BalticNOG

FRRouting: BGP Features You Haven’t Used (= Heard), Yet?

September 24-25, BalticNOG 2025
Vilnius, Lithuania

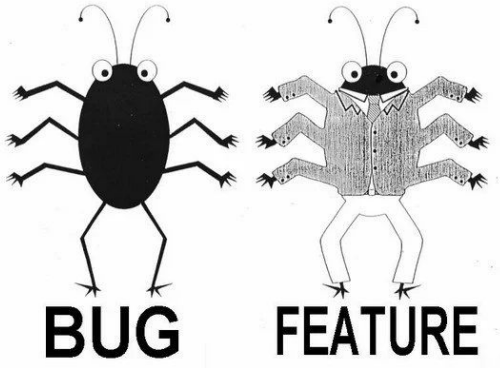
Donatas Abraitis

whoami

- Donatas Abraitis (ton31337)
- FRRouting maintainer since 2019
- Systems Engineer at
- Software Engineer at
- Co-founder at

Agenda

- Discover some recent FRRouting known unknowns
- Some handy not well-known BGP capabilities
- New drafts



0/8; 127/8; 240/4

- Shrink from 16M to a single address 0.0.0.0 for **0.0.0.0/8** draft-schoen-intarea-unicast-0)
- 16M to 65k addresses (127.0.0.0/16) for **127.0.0.0/8** (draft-schoen-intarea-unicast-127)
- 268M for **240.0.0.0/4** (draft-schoen-intarea-unicast-240) => allowed by default
- Roughly “new” ~300M addresses on the market
- “localhost” peering, e.g. 127.0.0.1 <---> 127.0.0.2
- allow-reserved-ranges (disabled by default)
- 240/4 used internally by some well-known companies

remote-as auto

- BGP message type (1 => **OPEN**) + Capability (4-bytes ASN)
- neighbor 127.0.0.2 remote-as (plain or asdot formats):
  - **65002** (2-bytes)
  - **4200000000** (4-bytes)
  - **64154.0** (asdot)
  - **internal** (or **external**)
  - **auto** => reduces configuration overhead, more flexible
- neighbor PEER remote-as auto

Link-Local Next Hop Capability

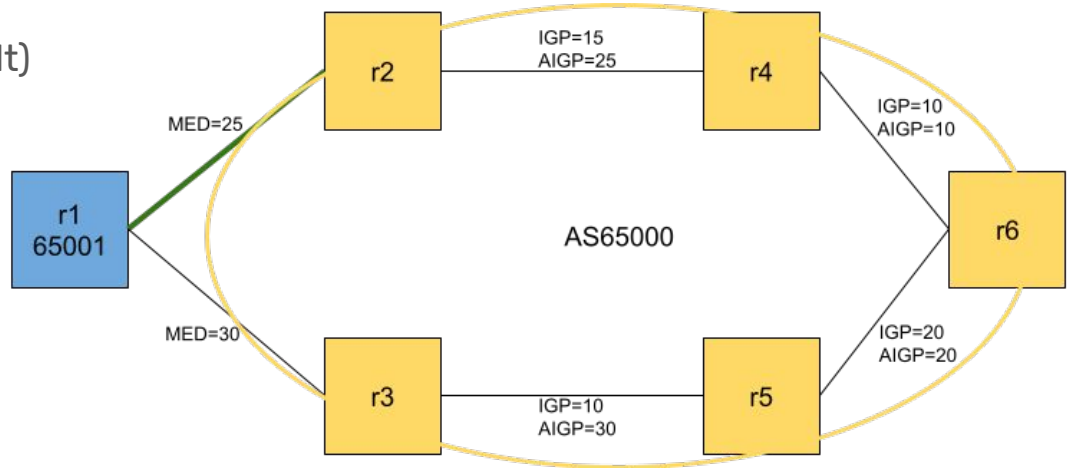
- “unnumbered” peering (interface-based)
- Interoperability issues handling 32-bytes next-hops
  - **fe80::dead** | fe80::dead
  - **::** | fe80::dead
- Send only 16-bytes next-hop (IPv6 Link-Local address)
- draft-ietf-idr-linklocal-capability (adopted by IETF IDR WG)
- neighbor PEER capability link-local (implicitly enabled for unnumbered peering)

Dynamic BGP Capability

- New BGP message type (6 => **DYNAMIC**)
- Reducing downtime (no session reset)
- Adjust Graceful-Restart (or Long-Lived Graceful-Restart) timers “on-the-fly” before going down
- Change BGP peering role
- Enable/disable ORF capability (Outbound Route Filtering)
- Facilitate non-disruptive capability changes
- Not all vendors support it - not yet RFC’ed (draft-ietf-idr-dynamic-cap)
- Enabled by default for the “datacenter” profile since 10.1 release

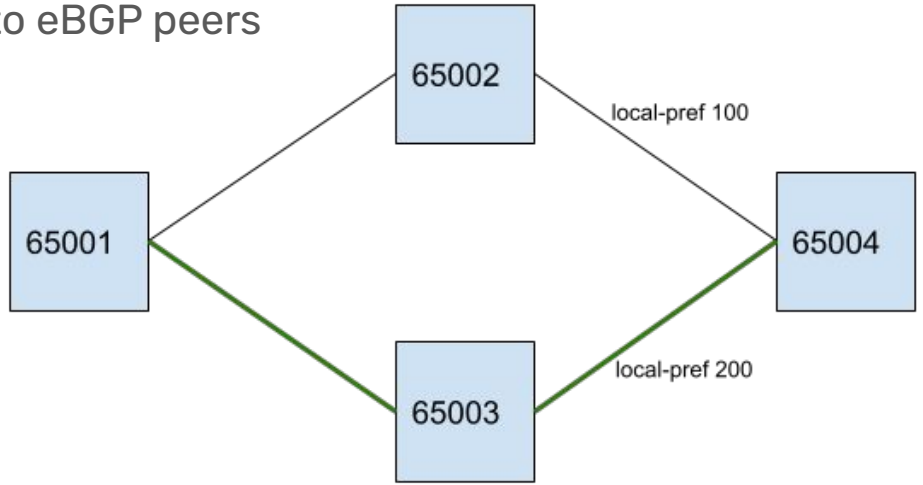
Accumulated IGP Metric for BGP (AIGP)

- BGP primarily relies on AS-path length
- Distance = sum(metrics of all the links that belong to that path)
- Optional, non-transitive BGP attribute to carry IGP metric (**NOT** “AI Grand Prix”)
- Only between internal BGP peers (iBGP) or eBGP-OAD
- bgp bestpath aigp (disabled by default)
  - Shortest AIGP wins
  - Preferred over AS-path length
- set metric aigp (eBGP)
- rfc7311



One Administrative Domain (OAD)

- Networks with multiple ASNs require flexible way to propagate the attributes
- Common administrative entity / domain
- rfc7938 (“large-scale”, “hyper-scale”, “warehouse-scale”)
- Allow non-transitive attributes passing to eBGP peers
  - draft-uttaro-idr-bgp-oad
- neighbor PEER oad
  - AIGP metric
  - local-preference



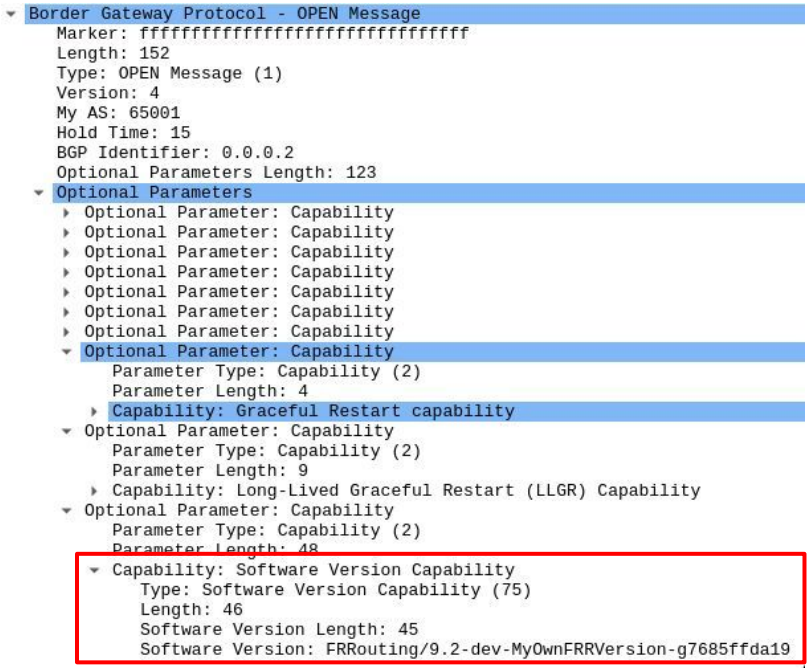
Paths Limit for Multiple Paths Capability

- BGP routers normally only advertise the **best** path to their neighbors
- Additional paths for the same prefix can be sent with ADD-PATH capability
- Reduce memory consumption and increase network stability (DoS attack)
    - Craziness => 100 paths per-prefix
- RSs are sending additional paths because most of the implementations do not support dropping RX flag for AddPath capability
- neighbor PEER disable-addpath-rx
- neighbor PEER addpath-rx-paths-limit
- draft-abraitis-idr-addpath-paths-limit

BalticNOG

Software Version Capability

- Modern data centers running (non-)conventional routers
- Fleet of different versions of the routing daemons
- Identify software mismatches causing issues
- Interoperability troubleshooting
- Detect outdated/vulnerable software versions
- “Twitter-over-BGP”
- neighbor PEER capability software-version
- draft-abraitis-bgp-version-capability



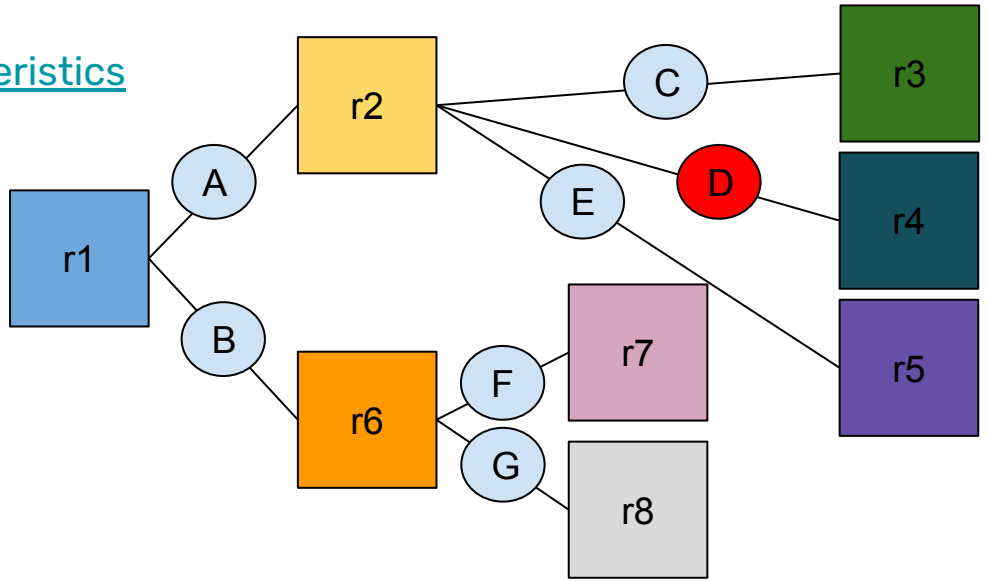
Notification Message Support for BGP Graceful Restart

- Normally => Send a NOTIFICATION and reset to handle error conditions
- Graceful Restart mechanism is ignored when NOTIFICATION sent/received
- rfc8538
- Routes retained during configuration changes (%NOTIFICATION sent/received)
- … or when Hold Time expires
- … or when Send Hold Time expires? (rfc9687)
- GR capability “N” bit —> “Hard Reset” vs. “Graceful Reset”
- bgp hard-administrative-reset (**clear ip bgp 127.0.0.2**)

BGP Next Hop Dependent Characteristics

- Advertise forwarding plane features
- Clos (ECMP/WCMP)
- **<next-hop**; **next-next-hop>** pair
- neighbor PEER send-nexthop-characteristics
- Congestion avoidance

```
r1# show bgp ipv4 10.0.0.1/32
…
  10.255.0.2 from 10.255.0.2 (10.255.0.2)
        Next-next Hop Nodes:
        10.254.0.3
        10.254.0.4
        10.254.0.5
  10.255.0.6 from 10.255.0.6 (10.255.0.6)
        Next-next Hop Nodes:
        10.254.0.7
        10.254.0.8
```



BalticNOG

Thank you!