

draft-hilliard-ix-bgp-route-server-operations

IXP Route Server Stuff



i n t e r n e t n e u t r a l e x c h a n g e

Nick Hilliard - INEX

Elisa Jasinska - Microsoft

Robert Raszuk - NTT I³

Niels Bakker - AMS-IX



i n e x
i n t e r n e t n e u t r a l e x c h a n g e

Problem Definition

Route Servers are widely used at IXPs for convenience

They allow IXP participants to implement peering with multiple other participants using a single BGP session

This introduces some interesting operational issues which need to be documented



i n e x
i n t e r n e t n e u t r a l e x c h a n g e

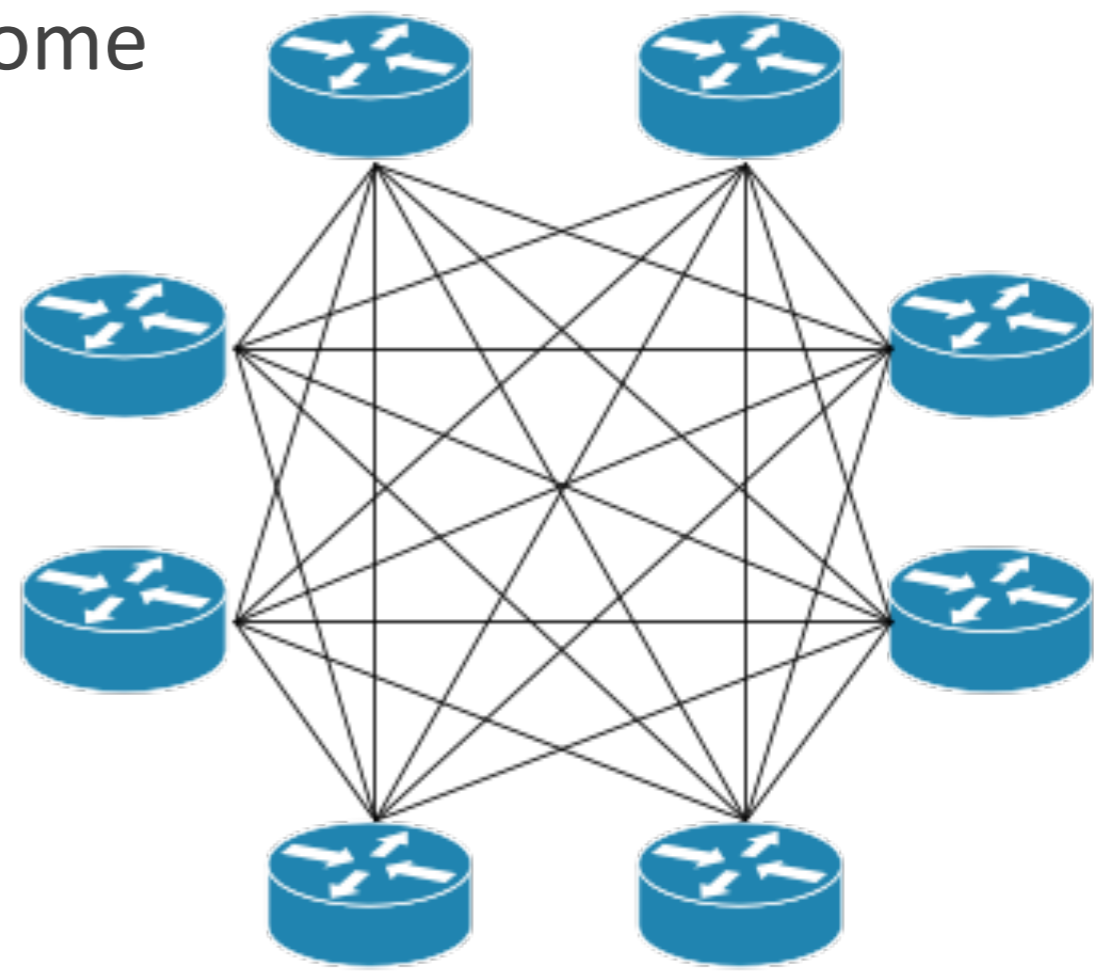
Why use Route Servers

Bilateral Peering is operationally troublesome

N^2 BGP sessions

Scope for admin error

Slow startup for new IXP participants





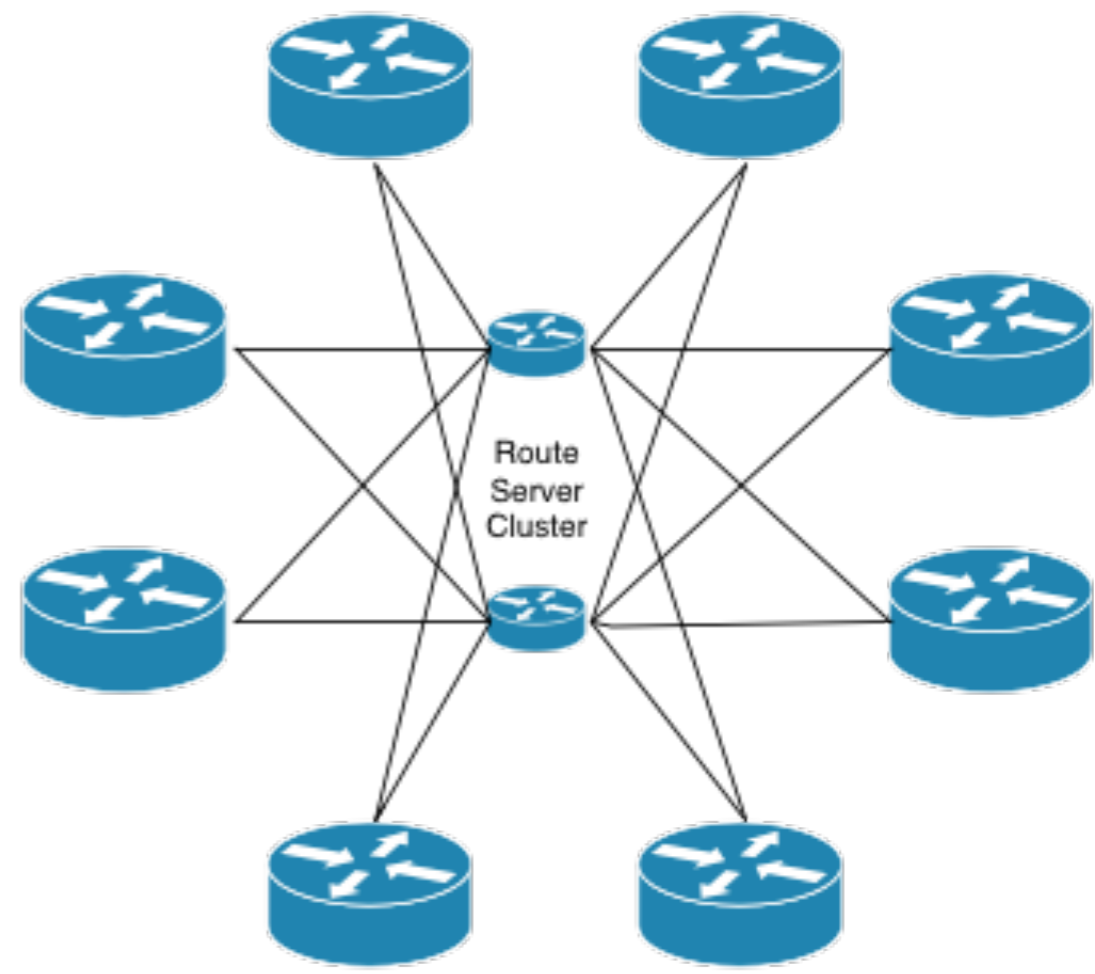
i n e x
i n t e r n e t n e u t r a l e x c h a n g e

Why use Route Servers

Route Servers allow interconnection to all other RS participants using a single BGP session

Operationally much easier for IXP users

Less overhead, quicker ramp-up time for new IXP users





How Route Servers Work

Regular BGP session

Outbound updates omit route server ASN

BGP next-hop is preserved

The intention is that Route Server clients see the correct next hop with the correct AS path

In the real world, per-client routing policy is required

Route-server participants do not necessarily want to peer with every other participant

If you don't do this, you can mix up transit and peering traffic

Worse still, you can cause BGP paths to be dropped to other IXP participants (path hiding)

Described thoroughly in draft-ietf-idr-ix-bgp-route-server



Scaling Problems

Route Server needs to calculate the BGP best path for each client

A direct consequence of the requirement to have per-client routing policy

In the worst case, scales according to N^2 where N is the number of RS participants

This affects CPU, memory and network bandwidth requirements of RS

The draft suggests ways of cutting resource requirements

View Merging and Decomposition - minimise the number of best path calculations required

Destination Splitting - split load over multiple route servers

No requirement for next-hop resolution - all next-hops on the same LAN anyway



Other Operational Issues

i n e x
i n t e r n e t n e u t r a l e x c h a n g e

Route Leakage

Almost all route server clients manually configure their routers which leads to prefix leakage on a regular basis

Prefix leakage on a route server is much worse than via bilateral peering

Both prefix filtering and prefix limits recommended

Export routing policies

There is no native BGP support for signalling routing policy changes via a neighbor session

Yet IXP participants need this functionality

This can be synthesised by various means (BGP communities, IRRDBs, IXP databases)

Layer 2 Reachability Problems

Route Servers will continue to announce prefixes if there are connectivity problems on the IXP

Currently no mechanisms to handle this



Summary

The draft summarises problems which IXP operators need to understand to run a reliable service

Also describes issues which need to be understood by route server BGP developers

Recommendations about best operational practice where they differ from normal BGP peering

Generally useful for IXPs