Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 7, 2009

P. Pillay-Esnault
Cisco Systems
P. Moyer
Juniper Networks
J. Doyle
Jeff Doyle and Associates
E. Ertekin
M. Lundberg
Booz Allen Hamilton

July 6, 2008

OSPFv3 as a PE-CE routing protocol
draft-pillay-esnault-moyer-ospfv3-pece-00

Status of This Memo

Abstract

Many Service Providers (SPs) offer the Virtual Private Network (VPN) services to their customers, using a technique in which Customer Edge (CE) routers are routing peers of Provider Edge (PE) routers. The Border Gateway Protocol (BGP) is used to distribute the customer's routes across the provider's IP backbone network, and Multiprotocol Label Switching (MPLS) is used to tunnel customer packets across the provider's backbone. This is known as a "BGP/MPLS IP VPN". Originally only IPv4 was supported and it was later extended to support IPv6 VPNs as well. Extensions were later added for the support of the Open Shortest Path First protocol version 2 (OSPFv2) as a PE-CE routing protocol for the IPv4 VPNs. This document extends those specifications to support OSPF version 3 (OSPFv3) as a PE-CE routing protocol. The OSPFv3 PE-CE functionality is identical to that of OSPFv2 except for the differences described in this document.

Table of Contents

1. Introduction

[rfc4364] offers Service Providers (SPs) a method for providing Layer-3 Virtual Private Network (VPN) services to subtending customer networks. Using the procedures defined in [rfc4364], provider edge (PE) routers separate customer VPN routing information into Virtual Routing and Forwarding (VRF) tables. The Border Gateway Protocol (BGP) is used to disseminate customer network VPN routes between PE VRFs configured in the same VPN.

Initially, the BGP/MPLS IP VPN specification enabled PE routers to learn routes within customer sites through static routing, or through a dynamic routing protocol instantiated on the PE-CE link. Specifically, [rfc4364] (and its predecessor, [rfc2547]) included support for dynamic routing protocols such as BGP, RIP, and OSPFv2. The OSPFv2 as the Provider/Customer Edge Protocol for BGP/MPLS IP Virtual Private Networks specification [rfc4577] further updates the operation of OSPFv2 as the PE-CE routing protocol by detailing additional extensions to enable intra-domain routing connectivity between OSPFv2-based customer sites.

While [rfc4364] was defined for IPv4 based networks, [rfc4659] extends the BGP/MPLS IP VPN framework to support IPv6 VPNs. This includes the capability to connect IPv6 based sites over an IPv4 or IPv6 SP backbone. It is expected that OSPFv3 will be used as the IGP for some IPv6 VPNs just as the OSPFv2 was used for IPv4 VPNs. The advantages of using OSPFv3 as a PE-CE protocol are the same as for the IPv4 VPN deployment.

This document defines the mechanisms required to enable the operation of OSPFv3 as the PE-CE Routing Protocol in BGP MPLS/IP VPNs. In doing so, it reuses, and extends where necessary, the "BGP/MPLS IP VPN" method for IPv6 VPNs defined in [rfc4659], and OSPFv2 as the PE-CE routing protocol defined in [rfc4577]. This document also includes the specifications for maintaining intra-domain routing connectivity between OSPFv3-based customer sites across a SP backbone.

We presuppose familiarity with the contents of [rfc4364], [rfc4659], [rfc4577], [rfc4576], [rfc2740] and [rfc2328].

2. Specification of Requirements

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].

3. Requirements

The benefits and considerations associated with deploying OSPFv3 as the PE-CE routing protocol are similar to those described in [rfc4577]. The requirements described in Section 3 of [rfc4577] remain semantically identical for the deployment of OSPFv3 for IPv6 VPNs.

[rfc2740] describes the modifications required to OSPF to support IPv6. In that specification, many of the fundamental mechanisms associated with OSPFv2 remain unchanged for OSPFv3. Consequently, the operation of OSPFv3 as the PE-CE routing protocol is very similar to OSPFv2 as the PE-CE protocol.

3.1. OSPFv3 Specificities

Section 2.0 of [rfc2740] describes differences between OSPFv3 and OSPFv2. Several of these changes will require modifications to the architecture described in [rfc4577]. These differences and their corresponding impact to [rfc4577] are described below:

New LSA types:

For an IPv6 MPLS/VPN architecture where customers interface to providers through OSPFv3, traditional BGP/OSPF interactions specify that VPN-IPv6 reachability information redistributed into OSPFv3 will be expressed as an AS-External OSPFv3 LSAs. Instead, it is desirable to view these LSAs as AS-internal (inter-area-prefix, and intra-area-prefix) LSAs. For the encoding of OSPFv3 LSAs, a new OSPFv3 Route Extended Community attribute is defined in Section 4.4.

Multiple instances over a link:

OSPFv3 operates on a per-link basis as opposed to OSPFv2, which operates on a per-IP-subnet basis. The support of multiple OSPFv3 protocol instances on a link changes the architecture described in [rfc4577]. [rfc4577] specifies that each interface belongs to no more than one OSPF instance. For OSPFv3, multiple instances can be established over a single interface, and associated with the same VRF. To distinguish between routes originated from different OSPFv3 instances, an Instance ID field is carried in the newly-defined OSPFv3 Route Extended Community attribute.

In addition to establishing multiple OSPFv3 instances over a single PE-CE link, multiple OSPFv3 instances can also be established across a sham link. This enables multiple OSPFv3 instances associated with a VRF to independently establish intra-

area connectivity to other OSPFv3 instances attached to a remote PE VRF. Support for multiple OSPFv3 instances across the sham link is described in Section 5.2.

4. BGP/OSPFv3 Interaction Procedures for PE Routers

4.1. VRFs and OSPFv3 Instances

The relationship between VRFs, interfaces, and OSPFv3 instances on a PE router is described in the following section.

As defined in [rfc4364], a PE router can be configured with one or more VRFs. Each VRF configured on the PE corresponds to a customer VPN, and retains the destinations that are reachable within that VPN. Each VRF may be associated with one or more interfaces, which allows multiple sites to participate in the same VPN. If OSPFv3 is instantiated on an interface associated with a VRF, the VRF will be populated with OSPFv3 routing information.

As OSPFv3 supports multiple instances on a single interface, it is therefore possible that multiple customer sites can connect to the same interface of a PE router (e.g., through a layer 2 switch) using distinct OSPFv3 instances. However, since a PE interface can be associated with only one VRF, all OSPFv3 instances running on the same interface MUST be associated with the same VRF.

Since multiple OSPFv3 instances can be associated with a single VRF, an additional mechanism is needed to demultiplex routes across these instances. When a PE supports multiple OSPFv3 instances in a VRF, a local Instance ID is assigned to the "link" that spans over the MPLS VPN backbone (PE-PE). By default, this Instance ID is set to NULL. The OSPFv3 Domain ID and local Instance ID associated with the MPLS backbone may be used to demultiplex routes for multiple instances. The detailed mechanism is described in Section 4.1.2.

4.1.1. Independent OSPFv3 Instances in PEs

Similar to [rfc4577], the PE must associate at least one OSPFv3 instance for each OSPFv3 domain to which it attaches, and each instance of OSPFv3 MUST be associated with a single VRF.

The support of multiple PE-CE OSPFv3 instances per PE interface does not change the paradigm that an OSPF instance can be associated with only a VRF. Furthermore, for each instance instantiated on the interface, the PE establishes adjacencies with corresponding CEs associated with the instance. Note that although multiple instances may populate a common VRF, they do not leak routes to one another, unless configured to do so.

4.1.2. OSPFv3 Domain and PE-PE Link Instance Identifiers

The OSPFv3 Domain ID describes the administrative domain of the OSPF instance which originated the route. It has an AS wide significance and is one of the parameters used to determine whether a VPN-IPv6 route should be translated as an Inter-area-prefix-LSA or External-LSA. Each OSPFv3 instance MUST have a primary Domain ID which is transported along with the VPN-IPv6 route in a BGP attribute over the MPLS VPN backbone. Each OSPFv3 instance may have a set of secondary Domain IDs which applies to other OSPFv3 instances within its administrative domain.

The primary Domain ID may either be configured or may be set to a value of NULL. The secondary Domain IDs are only allowed if a non-null primary Domain ID is configured. The Domain ID may be configured on a per-OSPFv3 Instance basis or per-VRF. If the Domain ID is configured on the VRF level, consequently all OSPFv3 instances associated with the VRF will share the same Domain ID.

The OSPFv3 PE-PE Link Instance ID has local significance for the PE-PE link over the MPLS VPN backbone within a VRF. This link Instance ID is used for the support of multiple OSPFv3 instances within the same VRF and it is also transported along with the VPN-IPv6 route in a BGP attribute over the MPLS VPN backbone. A PE-PE Link Instance ID is needed only if multiple OSPFv3 instances are supported, otherwise it is set to NULL. When multiple instances are associated with a VRF, each instance should have a unique PE-PE Link Instance ID.

The <Domain ID, Instance ID> tuple is used to determine whether an incoming VPN-IPv6 route belongs to the same Domain as in the receiving OSPFv3 instance. An incoming VPN-IPv6 route is said to belong to the same domain if both conditions below are met

1. The non-NULL incoming Domain ID matches either the local primary or one of the secondary Domain IDs. If the local Domain ID or incoming Domain ID is NULL, it is considered a match.

2. The non-NULL incoming Instance ID matches the local Instance ID. If the local Instance ID or incoming Instance ID is NULL, it is considered a match.

4.2. OSPFv3 Areas

Sections 4.1.4 and 4.2.3 of [rfc4577] describe the characteristics of a PE router within an OSPF domain. The mechanisms and expected behavior described in [rfc4577] are applicable to an OSPFv3 Domain.

4.3. VRFs and Routes

From the perspective of the CE, the PE appears as any other OSPFv3 neighbor. There is no requirement for the CE to support any mechanisms of IPv6 BGP/MPLS VPNs or for the CE to have any awareness of the VPNs, thereby enabling any OSPFv3 implementation to be used on a CE.

Because the export and import policies might cause different routes to be installed in different VRFs of the same OSPFv3 Domain, the MPLS VPN backbone cannot be considered as a single router from the perspective of the Domain's CEs. Rather, each CE should view its connected PE as a separate router.

The PE uses OSPFv3 to distribute routes to CEs, and MP-BGP [rfc2858] to distribute VPN-IPv6 routes to other (remote) PE routers as defined in [rfc4659]. An IPv6 prefix installed in the VRF by OSPFv3 is changed to a VPN-IPv6 prefix by the addition of an 8-octet Route Distinguisher (RD) as discussed in Section 2 of [rfc4659]. This VPN-IPv6 route can then be redistributed into MP-BGP according to an export policy that adds a Route Target Extended Communities (RT) attribute to the NLRI [rfc4360]. An IPv6 Address Specific BGP Extended Communities attribute as described in [BGP-EXTCOMM-IPV6] may also be attached to the route.

Domain IDs and Instance IDs are used to distinguish between OSPFv3 instances. When an OSPFv3-distributed route is redistributed into MP-BGP, the Domain ID, OSPFv3 Router ID, Area, OSPFv3 Route Type, External Route Type, and Intance ID are also carried in an attribute of the MP-BGP route.

A PE receiving a VPN-IPv6 NLRI from MP-BGP uses an import policy to determine, based on the RT, whether the route is eligible to be installed in one of its local VRFs. The BGP decision process selects which of the eligible routes are to be installed in the associated VRF, and the selected set of VPN-IPv6 routes are converted into IPv6 routes by removing the RD before installation.

An IPv6 route learned from MP-BGP and installed in a VRF might or might not be redistributed into OSPFv3, depending on the local configuration. For example, the PE might be configured to advertise only a default route to CEs of a particular OSPFv3 instance. Further, if the route is to be redistributed into multiple OSPFv3 instances, the route might be advertised using different LSA types in different instances.

If an IPv6 route learned from MP-BGP is to be redistributed into a particular OSPFv3 instance, the OSPFv3 Route Extended Community

attribute (Section 4.4) of the VPN-IPv6 route is used to determine whether the OSPFv3 instance from which the route was learned is the same as the OSPFv3 instance into which the route is to be redistributed.

4.3.1. OSPFv3 Routes on PE

VRFs may be populated by both OSPFv3 routes from a CE or VPN-IPv6 routes from other PEs via BGP. OSPFv3 routes are installed in a VRF using the OSPFv3 decision process. As described in [rfc4577], OSPF routes installed in a VRF may be redistributed into BGP and disseminated to other PEs participating in the VPN. At these remote PEs, the VPN-IPv6 routes may be imported into a VRF and redistributed into the OSPFv3 instance(s) associated with that VRF.

As specified in [rfc4659], routes imported and exported into a VRF are controlled by the Route Target (RT) Extended Communities attribute. OSPFv3 routes that are redistributed into BGP are given a RT that corresponds to the VRF. This RT is examined at remote PEs. In order to import a route, a VRF must have a RT that is identical to the routes RT. For routes which are eligible to be imported into the VRF, the standard BGP decision process is used to choose the "best" route(s).

When a route is advertised from a CE to a PE via OSPFv3 and that route installed in the VRF associated with the CE, the route is advertised to other locally attached CEs under normal OSPFv3 procedures.

The route is also redistributed into MP-BGP to be advertised to remote PEs. The information necessary for accurate redistribution back into OSPFv3 by the remote PEs is carried in an OSPFv3 Route Extended Communities attribute (Section 4.4). The relevant local OSPFv3 information encoded into the attribute is:

The Domain ID of the local OSPFv3 process. If no Domain ID is configured, the NULL identifier is used.

The Instance ID of the PE-PE link

The Area ID of the PE-CE link.

The PE's Router ID associated with the OSPFv3 instance. The Route Type, as determined by the LSA type from which the route was learned.

4.3.2. VPN-IPv6 Routes Received from MP-BGP

When a PE receives a valid VPN-IPv6 route from MP-BGP and has identified an association with a local VRF, it must determine:

Whether a route to the corresponding IPv6 prefix is to be installed in the VRF;

Whether the installed IPv6 route is to be redistributed to one or more local OSPFv3 instances; and

What OSPFv3 LSA type is to be used to advertise the route

An IPv6 route derived from a received VPN-IPv6 route is not installed in the associated local VRF if:

The BGP decision process identifies a better route to the destination NLRI

A configured import policy prohibits the installation of the route

The PE advertises the IPv6 route learned from MP-BGP to attached CEs via OSPFv3 if:

No configured filtering prohibits redistributing the route to OSPFv3

No configured policy blocks the route in favor of a less-specific summary route

No OSPFv3 route to the same prefix exists in the VRF, as discussed in Section 4.3.2.4.

The subsequent sections discuss the advertisement of routes learned from MP-BGP, and the rules for determining what LSA types and what CEs to advertise the routes to.

When the PE sends an LSA to a CE, it sets the DN bit in the LSA to prevent looping. The DN bit is discussed in Section 4.5.1.

4.3.2.1. OSPF Inter-Area Routes

A PE advertises an IPv6 route using an Inter-Area-Prefix (type 0x2003) LSA under the following circumstances:

The OSPFv3 domain from which the IPv6 route was learned is the same (as determined by the <Domain ID, Instance ID> tuple) as the domain of the OSPFv3 instance into which it is to be

redistributed; AND

The IPv6 route was advertised to a remote PE in an Intra-Area-Prefix (type 0x2009) OR an Inter-Area-Prefix (type 0x2003) LSA.

Note that under these rules the PE represents itself as an ABR regardless of whether or not the route is being advertised into the same area number from which the remote PE learned it (that is, whether the VPN-IPv6 route carries the same or different area numbers). This insures that if an area becomes partitioned, so that two areas with the same area ID are separated by the VPN MPLS backbone connectivity is maintained through an inter-area route.

4.3.2.2. OSPF Intra-Area Route

A route is advertised from a PE to a CE as an intra-area route using an Intra-Area-Prefix (type 0x2009) LSA only when sham links are used, as described in Section 5.2. Otherwise routes are advertised as either inter-area (Section 4.3.2.1) or external (Sections 4.3.2.3) routes.

4.3.2.3. OSPF External Routes And NSSA Routes

A PE considers an IPv6 route to be external under the following circumstances:

The OSPFv3 domain from which the route was learned is different (as determined by the <Domain ID, Instance ID> tuple) from the domain of the OSPFv3 instance into which it is redistributed; OR

The OSPFv3 Domain from which the route was learned is the same as the domain of the OSPFv3 instance into which it is redistributed AND it was advertised to the remote PE in an AS-External (type 0x4005) or a Type-7 (type 0x2007, NSSA) LSA; OR

The route was not learned from an OSPFv3 instance

To determine if the learned route is from a different domain, the <Domain ID, Instance ID> tuple associated with the VPN-IPv6 route (in the route OSPFv3 Route Extended Communities attribute or attributes) is compared with the local OSPFv3 Domain ID and Instance ID, if configured. Compared Domain IDs are considered identical if:

1. All six bytes are identical; or

2. Both Domain IDs are NULL (all zeroes).

Note that if the VPN-IPv6 route does not have a Domain ID in its

attributes, or if the local OSPFv3 instance does not have a configured Domain ID, in either case the route is considered to have a NULL Domain ID.

An IPv6 route that is determined to be external might or might not be advertised to a connected CE, depending on the type of area to which the PE-CE link belongs and whether there is a configured policy restricting its advertisement.

If there are multiple external routes to the same prefix, the standard OSPFv3 decision process is used to select the "best" route.

If the external route is to be advertised and the area type of the PE/CE link is NSSA, the PE advertises the route in a Type-7 (type 0x2007) LSA; otherwise the external route is advertised in an AS-External (type 0x4005) LSA.

The DN bit of the LSA advertising the external route MUST be set, as described in Section 4.5.1.

The PE sets the metric of the advertised external IPv6 route to the same value as the MED attribute of the VPN-IPv6 route from which the IPv6 route was derived. If the VPN-IPv6 route has no associated MED attribute, a default metric value is used.

If the VPN-IPv6 route indicates a route type of 1, the PE advertises the external route with that route type; otherwise the route type of the external IPv6 route is set to 2 by default.

4.4. OSPFv3 Route Extended Communities Attribute

OSPFv3 routes from one site are translated and delivered transparently to the remote site as BGP VPN-IPv6 routes. The original OSPFv3 routes carry OSPFv3 specific information which need to be communicated to the remote PE to ensure transparency. BGP extended communities are used to carry the needed information to enable the receiving side to reconstruct a database just as in the OSPFv2 case.

All OSPFv3 routes added to the VRF routing table on a PE router are examined to create a corresponding VPN-IPv6 route in BGP. Each of the OSPFv3 routes need to carry a BGP Extended Community Attribute which contains and preserves the OSPFv3 information attached to the original OSPFv3 route.

This document defines a new BGP attribute in the proposed "IPv6 Address Specific Extended Community" registry described in Section 3 of [BGP-EXTCOMM-IPV6]. The OSPFv3 Route Extended Community Attribute

has the Sub-type value of 0x0004. It carries an OSPFv3 Domain ID, OSPFv3 Router ID, OSPFv3 Area ID OSPFv3 Route type, Options, and an OSPFv3 Instance ID field.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| 0x00          |       4       |    OSPF Domain ID             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     OSPF Domain ID (Cont.)                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     OSPF Router ID                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Area ID                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Route Type  |    Options    |OSPF InstanceID|   UNUSED      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The OSPFv3 Route Extended Community Attribute

This attribute is MANDATORY for all OSPFv3 routes in a VRF instance on a PE router. The fields of this new BGP Extended Community attribute are described in the following sections.

OSPFv3 Domain IDs field : 6 bytes

Each OSPFv3 Instance within a VRF MUST have a Domain ID. The Domain ID may be configured at the VRF level or at the OSPFv3 Instance level. The OSPFv3 Domain ID is a 6-byte number and its default value if none is configured should be NULL.

OSPFv3 Router ID field : 4 bytes

The OSPFv3 Router ID is a 32 bit number as in OSPFv2. This field is OPTIONAL and may be set to 0.

OSPFv3 Area ID : 4 bytes

The Area ID field indicates the 32-bit Area ID to which the route belongs to.

OSPFv3 Route Types : 1 byte

To accommodate OSPFv3 LSA types, the OSPF Route Type field is encoded as follows:

Route Type Code	Route Type	LSA Type	Description
0x30	Inter-area	0x2003	Inter-area-prefix-LSA
0x50	External	0x2005	AS-external-LSA
0x70	NSSA	0x2007	NSSA-LSA
0x90	Intra-area-prefix	0x2009	Intra-area-prefix-LSA

The OSPFv3 Route Type Field Encoding

OSPFv3 Options : 1 byte

The Options field indicates if the route carries a type-1 or type-2 metric. Setting the least significant bit in the field indicates that the route carries a External type-2 metric.

OSPFv3 Instance ID field : 1 byte

The OSPFv3 Instance ID field is defined to carry the OSPFv3 Instance ID which is a one-byte number. The OSPFv3 Instance ID is configured for the "link" simulated by the MPLS VPN backbone. This attribute MAY be present whether several OSPFv3 instances are defined or not. The Instance ID default value is 0.

4.5. Loop Prevention Techniques

In some topologies, it is possible for routing loops to occur due to the nature and manner of route reachability propagation. One such example is the case of a dual homed CE router connected to two PEs; those PE routers would received this information both through their CE and their peer PE. As there is transparent transport of OSPFv3 routes over the BGP/MPLS backbone, it is not possible for the PE routers to determine whether they are within a loop.

The loop scenarios in OSPFv3 topologies are identical to those in the OSPFv2 topologies described in Section 4.2.5.1 and Section 4.2.5.2 of [rfc4577]. Of the two loop preventions mechanisms described in the sections aforementioned, only the DN bit option will be supported in the OSPFv3 implementation.

4.5.1. OSPFv3 Down Bit

Section 1 and Section 3 of [rfc4576] describe the usage of the DN-bit for OSPFv2 and are applicable for OSPFv3 for inter-area-prefix LSAs, NSSA LSAs and External LSAs. Similarly, the DN-bit must be set in inter-area-prefix-LSAs, NSSA-LSAa and AS-External-LSAs, when these are originated from a PE to a CE,to prevent those prefixes from being

re-advertised into BGP.

The DN bit MUST be clear in all other LSA types. The OSPFv3 DN-bit format is described in Appendix 4.1.1 of [OSPF-IPV6-DRAFT].

4.5.2. Other Possible Loops

The mechanism described in Section 4.5.1 of this document is sufficient to prevent looping if the DN bit information attached to a prefix is preserved in the OSPF domain. As described in Section 4.2.5.3 of [rfc4576], caution must be exercised if mutual redistribution is performed on a PE causing loss of loop prevention information.

5. OSPFv3 VRF Instance Processing

5.1.   OSPFv3 VRF LSA Handling From CE

Much like [rfc4577], any LSA with the DN bit set must not be used for route calculation. The DN bit for OSPFv3 LSAs is defined in [rfc2740].

Section 4.2.6 of [rfc4577] states that a PE router must create a VPN route in BGP for "every address prefix that was installed in the VRF by one of its associated OSPFv3 instances". This holds true for OSPFv3 as well.

Each VPN-IPv6 route created by a PE will carry an OSPFv3 Route Extended Community Attribute, as defined in Section 4.4. The Domain ID, Router ID, and area ID, Route Type and options fields within this extended community correspond to the attributes defined in [rfc4577], as they convey information about an OSPFv3 route in BGP. One new addition is the Instance ID field. This field is used to encode information about the OSPFv3 instances associated with a VRF.

Note that the new OSPFv3 Route Extended Community Attribute contains all extended community attributes specified in [rfc4577] and the OSPFv3 Instance ID but packs them into one attribute.

5.2. OSPFv3 Sham Links

This section modifies the specification of OSPFv2 sham links (defined in Section 4.2.7 of [rfc4577]) to support OSPFv3. Support for OSPFv3 sham links is an OPTIONAL feature of this specification.

Sham links are used to allow two sites that have an intra-area connection between them to act as a single VPN site that is multi-homed to the backbone. Figure 1 shows the instantiation of a sham

link between two VPN sites.

```
                             (VPN backbone)
        (site-1)      <-------- sham link -------->      (site-2)
         CE1 -------- PE1 -------- P ---------- PE2 -------- CE2
          |                                                   |
          |___________________________________________________|
               <------------ backdoor link -------------->
                        (OSPF intra-area link)
```

Sham Link

Much of the operation of sham links remains semantically identical to what was previously specified. There are, however, several differences that need to be defined to ensure the proper operation of OSPFv3 sham links.

One of the primary differences between sham links for OSPFv3 and sham links as specified in [rfc4577] are for configurations where multiple OSPFv3 instances populate a VRF. It may be desirable to provide separate intra-area links between these instances over the same sham link. To achieve this, multiple OSPFv3 instances may be established across the PE-PE sham link to provide intra-area connectivity between PE-CE OSPFv3 instances.

Note that even though multiple OSPFv3 instances may be associated with a VRF, a sham link is still thought of as a relation between two VRFs.

Another modification to OSPFv2 sham links is that OSPFv3 sham links are now identified by 128-bit endpoint addresses. Since sham links end-point addresses are now 128-bits, they can no longer default to the RouterID, which is 32-bits number. Sham link endpoint addresses MUST be configured.

Sham link endpoint addresses MUST be distributed by BGP as routeable VPN IPv6 addresses whose IPv6 address prefix is 128 bits long. As specified in [rfc4577], these endpoint addresses MUST NOT be advertised by OSPFv3.

If there is a BGP route to the remote sham link endpoint address, the sham link appears to be up. Conversely, if there is no BGP route to the sham link endpoint address, the sham link appears to be down.

5.2.1. Creating A Sham link

The procedures for creating an OSPFv3 sham link are identical to those specified in Section 4.2.7.2 of [rfc4577]. Note that the creation of OSPFv3 sham links requires the configuration of both local and remote 128-bit sham link endpoint addresses. The local Sham link endpoint address associated with a VRF MAY be used by all OSPFv3 instances that are attached to that VRF. The OSPFv3 PE-PE link Instance ID is used to demultiplex multiple OSPFv3 instance protocol packets exchanged over the sham link.

5.2.2. OSPF Protocol On Sham link

Much of the operation of OSPFv3 over a sham link is semantically the same as the operation of OSPFv2 over a sham link, as described in Section 4.2.7.3 of [rfc4577]. This includes the methodology for sending and receiving OSPFv3 packets over sham links, as well as Hello/Router Dead Intervals. Furthermore, the procedures associated with the assignment of sham link metrics adhere to those set forth for OSPFv2. OSPFv3 sham links are treated as on demand circuits.

Although the operation of the OSPFv3 protocol over the sham link is the same as OSPFv2, multiple OSPFv3 instances may be instantiated across this link. By instantiating multiple instances across the sham link, distinct intra-area connections can be established between PE-PE OSPFv3 instances associated with the endpoint addresses.

For example, if two OSPFv3 instances (O1, O2) attach to a VRF V1, and on a remote PE, two other OSPFv3 instances (O3, O4) attach to a VRF V2, it may be desirable to connect, O1 and O3 with an intra-area link, and O2 and O4 with an intra-area link. This can be accomplished by instantiating two OSPFv3 instances across the sham link, which connects V1 and V2. O1 and O3 can be mapped to one of the sham link OSPFv3 instances; O2 and O4 can be mapped to the other sham link OSPFv3 instance.

One difference from Section 4.2.7.3 of [rfc4577] is the addition of Type 0x2009 intra-area-prefix LSAs, and the flooding of these LSAs across the sham link. Furthermore, where OSPFv2 sham links are advertised in Type-1 LSAs, prefixes associated with OSPFv3 sham links are advertised as OSPFv3 Type 0x2009 LSAs. This change is required based on [rfc2740], which states that loopback interfaces are advertised in intra-area-prefix LSAs.

5.2.3. OSPF Packet Forwarding On Sham Link

The rules associated with route redistribution, stated in Section 4.2.7.4 of [rfc4577], remain unchanged in this specification.

Specifically:

> If the next hop interface for a particular route is a sham link, then the PE SHOULD NOT redistribute that route into BGP as a VPN-IPv6 route.
>
> Any other route advertised in an LSA that is transmitted over a sham link MUST also be redistributed (by the PE flooding the LSA over the sham link) into BGP.

When redistributing these LSAs into BGP, they are encoded with the OSPFv3 Route Extended Community, as defined in Section 4.4 of this document.

When forwarding a packet, if the preferred route for that packet has the sham link as its next hop interface, then the packet MUST be forwarded according to the corresponding BGP route (as defined in [rfc4364] and [rfc4659]).

6. Security Considerations

The extensions described in this document are specific to the use of OSPFv3 as the PE-CE protocol and do not introduce any concerns regarding the use of BGP as transport of IPv6 reachability over the MPLS Backbone. The Security considerations for the transport of IPv6 reachability information using BGP are discussed in Section 11 of [rfc4659] and are not altered.

The new extensions defined in this document do not introduce any new security concerns other than those already defined in Section 6 of [rfc4577].

7. IANA Considerations

This document defines a new BGP attribute in the proposed "IPv6 Address Specific Extended Community" registry described in Section 3 of [BGP-EXTCOMM-IPV6]. This document makes the following assignments in the "IPv6 Address Specific Extended Community" registry.

Name	Sub-type Value
OSPFv3 Route Attributes	0x0004

The OSPFv3 specific BGP Extended Community types

8. Contributors

Joe Lapolito

9. Acknowledgments

The authors would like to thank Kelvin Upson, Seiko Okano, and Dr. Vineet Mehta for their support of this work.

This document was produced using Marshall Rose's xml2rfc tool.

10. References

10.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFC's to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[rfc2328] Moy, J., "OSPF Version 2", RFC 2328, April 1998.

[rfc2547] Rosen, E. and Y. Rehkter, "OSPF Version 2", RFC 2547, March 1999.

[rfc2740] Moy, J., Ferguson, D., and R. Coltun, "OSPF for IPv6", RFC 2740, March 1997.

[rfc2858] Bates, T., Rehkter, Y., Chandra, R., and D. Katz, "Multiprotocol Extensions for BGP-4", RFC 2858, June 2000.

[rfc4360] Sangli, S., Tappan, D., and Y. Rehkter, "BGP Extended Communities Attribute", RFC 4360, February 2006.

[rfc4364] Rosen, E. and Y. Rehkter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, February 2006.

[rfc4576] Rosen, E., Psenak, P., and P. Pillay-Esnault, "Using a Link State Advertisement (LSA) Options Bit to Prevent Looping in BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4576, June 2006.

[rfc4577] Rosen, E., Psenak, P., and P. Pillay-Esnault, "OSPF as the Provider/Customer Edge Protocol for BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4577, June 2006.

[rfc4659]   De Clercq, J., Ooms, D., Carugi, M., and F. Lefaucheur, "BGP-MPLS IP Virtual Private Network (VPN) Extension for IPv6 VPN", RFC 4659, September 2006.

10.2. Informative References

[BGP-EXTCOMM-IPV6] Rehkter, Y., "IPv6 Address Specific BGP Extended Communities Attribute", October 2008, <http://www.ietf.org/internet-drafts/draft-rekhter-v6-ext-communities-02.txt>.

[OSPF-IPV6-DRAFT]   Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", May 2008, <http://www.ietf.org/internet-drafts/draft-ietf-ospf-ospfv3-update-23.txt>.

Authors' Addresses

Padma Pillay-Esnault
Cisco Systems
510 McCarty Blvd
Milpitas, CA 95035
USA

EMail: ppe@cisco.com

Peter Moyer
Juniper Networks
1194 N Mathilda Avenue
Sunnyvale, CA 94089
USA

EMail: pete@juniper.net

Jeff Doyle
Jeff Doyle and Associates
9878 Teller Ct.
Westminster, CO 80021
USA

EMail: jdoyle@doyleassociates.net

Emre Ertekin
Booz Allen Hamilton
5220 Pacific Concourse Drive
Los Angeles, CA 90045
USA

EMail: ertekin_emre@bah.com

Michael Lundberg
Booz Allen Hamilton
35 Corporate Dr.
Burlington, MA 01803
USA

EMail: lundberg_michael@bah.com

Full Copyright Statement

Intellectual Property