Internet Engineering Task Force (IETF)
Request for Comments: 7777
Category: Standards Track
ISSN: 2070-1721

S. Hegde
Juniper Networks, Inc.
R. Shakir
Jive Communications, Inc.
A. Smirnov
Cisco Systems, Inc.
Z. Li
Huawei Technologies
B. Decraene
Orange

March 2016

Advertising Node Administrative Tags in OSPF

Abstract

This document describes an extension to the OSPF protocol to add an optional operational capability that allows tagging and grouping of the nodes in an OSPF domain. This allows simplification, ease of management and control over route and path selection based on configured policies. This document describes an extension to the OSPF protocol to advertise node administrative tags. The node tags can be used to express and apply locally defined network policies, which are a very useful operational capability. Node tags may be used by either OSPF itself or other applications consuming information propagated via OSPF.

This document describes the protocol extensions to disseminate node administrative tags to the OSPFv2 and OSPFv3 protocol. It provides example use cases of administrative node tags.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc7777.

Table of Contents

1. Introduction

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

(a) Traffic Engineering (TE) 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 [LFA-MANAGE].

This document provides mechanisms to advertise node administrative tags in OSPF for route and path selection. Route and path selection functionality applies to both TE and non-TE applications; hence, a new TLV for carrying node administrative tags is included in Router Information (RI) Link State Advertisement (LSA) [RFC7770].

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].

2. OSPF Node Admin Tag TLV

An administrative tag is a 32-bit integer value that can be used to identify a group of nodes in the OSPF domain.

The newly defined TLV is carried within an RI LSA for OSPFV2 and OSPFV3. RI LSA [RFC7770] can have flooding scope at the link, area, or Autonomous System (AS) level. The choice of what scope at which to flood the group tags is a matter of local policy. It is expected that node administrative tag values will not be portable across administrative domains.

The TLV specifies one or more administrative tag values. An OSPF node advertises the set of groups it is part of in the OSPF domain (for example, all PE nodes are configured with a certain tag value, and all P nodes are configured with a different tag value in the domain). Multiple TLVs MAY be added in same RI LSA or in a different instance of the RI LSA as defined in [RFC7770].

2.1. TLV Format

[RFC7770] defines the RI LSA, which may be used to advertise properties of the originating router. The payload of the RI LSA consists of one or more nested Type/Length/Value (TLV) triplets.

Node administrative tags are advertised in the Node Admin Tag TLV. The format of the Node Admin Tag TLV is:

```
 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                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: OSPF Node Admin Tag TLV

Type: 10

Length: A 16-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 set of administrative tags. Each tag is a 32-bit integer value. At least one tag MUST be carried if this TLV is included in the RI LSA.

2.2. Elements of Procedure

2.2.1. Interpretation of Node Administrative Tags

The meaning of the node administrative tags is generally opaque to OSPF. Routers advertising the node administrative tag (or tags) may be configured to do so without knowing (or even without supporting processing of) the functionality implied by the tag. This section describes general rules, regulations, and guidelines for using and interpreting an administrative tag that will facilitate interoperable implementations by vendors.

Interpretation of tag values is specific to the administrative domain of a particular network operator; hence, tag values SHOULD NOT be propagated outside the administrative domain to which they apply. The meaning of a 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 or does not have a local policy corresponding to the tag, it ignores the specific tag and floods the RI LSA without any change as defined in [RFC7770].

The semantics of the tag order has no meaning. That is, 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 must be treated as an independent identifier that may be used in the policy to perform a policy action. Each tag carried by the Node Admin Tag TLV should be used to indicate a characteristic of a node that is independent of the characteristics indicated by other administrative tags. The administrative-tag list within the TLV MUST be considered an unordered list. While 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).

2.2.2. Use of Node Administrative Tags

The node administrative tags are not meant to be extended by future OSPF standards. New OSPF extensions are not expected to require use of node administrative tags or define well-known tag values. Node administrative tags are for generic use and do not require IANA registration. Future OSPF extensions requiring well-known values MAY define their own data signaling tailored to the needs of the feature or MAY use the capability TLV as defined in [RFC7770].

Being part of the RI LSA, the Node Admin Tag TLV must be reasonably small and stable. In particular, implementations supporting node administrative tags MUST NOT be used to convey attributes of the routing topology or associate tags with changes in the network topology (both within and outside the OSPF domain) or reachability of routes.

2.2.3. Processing Node Administrative Tag Changes

Multiple Node Admin Tag TLVs MAY appear in an RI LSA or multiple Node Admin Tag TLVs MAY be contained in different instances of the RI LSA. The 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 received RI LSA instances in the Link-State Database (LSDB) advertised by the corresponding OSPF router. When an RI LSA is received that changes the set of tags applicable to any originating node, which has features depending on node administrative tags, a receiving node MUST repeat any computation or processing that is based on those administrative tags.

When there is a change or removal of an administrative affiliation of a node, the node MUST re-originate the RI LSA with the latest set of node administrative tags. On the receiver, when there is a change in the Node Admin Tag TLV or removal/addition of a TLV in any instance of the RI LSA, implementations MUST take appropriate measures to update their state according to the changed set of tags. The exact actions needed depend on features working with administrative tags and are outside of scope of this specification.

3. Applications

This section lists several examples of how implementations might use the node administrative tags. These examples are given only to demonstrate the generic usefulness of the router tagging mechanism. Implementations supporting this specification are not required to implement any of these 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.

3.1. Service Auto-Discovery

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

For example, a service provider might desire to establish a full mesh of MPLS TE tunnels between all PE routers in the area of the 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 a new PE router is added to the area, all other PE devices will open TE tunnels to it without needing to reconfigure them.

3.2. Fast-Rerouting Policy

Increased deployment of Loop-Free Alternates (LFA) as defined in [RFC5286] poses operation and management challenges. [LFA-MANAGE] 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 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 that advertise the Core or Distribution tag. This requires minimal per-node configuration and the network automatically adapts when new links or routers are added.

(b) LFA calculation optimization

Calculation of LFA paths may require significant resources of the router. One execution of Dijkstra's algorithm is required for each neighbor eligible to become the next hop of repair paths. Thus, a router with a few hundred 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 that are known to provide no valid LFA (such as single-connected routers) may significantly reduce the number of Dijkstra algorithm runs.

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

3.3. Controlling Remote LFA Tunnel Termination

[RFC7490] defined a method of tunneling traffic to extend the basic LFA coverage after connection failure of a link and defined an algorithm to find tunnel tail-end routers meeting the LFA requirement. In most cases, the proposed algorithm finds more than one candidate tail-end router. In a 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 node (the router does not accept the targeted LDP sessions required to implement remote LFA tunneling) or due to administrative requirements (for example, it may be desirable to choose the tail-end router among colocated devices).

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

3.4. Mobile Backhaul Network Service Deployment

Mobile backhaul networks usually adopt a ring topology to save fibre resources; it is usually divided into the aggregate network and the access network. Cell Site Gateways (CSGs) connects the LTE Evolved NodeBs (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 ring will attach to one pair of aggregate site gateways (ASGs) and more than one aggregate ring 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 backhaul network with access rings and aggregate links is shown in the figure above. The mobile backhaul networks deploy traffic engineering due to strict Service Level Agreements (SLAs). 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 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 node administrative tags.

3.5. Explicit Routing Policy

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 not the case, for a variety of reasons (e.g., historic, fibre availability constraints, different distances between transit nodes, and 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, an operator may want to enforce the following high-level explicit routing policies:

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

- o 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 Constrained Shortest Path First (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".

4. Security Considerations

Node administrative tags may be used by operators to indicate geographical location or other sensitive information. As indicated in [RFC2328] and [RFC5340], OSPF authentication mechanisms do not provide confidentiality and the information carried in node administrative tags could be leaked to an IGP snooper. Confidentiality for the OSPF control packets can be achieved by either running OSPF on top of IP Security (IPsec) tunnels or by applying IPsec-based security mechanisms as described in [RFC4552].

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

[RFC4593] and [RFC6863] discuss the generic threats to routing protocols and OSPF, respectively. These security threats are also applicable to the mechanisms described in this document. OSPF authentication described in [RFC2328] and [RFC5340] or extended authentication mechanisms described in [RFC7474] or [RFC7166] SHOULD be used in deployments where attackers have access to the physical networks and nodes included in the OSPF domain are vulnerable.

5. Operational Considerations

Operators can assign meaning to the node administrative tags, which are local to the operator’s administrative domain. The operational use of 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 IS-IS prefix tags is also applicable to the usage of node administrative tags.

Defining language for local policies is outside the scope of this document. As is the 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.

6. Manageability Considerations

Node administrative tags are configured and managed using routing policy enhancements. The YANG data definition language is the latest model to describe and define configuration for network devices. The OSPF YANG data model is described in [OSPF-YANG] and the routing policy configuration model is described in [RTG-POLICY]. These two documents will be enhanced to include the configurations related to the node administrative tag.

7. IANA Considerations

This specification updates the "OSPF Router Information (RI) TLVs" registry. IANA has registered the following value:

Node Admin Tag TLV - 10

8. References

8.1. 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>.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, DOI 10.17487/RFC2328, April 1998, <http://www.rfc-editor.org/info/rfc2328>.

[RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, DOI 10.17487/RFC5340, July 2008, <http://www.rfc-editor.org/info/rfc5340>.

[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>.

[RFC7770] Lindem, A., Ed., Shen, N., Vasseur, JP., Aggarwal, R., and S. Shaffer, "Extensions to OSPF for Advertising Optional Router Capabilities", RFC 7770, DOI 10.17487/RFC7770, February 2016, <http://www.rfc-editor.org/info/rfc7770>.

8.2. Informative References

[LFA-MANAGE]
Litkowski, S., Decraene, B., Filsfils, C., Raza, K., Horneffer, M., and P. Sarkar, "Operational management of Loop Free Alternates", Work in Progress, draft-ietf-rtgwg-lfa-manageability-11, June 2015.

[OSPF-YANG]
Yeung, D., Qu, Y., Zhang, J., Bogdanovic, D., and K. Koushik, "Yang Data Model for OSPF Protocol", Work in Progress, draft-ietf-ospf-yang-03, October 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>.

[RFC4552] Gupta, M. and N. Melam, "Authentication/Confidentiality for OSPFv3", RFC 4552, DOI 10.17487/RFC4552, June 2006, <http://www.rfc-editor.org/info/rfc4552>.

[RFC4593] Barbir, A., Murphy, S., and Y. Yang, "Generic Threats to Routing Protocols", RFC 4593, DOI 10.17487/RFC4593, October 2006, <http://www.rfc-editor.org/info/rfc4593>.

[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>.

[RFC6863] Hartman, S. and D. Zhang, "Analysis of OSPF Security According to the Keying and Authentication for Routing Protocols (KARP) Design Guide", RFC 6863, DOI 10.17487/RFC6863, March 2013, <http://www.rfc-editor.org/info/rfc6863>.

[RFC7166] Bhatia, M., Manral, V., and A. Lindem, "Supporting Authentication Trailer for OSPFv3", RFC 7166, DOI 10.17487/RFC7166, March 2014, <http://www.rfc-editor.org/info/rfc7166>.

[RFC7474] Bhatia, M., Hartman, S., Zhang, D., and A. Lindem, Ed., "Security Extension for OSPFv2 When Using Manual Key Management", RFC 7474, DOI 10.17487/RFC7474, April 2015, <http://www.rfc-editor.org/info/rfc7474>.

[RTG-POLICY] Shaikh, A., Shakir, R., D'Souza, K., and C. Chase, "Routing Policy Configuration Model for Service Provider Networks", Work in Progress, draft-ietf-rtgwg-policy-model-00, September 2015.

Contributors

Thanks to Hannes Gredler for his substantial review, guidance, and editing of this document. Thanks to Harish Raguveer for his contributions to initial draft versions of this document.

Acknowledgements

Thanks to Bharath R, Pushpasis Sarakar, and Dhruv Dhody for useful input. Thanks to Chris Bowers for providing useful input to remove ambiguity related to tag ordering. Thanks to Les Ginsberg and Acee Lindem for the input. Thanks to David Black for careful review and valuable suggestions for the document, especially for the operations section.

Authors' Addresses

Shraddha Hegde
Juniper Networks, Inc.
Embassy Business Park
Bangalore, KA 560093
India

Email: shraddha@juniper.net

Rob Shakir
Jive Communications, Inc.
1275 W 1600 N, Suite 100
Orem, UT 84057
United States

Email: rjs@rob.sh

Anton Smirnov
Cisco Systems, Inc.
De Kleetlaan 6a
Diegem 1831
Belgium

Email: as@cisco.com

Li zhenbin
Huawei Technologies
Huawei Bld. No.156 Beiqing Rd
Beijing 100095
China

Email: lizhenbin@huawei.com

Bruno Decraene
Orange

Email: bruno.decraene@orange.com