Network Working Group
Internet Draft
Expiration Date: February 2002
File name: draft-ietf-isis-igp-p2p-over-lan-00.txt

Naiming Shen
Acee Lindem
Jenny Yuan
Redback Networks
Alex Zinin
Nexsi Systems
Russ White
Stefano Previdi
Cisco Systems
August 2001

Point-to-point operation over LAN in link-state routing protocols

draft-ietf-isis-igp-p2p-over-lan-00.txt

Status of this Memo

Abstract

The two predominant circuit types used by link state routing protocols are point-to-point and broadcast. It is important to identify the correct circuit type when forming adjacencies, flooding link state database packets, and representing the circuit topologically. This document describes a simple mechanism to treat

the broadcast network as a point-to-point connection from the standpoint of IP routing.

1. Introduction

Point-to-point and broadcast are the two predominant circuit types used by link state routing protocols such as IS-IS [ref1] [ref2] and OSPF [ref3]. They are treated differently with respect to establishing neighbor adjacencies, flooding of link-state information, representation of the topology, SPF calculation and protocol packets. The most important differences are that broadcast circuits utilize the concept of a designated router and are represented topologically as virtual nodes in the network topology graph.

Compared with broadcast circuits, point-to-point circuits afford more straightforward IGP operation. There is no designated router involved and there is no representation of the pseudo-node or network LSA in the link state database. For ISIS, there also is no periodic database synchronization. Conversely, if there are more than two routers on the LAN media, the traditional view of the broadcast circuit will reduce the routing information in the network.

When there are only two routers on the LAN, it makes more sense to treat the connection between the two routers as a point-to-point circuit. This document describes the mechanism to allow link state routing protocols to operate using point-to-point connections over a LAN under this condition. Some implications related to forwarding IP packets on this type of circuit are also discussed. We will refer to this as a p2p-over-lan circuit in this document.

2. Motivation

Even though a broadcast circuit is meant to handle more than two devices, there are cases where only two routers are connected over either the physical or logical LAN segment:

1. The media itself is being used for point-to-point operation between two routers. This is mainly for long-haul operation.
2. There are only two routers on the physical LAN.

3. There are only two routers on a virtual LAN (vLAN).

In any of the above cases, the link state routing protocols will normally still treat the media as a broadcast circuit. Hence, they will have the overhead involved with protocol LAN operation without the benefits of reducing routing information and optimized flooding.

Being able to treat a LAN as a point-to-point circuit provides the benefit of reduction in the amount of information routing protocols must carry and manage. DR/DIS election can be omitted. Flooding can be done as in p2p links without the need of using "LSA reflection" by the DR in OSPF or periodic CSNPs in ISIS.

Also, if a broadcast segment wired as a point-to-point link can be treated as a point-to-point link, only the connection between the two routers would need to be advertised as a topological entity.

Even when there are multiple routers on the LAN an ISP may want to sub-group the routers into multiple vLANs since this allows them to assign different costs to IGP neighbors. When there are only two routers in some of the vLANs, this LAN can be viewed by the IGP as a mesh of point-to-point connections.

As a side benefit, unnumbered interface can also be applied over p2p-over-lan circuits. The advantages of unnumbered point-to-point links are obvious in the current IP addressing environment where addresses are a scarce resource. Separating the concept of network type from media type will allow LANs, e.g. ethernet, to be unnumbered and realize the IP address space savings. Another advantage is in simpler network management and configuration.

3. IP multi-access subnets

When an IP network includes multi-access segments, each segment is usually assigned a separate subnet and each router connected to it is assigned a distinct IP address within that subnet. The role of the IP address assigned to a multi-access interface can be outlined as follows:

1. Source IP address - The interface address can be used by the router as the source IP address in locally originated IP packets destined for that subnet or having a best path

next hop on that subnet.

2. Destination IP address - The interface address can be used by other devices in the network as a destination address for packets to router applications (examples include telnet, SMTP, TFTP, OSPF, BGP, etc).

3. Next-hop identifier - If other routers connected to the same segment need to forward traffic through the router, the corresponding routes in their routing tables will include the router's interface IP address. This address will be used to find the router's MAC address using the ARP protocol. Effectively, the interface IP addresses help other routers find the data-link layer details that are required to specify the destination of the encapsulating data-link frame when it is sent on the segment.

The IP addressing scheme includes an option that allows the administrators to not assign any subnets to point-to-point links (links connecting only two devices and using protocols like PPP, SLIP or HDLC for IP encapsulation). This is possible, because the routers

do not need next-hop identifiers on point-to-point links (there is only one destination for any transmission), and an interface independent IP address can be used as the source and destination. Using the unnumbered option for a point-to-point link essentially makes it a purely topological entity used only to reach other destinations.

4. Point-to-point connection over LAN media

The idea is very simple: provide a configuration mechanism to inform the IGP that the circuit is type point-to-point irrespective of the physical media type. For the IGP, this implies that it will send protocol packets with the appropriate point-to-point information and expects to receive protocol packets as they would be received on a point-to-point circuit. Over LAN media, the MAC header must contain the correct multicast MAC address to be received by the other side of the connection. For vLAN environments, the MAC header must also contain the proper vLAN ID.

In order to allow LAN links used to connect only two routers to be

treated as unnumbered point-to-point interfaces, the MAC address resolution and nexthop IP address issues need to be addressed.

4.1 Operation of IS-IS

This p2p-over-lan circuit extension for IS-IS is only concerned in pure IP routing and forwarding operation.

Since the physically circuit is a broadcast one, the IS-IS protocol packets need to have MAC addresses for this p2p-over-lan circuit. From link layer point of view, those packets are IS-IS LAN packets. The Multi-destination address including AllISs, AllL1ISs and AllL2ISs defined in [ref1] can be used for link layer encapsulation, the use of AllISs is recommended.

The circuit needs to have IP address(es) and the p2p IIH over this circuit MUST include the IP interface address(es) as defined in [ref2]. The IP address(es) can be numbered or unnumbered.

4.2 Operation of OSPF

OSPF routers supporting the capabilities described herein should support an additional interface configuration parameter specifying the interface topology type. For a LAN (i.e., broadcast capable) interface, the interface may be viewed as a point-to-point interface. Both routers on the LAN will simply join the AllSPFRouters (224.0.0.5) multicast group and send all OSPF packets to 224.0.0.5. This is identical to operation over a physical point-to-point link as described in sections 8.1 and 8.2 of [ref3].

4.3 IP forwarding and ARP

Unlike normal point-to-point IGP circuit, the IP nexthop for the routes using this p2p-over-lan circuit as an outbound interface is not optional. The IP nexthop address has to be a valid interface or internal address on the adjacent router. This address is used by local router to obtain the MAC address for IP packet forwarding. Proxy ARP has to be enabled if the address is not the adjacent interface IP address.

In the case where unnumbered IP addresses are used for p2p-over-lan circuit, the source IP address of ARP request and the target

interface IP address are usually on different subnets. The ARP should reply only if this is a p2p-over-lan circuit and the source IP address of the ARP request is the same as the neighbor's interface IP address at the other end. The neighbor's address is learned from IGP hello exchanges over this circuit.

4.4 Other MAC address resolution mechanisms

In more general cases while p2p-over-lan circuit is used as an unnumbered link, other MAC address resolution mechanisms are needed for IP packet forwarding. For example, if link-state IGP is not configured over this p2p-over-lan link, or Proxy ARP is not enabled on the circuit. The following techniques can be used to acquire the MAC address and/or the next-hop IP address of the remote device on an unnumbered point-to-point LAN link.

1. Static configuration. A router can be statically configured with the MAC address that should be used as the destination MAC address when sending data out of the interface.

2. MAC address gleaning. If a dynamic routing protocol is running between the routers connected to the link, the MAC address of the remote device can be taken from a data-link frame carrying a packet of the corresponding routing protocol.

3. ARP for reference IP address. When a point-to-point link is configured as unnumbered, the router usually associates with it a "reference IP address", that is used as the source IP address in the packets originated for the unnumbered interface. When such an address is known to a router, the router may announce its MAC address by sending a gratuitous ARP message. This solution will also help in the situations where routers calculate the next-hop addresses for the routes through point-to-point interfaces. Since the source IP address in the received routing protocol packet is used as the next-hop address in the route, forwarding an IP packet along such a route will lead to an ARP request submission on the LAN link that will be answered by the remote device.

4. Broadcast/multicast/proprietary.

4.5 Detection of mis-configuration

With this p2p-over-lan extension, the difference between a LAN and a point-to-point circuit can be made purely by configuration. It is important to implement the mechanisms for early detection of mis-configuration.

If the circuit is configured as point-to-point type and receives LAN hello packets, the router MUST discard the incoming packets; If the circuit is a LAN type and receive point-to-point hello packets, it MUST discard the incoming packets. If the system ID or the router ID of incoming hello packet does not match the system ID or the router ID of already established adjacency over this p2p-over-lan circuit, it MUST discard the packet. The implementation should offer logging and debugging information of the above events.

5. Compatibility considerations

Both routers on a LAN must support the p2p-over-lan extension and both must have the LAN segment configured as a p2p-over-lan circuit for successful operation. Both routers MAY also support one of the above listed methods for mapping ip addresses on the link to MAC address, and MUST support proxy ARP on the link. If a proprietary method of IP address to MAC address resolution is used by one router, both routers must be capable of using the same method. Otherwise, the link should be configured as a standard LAN link, with traditional IGP LAN models used.

6. Scalability and deployment considerations

There is obvious advantage to use this extension on the LANs that are connected back-to-back or only contain two routers. However, there are tradeoffs when modeling a LAN as multiple vLANs and using this extension since one does sacrifice the inherent scalability benefits of multi-access networks. In general, it will increase the link-state database size, the amount of packets flooded and the route calculation overhead. Network design engineers should carefully balance between the associated overhead. The scalability impact is less of a concern if all the vLANs are within a single OSPF area or ISIS level.

Deployment of the described technique brings noticeable benefits from the perspective of IP address usage, the network management and the router configuration. Note, however, that use of the IP unnumbered option for point-to-point LAN links inherits the same problems as those present for serial links, i.e., not being able to ping or monitor a specific interface between routers.

7. Security Issues

This document does not introduce any new security issues to ISIS or OSPF. For ARP to support unnumbered IP interface addresses, it needs to verify the p2p-over-lan circuit type described in this document and to verify the ARP packet source interface address to match the IGP adjacency interface IP address. This is due to normal ARP sanity check for common subnet can not be applied in this case.

8. Acknowledgments

The authors would like to acknowledge the following individuals: (in last name alphabetical order) Pedro Marques, Christian Martin, Danny McPherson, Ajay Patel, Tony Przygienda and Alvaro Retana.

9. References

[ref1] ISO. Information Technology - Telecommunications and Information Exchange between Systems - Intermediate System to Intermediate System Routing Exchange Protocol for Use in Conjunction with the Protocol for Providing the Connectionless-Mode Network Service. ISO, 1990.

[ref2] R. Callon. Use of OSI ISIS for Routing in TCP/IP and Dual Environments. INTERNET-RFC, Internet Engineering Task Force, December 1990.

[ref3] J. Moy. OSPF Version 2. Technical Report RFC2328 Internet Engineering Task Force, 1998.

10. Authors' Addresses

Naiming Shen
Redback Networks
350 Holger Way

San Jose, CA, 95134 USA
naiming@redback.com

Acee Lindem
Redback Networks
102 Carric Bend Court
Apex, NC 27502 USA
acee@redback.com

Jenny Yuan
Redback Networks
350 Holger Way
San Jose, CA, 95134 USA
jenny@redback.com

Alex Zinin
Nexsi Systems
1959 Concourse Drive
San Jose, CA 95131
azinin@nexsi.com

Russ White
Cisco Systems, Inc.
7025 Kit Creek Rd.
Research Triangle Park, NC 27709
e-mail: riw@cisco.com

Stefano Previdi
Cisco Systems, Inc.
De Kleetlaan 6A
1831 Diegem - Belgium
email: sprevidi@cisco.com