Open Shortest Path First IGP
Internet-Draft
Intended status: Standards Track
Expires: December 3, 2015

S. Hegde
Juniper Networks, Inc.
H. Raghuveer

H. Gredler
Juniper Networks, Inc.
R. Shakir
British Telecom
A. Smirnov
Cisco Systems, Inc.
Z. Li
Huawei Technologies
B. Decraene
Orange
June 1, 2015

Advertising per-node administrative tags in OSPF
draft-ietf-ospf-node-admin-tag-02

Abstract

This document describes an extension to OSPF protocol [RFC2328] 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 OSPF protocol [RFC2328] to advertise per-node administrative tags. This optional operational capability allows to express and act upon locally-defined network policy which considers node properties conveyed by tags. Node tags may be used either by OSPF itself or by other applications consuming information propagated via OSPF.

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

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.

Copyright Notice

Table of Contents

1. Introduction

It is useful to assign a per-node administrative tag to a router in the OSPF domain and use it as an attribute associated with the node. The per-node administrative tag can be used in variety of applications, for ex: - Traffic-engineering applications to provide different path-selection criteria, - Prefer or prune certain paths in Loop Free Alternate (LFA) backup selection via local policies.

This document provides mechanisms to advertise per-node administrative tags in the OSPF. Path selection is a functional set which applies both to TE and non-TE applications and hence new TLV for carrying per-node administrative tags is included in Router Information LSA [RFC4970] .

2. Administrative 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 new TLV defined will be carried within an RI LSA for OSPFV2 and OSPFV3. Router information LSA [RFC4970] can have link, area or AS level flooding scope. Choosing the flooding scope to flood the group tags are defined by the policies and is a local matter.

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 certain tag value, all P-nodes are configured with a different tag value in a domain). Multiple TLVs MAY be added in same RI-LSA or in different instance of the RI LSA as defined in [I-D.acee-ospf-rfc4970bis].

3. OSPF per-node administrative tag TLV

3.1. TLV format

[RFC4970], defines Router Information (RI) LSA which may be used to advertise properties of the originating router. Payload of the RI LSA consists of one or more nested Type/Length/Value (TLV) triplets. Node administrative tags are advertised in the Node Administrative Tag TLV. The format of Node Administrative Tag TLV is:

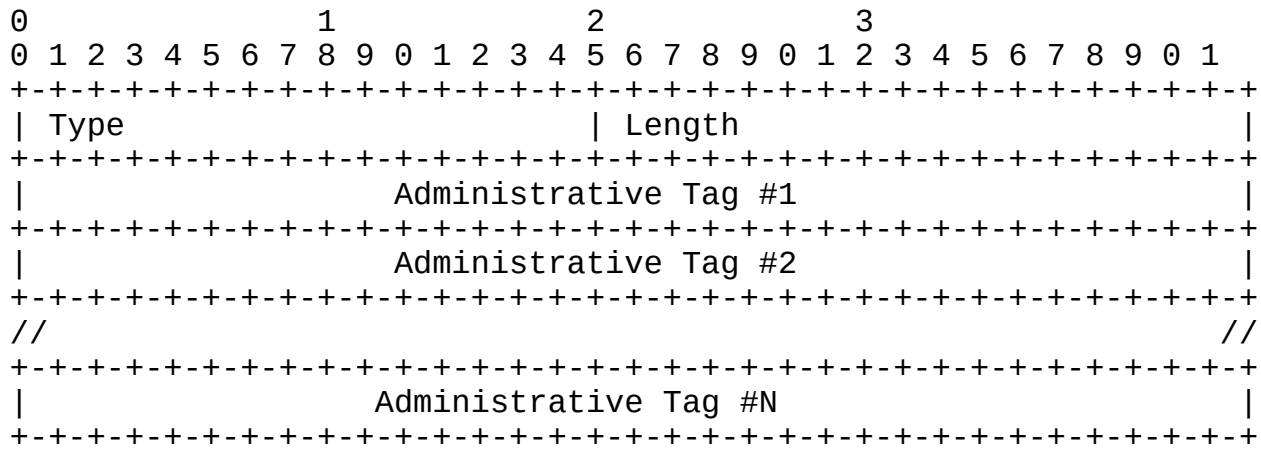

Figure 1: OSPF per-node Administrative Tag TLV

Type : TBA, Suggested value 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 sequence of multiple 4 octets defining the administrative tags. At least one tag MUST be carried if this TLV is included in the RI-LSA.

3.2. Elements of procedure

Meaning of the Node administrative tags is generally opaque to OSPF. 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 RI LSA without any change as defined in [RFC4970].

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 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 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).

To avoid incomplete or inconsistent interpretations of the per-node administrative tags the same tag value MUST NOT be advertised by a router in RI LSAs of different scopes. The same tag MAY be advertised in multiple RI LSAs of the same scope, for example, OSPF Area Border Router (ABR) may advertise the same tag in area-scope RI LSAs in multiple areas connected to the ABR.

The per-node administrative tags are not meant to be extended by the future OSPF standards. The new OSPF extensions MUST NOT require use of per-node administrative tags or define well-known tag values. Node administrative tags are for generic use and do not require IANA registry. The future OSPF extensions requiring well known values MAY define their own data signaling tailored to the needs of the feature or MAY use capability TLV as defined in [RFC4970].

Being part of the RI LSA, the per-node administrative tag 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 OSPF domain) or reachability of routes.

4. Applications

This section lists several examples of how implementations might use the 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.

4.1. Service auto-discovery

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.

4.2. Fast-Rerouting policy

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 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) LFA calculation optimization

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.

4.3. Controlling Remote LFA tunnel termination

[RFC7490] 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 node (the router does not 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 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.

4.4. 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.

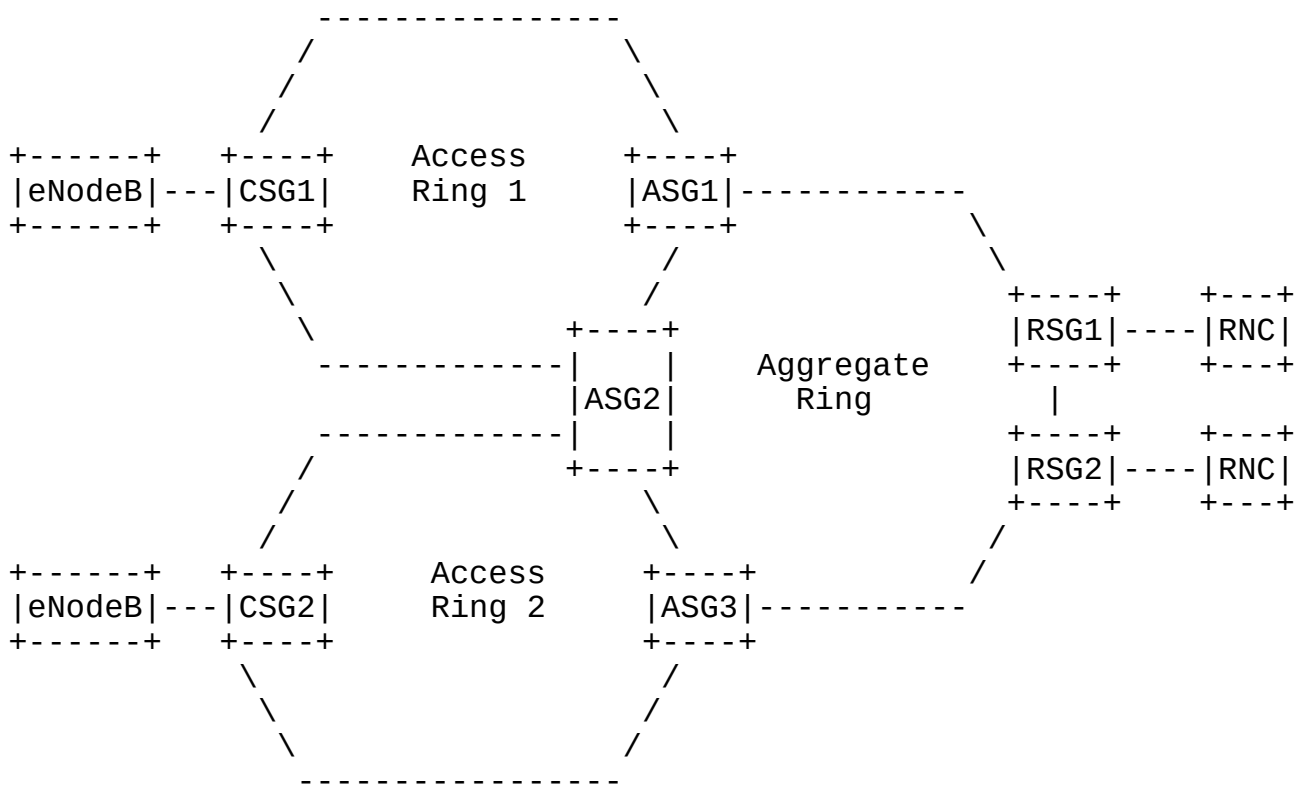

Figure 2: Mobile Backhaul Network

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 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.

4.5. Explicit routing policy

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.

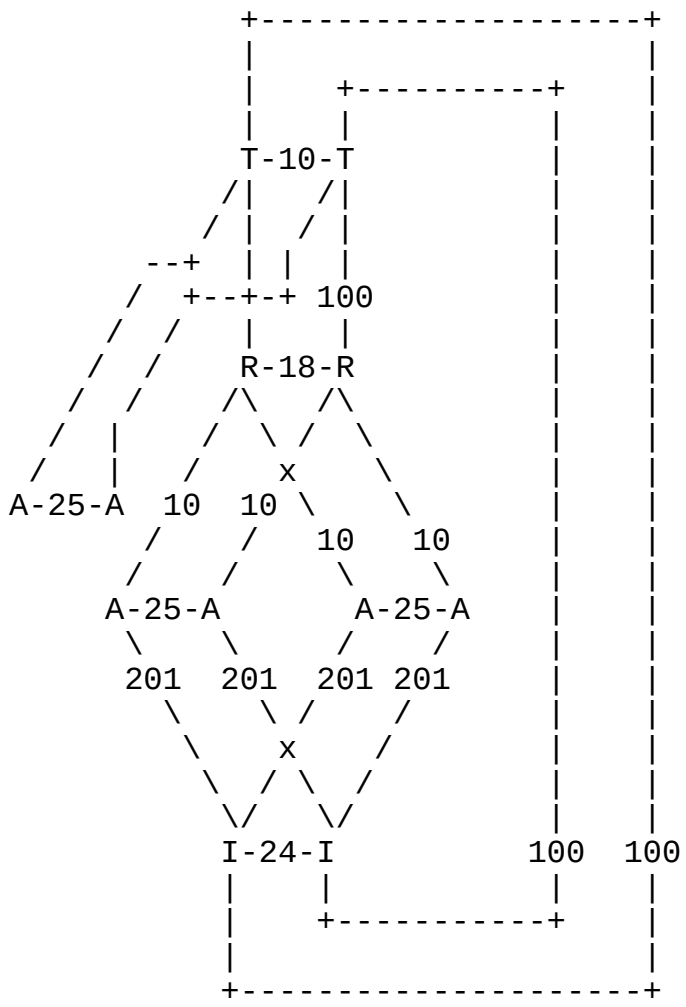

Figure 3: Explicit Routing topology

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 node admin tags, tag A (resp. I, R, T) can be configured on all A (resp. I, R, T) nodes to advertise their role. Then a generic

CSPF policy can be configured on all A nodes to enforce the above explicit routing objectives. (e.g. CSPF to destinations A exclude node with tags I).

5. Security Considerations

This document does not introduce any further security issues other than those discussed in [RFC2328] and [RFC5340].

6. IANA Considerations

This specification updates one OSPF registry: OSPF Router Information (RI) TLVs Registry

i) TBD - Node Admin tag TLV

7. Acknowledgments

Thanks to Bharath R, Pushpasis Sarakar and Dhruv Dhody for useful inputs. Thanks to Chris Bowers for providing useful inputs to remove ambiguity related to tag-ordering. Thanks to Les Ginsberg and Acee Lindem for the inputs.

8. References

8.1. Normative References

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

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.

[RFC4970] Lindem, A., Shen, N., Vasseur, JP., Aggarwal, R., and S. Shaffer, "Extensions to OSPF for Advertising Optional Router Capabilities", RFC 4970, July 2007.

[RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.

[RFC7490] Bryant, S., Filsfils, C., Previdi, S., Shand, M., and N. So, "Remote Loop-Free Alternate (LFA) Fast Reroute (FRR)", RFC 7490, April 2015.

8.2. Informative References

[I-D.acee-ospf-rfc4970bis]
Lindem, A., Shen, N., Vasseur, J., Aggarwal, R., and S. Shaffer, "Extensions to OSPF for Advertising Optional Router Capabilities", draft-acee-ospf-rfc4970bis-00 (work in progress), July 2014.

[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-08 (work in progress), March 2015.

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

Authors' Addresses

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

Email: shraddha@juniper.net

Harish Raghuveer

Email: harish.r.prabhu@gmail.com

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

Email: hannes@juniper.net

Rob Shakir
British Telecom

Email: rob.shakir@bt.com

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