Atlas, K-root and RIS updates

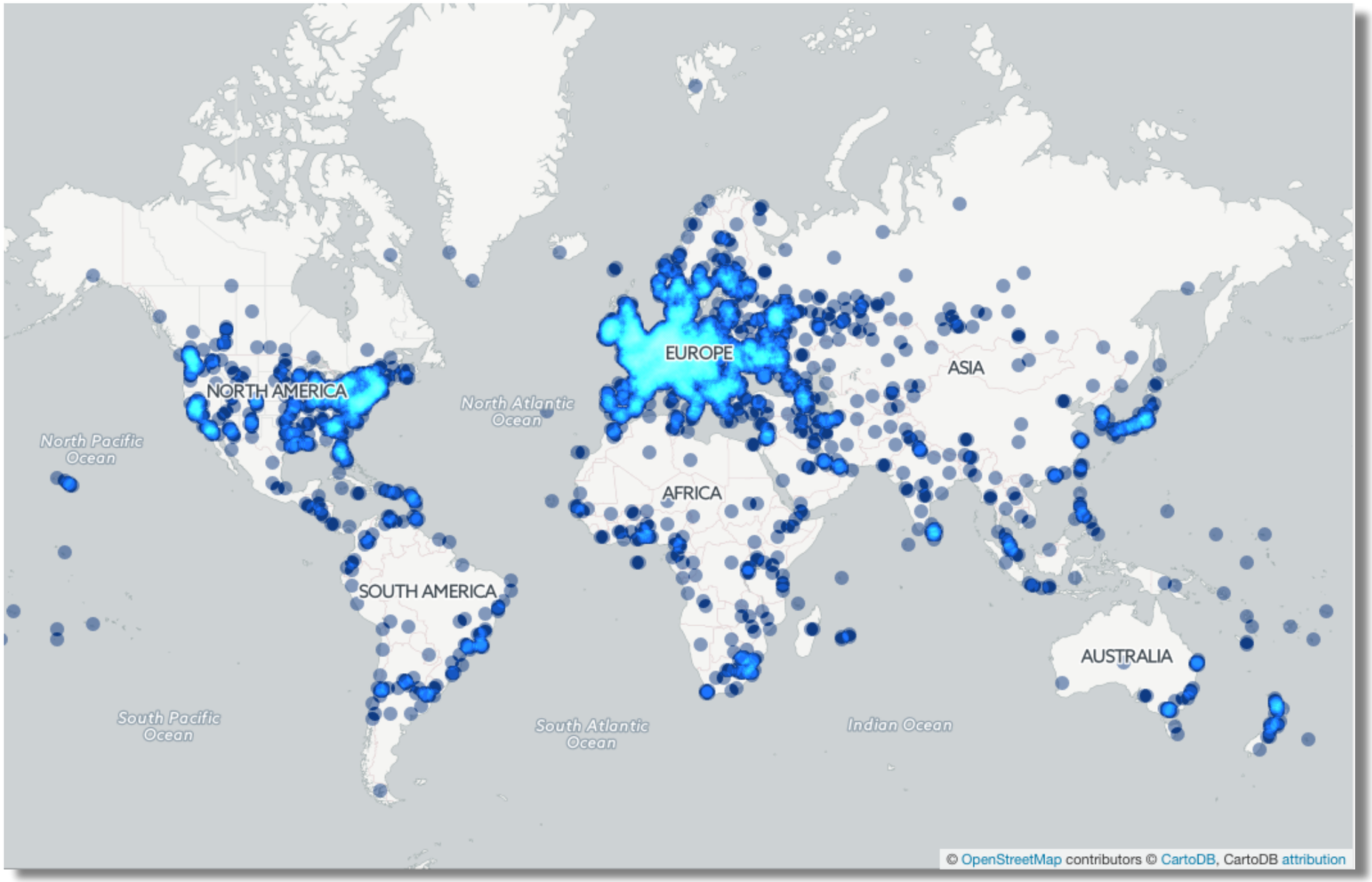
Iñigo Ortiz de Urbina Cazenave

ESNOG16 | Barcelona - Spain | 17 September 2015

RIPE Atlas

RIPE Atlas

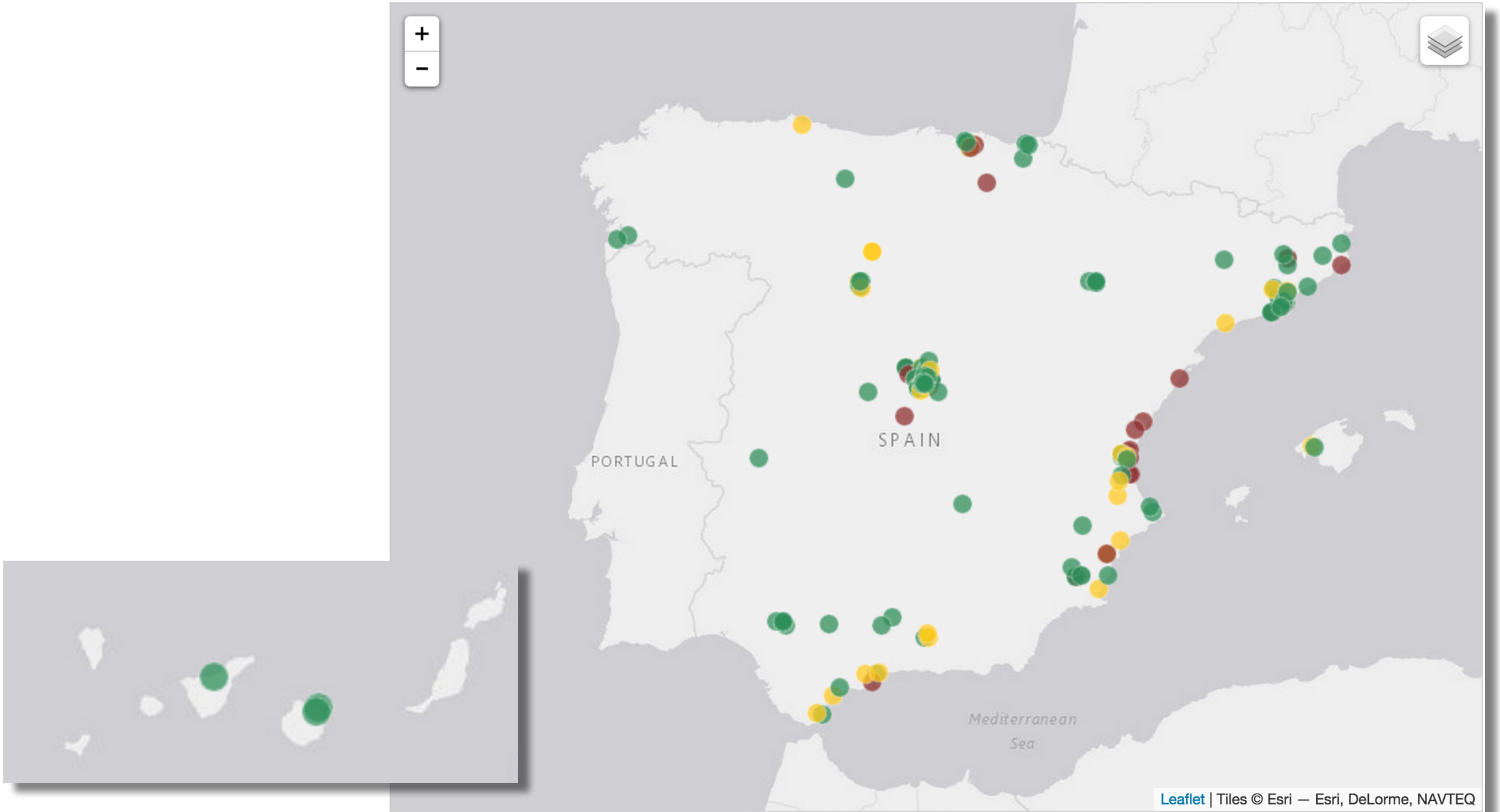
- Internet data plane monitoring **for** the community, **by** the community



https://astrikos.cartodb.com/viz/d9cdef7c-465a-11e5-83b0-0e4fddd5de28/public_map

RIPE NCC

RIPE Atlas in Spain



Connected: 136 Disconnected: 55 Abandoned: 33

Logistics and growth

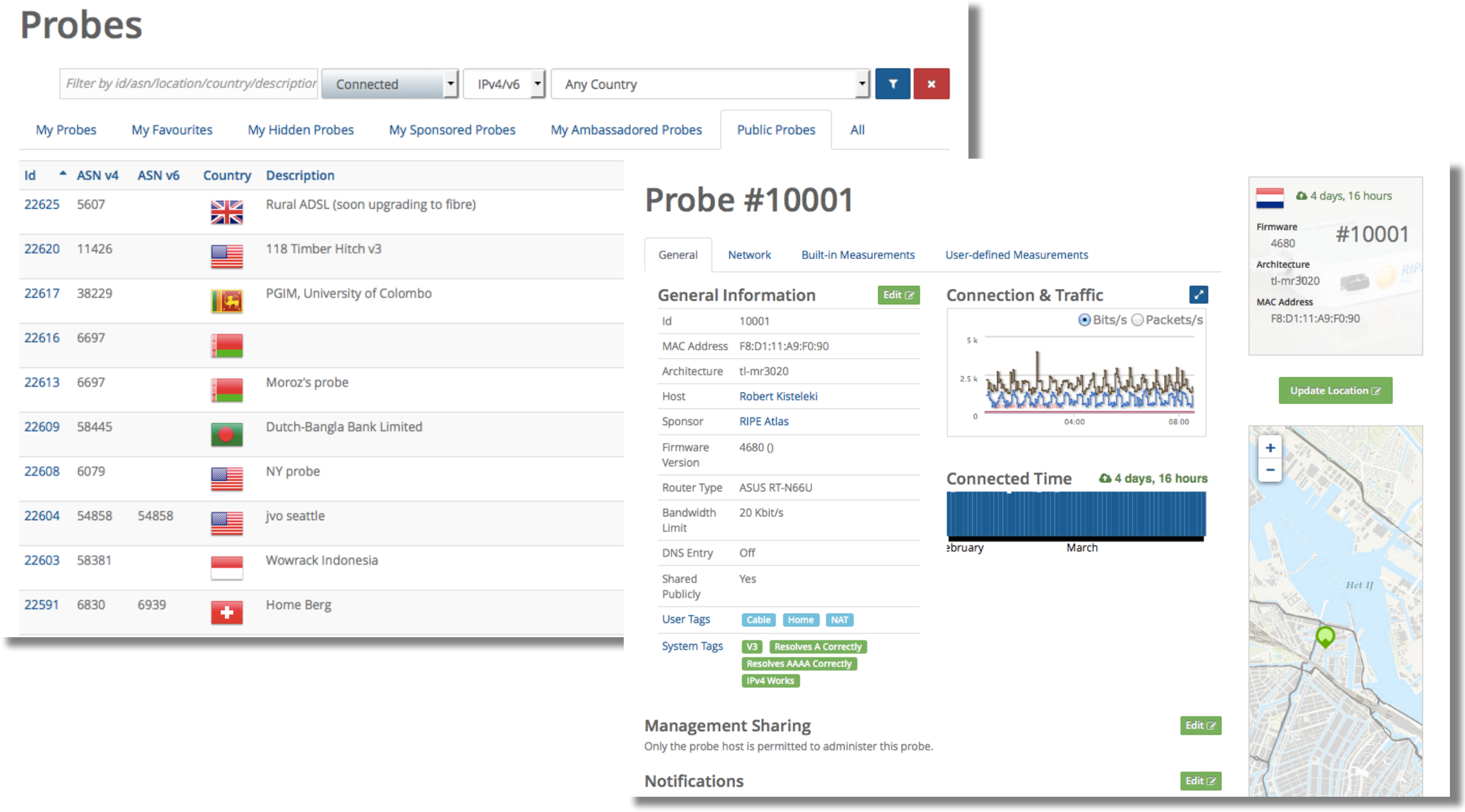
- Change in the distribution strategy
  - Focusing on RIPE NCC members without probes
  - Cooperation with other RIRs to reach out to their members
  - Only buying probes using sponsorship money
- Total:
  - 8.600+ active probes, 14.000 distributed
  - 135 active anchors, 200+ applications
  - almost 1.000.000 measurements in total!!!
    - 35.000 user-defined measurements weekly
- Growth:
  - One anchor activated, two new applications every week

New features

- New measurement types: HTTP, NTP, TLS
- Data streaming: results & probe connection status
- Better UIs and APIs
- Probe tagging

- Interesting use cases
  - Are the local paths staying local? What is the impact of IXPs? IXP-country-Jedi
  - Visualising network outages

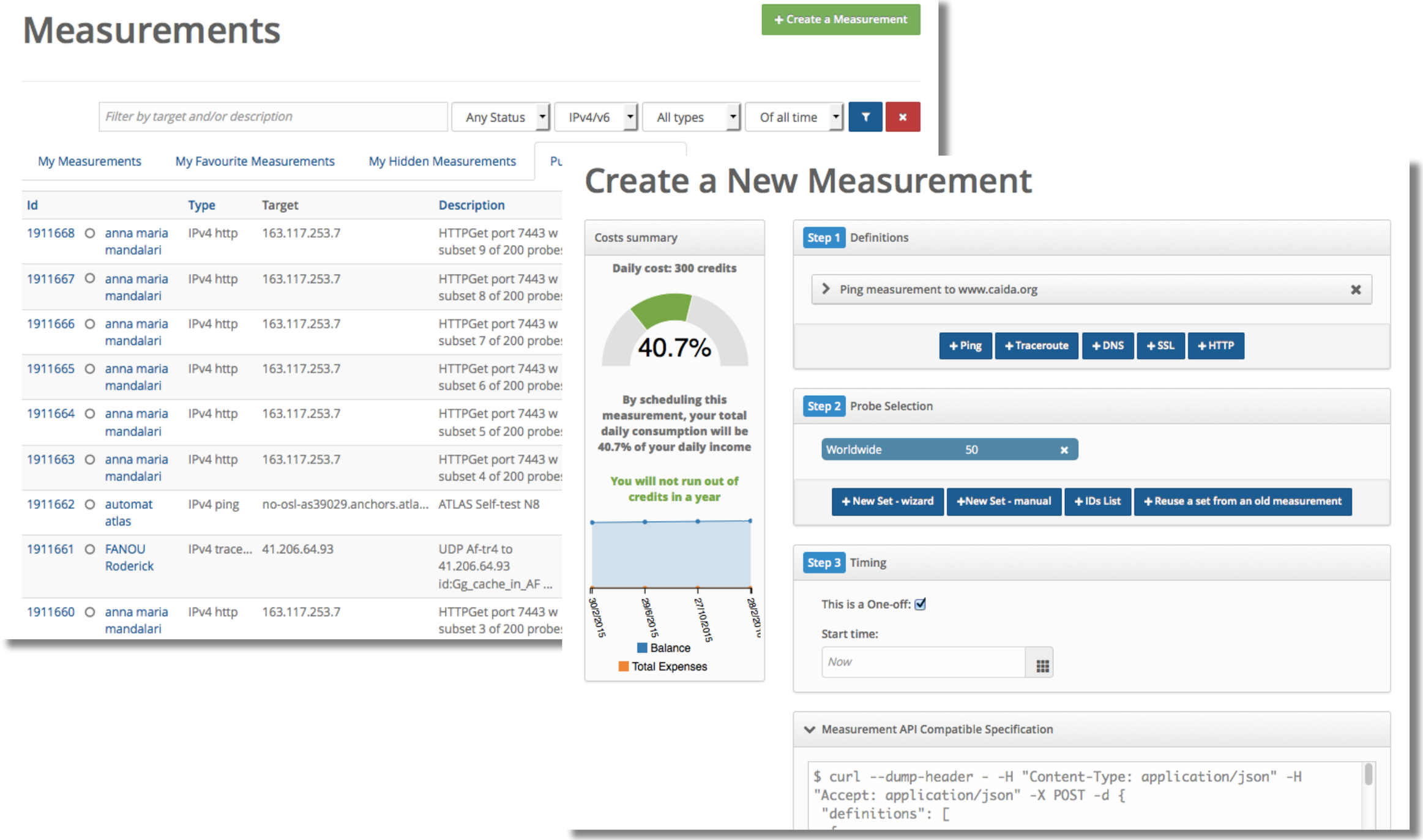
Better probe UI



- See https://atlas.ripe.net/probes/

RIPE NCC

Better measurement UI

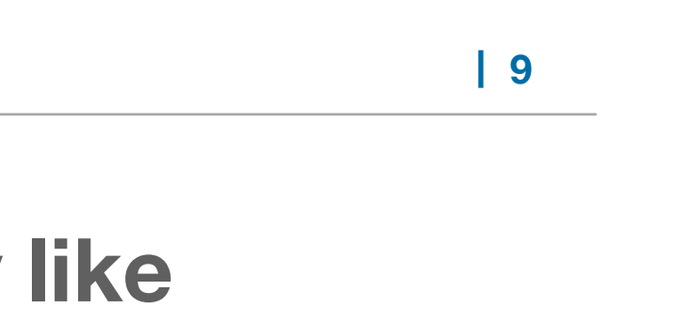


- See https://atlas.ripe.net/measurements/

RIPE NCC

Using probe tags

- Users can tag their probes **any way they like**
  - The commonly used tags are available to everyone
- The **system** also tags them **automatically**:
  - (Non)working IPv6, IPv4, DNS (A/AAAA)…
- **Use these tags** when scheduling measurements:
  - Measure from home or data centre probes
  - Measure from broken or working IPv6 probes
  - Combine this with other filters (eg. country)



- See https://atlas.ripe.net/docs/probe-tags/

New measurement types

- HTTP: query web servers (anchors only)
- NTP: query Network Time Protocol (NTP) servers
  - https://labs.ripe.net/Members/philip_homburg/ntp-measurements-with-ripe-atlas
- TLS check
  - Check for protocols, ciphers, certificates…

APIs, APIs, APIs

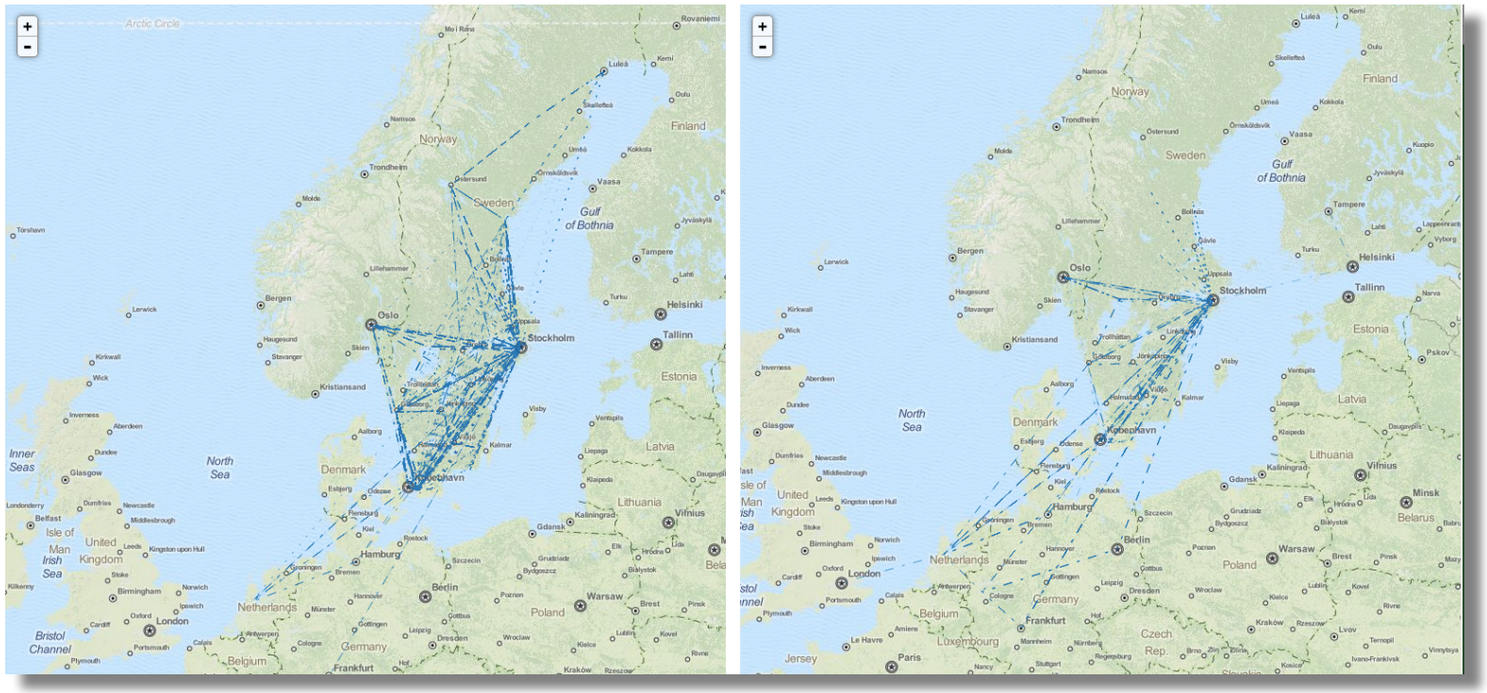
- Measurement API:
  - Query/search, create, change, stop, …
  - Download results, latest results, state checks, …
  - **Parse results**: https://atlas.ripe.net/docs/sagan/
- Probe API: query/search, probe archive (bulk access)
- **Result streaming**: results and probe connections
- See https://atlas.ripe.net/docs/

Data streaming APIs

- Data result streams
  - Real-time access to data (“drinking from the firehose”)
  - Can listen to the incoming data of public msms(s)
  - WebSocket clients + legacy support using polling
  - Allows for really cool visualisations
  - Has short-term memory and can also **replay historical data**, optionally at faster or slower speed (bullet time for RIPE Atlas data - yay!)
- Probe connection streams
  - Similar to results, but about probe connections/disconnections
  - Annotated by ASN/prefix/country…
- See https://atlas.ripe.net/docs/result-streaming/

Are local paths staying local?

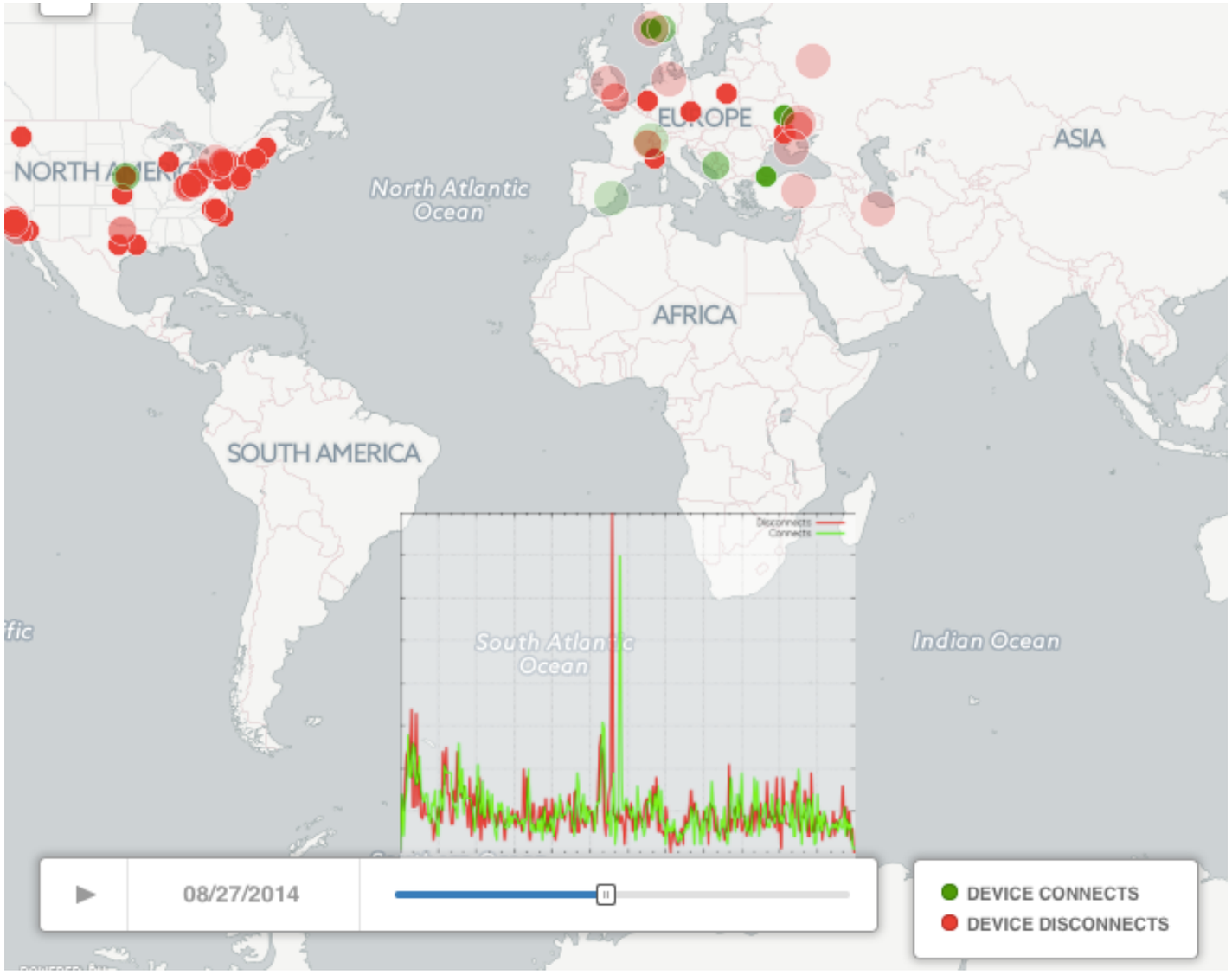
- IXP-Country-Jedi



- https://labs.ripe.net/Members/emileaben/measuring-ixps-with-ripe-atlas
- https://labs.ripe.net/Members/emileaben/measuring-countries-and-ixps-in-the-see-region
- https://github.com/emileaben/ixp-country-jedi
- https://github.com/RIPE-Atlas-Community/openipmap

RIPE NCC

Visualising network outages



- https://labs.ripe.net/Members/andreas_strikos/amsterdam-power-outage-as-seen-by-ripe-atlas
- https://labs.ripe.net/Members/emileaben/visualising-network-outages-with-ripe-atlas
- https://labs.ripe.net/Members/emileaben/facebookdown-and-what-internet-data

Plans for the rest of the year

- New measurement types
  - http measurements towards anchors
  - WiFi probe
- APIs for anchors, anchoring measurements
- Data streaming access to historical data
- Improve on OpenIPMap
- Security review
- Webinar coming up this summer
- Expansion goals: 150 anchors, 10,000 active probes

http://roadmap.ripe.net/ripe-atlas/

Learn more and get in touch

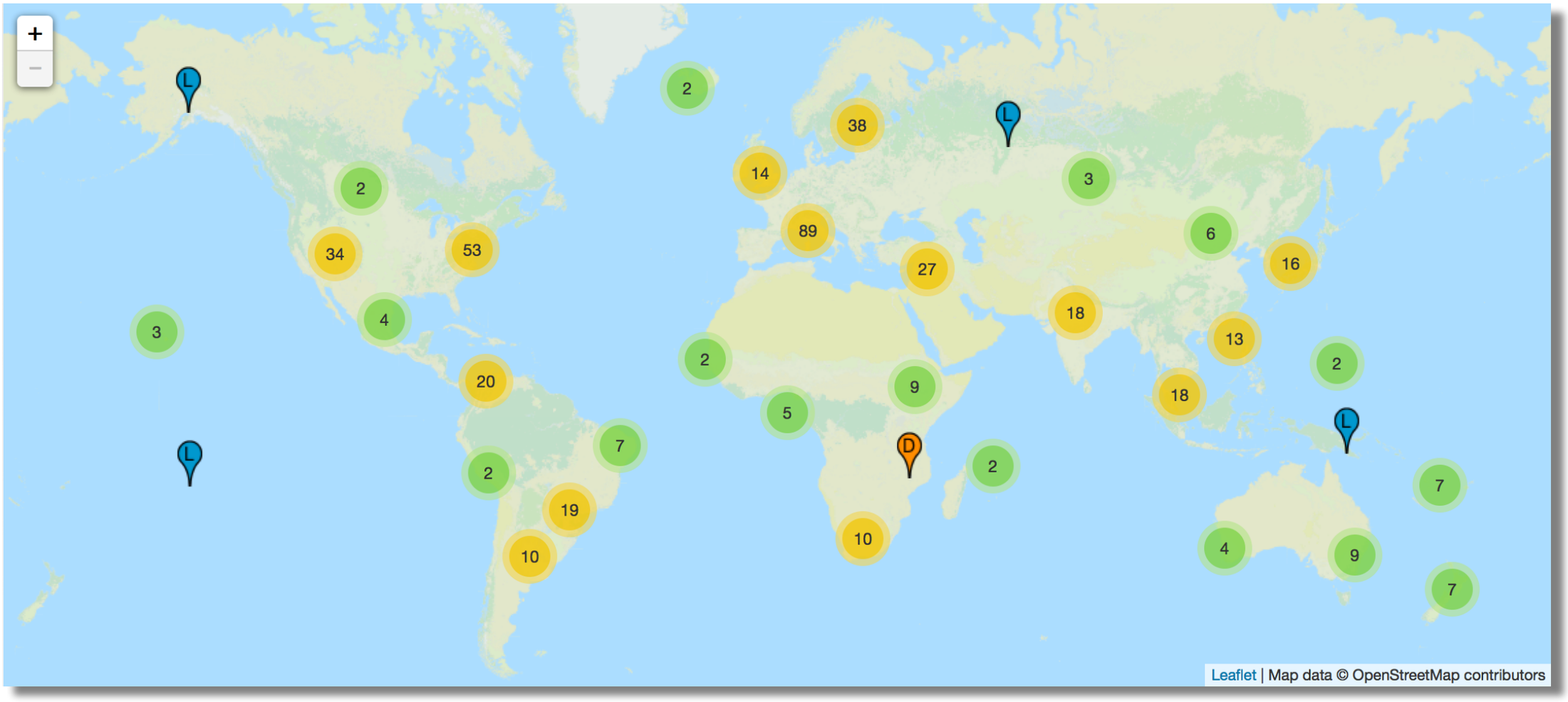
- Mailing list: ripe-atlas@ripe.net
- Blog: https://labs.ripe.net/atlas
- Twitter: @RIPE_Atlas
- Tickets: atlas@ripe.net
- Everything: https://atlas.ripe.net

k.root-servers.net

RIPE NCC

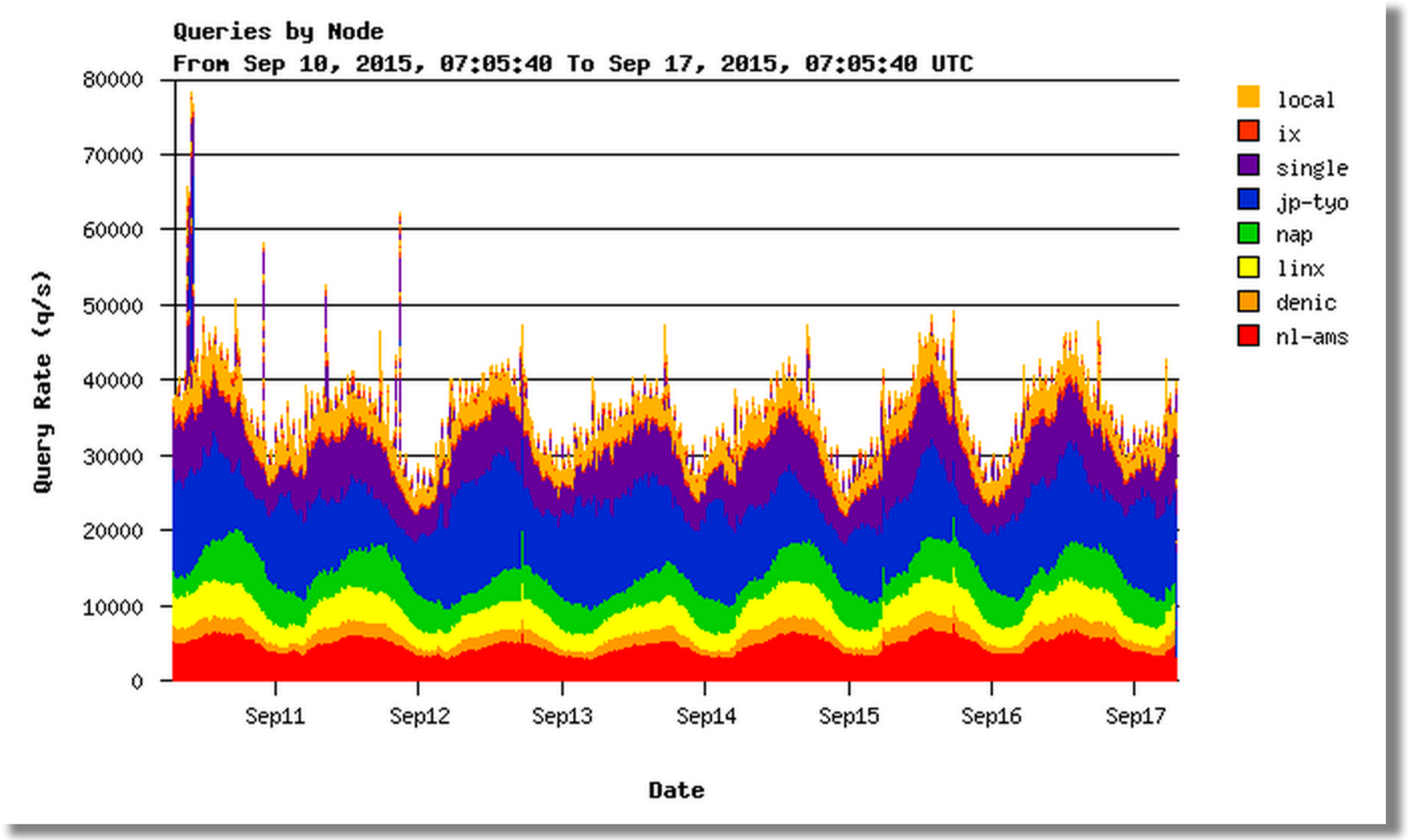
K-root (AS25152)

- One of the 13 Internet root name servers
- Operated by RIPE NCC since 1997



K-root (AS25152)

- One of the 13 Internet root name servers
- Operated by RIPE NCC since 1997



K-root (AS25152)

- Service provided by a set of distributed nodes anycasting:
  - 193.0.14.0/23
  - 2001:7fd::/32
- Each site, one or more servers running:
  - BIND
  - Knot
  - NSD

K-root (AS25152)

- RSSAC002 metrics
  - ICANN's Root Server System Advisory Committee
  - Available at www-static.ripe.net/dynamic/rssac002-metrics/
  - Currently publishing
    - load-time
    - rcode-volume
    - traffic-sizes
    - traffic-volume
    - unique-sources
    - zone-size

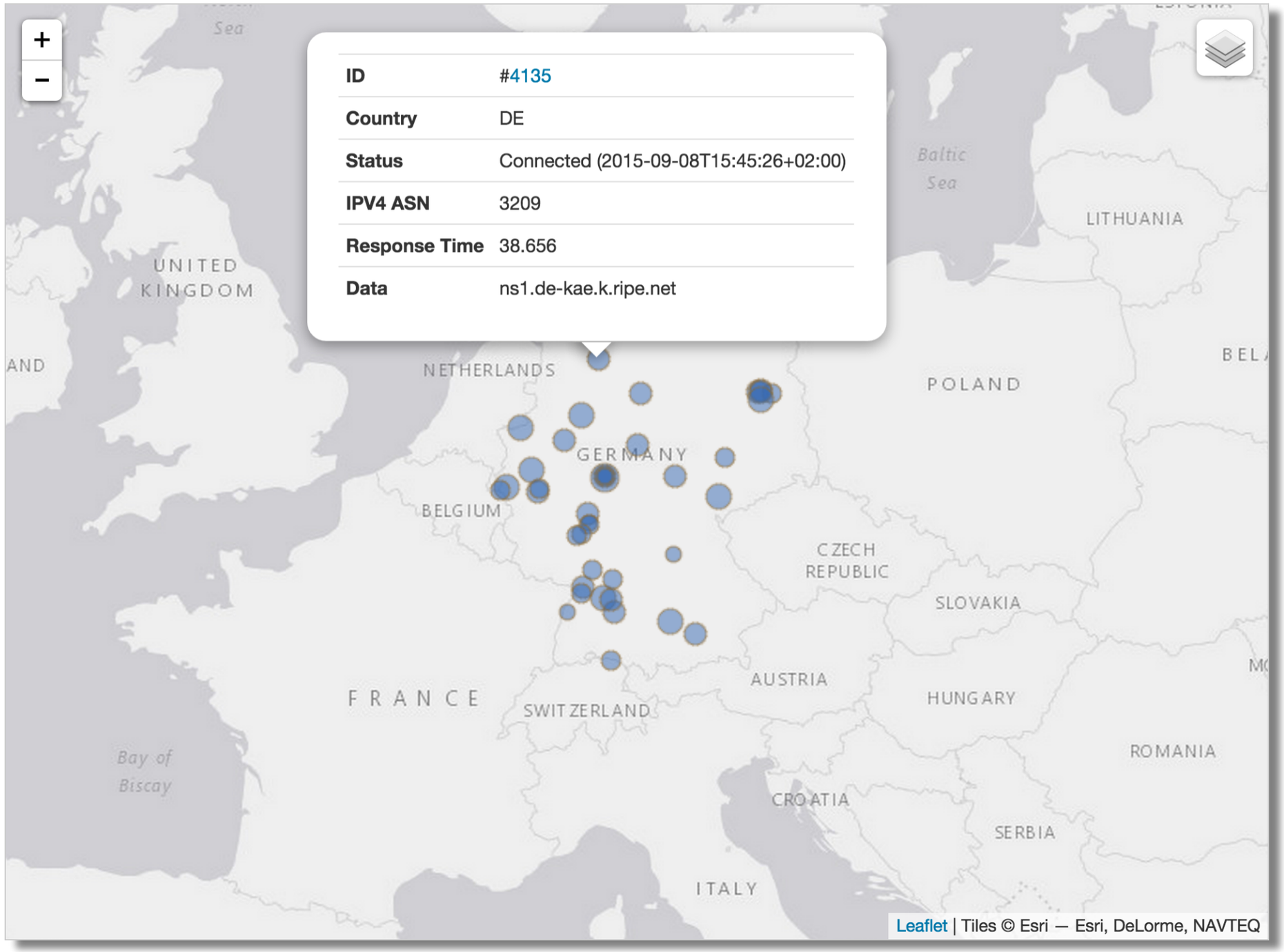
K-root (AS25152)

- Traditionally, two types of sites
  - Global
    - 5 sites: DE, JP, NL, UK, US
    - 3 servers per site
    - Router, switch, powerful hardware
  - Local
    - 10+ sites: AU, CH, IT, IS, PL, RU…
    - Router, modest hardware
    - Smaller footprint

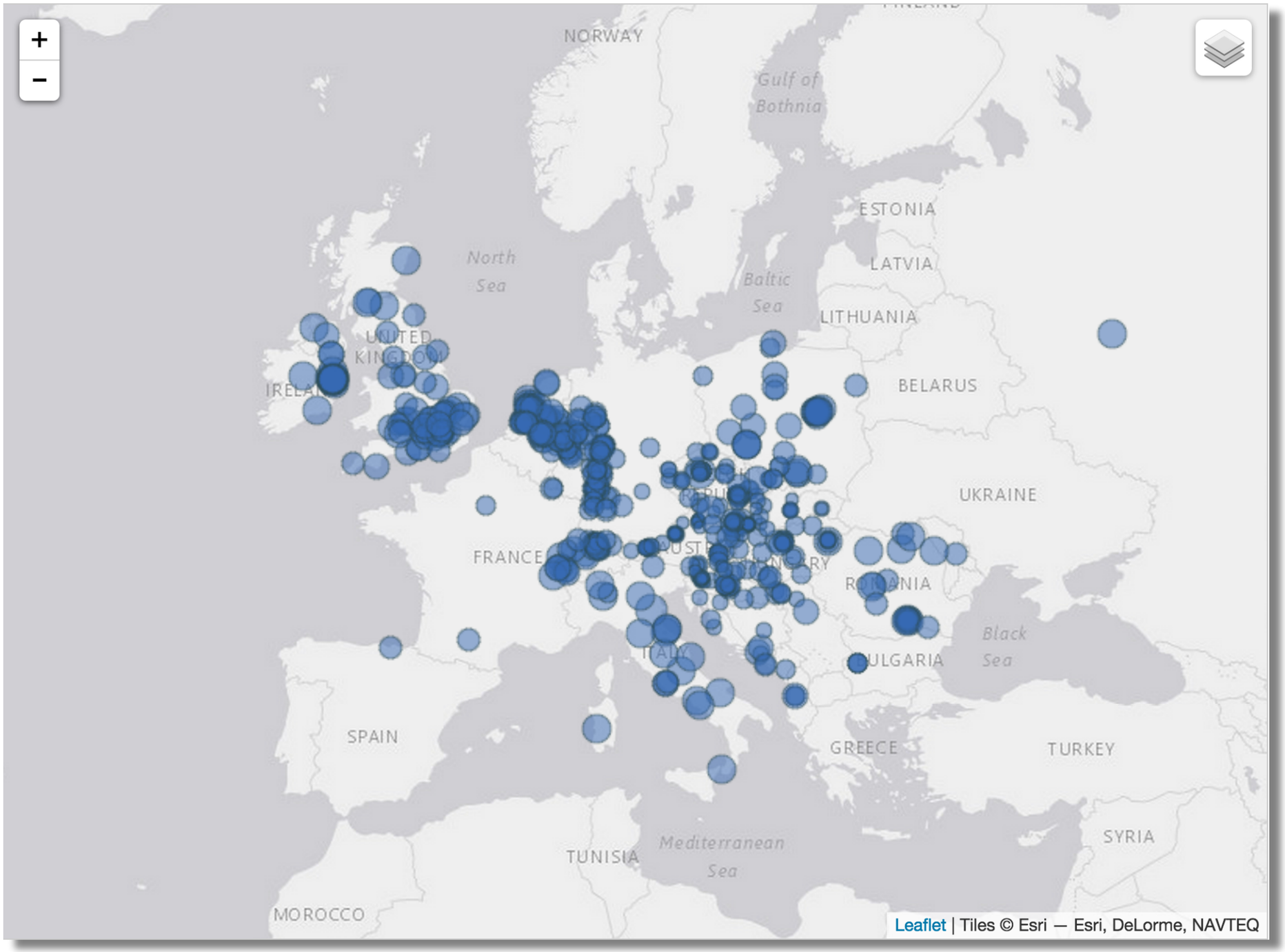
K-root (AS25152)

- Expansion plans add two new types
  - Hosted
  - IX
- Simplified peering management
  - Hosted: default+FHRP, single BGP session
  - IX: transit feed, route servers
- Varied footprint

K-root (AS25152)



K-root (AS25152)



K-root (AS25152)

- Requirements to host a new K-root node
  - Fill in questionnaire
  - Sponsor hardware and colo
  - Sponsor connectivity
    - OOB
      - Management interface
      - Service interface
  - Sign MoU
  - Handover to RIPE NCC/GII

K-root (AS25152)

- Questionnaire
  - Contact information
  - Hosting facilities
    - Physical security
    - Power supply
  - Hardware and network
    - Bandwidth requirements
    - Prefix propagation plans (hosted)
    - Route servers (IX)
    - Pingable target
    - dig -{6,4} @k.root-servers.net ch txt hostname.bind

Server minimum requirements:

1. Dell PowerEdge R2xx family server (current model is R220). Anything better, such as R320 or R420 is also acceptable.
2. At least 16 GB of RAM
3. At least a dual-core processor (but we prefer more cores to faster processors, so quad-core is better)
4. At least 2 Gigabit Ethernet ports (the two on-board ones will be fine)
5. No operating system (we will install CentOS)
6. PERC H310 hardware RAID adaptor
7. 2 x 500 GB 3.5" 7.2K RPM SATA drives
8. Rack mount rails
9. Appropriate power cord for the region where the server will be installed
10. iDRAC 7 enterprise (NOT Express)

Routing Information Service

RIPE NCC

RIPE RIS (AS12654)

- Route Collector system
  - Control plane info **for** the community, **by** the community
  - Data for RIPEstat: https://stat.ripe.net/
- We’re doing an overhaul
  - https://ripe70.ripe.net/archives/video/112/

OLD	NEW
Quagga based	ExaBGP based
11+1 route collectors (RRCs)	3 route collectors
MRT files	JSON (+ MRT for backward compat)
No room for growth	Plenty room for growth!
production	beta

Old collectors

- 12 active collectors
  - 1 multihop collector in Amsterdam
  - 11 local collectors at IXPs around the world
- Quagga-based
  - Store BGP updates in MRT format every 5 minutes
  - Store BGP table dumps in MRT format every 8 hours
  - Provide looking-glass query via RIPEstat
  - Data archived since 1999
- Has been static for some time
  - Most recent collector added in 2008

Old collector scaling issues

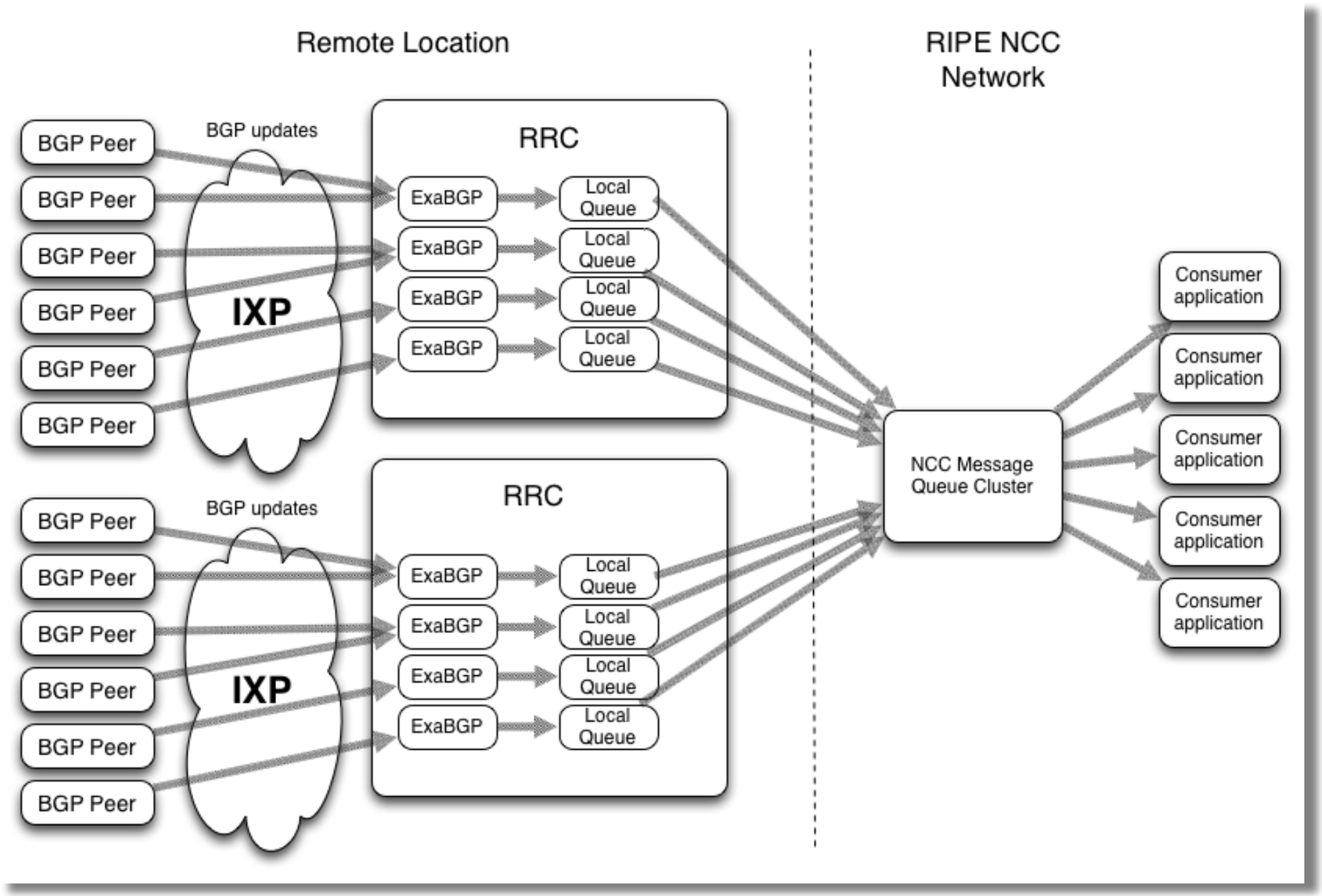
- Current Quagga implementation
- Single-threaded
  - Not as scalable on modern multi-core CPUs
- Locks updates during table-dump process
  - Requires that dump completes before the hold timer expires, or BGP session will drop
- Some data consistency issues
  - Sometimes updates are missing from the update dumps at the time of a table dump
  - This makes it difficult to accurately rebuild BGP state at an intermediate time, if updates are not reliable in-between

Ongoing work

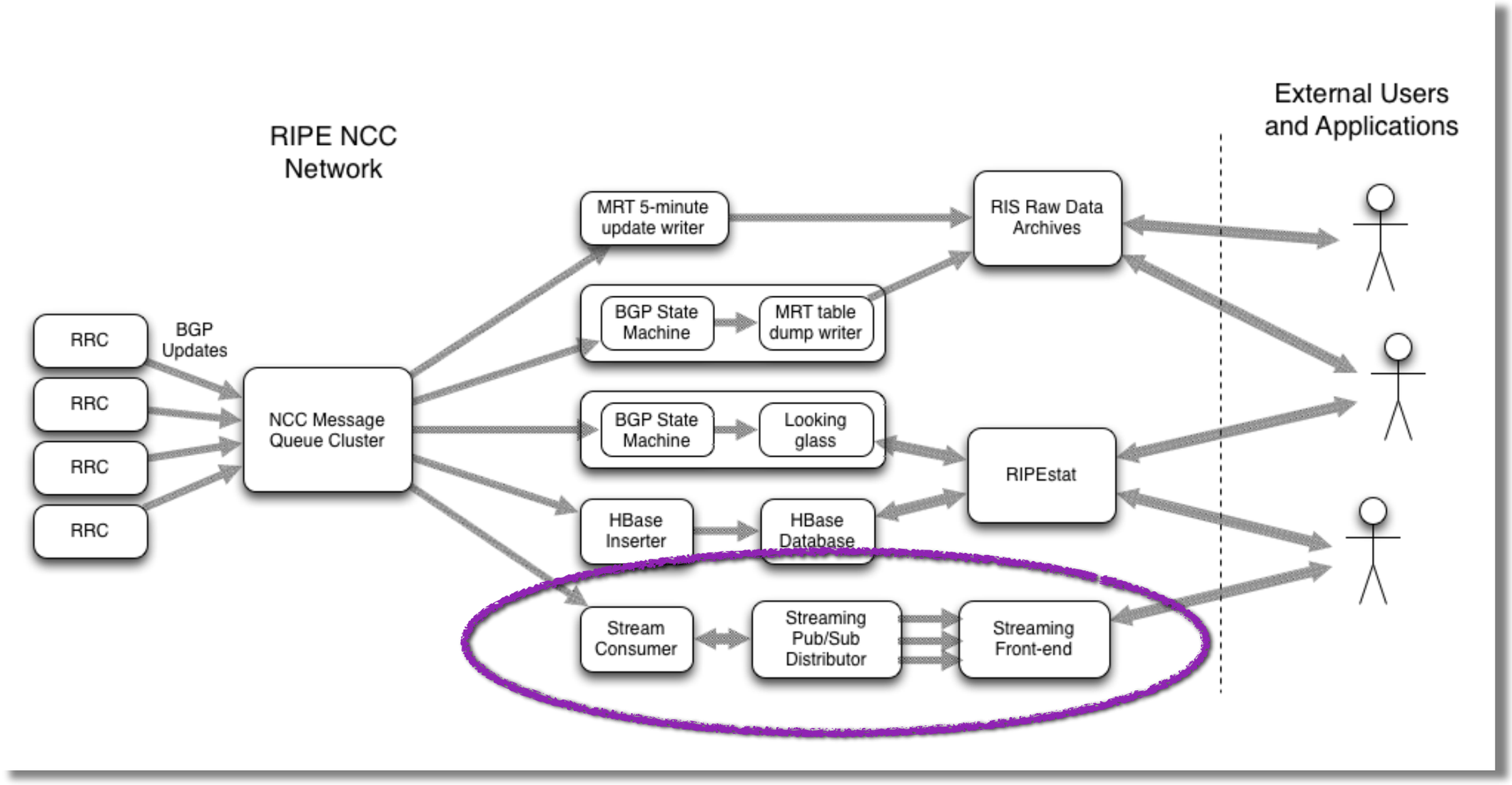
- Back-end replacement
  - Migration from old MySQL architecture
    - Scaling problems - needed a MySQL server per collector in some cases
    - Data retention - MySQL stored about 3 months per collector
  - Replaced with Hadoop
    - Horizontally-scalable processing and storage cluster
    - Map/Reduce performs data import, processing, and historical aggregations
    - HBase serves live queries from RIPEstat

“New RIS”

- Current status: In beta

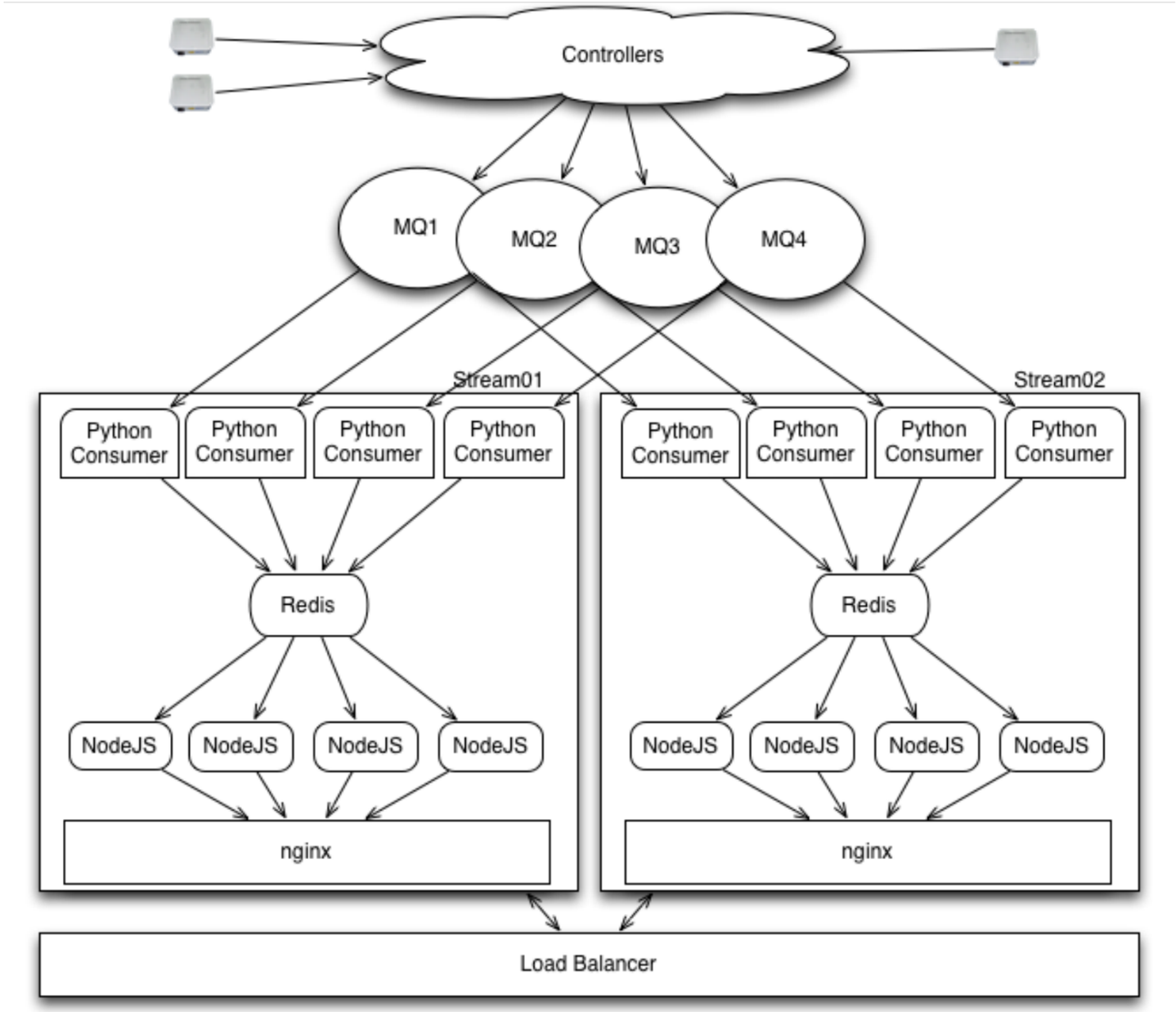


“New RIS”


RIPE NCC
Network
External Users
and Applications
RRC
RRC
RRC
RRC
BGP
Updates
NCC Message
Queue Cluster
MRT 5-minute
update writer
BGP State
Machine
MRT table
dump writer
BGP State
Machine
Looking
glass
HBase
Inserter
HBase
Database
Stream
Consumer
Streaming
Pub/Sub
Distributor
RIS Raw Data
Archives
RIPEstat
Streaming
Front-end

Streaming interface

- Same as the RIPE Atlas Streaming interface
- Diagram below from Massimo Candela's presentation on Monday Plenary session, RIPE 70



Use case: visibility (RIPEstat preview)

- Is your prefix/AS consistently visible?

Visibility Location Details of AS3333

RRC	IXP Location	Location	IPv4 peers seeing	IPv6 peers seeing	IPv4 Visibility	IPv6 Visibility
RRC00	RIPE-NCC Multihop	Amsterdam, Netherlands	15 of 15	10 of 10	100%	100%
RRC01	LINX	London, United Kingdom	10 of 10	9 of 9	100%	100%
RRC03	AMS-IX / NL-IX / GN-IX	Amsterdam, Netherlands	8 of 8	12 of 12	100%	100%
RRC04	CIXP	Geneva, Switzerland	7 of 7	4 of 4	100%	100%
RRC05	VIX	Vienna, Austria	6 of 6	7 of 7	100%	100%
RRC06	DIX-IE	Tokyo, Japan	1 of 1	1 of 1	100%	100%
RRC07	Netnod	Stockholm, Sweden	2 of 2	4 of 4	100%	100%
RRC10	MIX	Milan, Italy	10 of 10	7 of 7	100%	100%
RRC11	NYIIX	New York City, US	8 of 8	8 of 8	100%	100%
RRC12	DE-CIX	Frankfurt, Germany	15 of 15	20 of 20	100%	100%
RRC13	MSK-IX	Moscow, Russian Federation	12 of 12	5 of 5	100%	100%
RRC14	PAIX	Palo Alto, US	6 of 6	6 of 6	100%	100%
RRC15	PTTMetro	Sao Paulo, Brazil	11 of 11	8 of 8	100%	100%
RRC18	Test-RRC	Barcelona, Spain	2 of 2	1 of 1	100%	100%
RRC20	Test-RRC	Zurich, Switzerland	8 of 8	7 of 7	100%	100%
RRC21	Test-RRC	Paris, France	11 of 11	10 of 11	100%	91%

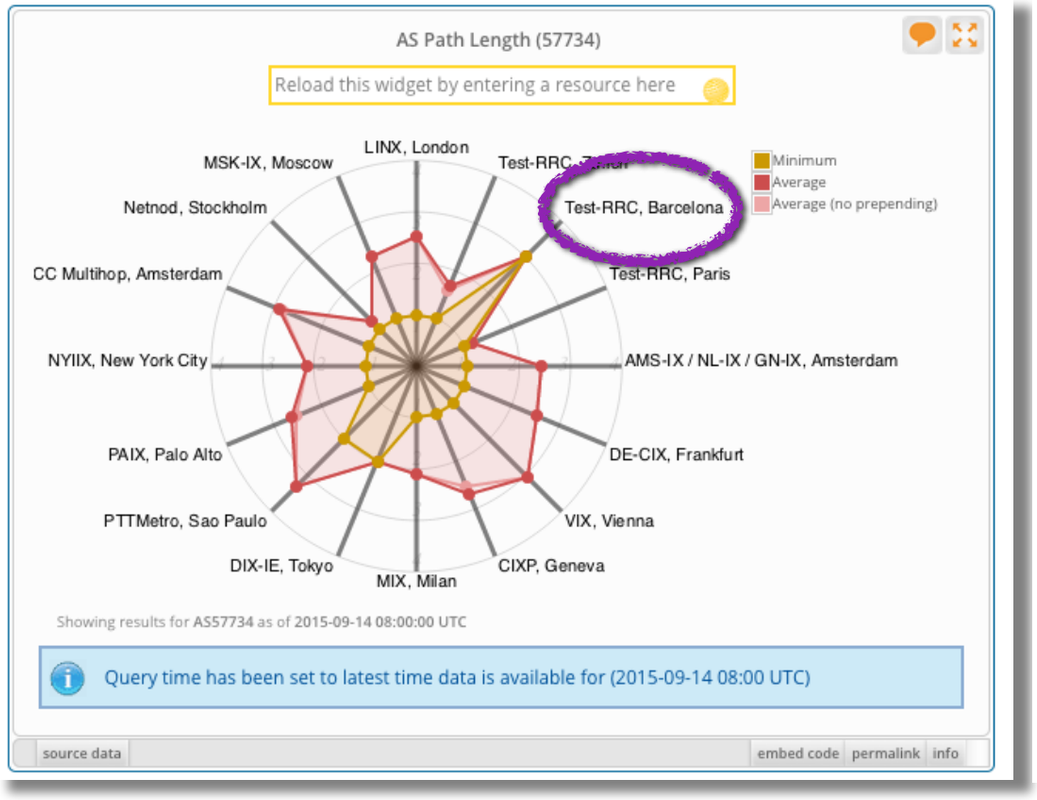
▼ List of RIS peers not seeing AS3333

Show 10 entries Search:

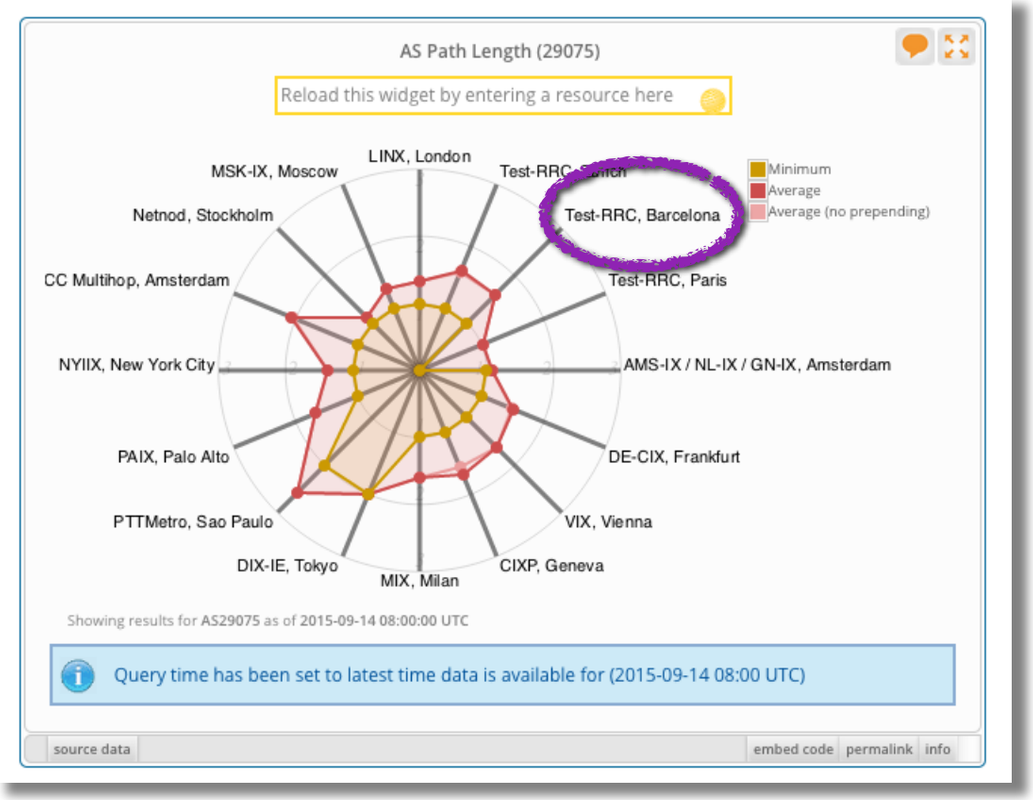
RRC	City	Peer AS	Peer IP	BGP Table	Prefix Count
RRC21	Paris, France	49463	2001:7f8:54::145	IPv6	23901

https://stat.ripe.net/widget/visibility

AS path length (RIPEstat preview)



FranceIX

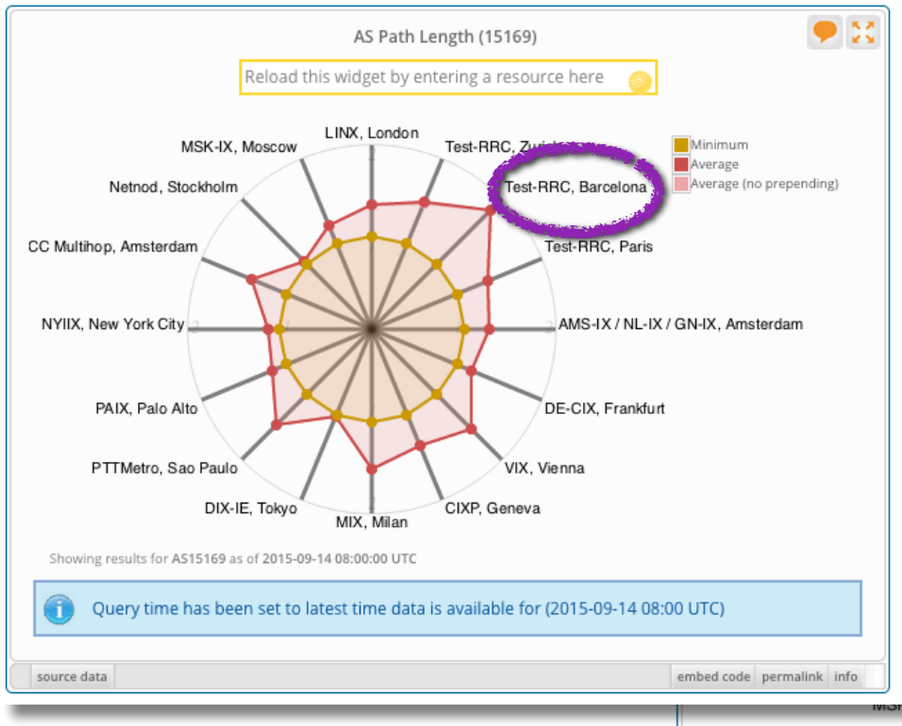


IELO

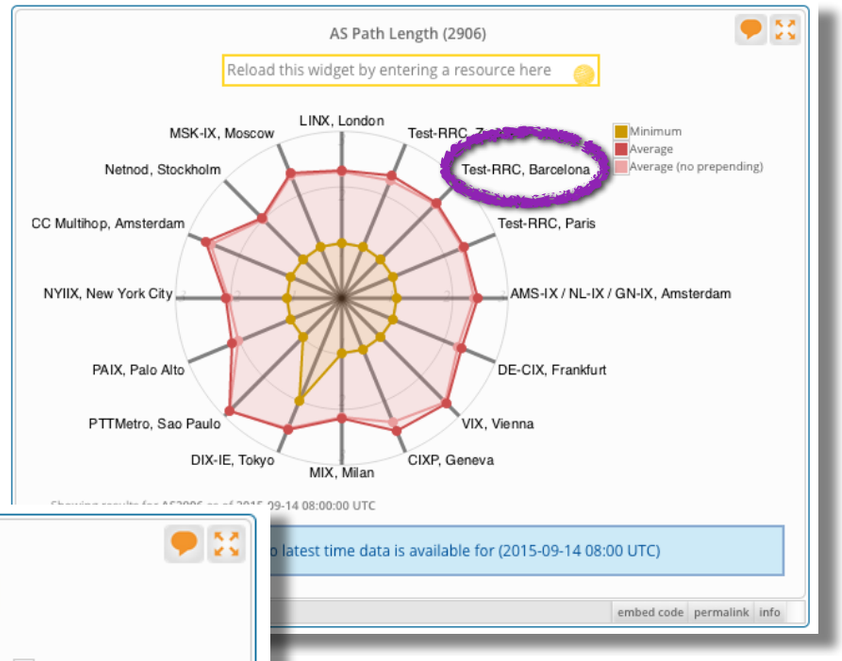
- Which location has short AS paths for a particular AS?

https://stat.ripe.net/widget/as-path-length

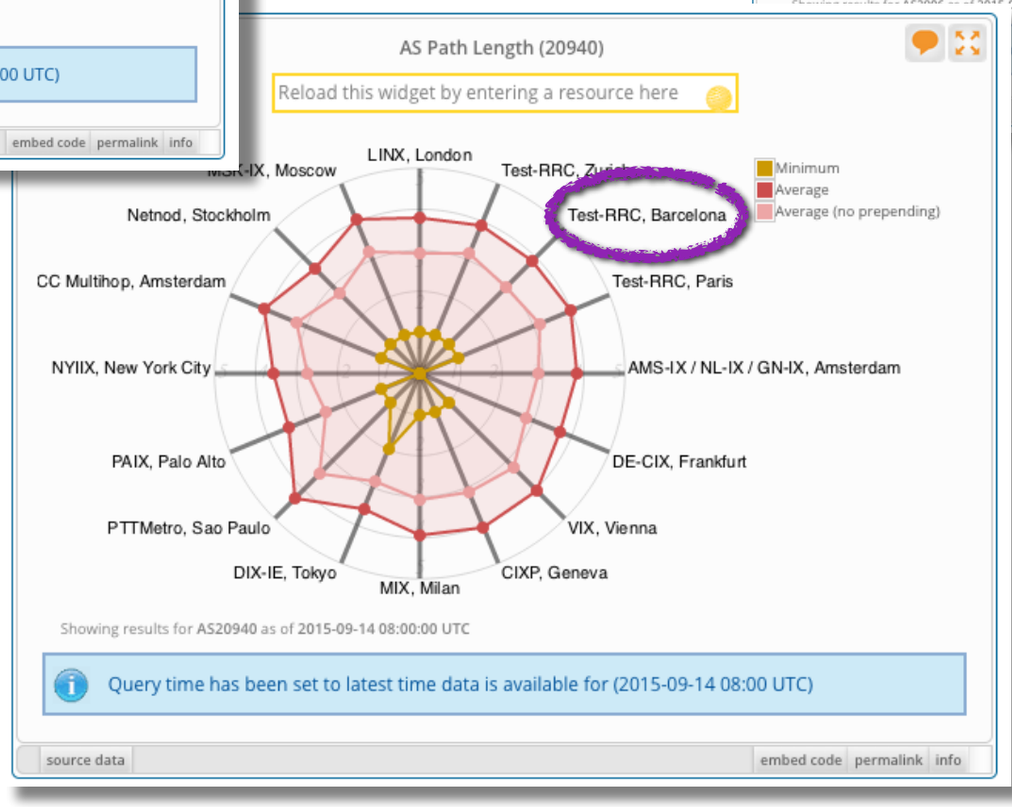
AS path length for Content



Google



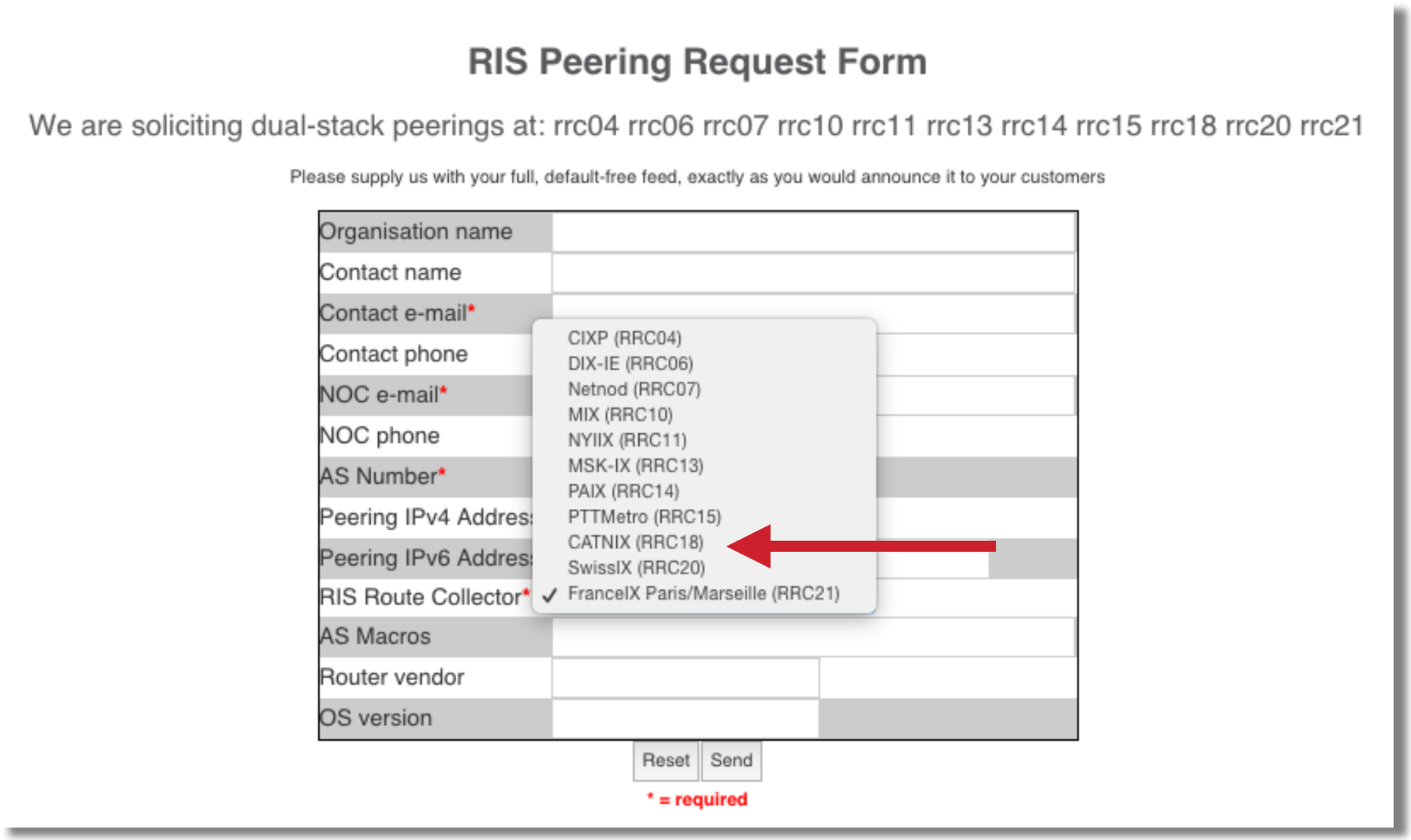
Netflix



Akamai

Come peer with us

- http://www.ris.ripe.net/cgi-bin/peerreg.cgi



What feed do you give us?

- In the past we've asked for "Full Table" (ie. as if you were giving RIS transit)
- Lots of RIS peers provide other feeds
    - Typically "peering"
- Result: Many different types of RIS peers
    - Harder to understand in data consumption/analysis
    - Potential fix for that in the works:
    - draft-ymbk-grow-bgp-collector-communities
        - differentiate "customer cone", "external", and "internal" routes, using additional BGP communities

Questions?