Flooding Over a Subset Topology
draft-ietf-ospf-subset-flood-00.txt

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

Abstract

This memo defines a method for limiting the flooding of OSPF LSAs to a configurable subset of the network topology. The following OSPF properties are maintained: (1) routers are omitted from the routing calculation until their link-state databases are synchronized and (2) links must be bidirectional before they can be used in the calculation. Backward-compatibility with unmodified OSPF routers is also provided.
1. Overview

Standard OSPF floods its LSAs over all links. This flooding logic is simple, robust, and auto-configuring. However, in highly meshed environments when many routers have a large number of neighbors, this flooding can be a burden on the router’s processing power.

For that reason, this memo suggests restricting flooding to a configured set of links. For backward-compatibility, and to enable the network operator to restore control connectivity from any location, a link is used for flooding if configured as such in *either* end. To prevent a link to be used for flooding, we use the technique from [Ref3], preventing its neighbor relationships from advancing past 2-way state.

These neighbor relationships that were artificially stopped at 2-way state, but would have advanced to Full state if Section 10.4 of [Ref1] were followed, are termed "forwarding adjacencies". We do not change the building of router-LSAs and network-LSAs, and instead report these forwarding adjacencies in a new set of LSAs, called router-addition-LSAs and network-addition-LSAs.

The standard OSPF routing calculation is then extended on an area-by-area basis to include the forwarding adjacencies, but only if both (a) all routers in the area support this memo and (b) both ends of the forwarding adjacency are reachable via the standard OSPF routing calculation.
2. Mechanisms

The descriptions of the required enhancements is split into the following sections. Section 2.1 describes how we prevent flooding on certain links by preventing their neighbor relationships from advancing past state 2-Way. These non-flooding relationships, called "forwarding adjacencies", are advertised in new LSAs as described in Section 2.2. Section 2.3 describes how these new LSAs are used in the routing calculation. The use of forwarding adjacencies requires that we perform the MTU check in the OSPF Hello protocol, as Section 2.4 explains.

2.1. Deciding to become adjacent

OSPF floods only to neighbors is state Exchange or greater. So we prevent flooding on a link by preventing the neighbor relationships on the link from advancing past 2-way, exactly as was done in [Ref3].

In particular, if Section 10.4 of [Ref1] indicates that the router should form an adjacency with a neighbor (transitioning the neighbor from 2-Way to ExStart state), the router should execute additional steps as follows:

1. If the interface type is Virtual Link, start forming the adjacency (we don’t allow you to disable flooding over virtual links).

2. If the neighbor is asking to form an adjacency (that is, we’re running the logic in Section 10.4 of [Ref1] because we have received a Database Description packet from the neighbor), start forming the adjacency. This is necessary for backward compatibility.

3. Otherwise, we’re running Section 10.4 of [Ref1] because either (i) we’ve just received a bidirectional Hello from the neighbor, (ii) there was an error in the previous Database Exchange over this link or (iii) an adjacency over an equivalent link has been lost (see Section 2.2). In this case, start forming the adjacency by transitioning the neighbor state to ExStart *only* if you have been configured to do so.

2.2. LSA origination

A router implementing the enhancements in this memo sets the FA bit in its router-LSA’s type field (Section A.1), and advertises its forwarding adjacencies in router-addition-LSAs and network-
2.2.1. Advertising forwarding adjacencies

Forwarding adjacencies, those bidirectional neighbors (neighbor state 2-Way) that would have been advertised in router-LSAs and network-LSAs had the router been configured to flood over them, are advertised instead in router-addition-LSAs and network-addition-LSAs.

The way a forwarding adjacency is advertised depends upon its associated interface type and the role that the router is playing on the associated segment.

- Neighbors that have been stopped at 2-Way state on point-to-point and point-to-multipoint interfaces are added to router-addition-LSAs as Type 1 links (point-to-point connection to another router), formatted according to Sections 12.4.1.1 and 12.4.1.4 of [Ref1].

- If the router is attached to a broadcast or NBMA segment, is not the DR, and its conversation with the DR has been limited to state 2Way, a Type 2 link (connection to a transit network) is added to a router-addition-LSA.

- If the router is the DR on an attached broadcast or NBMA segment, neighbor conversations that have been limited to state 2Way are added to network-addition-LSAs.

2.3. Modified routing calculation

If all the router-LSAs in Area A’s link-state database have the FA bit (Section A.1) set in their rtype field, then the OSPF routing calculation for Area A is modified as follows.

(1) The intra-area calculation for Area A, Section 16.1 of [Ref1], is run to determine which routers are reachable in Area A.

(2) The intra-area calculation is then rerun. However, this time when Section 16.1 of [Ref1] examines the router-LSA for router X, you must examine both the router-LSA originated by X *and* all the router-addition-LSAs that it has originated. Likewise, when 16.1 of [Ref1] examines network-LSAs for network N (defined by its Designated Router’s address), you must examine the network-LSA and also all matching network-addition-LSAs. For the
forwarding adjacencies listed in router-addition-LSAs and network-addition-LSAs, we substitute a different check for the bidirectional check in Step 2b of Section 16.1 of [Ref1]. In order to use a forwarding adjacency in the routing calculation, both router endpoints must have been found to be reachable.

2.4. MTU check

Links that are not used by the OSPF Database Exchange process are now included in the routing calculation. However, we still want links with MTU mismatches to be excluded from the routing calculation. For this reason we implement the MTU mismatch detection in OSPF’s Hello Protocol, exactly as was specified in Section 2.4 of [Ref3]. This logic prevents links with MTU mismatches from being declared bidirectional. See Section G.9 of [Ref3] for more details on MTU mismatches.

3. Backward compatibility

If the router’s neighbor requests to form a full adjacency, by sending a Database Description packet, the router must comply as long as a full adjacency is warranted according to Section 10.4 of [Ref1], and is the same backward-compatibility mechanism used by [Ref3].

Also, all routers within an OSPF area need to be capable of including forwarding adjacencies (advertised in router-additions-LSAs and network-additions-LSAs) in their routing calculations before any router in the area is allowed to. This is determined by checking to see that all router-LSAs in the area’s link-state database have the FA-bit set.

4. Notes

Note the following concerning the enhancements proposed by this memo.

- We do not recommend any particular configuration syntax. A vendor may decide to let you configure over which links to flood, or configure over which links not to flood. Or the vendor could combine with the functionality of [Ref3], and configure the Router IDs of the neighbors with which to flood (or not to flood).

- In the future, the routers may themselves choose which links to use in flooding. For example, if a distributed, stable algorithm were developed which produced a 2-connected spanning graph, that
might be used to autoconfigure the flooding links.

- If insufficient links are configured from flooding, some routers may become isolated from the flooding algorithm, and hence from the routing calculation. However, since a link’s flooding participation need only be configured in one endpoint, an operator would be able to reconfigure flooding and fix the problem remotely.

- Two Dijkstra calculations are employed by the enhanced routing calculation of this memo, the first to determine router reachability, and the second to include the forwarding adjacencies. However, since the first only deals with reachability, one does not need to perform its sorting phase.

References


A. New LSA formats

This memo requires that the router set an additional bit in its router-LSA’s rtype field (Section A.1) and that the router be capable of originating and processing two new LSAs, the router-additions-LSA (Section A.2) and the network-additions-LSA (Section A.2).
A.1 Router-LSA: rtype field

The format and building of the OSPF router-LSA remains unchanged, reflecting the router’s full adjacencies (neighbor state Full) as specified in Section 12.4.1 of [Ref1]. However, a new flag, called bit FA, is added to the rtype field of the router-LSA. A router sets this bit if and only if it is capable of using the new router-additions-LSAs and network-additions-LSAs in its routing calculations. Equivalently, bit FA is set when the router is capable of using forwarding adjacencies in the routing calculation. Setting bit FA also implies that the router is capable of handling Opaque-LSAs, as specified in [Ref2].
The rtype field
The router-additions-LSA is an area-scoped Opaque-LSA, having Opaque Type equal to TBD1. It is used to advertise forwarding adjacencies, and uses the same format as the router-LSA. The router’s collection of forwarding adjacencies can be listed in one or more router-additions-LSAs, with the Opaque ID field used to distinguish the LSAs. Rules for building the router-additions-LSA are described in Section 2.2.1 above.

The format of the router-additions-LSA is identical to the router-LSA, except for the following differences:

- Multiple router-addition-LSAs can be originated, distinguished by Opaque ID. The value of Opaque ID can be arbitrary. Note the similarity to the OSPF for IPv6 router-LSA [Ref5].
- The router-additions-LSA has no rtype field.
A.3 Network-additions-LSA

The network-additions-LSA is an area-scoped Opaque-LSA, having Opaque Type equal to TBD2. It is used by the Designated Router on a broadcast or NBMA segment to advertise its forwarding adjacencies on the segment, and uses a similar format to the network-LSA. The router’s collection of forwarding adjacencies can be listed in one or more network-additions-LSAs, with the Opaque ID field used to distinguish the LSAs. Rules for building the network-additions-LSA are described in Section 2.2.1 above.

```
+---------+-------+------+
| LS age  | Options| 10   |
+---------+-------+------+
| Opaque Type | Opaque ID |
+--------------------------+
| Advertising Router       |
+--------------------------+
| LS sequence number       |
+--------------------------+
| LS checksum | length |
+--------------------------+
| Network Address          |
+--------------------------+
| Network Mask             |
+--------------------------+
| Attached Router          |
| ...                      |
```

The format of the network-additions-LSA is identical to the network-LSA, except for the following differences:

- The IP address of the network segment is included in the body of the network-additions-LSA, in the "Network Address" field. As in standard OSPF, this is the IP address of the segment’s Designated Router.

- Multiple network-addition-LSAs can be originated, distinguished by Opaque ID. The value of Opaque ID can be arbitrary.
Security Considerations

This memo does not create any new security issues for the OSPF protocol. Security considerations for the base OSPF protocol are covered in [Ref1].

Authors’ Addresses

J. Moy
Sycamore Networks, Inc.
150 Apollo Drive
Chelmsford, MA 01824
Phone: (978) 367-2505
Fax: (978) 256-4203
email: jmoy@sycamorenet.com