A study in BGP Route Servers

1. Source:

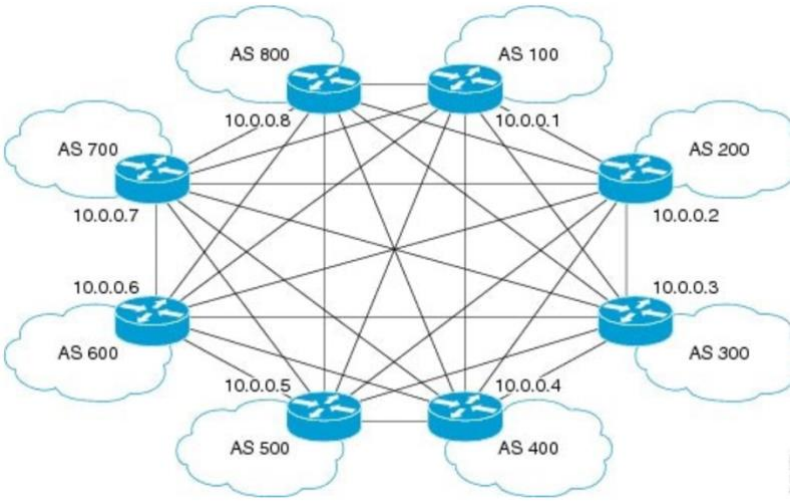
https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/iproute_bgp/configuration/xe-16/irg-xe-16-book/configuring-a-bgp-route-server.html

2. Preface

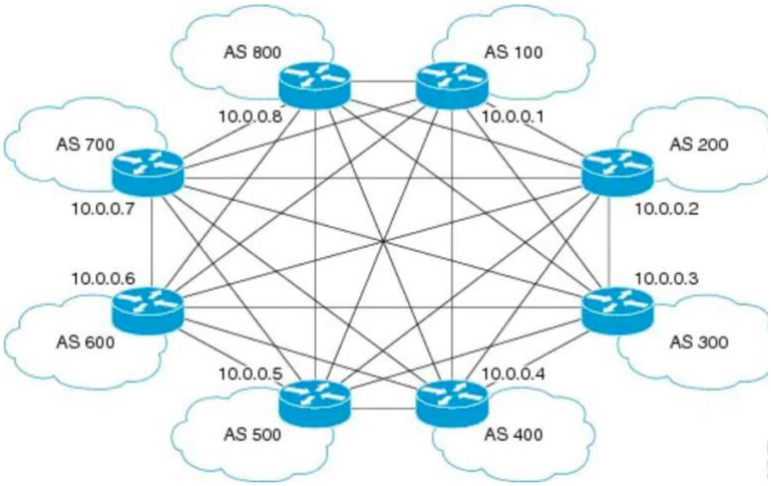
For a BGP router administrator on the Internet, having many peering agreements can be challenging, time consuming and costly to setup and maintain.

Hence just like a Route Reflector will allow you to have just a few iBGP peering agreements to serve many iBGP clients, having a BGP Route server will allow you to learn routes from MANY other Ass, without needing a to peer with each individual AS – Router.

So that will transform this:



To this:



Looks better doesn't it?! Notice that ALL routers have a direct route to all other Routers!

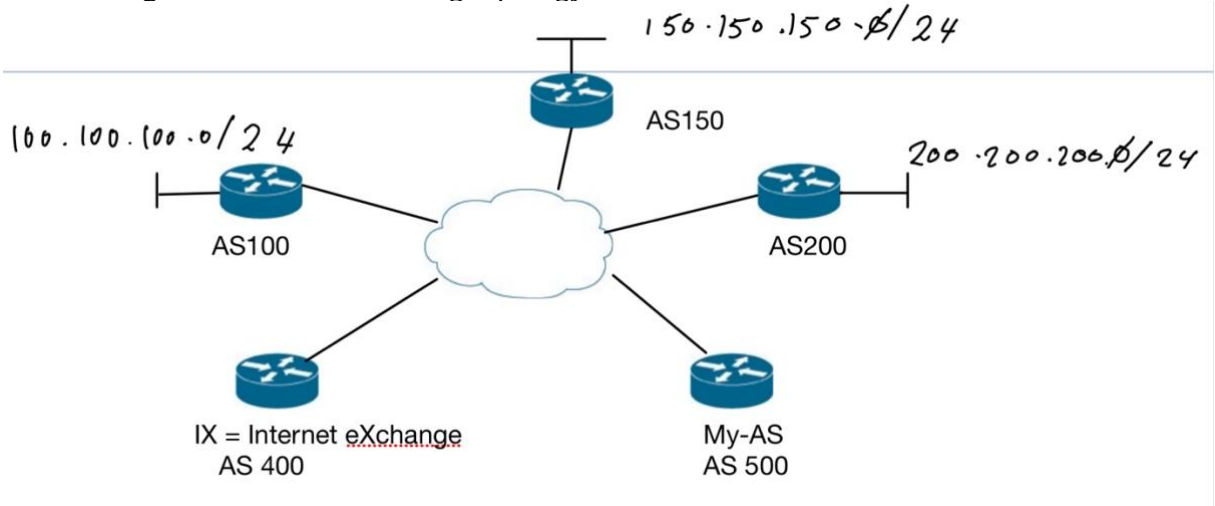
The Route-Server is like the "Route-Reflector for eBGP".

A study in BGP Route Servers

Your AS (eg 500) will still learn all the prefixes from all the other ASs, without the full – mesh peerings. You just need to peer with the Route Server, often an Internet Exchange Router (IX).

3. The topology used

Let's configure this in for following topology:



4. The standard configuration of the BGP routers

The BGP configuration of the three top routers is fairly basic; all three publish a (loopback) route into BGP and peer to the IX Router with AS 400 with IP address 10.0.0.13.

An example on BGP Router AS100:

```
router bgp 100
 bgp log-neighbor-changes
 network 100.100.100.0 mask 255.255.255.0
 neighbor 10.0.0.13 remote-as 400
```

The BGP Configuration on the IX Router (AS400) is very basic as well:

```
router bgp 400
 bgp log-neighbor-changes
 neighbor 10.0.0.1 remote-as 100
 neighbor 10.0.0.2 remote-as 150
 neighbor 10.0.0.3 remote-as 200
 neighbor 10.0.0.14 remote-as 500
```

Router "My-AS" (AS500) will only peer with the IX Router:

```
router bgp 500
 bgp log-neighbor-changes
 neighbor 10.0.0.13 remote-as 400
```

At this point in time, the regular BGP routes come in and NEXT AS within all AS-PATHS is 400, as the IX BGP router is the first AS in line:

```
My-AS(config-router)#do sh ip bgp
     Network          Next Hop            Metric LocPrf Weight Path
 *>  100.100.100.0/24 10.0.0.1                               0 400 100 i
 *>  150.150.150.0/24 10.0.0.2                               0 400 150 i
 *>  200.200.200.0    10.0.0.3                               0 400 200 i
```

Notice how the next-hop IPs are "directly reachable" IP addresses of the BGP routers for AS 100, 150 and 200, despite that fact that they come in via AS400. (IX) So no "next-hop self".

5. Convert the IX BGP Router to Route Server

On the IX BGP Route Server – Router, define the My-AS router as a route-server client in the BGP section:

```
IX(config-router)# address-family ipv4 unicast
IX(config-router)# neighbor 10.0.0.14 route-server-client
```

If you want to speed things up, you can restart the BGP peering in your lab.

6. Convert the BGP Router to be BGP Route Server – Client

At this stage the My-AS Router will go nuts as the prefixes that now come in do NOT have the AS set to the next – AS hop of IX:

"A **route server** does not put its own AS number in the AS-path; there is AS-path transparency. This means the route server client will receive updates in which the first AS number in the AS-path is not the sending router's AS number.

By default, a router denies an update received from an eBGP peer that does not list its AS number at the beginning of the AS-path in an incoming update. **Therefore, you must disable that behavior on the client in order for the client to receive the updates**."

Disable the first-AS check:

```
My-AS(config-router)# no bgp enforce-first-as
```

After this change, the peering is reset and the result can be observed from the Route-Server that publishes the prefixes to the "My-AS" Router:

```
IX# show ip bgp neighbors  10.0.0.14 advertised-routes

     Network          Next Hop            Metric LocPrf Weight Path
 *>  100.100.100.0/24 10.0.0.1                 0             0 100 i
 *>  150.150.150.0/24 10.0.0.5                 0             0 150 i
 *>  200.200.200.0/24 10.0.0.9                 0             0 200 I
```

As you can see, the AS number of the Route-Server (400) no longer appears in the AS Path.

As long as the next-hop IP addresses of the above prefixes are directly reachable for My-AS, the Route-Server routes will be imported into the BGP routing table:

```
My-AS(config-router)# do sh ip bgp
     Network          Next Hop            Metric LocPrf Weight Path
 *>  100.100.100.0/24 10.0.0.1                 0             0 100 i
 *>  150.150.150.0/24 10.0.0.2                 0             0 150 i
 *>  200.200.200.0    10.0.0.3                 0             0 200 i
```

And the proof is in the pudding:

```
My-AS# traceroute 100.100.100.1
VRF info: (vrf in name/id, vrf out name/id)
  1 10.0.0.1 0 msec *  2 msec
```

This concludes the study of the “Route-Server”.