RIPE NCC
RIPE NETWORK COORDINATION CENTRE

# RIPE Atlas infrastructure & Measurement Results Sharing

Robert Kisteleki | CAIDA AIMS 2019

# Part I

## RIPE Atlas changes

# RIPE Atlas current status

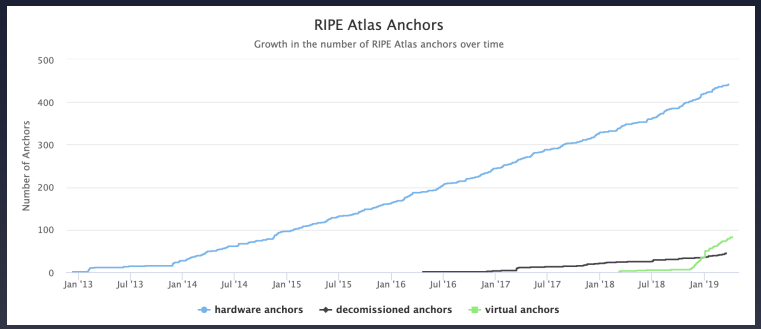
- About 10.500 probes are up
  - Of which 440 are anchors
    - Of which 80 are VMs
- Close to 20M measurements defined
  - About 20k running concurrently
- Collecting 7000-8000 data points / sec (~600M/day)

# RIPE Atlas current work items

- Use more client-side rendering relying more on APIs
- Strongly considering changing queues from RabbitMQ to Kafka
- Switching measurement metadata store from MySQL to ElasticSearch
- Switching main database from MySQL to PostgreSQL
- Python2 -> Python3 migration, other framework upgrades

# RIPE Atlas anchor VMs

- Pilot ran between May-October 2018
- Service is official since November 2018
- High interest since launch time
- Includes nodes in AWS
  - Talking to other cloud providers too
- ASN holder vs VM host needs attention
- VMs are virtually indistinguishable from HW anchors

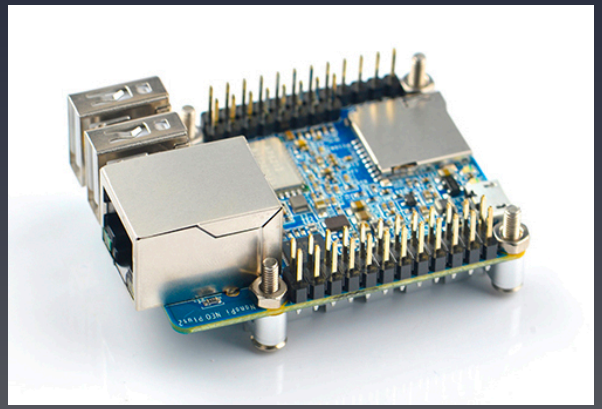


# RIPE Atlas software probes

- Probes are still hardware based (excl. anchor VMs)
- There's demand for an installable software package
- Would allow the next level of growth
- We're evaluating this at the moment
  - It has important consequences to operations
  - Now to encourage growth while encouraging spread?

# Other bits and blobs

- We're at version 4 of the probes (NanoPi based)
- The distributed nature brings some unique challenges
  - So does the volume of data we're handling



# Part II

## Measurement Results Sharing

# Downloadable results

- All public measurement results are available in bulk via FTP
  - Much easier/faster to download large amounts
  - Measurement metadata (via API) helps with indexing
- Download API is still the main path for individual msms
- Result streaming is still available and supported
  - Fully real-time
  - Has different availability characteristics

# RIPE Atlas data in the cloud

- Pushing data to BigQuery too
- Has unique advantages
- See a more detailed presentation by Stephen

# Measurement infrastructure collaboration

- “Can we lower the barrier for users (researchers or otherwise) to use multiple measurement systems without having to relearn or reimplement anything?”
- Different systems use dirrerent APIs and data formats
- Can we improve on this?
- Is this an incentive or a technology question?

# Questions