

Network Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: September 1, 2012

N. Sheth  
L. Wang  
J. Zhang  
Juniper Networks  
March 1, 2012

**OSPF Hybrid Broadcast and P2MP Interface Type  
draft-ietf-ospf-hybrid-bcast-and-p2mp-01.txt**

**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 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.

**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 September 2, 2012.

**Copyright Notice**

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of

publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.



## Table of Contents

<a href="#">1.</a>	Introduction . . . . .	<a href="#">4</a>
<a href="#">2.</a>	Motivation . . . . .	<a href="#">5</a>
<a href="#">3.</a>	Operation . . . . .	<a href="#">6</a>
<a href="#">3.1.</a>	Interface Parameters . . . . .	<a href="#">6</a>
<a href="#">3.2.</a>	Neighbor Data Structure . . . . .	<a href="#">6</a>
<a href="#">3.3.</a>	Neighbor Discovery and Maintenance . . . . .	<a href="#">6</a>
<a href="#">3.4.</a>	Database Synchronization . . . . .	<a href="#">6</a>
<a href="#">3.5.</a>	Generating Network LSAs . . . . .	<a href="#">6</a>
<a href="#">3.6.</a>	Generating Router and Intra-Area-Prefix-LSAs . . . . .	<a href="#">7</a>
<a href="#">3.6.1.</a>	Stub Links in OSPFv2 Router LSA . . . . .	<a href="#">7</a>
<a href="#">3.6.2.</a>	OSPFv3 Intra-Area-Prefix-LSA . . . . .	<a href="#">7</a>
<a href="#">3.7.</a>	Next-Hop Calculation . . . . .	<a href="#">7</a>
<a href="#">3.8.</a>	Graceful Restart . . . . .	<a href="#">8</a>
<a href="#">4.</a>	Compatibility Considerations . . . . .	<a href="#">9</a>
<a href="#">5.</a>	Scalability and Deployment Considerations . . . . .	<a href="#">10</a>
<a href="#">6.</a>	Management Considerations . . . . .	<a href="#">11</a>
<a href="#">7.</a>	Security Considerations . . . . .	<a href="#">12</a>
<a href="#">8.</a>	IANA Considerations . . . . .	<a href="#">13</a>
<a href="#">9.</a>	Acknowledgements . . . . .	<a href="#">14</a>
<a href="#">10.</a>	Normative References . . . . .	<a href="#">15</a>
	Authors' Addresses . . . . .	<a href="#">16</a>



## **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 and Backup Designated Router to keep the 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.

Operation on point-to-multipoint interfaces could require explicit configuration of the identity of its 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 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 layer 2 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. Motivation**

There are some layer 2 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 layer 2 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.





### **3. 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.

#### **3.1. Interface Parameters**

Routers MUST support configuration of the Router Priority for the interface.

The default value of the LinkLSASuppression is "disabled". It MAY be set to "enabled" via configuration.

#### **3.2. Neighbor Data Structure**

Routers MUST support an additional field called the Neighbor Output Cost. 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 which are outside the scope of this document, like static per-neighbor configuration, or any dynamic discovery mechanism that is supported by the underlying network.

#### **3.3. 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 [\[RFC5340\]](#).

#### **3.4. 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 [\[RFC5340\]](#).

#### **3.5. 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 SHOULD NOT generate a network LSA for the interface.



### **3.6. 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:

- o If a router is not the DR, it MUST NOT add any Type 1 links if it does not have a full adjacency to the DR.
- o 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.
- o 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 3.2](#)

#### **3.6.1. 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.

#### **3.6.2. OSPFv3 Intra-Area-Prefix-LSA**

Routers MUST add global 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 global 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 SHOULD 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.

### **3.7. 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 \[RFC2328\]](#).



### **3.8. 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, it MUST NOT examine the pre-restart network LSA for the interface in order to determine the previous set of adjacencies.
- o If a router is in state DROther on the interface, it MUST consider an adjacency to non-DR and non-BDR neighbors as reestablished when the neighbor state reaches 2-Way.



#### **4. 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.





## **5. 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.

## **6. Management Considerations**

The following MIB variable/value should be added to the appropriate OSPFv2 and OSPFv3 MIBs.

- o For ospfIfType/ospfv3IfType, a new value broadcast-p2mp-hybrid (6) for the hybrid interface type.
- o For ospfNbrEntry/ospfv3NbrEntry, an ospfNbrMetricValue/ospfv3NbrMetricValue attribute for the per-neighbor metrics. In case of non-hybrid interfaces, the value is the same as the interface metric.

## **7. Security Considerations**

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

## **8. IANA Considerations**

This document has no IANA considerations.

This section should be removed by the RFC Editor to final publication.

## **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.
- [RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", [RFC 5340](#), July 2008.
- [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.
- [RFC5643] Joyal, D. and V. Manral, "Management Information Base for OSPFv3", [RFC 5643](#), August 2009.





Authors' Addresses

Nischal Sheth  
Juniper Networks  
1194 N. Mathilda Ave.  
Sunnyvale, CA 94089  
US

Email: [nsheth@juniper.net](mailto:nsheth@juniper.net)

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

Email: [liliw@juniper.net](mailto:liliw@juniper.net)

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

Email: [zzhang@juniper.net](mailto:zzhang@juniper.net)

