Network Working Group
Request for Comments: 1745
Category: Standards Track
K. Varadhan
OARnet & ISI
S. Hares
NSFnet/Merit
Y. Rekhter
T.J. Watson Research Center, IBM Corp.
December 1994

BGP4/IDRP for IP---OSPF Interaction

Status of this Memo

Abstract

This memo defines the various criteria to be used when designing an Autonomous System Border Router (ASBR) that will run either BGP4 or IDRP for IP with other ASBRs external to the AS and OSPF as its IGP.

Table of Contents

1. Introduction

This document defines the various criteria to be used when designing an Autonomous System Border Router (ASBR) that will run BGP4 [RFC1654] or IDRP for IP [IDRP] with other ASBRs external to the AS, and OSPF [RFC1583] as its IGP.

All future references of BGP in this document will refer to BGP version 4, as defined in [RFC1654]. All reference to IDRP are references to the Inter-Domain Routing Protocol (ISO 10747) which has been defined by the IDRP for IP document [IDRP] for use in Autonomous Systems.

This document defines how the following fields in OSPF and attributes in BGP/IDRP are to be set when interfacing between BGP/IDRP and OSPF at an ASBR:

IDRP came out of the same work as BGP, and may be consider a follow on to BGP-3 and BGP-4. Most fields defined in the interaction between BGP and IDRP are named the same. Where different, the IDRP fields are shown separately.

```
        BGP/IDRP MULTI_EXIT_DISC

        BGP ORIGIN and AS_PATH/AS_SET     vs. OSPF tag
        IDRP EXT_INFO and RD_PATH/RD_SET

        BGP/IDRP NEXT_HOP                 vs. OSPF Forwarding Address

        BGP/IDRP LOCAL_PREF               vs. OSPF cost and type
```

IDRP contains RD_PATH and RD_SET fields which serves the same purpose as AS_PATH and AS_SET fields for IDRP for IP. In this document, we will use the terms PATH and SET to refer to the BGP AS_PATH and AS_SET, or the IDRP RD_PATH and RD_SET fields respectively, depending on the context of the protocol being used.

Both IDRP and BGP provide a mechanism to indicate whether the routing information was originated via an IGP, or some other means. In IDRP, if route information is originated by means other than an IGP, then the EXT_INFO attribute is present. Likewise, in BGP, if a route information is originated by means other than an IGP, then the ORIGIN attribute is set to <EGP> or <INCOMPLETE>. For the purpose of this document, we need to distinguish between the two cases:

(a) Route information was originated via an IGP,

(b) Route information was originated by some other means.

The former case is realized in IDRP by not including the EXT_INFO attribute, and in BGP by setting the BGP ORIGIN=<IGP>; The latter case is realized by including the EXT_INFO attribute in IDRP, and by setting the BGP ORIGIN=<EGP>. For the rest of the document, we will use the BGP ORIGIN=<IGP> to refer to the former scenario, and BGP ORIGIN=<EGP> to refer to the latter.

One other difference between IDRP and BGP remains. IDRP for IP identifies an autonomous system by an identifier of variable length that is syntactically identical to an NSAP address prefix, and explicitly embeds the autonomous system number [IDRP]. BGP identifies an autonomous system just by an autonomous system number. Since there is a one-to-one mapping between how an autonomous system is identified in IDRP and in BGP, in this document, we shall identify an autonomous system by its autonomous system number.

For a more general treatise on routing and route exchange problems, please refer to [ROUTE-LEAKING] and [NEXT-HOP] by Philip Almquist.

This document uses the two terms "Autonomous System" and "Routing Domain". The definitions for the two are below:

The term Autonomous System is the same as is used in the BGP RFC [RFC1267], given below:

> "The use of the term Autonomous System here stresses the fact that, even when multiple IGPs and metrics are used, the administration of an AS appears to other ASs to have a single coherent interior routing plan and presents a consistent picture of what destinations are reachable through it. From the standpoint of exterior routing, an AS can be viewed as monolithic: reachability to destinations directly connected to the AS must be equivalent from all border gateways of the AS."

The term Routing Domain was first used in [ROUTE-LEAKING] and is given below:

> "A Routing Domain is a collection of routers which coordinate their routing knowledge using a single [instance of a] routing protocol."

By definition, a Routing Domain forms a single Autonomous System, but an Autonomous System may be composed of a collection of Routing Domains.

BGP, IDRP and OSPF have the concept of a set of reachable destinations. This set is called NLRI or Network Layer Reachability Information. The set can be represented either as an IP address prefix, or an address, mask pair. Note that if the mask is contiguous in the latter, then the two representations are equivalent. In this document, we use the term "address/mask pair" in the context of OSPF, and "destination" or "set of reachable destinations" in the context of BGP or IDRP.

This document follows the conventions embodied in the Host Requirements RFCs [RFC1122, RFC1123], when using the terms "MUST", "SHOULD," and "MAY" for the various requirements.

A minimal implementation of BGP/IDRP OSPF exchange MUST not advertise a route containing a set of reachable destinations when none of the destinations in the address/mask pair is reachable via OSPF (section 2.1, bullet 3), MUST merge the PATH into a SET when multiple exit points exist within the same autonomous system for the same external destination (section 3), MUST set the OSPF tag accurately (section 4). This subset is chosen so as to cause minimal havoc to the Internet at large. It is strongly recommended that implementors implement more than a minimalistic specification.

2. Reachability Information Exchange

This section discusses the constraints that must be met to exchange the set of reachable destinations between an external BGP/IDRP peer from another AS and internal OSPF address/mask pairs.

2.1. Exporting OSPF information into BGP

1. The administrator MUST be able to selectively export address/mask pairs into BGP/IDRP via an appropriate filter mechanism.

   This filter mechanism MUST support such control with the granularity of an address/mask pair.

   This filter mechanism will be the primary method of aggregation of OSPF internal and type 1 and type 2 external routes within the AS into BGP/IDRP.

   Additionally, the administrator MUST be able to filter based on the OSPF tag and the various sub-fields of the OSPF tag. The settings of the tag and the sub-fields are defined in section 4 in more detail.

   - The default MUST be to export no routes from OSPF into BGP/IDRP. A single configuration parameter MUST permit all OSPF inter-area and intra-area address/mask pairs to be exported into BGP/IDRP.

     OSPF external address/mask pairs of type 1 and type 2 MUST never be exported into BGP/IDRP unless they are explicitly configured.

2. An address/mask pair having a non-contiguous mask MUST not be exported to BGP/IDRP.

3. When configured to export an address/mask pair from OSPF into BGP/IDRP, the ASBR MAY advertise the route containing the set of reachable destinations via BGP/IDRP as soon as at least one of the destinations in the address/mask pair is determined to be reachable via OSPF; it MUST stop advertising the route containing the set of reachable destinations when none of the destinations in the address/mask pair is reachable via OSPF.

4. The network administrator MUST be able to statically configure the BGP/IDRP attribute MULTI_EXIT_DISC attribute to be used for any route.

   - The default MUST be to omit the MULTI_EXIT_DISC in the route advertised via BGP/IDRP.

5. An implementation of BGP/IDRP and OSPF on an ASBR MUST have a mechanism to set up a minimum amount of time that must elapse between the learning of a new address/mask pair via OSPF and subsequent advertisement of the address/mask pair via BGP/IDRP to the external neighbours.

   - The default value for this setting MUST be 0, indicating that the address/mask pair is to be advertised to the neighbour BGP/IDRP peers instantly.

     Note that BGP and IDRP mandate a mechanism to dampen the inbound advertisements from adjacent neighbours. See the variable MinRouteAdvertisementInterval in section 9.2.3.1, [RFC1654] or in section 7.17.3.1, [IS10747].

6. LOCAL_PREF is not used when exporting OSPF information into BGP/IDRP, as it is not applicable.

2.2. Importing BGP/IDRP information into OSPF

1. BGP/IDRP implementations SHOULD allow an AS to control announcements of BGP/IDRP learned set of reachable destinations into OSPF. Implementations SHOULD support such control with the granularity of a single destination.

   Implementations SHOULD also support such control with the granularity of an autonomous system, where the autonomous system may be either the autonomous system that originated the information or the autonomous system that advertised the information to the local system (adjacent autonomous system).

   - The default MUST be to import nothing from BGP/IDRP into OSPF. Administrators must configure every destination they wish to import.

     A configuration parameter MAY allow an administrator to configure an ASBR to import all the set of reachable destinations from BGP/IDRP into the OSPF routing domain.

2. The administrator MUST be able to configure the OSPF cost and the OSPF metric type of every destination imported into OSPF. The OSPF metric type MUST default to type 2. If the LOCAL_PREF value is used to construct the OSPF cost, one must be extremely careful with such a conversion. In OSPF the lower cost is preferred, while in BGP/IDRP the higher value of the LOCAL_PREF is preferred. In addition, the OSPF cost ranges between 1 and 2^24, while the LOCAL_PREF value ranges between 0 and 2^32. Note that if ASBRs within a domain are configured to correlate BGP/IDRP and OSPF information (as described in Section 3), then the route selection by the ASBRs is determined solely by the OSPF cost, and the value carried by the LOCAL_PREF attribute has no impact on the route selection.

3. Information learned via BGP/IDRP from peers within the same AS MUST not be imported into OSPF.

4. The ASBR MUST never generate a default destination into the OSPF routing domain unless explicitly configured to do so.

   A default destination is a set of all possible destinations. By convention, it is represented as a prefix of 0 length or a mask of all zeroes.

   A possible criterion for generating default into an IGP is to allow the administrator to specify a set of (set of reachable destinations, PATH, default cost, default type) tuples. If the ASBR learns of at least one of the destinations in the set of reachable destinations, with the corresponding PATH, then it generates a default destination into the OSPF routing domain, with the appropriate cost and type. The lowest cost route will then be injected into the OSPF routing domain.

This is the recommended method for originating default destinations in the OSPF routing domain.

5. Note that [RFC1247] requires the network number to be used as the Link State ID. This will produce a conflict if the ASBR tries to import two destinations, differing only in their prefix length. This problem is fixed in [RFC1583], which obsoletes [RFC1247].

   An implementation conforming to the older [RFC1247] MUST, in this case, drop the more specific route, i.e. the route corresponding to the longer prefix in the reachability information.

6. MULTI_EXIT_DISC is not used to import BGP/IDRP information into OSPF, as it is not applicable.

3. BGP/IDRP Identifier and OSPF router ID

The BGP/IDRP identifier MUST be the same as the OSPF router id at all times that the router is up.

Note that [RFC1654] requires that the BGP identifier be an address assigned to the BGP speaker.

In the case of IDRP, the IDRP protocol does not explicitly carry the identity of the IDRP speaker. An implicit notion of the identity of the IDRP speaker can be obtained by examining the source address in the IP packets carrying the IDRP information. Therefore, all IDRP speakers participating in the OSPF protocol MUST bind the IDRP identifier to be the address of the OSPF router id.

This characteristic makes it convenient for the network administrator looking at an ASBR to correlate different BGP/IDRP and OSPF information based on the identifier. There is another more important reason for this characteristic.

Consider the scenario in which 3 ASBRs, RT1, RT2, and RT3, belong to the same autonomous system.

```
                                  +-----+
                                  | RT3 |
                                  +-----+
                                     |

                      Autonomous System running OSPF

                             /                \

                         +-----+          +-----+
                         | RT1 |          | RT2 |
                         +-----+          +-----+
```

Both RT1 and RT2 can reach an external destination X and import this information into the OSPF routing domain. RT3 is advertising this information about destination X to other external BGP/IDRP speakers. The following rule specifies how RT3 can generate the correct advertisement.

RT3 MUST determine which ASBR(s) it is using to reach destination X by matching the OSPF router ID for its route to destination with the BGP identifier of the ASBR(s), or the IP source address of the IDRP protocol packet from the ASBR(s).

- o   If RT3 has equal cost routes to X through RT1 and RT2, then, RT3 MUST merge the PATH through RT1 and RT2 into a SET.

- o   Otherwise, RT3 MAY merge the PATH through RT1 and RT2.

It MAY then generate the corresponding network layer reachability information for further advertisement to external BGP/IDRP peers.

4. Setting OSPF tags, ORIGIN and PATH attributes

The OSPF external route tag is a "32-bit field attached to each external route . . . It may be used to communicate information between AS boundary routers; the precise nature of such information is outside the scope of [the] specification" [RFC1583].

We use the external route tag field in OSPF to intelligently set the ORIGIN and PATH attributes in BGP/IDRP. These attributes are well-known, mandatory attributes in BGP/IDRP. The exact mechanism for setting the tags is defined in sections 4.2 and 4.3. Every combination of tag bits is described in two parts:

import This describes when an ASBR imports an AS external LSA into the OSPF domain with the given tag setting.

export This indicates how the BGP/IDRP path attribues should be formatted when an ASBR, having a given type 1 or type 2 OSPF external route in its routing table, decides to export according to the considerations in section 2.1.

The tag is broken up into sub-fields shown below. The various sub-fields specify the characteristics of the set of reachable destinations imported into the OSPF routing domain.

The high bit of the OSPF tag is known as the "Automatic" bit. Setting this bit indicates that the tag has been generated automatically by an ASBR.

When the network administrator configures the tag, this bit MUST be 0. This setting is the default tag setting, and is described in section 4.2.

When the tag is automatically generated, this bit is set to 1. The other bits are defined below:

```
 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 3 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|1|c|p l|     ArbitraryTag      |       AutonomousSystem        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

c   1 bit of Completeness information, set when the ORIGIN of the route is either <EGP> or <IGP>.

pl   2 bits of PathLength information; this field is set depending on the length of the PATH that the protocol could have carried when importing the reachability information into the OSPF routing domain.

ArbitraryTag
     12 bits of tag information, defaults to 0 but can be configured to anything else.

AutonomousSystem (or "AS")
     16 bits, indicating the AS number corresponding to the set of reachable destinations, 0 if the set of reachable destinations is to be considered as part of the local AS.

local_AS: The AS number of the local OSPF routing domain.

next_hop_AS: The AS number of an external BGP peer.

4.1. Configuration parameters for setting the OSPF tag

- o There MUST be a mechanism to enable automatic generation of the tag characteristic bits.

- o Configuration of an ASBR running OSPF MUST include the capability to associate a tag value, for the ArbitraryTag, or LocalInfo sub-field of the OSPF tag, with each instance of a routing domain.

- o Configuration of an ASBR running OSPF MUST include the capability to associate an AS number with each instance of a routing domain.

  Associating an AS number with an instance of an IGP is equivalent to flagging those set of reachable destinations imported from the IGP to be external destinations outside the local autonomous system.

4.2. Manually configured tags

```
 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 3 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|                          LocalInfo                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

import This tag setting corresponds to the administrator manually setting the OSPF tag bits.

export The route SHOULD be exported into BGP with the attributes ORIGIN=<EGP>, PATH=<local_AS>.

Nothing MUST inferred about the characteristics of the set of reachable destinations corresponding to this tag setting.

For backward compatibility with existing implementations of OSPF currently deployed in the field, this MUST be the default setting for importing destinations into the OSPF routing domain. There MUST be a mechanism to enable automatic tag generation for imported destinations.

4.3. Automatically generated tags

4.3.1. Tag = <Automatic = 1, Complete = 0, PathLength = 00>

```
 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 3 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|1|0|0|0|     ArbitraryTag      |       AutonomousSystem        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

import These are reachable destinations imported from routing protocols with incomplete path information and cannot or may not carry the neighbour AS or AS path (and hence the IDRP RD_PATH) as part of the routing information.

This setting SHOULD be used to import reachable destinations from an IGP that the network administrator has configured as external routes, without specifying the next_hop_AS.

export The route SHOULD be exported into BGP/IDRP with the attributes ORIGIN=<EGP>, PATH=<Local_AS>.

4.3.2. Tag = <Automatic = 1, Complete = 0, PathLength = 01>

```
 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 3 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|1|0|0|1|     ArbitraryTag      |       AutonomousSystem        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

import These are reachable destinations imported from routing protocols with incomplete path information. The neighbour AS (and therefore IDRP RD) is carried in the routing information.

This setting SHOULD be used for importing reachable destinations from EGP into the OSPF routing domain. This setting MAY also be used when importing reachable destinations from BGP/IDRP whose ORIGIN=<EGP> and PATH=<next_hop_AS>; if the BGP/IDRP learned route has no other transitive attributes, then its propagation via BGP/IDRP to ASBRs internal to the autonomous system MAY be suppressed.

export The route SHOULD be exported into BGP/IDRP with ORIGIN=<EGP> and PATH=<local_AS, next_hop_AS>.

4.3.3. Tag = <Automatic = 1, Complete = 0, PathLength = 10>

```
 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 3 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|1|0|1|0|     ArbitraryTag      |       AutonomousSystem        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

import These are reachable destinations imported from routing protocols with truncated path information.

These are imported by a border router, which is running BGP/IDRP to a stub domain, and not running BGP/IDRP to other ASBRs in the same autonomous system. This causes a truncation of the PATH. These destinations MUST not be re-exported into BGP/IDRP at another ASBR.

export The route MUST never be exported into BGP/IDRP.

4.3.4. Tag = <Automatic = 1, Complete = 1, PathLength = 00>

```
 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 3 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|1|1|0|0|     ArbitraryTag      |       AutonomousSystem        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

import These are reachable destinations imported from routing protocols with either complete path information or are known to be complete through means other than that carried by the routing protocol.

This setting SHOULD be used for importing reachable destinations into OSPF from an IGP.

export The route SHOULD be exported to BGP/IDRP with ORIGIN=<IGP>, PATH=<Local_AS>.

4.3.5. Tag = <Automatic = 1, Complete = 1, PathLength = 01>

```
 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 3 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|1|1|0|1|     ArbitraryTag      |       AutonomousSystem        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

import These are reachable destinations imported from routing protocols with either complete path information, or are known to be complete through means other than that carried by the routing protocol. The routing protocol also has additional information about the next hop AS or RD, the destination was learned from.

This setting SHOULD be used when the administrator explicitly associates an AS number with an instance of an IGP. This setting MAY also be used when importing reachable destinations from BGP/IDRP whose ORIGIN=<IGP> and PATH=<next_hop_AS>; if the BGP/IDRP learned route has no other transitive attributes, then its propagation via BGP/IDRP to other ASBRs internal to the autonomous system MAY be suppressed.

export OSPF routes with this tag setting SHOULD be exported with the BGP/IDRP attributes, ORIGIN=<IGP>, PATH=<local_AS, next_hop_AS>.

4.3.6. Tag = <Automatic = 1, Complete = 1, PathLength = 10>

```
 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 3 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|1|1|1|0|     ArbitraryTag      |       AutonomousSystem        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

import These are reachable destinations imported from routing protocols with complete path information and carry the AS path information as part of the routing information.

These destinations MUST not be exported into BGP/IDRP because these are destinations that are already imported from BGP/IDRP into the OSPF RD. Hence, it is assumed that the BGP/IDRP speaker will convey these routes to other BGP/IDRP speakers within the same autonomous system via BGP/IDRP. An ASBR learning of such a destination MUST wait for the BGP update from its internal neighbours before advertising it to external BGP/IDRP peers.

export These routes MUST not be exported into BGP/IDRP.

4.4. Miscellaneous tag settings

```
 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 3 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|1|x|1|1|              Reserved  for  future  use               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The value of PathLength=11 is reserved during automatic tag generation. Routers MUST NOT generate such a tag when importing reachable destinations into the OSPF routing domain. ASBRs must ignore tags which indicate a PathLength=11.

5. Setting OSPF Forwarding Address and BGP/IDRP NEXT_HOP attribute

Forwarding addresses are used to avoid extra hops between multiple routers that share a common network and that speak different routing protocols with each other on the common network.

Both BGP/IDRP and OSPF have equivalents of forwarding addresses. In BGP and IDRP, the NEXT_HOP attribute is a well-known, mandatory attribute. OSPF has a Forwarding address field. We will discuss how these are to be filled in various situations.

Consider the 4 router situation below:

RT1 and RT2 are in one autonomous system, RT3 and RT4 are in another. RT1 and RT3 are talking BGP/IDRP with each other. RT3 and RT4 are talking OSPF with each other.

```
+-----+                 +-----+
| RT1 |                 | RT2 |
+-----+                 +-----+
   |                       |            common network
---+-----------------------+-------------------------
<BGP/IDRP> |                       |
        +-----+     <OSPF>      +-----+
        | RT3 |                 | RT4 |
        +-----+                 +-----+
```

- Importing a reachable destination into OSPF:
  When importing a destination from BGP/IDRP into OSPF, RT3 MUST always fill the OSPF Forwarding Address with the BGP/IDRP NEXT_HOP attribute for the destination.

- Exporting a reachable destination into BGP:
  When exporting set of reachable destinations internal to the OSPF routing domain from OSPF to BGP/IDRP, if all the destinations in the set of reachable destinations are through RT4, then RT3 MAY fill the NEXT_HOP attribute for the set of reachable destinations with the address of RT4. This is to avoid requiring packets to take an extra hop through RT3 when traversing the AS boundary. This is similar to the concept of indirect neighbour support in EGP [RFC888, RFC827].

6. Changes from the BGP 3 - OSPF interactions document

o The use of the term "route" has attained a more complicated structure in BGP 4. This document follows the constraint in the manner shown below:

  - The term "set of reachable destinations" is called a NLRI in [RFC1654].

  - The term "route" in the BGP context refers to a set of reachable destinations, and the associated attributes for the set.

  - The term "route" in the OSPF context refers to the set of reachable destinations, and the cost and the type to reach destinations. This is to keep the definitions consistent in the document.

o The notion of exchanging reachability information between BGP 4 and OSPF has been updated to handle variable length net mask information.

o The previous term INTER_AS_METRIC in BGP 3 has now been changed to MULTI_EXIT_DISC.

o The default metric to be used for importing BGP information into the OSPF RD is now the LOCAL_PREF attribute, instead of a constant value.

o Section 3 which requires an ASBR to match the OSPF tag corresponding to a route to the BGP Identifier, can cause potential loops if the AS has equal cost multipath routing amongst the ASBRs. This scenario is outlined in the Appendix below. This is fixed in BGP4 by requiring the ASBR seeing equal cost multi-path routes to merge the PATHs through the various ASBRs into appropriate SETs.

- o BGP 4 requires that the BGP identifier be an address assigned to the BGP speaker. This is dealt with in section 3.

- o Section 5 on setting NEXT_HOP attributes and Forwarding Address field has been updated to account for variable length net information.

- o This section, 6, has been added.

7. Security Considerations

Security issues are not discussed in this memo.

8. Acknowledgements

We would like to thank Jeff Honig (Cornell University), John Moy (Cascade Communications Corp.), Tony Li (cisco Systems), Rob Coltun (Consultant), Dennis Ferguson (ANS, Inc.), Phil Almquist (Consultant), Scott Bradner (Harvard University), and Joel Halpern (Newbridge Networks Inc.) for their help and suggestions in writing this document. Cengiz Aleattinoglu (USC/ISI) and Steve Hotz (USC/ISI) provided fresh insights into the packet looping problem described in the appendix.

We would also like to thank the countless number of people from the OSPF and BGP working groups who have offered numerous suggestions and comments on the different stages of this document.

Thanks also to Bob Braden (ISI), whose suggestions on the earlier BGP-OSPF document, [RFC1403] were useful even for this one, and have been carried through.

We would also like to thank OARnet, where one of the authors did most of his work on this document, before moving to USC to resurrect his PhD.

9. Bibliography

[RFC827] Rosen, E., "Exterior Gateway Protocol (EGP)", RFC 827, BBN, October 1982.

[RFC888] Seamonson, L., and E. Rosen, "'STUB' Exterior Gateway Protocol", RFC 888, BBN, January 1984.

[RFC1058] Hedrick, C, "Routing Information Protocol", RFC 1058, Rutgers University, June 1988.

[RFC1122] Braden, R., Editor, "Requirements for Internet Hosts - Communication Layers", STD 3, RFC 1122, USC/Information Sciences Institute, October 1989.

[RFC1123] Braden, R., Editor, "Requirements for Internet Hosts - Application and Support", STD 3, RFC 1123, USC/Information Sciences Institute, October 1989.

[RFC1247] Moy, J., "The OSPF Specification Version 2", RFC 1247, Proteon, January 1991.

[RFC1403] Varadhan, K., "BGP OSPF Interaction", RFC 1403, OARnet, January 1993.

[RFC1519] Fuller, V., Li, T., Yu, J., and K. Varadhan, "Supernetting: an Address Assignment and Aggregation Strategy", RFC 1519, BARRNet, cisco, Merit, OARnet, September 1993.

[RFC1583] Moy, J., "The OSPF Specification Version 2", RFC 1583, (Obsoletes [RFC1247]), Proteon, March 1994.

[RFC1654] Rekhter, Y., and T. Li, Editors, "A Border Gateway Protocol 4 (BGP-4)", RFC 1654, T.J. Watson Research Center, IBM Corp., cisco Systems, July 1994.

[ROUTE-LEAKING] Almquist, P., "Ruminations on Route Leaking", Work in Progress.

[NEXT-HOP] Almquist, P., "Ruminations on the Next Hop, Work in Progress.

[IDRP] Hares, S., "IDRP for IP", Work in Progress.

[IS10747] ISO/IEC IS 10747 - Information Processing Systems - Telecommunications and Information Exchange between Systems - Protocol for Exchange of Inter-domain Routeing Information among Intermediate Systems to Support Forwarding of ISO 8473 PDUs, 1993.

10. Appendix

This is an example of how the two routing protocols, BGP/IDRP and OSPF, might both be consistent in their behaviour, and yet packets from a source domain, S, to a destination in domain X might be stuck in a forwarding loop.

```
                                    +--------+
                        X ----------| C1     |
                        |           |Domain C|
                        |           | C3  C2 |
                        |           +--------+
                        B             /   \
                         \           /     \
                          \         /      S
                           \       /      /
                            \     /      /
                          +--------+    /
                          | A1  A2 |   /
                          |Domain A|  /
                          |     A3 |-/
                          +--------+
```

Given the domains, X, A, B, C and S, let domains A and C be running OSPF, and ASBRs A3 and C3 have equal cost multipath routes to A1, A2 and C1, C2 respectively. The picture above shows the internal structure of domains A and C only.

During steady state, the following are the route advertisements:

- o Domain B advertises to A path <B,X>

- o ASBR A3 in domain A advertises path <A,B,X> to domain C, at ASBR C2.

- o Domain C has two equal cost paths to X: one direct <C,X>, and another through A <C,A,B,X>

- o BR C3 in domain C advertises to A2 path <C,X>

- o Domain A has two equal cost paths to X: <A,C,X> and <A,B,X>

Both C1 and C2 inject a route to X within the domain C, and likewise A1 and A2 inject a route to X within the domain A. Since A3 and C3 see equal cost routes to X through A1, A2 and C1, C2 respectively, a stable loop through ASBRs <A3, A2, C3, C2, A3> exists.

Section 4 specifies that A3 and C3 that advertise a PATH to destination X, MUST aggregate all the PATHs through A1 and A2, and C1 and C2 respectively. This has the consequence of constraining the BGP/IDRP speaker in either domain A or domain C from choosing multiple routes to destination X, and importing only one route into OSPF. This breaks the multiple paths seen in one domain. The exact domain in which the multiple paths are broken is nondeterministic.

11. Authors' Present Addresses

Kannan Varadhan
USC/Information Sciences Institute
4676 Admiralty Way
Marina Del Rey, CA 90292-6695

Phone: +1 310 822 1511 x 402
EMail: kannan@isi.edu

Susan Hares
Merit, Inc.
1071 Beal Avenue,
Ann Arbor, MI 48109

Phone: +1 313 936 2095
EMail: skh@merit.edu

Yakov Rekhter
T.J. Watson Research Center, IBM Corporation
P.O. Box 704,
Yorktown Heights, NY 10598.

Phone: +1 914 784 7361
EMail: yakov@watson.ibm.com