IETF 113 – Vienna
Mar 2022

BGP Color-Aware Routing (CAR)

draft-dskc-bess-bgp-car-03

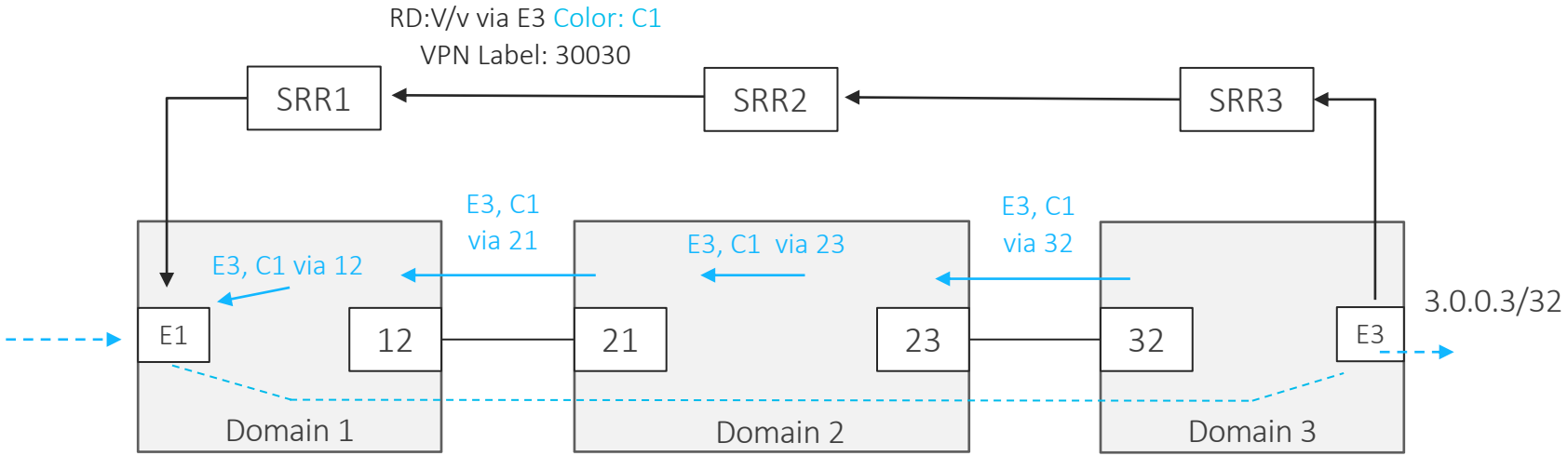
D. Rao, Cisco Systems
S. Agrawal, Cisco Systems
C. Filsfils, Cisco Systems
K. Talaulikar
L. Jalil, Verizon
Y. Su, Alibaba
D. Steinberg, Steinberg Consulting
K. Patel, Arrcus
H. Wang, Huawei
J. Guichard, Futurewei
Bruno Decraene, Orange

BGP Color-Aware Routing

- Define BGP based routing solution to establish end-to-end intent-aware paths across a multi-domain network environment
  - Intent : Example – low-latency path between two PEs

- Color represents intent in signaling
  - draft-ietf-spring-segment-routing-policy
  - draft-ietf-idr-segment-routing-te-policy
  - Color is the standard way to represent intent
    - Carried in BGP Color Extended-Community in BGP service routes (L3VPN, EVPN etc)

BGP Color-Aware Route & Automated Steering



- E3, C1 is a Color-Aware BGP route in underlay that provides intent-aware path to E3
- A C1 Colored service route RD:V/v from E3 is automatically steered onto a Color-Aware path (E3, C1)
  - Color: C1 carried in BGP Color Extended-Community attached to RD:V/v
- Steering for all services (L3VPN, EVPN, Internet/global table etc)

© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Confidential.

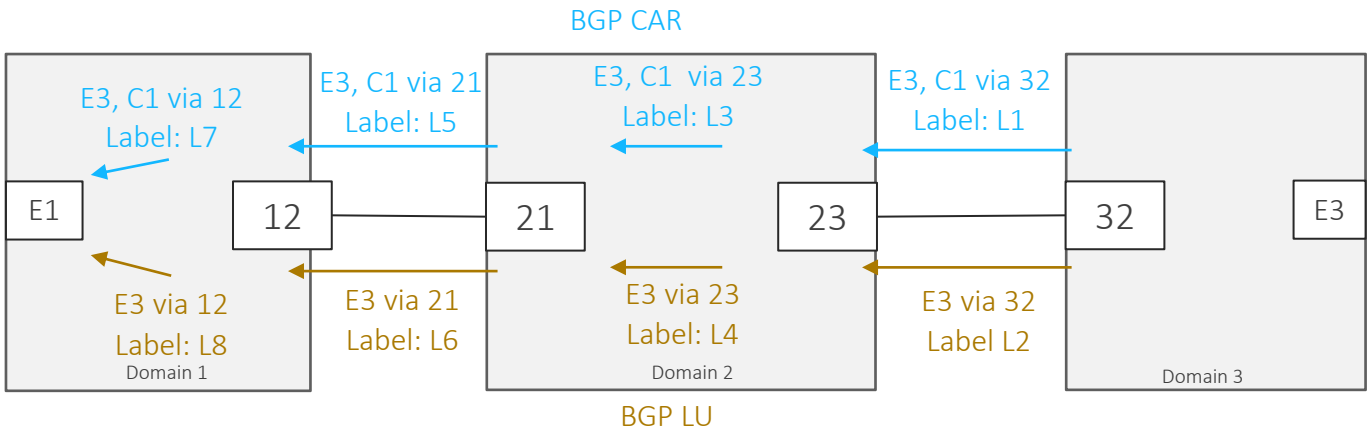
BGP CAR Overview (Refresher)

- Solution draft describe the following aspects
  - New SAFI in BGP
  - Desired Data Model
  - Multiple encapsulations, their signaling and validation
  - Efficient and extensible NLRI
  - Handling of multiple color domains
  - Route resolution & steering mechanisms
  - Scale Analysis

© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Confidential.

New SAFI in BGP

- Need ability to signal multiple instances of the same prefix for each color (i.e., intent)
- Evolution of best effort BGP-LU SAFI (RFC 3107/8277)
  - Modernize, Address some of the limitations with BGP-LU
- Maintain functional and operational consistency with BGP-LU
- No need to use VPN constructs and machinery at every transport hop



© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Confidential.

CAR NLRI Proposal

- NLRI Key – E, C
  - E : IPv4 or IPv6 Endpoint Prefix (Network-wide Unique)
  - Color : 32-bit value (same as SR-TE Policy)

- Color distinguishes per-intent instances of same prefix
- Color also indicates intent provided by route
- Color is same as in BGP Color Extended-Community
- Color is consistent across devices within a “color domain”
- Optimal for 99% deployments under admins with consistent color mapping

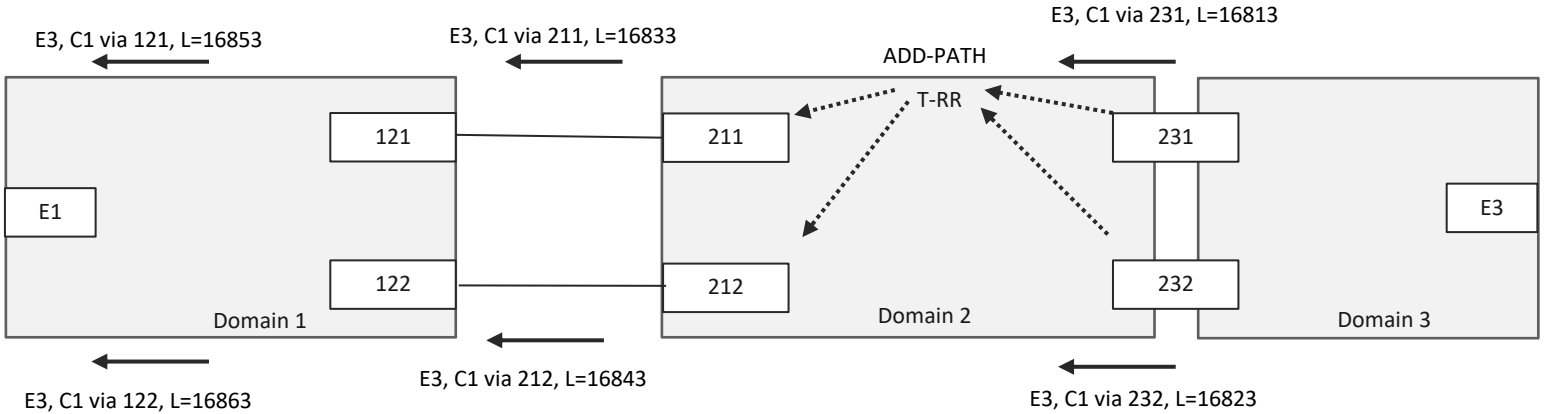
© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Confidential.

CAR NLRI – E, C

- Simplest data model, precise
- Identical routing semantics as BGP IPv4/v6, BGP-LU
  - Efficient route processing, storage
  - No need for VPN import/export each underlay hop
- Inherently provides ECMP-aware/backup paths at every hop
  - Faster, localized convergence
  - No need for VPN import to bring diverse path together with complex workarounds
- Most efficient for subscription
  - [E, C] direct lookup
- Consistent with SR Policy data model

© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Confidential.

Path Availability & Domain-local Convergence



- (E, C) NLRI provides ECMP or backup paths at each hop (single label entry)
- Localized convergence with Next-Hop Self
  - E.g., 231 failure is handled locally within domain, churn is not propagated beyond 212 and 211
- BGP ADD-PATH at T-RR for redundant path availability
- Note: BGP-CT fails to provide domain-local convergence and BR failure churn suppression

© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Confidential.

Extensible, Future-Proof NLRI Encoding

- New SAFI allows opportunity for better NLRI design
  - Existing SAFIs carry key (prefix) and non key information(eg: label in VPN,BGP-LU, EVPN)
    - Hard coded in per SAFI specification
  - BGP CAR provides structure to this non-key information for future extensibility and flexibility
    - No good reason to inherit constraints of current SAFIs, e.g., only a MPLS label field in NLRI
- Encode a NLRI (Route) Type
- Encode a key length
- Encode non-key TLVs
- Per route unique data in NLRI non-key TLVs; rest in Attribute
  - Provides packing efficiency for BGP updates

© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Confidential.

Encapsulations

- Ability to signal multiple encapsulations for a CAR route
  - Signaled via Non-Key TLVs
    - MPLS Label(s), Label-Index, SRv6 SID(s) etc
  - Can signal separate label (or equivalent) values for different encapsulations
  - Efficient, preserves packing – e.g., label-index
  - Beneficial for co-existence, migration & interworking
    - Efficient signaling, automated migration handling, operational simplicity
      - Avoids duplicate routes for each encap
      - Avoids separate control plane planes for distribution

© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Confidential.

Observation

- Discussion at interim acknowledged limitations with BGP-LU/VPN SAFI NLRI (RFC8277 encoding), desirability to address them
  - Including from BGP-CT co-authors
- BGP-CT inherits the same limitations
- BGP CAR addresses these limitations, and more

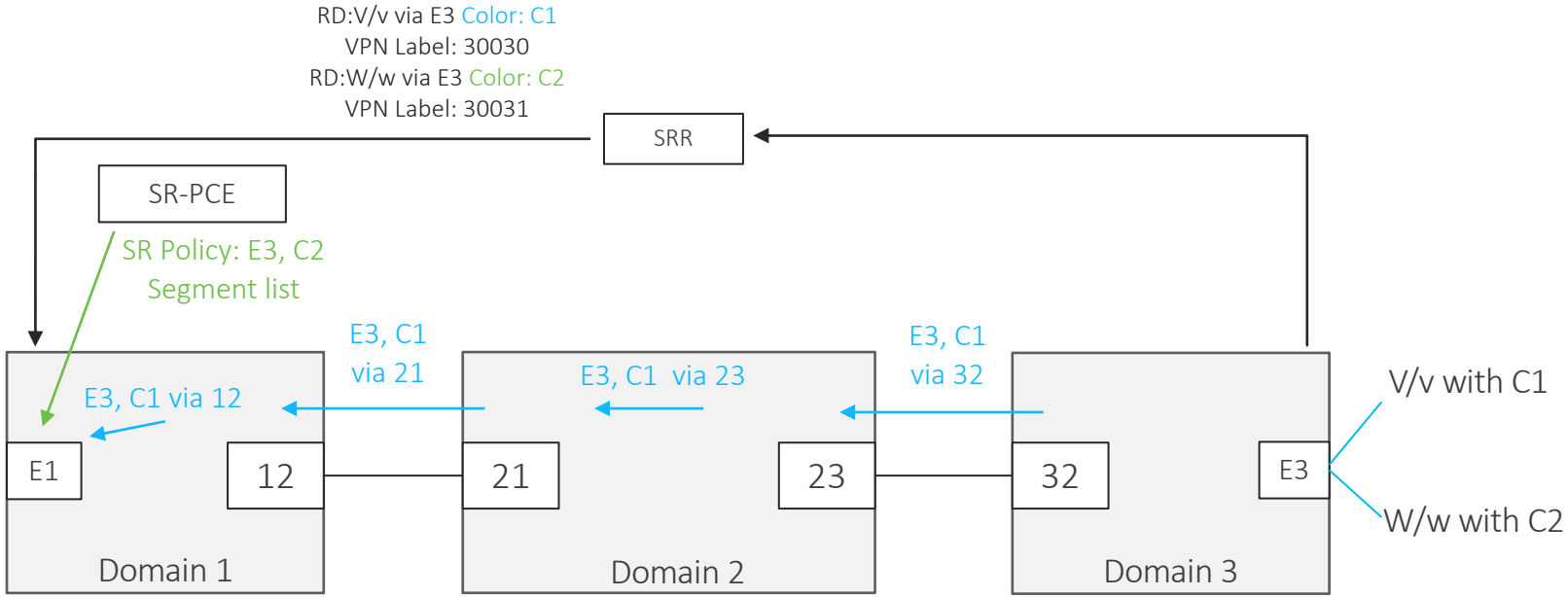
© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Confidential.

CAR Next-Hop Resolution

- Resolution is recursive and color-aware
  - (E, C) via (N, C)
- (N, C) provided by other color-aware mechanisms
  - SR Policy, IGP Flex-Algo, or BGP CAR itself
- Resolution will also be mapped to traditional mechanisms
  - RSVP-TE
  - IGP/LDP
  - BGP-LU
  - Supports brownfield, incremental deployment

© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Confidential.

Seamless BGP CAR and SR Policy co-existence with E,C model



- E1 automatically steers the received colored service routes as follows:
  - V/v via (E3, C1) provided by BGP CAR
  - W/w via (E3, C2) provided by SR Policy
- Seamless compatibility with SR-Policy architecture - color, steering, fallback etc
  - Supported in multiple implementations, deployed

© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Confidential.

Multiple Color Domains

- Network domains where color-intent mappings are different
- Local-Color-Mapping (LCM) Extended Community
  - Optional, only used if routes go across a color domain boundary
  - Color re-mapped and rewritten into receiving domain's color at a color domain boundary
  - Color Ext-Comm sent with service routes also gets re-mapped in parallel
- CAR NLRI (E, C) is immutable, preserved e2e
  - Eases tracking of route
- E (Prefix) is unique in inter-domain transport network (e.g., PE)
  - Makes (E, C) unique e2e even if C is local to a color domain

© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Confidential.

Reiteration

- CAR NLRI does not get rewritten across color domain boundaries
- Semantics of color defined are clear – in NLRI for same color domain, in LCM for routes from a different color domain
- CAR does not need eBGP add-path to get individual CAR routes across eBGP boundary
  - Color in NLRI provides distinction needed for specific intent
- CAR supports both IP host and prefix

© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Confidential.

Additional aspects in draft

- CAR provides flexibility to operator for various dataplane designs, flat and hierarchical for scaling
  - CAR draft provides analysis of trade-offs and optimizations
- CAR design enables extension of intent-aware routing to PE-CE networks (VPN CAR)

© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Confidential.

Summary

- CAR is evolution of BGP-LU, but intent-aware
- CAR is extensible, built to accommodate new use-cases, multiple encapsulations efficiently
  - Defines a base framework that can be extended with low overhead
- CAR definitions focus on better protocol performance & scaling
  - Preserves update packing efficiency, memory storage
  - Avoids route duplication during migration
- CAR works seamlessly across traditional networks (LDP/RSVP-TE/BGP-LU)
- It is also totally compatible and consistent with SR-Policy/IGP-FA solution

© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Confidential.

BGP CAR sample output (IOS-XR)

```
RP/0/0/CPU0:ABR# show bgp ipv4 car unicast 10.11.12.13/32 color 3
Sat Jan 22 12:32:46.172 PST
BGP routing table entry for [1][32][10.11.12.13][3]/72
Versions:
  Process          bRIB/RIB  SendTblVer
  Speaker              2          2
Paths: (1 available, best #1)
  Not advertised to any peer
  Path #1: Received by speaker 0
  Not advertised to any peer
  Local
    192.168.0.3 (metric 30) from 192.168.0.2 (192.168.0.3)
      Received Label 3
      Origin IGP, localpref 100, valid, internal, best, group-best
      Received Path ID 0, Local Path ID 1, version 2
      Originator: 192.168.0.3, Cluster list: 192.168.1.2
      CAR Non-Key TLVs:
       Label-Index 801
       Unknown TLV:
        04020101

Update dump:
Dump of the BGP CAR update
ffff ffff ffff ffff ffff ffff ffff ffff
0083 0200 0000 6c90 0e00 4c00 0153 04c0
a800 0300 1a09 0120 0e0e 0e0e 0000 0003
0103 0000 3102 0400 0003 2103 0201 0116
0901 200e 0e0e 0e00 0000 0201 0300 0031
0204 0000 0321 1009 0120 0e0e 0e0e 0000
0001 0103 0000 3140 0101 0040 0200 4005
0400 0000 6480 0a04 c0a8 0102 8009 04c0
a800 03
```

© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Confidential.

Next Steps

- Continue to address use-cases & requirements listed in problem statement
- Request collaboration & review from Working Group
- Problem statement drafts merge effort is ongoing
- 2 known implementations, with interoperability
- Ready for WG adoption

© 2020 Cisco and/or its affiliates. All rights reserved. Cisco Confidential.