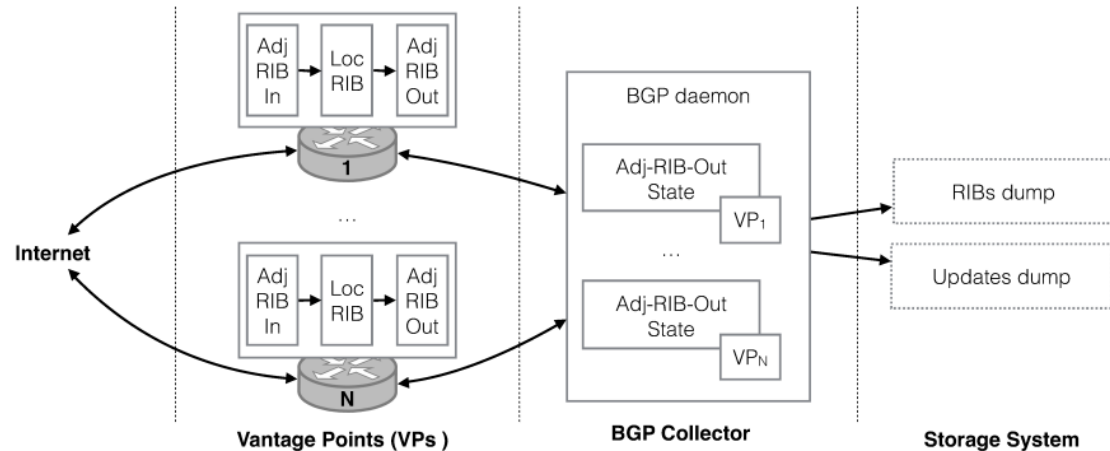# BGP Collector Archives Growth

BGP collection process

Figure 1: **BGP collection process.** Once a BGP collector establishes a BGP session with a VP, it maintains a state and an image of the VP's Adj-RIB-out table derived from the updates received through the session. With different periodicity, it dumps (i) a snapshot of all the VP Adj-RIB-out tables (RIB dump) and (ii) the update messages received within that period from all the VPs (Updates dump).

BGPStream sorting

From: https://bgpstream.caida.org/bundles/caidabgpstreamwebhomepage/pubs/bgpstream.pdf
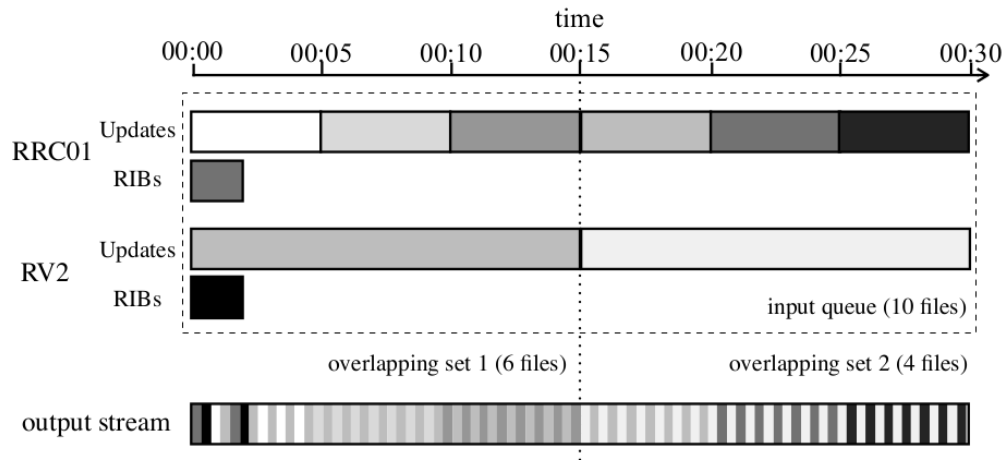

Figure 3: **Intra- and inter-collector sorting in libBGP-Stream.** An example showing how RIB and Updates dumps generated by a RIPE RIS collector (RRC01) and a Route-Views collector (RV2) are interleaved into a sorted stream. The 30 minutes (10 files) of BGP data are first separated into two disjoint sets (of 6 and 4 files) based on overlapping file time intervals. Then a multi-way merge is applied separately to the two sets, yielding the stream depicted at the bottom.

Collector archive growth

- What is normal?
- Comparison of Routeviews and RIPE RIS
  - What can we learn from one about the other?
  - Main differences
    - RIB dump interval: 8 hours RIS, 2 hours RV
    - Upd dump interval: 5 mins RIS, 15 mins RV
    - Compression: gz RIS, bz2 RV
- Two metadata datasets
  - CAIDA bgp-archive (used by BGPStream broker)
  - CAIDA bgp2go (only Routeviews, only updates)

Growth
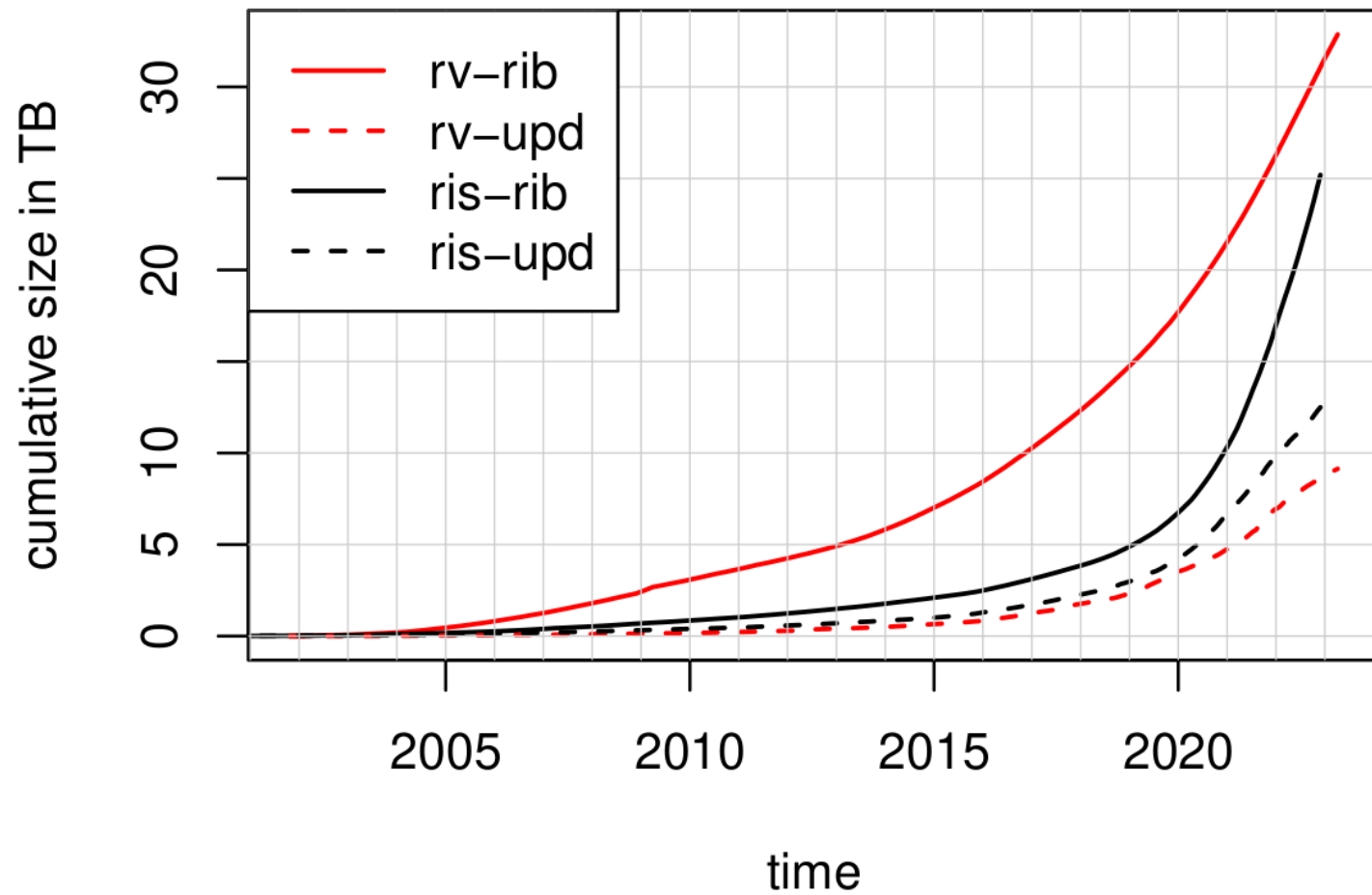
RIBs: avg file size
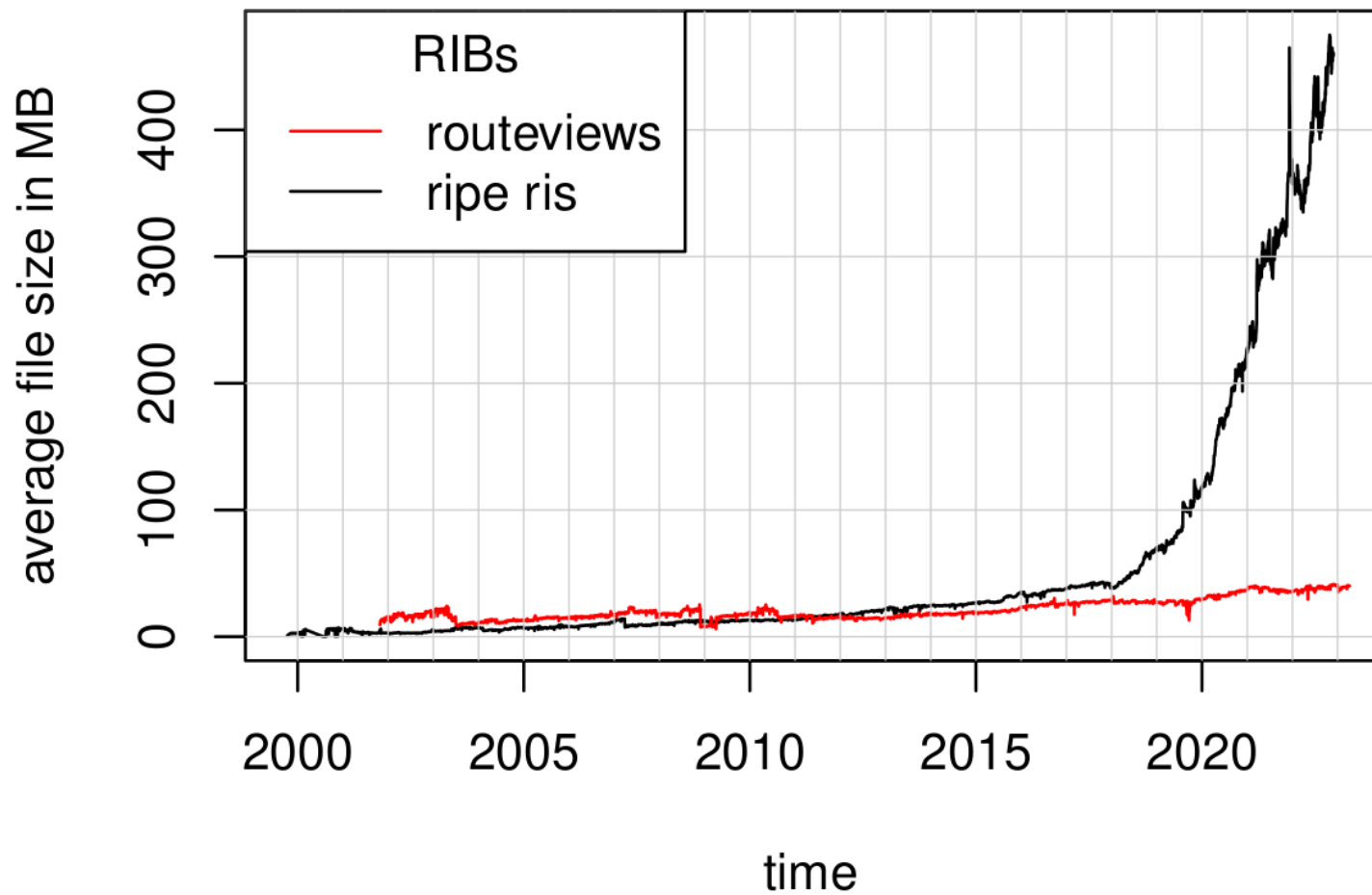
RIPE RIS growth

Index of /rrc00/2023.04/

../
bview.20230401.0000.gz   01-Apr-2023 02:47   2G
bview.20230401.0800.gz   01-Apr-2023 10:09   2G
bview.20230401.1600.gz   01-Apr-2023 18:12   2G
bview.20230402.0000.gz   02-Apr-2023 02:43   2G
bview.20230402.0800.gz   02-Apr-2023 10:05   2G
bview.20230402.1600.gz   02-Apr-2023 18:35   2G
bview.20230403.0000.gz   03-Apr-2023 00:10   389M
bview.20230403.0800.gz   03-Apr-2023 08:09   383M
bview.20230403.1600.gz   03-Apr-2023 16:09   382M
bview.20230404.0000.gz   04-Apr-2023 00:09   383M
bview.20230404.0800.gz   04-Apr-2023 08:09   384M
bview.20230404.1600.gz   04-Apr-2023 16:09   383M
bview.20230405.0000.gz   05-Apr-2023 00:10   384M
bview.20230405.0800.gz   05-Apr-2023 08:10   387M
bview.20230405.1600.gz   05-Apr-2023 16:09   388M
bview.20230406.0000.gz   06-Apr-2023 00:09   388M
bview.20230406.0800.gz   06-Apr-2023 08:09   385M
                        06-Apr-2023 16:09   386M
                        07-Apr-2023 00:09   386M

Bug causing explosion of RIB files has been fixed.

https://data.ris.ripe.net/rrc00/2023.04/

RIPE RIS growth

```
thomas@lemur:/tmp$ ll bview.* -h
-rw-rw-r-- 1 thomas thomas  56M Mar 31  2017 bview.20170401.0000.gz
-rw-rw-r-- 1 thomas thomas 1.7G Mar 31 19:47 bview.20230401.0000.gz
-rw-rw-r-- 1 thomas thomas 389M Apr  2 17:10 bview.20230403.0000.gz
-rw-rw-r-- 1 thomas thomas  98M Mar 31 17:00 rib.20230401.0000.bz2

thomas@lemur:/tmp$ bgpdump -m -l bview.20170401.0000.gz | wc
 9561795  68882417 1390857345
thomas@lemur:/tmp$ bgpdump -m -l bview.20230401.0000.gz | wc
55897049 465970556 9323050394
thomas@lemur:/tmp$ bgpdump -m -l bview.20230403.0000.gz | wc
56106002 468473147 9367422094
thomas@lemur:/tmp$ bgpdump -m -l rib.20230401.0000.bz2 | wc
26243007 189645660 3927967104
```

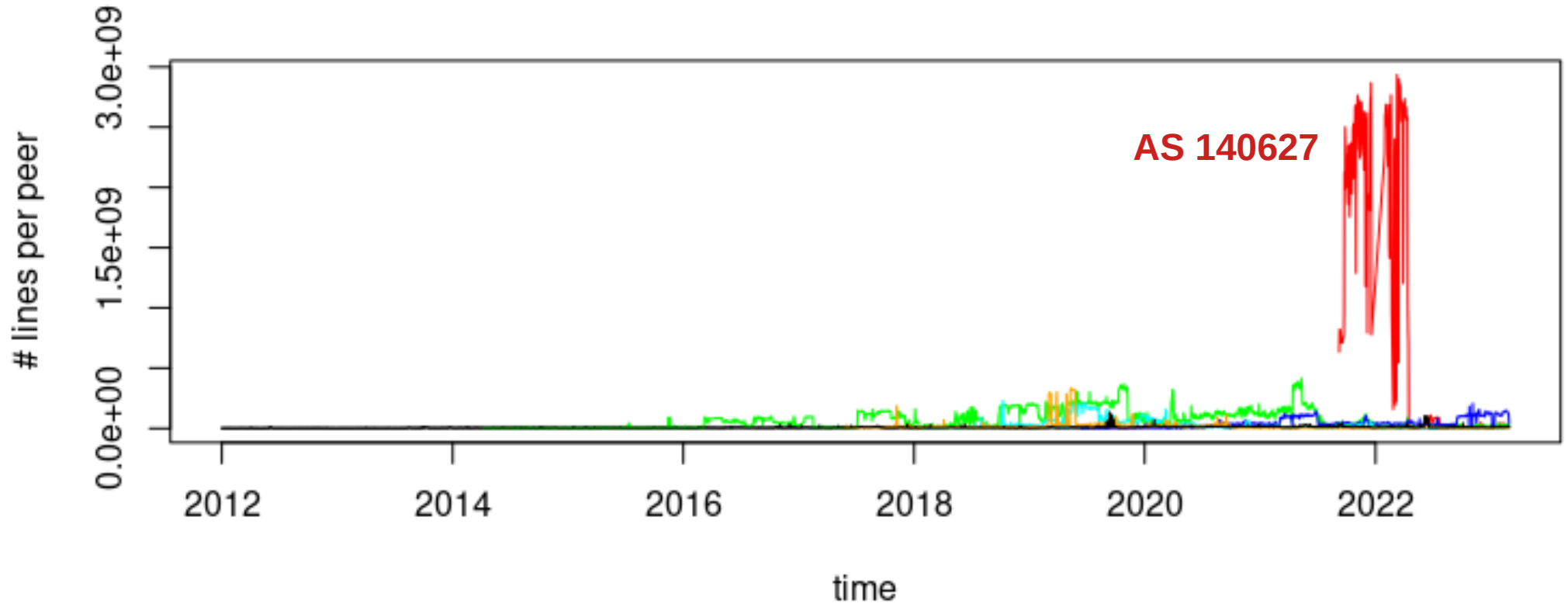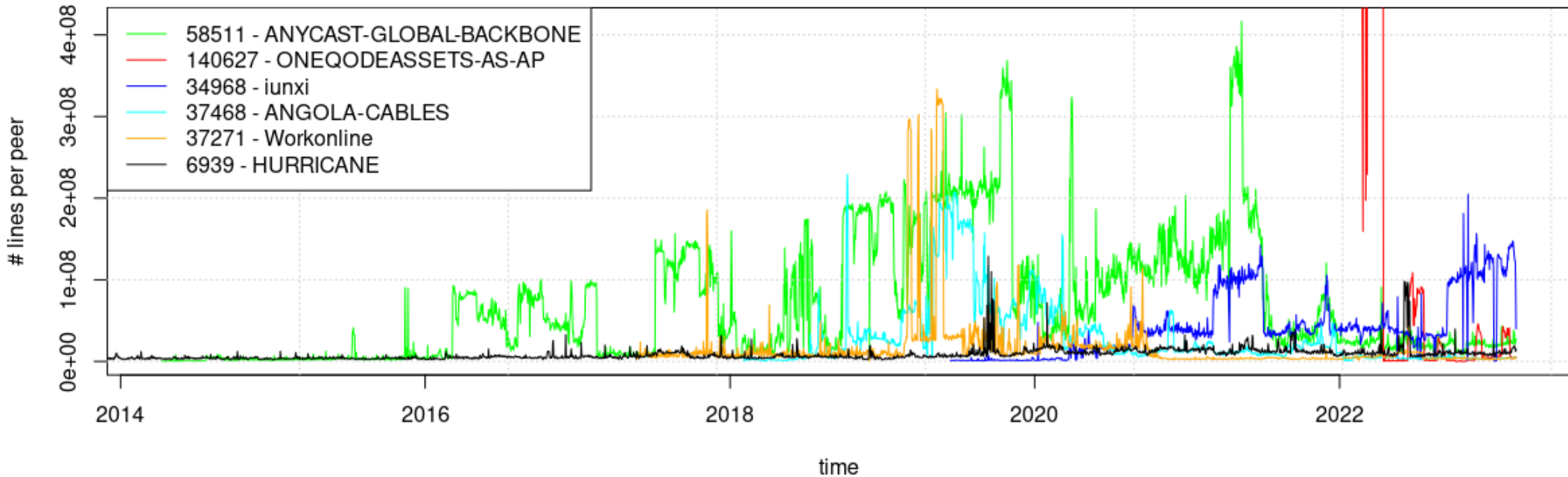**Bug causing explosion of RIB files has been fixed.**

Updates: avg file size
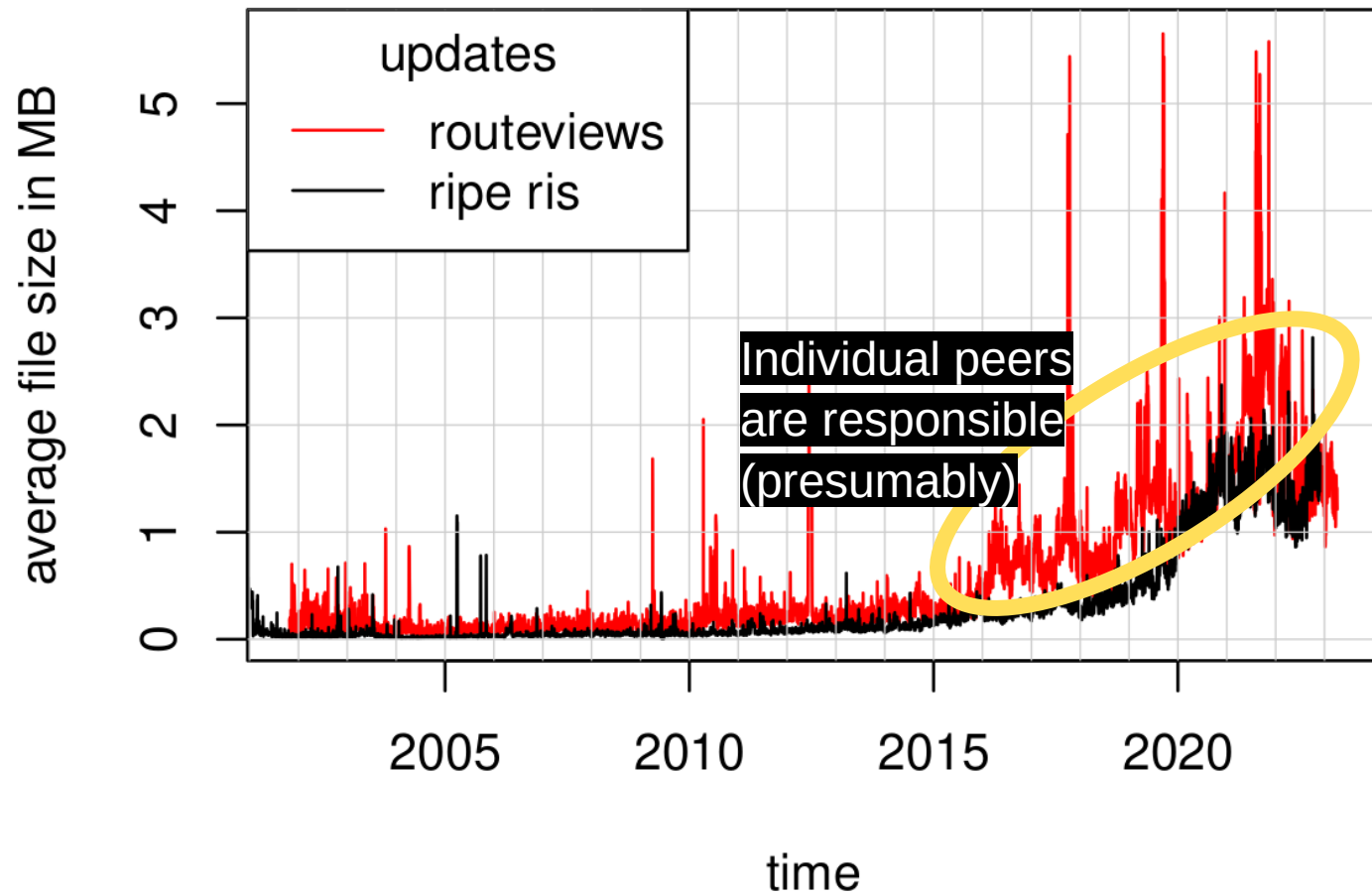
Update file aggregates:
RIS: 5 minutes
RV: 15 minutes
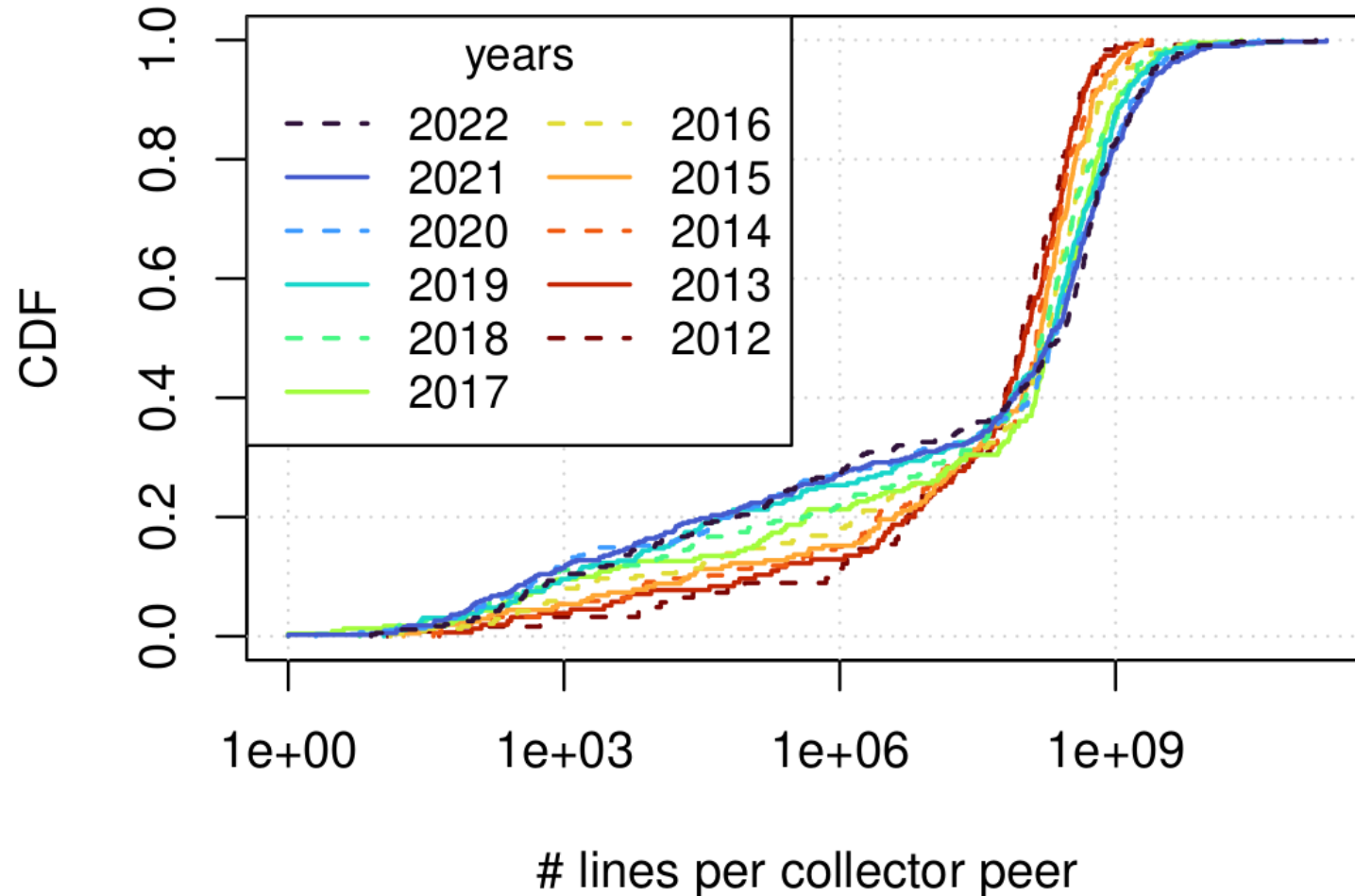
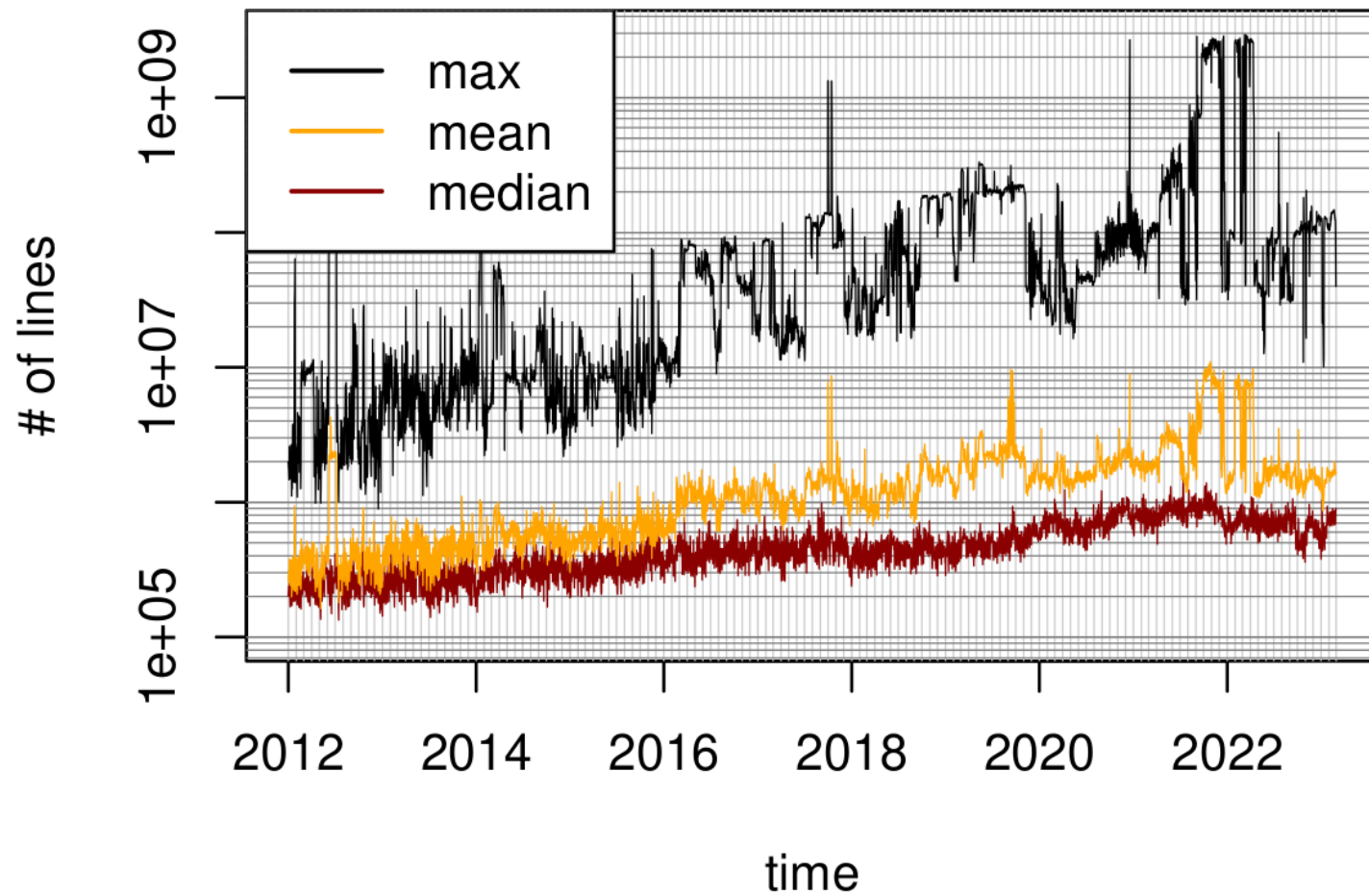More variation
compared to RIBs

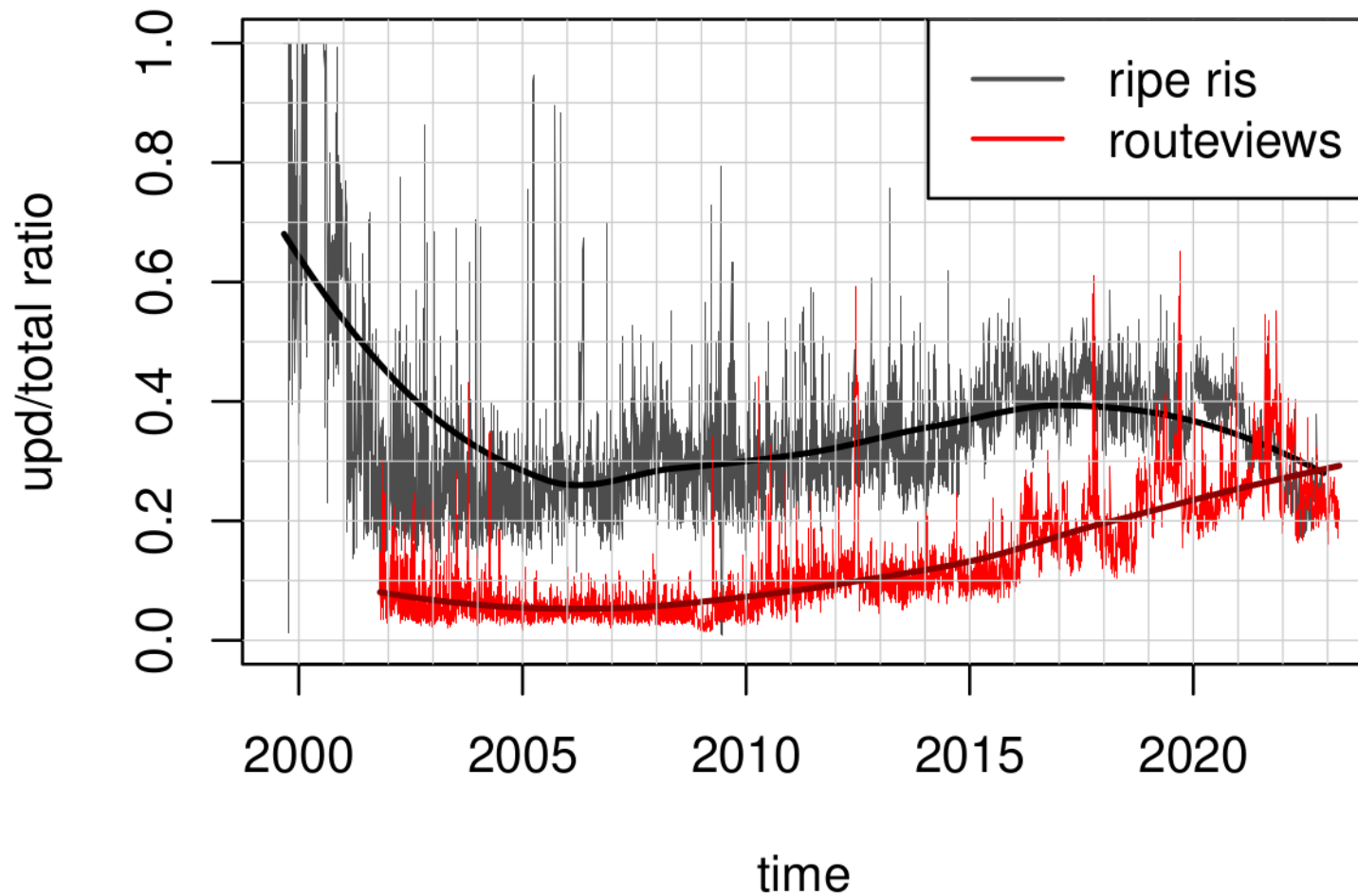# Contribution of collector peers

AS 140627

Contribution of collector peers

Collector peers

CDF vs # lines per collector peer

years: 2022, 2021, 2020, 2019, 2018, 2017, 2016, 2015, 2014, 2013, 2012

Ratio: Update / total size