Open Shortest Path First IGP
Internet-Draft
Intended status: Standards Track
Expires: December 28, 2014

S. Hegde
H. Raghuveer
H. Gredler
Juniper Networks, Inc.
R. Shakir
British Telecom
A. Smirnov
Cisco Systems, Inc.
Z. Li
Huawei Technologies

June 26, 2014

Advertising per-node administrative tags in OSPF
draft-hegde-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 the protocol extensions to disseminate per-node admin-tags to the OSPFv2 and OSPFv3 protocol.

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 December 28, 2014.

Table of Contents

1. Introduction

This document provides mechanisms to advertise per-node administrative tags in the OSPF Router Information LSA [RFC4970]. In certain path-selection applications like for example in traffic-engineering or LFA backup 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. Applicability

For the purpose of advertising per-node administrative tags within OSPF a new TLV is proposed. Because path selection is a functional

set which applies both to TE and non-TE applications, this new TLV is carried in the Router Information LSA (RI LSA) [RFC4970]

3. 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). The total number of admin tags that a given router can advertise at one time is restricted to 64. If more tags are needed in future, multi-instancing of the RI LSA [RFC4970] may be required.

4. OSPF per-node administrative tag TLV

4.1. TLV format

The format of the TLVs within the body of an RI LSA is the same as the format used by the Traffic Engineering Extensions to OSPF [RFC3630].

The LSA payload consists of one or more nested Type/Length/Value (TLV) triplets. The format of each TLV is:

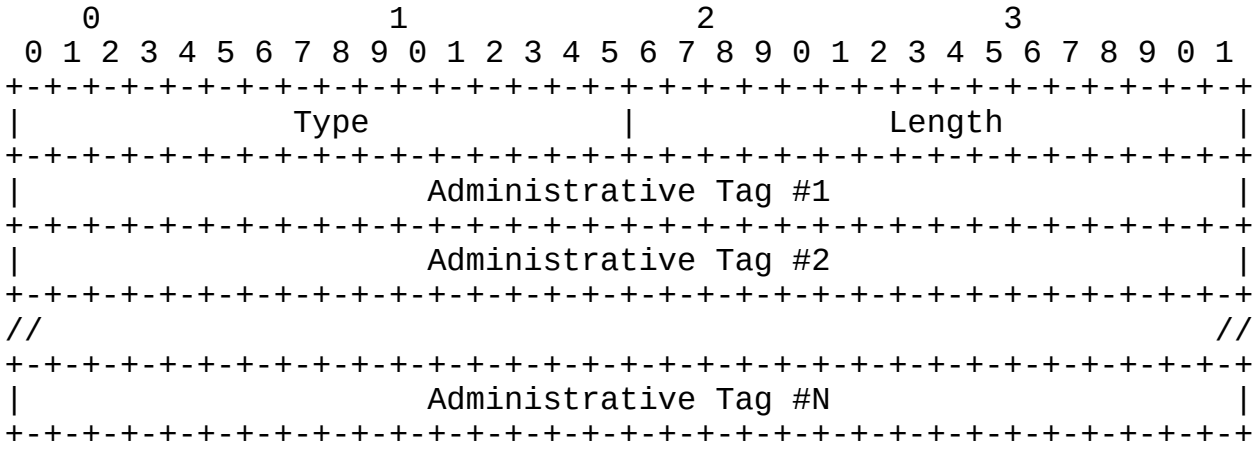

Figure 1: OSPF per-node Administrative Tag TLV

Type : TBA

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. The number of tags carried in this TLV is restricted to 64.

4.2. Elements of procedure

Meaning of the Node administrative tags is generally opaque to OSPF. 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 the tag values is implementation-specific. The meaning of a Node administrative tag is defined by the network local policy and is controlled via the configuration. There are no tag values defined by this specification.

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. Whether or not tag A precedes or succeeds tag B SHOULD not change the meaning of the tag set.

To avoid incomplete or inconsistent interpretations of the 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 Node administrative tags are not meant to be extended by the future OSPF standards. The new OSPF extensions MUST NOT require use of Node administrative tags or define well-known tag values. Instead, the future OSPF extensions must define their own data signaling tailored to the needs of the feature.

Being part of the RI LSA, the Node administrative tag TLV must be reasonably small and stable. In particular, but not limited to, implementations supporting the 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.

5. 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.

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.

2. Fast-Rerouting policy

   Increased deployment of Loop Free Alternates (LFA) as defined in [RFC5286] poses operation and management challenges.

[I-D.litkowski-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.

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 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 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.

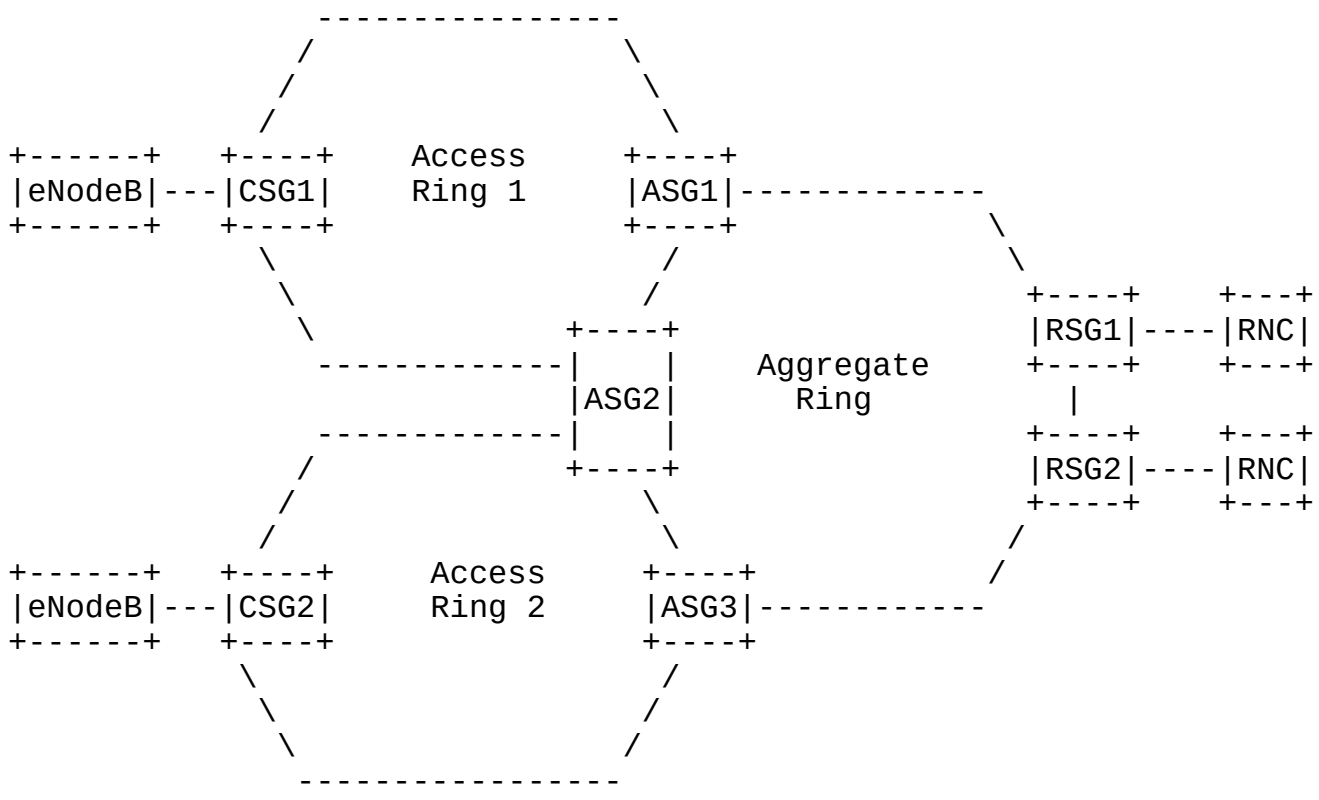

Figure 2: Mobile Backhaul Network

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 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.

6. Security Considerations

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

7. IANA Considerations

IANA maintains the registry for the TLVs. OSPF Administrative Tags will require one new type code for the TLV defined in this document.

8. Acknowledgments

Thanks to Bharath R and Pushpasis Sarakar for useful inputs. Thanks to Chris Bowers for providing useful inputs to remove ambiguity related to tag-ordering.

9. References

9.1. 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.

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

[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.

9.2. Informative References

[I-D.ietf-rtgwg-remote-lfa] Bryant, S., Filsfils, C., Previdi, S., Shand, M., and S. Ning, "Remote LFA FRR", draft-ietf-rtgwg-remote-lfa-02 (work in progress), May 2013.

[I-D.litkowski-rtgwg-lfa-manageability] Litkowski, S., Decraene, B., Filsfils, C., and K. Raza, "Operational management of Loop Free Alternates", draft-litkowski-rtgwg-lfa-manageability-01 (work in progress), February 2013.

[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
Juniper Networks, Inc.
Embassy Business Park
Bangalore 560093
India

Email: hraghuveer@juniper.net

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