Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 30, 2017

Yiqun. Cai
Heidi. Ou
Alibaba Group
Sri. Vallepalli
Mankamana. Mishra
Stig. Venaas
Cisco Systems
Andy. Green
British Telecom
June 28, 2017

PIM Designated Router Load Balancing
draft-ietf-pim-drlb-06

Abstract

On a multi-access network, one of the PIM routers is elected as a Designated Router (DR). On the last hop LAN, the PIM DR is responsible for tracking local multicast listeners and forwarding traffic to these listeners if the group is operating in PIM-SM. In this document, we propose a modification to the PIM-SM protocol that allows more than one of these last hop routers to be selected so that the forwarding load can be distributed among these routers.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

On a multi-access LAN such as an Ethernet, one of the PIM routers is elected as a DR. The PIM DR has two roles in the PIM-SM protocol. On the first hop network, the PIM DR is responsible for registering an active source with the Rendezvous Point (RP) if the group is operating in PIM-SM. On the last hop LAN, the PIM DR is responsible for tracking local multicast listeners and forwarding to these listeners if the group is operating in PIM-SM.

Consider the following last hop LAN in Figure 1:

```
                     ( core networks )
                       |     |     |
                       |     |     |
                      R1    R2     R3
                       |     |     |
                    --(last hop LAN)--
                             |
                             |
                     (many receivers)
```

Figure 1: Last Hop LAN

Assume R1 is elected as the Designated Router. According to [RFC4601], R1 will be responsible for forwarding traffic to that LAN on behalf of any local members. In addition to keeping track of IGMP and MLD membership reports, R1 is also responsible for initiating the creation of source and/or shared trees towards the senders or the RPs.

Forcing sole data plane forwarding responsibility on the PIM DR proves a limitation in the protocol. In comparison, even though an OSPF DR, or an IS-IS DIS, handles additional duties while running the OSPF or IS-IS protocols, they are not required to be solely responsible for forwarding packets for the network. On the other hand, on a last hop LAN, only the PIM DR is asked to forward packets while the other routers handle only control traffic (and perhaps drop packets due to RPF failures). The forwarding load of a last hop LAN is concentrated on a single router.

This leads to several issues. One of the issues is that the aggregated bandwidth will be limited to what R1 can handle towards this particular interface. These days, it is very common that the last hop LAN usually consists of switches that run IGMP/MLD or PIM snooping. This allows the forwarding of multicast packets to be restricted only to segments leading to receivers who have indicated their interest in multicast groups using either IGMP or MLD. The emergence of the switched Ethernet allows the aggregated bandwidth to exceed, some times by a large number, that of a single link. For example, let us modify Figure 1 and introduce an Ethernet switch in Figure 2.

```
                              ( core networks )
                                 |     |      |
                                 |     |      |
                                R1    R2     R3
                                 |     |      |
                             +=gi0===gi1===gi2=+
                             +                 +
                             +      switch     +
                             +                 +
                             +=gi4===gi5===gi6=+
                                 |     |      |
                                H1    H2     H3
```

Figure 2: Last Hop Network with Ethernet Switch

Let us assume that each individual link is a Gigabit Ethernet. Each router, R1, R2 and R3, and the switch have enough forwarding capacity to handle hundreds of Gigabits of data.

Let us further assume that each of the hosts requests 500 Mbps of data and different traffic is requested by each host. This represents a total 1.5 Gbps of data, which is under what each switch or the combined uplink bandwidth across the routers can handle, even under failure of a single router.

On the other hand, the link between R1 and switch, via port gi0, can only handle a throughput of 1Gbps. And if R1 is the only router, the PIM DR elected using the procedure defined by [RFC4601], at least 500 Mbps worth of data will be lost because the only link that can be used to draw the traffic from the routers to the switch is via gi0. In other words, the entire network’s throughput is limited by the single connection between the PIM DR and the switch (or the last hop LAN as in Figure 1).

The problem may also manifest itself in a different way. For example, R1 happens to forward 500 Mbps worth of unicast data to H1, and at the same time, H2 and H3 each requests 300 Mbps of different multicast data. Once again packet drop happens on R1 while in the mean time, there is sufficient forwarding capacity left on R2 and R3 and link capacity between the switch and R2/R3.

Another important issue is related to failover. If R1 is the only forwarder on the last hop router for shared LAN, in the event of a failure when R1 goes out of service, multicast forwarding for the entire LAN has to be rebuilt by the newly elected PIM DR. However, if there was a way that allowed multiple routers to forward to the LAN for different groups, failure of one of the routers would only lead to disruption to a subset of the flows, therefore improving the overall resilience of the network.

In this document, we propose a modification to the PIM-SM protocol that allows more than one of these routers, called Group Designated Router (GDR) to be selected so that the forwarding load can be distributed among a number of routers.

2. Terminology

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] .

With respect to PIM, this document follows the terminology that has been defined in [RFC4601] .

This document also introduces the following new acronyms:

- GDR: GDR stands for "Group Designated Router". For each multicast flow, either a (*,G) for ASM, or an (S,G) for SSM, a hash algorithm (described below) is used to select one of the routers as a GDR. The GDR is responsible for initiating the forwarding tree building for the corresponding multicast flow.

- GDR Candidate: a last hop router that has potential to become a GDR. A GDR Candidate must have the same DR priority and must run the same GDR election hash algorithm as the DR router. It must send and process new PIM Hello Options as defined in this document. There might be more than one GDR Candidate on a LAN. But only one can become GDR for a specific multicast flow.

3. Applicability

The proposed change described in this specification applies to PIM-SM last hop routers only.

It does not alter the behavior of a PIM DR on the first hop network This is because the source tree is built using the IP address of the sender, not the IP address of the PIM DR that sends the registers towards the RP. The load balancing between first hop routers can be achieved naturally if an IGP provides equal cost multiple paths (which it usually does in practice). And distributing the load to do registering does not justify the additional complexity required to support it.

4. Functional Overview

In the existing PIM DR election, when multiple last hop routers are connected to a multi-access LAN (for example, an Ethernet), one of them is selected to act as PIM DR. The PIM DR is responsible for sending local Join/Prune messages towards the RP or source. To elect the PIM DR, each PIM router on the LAN examines the received PIM Hello messages and compares its DR priority and IP address with those of its neighbors. The router with the highest DR priority is the PIM DR. If there are multiple such routers, their IP addresses are used as the tie-breaker, as described in [RFC4601].

In order to share forwarding load among last hop routers, besides the normal PIM DR election, the GDR is also elected on the last hop multi-access LAN. There is only one PIM DR on the multi-access LAN, but there might be multiple GDR Candidates.

For each multicast flow, that is (*,G) for ASM and (S,G) for SSM, a hash algorithm is used to select one of the routers to be the GDR. A new DR Load Balancing Capability (DRLBC) PIM Hello Option, which contains hash algorithm type, is announced by routers on interfaces where this specification is enabled. Last hop routers with the new DRLBC Option advertised in its Hello, and using the same GDR election hash algorithm and the same DR priority as the PIM DR, are considered as GDR Candidates.

Hash Masks are defined for Source, Group and RP separately, in order to handle PIM ASM/SSM. The masks, as well as a sorted list of GDR Candidates' Addresses are announced by DR in a new DR Load Balancing GDR (DRLBGDR) PIM Hello Option.

For each multicast flow, a hash algorithm is used to select one of the routers to be the GDR. The masks are announced in PIM Hello by DR as a DR Load Balancing GDR (DRLBGDR) Hello Option. Besides that, a DR Load Balancing Capability (DRLBC) Hello Option, which contains hash algorithm type, is also announced by the router on interfaces where this specification is enabled. Last hop routers with the new DRLBC Option advertised in its Hello, and using the same GDR election hash algorithm and the same DR priority as the PIM DR, are considered as GDR Candidates.

A hash algorithm based on the announced Source, Group or RP masks allows one GDR to be assigned to a corresponding multicast state. And that GDR is responsible for initiating the creation of the multicast forwarding tree for multicast traffic.

4.1. GDR Candidates

GDR is the new concept introduced by this specification. GDR Candidates are routers eligible for GDR election on the LAN. To become a GDR Candidate, a router MUST support this specification, have the same DR priority and run the same GDR election hash algorithm as the DR on the LAN.

For example, assume there are 4 routers on the LAN: R1, R2, R3 and R4, which all support this specification on the LAN. R1, R2 and R3 have the same DR priority while R4’s DR priority is less preferred. In this example, R4 will not be eligible for GDR election, because R4 will not become a PIM DR unless all of R1, R2 and R3 go out of service.

Further assume router R1 wins the PIM DR election, and R1, R2 run the same hash algorithm for GDR election, while R3 runs a different one. Then only R1 and R2 will be eligible for GDR election, R3 will not.

As a DR, R1 will include its own Load Balancing Hash Masks, and also the identity of R1 and R2 (the GDR Candidates) in its DRLBGDR Hello Option.

4.2. Hash Mask and Hash Algorithm

A Hash Mask is used to extract a number of bits from the corresponding IP address field (32 for v4, 128 for v6), and calculate a hash value. A hash value is used to select a GDR from GDR Candidates advertised by PIM DR. For example, 0.0.255.0 defines a Hash Mask for an IPv4 address that masks the first, the second and the fourth octets.

There are three Hash Masks defined,

o RP Hash Mask

o Source Hash Mask

o Group Hash Mask

The hask masks need to be configured on the PIM routers that can potentially become a PIM DR, unless the implementation provides default hash mask. An implementation SHOULD provide masks with default values 255.255.255.255 (IPv4) and FFFF:FFFF:FFFF:FFFF:FFFFF:FFFF:FFFF:FFFF (IPv6).

- o If the group is ASM, and if the RP Hash Mask announced by the PIM DR is not 0, calculate the value of hashvalue_RP [Section 4.3] to determine GDR.

- o If the group is ASM and if the RP Hash Mask announced by the PIM DR is 0, obtain the value of hashvalue_Group [Section 4.3] to determine GDR.

- o If the group is SSM, use hashvalue_SG [Section 4.3] to determine GDR.

A simple Modulo hash algorithm will be discussed in this document. However, to allow other hash algorithm to be used, a 4-bytes "Hash Algorithm Type" field is included in DRLBC Hello Option to specify the hash algorithm used by a last hop router.

If different hash algorithm types are advertised among last hop routers, only last hop routers running the same hash algorithm as the DR (and having the same DR priority as the DR) are eligible for GDR election.

4.3. Modulo Hash Algorithm

Modulo hash algorithm is discussed here as an example, with detailed description on hashvalue_RP.

- o For ASM groups, with a non-zero RP_hash mask, hash value is calculated as:

  hashvalue_RP = (((RP_address & RP_hashmask) >> N) & 0xFFFF) % M

  RP_address is the address of the RP defined for the group. N is the number of zeros, counted from the least significant bit of the RP_hashmask. M is the number of GDR Candidates.

  For example, Router X with IPv4 address 203.0.113.1, receives a DRLBGDR Hello Option from the DR, which announces RP Hash Mask 0.0.255.0, and a list of GDR Candidates, sorted by IP addresses from high to low, 203.0.113.3, 203.0.113.2 and 203.0.113.1. The ordinal number assigned to those addresses would be:

  0 for 203.0.113.3; 1 for 203.0.113.2; 2 for 203.0.113.1 (Router X)

  Assume there are 2 RPs: RP1 192.0.2.1 for Group1 and RP2 198.51.100.2 for Group2. Following the modulo hash algorithm:

N is 8 for 0.0.255.0, and M is 3 for the total number of GDR Candidates. The hashvalue_RP for RP1 192.0.2.1 is:

(((192.0.2.1 & 0.0.255.0) >> 8) & 0xFFFF % 3) = 2 % 3 = 2

matches the ordinal number assigned to Router X. Router X will be the GDR for Group1, which uses 192.0.2.1 as the RP.

The hashvalue_RP for RP2 198.51.100.2 is:

(((198.51.100.2 & 0.0.255.0) >> 8) & 0xFFFF % 3) = 100 % 3 = 1

which is different from Router X's ordinal number 2, hence, Router X will not be GDR for Group2.

- o If RP_hashmask is 0, a hash value for ASM group is calculated using the group Hash Mask:

  hashvalue_Group = (((Group_address & Group_hashmask) >> N) & 0xFFFF) % M

  Compare hashvalue_Group with Ordinal number assigned to Router X, to decide if Router X is the GDR.

- o For SSM groups, a hash value is calculated using both the source and group Hash Mask

  hashvalue_SG = ((((Source_address & Source_hashmask) >> N_S) & 0xFFFF) ^ (((Group_address & Group_hashmask) >> N_G) & 0xFFFF)) % M

4.4. PIM Hello Options

When a last hop PIM router sends a PIM Hello from an interface with this specification enabled, it includes a new option, called "Load Balancing Capability (DRLBC)".

Besides this DRLBC Hello Option, the elected PIM DR also includes a new "DR Load Balancing GDR (DRLBGDR) Hello Option". The DRLBGDR Hello Option consists of three Hash Masks as defined above and also the sorted list of all GDR Candidates' Address on the last hop LAN.

The elected PIM DR uses DRLBC Hello Option advertised by all routers on the last hop LAN to compose its DRLBGDR . The GDR Candidates use DRLBGDR Hello Option advertised by PIM DR to calculate hash value.

5. Hello Option Formats

5.1. PIM DR Load Balancing Capability (DRLBC) Hello Option

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Type = TBD          |         Length = 4            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Hash Algorithm Type                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: Capability Hello Option

Type: TBD.

Length: 4 octets

Hash Algorithm Type: 0 for Modulo hash algorithm

This DRLBC Hello Option SHOULD be advertised by last hop routers from interfaces with this specification enabled.

5.2. PIM DR Load Balancing GDR (DRLBGDR) Hello Option

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Type = TBD          |         Length                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Group Mask                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Source Mask                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          RP Mask                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  GDR Candidate Address(es)                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: GDR Hello Option

Type: TBD

Length: 3 x (4 byte or 16 byte) + n x (4 byte or 16 byte) where n is number of GDR candidate

Group Mask (32/128 bits): Mask

Source Mask (32/128 bits): Mask

RP Mask (32/128 bits): Mask

All masks MUST be in the same address family as the Hello IP header.

GDR Address (32/128 bits): Address(es) of GDR Candidate(s)

All addresses must be in the same address family as the Hello IP header. The addresses are sorted from high to low. The order is converted to the ordinal number associated with each GDR candidate in hash value calculation. For example, addresses advertised are R3, R2, R1, the ordinal number assigned to R3 is 0, to R2 is 1 and to R1 is 2.

If "Interface ID" option, as described in [RFC6395], presents in a GDR Candicate’s PIM Hello message, and the "Router ID" portion is non-zero,

+ For IPv4, the "GDR Candidate Address" will be set directly to "Router ID".

+ For IPv6, the "GDR Candidate Address" will be set to the IPv4-IPv6 translated address of "Router ID", as described in [RFC4291] , that is the "Router-ID" is appended to the prefix of 96-bits zeros.

If the "Interface ID" option is not present in a GDR Candidate’s PIM Hello message, or if the "Interface ID" option is present, but"Router ID" field is zero, the "GDR Candidate Address" will be the IPv4 or IPv6 source address from PIM Hello message.

This DRLBGDR Hello Option SHOULD only be advertised by the elected PIM DR.

6. Protocol Specification

6.1. PIM DR Operation

The DR election process is still the same as defined in [RFC4601]. A DR that has this specification enabled on the interface, advertises the new LBGRD Hello Option, which contains value of masks from user configuration, followed by a sorted list of all GDR Candidates' Addresses, from high to low. Moreover, same as non-DR routers, DR also advertises DRLBC Hello Option to indicate its capability of supporting this specification and the type of its GDR election hash algorithm.

If a PIM DR receives a PIM Hello with DRLBGRD Option, the PIM DR SHOULD ignore the TLV.

If a PIM DR receives a neighbor DRLBC Hello Option, which contains the same hash algorithm type as the DR, and the neighbor has the same DR priority as the DR, PIM DR SHOULD consider the neighbor as a GDR Candidate and insert the GDR Candidate's Address into the sorted list of DRLBGRD Option.

6.2. PIM GDR Candidate Operation

When an IGMP/MLD join is received, without this proposal, only PIM DR will handle the join and potentially run into the issues described earlier. Using this proposal, a hash algorithm is used on GDR Candidate to determine which router is going to be responsible for building forwarding trees on behalf of the host.

A router which supports this specification, a interface where this protocol is enabled advertises DRLBC Hello Option in its PIM Hello, even if the router may not be considered as a GDR Candidate, for example, due to low DR priority. once DR election is done, DRLBGDR Hello option would be received from the current PIM DR on link.

A GDR Candidate may receive a DRLBGDR Hello Option from PIM DR, with different Hash Masks from those configured on it, The GDR Candidate must use the Hash Masks advertised by the PIM DR to calculate the hash value.

A GDR Candidate may receive a DRLBGDR Hello Option from a PIM router which is not DR. The GDR Candidate MUST ignore such DRLBGDR Hello Option.

A GDR Candidate may receive a Hello from the elected PIM DR, and the PIM DR does not support this specification. The GDR election described by this specification will not take place, that is only the PIM DR joins the multicast tree.

A router only act as GDR if it is included in the GDR list of DRLBGDR Hello Option

6.2.1. Router receives new DRLBGDR

When a router receives new DRLBGDR from the current PIM DR, it need to process and check if router is in list of of GDR

1. If router is not listed as a GDR candidate in DRLBGDR , no action needed.

2. If router is listed as a GDR candidate in DRLBGDR, then it need to process each of the groups in the IGMP/MLD reports. The masks are announced in the PIM Hello by DR as DRLBGDR Hello option. For each of groups in the reports it need to run hash algorithem (described in section 4.3) based on the announced Source, Group or RP masks to determine if it is GDR for specified group. If hash result is to be GDR for multicast flow, it does build multicast forwarding tree. if it is not GDR for flow, no action is needed.

6.2.2. Router receives updated DRLBGDR

If router (GDR or non GDR) receives an unchanged DRLBGDR from the current PIM DR, no action needed.

If router (GDR or non GDR) receives a new or modified DRLBGDR from the current PIM DR. It requires processing as described below

1. If it was GDR and still included in current GDR list: It need to process each of the groups, run hash algorithem to check if it is still GDR for given group.

    If it was GDR for group earlier. and even new hash choose it as GDR, no processing required.

    If it was GDR for group earlier and now it is no more GDR, then it sets its assert metric for this flow to be (PIM_ASSERT_INFINITY - 1), as explained in Sec 6.3

    If it was not GDR for group earlier, and even new hash does not make it GDR no processing required.

    If it was not GDR earlier and now becomes GDR, it starts building multicast forwarding tree for this flow.

2. If it was non GDR , and updated DRLBGDR from current PIM DR contains this router as one of the GDR. In this case this router

being new GDR candiate MUST run hash algorithem for each of the groups (multicast flows) and for given group,

If it is not GDR, no processing is required.

If it is hased as GDR , it need to build multicast forwarding tree.

3. If a router receives IGMP/MLD report for flow for which the router has been the GDR AND the DRLBGDR has changed since last report for this flow, then the router MUST determine if it is still the GDR. if it is, no action needed. if it is not, then the router sets its assert metric for this flow to be (PIM_ASSERT_INFINITY - 1) as explained in Sec 6.3.

6.3. PIM Assert Modification

It is possible that the identity of the GDR might change in the middle of an active flow. Examples this could happen include:

When a new PIM router comes up

When a GDR restarts

When the GDR changes, existing traffic might be disrupted. Duplicates or packet loss might be observed. To illustrate the case, consider the following scenario: there are two streams G1 and G2. R1 is the GDR for G1, and R2 is the GDR for G2. When R3 comes up online, it is possible that R3 becomes GDR for both G1 and G2, hence R3 starts to build the forwarding tree for G1 and G2. If R1 and R2 stop forwarding before R3 completes the process, packet loss might occur. On the other hand, if R1 and R2 continue forwarding while R3 is building the forwarding trees, duplicates might occur.

This is not a typical deployment scenario but it still might happen. Here we describe a mechanism to minimize the impact. The motivation is that we want to minimize packet loss. And therefore, we would allow a small amount of duplicates and depend on PIM Assert to minimize the duplication.

When the role of GDR changes as above, instead of immediately stopping forwarding, R1 and R2 continue forwarding to G1 and G2 respectively, while at the same time, R3 build forwarding trees for G1 and G2. This will lead to PIM Asserts.

With introduction of GDR, the following modification to the Assert packet MUST be done: if a router enables this specification on its downstream interface, but it is not a GDR (before network event it was GDR), it would adjust its Assert metric to (PIM_ASSERT_INFINITY - 1).

Using the above example, for G1, assume R1 and R3 agree on the new GDR, which is R3. R1 will set its Assert metric as (PIM_ASSERT_INFINITY - 1). That will make R3, which has normal metric in its Assert as the Assert winner.

For G2, assume it takes a little bit longer time for R2 to find out that R3 is the new GDR and still thinks itself being the GDR while R3 already has assumed the role of GDR. Since both R2 and R3 think they are GDRs, they further compare the metric and IP address. If R3 has the better routing metric, or same metric but better tie-breaker, the result will be consistent with GDR selection. If unfortunately, R2 has the better metric or same metric but better tie-breaker R2 will become the Assert winner and continues to forward traffic. This will continue until:

The next PIM Hello option from DR is seen that selects R3 as the GDR. R3 will then build the forwarding tree and send an Assert.

The process continues until R2 agrees to the selection of R3 as being the GDR, and set its own Assert metric to (PIM_ASSERT_INFINITY - 1), which will make R3 the Assert winner. During the process, we will see intermittent duplication of traffic but packet loss will be minimized. In the unlikely case that R2 never relinquishes its role as GDR (while every other router thinks otherwise), the proposed mechanism also helps to keep the duplication to a minimum until manual intervention takes place to remedy the situation.

7. Compatibility

In case of hybrid Ethernet shared LAN (where some PIM router enables specification defined in this draft and some do not enable)

- o If router which does not support specification defined in this draft becomes DR on link, it MUST be only DR on link as [RFC4601] and there would be no router which would act as GDR.

- o If router which does not support specification defined in this draft becomes non DR on link, then it should act as non-DR defined in [RFC4601].

8. Manageability Considerations

- o All of the router in LAN who are supporting this specification MUST use identical Hash Algorithm Type (described in section 5.1). In case of hybrid Hash Algorithm Type router must go backward to use DR election method defined in PIM-SM [RFC4601]. Migration between different algorithem type is out of scope of this document.

9. IANA Considerations

Two new PIM Hello Option Types have been assigned to the DR Load Balancing messages. [HELLO-OPT], this document recommends 34(0x22) as the new "PIM DR Load Balancing Capability Hello Option", and 35(0x23) as the new "PIM DR Load Balancing GDR Hello Option".

10. Security Considerations

Security of the new DR Load Balancing PIM Hello Options is only guaranteed by the security of PIM Hello message, so the security considerations for PIM Hello messages as described in PIM-SM [RFC4601] apply here.

11. Acknowledgement

The authors would like to thank Steve Simlo, Taki Millonis for helping with the original idea, Bill Atwood, Bharat Joshi for review comments, Toerless Eckert and Rishabh Parekh for helpful conversation on the document.

12. References

12.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, DOI 10.17487/RFC4291, February 2006, <http://www.rfc-editor.org/info/rfc4291>.

[RFC4601] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", RFC 4601, DOI 10.17487/RFC4601, August 2006, <http://www.rfc-editor.org/info/rfc4601>.

[RFC6395] Gulrajani, S. and S. Venaas, "An Interface Identifier (ID) Hello Option for PIM", RFC 6395, DOI 10.17487/RFC6395, October 2011, <http://www.rfc-editor.org/info/rfc6395>.

12.2. Informative References

[HELLO-OPT]
IANA, "PIM Hello Options", IANA PIM-HELLO-OPTIONS, March 2007.

Authors' Addresses

Yiqun Cai
Alibaba Group

Email: yiqun.cai@alibaba-inc.com

Heidi Ou
Alibaba Group

Sri Vallepalli
Cisco Systems
3625 Cisco Way,
Sanjose, CALIFORNIA 95134
UNITED STATES

Email: svallepa@cisco.com

Mankamana Prasad Mishra
Cisco Systems
821 Alder Drive,
MILPITAS, CALIFORNIA 95035
UNITED STATES

Email: mankamis@cisco.com

Stig Venaas
Cisco Systems
821 Alder Drive,
MILPITAS, CALIFORNIA 95035
UNITED STATES

Email: stig@cisco.com

Andy Green
British Telecom
Adastral Park
Ipswich IP5 2RE
United Kingdom

Email: andy.da.green@bt.com