NSWI184 – Řízení počítačových sítí

Přednáška sedmá

Ondřej Zajíček, Kateřina Kubecová

2025-11-19

Internal BGP (Recap)

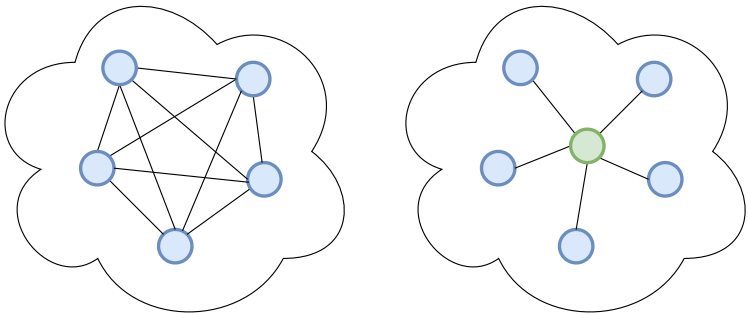
- Propagate paths received from EBGP to all routers in AS
- Do **NOT** propagate paths received from IBGP
- Therefore require full BGP mesh between routers in AS
- Generally keep path attributes unmodified, including BGP Next Hop
- BGP Next Hop points to BGP border routers or its EBGP neighbors
- BGP Next Hop resolved through IGP routes

IBGP Alternative Approaches

- Full BGP mesh does not scale well
- BGP route reflectors
- BGP confederations

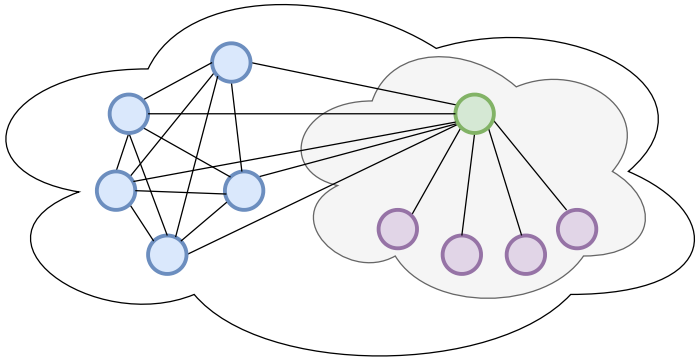
BGP Route Reflectors

- RFC 4456
- Central hubs for route distribution
- Replace full BGP mesh with hub-and-spoke topology
- Reduce IBGP sessions from $O(n^2)$ to $O(n)$
- Do not need to be real routers / to have data plane



BGP Route Reflector Concepts

- Route Reflector: The central hub
- RR clients: Routers that rely on RR to exchange paths with rest of AS
- Non-clients: Regular IBGP neighbors of RR
- Cluster: RR and its clients
- Cluster ID: Identifier of cluster, often Router ID of RR



BGP Route Reflector Operation

- Route from non-client $\rightarrow$ propagate to all clients
- Route from client $\rightarrow$ propagate to both clients and non-clients
- Route from EBGP $\rightarrow$ propagate to both clients and non-clients
- First two cases are route reflection, third is just regular EBGP to IBGP

BGP Route Reflector Path Attributes

ORIGINATOR_ID

- Router ID of ASBR introducing route into AS
- Necessary for consistent best path selection

CLUSTER_LIST

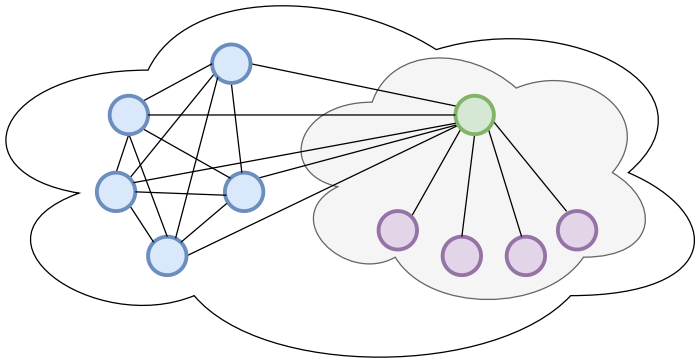
- Sequence of cluster IDs as route is propagated within AS
- RR prepends its cluster ID to reflected routes
- RR rejects received routes with its own cluster ID in CLUSTER_PATH
- Analogous to AS_PATH
- Loop prevention mechanism

Complex Route Reflector Topologies

- Route reflector with clients and non-clients
- Multiple route reflector clusters
- Redundant route reflectors
- Multiple RR clusters, each with redundant RRs

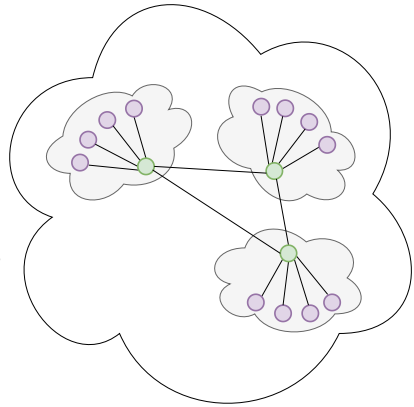
Route Reflector with Clients and Non-clients

- Some IBGP neighbors connected to RR as non-clients
- Route reflector and non-clients form full mesh
- Purpose: Transition from full mesh to RR-based setup



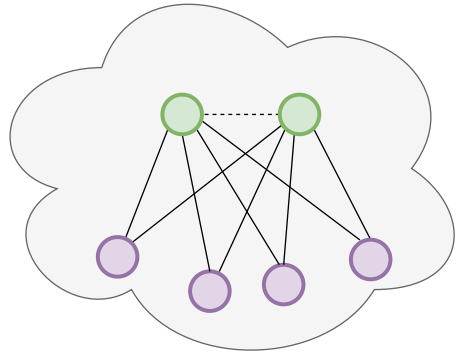
Multiple Route Reflector Clusters

- Multiple route reflectors, each forms one cluster
- Configured with different cluster IDs
- Must have disjoint set of clients
- Each client connected to only one RR
- RRs mutually connected as full mesh as non-clients
- Purpose: Split load between multiple RRs
- Purpose: Align RRs with network topology or administrative boundaries



Redundant Route Reflectors

- Multiple route reflectors within one cluster
- RRs configured with the same cluster ID
- RRs must have the same set of clients
- Clients connected to all RRs within cluster
- Purpose: Avoid single point of failure



Route Reflector Considerations

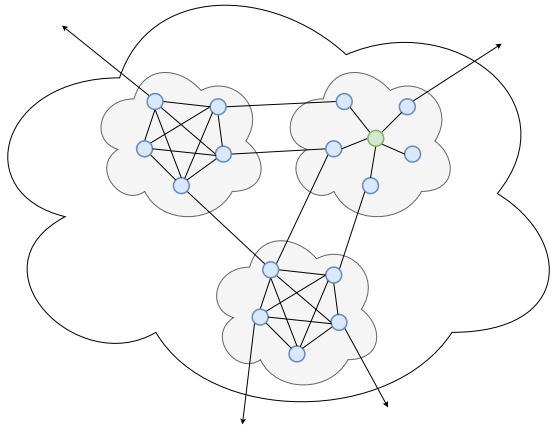
- Route reflectors do best path selection and only reflect best paths
- Best path selection may depend on IGP metric
- Also possible problems with MED attribute
- May lead to suboptimal routings

RR Suboptimal Routing Solutions

- Make clusters congruent with network topology
- I.e., intra-cluster distances $<$ inter-cluster distances
- BGP ADD_PATH – Announce all paths from RR to clients
- Optimal Route Reflection – Compute IGP metrics based on client location

BGP Confederations

- RFC 5065
- AS appears as single entity to external peers
- But internally divided into multiple sub-ASes
- Internal BGP only within each sub-AS
- Sub-ASes connected with External-ish BGP



BGP Confederation Concepts

- Confederation: Group of ASes that appears as single AS from outside
- Confederation ID: Common external ASN
- Member AS: One of ASes that form Confederation
- Member ASN: Internal ASN, usually from private range

Inter-member BGP

- BGP sessions between different member ASes within confederation
- Mostly behave as external BGP on member ASNs
- Special handling of AS_PATH attribute
- LOCAL_PREF, MED, and NEXT_HOP attributes can be passed unmodified

Confederation Segments in AS_PATH

- AS_PATH is not just list of ASNs
- It is list of typed segments, each of them is list of ASNs
- The common segment type is AS_SEQUENCE
- Confederations introduced new segment type AS_CONFED_SEQUENCE
- (There are also deprecated types AS_SET and AS_CONFED_SET)
- When route is propagated by inter-member BGP, member ASN is prepended as AS_CONFED_SEQUENCE
- When route left confederation, recent AS_CONFED_SEQUENCE is replaced with confederation ID

Confederation Segments in AS_PATH – Example

- Confederation 40
- Entering: SEQ[30 20 10]
- Inside: CONFED_SEQ[65530 65520 65510] SEQ[30 20 10]
- Leaving: SEQ[40 30 20 10]

Confederations with Shared IGP

- One IGP for whole confederation
- NEXT_HOP should be passed unmodified on inter-member BGP
- Optimal routing through whole confederation

Confederations with Per-member IGP

- Each member AS has its own IGP
- NEXT_HOP is reset on inter-member BGP like on EBGP
- Better isolation and control, but suboptimal routing

Accumulated IGP Metric

- RFC 7311
- When next hop is reset, add IGP metric to AIGP attribute
- Compare routes based on sum of local IGP metric and AIGP attribute
- Allow optimal routing through multiple IGP domains with consistent metrics

Inspecting BGP

- MRT dumps every 5 minutes: `https://ris.ripe.net/docs/mrt/`
- BMP: send online BGP data to a listener
- BGP global looking glass: `https://bgp.tools/`
- IXP looking glasses: `https://lg.de-cix.net/`

Non-standard BGP

- Anycast: the same prefix announced from many places
- Backup link: using AS Path stuffing + communities
- IXP outreach: linking a little local IXP to a big remote one
- Feed only: to probes checking the world BGP status