IS-IS for IP Internets
Internet-Draft
Intended status: Standards Track
Expires: June 25, 2015

P. Sarkar, Ed.
H. Gredler
S. Hegde
Juniper Networks, Inc.
S. Litkowski
B. Decraene
Orange
Z. Li
Huawei Technologies
E. Aries
R. Rodriguez
Facebook
H. Raghuveer

December 22, 2014

Advertising Per-node Admin Tags in IS-IS
draft-ietf-isis-node-admin-tag-00

Abstract

This document describes an extension to IS-IS protocol [ISO10589], [RFC1195] to add an optional operational capability, that allows tagging and grouping ofthe nodes in an IS-IS domain. This allows simple management and easy control over route and path selection, based on local configured policies.

This document describes the protocol extensions to disseminate per-node administrative tags in IS-IS protocols.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Copyright Notice

Table of Contents

1. Introduction

This document provides mechanisms to advertise per-node administrative tags in the IS-IS Link State PDU [RFC1195]. In certain path-selection applications like for example in traffic-engineering or LFA [RFC5286] selection there is a need to tag the nodes based on their roles in the network and have policies to prefer or prune a certain group of nodes.

2. Administrative Tag

For the purpose of advertising per-node administrative tags within IS-IS, a new sub-TLV to the IS-IS Router Capability TLV-242 that is defined in [RFC4971] is proposed. Because path selection is a functional set which applies both to TE and non-TE applications the same has not been added as a new sub-TLV in the Traffic Engineering TLVs [RFC5305].

An administrative Tag is a 32-bit integer value that can be used to identify a group of nodes in the IS-IS domain. The new sub-TLV specifies one or more administrative tag values. An IS-IS router advertises the set of groups it is part of in the specific IS-IS level. As an example, all PE-nodes may be configured with certain tag value, whereas all P-nodes are configured with a different tag value in.

The new sub-TLV defined will be carried inside the IS-IS Router Capability TLV-242 (defined in [RFC4971]) in the Link State PDUs originated by the router. Link State PDUs [ISO10589] that has either level-wise (i.e. L1 or L2) or domain-wide flooding scope. Choosing the flooding scope to flood the group tags are defined by the needs of the operator's usage and is a matter of local policy or configuration.

Operator may choose to advertise a set of per-node administrative tags across levels and another set of per-node administrative tags within the specific level. But evidently the same set of per-node administrative tags cannot be advertised both across levels and within a specific level. A receiving IS-IS router will not be able to distinguish between the significance of a per-node administrative tag advertised globally from that of a administrative tag advertised locally if they have the same value associated but different significance across different scopes.

Implementations SHOULD allow configuring one or more 'global' as well as 'level-wide' administrative tags. A operator may only need to advertise and flood a specific per-node administrative tag, either across all levels, or only within a specific level. Hence implementations MUST NOT allow configuring the same per-node administrative tag values in both 'global' and 'level-wide' scopes. However the same administrative tag value MAY be allowed to be configured and advertised for multiple levels with 'level-wide' flooding scope.

The 'global' per-node administrative tags shall have significance across the entire administrative domain and hence MUST be advertised in a Router-Capability TLV with 'global' scope (i.e. S-bit set to

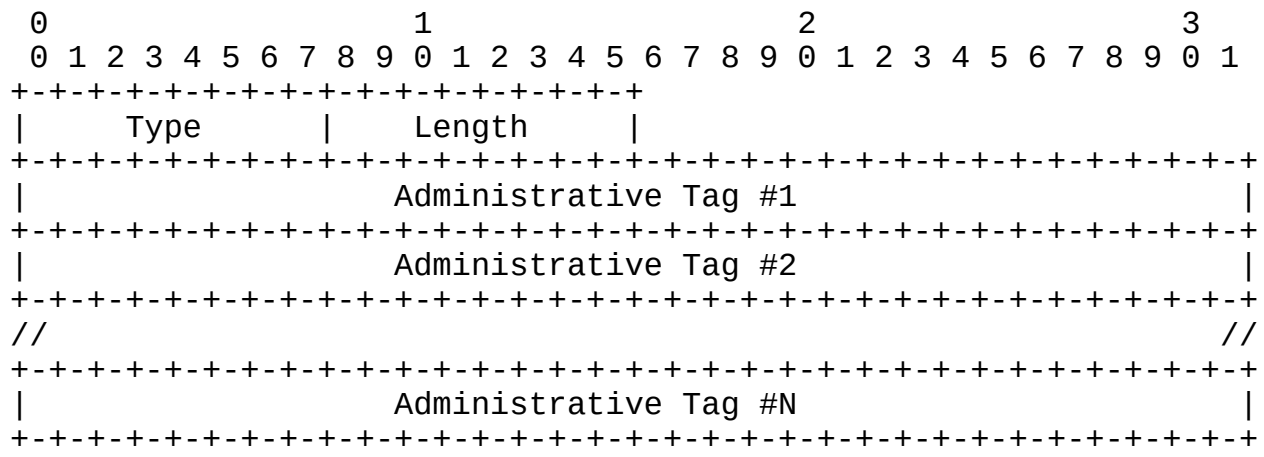
1), and inserted in the LSP PDUs generated for all levels applicable. The 'level-wide' administrative tags should be copied in to a Router-Capability with 'level-wide' scope only (i.e S-bit reset to 0) and copied into the LSP PDU for the specific level.

In deployments using multi-topology routing [RFC5120], since multiple topologies within same IS-IS level share the same flooding scope configuring the same per-node administrative tag across different topologies, SHOULD NOT be allowed. Advertising the same tag value across multiple topologies will lead to same inconsistencies as with the case of advertising same tag value across 'global' and 'level-wide' flooding scope. If there is need to distinguish between the per-node administrative tags used for one topology to another, operators are advised to use disjoint sets of per-node administrative tags across such topologies.

3. TLV format

3.1. Per-node Admin Tag sub-TLV

The new Per-node Administrative Tag sub-TLV, like other ISIS Capability sub-TLVs, is formatted as Type/Length/Value (TLV)triplets. Figure 1 below shows the format of the new sub-TLV.



Type : TBA

Length: A 8-bit field that indicates the length of the value portion in octets and will be a multiple of 4 octets dependent on the number of tags advertised.

Value: A sequence of multiple 4 octets defining the administrative tags.

Figure 1: IS-IS Per-node Administrative Tag sub-TLV

The ’Per-node Admin Tag’ sub-TLV may be generated more than once by an originating router. This MAY happen if a node carries more than 63 per-node administrative groups and a single sub-TLV does not provide sufficient space. As such occurence of the ’Per-node Admin Tag’ sub-TLV does not cancel previous announcements, but rather is cumulative.

4. Elements of Procedure

Meaning of the Per-node administrative tags is generally opaque to IS-IS. Router advertising the per-node administrative tag (or tags) may be configured to do so without knowing (or even explicitly supporting) functionality implied by the tag.

Interpretation of tag values is specific to the administrative domain of a particular network operator. The meaning of a per-node administrative tag is defined by the network local policy and is controlled via the configuration. If a receiving node does not understand the tag value, it ignores the specific tag and floods the Router Capability TLV without any change as defined in [RFC4971].

The semantics of the tag order has no meaning. There is no implied meaning to the ordering of the tags that indicates a certain operation or set of operations that need to be performed based on the ordering.

Each tag SHOULD be treated as an independent identifier that MAY be used in policy to perform a policy action. Tags carried by the administrative tag TLV SHOULD be used to indicate independent characteristics of a node. The TLV SHOULD be considered as an unordered list. Whilst policies may be implemented based on the presence of multiple tags (e.g., if tag A AND tag B are present), they MUST NOT be reliant upon the order of the tags (i.e., all policies should be considered commutative operations, such that tag A preceding or following tag B does not change their outcome).

As mentioned earlier, to avoid incomplete or inconsistent interpretations of the per-node administrative tags the same tag value MUST NOT be advertised by a router in Router Capabilities of different scopes. Implementations MUST NOT allow configuring the same tag value across domain-wide and 'level-wide' scopes. The same tag value MAY be allowed to be configured and advertised under 'level-wide' scope for all levels. A IS-IS Area Border Routers (ABR) participating in both levels 1 and 2 MAY advertise the same tag value in the level-specific Router Capability TLVs with 'level-wide' scope (S-bit reset to 0) generated by it. But the same tag value MUST not be advertised in any of level 1 or level 2 Router-Capability TLV with 'global' scope (S-bit set to 1).

The per-node administrative tags are not meant to be extended by the future IS-IS standards. The new IS-IS extensions MUST NOT require use of per-node administrative tags or define well-known tag values. Per-node administrative tags are for generic use and do not require IANA registry. The future IS-IS extensions requiring well known values MAY use new Capability sub-TLVs tailored to the needs of the feature, as defined in [RFC4971].

Being part of the Router Capability TLV, the per-node administrative tag sub-TLV MUST be reasonably small and stable. In particular, but not limited to, implementations supporting the per-node administrative tags MUST NOT tie advertised tags to changes in the network topology (both within and outside the IS-IS domain) or reachability of routes.

5. Applications

This section lists several examples of how implementations might use the Per-node administrative tags. These examples are given only to demonstrate generic usefulness of the router tagging mechanism.

Implementation supporting this specification is not required to implement any of the use cases. It is also worth noting that in some described use cases routers configured to advertise tags help other routers in their calculations but do not themselves implement the same functionality.

1. Auto-discovery of Services

   Router tagging may be used to automatically discover group of routers sharing a particular service.

   For example, service provider might desire to establish full mesh of MPLS TE tunnels between all PE routers in the area of MPLS VPN network. Marking all PE routers with a tag and configuring devices with a policy to create MPLS TE tunnels to all other devices advertising this tag will automate maintenance of the full mesh. When new PE router is added to the area, all other PE devices will open TE tunnels to it without the need of reconfiguring them.

2. Policy-based Fast-Reroute

   Increased deployment of Loop Free Alternates (LFA) as defined in [RFC5286] poses operation and management challenges. [I-D.ietf-rtgwg-lfa-manageability] proposes policies which, when implemented, will ease LFA operation concerns.

   One of the proposed refinements is to be able to group the nodes in IGP domain with administrative tags and engineer the LFA based on configured policies.

   (a) Administrative limitation of LFA scope

      Service provider access infrastructure is frequently designed in layered approach with each layer of devices serving different purposes and thus having different hardware capabilities and configured software features. When LFA repair paths are being computed, it may be desirable to exclude devices from being considered as LFA candidates based on their layer.

      For example, if the access infrastructure is divided into the Access, Distribution and Core layers it may be desirable for a Distribution device to compute LFA only via Distribution or Core devices but not via Access devices. This may be due to features enabled on Access routers; due to capacity limitations or due to the security requirements. Managing

such a policy via configuration of the router computing LFA is cumbersome and error prone.

With the Per-node administrative tags it is possible to assign a tag to each layer and implement LFA policy of computing LFA repair paths only via neighbors which advertise the Core or Distribution tag. This requires minimal per-node configuration and network automatically adapts when new links or routers are added.

(b) Optimizing LFA calculations

Calculation of LFA paths may require significant resources of the router. One execution of Dijkstra algorithm is required for each neighbor eligible to become next hop of repair paths. Thus a router with a few hundreds of neighbors may need to execute the algorithm hundreds of times before the best (or even valid) repair path is found. Manually excluding from the calculation neighbors which are known to provide no valid LFA (such as single-connected routers) may significantly reduce number of Dijkstra algorithm runs.

LFA calculation policy may be configured so that routers advertising certain tag value are excluded from LFA calculation even if they are otherwise suitable.

3. Controlling Remote LFA tunnel termination

[I-D.ietf-rtgwg-remote-lfa] proposed method of tunneling traffic after connected link failure to extend the basic LFA coverage and algorithm to find tunnel tail-end routers fitting LFA requirement. In most cases proposed algorithm finds more than one candidate tail-end router. In real life network it may be desirable to exclude some nodes from the list of candidates based on the local policy. This may be either due to known limitations of the per-node (the router does accept targeted LDP sessions required to implement Remote LFA tunneling) or due to administrative requirements (for example, it may be desirable to choose tail-end router among co-located devices).

The Per-node administrative tag delivers simple and scalable solution. Remote LFA can be configured with a policy to accept during the tail-end router calculation as candidates only routers advertising certain tag. Tagging routers allows to both exclude nodes not capable of serving as Remote LFA tunnel tail-ends and to define a region from which tail-end router must be selected.

4. Mobile backhaul network service deployment

The topology of mobile backhaul network usually adopts ring topology to save fiber resource and it is divided into the aggregate network and the access network. Cell Site Gateways(CSGs) connects the eNodeBs and RNC(Radio Network Controller) Site Gateways(RSGs)connects the RNCs. The mobile traffic is transported from CSGs to RSGs. The network takes a typical aggregate traffic model that more than one access rings will attach to one pair of aggregate site gateways(ASGs) and more than one aggregate rings will attach to one pair of RSGs.

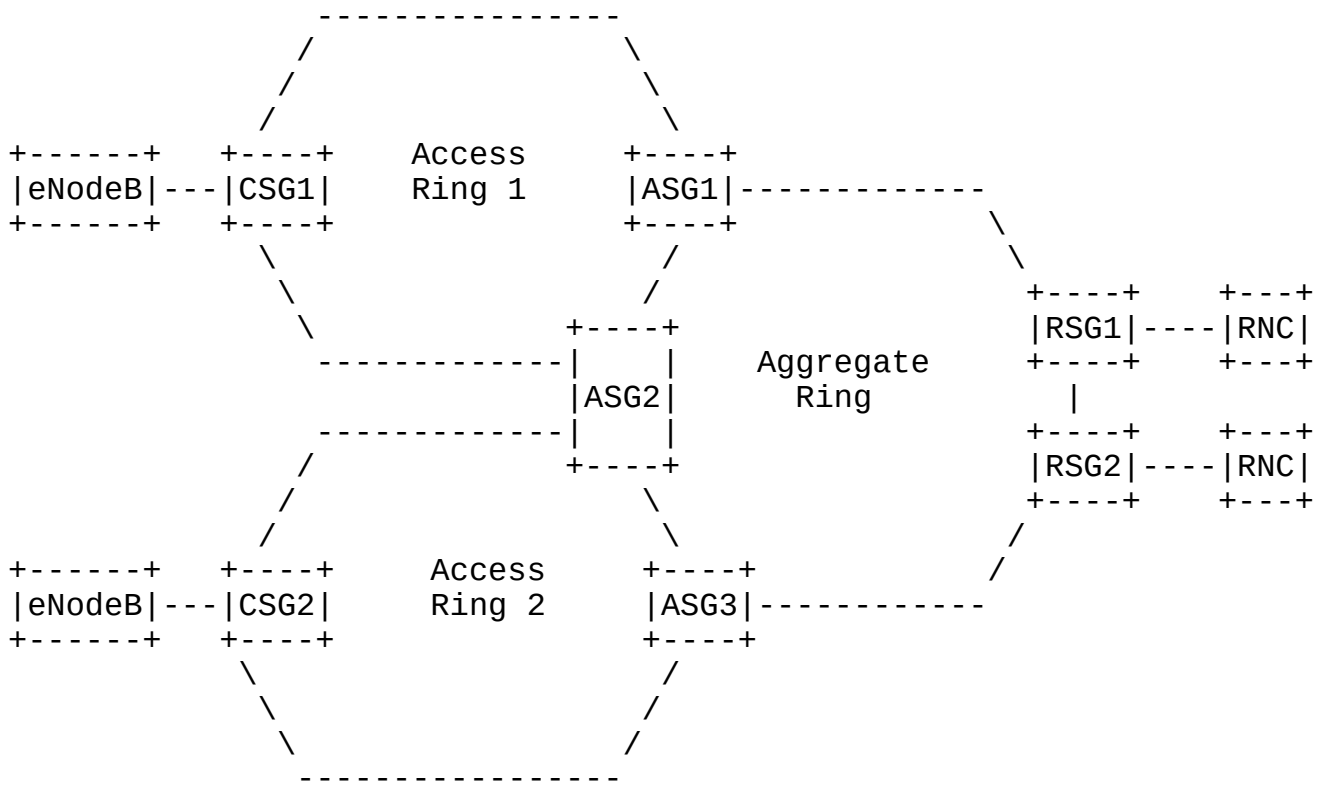


Figure 2: Mobile Backhaul Network

A typical mobile backhaul network with access rings and aggregate links is shown in figure above. The mobile backhaul networks deploy traffic engineering due to the strict Service Level Agreements(SLA). The TE paths may have additional constraints to avoid passing via different access rings or to get completely disjoint backup TE paths. The mobile backhaul networks towards the access side change frequently due to the growing mobile traffic and addition of new eNodeBs. It’s complex to satisfy the

requirements using cost, link color or explicit path configurations. The per-node administrative tag defined in this document can be effectively used to solve the problem for mobile backhaul networks. The nodes in different rings can be assigned with specific tags. TE path computation can be enhanced to consider additional constraints based on per-node administrative tags.

5. Policy-based Explicit Routing

Partially meshed network provides multiple paths between any two nodes in the network. In a data center environment, the topology is usually highly symmetric with many/all paths having equal cost. In a long distance network, this is usually less the case for a variety of reasons (e.g. historic, fiber availability constraints, different distances between transit nodes, different roles ...). Hence between a given source and destination, a path is typically preferred over the others, while between the same source and another destination, a different path may be preferred.

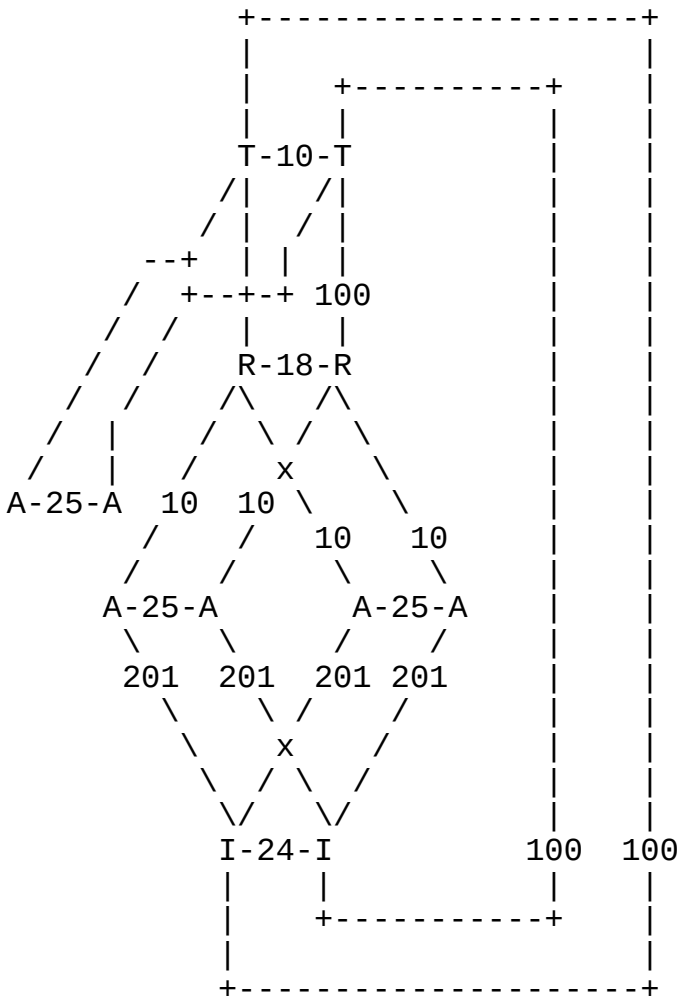


Figure 3: Explicit Routing topology

In the above topology, operator may want to enforce the following high level explicitly routed policies: - Traffic from A nodes to A nodes must not go through I nodes - Traffic from A nodes to I nodes must not go through R and T nodes with per-node administrative tag, tag A can be configured on all A nodes, (similarly I, R, T), and then configure this single CSPF policy on all A nodes to avoid I nodes for path calculation.

6. Security Considerations

This document does not introduce any further security issues other than those discussed in [ISO10589] and [RFC1195].

7. IANA Considerations

IANA maintains the registry for the Router Capability sub-TLVs. IS-IS Administrative Tags will require new type code for the following new sub-TLV defined in this document.

i) Per-Node-Admin-Tag Sub-TLV, Type: TBD

8. Acknowledgments

Many thanks to Les Ginsberg, Dhruv Dhody, Uma Chunduri for useful inputs. Thanks to Chris Bowers for providing useful inputs to remove ambiguity related to tag-ordering.

9. References

9.1. Normative References

[ISO10589] "Intermediate system to Intermediate system intra-domain routeing information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode Network Service (ISO 8473), ISO/IEC 10589:2002, Second Edition.", Nov 2002.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

9.2. Informative References

[I-D.ietf-rtgwg-lfa-manageability] Litkowski, S., Decraene, B., Filsfils, C., Raza, K., Horneffer, M., and p. psarkar@juniper.net, "Operational management of Loop Free Alternates", draft-ietf-rtgwg-lfa-manageability-04 (work in progress), August 2014.

[I-D.ietf-rtgwg-remote-lfa] Bryant, S., Filsfils, C., Previdi, S., Shand, M., and N. So, "Remote LFA FRR", draft-ietf-rtgwg-remote-lfa-09 (work in progress), December 2014.

[RFC1195] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", RFC 1195, December 1990.

[RFC4971] Vasseur, JP., Shen, N., and R. Aggarwal, "Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information", RFC 4971, July 2007.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, February 2008.

[RFC5286] Atlas, A. and A. Zinin, "Basic Specification for IP Fast Reroute: Loop-Free Alternates", RFC 5286, September 2008.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, October 2008.

Authors' Addresses

Pushpasis Sarkar (editor)
Juniper Networks, Inc.
Electra, Exora Business Park
Bangalore, KA 560103
India

Email: psarkar@juniper.net

Hannes Gredler
Juniper Networks, Inc.
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
US

Email: hannes@juniper.net

Shraddha Hegde
Juniper Networks, Inc.
Electra, Exora Business Park
Bangalore, KA 560103
India

Email: shraddha@juniper.net

Stephane Litkowski
Orange

Email: stephane.litkowski@orange.com

Bruno Decraene
Orange

Email: bruno.decraene@orange.com

Li Zhenbin
Huawei Technologies
Huawei Bld. No.156 Beiqing Rd
Beijing, KA 100095
China

Email: lizhenbin@huawei.com

Ebben Aries
Facebook

Email: exa@fb.com

Rafael Rodriguez
Facebook

Email: rafael@fb.com

Harish Raghuveer

Email: harish.r.prabhu@gmail.com