KIRIN: Hitting the Internet with Distributed BGP Announcements

Lars Prehn

Max Planck Institute for Informatics

Pawel Foremski

IITiS PAN / DomainTools

Oliver Gasser

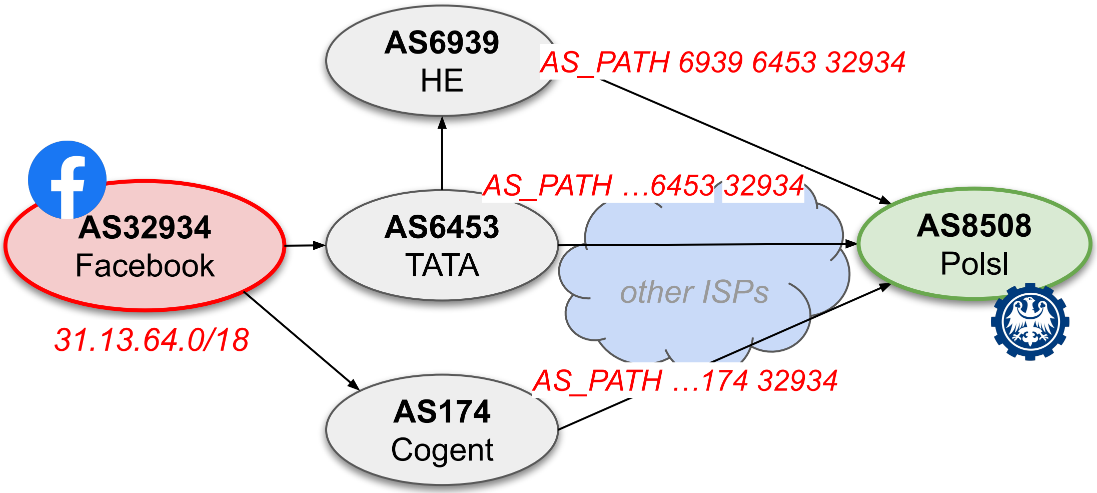
IPinfo / Max Planck Institute for Informatics

ACM AsiaCCS 2024

1 - 5 July, Singapore

BGP Background

- **Border Gateway Protocol runs the Internet routing**



- **BGP ~ selective broadcast via graph**
  - Internet routers highly interconnected (IXPs)
  - Transit, peering, customer links (p2p / p2mp)
  - Usually the shortest path selected

- **Routes stored in FIB and RIB tables**
  - FIB: Forwarding Information Base = selected
  - RIB: Routing Information Base = available
  - Both have limited capacity

What if too many routes announced?

🤔

Good tutorial: BGP for All

Prefix De-aggregation Attack

- **Well-known idea:** split large prefixes, overwhelm BGP neighbors

192.168.0.0/**16** (1)

↓

192.168.0.0/**24**, 192.168.1.0/**24**, 192.168.2.0/**24**, …, 192.168.254.0/**24**, 192.168.255.0/**24** (256)

+

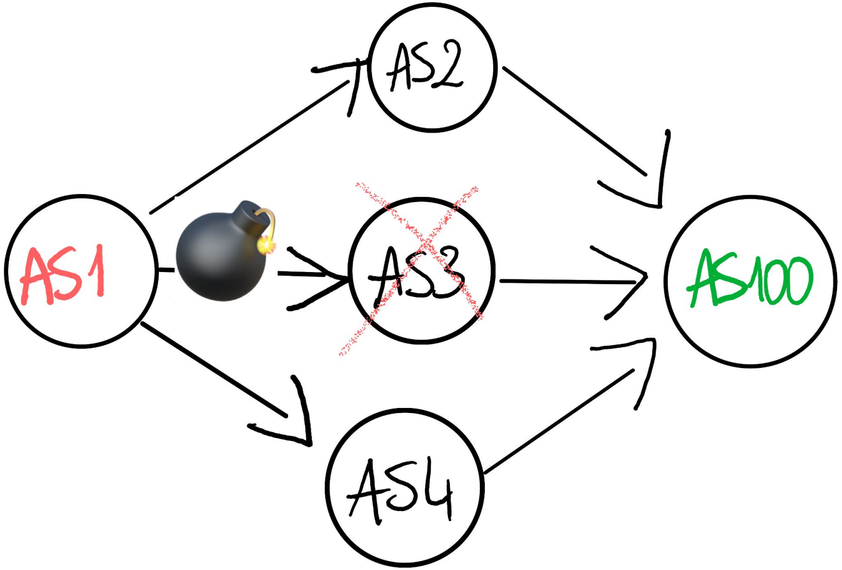
192.168.0.0/**23** … 192.168.254.0/**23** (128)

+

192.168.0.0/**22** … 192.168.252.0/**22** (64)

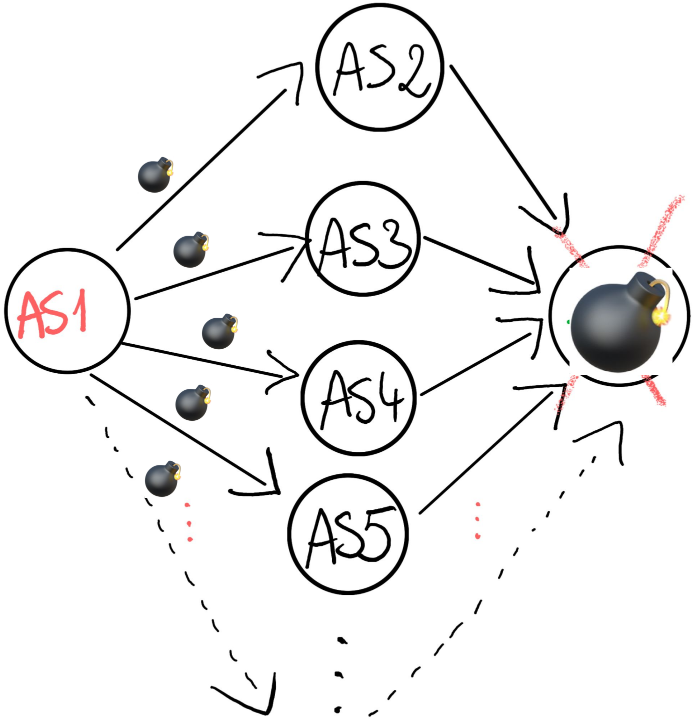
…

- **Protection:**
  - BGP session max-prefix limits
  - Route aggregation
  - More router memory
    (# of routes: 970k IPv4 + 210k IPv6)



Distributed Prefix De-aggregation Attack

- **KIRIN** revisits the attack in modern context:
  - **Remote peering** is increasingly popular
  - **IPv6** is widely deployed and available
- **Key Ideas:**
  - **Many distributed sessions:**
    workaround for max-prefix limits
    and route aggregation
    (announce unique, disjoint sets of routes)
  - **Instant and cheap remote BGP sessions**:
    no need for physical presence, automated setup
  - **IPv6 with RPKI maxLength:**
    eg. easy /29 = 1 million RPKI-valid sub-prefixes
    (global propagation with route filter accept)



Results: Theoretical Feasibility #1

- ILP solver for real-world Internet topology
  - #1 Transit Scenario
  - #2 Peering Scenario

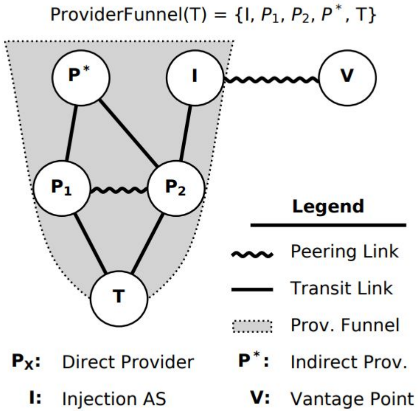


Figure 1: Provider funnel example.

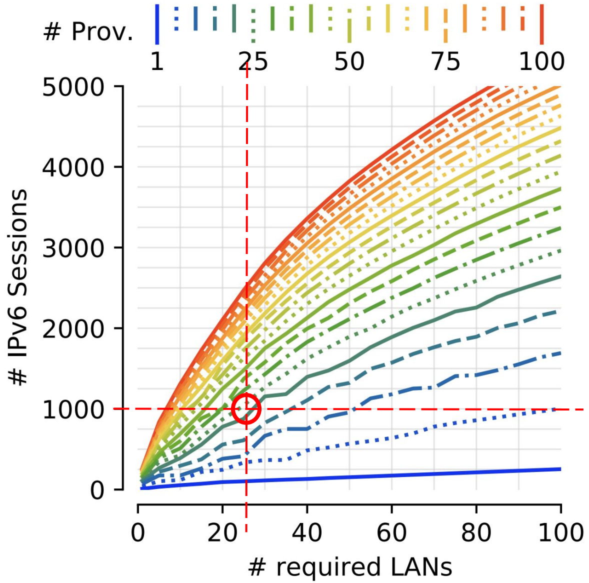


Figure 2: Transit Scenario: trade-off landscape.

Attack feasible

(20 providers @ 25 points → 1M prefixes)

Results: Theoretical Feasibility #2

- ILP solver for real-world Internet topology
  - #1 Transit Scenario
  - #2 Peering Scenario

Peering alone requires unrealistic resources

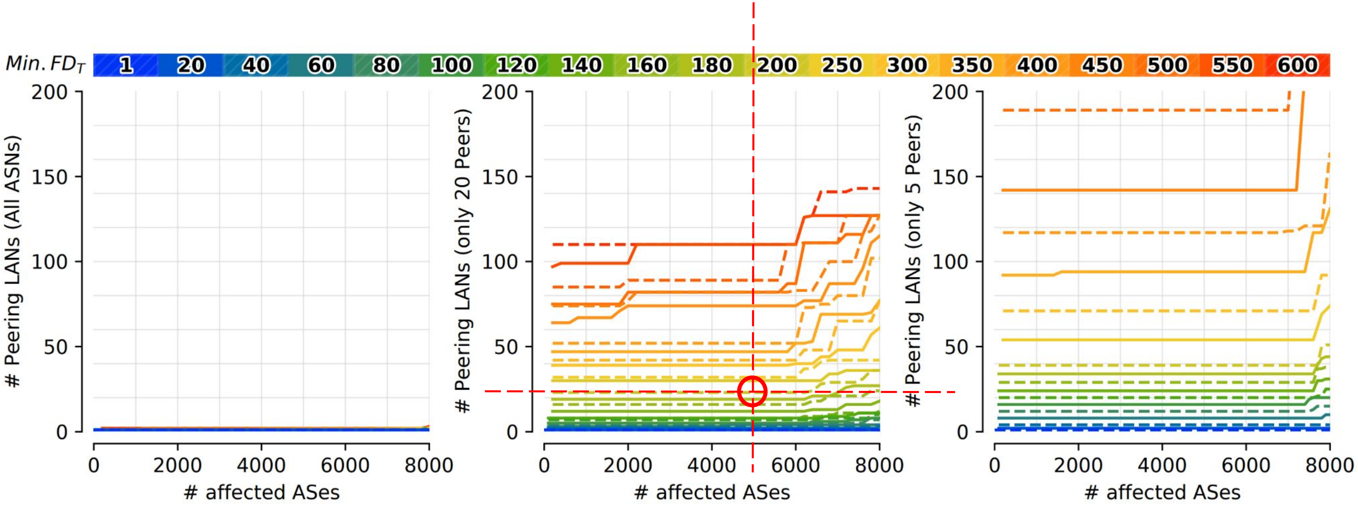
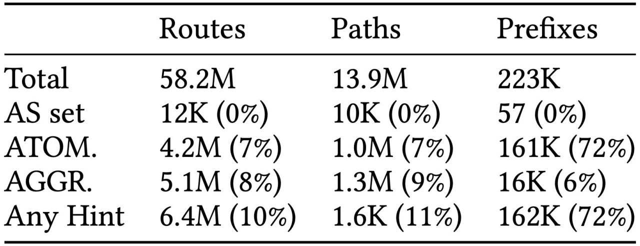


Figure 3: Peering Scenario: trade-off landscape for I_{all} (left), I_{20} (middle), and I_5 (right).

Results: Practical Experiments

- Built real-world BGP testbed
- Tested Kirin on micro-scale
- Route Aggregation is rare and can be circumvented
- Validated route propagation assumptions
- Tested real BGP routers memory usage



	Routes	Paths	Prefixes
Total	58.2M	13.9M	223K
AS set	12K (0%)	10K (0%)	57 (0%)
ATOM.	4.2M (7%)	1.0M (7%)	161K (72%)
AGGR.	5.1M (8%)	1.3M (9%)	16K (6%)
Any Hint	6.4M (10%)	1.6K (11%)	162K (72%)

Table 1: Results of aggregation analysis.

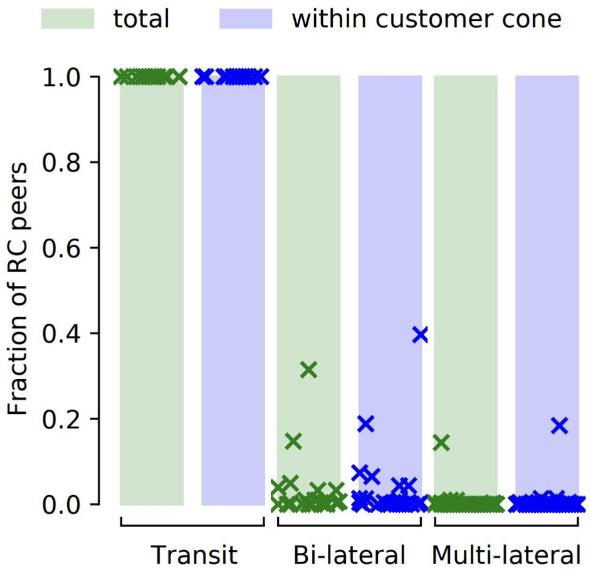


Figure 6: Redistribution behavior of different session types.

Potential Defenses & Operator Response

- Dynamic and Tight Max-Prefix Limits
  - Small but possibly growing <1.5x per day
- Per-Origin and Per-Block Prefix Limits
  - Open-Source implementation: bgpipe.org
  - Presented at RIPE88 conference
- Monitoring, Filtering, Adding Delay
  - Be careful with automated filter lists
  - Monitor for novel prefixes

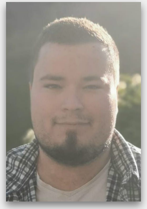
- Responsible Disclosure
  - Private (IXPs, Tier-1s, Clouds, etc.)
  - Public (mailing lists, blog posts, IETF)
- **Operators deployed protections, eg:**
  - 2 Tier-1 ASes
  - 3 Cloud Providers
  - Various smaller networks

🎓 🎉

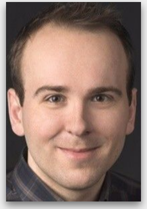
Thank you!

kirin-attack.github.io

Lars Prehn
lprehn@mpi-inf.mpg.de
@mydamnhandle1



Pawel Foremski
pjf@iitis.pl
@pforemski



Oliver Gasser
oliver@ipinfo.io

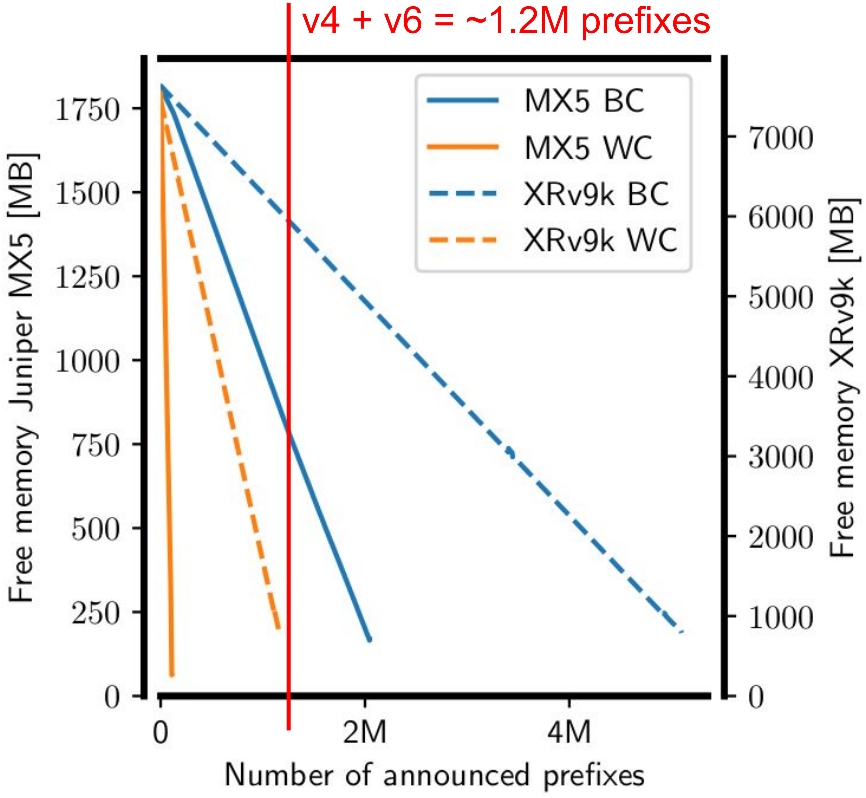
Backup Slides

Results: Lab Experiments

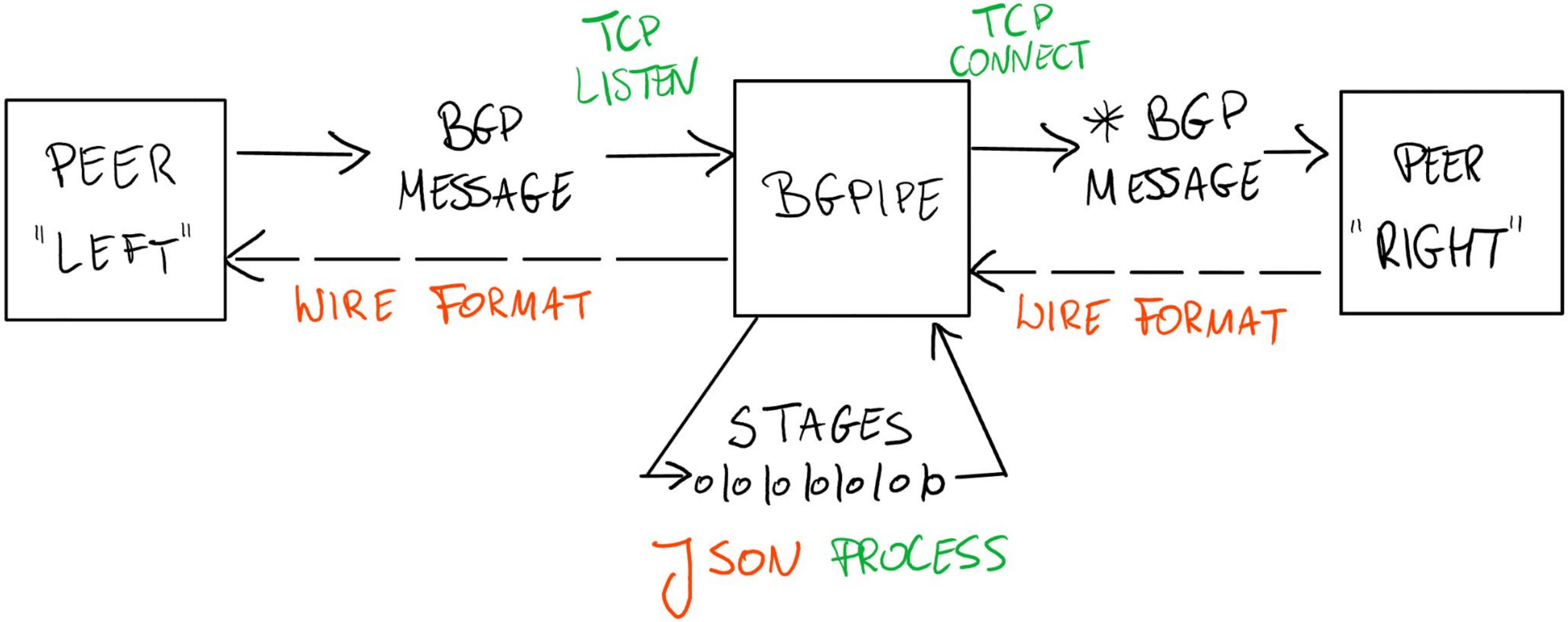
- Track memory usage for 2 BGP routers
  - Juniper MX5 —— ——
  - Cisco XRv9k - - - - - -

- Announce non-aggregatable IPv6 routes:
  - **BC = Best-Case Scenario:**
    shortest AS_PATH, no BGP communities
  - **WC = Worst-Case Scenario:**
    longest AS_PATH, full Large BGP communities

$$
\begin{aligned}
MEM = \; & (PREFIX\_SIZE + (255 \times ASN\_SIZE) + \\
& (255 \times COMM\_SIZE)) \times NUM\_PFX
\end{aligned}
$$

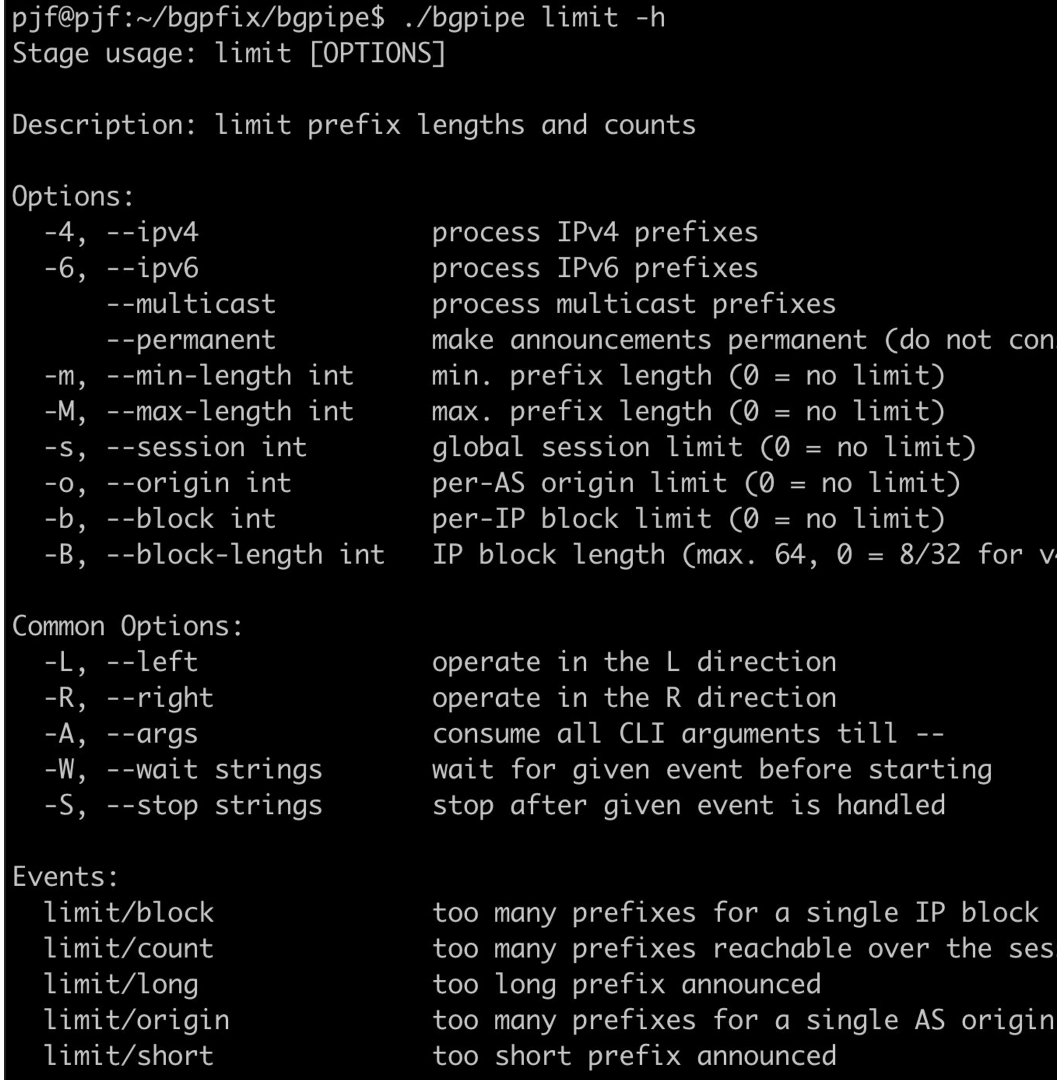

v4 + v6 = ~1.2M prefixes

bgpipe.org overview



bgpipe.org: limit

- More advanced max-prefix limits:
  - per-session (classic)
  - per-IP block (eg. 10k per each ::/32)
  - per-AS origin (eg. 15k for any ASN)
- Implemented as a Stage: see limit.go



```
pjf@pjf:~/bgpfix/bgpipe$ ./bgpipe limit -h
Stage usage: limit [OPTIONS]

Description: limit prefix lengths and counts

Options:
  -4, --ipv4                process IPv4 prefixes
  -6, --ipv6                process IPv6 prefixes
      --multicast           process multicast prefixes
      --permanent           make announcements permanent (do not con
  -m, --min-length int      min. prefix length (0 = no limit)
  -M, --max-length int      max. prefix length (0 = no limit)
  -s, --session int         global session limit (0 = no limit)
  -o, --origin int          per-AS origin limit (0 = no limit)
  -b, --block int           per-IP block limit (0 = no limit)
  -B, --block-length int    IP block length (max. 64, 0 = 8/32 for v

Common Options:
  -L, --left                operate in the L direction
  -R, --right               operate in the R direction
  -A, --args                consume all CLI arguments till --
  -W, --wait strings        wait for given event before starting
  -S, --stop strings        stop after given event is handled

Events:
  limit/block               too many prefixes for a single IP block
  limit/count               too many prefixes reachable over the ses
  limit/long                too long prefix announced
  limit/origin              too many prefixes for a single AS origin
  limit/short               too short prefix announced
```

ONE DOES NOT SIMPLY
ANNOUNCE A MILLION NEW ROUTES
imgflip.com