

Network Working Group
Internet Draft
Expiration Date: August 2001
File name: [draft-ietf-ospf-ppp-flood-01.txt](#)

J. Moy
Sycamore Networks, Inc.
February 2001

Flooding over parallel point-to-point links

[draft-ietf-ospf-ppp-flood-01.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

The OSPF routing protocol synchronizes its link-state database over all links. However, when multiple point-to-point links connect a pair of OSPF routers, it is only necessary to flood over one of the parallel links. This can be done in a backward-compatible fashion, without requiring negotiation between neighboring routers, as described in this memo.

Table of Contents

1	Overview	2
2	Implementation	3
2.1	Whether to become adjacent	3
2.2	Lost adjacencies	4
2.3	Originating router-LSAs	4
2.4	Next hop calculation	5
2.5	MTU check	5
3	Backward compatibility	5
4	Example	6
5	Notes	6
	References	7
	Security Considerations	7
	Authors' Addresses	8

[1](#). Overview

When multiple "equivalent" links connect a pair of OSPF routers, database synchronization (both initial via the Database Exchange process and ongoing via flooding, also called adjacency formation and maintenance) need only be performed over one of the links. The key reason for this is that remote routers only care that at least one link is advertised in the two routers' router-LSAs; advertisement of additional links is redundant.

The definition of "equivalent" links is as follows. A set of links are equivalent if they (a) are all point-to-point links, (b) all connect the same pair of OSPF routers, and (c) all belong to the same OSPF area.

The organization of this memo is as follows. [Section 2](#) describes the implementation in detail. In a nutshell, the changes required to implement the reduction in adjacencies are: ([Section 2.1](#)) The router with the higher OSPF router ID chooses which of the equivalent links to form adjacencies over; the remaining equivalent links stay in state 2-Way. ([Section 2.2](#)) When an existing adjacency is lost, the router with the higher Router ID forms an adjacency over one of the other equivalent links. ([Section 2.3](#)) Router-LSAs advertise at most one Type 1 router link (point-to-point connection to another router) for the entire collection of equivalent links, with the advertised cost equal to the smallest cost of any of the 2-Way links. ([Section 2.4](#)) The routing calculation in the routers at either end of the equivalent links is modified to include the least cost 2-Way links as next hops. ([Section 2.5](#)) The MTU check is performed as part of Hello processing, since it is now required on 2-Way links as well as adjacencies.

[Section 3](#) addresses backward compatibility with implementations of the OSPF specification [[Ref1](#)]. A simple example of the adjacency reduction is given in [Section 4](#). Additional information concerning the adjacency reduction, including anomalies and possible enhancements, are provided in [Section 5](#).

2. Implementation

This section discusses the implementation of the adjacency reduction in detail, identifying the sections of the base OSPF protocol [[Ref1](#)] which must be modified.

2.1. Whether to become adjacent

The decision as to whether to become adjacent with a neighbor is covered by [Section 10.4](#), "Whether to become adjacent", of the OSPF specification [[Ref1](#)]. That section must be modified to implement the following idea: "When there are multiple equivalent links attaching a pair of OSPF Routers, the Router with the higher OSPF Router ID decides which links will form adjacencies".

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 other than point-to-point, start forming the adjacency.
- (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:
 - (a) If the router has a smaller Router ID than the neighbor, leave the neighbor in state 2-Way. The neighbor will decide over which of the equivalent links adjacencies should form.

- (b) If the router's Router ID is larger, examine all the equivalent links to the neighbor. If one or more of them are adjacent (neighbor state Full) or are in the process of becoming adjacent (neighbor state greater than or equal to ExStart) leave the neighbor state on the current link in state 2-Way. Otherwise, start forming the adjacency by transitioning the neighbor state to ExStart.

2.2. Lost adjacencies

If the router with the higher OSPF Router ID notices that the single adjacency in a collection of equivalent links has gone down, it must replace it by forming an adjacency on one of the equivalent links.

To be more precise, Section 10.3 of [\[Ref1\]](#) must be modified as follows. If a neighbor in state ExStart or greater transitions to a state of 2-Way or lower, and (a) the router has a larger OSPF Router ID than the neighbor, (b) the link associated with the failed adjacency is one of a collection of equivalent links, and (c) none of the other equivalent links are in state ExStart or greater, then the router must start forming an adjacency on one of the equivalent 2-Way links (if any) by transitioning that link's neighbor's state to ExStart, which starts the Database Exchange process on that link.

2.3. Originating router-LSAs

Section 12.4.1.1 of [\[Ref1\]](#), "Describing point-to-point interfaces in the router-LSA", is changed as follows. If one or more of the equivalent links is fully adjacent (neighbor state Full), a single Type 1 link (point-to-point connection to another router) is added to the router-LSA. The advertised metric is set equal to the smallest cost of any of the equivalent links which are in state 2-Way or greater. In this way, in addition to the main benefit of reducing flooding traffic, this memo also reduces the size of the router-LSA by suppressing redundant link advertisements.

Type 3 links (connection to stub networks) continue to be added to the router-LSA as specified in Section 12.4.1.1 of [\[Ref1\]](#). Up to one of these links will be added for each of the equivalent links.

Now, in addition to the events listed in Section 12.4 of [\[Ref1\]](#), the transition of a point-to-point link to/from neighbor state 2-Way can cause a router-LSA to be reoriginated. Such a state

transition may change the cost that is advertised for the equivalent links' Type 1 link.

2.4. Next hop calculation

We must change routing calculation in the routers at the end of the equivalent links, allowing 2-Way interfaces to be installed as next hops as long as at least one equivalent link is fully adjacent (neighbor state Full).

To this effect, Section 16.1.1 of [Ref1] is changed as follows. When installing a next hop to a directly connected router, through a point-to-point interface, all least cost equivalent links to the neighbor in state 2-Way or greater should be added as equal-cost next hops.

Even if it doesn't cause the contents of the link-state database to change, the transition of a point-to-point link to/from neighbor state 2-Way may change the next hops of routing table entries, forcing rerunning of the routing calculation.

2.5. MTU check

Since you are now adding certain 2-way, but non-adjacent, links as next hops in the routing table entries ([Section 2.4](#)), the MTU mismatch detection must be implemented in OSPF Hello packets sent over point-to-point links. To this end, Hello packets sent over point-to-point links (Section 9.5 of [Ref1]) have their Designated Router field set to the MTU of the point-to-point interface. Upon receiving an Hello on a point-to-point interface (Section 10.5 of [Ref1]), the new MTU field is examined. If it is greater than the interface's MTU, the Hello is discarded, preventing the neighbor relationship from forming and the interface from being installed as a next hop in the routing table (see Section G.9 of [Ref3] for more details on MTU mismatches).

3. Backward compatibility

This memo is backward compatible with implementations of the OSPF specification in [Ref1]. No negotiation between neighbors is required. If the neighbor runs [Ref1] but not the enhancements in this memo, adjacencies will form over all links, because of Step 2 in [Section 2.1](#).

4. Example

Suppose there are six point-to-point links connecting Routers A and B. Router A has the higher OSPF Router ID. The first two links (IfIndex 1 and 2 on the Router A end) belong to Area 0.0.0.0. The last four (IfIndexes 3-6 in Router A) belong to Area 0.0.0.1. There are then two sets of equivalent links, one for each area.

In all cases, OSPF Hellos will always be sent over all links. Assuming the links are all operational, they will all attain a neighbor state of 2-Way.

There are then three cases of interest.

Case 1:

A and B running enhancements defined in this memo. In this case, B will let A choose one link in each area over which to form an adjacency. Suppose these are the links corresponding to IfIndexes 1 and 3. If the link corresponding to IfIndex 3 later fails, A will choose a different link (say IfIndex 4) over which to form an adjacency. Suppose that IfIndexes 5 and 6 have been assigned the smallest costs, each with cost 10. As long as IfIndexes 5 and 6 are bidirectional (in neighbor state 2-Way or greater), A's router-LSA for area 0.0.0.1 will include a single Type 1 link to B with cost 10, and the outgoing interfaces for routing table entries through B will be the combination of IfIndexes 5 and 6. This will be true both before and after the failure of IfIndex 3.

Case 2:

Only A runs the enhancements in this memo. A will receive requests to form adjacencies on all links (that is, Database Description packets from B) and will cooperate by establishing adjacencies over all links.

Case 3:

Only B runs the enhancements in this memo. The mirror image of Case 2; adjacencies again form over all links.

5. Notes

Here is additional information on the enhancements provided by this memo.

- (1) The biggest code change required by this memo is to base the decision to form an adjacency on whether a Database Description packet has just been seen from the neighbor (Step 2 of [Section 2.1](#)). However, this distinction is useful for

other reasons; for example, in rate-limiting the number of concurrent Database Exchange sessions (see Section 8.3 of [\[Ref2\]](#)).

- (2) Why not include Point-to-MultiPoint links in the equivalent links definition? Because they can't be excluded from the router-LSA, as they are necessary for the next hop calculation.
- (3) When the single adjacency goes down, packets will not be forwarded between the neighbors until a new adjacency is formed. To get around this problem, you can introduce a new parameter, NumFloodingLinks, and require that that many adjacencies be formed within each set of equivalent links. This is equivalent to OSPF's Backup Designated Router on broadcast subnets.
- (4) Whenever you are limiting the number of adjacencies, you should timeout adjacencies that are not progressing towards Full state. See Section 8.3 of [\[Ref2\]](#) for details.
- (5) If a router running the enhancements in this memo restarts, and chooses not to form an adjacency over a given point-to-point link, its neighbor may mistakenly believe that an adjacency still exists: there may have been an adjacency before the restart, and either the router did not send an empty Hello Packet out the interface after restart, or the Hello was dropped for some reason. The router will eventually notice its neighbor's confusion when the neighbor sends a Link State Update packet over the former adjacency. At this time the router should tell the neighbor that the adjacency no longer exists by responding with an empty Hello Packet.

References

[Ref1] Moy, J., "OSPF Version 2", [RFC 2328](#), April 1998.

[Ref2] Moy, J., "OSPF Complete Implementation", Addison-Wesley, October 2000.

[Ref3] Moy, J., "OSPF Version 2", [RFC 2178](#), July 1997.

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