

CAR-CT Adoption Poll: Part 1 – Information [Chair’s notes] (7/6 to 7/25)

General comments:

Response to Gyan’s definition.

[Ketan]: The Gold, Bronze and Silver classes of the past may be somewhat limited and may not cover a wider set of “intents” that are now required [to be] address[ed]. When referring to QoS, perhaps PHB is what is being referred to? Which is again limited. Not all intents are related to QOS.

[Ketan]: The notion of color was introduced in BGP by RFC5512 and predates SR. That said, I do agree that “color” based steering of BGP services over SR Policies using Color Ext-Comm is probably the most widely implemented and deployed solution that leveraged “color”.

[Luay]: These are congestion control mechanisms and not sufficient for today's needs. We will always have QoS but it's one piece of the pie since many of the steering use cases needed are not QoS dependent in an IETF-based network

Response on Sue’s view on Intent: There is no clear and universal intent-based language in 2022.

[Kaliraj:] BGP implements the Internet’s policy and economic infrastructure. Enabling customers and providers to express SLA in BGP allows for Automation and further facilitates such economic interactions. It allows [networks] to create a slightly better feedback loop between inter-domain networks and avoids over provisioning of provider networks. 5G topology is one use case.

[Ketan]: I agree [with Sue on intent] and this is very important. We need to be careful when using the terms “intent-aware” or “intent-based” in IDR documents.

[Bruno]: +1 (on no clear language for intent.)

[Aravind]: There are several applications I have heard from my customers I work with where intent (color) is expressed to treat traffic flows in a particular way. For example, in order to service chain or to offer security services we would need to redirect them to the security component (firewall). In order to [handle] this at a granular level we can leverage BGP flowspec (as Kaliraj points out in 2c on ingress). The path to service chain component itself can be on any transport (RSVP-TE/SR-TE) and the SLA treatment is given to a particular flow. So I (and my customers) think Flowspec redirecting traffic over transport-class aware paths is a useful and desired feature.

[Natrajan (NV)] I agree that intent is more than a color. Seamless MPLS networks have been able to guarantee intent using SRLG/admin-groups within a single TE domain. In Brownfield RSVP-TE deployment, there are adequate ways to do the above and I see Greenfield SPRING networks catching up to the same. However, BGP did not produce a holistic way of cross cross-connecting these TE characteristics that are necessary for guaranteeing the SLA end to end until now. BGP needs an identifier to inter marry these TE characteristics across multiple TE domains. That identifier is the color. The second part is how flexible it [Color] is for services to express intent (language)?

[Sue]: IRTF document on definitions of Intent
(<https://datatracker.ietf.org/doc/draft-irtf-nmrg-ibn-concepts-definitions/>)

So let us use color as BGP defined value and transport class as BGP defined value.

[Ketan]: Correct. So, color is an existing BGP-specific abstraction of an “intent” (English word). [BGP-specific color exists in] ... SR Technologies (implemented, deployed, and understood [by those deploying

Summary of CAR-CT Adoption Poll – Part 1 from 7/6 to 7/25 (Weeks 1-3)

it]). However, there is nothing that precludes “color” based steering of BGP services over other tunneling or encapsulation technologies.

[Bruno]: +1 (plus BGP already has a color extended community, and BGP-SR policy uses color so I think there is not a need to define a different term.

[SA] Yes. Color has been used to both express the desire to go over certain tunnel by attaching color extended community to service route (colored route) and to create paths/tunnels that satisfy requested intent called color aware route (example SR policy architecture).

Q1: What is the customer need driving the use of Color to express Customer Intent?

General responses:

[Halpern] (indirect) Want better mechanism to set-up paths across AS for a single Intent/QoS.

[Moses]: I believe the primary motivation here is the need for the ability to program the network (Network Programmability). Service providers need the ability to program different network paths that are optimized for Service Level Objectives(SLOs) and Service Level Expectations (SLEs) contracted by the customers.

- SLOs are directly measurable indicators such as guaranteed bandwidth, latency, packet loss rate etc.
- whereas SLEs are not directly measurable attributes such as path diversity/”disjointness”, security, encryption, isolation, geographic restrictions, etc.

Providers are interested in developing the capabilities to provide differential carriage treatment to achieve the SLOs and SLEs within a single domain, across multiple domains and across multiple autonomous systems. Tagging the service routes can be performed by the provider's service nodes customer at attachment points, customer end devices or even by a SDN controller that monitors the end-to-end performance of service/application.

[Moses]: Some of the use cases:

- Transport Network Slicing to support 5G Slicing (URLLC, eMMB, mMTC, etc.)
- Dual-plane/ Multiplane forwarding for isolation between primary and secondary traffic flows
- Dedicated network slice for high value wholesale customers
- Carriage of Edge Compute services with stringent SLAs

[Luay]: [this motivation] Could be driven by customer or by the infrastructure to treat traffic with specific SLOs. Network slicing is driving more of these use cases. However, it's more than QoS and most of your questions assume QoS is the main mechanism.

[Srihari] with new applications looking for specific SLA/SLO (many of which are being defined at TEAS) I have heard customers asking for some of this which can be generalized as intent-aware end-to-end paths. It will be more than cos or phb treatment.

[Shradda]: I have interacted with a number of operators as part of defining requirements for inter-domain intent-aware routing as described in

- <https://datatracker.ietf.org/doc/draft-hr-spring-intentaware-routing-using-color/>
- The main driving factor as per my knowledge is that the customer networks are segregated into multiple IGP domains in order to scale the network. There is a need to establish per-intent e2e

Summary of CAR-CT Adoption Poll – Part 1 from 7/6 to 7/25 (Weeks 1-3)

paths across these domains. A number of mechanisms exist today for establishing intent-aware paths across multiple IGP domains but the main driver for a BGP based solution is the high scalability and loose coupling across domains that BGP provides.

- The service routes tagged with color helps to steer traffic onto the right intent path through the network.
- The QOS/SLA measurements of traffic between data center application to user, has a much larger scope. The path may span multiple networks that are operated by completely different organizations. In my knowledge, today the DC application to user, path-SLA measurements are done using overlay applications. I haven't heard the strong need to replace the application layer measurements with network layer protocol doing it. If we are going to consider such a requirement a number of aspects need discussion such as scalability, network stability, security, authentication issues etc.

Question 1-a: Are applications requesting to be able to tag their routes with SLAs (color) at the service level?

[Ketan:] Yes. However, the tagging may be done by customers OR by operators for their customers/applications. This is what is widely implemented and deployed today with SR]

[Bruno:] Different applications/customers have different requirements in terms of routing paths. [For example, they] may optimize for bandwidth, vs. link delay, etc. [One usual niche use case] is two different planes.

- [Bruno] Note that we already have the need and solution within the single IGP (e.g. FlexAlgo, SR policies). IMHO the question is extending this to BGP, in a way allowing each AS to use its own solution for its Intra-AS Routing. Such AS RSVP-TE in one AS, FlexAlgo in another AS, [and] SR-policy in another AS.

[Natrajan (NV):] Both a and b are true.

[Swadesh (SA)] Yes, routes are getting tagged to send traffic over path that satisfy some intent. This tagging can be done by operators or by their customers. Intents means more than just QOS. It could be low latency, avoidance of resource (link, node or domain) etc. Measurement may be based on intent type.

[Gyan:] Yes.

- Tagging has been done via [the] QOS model providing Value added service tiering for 20 years with Gold, Bronze, Silver SLA based strictly on traffic classification and PHB queuing mechanisms.
- SLA based on “coloring” meaning steering overlay to underlay slice underpinning had started initially with RSVP-TE steering.
- L3 VPN services with per VRF TE mapping next hop re-write which requires a separate loopback per SLA / Color.
- With the advent of SR-TE the concept of VPN service route coloring expanded to micro and macro flow “intent based” steering.
- Network slicing is a major use case that spans beyond just 5G to Enhanced VPN+ service offering which can use SR-TE / Flex Algo to color the underlay.

[Shraddha Hedge:] I have interacted with a number of operators as part of defining requirements for inter-domain intent-aware routing as described in

<https://datatracker.ietf.org/doc/draft-hr-spring-intentaware-routing-using-color/>

- The main driving factor as per my knowledge is that the customer networks are segregated into multiple IGP domains in order to scale the network. There is a need to establish per-intent e2e

Summary of CAR-CT Adoption Poll – Part 1 from 7/6 to 7/25 (Weeks 1-3)

paths across these domains. A number of mechanisms exist today for establishing intent-aware paths across multiple IGP domains but the main driver for a BGP based solution is the high scalability and loose coupling across domains that BGP provides. The service routes tagged with color helps to steer traffic onto the right intent path through the network.

- The QOS/SLA measurements of traffic between data center application to user, has a much larger scope. The path may span multiple networks that are operated by completely different organizations. In my knowledge, today the DC application to user, path-SLA measurements are done using overlay applications . I haven't heard the strong need to replace the application layer measurements with network layer protocol doing it. If we are going to consider such a requirement a number of aspects need discussion such as scalability, network stability, security, authentication issues etc.

[Mahesh]: As Ketan points out tagging can be done by end customer or by operators in their own applications. I believe this is what is implemented using SR.

Question 1-b: If so, is it due to QoS/SLA measurements on traffic between Data Center applications and user applications (such as applications on phone)?

[Ketan]: We have requirements (“intents”) like low latency, low loss, best-effort, specific bandwidth guarantee, avoidance of certain parts of the network, etc. Not all “intents” need to manifest in QoS Policy or require measurements.

[Bruno]: This may either dynamic measurement or static attributes. Immediate use cases use static attributes.

[Natrajan (NV)]: Both a and b are true.

[Swadesh (SA)] ... Intents means more than just QOS. It could be low latency, avoidance of resource (link, node or domain) etc. Measurement may be based on intent type.

[Gyan]: Yes. What operators have found is that QOS traffic classification is not enough and being able to provide a way to take VPN service routes and somehow map them to different paths based on cSPF TE constrains such as delay, bandwidth etc is critical for voice and video delay sensitive and mission critical customer traffic.

[Mahesh Jethanandani]: I think there is a need to separate requirements of applications like latency, loss, bw, and best effort from how it maps to the underlying network. Not all the requirements can be mapped into a QoS policy, e.g. network paths to avoid.

Question 2: In the distant past QoS was hard to set-up seamlessly as a QoS pathway across multiple Autonomous systems (AS).

General comments:

[Bruno]: We probably need to distinguish Multiple AS versus multiple Administrative Domains. A single Administrative Domain may cross multiple ASes, e.g. due to organization boundaries or technical scaling. I would hope that agreeing on a QoS/color within single administrative domain should be doable regardless of the number of ASes. Plus, in my view, a color does not match to packet forwarding QoS

Summary of CAR-CT Adoption Poll – Part 1 from 7/6 to 7/25 (Weeks 1-3)

treatment ('a la diffserv) but rather to a to a [set of] different routing paths toward the same destination (NLRI).

[Luay (LJ)]: Seamless could be accomplished in many ways, signaling is a big part of it. It's not a QoS pathway, I'd prefer to call it the SLO path that would have QoS treatments at different points. These questions are still QoS centric so nothing more to comment].

[Moses ([MN)] As I mentioned above, the intent is more than QoS. It can also be Service Level Expectations such as path diversity/"disjointness", security, encryption, isolation, geographic restrictions etc. E.g.: A media production company can stream multiple copies of the same live streams to the production site for resiliency. The intent can be signaled as stream A and Stream B (and Stream C) should be transported via mutually disjointed paths (within each AS's and/or across multiple AS's).

[Swadesh (SA)]: Yes it makes sense and already has deployment experience with SR policies.

[Luay (LJ)]: Seamless could be accomplished in many ways, signaling is a big part of it. It's not a QoS pathway, I'd prefer to call it the SLO path that would have QoS treatments at different points. These questions are still QoS centric so nothing more to comment.

[Srihari] for inter-domain connectivity BGP is best suited (only scalable) so extending BGP to solve that problem is worth it. It would be nice to leverage all existing tools and operational mechanisms to provide this additional path establishment.

[Shraddha Hedge]: Yes, the intent must be sent in BGP routing update between ASes. The different ASes here are closely co-operating domains such as same ownership but different teams administering different part of the network such as Global network/domestic or backbone/metro network as described in sec 4.1

of <https://datatracker.ietf.org/doc/draft-hr-spring-intentaware-routing-using-color/>.

The purpose of color carried in service prefix is for steering the traffic into the transport. The color carried in transport is to identify the intent the particular transport satisfies.

[Mahesh Jethanandani]: I don't understand what is meant by "QoS pathway" either (unless it has been answered in one of the subsequent e-mails, which I have not read).

Question 2a: Should Customer intent that expresses pathway "QoS" be passed in BGP routing updates sent between Autonomous Systems?

[Kaliraj: Yes]

[Ketan:] I'm not sure what is meant by QoS and "QoS pathway" here. Can you clarify what specific QoS attributes are being referred to?

[Ketan]: The point of setting up a color-aware transport path between points A and B in the network is to enable a BGP Service like L3VPN to operate between those two points (e.g. [so] A and B are seen [operate] BGP next-hops for the SAFI 128. This [design] puts [color-aware] "multiple-AS" [paths] into the right perspective – as in option C. Otherwise, Option A/B is used for services between ASes and CAR/CT reachability ends at the AS boundary.

[Bruno]: Yes IMHO for:

- the service routes (e.g. BGP or "internet" to express the requirement, and
- the transport routes fulfilling the requirement.

Summary of CAR-CT Adoption Poll – Part 1 from 7/6 to 7/25 (Weeks 1-3)

Note that the above already requires agreement on the meaning of Color between ASes and between administrative domains. (and many more than the/[this?] number of ASes may be involved to set-up interop domain transport colors.

[Natrajan (NV)]: yes

[Swadesh (SA)] Yes. [It] makes sense and already has deployment experience with SR policies. I don't see anything specific for flow specification. We should allow actions to redirect traffic onto these color-aware routes.

[Gyan]: Yes. ... [And] this is possible today with SR-TE policy candidate path coloring encoded in BGP TEA not NLRI. I think this is an important aspect to be able to provide inter-as ... [for] end VPN service route coloring to underlay path being instantiated

Question 2b: Is it the purpose of color or transport class to allow automatic steering of traffic on into an “QoS” path (across different technologies)?

[Kaliraj]: Yes, See answers to question 1 [expression of SLAS, better automation, feedback loop].

- [Important] to cater to brownfield deployments with different technologies,
- [important] Preserve and extend ROI on time tested technologies and training.

[Ketan]: Color is an abstraction that is being passed in BGP updates. Whether or not it also indicates certain QoS parameters depends on the operator and their deployment design.

[Bruno]: Yes

Natrajan (NV): Yes. However, transport-classes provides a way to organize “QoS” path into separate buckets for easier service resolution,

[Gyan]: Yes. The goal of CT is to be provide a seamless mechanism of providing VPN overlay coloring to underlay loose underpinning to all steering technologies for both green and brownfield deployments.

[Mahesh Jethanandani]: QoS will allow for traffic to be steered towards a certain path, but that may be dependent on the current state of the network also.

Question 2c: How should this automatic steering interact with flow specification?

[Kaliraj]: Flow specification should be able to express SLAs just like other service families (IPv4 or LV3PN). For example, Color extended community carried along with Redirect-IP extended community.

[Ketan]: For this topic, we [IDR] should consider them [flow specification and color] orthogonal.

[Bruno]: On my side, [the] initial use case does not involve BGP Flowspec. Automatic steering is expressed by the service/VPN routes attaching a color to express the color/path it wants to follow.

[Natrajan (NV)]: BGP features need to interoperate with each other seamlessly and allow [for] a deployment to leverage the benefits of the features being used. So, this is not only limited to BGP flowspec but to other BGP features as well.

Summary of CAR-CT Adoption Poll – Part 1 from 7/6 to 7/25 (Weeks 1-3)

[Swadesh (SA)]: ... I don't see anything specific for flow specification. We should allow actions to redirect traffic onto these color-aware routes.

[Gyan]: FS Redirect-IP Extended community plays an important role as a traffic classification tool that can be leveraged to map overlay VPN service route micro or macro flows to underlay constraint based paths.

[Mahesh Jethanandani]: Agree that these two are not necessarily related, and should be considered separately.

Question 3: For those who believe that BGP should set-up a seamless path across multiple Autonomous Systems for a single Intent/QoS, do the exact mechanisms matter or do you simply want an interoperable solution? If they matter, please describe what matters.

Do not matter:

[Halpern:] Exact mechanism does not matter. I do want interoperability. Prefer one mechanism.]

[Bruno]: Primarily, we need interoperability. Then Automated steering should be simple and inline with what we already have with SR-policy and what we have within one AS. BGP encoding is less important, but for sure, it seems like different people have different opinions.

[Luay (LJ)]: I believe Bruno brought this up, we have different ASes under one administrative control on top of different ASes of different administrative controls. Leaving QoS aside again, we want to be able to build an SLO path across the same AS, different ASes under the same administrative controls, and lastly ASes under different administrative controls.

Do matter:

[Kaliraj:] How, [BGP] carries this SLA information matters. SLA/Intent being an adjective should not be put in NLRI [since it is problematic] for TE,

[Ketan]: BGP signaling (and other mechanism) matters for interoperability. While some like QoS config and other implementation aspects like policies may not [matter].

[Bruno] Not having to manage unnecessary indirections or ID spaces would help IMO.

[Moshiko] Important things are:

- Easy to [experiment in lab with prior to deployment],
- Easy set-up with no/un-noticeable interruption to service traffic,
- Easy to troubleshoot and visualize
- Use of known mechanisms [e.g. RD and extended community] helps in traffic manipulation using regular policies, filters and roadmaps.

[Natrajan]: The summary of the IDR chairs is that BGP-CT and BGP-CAR are functionally similar. What matters most to the customer are the following:

- Do customers wish to carry color as a key in NLRI?

Summary of CAR-CT Adoption Poll – Part 1 from 7/6 to 7/25 (Weeks 1-3)

- Do customers wish to carry non-key “forwarding information” as part of NLRI?
- How would customers wish for their services to “Express Intent” for individual flows.

[Moses] Mechanism matters.

- Simplicity at the same time wide coverage of use cases - The solution framework must support all the current and future use cases in a simple manner. (It shouldn't be addressing only the typical use cases of subset of customers / 'lead operators').
- Mechanism should preferably leverage the familiarity of existing standards and mechanisms deployed by the customers.
- Mechanism should focus on the transport layer intent use cases rather than keeping the scope wide open for other non-transport use cases (e.g. allowing service layer intents to be mixed with transport layer intents)
- Mechanism should not allow and open a pathway to add any unknowns in future that could impact the stability of BGP.

[Sue]: Thank you for letting us know how key stability of BGP for routing protocols is to you. Your feedback seems to ask us to re-examine our decisions on using BGP for information passing as well as routing functions. Is this correct?

[Moses] My comment was about maintaining the functional separation and restricting the usage to the intended purpose rather than keeping it open.

- The purpose of BGP-CT/CAR AF is to exchange the intent-aware transport routes across domains. The mechanism shouldn't open access to the usage for other purposes.
- I do believe that BGP shouldn't be used for exchanging any unstable information.

[Swadesh (SA)]: Currently operators deploy BGP LU for best effort inter domain paths. Exact mechanism matter such that existing routing model, functionality and operational experience provided by BGP LU is maintained and enhanced.

[Gyan]: Yes. the mechanisms matter especially for OPEX and MTTR and network availability. Mechanisms used should allow for scalability as well as ease of use for MTTR and network availability. Use of TEA and path attributes is always better and more scalable when extended to CE as compared to NLRI encoding which can be problematic at CE layer adding too much complexity.

[Jim Uttaro]: R.Raszuk “...Let the market decide not the mailing list then move one of them to Standards Track and the other one to Historic. “ I agree (with Robert). Currently there are two solutions although they are “functionally identical” there are important reasons why operators may opt for one or the other.

- I am a bit confused as to the driver “We need one interoperable solution”. Kompella and EVPN both can be used to create PWs, Multicast can be deployed using ingress replication, Rosen etc....
- the time to drive towards one interoperable solution is before the technologies are specified..
- I look forward to a discussion of the use cases addressed in order to better inform architect/designers as to which approach is most appropriate for them. Letting the market decide will provide needed perspective to resolve this in favor of one or both specifications.

[Luay]: I agree with Jim that as an operator I like to have options and let market needs,

- use cases drive them. I do also understand sometimes this causes more development in the vendor community to develop both plus interoperability, but that's their bread & butter ☺

Summary of CAR-CT Adoption Poll – Part 1 from 7/6 to 7/25 (Weeks 1-3)

- It seems that we keep putting ourselves in this impasse because of the "One" rule. In my mind, if the community sees 2 "good" solutions then instead putting them against each other, promote them. You never know what we will find in the future with more operational experience and new use cases that could make one solution fits better in certain use cases. I'm not a big fan of the "One" basket as you can see 😊

[Gyan]: *I agree as well with Jim & Luay being* against the one size fits all "single" basket approach. Both solutions are functionally similar with different approaches.

- As Luay stated, having a two vendor solution does lead to more development work on the vendor side to develop both solutions and ensure interoperability, but that is indeed their business to develop new innovative solutions and with that, possible ROI in development of both solutions.
- This is not the first time we have been in this situation, where we are pitching both vendors against each other trying to get to a one solution model to ensure interoperability.
- For operators, if we have two good solid solutions, why not promote both even if they are functionally similar but operationally different.
- It is very difficult to make a decision now without letting the solutions bake off in the industry and see which one prevails or it might be they both have their own niche use cases and they both may prevails which I can see that happening in the long run.
- Flexibility is important which being the consumers of the solution and ultimate beneficiary, it makes sense to develop both solutions and let the industry decide.

[Srihari] it would really benefit the community if the solution to solve this problem is easy to manage, incrementally deploy and leverage existing toolsets

[Joel Halpern]: Really, you want two non-interoperable solutions?

- So what happens if you are peering with operator 1 who has chosen solution 1, and also peering with operator 2 who has chosen solution 2. Even if you support both solutions in your network, you can not offer paths that traverse yourself and both peers. And, if you want to deploy both solutions then not only do the vendors have to support both. you have to test both.
- I suppose you could choose only one of the two solutions. And simply be unable to have intent paths that cross into other operators who made a different choice? This seems like exactly what standards are supposed to avoid / prevent. It may be nice for python to have 6 ways to do everything. It is not good for the Internet.

[Robert]: What exactly do you mean in the light of the subject topic? Keep in mind that both solutions work **only** under the same (or closely cooperating) administration. So actually having both solutions documented as experimental IMO is not that bad.

- I have just asked in the other thread what is expected signaling outside of same or cooperating administration. I assume none.
- So naturally you can still provide color aware transit but you just do the mapping/translation to your own flavours. In fact that is the case for both CAR and CT between external parties.

[Gyan]: In my mind the SDOs job is to ensure interoperability.

- If the one solution is adopted then that one solution must be implemented by all vendors that choose to support the solution.
- If two solutions are adopted and both solutions are implemented by all providers that choose to support the solution, there should not be any interoperability issues either.
- If it's within the same administrative domain coordination is possible to use the matching solution 1 or Solution 2.
- If it's between administrative domains between operator 1 and 2 there is a chance or mismatch and that would require a coordinated effort to deploy the same solution.

Summary of CAR-CT Adoption Poll – Part 1 from 7/6 to 7/25 (Weeks 1-3)

- I would not call that an interoperability issue as Solution 1 and Solutions 2 are both progressed and implemented by all vendors and it's a matter of having the operators that require interoperability to pick a common solution 1 or 2 that would be applicable to their use case.

I don't think here we want or need to go the experimental track here, but rather progress both drafts and let the market decide.

[Kaliraj]: I agree with Joel.

- The WG can save lot of effort, money and confusion for the industry (operators, vendors) by choosing one solution that is complete. Of course, the industry is free to decide and implement whichever solution it wants, but it expects reasonable guidance from the WG that the WG has analyzed the correctness and completeness aspect of the solutions being adopted by the WG.
- In absence of such contract, only IANA is needed to assign non-conflicting code-points, and customers can decide the rest. Sorry to say that.
- Having said that, I want the WG to pay attention to the problems being discussed in Part3 of this Adoption call. For e.g. we have indicated possibilities of mis-routing with CAR: (<https://mailarchive.ietf.org/arch/msg/idr/OOZOBSyjdAYBar8NxvOqo6-5fAc/>)
- A solution with correctness issues should not be adopted by IDR WG. It will result in unsuspecting customers burning their fingers in the field, and question why the WG allowed progressing the draft if these problems were already known.

[Mahesh Jethanandani]: Prefer an interoperable solution (with use of BGP signaling). As mentioned, BGP signaling matters.

Question 4: RFC7606 focused on error handling in which the MP-NLRI focuses on destination keys (RD and Prefix) plus non-key material (Labels, SIDS). Attributes (generally) apply to all NLRI. For example, MED applies to all NLRIs in the packet.

Question 4a: Would error handling be better for color-aware routing if attributes relevant to a specific color/class be grouped in a MP-Color-Attribute?

Yes

[Kaliraj]: Yes. Not carrying “non-key fields” in NLRI helps in better error handling, ... [and] in keeping BGP modular [one AF, one modular function]. [Without] this focus on modularity [new AFIs tend to include previous AFIs] in to “God’-AFI” ... with the subsumption of other functionality into “God’-AFI.

[Bruno]: Probably, error handling would be better with a simple NLRI. However, this is a complex subject, and it really depends on the type of error and whether the length field match[es] the size of the value field (or not). However, there are other tradeoff[s] involve[d] (e.g. efficiency, extensibility/generality). Note that we have already encoded non-key data in the NLRI (e.g. label, label stack) so it may not be black and white.

[Natrajan (NV)]: Yes [to a and b] – BGP- BGP-CT team supports this approach as it confines the ... error handling to the attribute. Such an attribute is already being proposed by the BGP-CT team. I am bringing the same to the notice of both customers and the IDR chairs alike

<https://datatracker.ietf.org/doc/draft-kaliraj-idr-multinexthop-attribute/>

[Gyan]: Yes, that makes sense for BGP update packing efficiencies RFC 4277.

[Srihari] Yes. Grouping forwarding related information into one bucket seems like a Logical thing to do. Today we have many places to embed this information (Nkrumah, attributes). TEA was one attempt, but simpler attribute can go a long way to simplify.

No:

[Ketan]: No

[Swadesh (SA)]: SA] Not sure if I see advantage of such attribute from error handling point of view. Each color aware route has associated attributes and hence color scoped. It provide error handling at granularity of each route.

Question 4b: Should IDR consider future work on a MP-Color Attribute?

Yes: (4b)

[Kaliraj]: Yes. [see] draft-kaliraj-idr-multinexthop-attribute which can serve that purpose.

[Gyan]: As “coloring” VPN overlay to underlay colored path instantiation is critical for operators as they migrate to SR, having an MP-Color path attribute would be very helpful and could provide an SR-TE automated steering optimization and could help reduce SR Policy complexity as well as help with MTTR and availability.

No: (4b)

[Ketan]: I do not believe there is such a requirement for Color-Attribute. As a reminder, “color” is an abstraction.

No comment: (4b)

[Moses] No specific comments as I am not familiar with BGP error handling / RFC7606.

Bruno Discussion leading to a generic solution:

[Bruno]: Not sure what you have in mind. I don’t think that encoding the color is the issue as this is a short, constant size value. To me the difficult question is more about encoding some NLRI specific data (e.g. label for MPLS, but these days we may need to be more general) in order to avoid breaking packing (in which case, the question is not completely specific to BGP colors).

[Sue]: The question was on encoding some NLRI specific data (the second half of your comments). To [further explain this] consider:

- MP-REACH/MP-UNREACH - group NLRIs by AFI/SAFI
- An MP-Color attribute would group attributes by color and an An MP-NH attribute could group by NHs.

Example of an MP-Color attribute:

Summary of CAR-CT Adoption Poll – Part 1 from 7/6 to 7/25 (Weeks 1-3)

- [MP-Color-Green: {MED, Tunnels, non-key data (Next-Hop label-pairs, etc)}]
- [MP-color-Gold: (MED, Tunnels, non-key data (NH, Labels, backup, preferences)]

Example of an MP-NH attribute would group this by Next-Hops

[Next-hop-1: Color: Gold, tunnels, non-key data (labels, other things), preferences]

[Next-hop-2: Color Green, etc.]

Please do not take my examples as anything other than vague directions IDR might go into.

[Bruno2]: Thanks for the clarification. The above proposal could help; however IMHO one difficulty is the selection of this additional "group by" key. You provided two (Next-Hop, Color), but there could be multiple. IMHO, the one which is likely to be the most useful (based on history and current discussion) is to be able to advertise non-key data per NLRI.

[Bruno2] Personally, in order to be generic (and hopefully benefit from a single code base to debug) I would propose a new MP_REACH_NLRI attribute carrying sets of (NLRI, non-key data), compared to (NLRI) currently. Or may be even better for error handling two lists: a list of NLRI, followed by a list of non-key data (in the same order). But this may be seen as a larger effort compared to BGP-CAR which only defines a new SAFI.

[Jeff in Response to Bruno2]: This would nicely address the error handling point I'd raised on -car that Dhananjaya has responded to. The idea of a generic key that can safely make it through a route reflector without need to upgrade the reflector has strong appeal. We're seeing that it's challenging to do.

Also see Jeff Haas discussion on CAR encoding issue with Dhananjaya (DJ) Roa.

<https://mailarchive.ietf.org/arch/msg/idr/qorZr8Y7Cpn4cv-BkH3LhBL391I/>

(see text in full below with DJ's comments include]

[Mahesh Jethanandani] If we agree that color is an abstraction, then there is no requirement for a Color-Attribute. BTW, Arrcus has an implementation for CAR.

Jeff Haas Comment on CAR Error handling: (thread at)

Jeff's first post: <https://mailarchive.ietf.org/arch/msg/idr/qorZr8Y7Cpn4cv-BkH3LhBL391I/>

DJ's response: <https://mailarchive.ietf.org/arch/msg/idr/MddptzPITu1FxAgALAY-9liADgY/>

Jeff's second post: <https://mailarchive.ietf.org/arch/msg/idr/qorZr8Y7Cpn4cv-BkH3LhBL391I/>

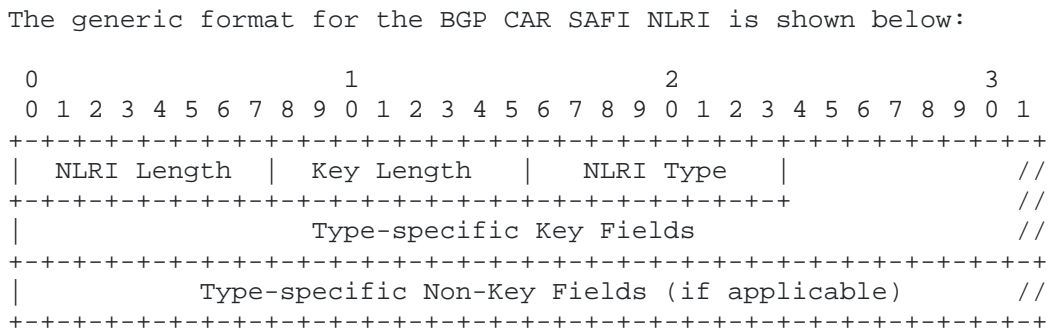
BGP-CAR authors,

Prompted somewhat by Sue's comment in the Informational Questions:

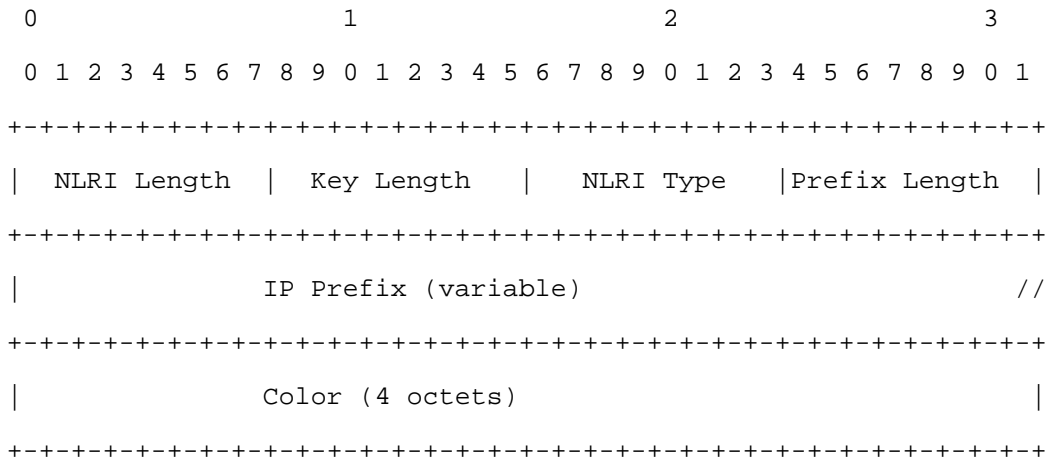
4. RFC7606 focused on error handling in which the MP-NLRI focuses on destination keys (RD and Prefix) plus non-key material (Labels, SIDS). Attributes (generally) apply to all NLRI. For example, MED applies to all NLRIs in the packet.

I was looking through the encodings and the related error handling text in the -05 version of the -car document. In section 2.9.1, we have the generic header:

The generic format for the BGP CAR SAFI NLRI is shown below:



The type 1 format in section 2.9.2 builds on that:



Summary of CAR-CT Adoption Poll – Part 1 from 7/6 to 7/25 (Weeks 1-3)

Followed by optional TLVs encoded as below:

```
+-----+
|   Type   | Length | Value (variable) //
+-----+
```

It creates the definition:

- : o Key Length: variable. It indicates the total length comprised of
- : the Prefix Length field, IP Prefix field, and the Color field, as
- : described below. For IPv4 (AFI=1), the minimum length is 5 and
- : maximum length is 9. For IPv6 (AFI=2), the minimum length is 5
- : and maximum length is 21.

Section 2.10 covers the error handling procedures, which I quote:

- : When the error determined allows for the router to skip the malformed
- : NLRI(s) and continue processing of the rest of the update message,
- : then it MUST handle such malformed NLRIs as 'Treat-as-withdraw'. In
- : other cases, where the error in the NLRI encoding results in the
- : inability to process the BGP update message, then the router SHOULD
- : handle such malformed NLRIs as 'AFI/SAFI disable' when other AFI/SAFI
- : besides BGP-CAR are being advertised over the same session.
- : Alternately, the router MUST perform 'session reset' when the session
- : is only being used for BGP-CAR.

The general idea makes sense here: Use the key field and as long as you trust its contents, that's enough to do RFC 7606 procedures.

- : Following errors result in 'AFI/SAFI disable' or 'session reset':
- : o Minimum NLRI length check error.

Summary of CAR-CT Adoption Poll – Part 1 from 7/6 to 7/25 (Weeks 1-3)

- : o NLRI length conflict with key length.
- : o Key length encoding errors (such as minimum, maximum and conflict
- : with prefix length).

For Type 1, these also make sense.

- : There can be cases where the NLRI length value is in conflict with
- : the enclosed non-key TLVs, which themselves carry length values.
- : Either the length of a TLV would cause the NLRI length to be exceeded
- : when parsing the TLV, or fewer than 2 bytes remain when beginning to
- : parse the TLV.

For non-key, overclaim and too-short can cause the TLV set to be malformed. However, this is in the optional stuff, so aside from avoiding issues that TLV protocols introduce of parsing past the content, or interpreting content outside of the containing TLV for a truncated sub-TLV, it's still clear.

I think some additional text covering length checks may still be appropriate. There's particular motivation between the normative RFC 2119 keywords for whether the MUST in question for a specific description of procedure implies malformed packets or not. It might be worth tagging each element of procedure where violations may imply the reset or not.

[DJ (CAR author)]: We will take a second look.

[Jeff]: Thanks!

(Note that I'll flag items as needing session reset for brevity. The text already discusses afi/safi shutdown.)

Here's some notes working through the encoding:

Generically:

- The minimum NLRI Length is 2 for unknown types. A key length and an NLRI Type are required fields. If shorter, the NLRI are malformed and we need a session reset.

[DJ]: We could explicitly state the minimum

[DJ]: For the General error handling comment, it's not entirely clear if it needs to be repeated explicitly everywhere, or is [it] sufficient to capture in the Error handling section. But we will discuss.

[Jeff2:] I understand it may need a bit of trial and error in the document to find a format that works best. The main challenge I was noting is that our RFC2119 keywords vary in this circumstances between “fatal” and “non-fatal” behaviors with regard to the session. I

[Jeff2:] It is the usual problem of overloaded exception handling. 😊

Why state this when the next requirement implies it? Clarity and a desire to help people avoid underflow of unsigned integers. :-)

- Key Length MUST be at least two less than NLRI Length. (Already stated.) If not, the NLRI are malformed and we need a session reset.
- The NLRI Type may contain unknown fields. The intent is to permit a route reflector to transparently carry unknown types. As will be noted further down, once we hit a device that understands the contents, the downstream BGP Speaker may hit a failure condition that requires session reset. This will be illustrated using the documented Type 1's as an example. I think this may violate the spirit of RFC 7606 where our old Path Attribute form of "optional, transitive nonsense" gets a new life in this kind of NLRI.

[DJ]: Please see response at the end.

For Type 1:

- Key Length gets additional semantic checks. This is our first example that validation may change depending on whether the type is known vs. unknown. Failure of those minimums require session resets.

This likely motivates the error handling text being specific about general behavior (small set of things) vs. type-specific. The text is mostly structured this way, but clarifying that we've moved into type-specific validation would be good for readability.

[DJ]: We will check.

- Prefix Length specifies the usual limits for IPv4/IPv6. It'd be good to make sure the definition section notes that violating length is worth a session reset.
- IP Prefix: Consider borrowing the text from the main BGP RFC. This would also cover what to do about "trailing bits".

I suspect the usual caveats in RFC 4271, §6.3 apply about semantically incorrect addresses being ignored?

[DJ]: We believe so. We will check if we can borrow the text.

[Jeff-2]: Thanks. I think the somewhat novel consideration, going with the novel encoding, is that with the intent of targeting generic reflectors is that situations like the "trailing bits" accommodation may be worth abandoning.

[Jeff-2] The path that got us to our behaviors in RFC 4271 were based on a certain level of forgiveness (Postel's Maxim) with regard to trailing bits based on long history. But even with long history, when implementations do have accidental trailing bits, we still see bugs in the field with regard to processing them.

[Jeff-2]: In formats like BGP-LS, BGP Flowspec, and structured NLRI as seen in many of the BESS protocols, we have additional issues that "don't care" bits become problematic because it's not clear when an implementation should or should not ignore

them for key considerations. Most implementations fall back to simple memcmp() semantics without a canonicalization step. Thus, two theoretically comparable NLRI may not compare identically depending on don't-care semantics.

[Jeff-2]: As a new NLRI, you do have the option to make a choice in the matter. The challenge is when that choice conflicts with established practice (existing library code!), even if for good reasons.

Not specific to the key fields:

- The optional TLVs are largely targeted toward forwarding behaviors. It's not clear what should be done if more than one TLV is carried at the same time.

[DJ]: This is a matter of local policy to decide which one or more forwarding types we can make it explicit.

The above comments on the key field validation is, I think, the essence of whether the NLRI keys can be safely used or not. Making sure that the violations and their impact are close together in the text would be helpful.

Where we see a more general issue is the transitional stage of going from an unknown type to a known type could result in downstream session resets.

The goal was to provide transparent carriage of unknown types, but unless I'm missing something, we didn't quite succeed at this when errors are present and raise issues similar to those we were trying to mitigate in RFC 7606.

[DJ]: Since the intent itself is to allow new route types to transit through an RR transparently without requiring an upgrade, the RR cannot detect an error in that NLRI. It is expected though that the client of the RR (for example, an ABR) which is supposed to consume this route for installation will be upgraded, and hence can do the error handling. The route will therefore not propagate any further. This is also true if the ABR is not upgraded and does not recognize the new route type. It will not install and hence not propagate any further.

[Jeff-2] Agreed. However, that means that like the situations that motivated RFC 7606, the session reset happens "far away" from the route origination.

[DJ]: One possible option is to enable the operator to make the decision on whether they want to allow this propagation through an RR; or enforce the validation at the RR, in which case the RR will need to be upgraded. We could describe this consideration and option in the Deployment section. Does that look useful to you?

[Jeff-2] Possibly so. Minimally, the Error Handling section could use the text. One additional way to consider addressing this is to capability negotiate whether a given NLRI type for the afi/safi is understood. Implementations may thus have the operational ability to choose whether they accept unknown NLRI types for that afi/safi or not.

[Jeff-2]: You may find my somewhat recent flowspec capability bits draft a possible example for encoding such a thing.

Section 3: Customer Use cases:

Re: [\[Idr\] Part 1 of CAR/CT Adoption call \(7/6 to 7/20\) - Informational Questions \(ietf.org\)](#)

https://mailarchive.ietf.org/arch/msg/idr/cxPmHvklLddwJn0iw_RKTMpyxQA/

[Mosikho Nayman (Juniper)]

My customer will reply tomorrow.

The use case is when their customer ACME buys an infrastructure and would like to have control on paths going through this infrastructure (metro-e). Ability to preserve different coloring in the same purchased service across the 3rd party network.

That allows ACME to have additional control on how their data traverse a 3rd party infrastructure.

For example, ACME buy infrastructure from metro-e company and signal two colors as per below

Voice group A --- RED 0:100

Voice group B --- BLUE 0:200

RED regular latency

BLUE Low latency

With BGP-CT, ACME can swap between the signaled color and prefer Group A over B.

This works when metro-e split across multi-as OR when metro-e decide to acquire another metro-e company with different AS.

Yes, they do have multiple AS numbers.

And of course, any other company acquisition may add more AS numbers.

"Nagarajah, Moses" <Moses.Nagarajah@team.telstra.com>

Some of the use cases:

Transport Network Slicing to support 5G Slicing (URLLC, eMMB, mMTC, etc.)

Dual-plane/ Multiplane forwarding for isolation between primary and secondary traffic flows

Dedicated network slice for high value wholesale customers

Carriage of Edge Compute services with stringent SLAs