BGP Color-Aware Routing (CAR)

draft-ietf-idr-car-02

D. Rao, Cisco Systems
S. Agrawal, Cisco Systems
C. Filsfils, Cisco Systems
Bruno Decraene, Orange
L. Jalil, Verizon
Y. Su, Alibaba
K. Patel, Arrcus
J. Uttaro
K. Talaulikar, Cisco Systems
H. Wang, Huawei
J. Guichard, Futurewei
R. Rokui, Ciena
D. Steinberg, Steinberg Consulting
Context

• Adopted by WG at IETF 114

• Broad multi-vendor and operator support during adoption and WGLC

• Two interoperable implementations

• In progress
  – Draft updates
    > Address Routing Directorate and Ops Directorate/NM review
    > Address shepherd’s review feedback

• Implementation update
  – Route Type 2 support
Focus for this discussion

1. Motivation of the two route-types in CAR SAFI / VPN CAR SAFI
2. Procedures for type 1 and type 2 Routes for CAR SAFI
3. Use of CAR SAFI / VPN CAR SAFI for infrastructure routes
4. SAFIs for Service Routes steered over BGP CAR
BGP CAR SRv6 – Motivation for IP Prefix route

Service steering over BGP CAR SRv6 established intent aware multi-domain paths (Section 9)

Cases as per RFC 9252

- **Non-routed Service SID**: SRv6 Service SI allocated by an egress PE is not routed. Intent aware path is provided by SR policy to egress PE.
  - SR-Policy redistributed into BGP CAR (E, C) – Type 1 route

- **Routed Service SID**: Services are steered to SRv6 Service SID that is allocated by egress PE from intent aware locator
  - Intent-aware SRv6 Locator or Summary IPv6 prefix reachability at ingress PE - Type 2 route
  - BGP specific or IGP-FlexAlgo shared per-intent locators
  - Advertisement from BRs (redistribution from IGP-FlexAlgo), or from PE
IPv6 Prefix Route

- IP Prefix == Intent
- Routable, installed in IPv6 forwarding table
- Route processing semantics same as RFC 4271, RFC 2545 (IPv6 Unicast)
- Consistent CAR semantics & procedures for color-aware next-hop/SID selection, resolution, route policies, AIGP - same as Type-1 route
  - Route E2E color carried in LCM-EC
  - Section 10
Motivation for two route-types in CAR SAFI

Type 1 - NLRI key: E,C
- (E,C) route is used when same IP address (E) on a router needs to be used to create multiple intent aware paths in transport

Type 2 – NLRI key: IP Prefix (E)
- Used when intent and IP address has a 1:1 relationship such as in SRv6 where a unique locator is assigned for a given intent.
  > No requirement to have multiple instances of the locator with different colors

- Complete reuse of well-known, existing BGP routing and CAR mechanisms across both route types
- Operational consistency and simplicity with common procedures in same SAFI
- Single SAFI sufficient to support both SR/MPLS and SRv6

- Both route type 1 and 2 are relevant for SRv6 intent-aware routing
  • There are also use cases both are used simultaneously. (Eg. Section 9.2 and Appendix C.3)
Flat Routing model as BGP IP/LU for both route types

- Both route types maintain functional and operational consistency with BGP-IP/LU
Choice of route-type

https://mailarchive.ietf.org/arch/msg/idr/lmqA2VSozVemjx8lO0WAol37mXCM/

Jeff Haas query - “Did the authors consider simply designating a targeted color, say 0 on the existing type 1, to avoid creating a new type for the NLRI? Compare vs. the best effort transport-class in the more recent -ct work.”

• The alternative i.e. reuse Type-1 (E,C route) with color 0 overloaded the type-1 route semantics and creates exceptions based on color values, hence the choice to use a distinct route-type

• Observation:
  - Regardless of encoding, it doesn’t change actual semantics of the route
  - Hence, the issues being raised are not specific to use of a separate route-type for IP Prefix
  - Or even specific to CAR SAFI alone
Issues raised about IP prefix route

1. Question about clarity of procedures

Clarification:

Both route types use exact same procedures defined in CAR draft

- Route origination (section 2.3)
- Route validation (section 2.4)
- Color-aware recursive resolution (section 2.5)
- AIGP metric computation (section 2.6)
- Path availability (section 2.7)
  > ECMP/primary-backup
- BGP Color EC handling when present (section 2.5, 2.10)
- LCM-EC handling when present (section 2.8, 2.10)

- Shepherd's review provided feedback to enhance the introduction section to refer to both route types and some reordering of text, references etc. Being addressed in next version.
Issues raised about IP prefix route

2. Extension for CAR IP prefix route is relatively recent

Clarification:

- IP prefix route - so same semantics of RFC 4271/2545
  - Reuses the mature IP unicast implementation/design

- Additionally, it's part of CAR SAFI, so already defined semantics for Type-1 such as BGP color EC, LCM EC etc. apply (previous slide)

- Been reviewed heavily

- Implementation for a few months
  - Experience confirms what was stated during WGLC - the semantics/procedures of the route type are not new but reuse well-established ones.
Issues raised about IP prefix route

3. Whether security risk/”walled garden” leak caused by this route

Clarifications:

- Security section of CAR draft: “Since CAR SAFI is a separate BGP SAFI that carries transport routes for routers in the operator network, it provides automatic separation of infrastructure routes from service routes that may be carried in existing BGP SAFIs such as BGP IPv4/IPv6 (SAFI=1), and BGP-LU (SAFI=4).”

- Further, section 9.3 clearly states that using CAR SAFI provides:
  
  “Automatic separation of SRv6 locator (transport) routes from Internet (service) routes

  1. Preventing inadvertent leaking of routes

  2. Avoiding need to configure specific route filters for locator routes”

- Important operational benefit CAR SAFI (Automatic separation of infra routes in CAR SAFI and avoid leak into IP unicast and vice versa)

- So CAR SAFI in fact does the opposite – it provides the separation of infrastructure and service routes, reinforcing the “walled garden”

- Shepherd has asked to be explicit about advertising infrastructure routes across sections… will add.
4. Comparison to EVPN Type-5 and whether any related concerns apply

Clarifications:

- CAR carries infra routes, not service routes
- Secondly as mentioned earlier, no redistribution or leaking of IP unicast SAFI routes into CAR or vice versa.
- If redistribution is done, this case is same as any two BGP SAFIs, not specific to CAR route-type 2
  - Example – between SAFI 1 and SAFI 4, OR between SAFI 4 and BGP-CT
  - In CAR, as with these other SAFIs, the path attributes such as as-path and cluster list are used and propagated. EBGP/IBGP semantics are followed. So, there should be no issue with loops specific to CAR route type-2 or CAR SAFI

Issues raised about IP prefix route
5. Which SAFIs used for service routes?

Clarifications:

- Section 3 describes the SAFIs that are steered into BGP CAR such as IPv4/IPv6, L3VPN, PW, EVPN, FlowSpec, and BGP-LU.
  - Same as steering over SR-TE policy (RFC9256) or IGP Flex Algo
  - Service routes are carried in existing SAFIs
- As per shepherd’s suggestion, we will add explicit SAFI values.
Issues raised

6. Use of VPN CAR SAFI

Clarifications:

• VPN CAR enables signaling of intent awareness end-to-end: customer site to customer site across provider networks
  - Provides ability for a customer site to select transport paths through specific PEs
  - Carries infrastructure routes
  - Use case: CSC like deployment

• VPN CAR is separate SAFI
  - Defined since initial version of CAR draft (Section 8)

• Follows same VPN semantics as defined in [RFC4364], the difference being that the advertised routes carry CAR NLRI with the VPN RD
  - VPN RD distinguish CAR routes from different customers
  - CAR SAFI procedures apply

Shepherd has asked to add specific section references for CAR SAFI procedures applicable, will add
Summary

• Flat Routing model as BGP IP/LU for both route types
• Same CAR procedures, extensions, extensible encodings apply to both the route types
• Route type 2 is IP prefix
  • Same semantics of RFC4271/RFC2545
  • Semantics remain same regardless of NLRI encoding format
  • Most issues raised are not related to format choice or specific to CAR SAFI
  • CAR SAFI prevents inadvertent leaking between infrastructure routes and service routes
Next Steps

- Publish -03 version with updates based on review feedback