



# BGP/MRT/RPKI Data Processing In Python with PyBGPKIT

July 5th, 2022  
Mingwei Zhang  
Founder of BGPKIT, Senior Systems Engineer at Cloudflare  
([mingwei@bgpkit.com](mailto:mingwei@bgpkit.com))



# BGPKIT Overview

- Modern, performant, open-source BGP data processing tools
  - <https://github.com/bgpkit>
  - Personal projects turned serious
- bgpkit-parser – Rust-based MRT Data Parsing
  - <https://github.com/bgpkit/bgpkit-parser>
  - Alternatives Rust-based solution to
    - libparsebgp (CAIDA BGPStream)
    - bgpdump (RIPE)
    - bgpscanner (Isolario)
    - Micro BGP suite (DoubleForteen Code Forge)
  - Performant, easy to extend and compose, covers more RFCs and continue updating
- bgpkit-broker – public MRT data files indexer
  - <https://github.com/bgpkit/bgpkit-broker>
  - Find MRT files available with one API call. Rust, Python, RESTful access provided.
- bgpkit-roas – a historical RPKI data lookup service
  - <https://api.roas.bgpkit.com/docs>
  - What prefixes did Google have on RPKI last year? How long has this prefix been on RPKI and since when? Is this prefix-origin pair valid based on today's data?



# BGPKIT Overview

- bgpkit-peerstat – daily summarization of the BGP peers for RouteViews and RIPE RIS
  - <https://github.com/bgpkit/peer-stats>
  - It parses one RIB dump file from every collector and produce peers information archive
  - Ip, asn, connected\_asns, v4 and v6 prefix count
- valley-free – a BGP message propagation simulation software
  - <https://github.com/bgpkit/valley-free>
  - If AS1 announces a prefix, what paths would it take to reach it from other ASes?
  - Assume valley-free model is followed
- bgp-models – a base Rust crate that defines BGP concepts
  - <https://github.com/bgpkit/bgp-models>
- pybgpkit – a all-covering Python binding for BGPKIT Rust tools
  - <https://github.com/bgpkit/pybgpkit>
  - Parser, broker, roas are currently included
- pybgpkit-api – RESTful API for pybgpkit
  - <https://github.com/bgpkit/pybgpkit-api>
  - Building block for BGP data processing pipeline
  - Useful for building user interfaces

```
100 302 100 302 0 0 172
{
  "collector": "route-views.amsix",
  "peers": {
    "185.1.167.24": {
      "asn": 3214,
      "ip": "185.1.167.24",
      "num_connected_asns": 1331,
      "num_v4_pfxs": 892093,
      "num_v6_pfxs": 0
    },
    "185.1.167.45": {
      "asn": 211398,
      "ip": "185.1.167.45",
      "num_connected_asns": 67,
      "num_v4_pfxs": 917523,
      "num_v6_pfxs": 0
    },
    "185.1.167.76": {
```



# PyBGPKIT – Control Plane Data Collector

- Python binding for the Rust code from bgpkit
  - Bgpkit-parser: parsing binary MRT files
  - Bgpkit-broker: searching available MRT files from RouteViews and RIPE RIS
  - Bgpkit-roas: historical RPKI data search using RIPE RIS daily ROV dump
- Allows easy access to public available BGP data
- All open-source and available at <https://github.com/bgpkit>
- Runnable on any Python script or Jupyter notebooks or Google Colab



# PyBGPKIT – Control Plane Data Collector

- PyBGPKIT running on Google Colab demo
  - <https://shorturl.at/dnruN>

▶ Install Package  
[ ] ↪ 1 cell hidden

▶ Parse single file with filter  
▶ ↪ 1 cell hidden

▶ Search for available files  
[ ] ↪ 1 cell hidden

▶ Check Historical RPKI Data  
[ ] ↪ 1 cell hidden



# PyBGPKIT API

- Parse files directly by URL:
  - E.g. <https://data.ripe.net/rrc00/2022.07/bview.20220701.0000.gz> (1.5GB)
- Search files:
  - E.g. get me all MRT files that contains messages of a certain time range
  - Used for data analysis a lot
- Search and parse
  - One stop query to get message from a list of filtering criteria
- Case study: YouTube 2008 Prefix Hijacking
  - Demo API (temporary link):  
[https://alpha.api.parser.bgpkit.com/docs#/default/search\\_files\\_files\\_get](https://alpha.api.parser.bgpkit.com/docs#/default/search_files_files_get)
  - <https://www.ripe.net/publications/news/industry-developments/youtube-hijacking-a-ripe-ncc-riis-case-study>
  - 2008-02-24T18:30:00 to 2008-02-24T23:00:00
  - 5 seconds, 56 MRT files from rrc00, 57 relevant prefix announcements found
- Try it out: <https://github.com/bgpkit/pybgpkit-api>

Thank you!

Mingwei Zhang  
Founder and Maintainer of [BGPKIT.com](https://bgpkit.com)  
Senior Systems Engineer at Cloudflare  
[mingwei@bgpkit.com](mailto:mingwei@bgpkit.com)

