



RIPE NCC
RIPE NETWORK COORDINATION CENTRE

Developing the RIPE Routing Information Service (RIS)

Ondřej Caletka, Emile Aben | 8 September 2020 | CSNOG 2020

What is RIS?



- Collects BGP data from a diverse set of networks on the Internet
- Many interfaces (<https://ris.ripe.net>)
 - RIPEstat, RIS Live, raw data
 - Internet outages analyses on RIPE Labs
 - RPKI stats

Route collectors

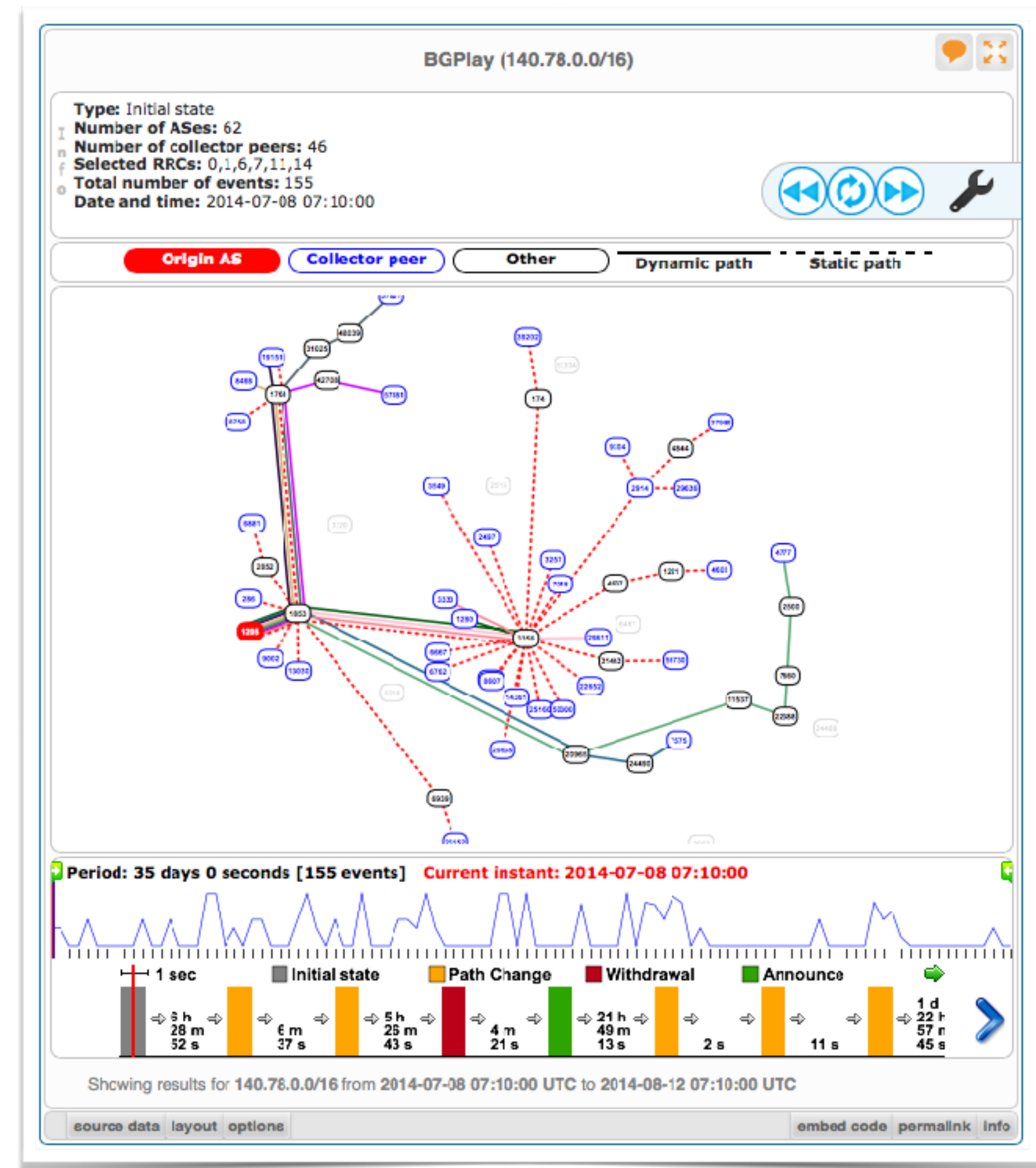


- In Internet Exchange Points
 - Peering within the IXP
 - Short and reliable path
 - 18 collectors around the world
- Multi-hop collectors
 - Long-distance BGP session over the Internet
 - Less stable and reliable
 - Two collectors: NL-AMS and UY-MVD



What is RIPLEstat?

- One-stop shop for lots of data about IP networks
- Presented using variety of widgets
- Many widgets **visualise RIS data**
- The most famous is **BGPlay**
- **JSON-based data API**



RIS Live



- Real-time public feed of BGP updates
- JSON over WebSockets
- Server-side filtering options
- Open for third-party utilities



BGPalerter APP 19:17

visibility

The prefix 165.254.225.0/24 (description 1) has been withdrawn. It is no longer visible from 4 peers.

visibility

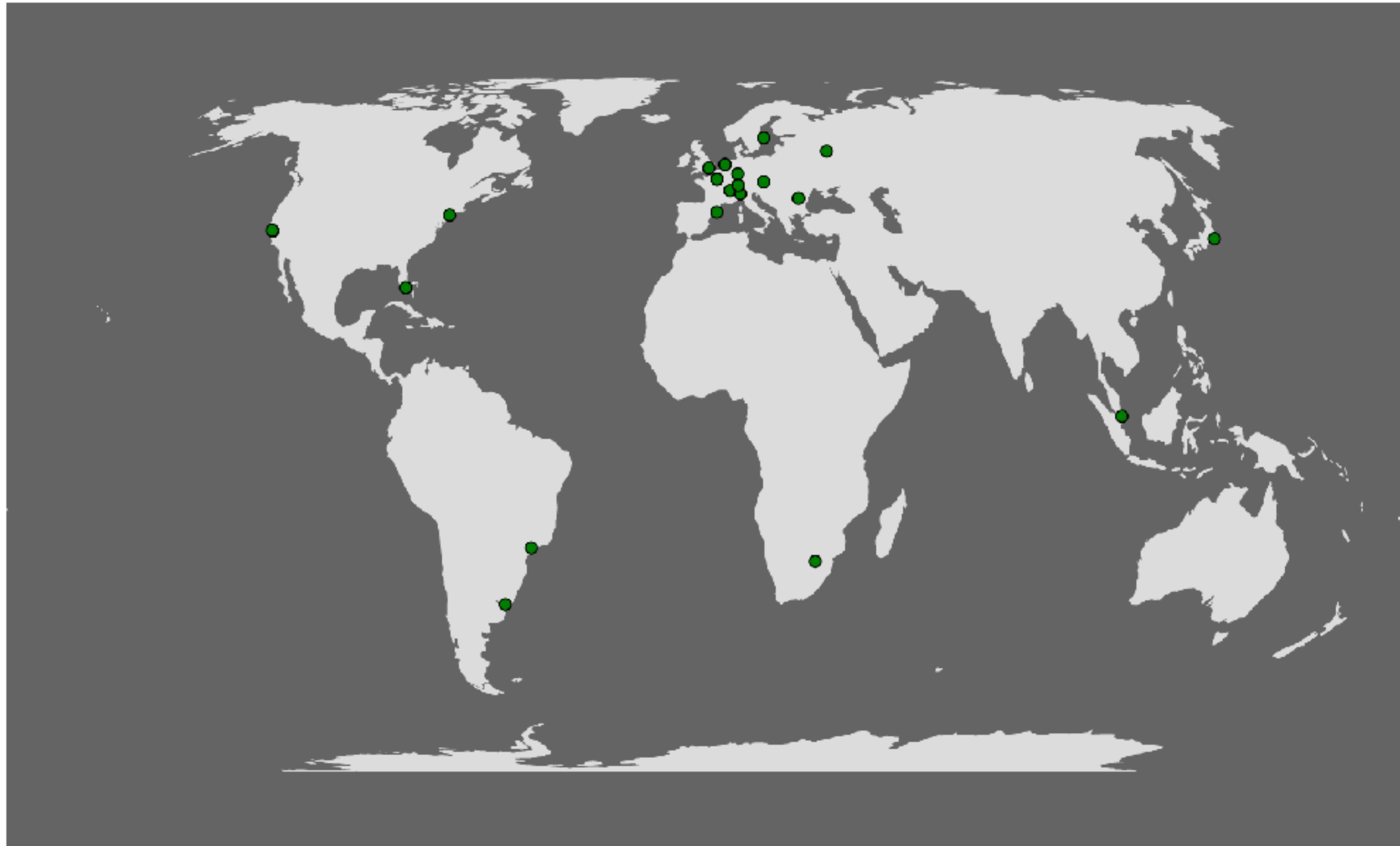
The prefix 2a00:5884::/32 (alarig fix test) has been withdrawn. It is no longer visible from 4 peers.

hijack

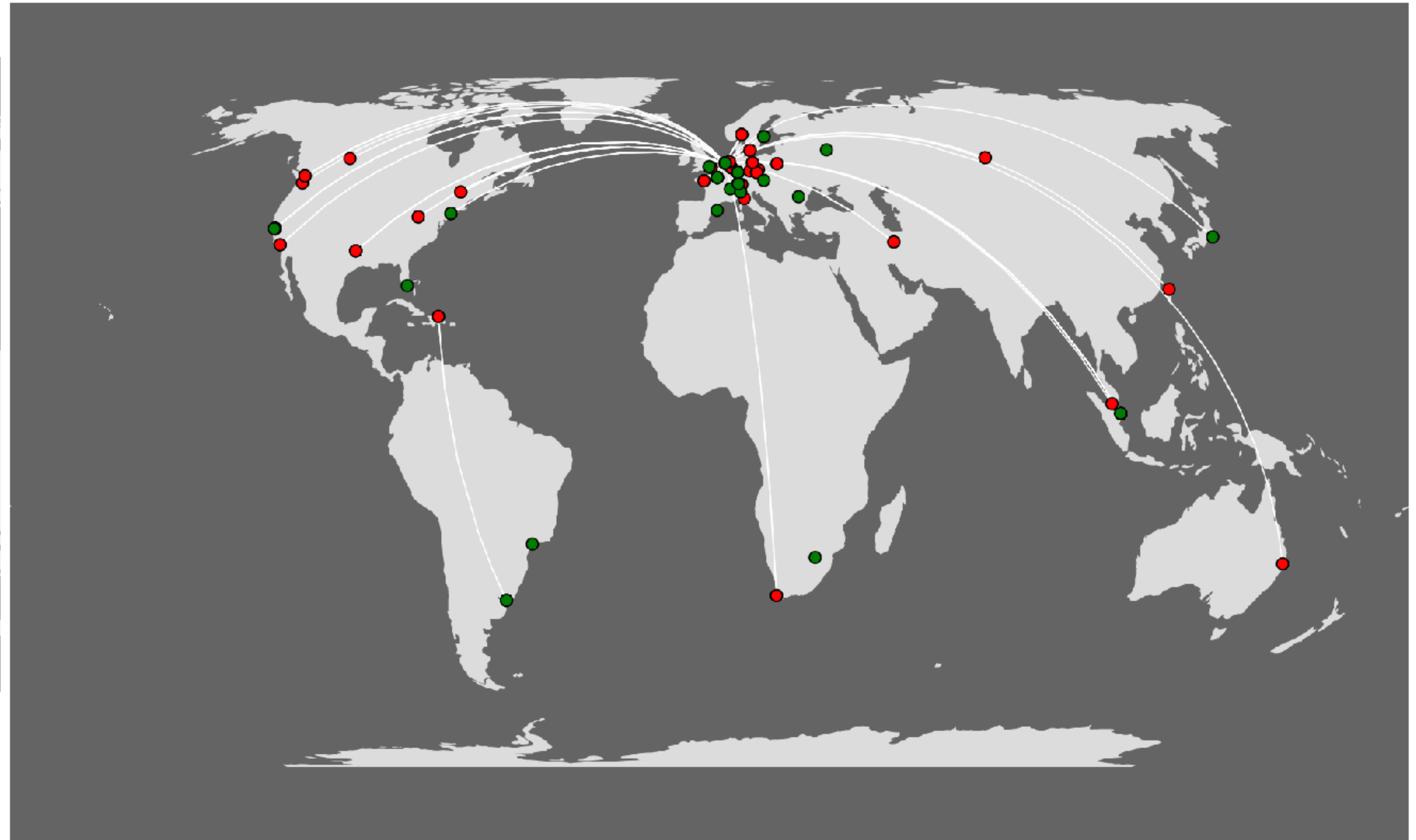
A new prefix 165.254.255.0/25 is announced by AS4, and AS15562. It should be instead 165.254.255.0/24 (description 2) announced by AS15562

<https://github.com/nttgin/BGPalerter>

What is RIS? (in pictures)



What is RIS? (in pictures)



RIS is Growing!



```
199M rrc00/2019.01/bview.20190101.0000.gz
103M rrc01/2019.01/bview.20190101.0000.gz
102M rrc03/2019.01/bview.20190101.0000.gz
32M rrc04/2019.01/bview.20190101.0000.gz
32M rrc05/2019.01/bview.20190101.0000.gz
16M rrc06/2019.01/bview.20190101.0000.gz
32M rrc07/2019.01/bview.20190101.0000.gz
82M rrc10/2019.01/bview.20190101.0000.gz
48M rrc11/2019.01/bview.20190101.0000.gz
120M rrc12/2019.01/bview.20190101.0000.gz
55M rrc13/2019.01/bview.20190101.0000.gz
45M rrc14/2019.01/bview.20190101.0000.gz
121M rrc15/2019.01/bview.20190101.0000.gz
29M rrc16/2019.01/bview.20190101.0000.gz
15M rrc18/2019.01/bview.20190101.0000.gz
37M rrc19/2019.01/bview.20190101.0000.gz
148M rrc20/2019.01/bview.20190101.0000.gz
110M rrc21/2019.01/bview.20190101.0000.gz
4.0K rrc22/2019.01/bview.20190101.0000.gz
22M rrc23/2019.01/bview.20190101.0000.gz
1.4G total
```

```
586M rrc00/2020.01/bview.20200101.0000.gz
214M rrc01/2020.01/bview.20200101.0000.gz
217M rrc03/2020.01/bview.20200101.0000.gz
30M rrc04/2020.01/bview.20200101.0000.gz
45M rrc05/2020.01/bview.20200101.0000.gz
20M rrc06/2020.01/bview.20200101.0000.gz
46M rrc07/2020.01/bview.20200101.0000.gz
105M rrc10/2020.01/bview.20200101.0000.gz
57M rrc11/2020.01/bview.20200101.0000.gz
264M rrc12/2020.01/bview.20200101.0000.gz
70M rrc13/2020.01/bview.20200101.0000.gz
53M rrc14/2020.01/bview.20200101.0000.gz
147M rrc15/2020.01/bview.20200101.0000.gz
22M rrc16/2020.01/bview.20200101.0000.gz
14M rrc18/2020.01/bview.20200101.0000.gz
57M rrc19/2020.01/bview.20200101.0000.gz
205M rrc20/2020.01/bview.20200101.0000.gz
175M rrc21/2020.01/bview.20200101.0000.gz
21M rrc22/2020.01/bview.20200101.0000.gz
22M rrc23/2020.01/bview.20200101.0000.gz
16M rrc24/2020.01/bview.20200101.0000.gz
2.4G total
```

Downside: analysis takes twice as long

Redundancy



- Do we have redundancies in the data?
- Is RIS diverse?
 - What does this mean for BGPlay, RIS Live?
- Current expansion:
 - Add route collectors (RRCs) at IXPs
 - Add peers at multi-hop or 'local' RRCs
- Can we think of strategies for better diversity
 - Less data processing, more signal, **shorter time to insight**

Diversity and Bias



- Is RIS (or any route collector project) representative of the Internet?
- The way we “sample the Internet” suggests it is biased
- Value for RIS peers:
 - For the good of the Internet
 - “I look better in Internet rankings”
- Are we in a “Filter Bubble”?



Convenience sampling

From Wikipedia, the free encyclopedia

Convenience sampling (also known as **grab sampling**, **accidental sampling**, or **opportunity sampling**) is a type of **non-probability sampling** that involves the **sample** being drawn from that part of the population that is close to hand. This type of sampling is most useful for **pilot testing**.

Advantages [\[edit\]](#)

Convenience sampling can be used by almost anyone and has been around for generations. One of the reasons that it is most often used is due to the numerous advantages it provides. This method is extremely speedy, easy, readily available, and cost effective, causing it to be an attractive option to most researchers.^[2]

Disadvantages [\[edit\]](#)

Even though convenience sampling can be easy to obtain, its disadvantages usually outweigh the advantages. This sampling technique may be more appropriate for one type of study and less for another.

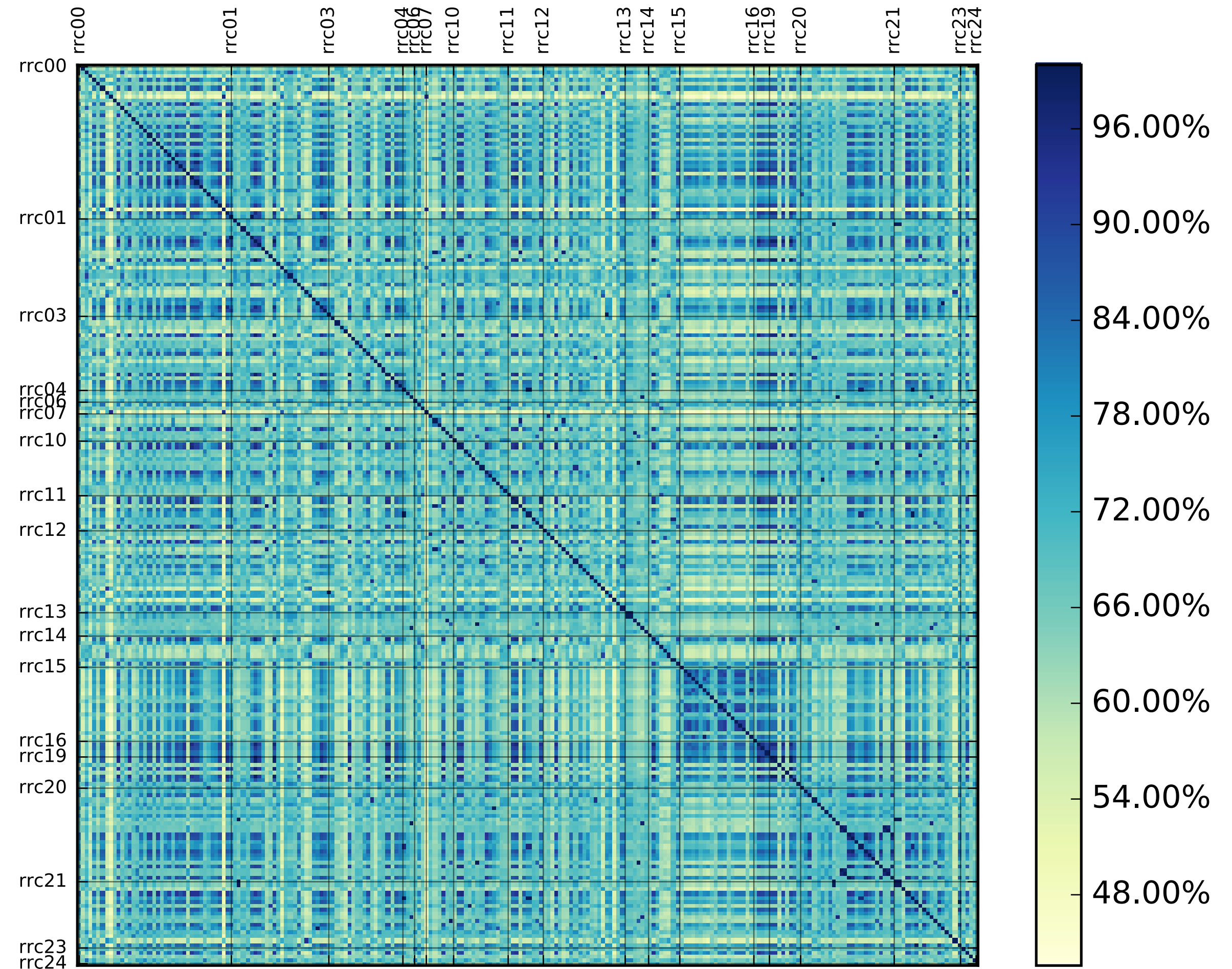
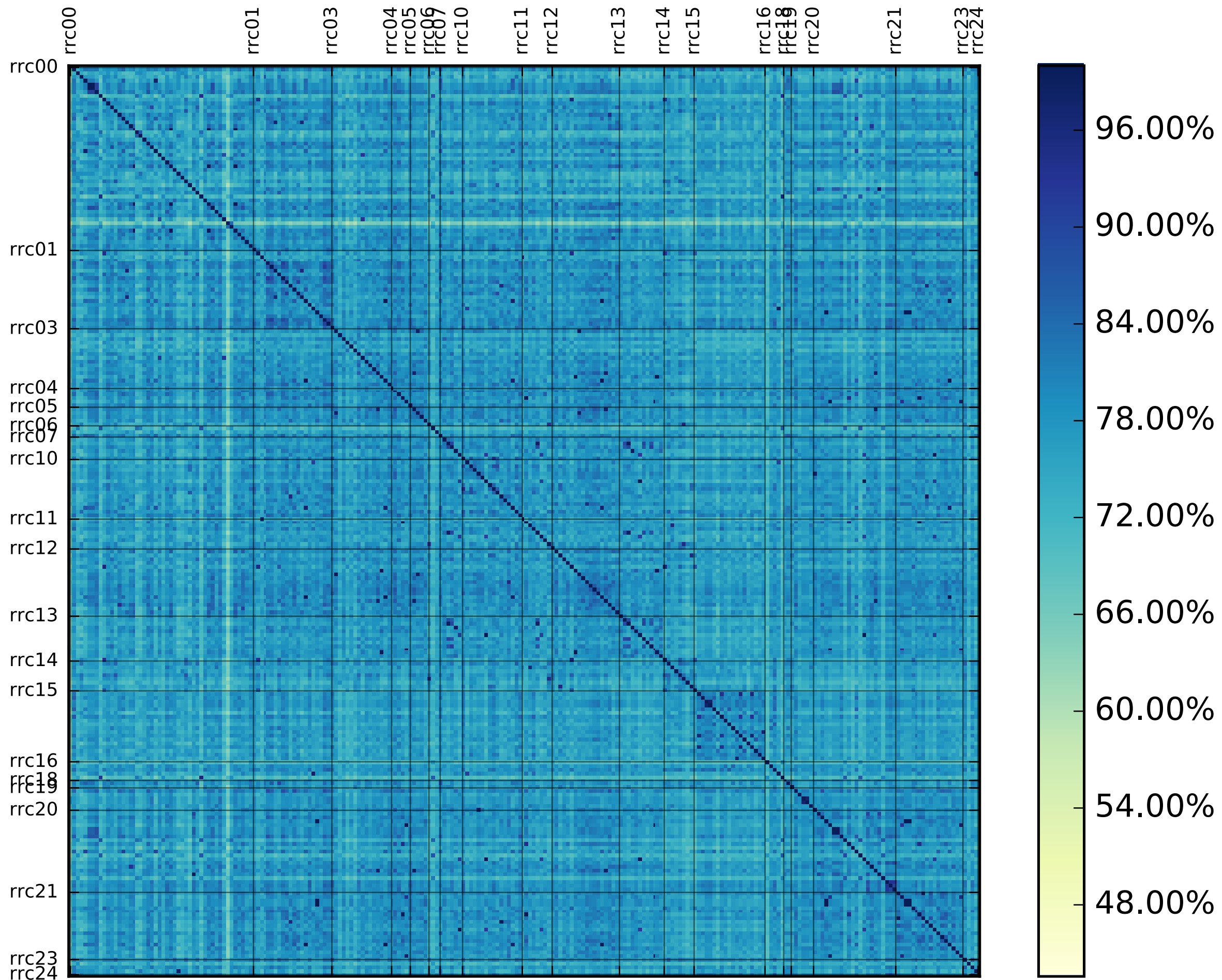
Bias

The results of the convenience sampling cannot be generalized to the target **population** because of the potential **bias** of the sampling technique due to under-representation of subgroups in the sample in comparison to the **population** of interest. The **bias** of the sample cannot be measured. Therefore, inferences based on the convenience sampling should be made only about the sample itself.^[9]

Power

Convenience sampling is characterized with insufficient **power** to identify differences of population subgroups.^[10]

Diversity in RIS



IPv4

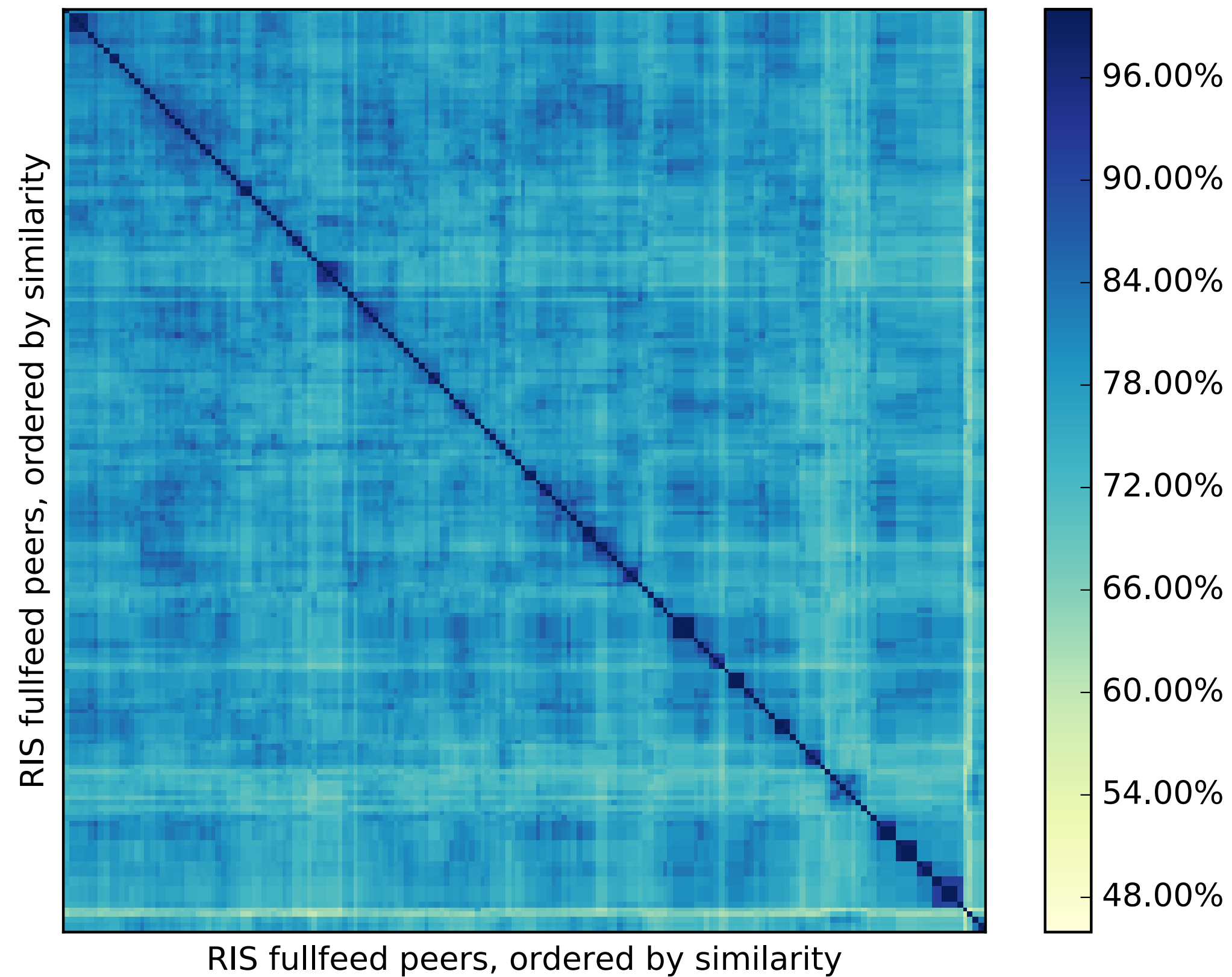
<https://labs.ripe.net/Members/emileaben/how-diverse-is-ris>

IPv6

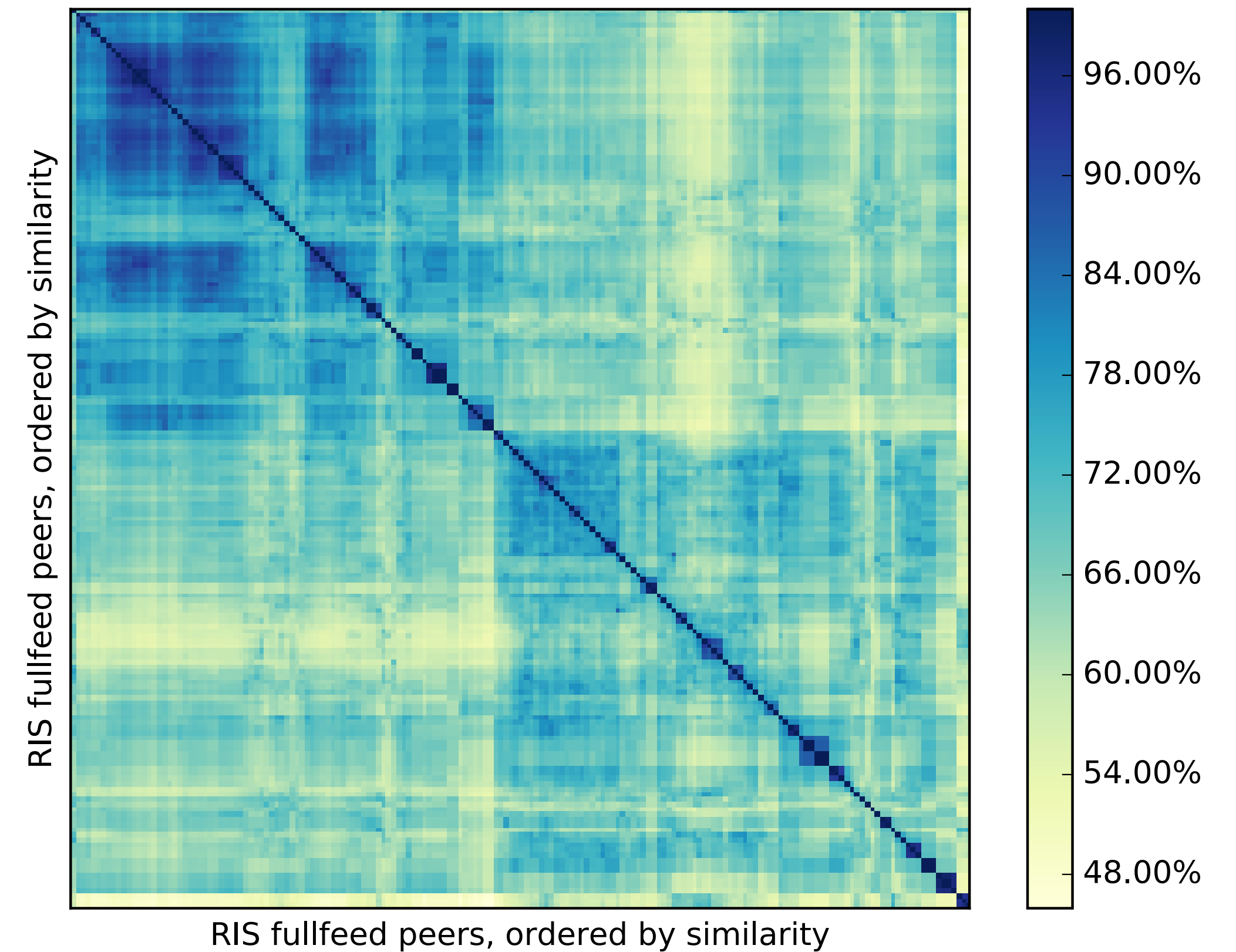
Measuring Diversity in RIS



Similarity matrix for RIS peers IPv4



Similarity matrix for RIS peers IPv6



adapted from:

<https://labs.ripe.net/Members/emileaben/how-diverse-is-ris>

Example: BGP Hijacks



- By making RIS more diverse, we'll be able to see hijacks that currently fly under the radar
- If we cover enough ASNs that are central to a certain region/country we can quantify the effect of a hijack on that region/country
- Detecting local (scoped) events needs diversity

Open Questions



- What would a peering strategy for RIS look like?
- What data ingress (peers) to focus on?
 - Tier1-like networks (central for whole Internet)
 - Locally influential networks (those central to transit for a given region/country)
 - Very local, high interconnect density (IXPs)
 - Route-servers?
 - 'Local' or multi-hop RRCs
 - Only local tables?



Conclusion

- RIS: very successful in growing
- Does this provide more value to our community?
- We don't know how well does it represent the Internet
- Growth vs. insight we get from data
 - What (type of) peers should we add?
- How should we develop RIS?
 - Let us know at: ris-users@ripe.net



Questions



ondrej.caletka@ripe.net
ris-users@ripe.net