IS-IS for IP Internets
Internet-Draft
Intended status: Standards Track
Expires: June 3, 2016

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 1, 2015

Advertising Per-node Admin Tags in IS-IS
draft-ietf-isis-node-admin-tag-07

Abstract

This document describes an extension to the IS-IS routing protocol to add an optional operational capability, that allows tagging and grouping of the 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

Copyright Notice

Table of Contents

1. Introduction

It is useful to assign a per-node administrative tag to a router in the IS-IS domain and use it as an attribute associated with the node. The per-node administrative tag can be used in variety of applications, for example:

(a) Traffic-engineering applications to provide different path-selection criteria.

(b) Prefer or prune certain paths in Loop Free Alternate (LFA) backup selection via local policies as defined in [I-D.ietf-rtgwg-lfa-manageability].

This document provides mechanisms to advertise per-node administrative tags in IS-IS for route and path selection. Route and path selection functionality applies to both to Traffic Engineering(TE) and non-TE applications. Hence the new TLV for carrying per-node administrative tags is included in Router Capability TLV [RFC4971].

2. Per-Node Administrative Tags

An administrative Tag is a 32-bit integer value that can be used to identify a group of nodes in the IS-IS domain. An IS-IS router SHOULD advertise 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.

3. Per-Node Administrative Tag Sub-TLV

The new sub-TLV defined will be carried inside the IS-IS Router Capability TLV-242 [RFC4971]) in the Link State PDUs originated by the router. The new sub-TLV specifies one or more administrative tag values. TLV 242 can be either specified to be flooded within the specific level in which the same has been originated, or they can be specfied to be relayed from originating level to the other levels as well. Per-node administrative tags that are included in a 'level-specific' TLV 242 have a 'level-wide' flooding scope associated. On the other hand, per-node administrative tags included in a 'domain-wide' TLV 242 have 'domain-wide' flooding scope associated. For details on how TLV 242 are flooded and relayed in the entire network please, refer to [RFC4971]. Choosing the flooding scope to be associated with group tags, is defined by the needs of the operator's usage and is a matter of local policy or configuration. Operator MAY choose to advertise a different set of per-node administrative tags across levels and another set of per-node administrative tags within the specific level. Alternatively, the operator may use the same per-node administrative tags both within the 'domain-wide' flooding scope as well as within one or more 'level-wide' flooding scope.

Implementations SHOULD allow configuring one or more per-node administrative tags to be advertised from a given device along with the flooding scope associated with the same. It SHOULD allow provisioning a set of per-node administrative tags having a 'domain-wide' flooding scope, as well as, a set of per-node administrative tags with 'level-wide' flooding scope only. A given per-node administrative tag MAY be advertised within one or more 'level-wide' flooding scopes and/or within the 'domain-wide' scope.

The format of per-node Administrative Tag sub-TLV (see Section 3.1) does not include a topology identifier. Therefore it is not possible to indicate a topology specific context when advertising per-node admin tags. Hence, in deployments using multi-topology routing [RFC5120], advertising a separate set of per-node administrative tags for each topology SHOULD NOT be supported.

3.1. TLV format

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.

```
     0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |     Type      |    Length     |
    +- +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                   Administrative Tag #1                       |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                   Administrative Tag #2                       |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    //                                                             //
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                   Administrative Tag #N                       |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

    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 occurrence of the ’Per-node Admin Tag’ sub-TLV does not cancel previous announcements, but rather is cumulative.

4. Elements of Procedure

4.1. Interpretation of Per-Node Administrative Tags

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. Each tag carried by the The Per-Node Administrative Tag TLVs should be used to indicate a characteristic of a node that is independent of the characteristics indicated by other adminsitrative tags within the same or another instance of a Per-node Administrative Tag sub-TLV. The list of Per-node administrative tags carried in a Per-Node Administrative Tag sub-TLV MUST 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).

4.2. Use of Per-Node Administrative Tags

The per-node administrative tags are not meant to be extended by future IS-IS standards. New IS-IS extensions are not expected to 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. Future IS-IS extensions requiring well known values MAY define their own data signalling tailored to the needs of the feature or MAY use the capability TLV 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 associate advertised tags to changes in the network topology (both within and outside the IS-IS domain) or reachability of routes.

4.3. Processing Per-Node Administrative Tag changes

Multiple Per-Node Administrative Tag sub-TLVs MAY appear in a Router Capability TLV(TLV-242) or Per-Node Administrative Tag sub-TLVs MAY be contained in different instances of Router Capability TLVs. The Per-node administrative tags associated with a node that originates tags for the purpose of any computation or processing at a receiving node SHOULD be a superset of node administrative tags from all the TLVs in all the instances of Router Capability TLVs received in the Link-State PDU(s) advertised by the corresponding IS-IS router. When an Router Capability TLV is received that changes the set of per-node administrative tags applicable to any originating node, a receiving node MUST repeat any computation or processing that makes use of per-node administrative tags.

When there is a change or removal of an administrative affiliation of a node, the node MUST re-originate the Router Capability TLV(s) with the latest set of per-node administrative tags. On a receiving router, on detecting a change in contents (or removal) of existing Per-Node Administrative Tag sub-TLV(s) or addition of new Per-Node Administrative Tag sub-TLV(s) in any instance of Router Capability TLV(s), implementations MUST take appropriate measures to update their state according to the changed set of per-node administrative tags. The exact actions needed depend on features working with per-node administrative tags and is outside of scope of this specification.

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. An 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-Re-Route(FRR)

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 an 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 a 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

[RFC7490] defined 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 back-haul network service deployment

The topology of mobile back-haul networks 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.

```
                    ----------------
                   /                \
                  /                  \
                 /                    \
    +------+   +-----+    Access     +-----+
    |eNodeB|---|CSG1|    Ring 1     |ASG1|-------------
    +------+   +-----+               +-----+            \
                 \                    /                  \
                  \                  /                    +-----+   +---+
                   \             +-----+                  |RSG1|----|RNC|
                    -------------|     |   Aggregate     +-----+   +---+
                                 |ASG2|      Ring          |
                    -------------|     |                 +-----+   +---+
                   /             +-----+                 |RSG2|----|RNC|
                  /                  \                   +-----+   +---+
                 /                    \                  /
    +------+   +-----+    Access     +-----+           /
    |eNodeB|---|CSG2|     Ring 2     |ASG3|------------
    +------+   +-----+               +-----+
                \                     /
                 \                   /
                  \                 /
                   -----------------
```

Figure 2: Mobile Backhaul Network

A typical mobile back-haul network with access rings and aggregate links is shown in figure above. The mobile back-haul 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 back-haul networks towards the access side change frequently due to the growing mobile traffic and addition of new LTE Evolved NodeBs (eNodeB). 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 back-haul 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

A partially meshed network provides multiple paths between any two nodes in the network. In a data centre 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, fibre 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.

```
        +----------------------+   +----------------+
        |                       \ /                 |
        |   +-----------------+  x   +---------+    |
        |   |                  \/  \/          |    |
        |   |                +-T-10-T          |    |
        |   |               /  |   /|          |    |
        |   |              /  100 / |          |    |
        |   |             /    | | 100         |    |
        |   |            /   +-+-+  |          |    |
        |   |           /   /  |    |          |    |
        |   |          /   /   R-18-R          |    |
        |   |        10   10  /\   /\          |    |
        |   |        /   /   /  \ /  \         |    |
        |   |       /   /   /    x    \        |    |
        |   |      /   /   10  10 \    \       |    |
        |   |     /   /   /    /   10   10     |    |
        |   |    /   /   /    /     \    \     |    |
        |   |   A-25-A  A-25-A       A-25-A    |    |
        |   |   |    |   \    \     /    /     |    |
        |   |   |    |   201  201  201 201     |    |
        |   |   |    |     \    \ /    /       |    |
        |   |  201  201     \    x    /        |    |
        |   |   |    |       \  / \  /         |    |
        |   |   |    |        \/   \/          |    |
        |   |   I-24-I        I-24-I          100  100
        |   |  /    /         |    |           |    |
        |   +-+    /          |    +-----------+    |
        +---------+           +---------------------+
```

Figure 3: Explicit Routing topology

In the above topology, operator may want to enforce the following high level explicit routing policies:

1. - Traffic from A nodes to A nodes should preferably go through R or T nodes (rather than through I nodes).

2. - Traffic from A nodes to I nodes must not go through R and T nodes.

With node admin tags, tag A (resp. I, R, T) can be configured on all A (resp. I, R, T) nodes to advertise their role. The first policy is about preferring one path over another. Given the chosen metrics, it is achieved with regular SPF routing. The second policy is about prohibiting (pruning) some paths. It requires an explicit routing policy. With the use of node tags, this may be achieved with a generic CSPF policy configured on A nodes: for destination nodes having the tag "A" runs a CSPF with the exclusion of nodes having the tag "I".

6. Security Considerations

Node administrative tags may be used by operators to indicate geographical location or other sensitive information. The information carried in node administrative tags could be leaked to an IGP snooper. This document does not introduce any new security issues. Security concerns for IS-IS are already addressed in [ISO10589], [RFC5304], and [RFC5310] and are applicable to the mechanisms described in this document. Extended authentication mechanisms described in [RFC5304] or [RFC5310] SHOULD be used in deployments where attackers have access to the physical networks and nodes included in the IS-IS domain are vulnerable.

Advertisement of tag values for one administrative domain into another invloves the risk mis-interpretation of the tag values (if the two domains have assigned different meanings to the same values), which may have undesirable and unanticipated side effects.

7. Operational Considerations

Operators can assign meaning to the per-node administrative tags which is local to the operator’s administrative domain. The operational use of per-node administrative tags is analogical to the IS-IS prefix tags [RFC5130] and BGP communities [RFC1997]. Operational discipline and procedures followed in configuring and using BGP communities and ISIS Prefix tags is also applicable to the usage of per-node administrative tags.

Defining language for local policies is outside the scope of this document. As in case of other policy applications, the pruning policies can cause the path to be completely removed from forwarding plane,and hence have the potential for more severe operational impact (e.g., node unreachability due to path removal) by comparison to preference policies that only affect path selection.

8. Manageability Considerations

Per-node administrative tags are configured and managed using routing policy enhancements. YANG data definition language is the latest model to describe and define configuration for network devices. IS-IS YANG data model is described in [I-D.ietf-isis-yang-isis-cfg] and routing policy configuration model is described in [I-D.ietf-rtgwg-policy-model]. These two documents will be enhanced to include the node administrative tag related configurations.

9. 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

10. Acknowledgments

Many thanks to Les Ginsberg, Dhruv Dhody, Uma Chunduri and Chris Bowers for providing useful inputs.

11. References

11.1. 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, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC4971] Vasseur, JP., Ed., Shen, N., Ed., and R. Aggarwal, Ed., "Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information", RFC 4971, DOI 10.17487/RFC4971, July 2007, <http://www.rfc-editor.org/info/rfc4971>.

[RFC7490] Bryant, S., Filsfils, C., Previdi, S., Shand, M., and N. So, "Remote Loop-Free Alternate (LFA) Fast Reroute (FRR)", RFC 7490, DOI 10.17487/RFC7490, April 2015, <http://www.rfc-editor.org/info/rfc7490>.

11.2. Informative References

[I-D.ietf-isis-yang-isis-cfg]
Litkowski, S., Yeung, D., Lindem, A., Zhang, J., and L. Lhotka, "YANG Data Model for IS-IS protocol", draft-ietf-isis-yang-isis-cfg-07 (work in progress), November 2015.

[I-D.ietf-rtgwg-lfa-manageability]
Litkowski, S., Decraene, B., Filsfils, C., Raza, K., Horneffer, M., and P. Sarkar, "Operational management of Loop Free Alternates", draft-ietf-rtgwg-lfa-manageability-11 (work in progress), June 2015.

[I-D.ietf-rtgwg-policy-model]
Shaikh, A., rjs@rob.sh, r., D’Souza, K., and C. Chase, "Routing Policy Configuration Model for Service Provider Networks", draft-ietf-rtgwg-policy-model-00 (work in progress), September 2015.

[RFC1997] Chandra, R., Traina, P., and T. Li, "BGP Communities Attribute", RFC 1997, DOI 10.17487/RFC1997, August 1996, <http://www.rfc-editor.org/info/rfc1997>.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, DOI 10.17487/RFC5120, February 2008, <http://www.rfc-editor.org/info/rfc5120>.

[RFC5130] Previdi, S., Shand, M., Ed., and C. Martin, "A Policy Control Mechanism in IS-IS Using Administrative Tags", RFC 5130, DOI 10.17487/RFC5130, February 2008, <http://www.rfc-editor.org/info/rfc5130>.

[RFC5286] Atlas, A., Ed. and A. Zinin, Ed., "Basic Specification for IP Fast Reroute: Loop-Free Alternates", RFC 5286, DOI 10.17487/RFC5286, September 2008, <http://www.rfc-editor.org/info/rfc5286>.

[RFC5304] Li, T. and R. Atkinson, "IS-IS Cryptographic Authentication", RFC 5304, DOI 10.17487/RFC5304, October 2008, <http://www.rfc-editor.org/info/rfc5304>.

[RFC5310] Bhatia, M., Manral, V., Li, T., Atkinson, R., White, R., and M. Fanto, "IS-IS Generic Cryptographic Authentication", RFC 5310, DOI 10.17487/RFC5310, February 2009, <http://www.rfc-editor.org/info/rfc5310>.

Authors’ Addresses

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

Email: psarkar@juniper.net; pushpasis.ietf@gmail.com

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

Email: hannes@gredler.at

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
1 Hacker Way
Menlo Park, CA 94025
US

Email: exa@dscp.org

Rafael Rodriguez
Facebook
1 Hacker Way
Menlo Park, CA 94025
US

Email: rafael@fb.com

Harish Raghuveer

Email: harish.r.prabhu@gmail.com