

[RFC 8950] BGP IP Peer Scalability At CXC-IX

IPv4 sometimes cannot scale ?

RFC 8950

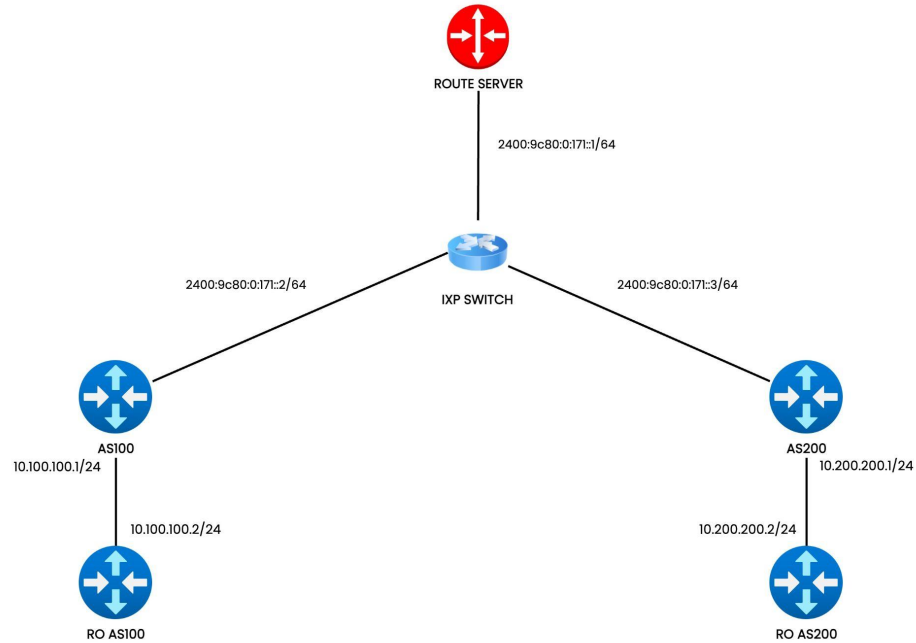
Advertising IPv4 Network Layer Reachability Information (NLRI) with an IPv6 Next Hop

<https://datatracker.ietf.org/doc/html/rfc8950>

Configuration Example

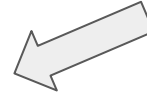
IXP and Peers

Example Topology



Route Server Configuration

```
router bgp 133339
.....
neighbor 2400:9c80:0:171::2 remote-as 100
neighbor 2400:9c80:0:171::2 description AS100
neighbor 2400:9c80:0:171::2 capability extended-nexthop
neighbor 2400:9c80:0:171::3 remote-as 200
neighbor 2400:9c80:0:171::3 description AS200
neighbor 2400:9c80:0:171::3 capability extended-nexthop
.....
```

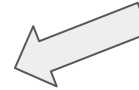


Enable extended-nexthop capability
to use RFC8590

(accept/send ipv4 prefix with ipv6 nexthop)

Peers (Members) Configuration

```
router bgp 100
.....|
neighbor 2400:9c80:0:171::1 remote-as 133339
neighbor 2400:9c80:0:171::1 description rs1
neighbor 2400:9c80:0:171::1 capability extended-nexthop
.....
```



Enable extended-nexthop capability
to use RFC8590

(accept/send ipv4 prefix with ipv6 nexthop)

```
...
route-map out-cxc permit 10
match ip address prefix-list my_prefix
set ipv6 next-hop global 2400:9c80:0:171::2
set ipv6 next-hop prefer-global
exit
...|
```



Set route-map ipv4 prefix list with ipv6
nexthop (prefer-global)

Received Route

```
frr# show ip bgp
BGP table version is 14, local router ID is 103.225.171.2, vrf id 0
Default local pref 100, local AS 100
Status codes: s suppressed, d damped, h history, * valid, > best, = multipath,
               i internal, r RIB-failure, S Stale, R Removed
Nexthop codes: @NNN nexthop's vrf id, < announce-nh-self
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 10.100.100.0/24	0.0.0.0	0		32768	i
*> 10.200.200.0/24	2400:9c80:0:171::3	0		0 200	i

Displayed 2 routes and 2 total paths

```
frr# sh ip bgp summary
```

```
IPv4 Unicast Summary (VRF default):
BGP router identifier 103.225.171.2, local AS number 100 vrf-id 0
BGP table version 14
RIB entries 3, using 552 bytes of memory
Peers 1, using 716 KiB of memory
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd	PfxSnt	Desc
2400:9c80:0:171::1	4	133339	86	84	0	0	0	00:37:10	1	1	rs1

Total number of neighbors 1

```
frr#
```

AS100 receive ipv4 route with ipv6
nexthop address

Reachability Test

```
[admin@MikroTik] > tool traceroute 10.200.200.2 src-address=10.100.100.2
# ADDRESS          LOSS SENT    LAST    AVG    BEST    WORST
1 10.100.100.1      0%   7    3.1ms  3.4    3.1    4.1
2 10.200.200.1     0%   7     6ms   6.1    5.8    6.7
3 10.200.200.2     0%   7     9ms   10     9     11.8
- [Q quit|D dump|C-z pause]
```




Test from AS100 to AS200

```
[admin@MikroTik] > tool traceroute 10.100.100.2 src-address=10.200.200.2
# ADDRESS          LOSS SENT    LAST    AVG    BEST    WORST
1 10.200.200.1     0%   6    2.8ms  3.2    2.8    3.6
2 10.100.100.1     0%   6    5.8ms  5.8    5.6    6.2
3 10.100.100.2     0%   6    8.7ms   9     8.6    9.8
- [Q quit|D dump|C-z pause]
```

Test from AS200 to AS100

BGP IP Peer on IXP only use
IPv6 address 🤪

Extended Nexthop Capability

	Mikrotik (ROS 6.49.6)	FRR 8.2.2	Bird 2.0.12
<i>Extended Nexthop Capability</i>			

Conclusion

Pros:

- ❑ Good for new IXP or new Cluster/Node/Metro ✓
- ❑ No need to renumbering in future, one /64 can handle it at scale 💪
- ❑ Quite straightforward configuration ⚙️

Cons:

- ❑ Not supported on legacy router/software ▼
- ❑ Not suitable for existing large IXP ⚠️
- ❑ Need to educate member to use RFC 8950 in their configuration 📖

Thank you

Ari, CXC-IX

ari@cloudxchange.id

APIX Meeting #28

31 Oct 2023

#APIX28