Internet Engineering Task Force (IETF)
Request for Comments: 6845
Updates: 2328, 5340
Category: Standards Track
ISSN: 2070-1721
January 2013

N. Sheth
Contrail Systems
L. Wang
J. Zhang
Juniper Networks

OSPF Hybrid Broadcast and Point-to-Multipoint Interface Type

Abstract

This document describes a mechanism to model a broadcast network as a hybrid of broadcast and point-to-multipoint networks for purposes of OSPF operation. Neighbor discovery and maintenance as well as Link State Advertisement (LSA) database synchronization are performed using the broadcast model, but the network is represented using the point-to-multipoint model in the router-LSAs of the routers connected to it. This allows an accurate representation of the cost of communication between different routers on the network, while maintaining the network efficiency of broadcast operation. This approach is relatively simple and requires minimal changes to OSPF.

This document updates both OSPFv2 (RFC 2328) and OSPFv3 (RFC 5340).

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/rfc6845.

Copyright Notice

Table of Contents

1. Introduction

OSPF [RFC2328] operation on broadcast interfaces takes advantage of the broadcast capabilities of the underlying medium for doing neighbor discovery and maintenance. Further, it uses a Designated Router (DR) and Backup Designated Router (BDR) to keep the Link State Advertisement (LSA) databases of the routers on the network synchronized in an efficient manner. However, it has the limitation that a router cannot advertise different costs to each of the neighboring routers on the network in its router-LSA.

Consider a radio network that supports true broadcast, yet the metrics between different pairs of terminals could be different for various reasons (e.g., different signal strength due to placement). When running OSPF over the radio network, for a router to advertise different costs to different neighbors, the interface must be treated as point-to-multipoint (P2MP), even though the network has true broadcast capability.

Operation on point-to-multipoint interfaces could require explicit configuration of the identity of neighboring routers. It also requires the router to send separate Hellos to each neighbor on the network. Further, it mandates establishment of adjacencies to all configured or discovered neighbors on the network. However, it gives the routers the flexibility to advertise different costs to each of the neighboring routers in their router-LSAs.

This document proposes a new interface type that can be used on networks that have broadcast capability. In this mode, neighbor discovery and maintenance, as well as database synchronization are performed using existing procedures for broadcast mode. The network is modeled as a collection of point-to-point links in the router-LSA, just as it would be in point-to-multipoint mode. This new interface type is referred to as hybrid-broadcast-and-P2MP in the rest of this document.

2. 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 [RFC2119].

3. Motivation

There are some networks that are broadcast capable but have a potentially different cost associated with communication between any given pair of nodes. The cost could be based on the underlying topology as well as various link quality metrics such as bandwidth, delay, and jitter, among others.

It is not accurate to treat such networks as OSPF broadcast networks since that does not allow a router to advertise a different cost to each of the other routers. Using OSPF point-to-multipoint mode would satisfy the requirement to correctly describe the cost to reach each router. However, it would be inefficient in the sense that it would require forming O(N^2) adjacencies when there are N routers on the network.

It is advantageous to use the hybrid-broadcast-and-P2MP type for such networks. This combines the flexibility of point-to-multipoint type with the advantages and efficiencies of broadcast interface type.

4. Operation

OSPF routers supporting the capabilities described herein should have support for an additional hybrid-broadcast-and-P2MP type for the Type data item described in Section 9 of [RFC2328].

The following sub-sections describe salient aspects of OSPF operation on routers configured with a hybrid-broadcast-and-P2MP interface.

4.1. Interface Parameters

The "Router Priority" interface parameter as specified in OSPFv2 [RFC2328] and OSPFv3 [RFC5340] applies to a hybrid-broadcast-and-P2MP interface.

The "LinkLSASuppression" interface parameter as specified in OSPFv3 [RFC5340] applies to a hybrid-broadcast-and-P2MP interface. The default value is "disabled". It may be set to "enabled" via configuration.

4.2. Neighbor Data Structure

An additional field called the Neighbor Output Cost is added to the neighbor data structure. This is the cost of sending a data packet to the neighbor, expressed in the link state metric. The default value of this field is the Interface output cost. It may be set to a different value using mechanisms that are outside the scope of this document, like static per-neighbor configuration, or any dynamic discovery mechanism that is supported by the underlying network.

4.3. Neighbor Discovery and Maintenance

Routers send and receive Hellos so as to perform neighbor discovery and maintenance on the interface using the procedures specified for broadcast interfaces in [RFC2328] and [RFC5340].

4.4. Database Synchronization

Routers elect a DR and BDR for the interface and use them for initial and ongoing database synchronization using the procedures specified for broadcast interfaces in [RFC2328] and [RFC5340].

4.5. Generating Network-LSAs

Since a hybrid-broadcast-and-P2MP interface is described in router-LSAs using a collection of point-to-point links, the DR MUST NOT generate a network-LSA for the interface.

4.6. Generating Router and Intra-Area-Prefix-LSAs

Routers describe the interface in their router-LSA as specified for a point-to-multipoint interface in Section 12.4.1.4 of [RFC2328] and Section 4.4.3.2 of [RFC5340], with the following modifications for Type 1 links:

- If a router is not the DR and does not have a full adjacency to the DR, it MUST NOT add any Type 1 links.

- If a router is not the DR and has a full adjacency to the DR, and both the DR and this router agree on the DR role, it MUST add a Type 1 link corresponding to each neighbor that is in state 2-Way or higher and to which the DR’s router-LSA includes a link.

- The cost for a Type 1 link corresponding to a neighbor SHOULD be set to the value of the Neighbor Output Cost field as defined in Section 4.2.

4.6.1. Stub Links in OSPFv2 Router-LSA

Routers MUST add a Type 3 link for their own IP address to the router-LSA as described in Section 12.4.1.4 of [RFC2328]. Further, they MUST also add a Type 3 link with the Link ID set to the IP subnet address, Link Data set to the IP subnet mask, and cost equal to the configured output cost of the interface.

4.6.2. OSPFv3 Intra-Area-Prefix-LSA

Routers MUST add globally scoped IPv6 addresses on the interface to the intra-area-prefix-LSA as described for point-to-multipoint interfaces in Section 4.4.3.9 of [RFC5340]. In addition, they MUST also add all globally scoped IPv6 prefixes on the interface to the LSA by specifying the PrefixLength, PrefixOptions, and Address Prefix fields. The Metric field for each of these prefixes is set to the configured output cost of the interface.

The DR MUST NOT generate an intra-area-prefix-LSA for the transit network for this interface since it does not generate a network-LSA for the interface. Note that the global prefixes associated with the interface are advertised in the intra-area-prefix-LSA for the router as described above.

4.7. Next-Hop Calculation

Next-hops to destinations that are directly connected to a router via the interface are calculated as specified for a point-to-multipoint interface in Section 16.1.1 of [RFC2328].

4.8. Graceful Restart

The following modifications to the procedures defined in Section 2.2, item 1, of [RFC3623] are required in order to ensure that the router correctly exits graceful restart.

- o If a router is the DR on the interface, the pre-restart network-LSA for the interface MUST NOT be used to determine the previous set of adjacencies.

- o If a router is in state DROther on the interface, an adjacency to a non-DR or non-BDR neighbor is considered as reestablished when the neighbor state reaches 2-Way.

5. Compatibility Considerations

All routers on the network must support the hybrid-broadcast-and-P2MP interface type for successful operation. Otherwise, the interface should be configured as a standard broadcast interface.

If some routers on the network treat the interface as broadcast and others as hybrid-broadcast-and-P2MP, neighbors and adjacencies will still get formed as for a broadcast interface. However, due to the differences in how router and network-LSAs are built for these two interface types, there will be no traffic traversing certain pairs of routers. Note that this will not cause any persistent loops or black-holing of traffic.

To detect and flag possible mismatched configurations, an implementation of this specification SHOULD log a message if a network-LSA is received for a locally configured hybrid interface.

6. Scalability and Deployment Considerations

Treating a broadcast interface as hybrid-broadcast-and-P2MP results in O(N^2) links to represent the network instead of O(N), when there are N routers on the network. This will increase memory usage and have a negative impact on route calculation performance on all the routers in the area. Network designers should carefully weigh the benefits of using the new interface type against the disadvantages mentioned here.

7. Management Considerations

The following MIB variable/value should be added to the appropriate OSPFv2 and OSPFv3 MIBs ([RFC4750], [RFC5643]).

- o For ospfIfType/ospfv3IfType, a new value broadcast-P2MP-hybrid (X) for the hybrid interface type (X to be defined when the revised MIB documents are approved).

- o For ospfNbrEntry/ospfv3NbrEntry, an ospfNbrMetricValue/ ospfv3NbrMetricValue attribute for per-neighbor metrics. In case of non-hybrid interfaces, the value is the same as the interface metric.

This section is not normative.

8. Security Considerations

This document raises no new security issues for OSPF. Security considerations for the base OSPF protocol are covered in [RFC2328], [RFC5340], and [RFC6506].

9. Acknowledgements

The authors would like to thank Acee Lindem and Richard Ogier for their comments and suggestions.

10. 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.

[RFC3623] Moy, J., Pillay-Esnault, P., and A. Lindem, "Graceful OSPF Restart", RFC 3623, November 2003.

[RFC4750] Joyal, D., Galecki, P., Giacalone, S., Coltun, R., and F. Baker, "OSPF Version 2 Management Information Base", RFC 4750, December 2006.

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

[RFC5643] Joyal, D. and V. Manral, "Management Information Base for OSPFv3", RFC 5643, August 2009.

[RFC6506] Bhatia, M., Manral, V., and A. Lindem, "Supporting Authentication Trailer for OSPFv3", RFC 6506, February 2012.

Authors' Addresses

Nischal Sheth
Contrail Systems
2350 Mission College Blvd, #1140
Santa Clara, CA 95054
US

EMail: nsheth@contrailsystems.com

Lili Wang
Juniper Networks
10 Technology Park Dr.
Westford, MA 01886
US

EMail: liliw@juniper.net

Jeffrey Zhang
Juniper Networks
10 Technology Park Dr.
Westford, MA 01886
US

EMail: zzhang@juniper.net