

# **Using BGP to Bind MPLS Labels to Address Prefixes**

## **draft-rosen-idr-rfc3107bis-00**

Eric Rosen  
(presented by Ross Callon)

# RFC 3107

- Specifies how to use BGP to bind MPLS labels to IPv4/IPv6/VPN-IPv4/VPN-IPv6 prefixes  
(Doesn't mention VPN-IP prefixes, but applies to them as well)
- Technique:
  - Both label and prefix encoded in NLRI, creating the *labeled address families*:
    - SAFI 4: prefix is IP, AFI determines v4 or v6
    - SAFI 128: prefix is VPN-IP, AFI determines v4 or v6
  - Label is “owned” by the next hop
- Allows NLRI to encode sequence of labels (*Multiple Labels*), representing contiguous portion of label stack

# Why Do We Need an RFC3107bis?

- Errors and underspecification
  - Many issues where details are missing, subject to interpretation
  - Different interpretations don't always interoperate well
  - (also, many errata)
- Examples
  - Unclear about semantics of multiple labels (stack) in NLRI.
  - Silent/unclear about when two routes are comparable
  - Rules for withdrawing bindings
  - Issue of multiple paths with same next hop and prefix, but different labels

# What Can and Can't We Do in RFC3107bis?

- RFC 3107 has some multi-vendor interop issues
  - No real way to fix these now, no point arguing about whose interpretation is most valid
  - Suggested approach: document the interop issues, do not favor one implementation over another, try not to make existing implementations non-compliant
- What we can do:
  - Make things easier for future implementers,
  - Get “multiple labels” feature working for the first time
  - Make some sense out of the “multiple paths with different labels” issue, integrating the use of add-paths

# Binding Multiple Labels to a Prefix

- Original intention was just to use BGP to bind a single label to a prefix (like LDP).
- RFC allows multiple labels (a stack) to be bound to a prefix
  - Aware of only one implementation, quite recent
  - Many implementations **assume** there's only a single label, and hence won't interoperate correctly if there are multiple
  - RFC doesn't say what to do when you set *next hop self* and then propagate a route that was received with multiple labels

# Binding Multiple Labels to a Prefix: Semantics

- When propagating route after setting *next hop self*, replace original set of labels (*Set1*) with set of one or more labels (*Set2*)
- Possible use cases discussed in the draft
- Note: no change to MPLS data plane semantics

# Binding Multiple Labels to a Prefix: Syntax

- Original encoding for determining the number of labels in the NLRI is “non-optimal”
- Most implementations ignore it anyway, assume one label
- Therefore 3107bis specifies that the use of multiple labels be controlled by a BGP Capability
  - Preserves compatibility with existing “single label” implementations
  - Capability should also specify maximum number of labels supported for each address family
  - Opportunity to create a more optimal encoding
    - Maybe NLRI length doesn’t have to be expressed in bits!

# Coexistence of Labeled and Unlabeled Route to Same Prefix

- What does it mean if you have an unlabeled route to prefix P as well as a labeled route to prefix P?
  - Does one invalidate/replace the other?
    - If so, do you get reasonable and predictable behavior?
    - If not, which do you use when? Multipath? For what traffic?
- Different vendors have taken different approaches
- 3107bis does not attempt to fix this or make judgments:
  - Suggests coexistence of labeled/unlabeled routes with same prefix be used with caution
  - Behavior is matter of local policy, unpredictable multi-vendor interop.



# How To Withdraw a Labeled Route

- RFC3107 says to withdraw a labeled route, you can either specify the label+prefix, or you can specify the prefix, with 0x800000 in the “labels” field
  - This had been 0x000001, things got changed between the last internet-draft and the RFC!
  - 3107bis suggests:
    - Put 0x800000 in the field when sending a withdraw
    - Ignore field when receiving a withdraw
    - Same withdraw works, whether one label had been assigned, or many
- Whether an unlabeled route for a given prefix withdraws a label binding is a matter of local policy

# Multiple Routes with Same Prefix, Same NH, Different Labels?

- Might receive, on different sessions, two routes with same prefix and next hop
- Might want to propagate as two routes with same prefix and next hop self, but different label
- Proposal:
  - Do not allow propagation of both except via *add-paths*
  - Even though explicit withdraw does not specify label (per previous slide), can withdraw one of these routes by using *add-paths* path identifier plus prefix

# Summary

- RFC 3107
  - Multiple implementations, widely deployed
  - An update is clearly needed (underspecified details, errata)
- draft-rosen-idr-rfc3107bis-00 is a very good start on an update
  - We are not finished, more discussion needed
  - But document is ready for WG adoption