OSPF
Internet Draft
Intended status: Internet Draft February 18, 2008
Expires: August 2008

R. Perlman
Sun microsystems
A. Roy
Cisco Systems
M. Thomas
University of Essex

Automatic Method for Minimization of Extraneous Pseudonodes in OSPF

draft-ospf-pseudo-perlman-roy-thomas-00.txt

Status of this Memo

Copyright Notice

Abstract

An automatic method is recommended for a link state protocol to automatically decide whether an Ethernet link should be represented by a pseudonode or not. Included is encoding recommendations for IS-IS and recommendations for OSPF for implementing this feature.

Conventions used in this document

In examples, "C:" and "S:" indicate lines sent by the client and server respectively.

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

Table of Contents

1. Introduction

When Ethernets were originally introduced as a type of link, the assumption was that an "Ethernet" was assumed to be a large cloud with many directly connected nodes. Given that the overhead of a link state protocol (such as IS-IS or OSPF) is proportional to the number of links in the topology, a cloud with a lot of neighbors, represented in the most straightforward way, as having all nodes on the cloud being fully connected, would represent n^2 links. Thus to make large shared links scalable, the concept of a "pseudonode" (Network-LSA) was introduced. If there are n neighbors on an Ethernet, rather than representing the connectivity between the neighbors as n^2 links, one router on the Ethernet is elected Designated Router, and that router impersonates the cloud itself, issuing link state information on behalf of the cloud, and each router on the cloud (in addition to the Designated Router) claims only the pseudonode as a neighbor. The "pseudonode" is the cloud itself. In this way the pseudonode will have n router neighbors, but each router will only report a single neighbor (the pseudonode).

However, Ethernet has evolved into "switched Ethernet", where instead of Ethernet being a large cloud, it is often as small as a single pt-to-pt link connecting exactly two neighbors. An Ethernet might even be a pt-to-pt link between a router and an endnode. It would be wasteful to represent such a link as a pseudonode.

This paper presents a method that allows a Designated Router running OSPF to decide when it is appropriate to represent an Ethernet as a pseudonode, and generate a Network-LSA and when it is appropriate instead to represent the topology as direct connections between all the nodes on the cloud. This is automatic and non disruptive to the rest of the network if routers on the cloud fail and reappear.

1.1. Method

To implement this technique, each link state router on a single broadcast segment would set a bit indicating it supports the technique as described in this document

If every router on the link claims the ability to represent the broadcast link without a pseudonode, then the Designated Router specifies to the routers on the link whether the link will be represented by a pseudonode or not. The actual algorithm chosen by the DR need not be the same for all routers. A router MAY be configured with a different algorithm, or experience may show that a different algorithm might be more optimal. Since the DR decides on behalf of the link, whether it will be represented with a pseudonode,

or as fully connected direct links between each pair of routers, there is no need for all routers to have the same algorithm. The only requirement is that a router follow the mandate of the DR.

The algorithm in this document is that the DR mandates use of a pseudonode if there are at least 4 routers on the link (including the DR), and mandates no pseudonode if there are only 2 routers on the link (including the DR). If there are 3 routers on the link (including the DR), the DR does not change the state of the link. So if there are 4 routers on the link, the link will be represented by a pseudonode. If one of those routers go down, the link will remain represented by a pseudonode. Only when two of the four routers go down, so that the DR has only one neighbor, will the DR revert to "no pseudonode". When the DR has specified that the link is not to be represented by a pseuonode, the link will remain as "no pseudonode" until there are again 4 or more routers on the link.

2. Generic changes

1. The Hello message on a broadcast link must contain a flag specifying the ability of the transmitting router to do pseudonode minimization.

2. The Hello message on a broadcast link must contain a flag set by the DR to indicate that the link is not to be represented by a pseudonode. If the link is not to be represented by a pseudonode, then there is no need to give a name to the pseudonode.

3. The DR must check the "pseudonode minimization capable" of all routers on the link. If any router does not advertise this ability, the DR MUST represent the link with a pseudonode.

4. If the link is not to be represented by a pseudonode, then all routers on the link will report direct pt-to-pt / p2mp links about each other router on the link. The DR remains in place and flooding continues on the network itself as normal.

5. If all routers on the link are "pseudonode minimization capable", then the DR specifies that the link will not use a pseudonode if the number of routers on the link is 2 or fewer (i.e. the DR has at most one router neighbor).

6. If there are at least 4 routers on the link (i.e., the DR has at least 3 router neighbors), then the DR specifies use of a pseudonode for the link.

7. If there are 3 routers on the link, then the DR does not change the state of the link.

8. If a non compatible router appears on the link then the DR signals to the other routers to no longer perform pseudonode minimization

3. OSPF Considerations and process

3.1. Extended Options TLV for OSPF V2 and V3

In order to indicate compliance with this draft, a router needs to also be compliant with link local signaling [zinin2007]. This draft provides for TLV extensions incorporated into an LLS data block at the end of the Hello packet. [zinin2007] recommends the registration of a particular TLV called the EO-TLV (Extended Options TLV). See Appendix A.1.1 It is our recommendation in this draft that we register bit 3 (NS-Bit) and bit 4 (SS-Bit), for use within this Extended options field.

3.2. Initial DR Election

It is envisioned that this will stay the same except for the following setting of two newly allocated flags in the EO-TLV field.

1. The NS-Bit; Network-LSA suppression capability supported bit, is set in the Extended Options field (EO-TLV) using Link Local signaling [Zinin2007]. This EO-TLV field is added to the Broadcast Hello. Setting of the NS-Bit implies compliance with this pseudonode reduction draft

2. The SS-Bit; Network-LSA suppression signal, is set in the EO-TLV field to allow the DR to indicate to the routers on the Broadcast LAN that network-LSA reduction is required.

Hellos are sent as usual with DR election taking place. If all of the routers on the LAN signify via NS-Bit that they are capable of this feature, AND if the DR decides to implement network-LSA suppression, then the DR signals this to the routers on the Ethernet via the setting of the SS-Bit. If the other routers all respond with the SS-Bit set, then the router LSA migration starts.

3.2.1. Changing link types to P2P/P2MP

The other routers must follow the DR and reset the network type advertised in their router LSAs to P2P for each link (as when using P2MP).

3.3. Overview of OSPF States

State1: OSPF has elected a DR. The DR noted that the NS-Bit was set on all neighbors indicating compliance with this draft. A network LSA has NOT yet been generated or advertised by the Designated Router.

State2: If required, the DR signals network-LSA suppression by setting the SS-Bit. If ALL of the routers respond to the DR by setting the SS-Bit, then state 3.

State3:

- All routers attached to the network advertise the network type in their router LSAs as multiple P2P connections(as in P2MP) and not broadcast. (This ensures that remote routers are not expecting a network-LSA. This allows the feature to be compatible with non network-LSA reduction capable routers.)

- The DR suppresses advertising a network-LSA for the network.

- Despite remotely representing the Ethernet as a series of P2P connections, the network itself still operates as a broadcast Ethernet. The DR and BDR continue to control the flooding process and the OSPF hellos on the network remain unchanged.

3.3.1. A router changes advertised support for Network-LSA reduction

Any changes to the advertised support for Network-LSA reduction /EO-TLV fields should be handled in the same way as changes to the options field [OSPFV2] section 10.2

3.3.2. Addition of a router that supports network-LSA reduction

If a further router is added to the network and indicates compliance with network-LSA reduction by the setting of the NS-Bit in the hello packets, then the DR responds with the SS-Bit set as expected. If the new neighbor responds with the SS-Bit also set then operations continue as in state 3. No timer is required, with the router advertising the multiple P2P connections in the network type in the router LSA.

3.3.3.

If a new router is added to a network that is implementing network-LSA reduction and does NOT support this feature then:

The DR signals to the other routers on the network the resetting of the SS-Bit flag. A 30 second LSA-reduction timer starts. Routers signify complicity with the change back to Broadcast type by resetting their SS-Bit in accordance with the DR. The current graph is retained for routing calculations while a new graph is built.

If after expiration of the LSA-reduction timer all routers are still advertising reset SS-Bit in agreement, then the migration back is considered as stable, and the network type is advertised simultaneously in the router LSAs as a broadcast OSPF type.

If a router has failed to reset the SS-Bit flag within the LSA-reduction timer, then neighbor state with the router is reset as per Changes in advertised options Section 10.2 [OSPFV2]

4. Security Considerations

A group of 3 routers could maliciously come up and down as a group, or a single router could pretend to be 3 routers, and cause the pseudonode state of the link to oscillate. Also, a malicious DR can oscillate the state of the link. A single router could oscillate the state of the link by advertising first that it is capable of pseudonode minimization, and then advertising that it is not capable. However, any disruption to routing by a router using any of these strategies can already be done trivially by, for instance, pretending to be highest priority and taking over as DR, and then changing identity and the pseudonode name. So the capability in this document does not make it any easier for a malicious router on a link to cause disruption.

5. IANA Considerations

The Assignment of bits 3 and 4 of the link local signaling EO-TLV.

6. Acknowledgments

We thank Mike Shand for doing the math to show that the link state database in IS-IS becomes smaller with use of a pseudonode on a link with at least 4 routers.

7. References

7.1. Normative References

[zinin2007] Zinin,A., Roy, A,. Nuguyen, L,. Friedman, B,. and Yeung, D,. "OSPF Link-local signaling draft-ietf-ospf0lls-03.txt", IETF working Group document August 2007

[OSPFV2] Moy, J., "OSPF Version 2", RFC 2328, April 1998.

[OSPFV3] Coltun, R., Ferguson, D., and J. Moy, "OSPF for IPv6", RFC 2740, December 1999.

[IANA] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", RFC 2334, October 1998.

[KEY] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

7.2. Informative References

APPENDIX A: Packet Structures

A.1. OSPF hello packet structure and flag locations

A.1.1. Extended Option LLS Data block TLV

LLS Data Block in OSPFv2 and OSPFv3 as described in [zinin2007]

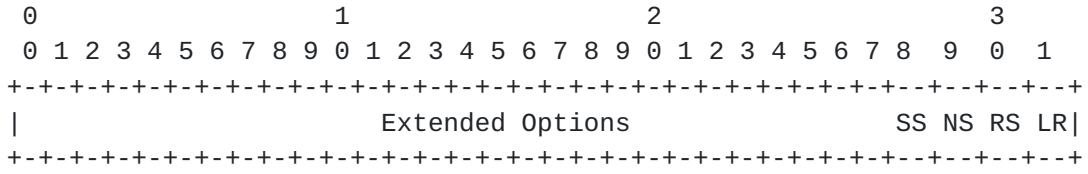

Figure X: Format of Extended Options in EO TLV

NS-Bit: Network-LSA suppression capability supported
SS-Bit: Network-LSA suppression signal. If this bit is set, the DR is requesting all of it's neighbors to use Network-LSA suppression mode on the link

Figure 2: Format of EO-TLV showing proposed Flags

Author's Addresses

Radia Perlman
Sun Microsystems

Email: radia.perlman@sun.com

Abhay Roy
Cisco Systems

E-mail: akr@cisco.com

Matthew Thomas
University of Essex

Email: mrthom@essex.ac.uk

Intellectual Property Statement

Disclaimer of Validity

Copyright Statement

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.