Inter-Domain Routing   P. Sarkar, Ed.
Internet-Draft   Arrcus, Inc.
Intended status: Standards Track   H. Gredler
Expires: January 26, 2018   RtBrick, Inc.
                                                            S. Litkowski
                                                                  Orange
                                                           July 25, 2017

        Advertising Node Admin Tags in BGP Link-State Advertisements
            draft-ietf-idr-bgp-ls-node-admin-tag-extension-02

Abstract

   This document describes the protocol extensions to collect node
   administrative tags adevertised in IGP Link State advertisements and
   disseminate the same in BGP Link-State advertisement protocol, to
   facilitate inter-AS TE applications that may need the same node
   administrative tags to associate a subset of network devices spanning
   across more than one AS with a specific functionality.

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/.

   Internet-Drafts are draft documents valid for a maximum of six months
   and may be updated, replaced, or obsoleted by other documents at any
   time. It is inappropriate to use Internet-Drafts as reference
   material or to cite them other than as "work in progress."

   This Internet-Draft will expire on January 26, 2018.

Table of Contents

1. Introduction

   Advertising Node Administrative Tags in Link State protocols like IS-
   IS [RFC7917] and OSPF [RFC7777] defines an optional operational
   capability, that allows tagging and grouping of the nodes in a IGP
   domain. This, among other applications, allows simple management and
   easy control over route and path selection, based on local configured
   policies. However, node administrative tags advertised in IGP
   advertisements let network operators associate nodes within a single
   AS (if not a single area). This limits the use of such node
   administrative tags and applications that need to associate a subset

of network devices spanning across multiple AS with a specific
functionality cannot use them.

To address the need for applications that require visibility into
Link State Databases (LSDBs) across IGP areas, or even across ASes,
the BGP-LS address-family/sub-address-family have been defined that
allows BGP to carry LSDB information. The BGP Network Layer
Reachability Information (NLRI) encoding format for BGP-LS and a new
BGP Path Attribute called BGP-LS attribute are defined in [RFC7752].
Please refer to [RFC7752] for more details.

For the purpose of advertising node administrative tags within BGP
Link-State advertisements, a new Node Attribute TLV to be carried in
the corresponding BGP-LS Node NLRI is proposed. For more details on
the Node Attribute TLVs please refer to section 3.3.1 in [RFC7752]

2. Per-Node Administrative Tag

An administrative Tag is a 32-bit integer value that can be used to
identify a group of nodes in the entire routing domain. The new TLV
and sub-TLV proposed in IS-IS [RFC7917] and OSPF [RFC7777]
respectively, specifies one or more administrative tag values. A BGP
Link-State speaker that also participates in the IGP link state
advertisements exchange may learn one or more node administrative
tags advertised by another router in the same IGP domain. Such BGP-
LS speaker shall encode the same set of node administrative tags in
the corresponding Node Attribute TLV representing the network device
that originated the node administrative tags.

The node administrative tags advertised in IGP link state
advertisements will have either per-area(or per-level in IS-IS)scope
or 'global' scope. An operator may choose to advertise one set of
node administrative tags across areas (or levels in IS-IS) and
advertise another set of node administrative tags within a specific
area (or level). But evidently two areas within the same AS or two
different AS's may use the same node administrative tag for different
purposes. In such a case, applications will need to distinguish
between the per-area(or per-level) scoped administrative tags
originated from a specific node against those originated from the
same node with 'global' scope.

A BGP-LS router in a given AS while copying the node administrative
tags learnt from IGP link-state advertisements, MUST also copy the
scope associated with the node administrative tags. Refer to
Section 3.1 for how to encode the associated scope of a node
administrative tags as well.

To be able to distinguish between the significance of an
administrative tag learnt in one area, from that advertised in
another area, or another AS, any applications receiving such a BGP-LS
advertisement MUST consider the scope associated with each node
administrative tag along with the area(or level in IS-IS) and the AS
number of the originating node associated with corresponding IGP link
state advertisement. The area(or level) associated with the
corresponding IGP link state advertisement and the AS number
associated with the originating node can be derived from appropriate
node attributes (already defined in BGP-LS [RFC7752]) attached with
the corresponding Node NLRI. [RFC7752] specifies that ISIS level
information be encoded in Node NLRI [1] and OSPF Area Identifiers be
encoded in Node Descriptor Sub-TLVs [2].

3. BGP-LS Extensions for Per-Node Administrative Tags

The BGP-LS NLRI can be a node NLRI, a link NLRI or a prefix NLRI.
The corresponding BGP-LS attribute is a node attribute, a link
attribute or a prefix attribute. BGP-LS [RFC7752] defines the TLVs
that map link-state information to BGP-LS NLRI and BGP-LS attribute.
This document adds an new Node Attribute TLV called 'Node Admin Tag
TLV' to encode node administrative tags information.

[RFC7917] defines the 'Node Admin Tag' sub-TLV in the Router
Capability TLV (type 242) in IS-IS Link State PDUs to encode node
administrative tags. Similarly [RFC7917] defines the 'Node
Administrative Tag' TLV in OSPF Router Information LSAs to encode
node administrative tags in OSPF Link State update packets. The node
administrative tags TLVs learnt from the IGP link state
advertisements of a specific node will all be inserted in a new Node
Admin Tag TLV and added to the corresponding Node are mapped to the
corresponding BGP-LS Node NLRI. Node administrative tags from IGP
advertisements are mapped to the corresponding Node Admin Tag TLV in
the following way.

TLV Code Point	Description	Length	IS-IS TLV /sub-TLV	OSPF LSA/TLV
TBD	Node Admin Tag TLV	Variable	242/21 [3]	RI-LSA/10 [4]

Table 1: Node Admin Tag TLV Mapping from IGP

3.1. Node Admin Tag TLV

   The new Node Administrative Tag TLV, like other BGP-LS Node Attribute
   TLVs, is formatted as Type/Length/Value (TLV)triplets. Figure 1
   below shows the format of the new 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           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Flags                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Administrative Tag #1                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Administrative Tag #2                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                                                               //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Administrative Tag #N                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
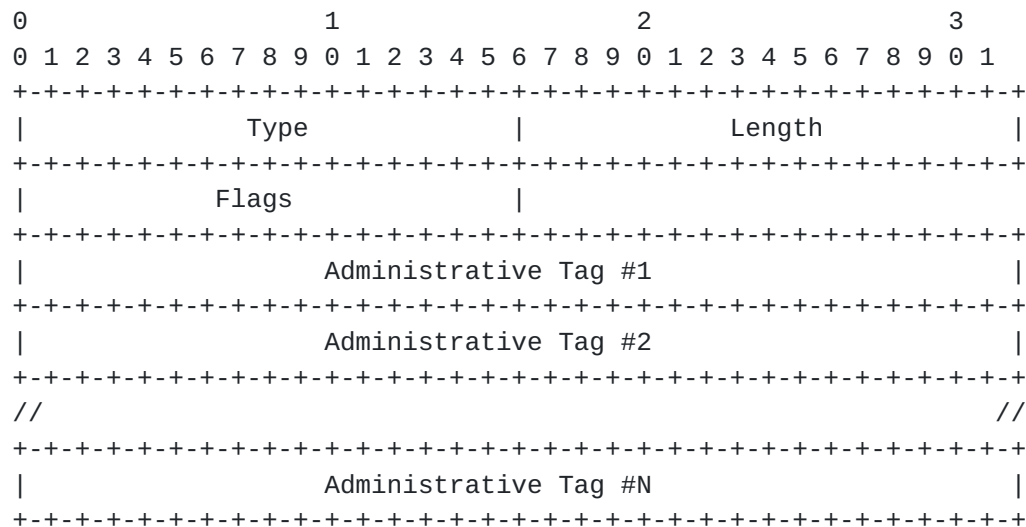
Type : A 2-octet field specifiying code-point of the new
        TLV type. Code-point: TBA (suggested 1040)

Length: A 2-octet 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 2-octet 'Flags' field, followed by a sequence of multiple
        4 octets defining the administrative tags.

Flags: A 2-octet field that carries flags associated with
        all the administrative flags encoded in this TLV.
        Following is the format of this field.

```
         0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
        +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
        |L|            Reserved         |
        +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

        The following bit flags are defined:

        L bit : If the L bit is set (1), it signifies that
                all administrative flags encoded in this
                TLV has per-area(or level in IS-IS) scope,
                and should not be mixed with ones with same
                value but with 'global' scope (L bit reset
                to 0).

            Figure 1: BGP Link-State Node Administrative Tag TLV

This new type of 'Node Admin Tag' TLVs can ONLY be added to the Node
Attribute associated with the Node NLRI that originates the
corresponding node administrative tags in an IGP domain.

All the node administrative tags with 'per-area' (or per-level)
scope, originated by a single node in an IGP domain SHALL be re-
originated in a single 'Node Admin Tag' TLV and inserted in the Node
NLRI generated for the same node. Similarly, all the node
administrative tags with 'global' scope originated by the same node
in IGP domain SHALL be re-originated in another 'Node Admin Tag' TLV
and inserted in the same Node NLRI generated for the originating
node. Multiple instances of a TLV may be generated by the BGP-LS
router for a given node in the IGP domain. This MAY happen if the
original node's link state advertisement carries more than 16383 node
administrative groups and a single TLV does not provide sufficient
space. As such multiple occurence of the 'Node Admin Tag' TLVs under
a single BGP LS NLRI is cumulative.

While copying node administrative tags from IGP link-state
advertisements to corresponding BGP-LS advertisements, the said BGP-
LS speaker MAY run all the node administrative flags through a
locally configured policy that selects which ones should be exported
and which ones not. And then the node administrative tag is copied
to the BGP-LS advertisement if it is permitted to do so by the said
policy. Definition of such a policy is outside the scope of this
document.

4. Elements of Procedure

Meaning of the Node administrative tags is generally opaque to the
BGP Link-State protocol. A router advertising the 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 node
administrative tag is defined by the network local policy. However
multiple administrative domain owners may agree on a common meaning
implied by an administrative tag for mutual benefit.

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. Node administrative tags

carried by the Node Admin Tag TLV SHOULD be used to indicate
independent characteristics of the node in the IGP domain that
originated it. 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).

For more details on guidance regarding usage of node administrative
tags please refer to section 4 [5] in [RFC7917] or section 2.2.1 [6]
in [RFC7777].

5. Applications

[RFC7917] and [RFC7777] present some applications of node
administrative tags.

The Policy-based Explicit routing use case can be extended to inter-
area or inter-AS scenarios where an end to end path needs to avoid or
include nodes that have particular properties. Following are some
examples.

1. Geopolitical routing : preventing traffic from country A to
    country B to cross country C. In this case, we may use node
    administrative tags to encode geographical information (country).
    Path computation may be required to take into account node
    administrative tag to permit avoidance of nodes belonging to
    country C.

2. Legacy node avoidance : in some specific cases, it is interesting
    for a service-provider to force some traffic to avoid legacy
    nodes in the network. For example, legacy nodes may not be
    carrier class (no high availability), and a service provider may
    want to ensure that critical traffic only uses nodes that are
    providing high availability.

In case of inter-AS Traffic-Engineering applications, different ASes
SHOULD share their administrative tag policies. They MAY also need
to agree upon some common tagging policy for specific applications.

For more details on some possible applications with node
administrative tags please refer to section 3 [7] in [RFC7777].

6. IANA Considerations

   This document requests assigning code-points from the registry for
   BGP-LS attribute TLVs based on Table 2.

7. Manageability Considerations

   This section is structured as recommended in [RFC5706].

7.1. Operational Considerations

7.1.1. Operations

   Existing BGP and BGP-LS operational procedures apply. No new
   operational procedures are defined in this document.

8. TLV/Sub-TLV Code Points Summary

   This section contains the global table of all TLVs/Sub-TLVs defined
   in this document.

```
         +----------------+----------------+----------+
         | TLV Code Point | Description    | Length   |
         +----------------+----------------+----------+
         |      1040      | Node Admin Tag | variable |
         +----------------+----------------+----------+
```

            Table 2: Summary Table of TLV/Sub-TLV Codepoints

9. Security Considerations

   Procedures and protocol extensions defined in this document do not
   affect the BGP security model. See the 'Security Considerations'
   section of [RFC4271] for a discussion of BGP security. Also refer to
   [RFC4272] and [RFC6952] for analysis of security issues for BGP.

10. Acknowledgements

   TBA.

11. References

11.1. Normative References

   [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>.

   [RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A
              Border Gateway Protocol 4 (BGP-4)", RFC 4271,
              DOI 10.17487/RFC4271, January 2006,
              <http://www.rfc-editor.org/info/rfc4271>.

   [RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and
              S. Ray, "North-Bound Distribution of Link-State and
              Traffic Engineering (TE) Information Using BGP", RFC 7752,
              DOI 10.17487/RFC7752, March 2016,
              <http://www.rfc-editor.org/info/rfc7752>.

11.2. Informative References

   [RFC4272] Murphy, S., "BGP Security Vulnerabilities Analysis",
              RFC 4272, DOI 10.17487/RFC4272, January 2006,
              <http://www.rfc-editor.org/info/rfc4272>.

   [RFC5706] Harrington, D., "Guidelines for Considering Operations and
              Management of New Protocols and Protocol Extensions",
              RFC 5706, DOI 10.17487/RFC5706, November 2009,
              <http://www.rfc-editor.org/info/rfc5706>.

   [RFC6952] Jethanandani, M., Patel, K., and L. Zheng, "Analysis of
              BGP, LDP, PCEP, and MSDP Issues According to the Keying
              and Authentication for Routing Protocols (KARP) Design
              Guide", RFC 6952, DOI 10.17487/RFC6952, May 2013,
              <http://www.rfc-editor.org/info/rfc6952>.

   [RFC7777] Hegde, S., Shakir, R., Smirnov, A., Li, Z., and B.
              Decraene, "Advertising Node Administrative Tags in OSPF",
              RFC 7777, DOI 10.17487/RFC7777, March 2016,
              <http://www.rfc-editor.org/info/rfc7777>.

   [RFC7917] Sarkar, P., Ed., Gredler, H., Hegde, S., Litkowski, S.,
              and B. Decraene, "Advertising Node Administrative Tags in
              IS-IS", RFC 7917, DOI 10.17487/RFC7917, July 2016,
              <http://www.rfc-editor.org/info/rfc7917>.

11.3. URIs

   [1] http://tools.ietf.org/html/rfc7752#section-3.2

   [2] http://tools.ietf.org/html/rfc7752#section-3.2.1.4

   [3] http://tools.ietf.org/html/rfc7917#section-3.1

   [4] http://tools.ietf.org/html/rfc7777#section-2.1

    [5] http://tools.ietf.org/html/rfc7917#section-4

    [6] http://tools.ietf.org/html/rfc7777#section-2.2.1

    [7] http://tools.ietf.org/html/rfc7777#section-3

Authors' Addresses

    Pushpasis Sarkar (editor)
    Arrcus, Inc.

    Email: pushpasis.ietf@gmail.com

    Hannes Gredler
    RtBrick, Inc.

    Email: hannes@rtbrick.com

    Stephane Litkowski
    Orange

    Email: stephane.litkowski@orange.com