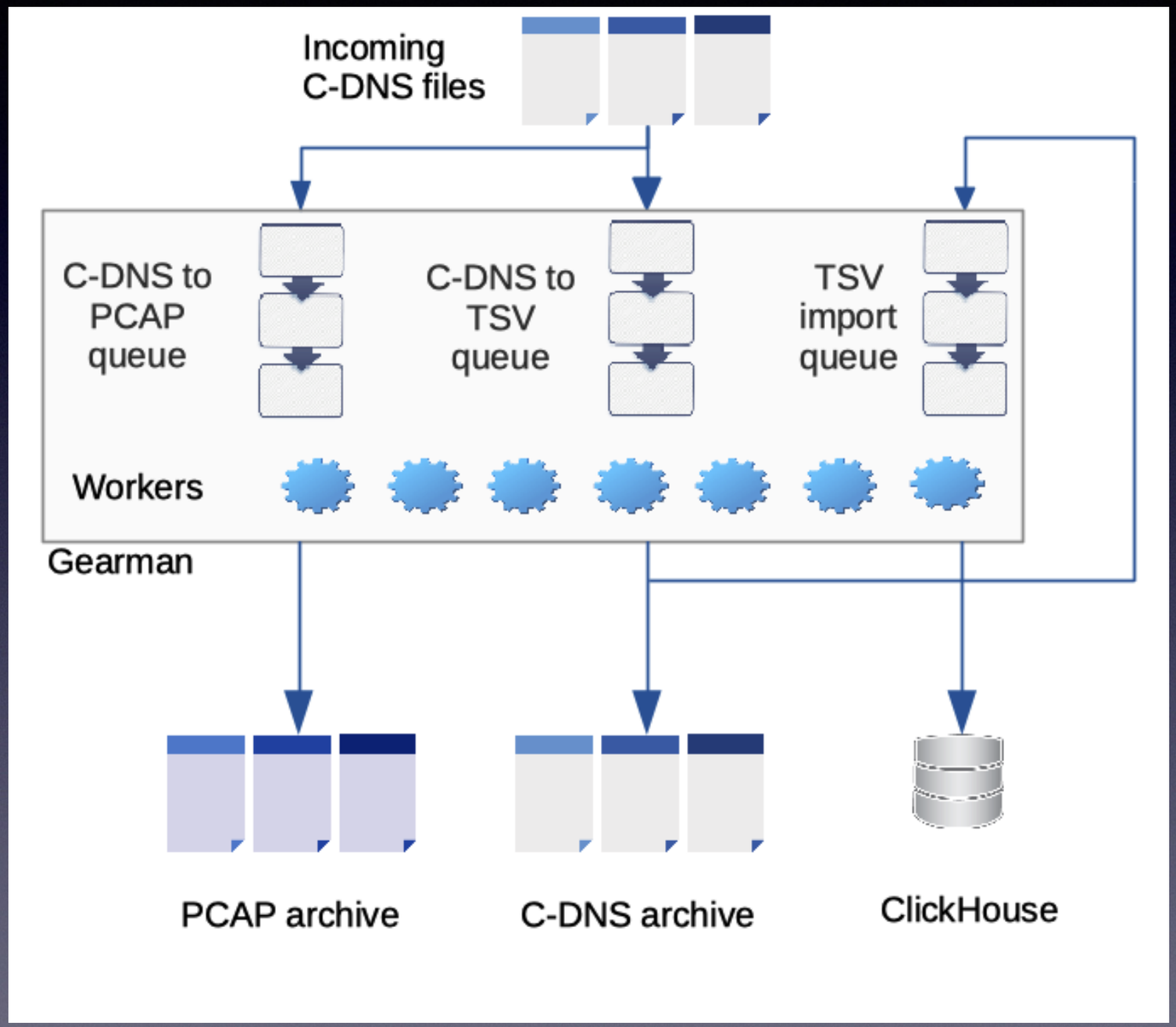


dns-stats compactor Deployment

- **compactor** constrained to **1 CPU** on the DNS server
  - Collects all data specified in C-DNS (query + response)
    - Note: RFC format allows almost all elements to be optional
  - Writes xz compressed files to local storage
  - Output file rotated every 5 minutes (configurable)
  - Handles query rates of up to 80 kqps depending on core and compression level

- Periodically files uploaded to central collection server

dns-stats inspector Deployment



- Uploaded C-DNS files queued for processing using **Gearman** job server and suite of Python programs
- Separate queues
  - Convert C-DNS to Tab-Separated-Value (TSV) files
  - Import TSV into ClickHouse database
  - Optionally convert C-DNS to anonymised PCAP e.g. for DITL

ClickHouse deployment

- ClickHouse is an open source time series SQL column database with Grafana plugin (other plugins are available!)
- Used by various other DNS projects (CloudFlare, NIC Chile)

- C-DNS schema:
  - **Main table**: holds raw C-DNS data - per q/r pair data
  - **Aggregation tables:** Does ON INSERT **aggregation** of data into separate 1 second and 5 minute tables
    - Aggregation is simple SQL MATERIALIZED VIEW with specialised storage engine (more here)

ClickHouse deployment

- 6 server cluster
- Import process handles ~17 billion records per day (~200 kqps)
- Disc usage 1Tb per ~39 billion records (2+ days of raw data)
- Management tools provide option to retain configurable amount of each type of data (raw vs 1s vs 5m)
- Serves multiple **Grafana** front ends and can be used for ad-hoc queries for data analysis

ClickHouse numbers

- Sample query speed: **count all AAAA queries in a week**
  - Raw data is 200 kqps i.e. a packet every ~5 micro sec
  - Table sizes are for full set of DSC like aggregations

Data Type	Query Speed (s)	Rows processed	Data size (1 week)
Raw	22	123 billion	4 Tb
1 sec agg	1.6	760 million	~1 Tb
5 min agg	0.13	3 million	~0.1 Tb

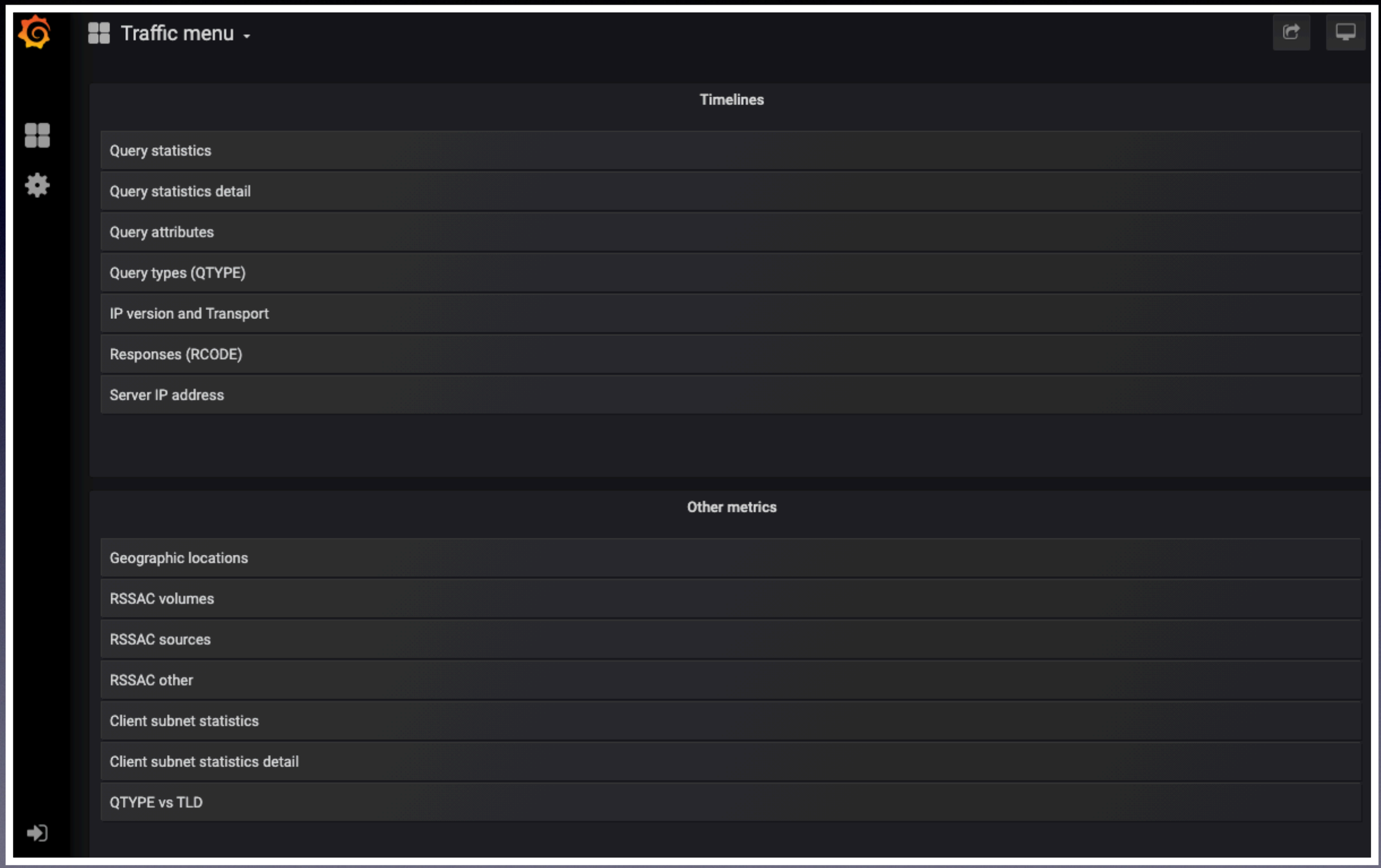
- Orders of magnitude reductions in query time and storage

Grafana deployment

- Web-based visualisation platform with various plot types:
  - Time series
  - Bar chart (using Sinodun modified plugin based on Plotly)
  - Map (using standard plugin)
  - Other plugins: ClickHouse data access, Image rendering

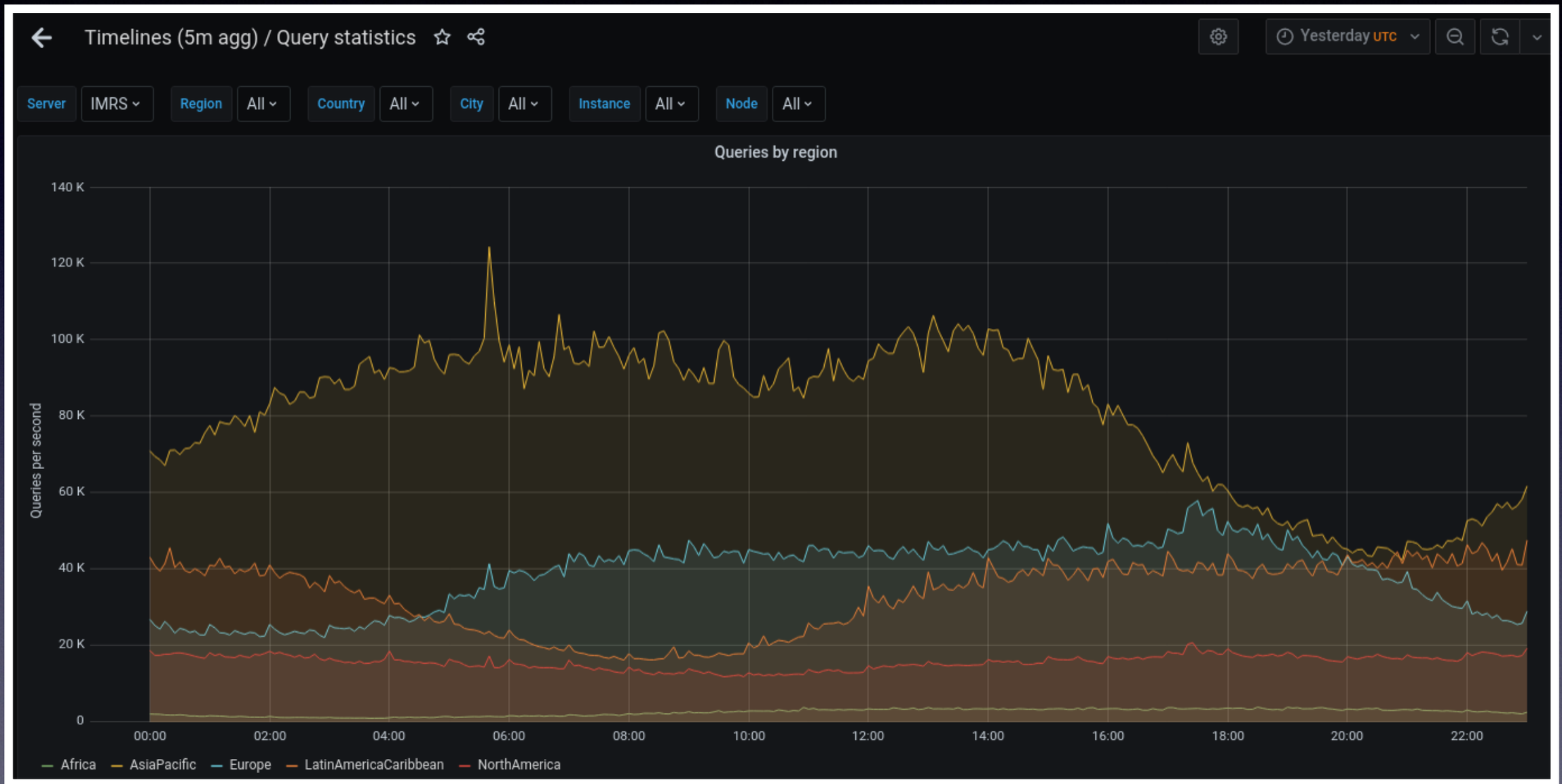
- ICANN public Grafana interface https://stats.dns.icann.org
  - Reproduces the various DSC like plots
  - Exposes just the 5 minute data with max time window
- Full data available via customised Grafana to ICANN staff

Grafana dashboard



Timeseries graph

Query Statistics



Timeseries graphs

Query Attributes

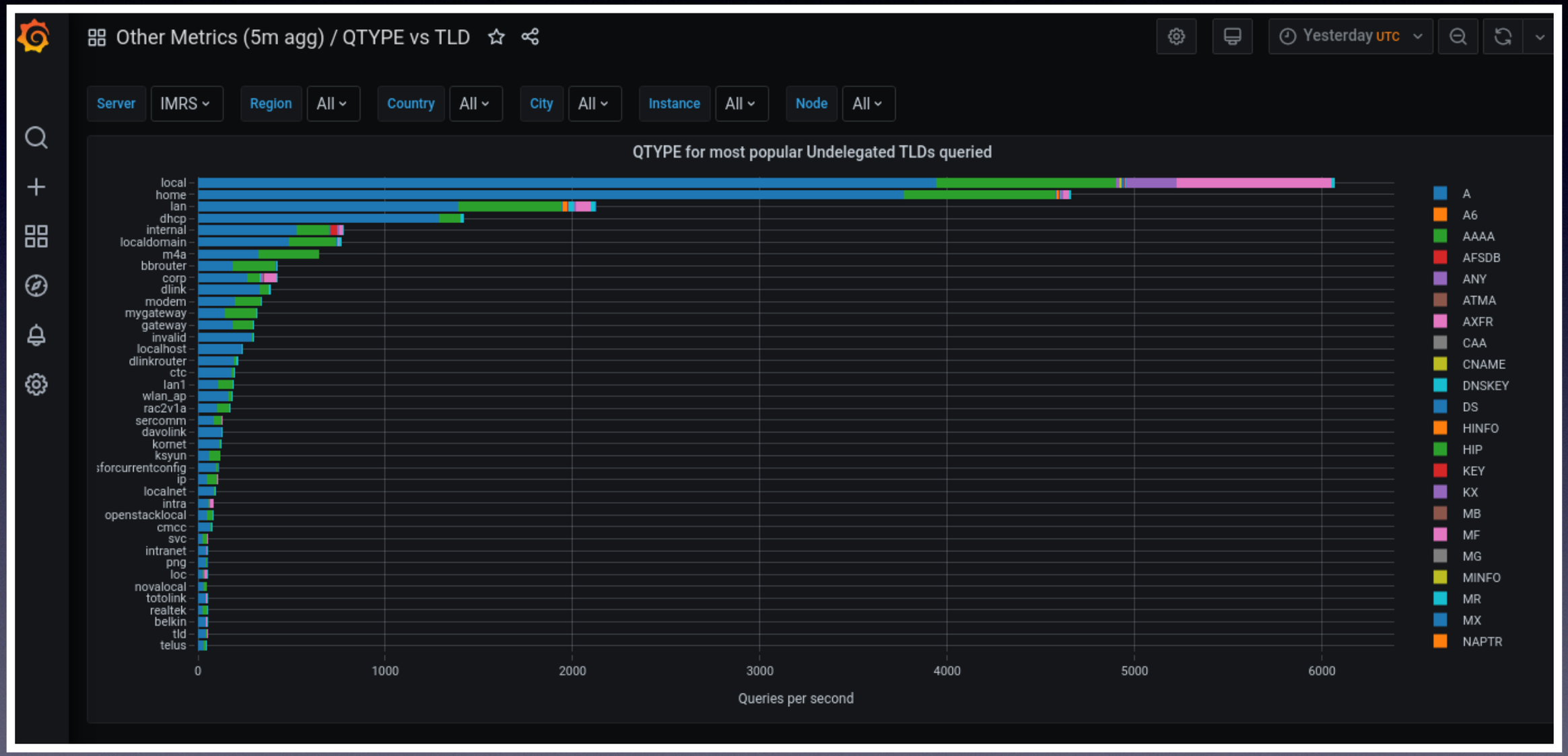
Simple bar chart

Client subnet statistics

inspector template output modifiers provide geo location and ASN lookup with MaxMind GeoLite data

More complex bar chart

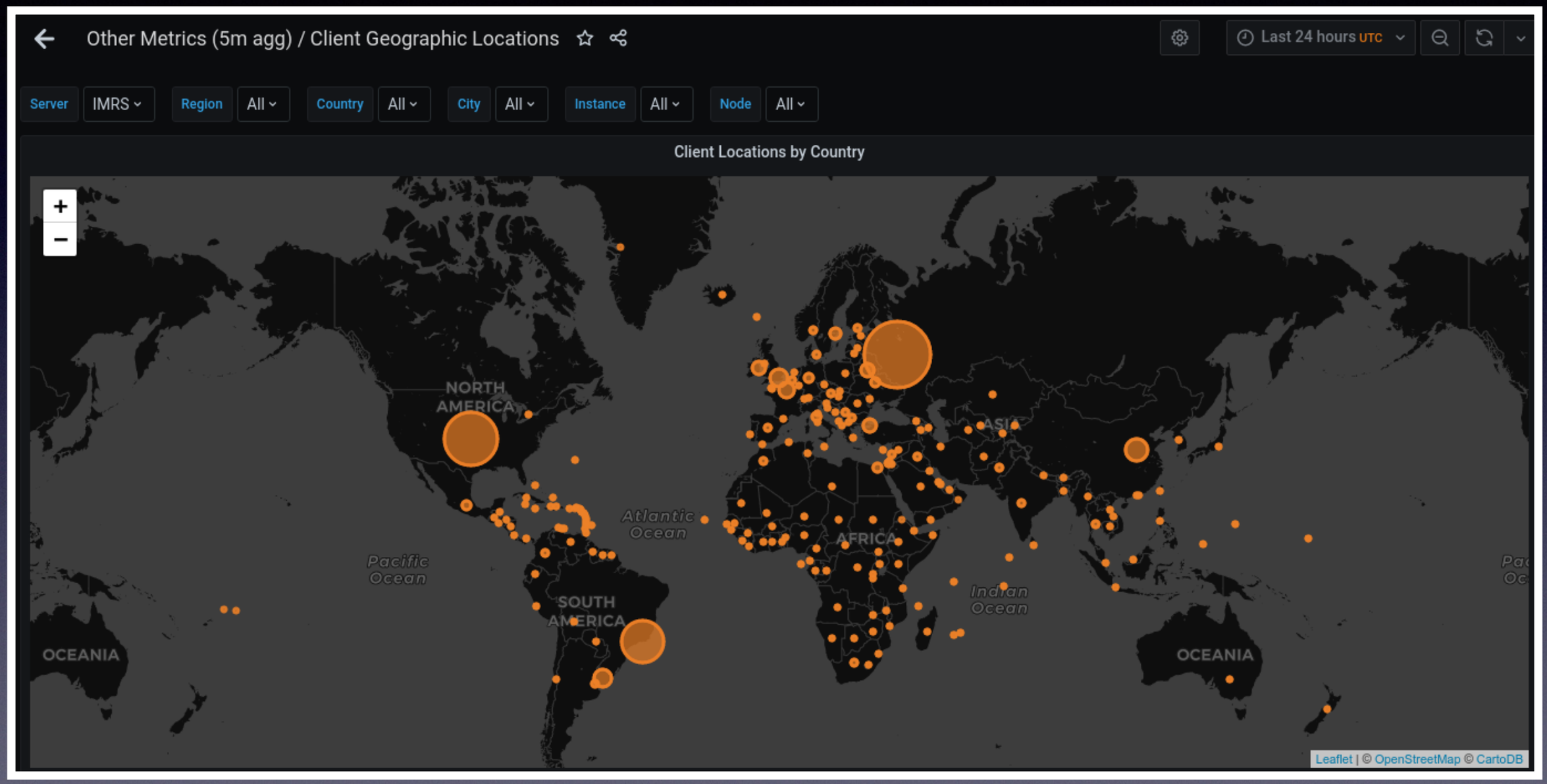
QTYPE vs TLD



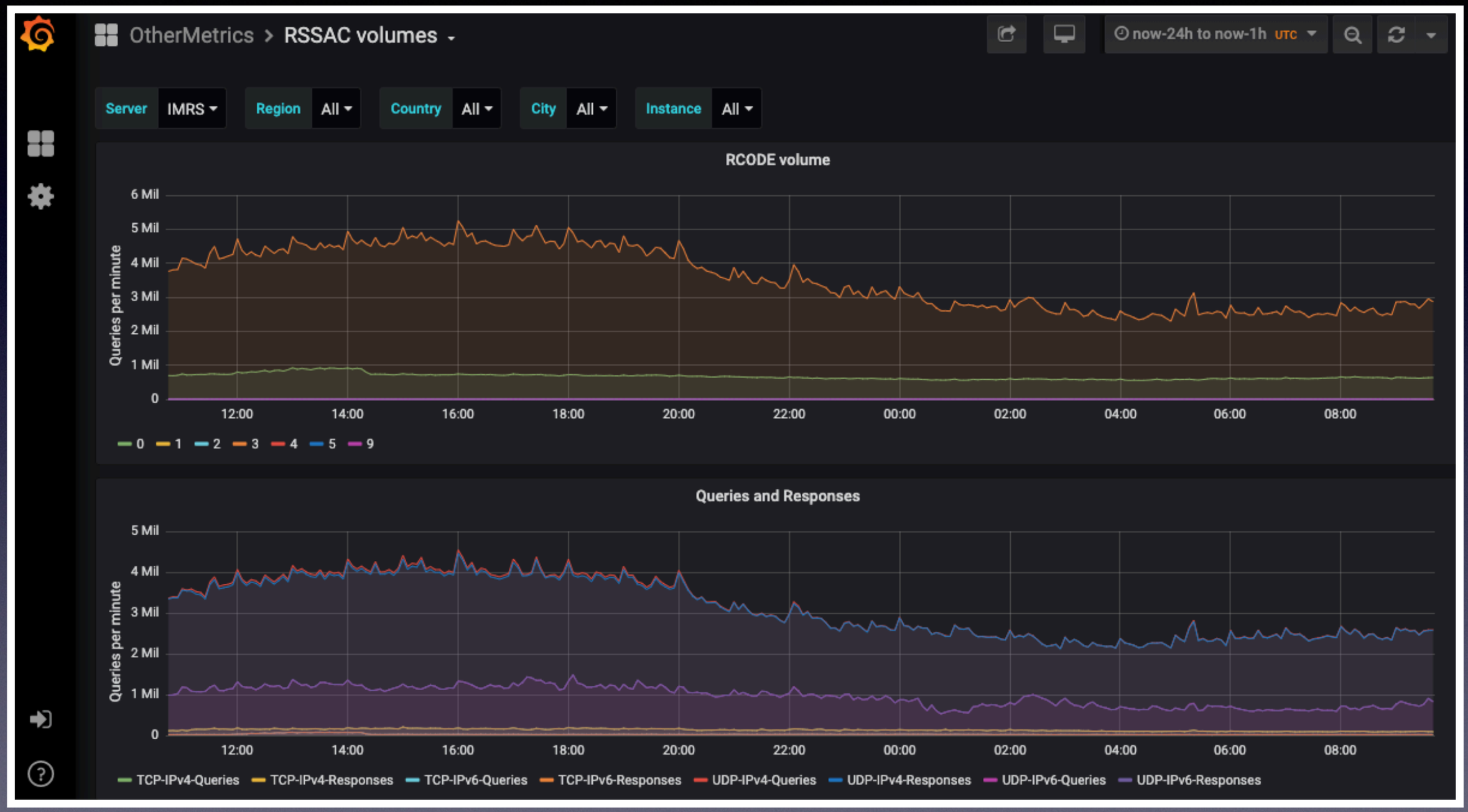
Using Sinodun modified plugin based on Plotly

Map based graph

Client geographic locations



RSSAC graphs



RSSAC reports generated by management tools

Summary

- C-DNS, ClickHouse and Grafana provide nice package for traffic capture and visualisation
- ClickHouse aggregations allows for flexibility to choose trade-offs between storage and performance
- Grafana can reproduce DSC like graphs with the right plugins…
- Management tools and schema not yet open sourced… will be shortly!

Thank you!

Any questions?

Offline questions to either

- SaNE (noc@dns.icann.org) or
- Sinodun (jim@sinodun.com)