## PIM DR Improvement
### draft-zhang-pim-dr-improvement-01.txt

Abstract

   PIM is worldly deployed multicast protocol.  This document will
   improve the stability of PIM protocol, decrease the lost of multicast
   packets when the PIM DR (Designed Router) is down.

Status of This Memo

Copyright Notice

Table of Contents

## 1.  Introduction

   Multicast technology is used widely.  Many modern technology use PIM
   technology, such as IPTV, Net-Meeting, and so on.  There are many
   events that will influence the quality of multicast services.  Except
   the unicast routes changing will cause the lost of multicast packets.
   The change of DR cause the lost of multicast packets too.

   When a DR on a share-media LAN is down, other routers will elect a
   new DR until the expiration of Hello-Holdtime.  The default value of
   Hello-Holdtime is 105 seconds.  Although the value of Hello-Holdtime
   can be changed by manual, when the DR is down, there are still many
   multicast packets will be lost.  The quality of IPTV and Net-Meeting
   will be influenced.

```
                 \                              /
                  \                            /
              -------                      -------
              |  A  |                      |  B  |
              -------                      -------
                 | DR                         |
                 |                            |
              -------                      -------
              | SW  |----------------------------| SW  |
              -------                      -------
                 |                            |
             Figure 1: An example of multicast network
```
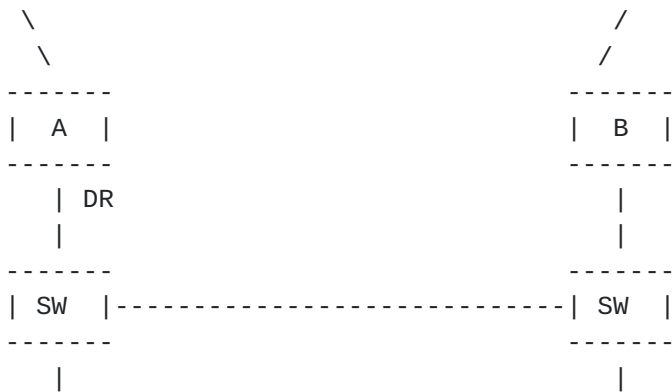
   For example, there were two routers on one Ethernet.  RouterA was
   elected to DR.  When RouterA is down, the multicast packets are

discarded until the RouterB is elected to DR and RouterB imports the
multicast flows successfully.

We suppose that there is only a RouterA in the Ethernet at first in
Figure 1.  RouterA is the DR who is responsible for forwarding
multicast flows.  When RouterB connects the Ethernet, RouterB will be
elected to DR because a high priority.  So RouterA will stop
forwarding multicast packets.  The multicast flows will not recover
until RouterB joins the multicast group after it is elected to DR.

## 2.  Terminology

Backup Designated Router (BDR): A shared-media LAN like Ethernet may
have multiple PIM-SM routers connected to it.  Except for DR, a other
router who will act on behalf of directly connected hosts with
respect to the PIM-SM protocol.  But BDR will not forward the flows.
When DR is down, the BDR will forward multicast flows immediately.  A
single BDR is elected per interface like the DR.

## 3.  PIM hello message format

In RFC4601, the PIM hello message format was defined.  In this
document, we define two new option values which are including Type,
Length, and Value.

```
     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
    +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
    |                   Hello message format                        |
    |                                                               |
    +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
    |          OptionType            +        OptionLength           |
    +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
    |                         OptionValue                           |
    |                                                               |
    +++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
                  Figure 2: Hello message format
```
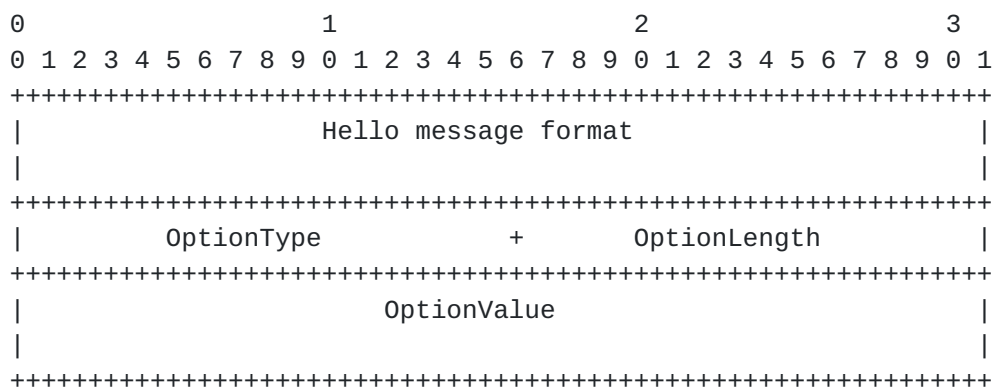
### 3.1.  DR Option format

o  OptionType : The value is TBD.

o  OptionLength: If the network is support IPv4, the OptionLength is
   4 octets.  If the network is support IPv6, the OptionLength is 16
   octets.

   o  OptionValue: The OptionValue is IP address of DR.  If the network
      is support IPv4, the value is IPv4 address of DR.  If the network
      is support IPv6, the value is IPv6 address of DR.

## 3.2.  BDR Option format

   o  OptionType : The value is TBD.

   o  OptionLength: If the network is support IPv4, the OptionLength is
      4 octets.  If the network is support IPv6, the OptionLength is 16
      octets.

   o  OptionValue: The OptionValue is IP address of BDR.  If the network
      is support IPv4, the value is IPv4 address of BDR.  If the network
      is support IPv6, the value is IPv6 address of BDR.

## 4.  The Protocol Treatment

   A new router starts to send hello messages with the values of DR and
   BDR are all set to 0 after its interface is enabled in PIM on a
   share-media LAN.  When the router receive hello messages from other
   routers on the same share-media LAN, the router will check if the
   value of DR or BDR is filled.  If the value of DR or BDR is filled
   with IP address of router who is sending hello messages, the router
   will store the IP address.

   Then the new router compare the priority and IP address itself to the
   stored IP address of DR and BDR accord to the algorithm of RFC 4601.
   If the new router notices that it is better to be DR than the existed
   DR or BDR.  The router will make itself the BDR, and send new hello
   messages with its IP address as BDR and existed DR.  If the router
   notices that the existed DR is most priority in the share-media LAN,
   but the existed BDR is set to 0x0 in the received hello messages, or
   the existed BDR is not better than the new router to be DR except
   existed DR, the router will elect itself to BDR.  If the router
   notices that it is not better to be DR than existed DR and BDR, the
   router will respect the PIM protocol.

   When the new router becomes the new BDR, the router will join the
   existed multicast groups, import multicast flows from upstream
   routers.  But the BDR MUST not forward the multicast flows to avoid
   the duplicate multicast packets in the share-media LAN.  The new
   router will monitor the DR.  When the DR becomes unavailable because
   of the down or other reasons, the BDR will forward multicast flows
   immediately.

## 4.1.  Sending Hello Messages

   When a new router's interface is enabled in PIM protocol, the router
   send hello messages with the values of DR and BDR are filled with
   0x0.

   When a new router sets itself BDR after receive hello messages from
   other routers, the router send hello messages with the value of DR is
   set to the IP address of existed DR and the value of BDR is set to
   the IP address of the router itself.

   When a new router notices the existed DR and BDR is more priority
   than itself.  The router will send hello messages with the values of
   DR and BDR are filled with existed DR and BDR.

   When a existed router sets itself non DR and non BDR after receive
   hello messages from other routers, the router will send hello
   messages with the value of DR is set to existed DR and the value of
   BDR is set to new BDR.

## 4.2.  Receiving Hello Messages

   When the values of DR and BDR which are carried by hello messages are
   received is all set to 0x0, the router MUST elect the DR due to the
   algorithm of RFC4601.  And elect a new BDR which are the best choice
   except DR.

   When the value of DR which is carried by received hello messages is
   not 0x0, and the value of BDR is set to 0x0, the router will elect
   itself to BDR.

   When the values of DR and BDR that carried by received hello messages
   are all larger than 0x0.  The router will mark the existed DR, and
   compare itself and the BDR in message.  When the router notice that
   it is better to be DR than existed BDR.  The router will elect itself
   to the BDR.

   When a router receives a new hello message with the values of DR and
   BDR are set to 0x0.  The router will compare the new router with
   itself.  If the router noticed that the new router is better to be DR
   than itself, the router will set the BDR to the new router.

## 4.3.  The election of DR and BDR

   When all the routers on a shared-media LAN are start to work on the
   same time, the election of DR is same as RFC4601.  And all the
   routers will elect a BDR which is suboptimum to DR.  The hello

messages sent by all the routers are same with the value of DR and
BDR are all set.

When a new router start to work on a shared-media LAN and receive
hello messages from other routers that the value of DR is set at
least.  The new router will not change the existed DR even if it is
superior to the existed DR.  If the new router is superior to existed
BDR, the new router will replace the place of BDR on the LAN.

When an existed router receives hello messages from a new router, and
the existed router is DR on the LAN, the existed DR router will
compare the new router and all the other routers on the LAN.  If the
new router is superior to all the other routers, the existed DR
router will treat the new router as new BDR.

When an existed router receives hello messages from a new router, and
the existed router is BDR on the LAN, the existed BDR will compare
itself and the new router.  If the new router is superior to itself,
the existed BDR will elect the new router as new BDR, and set itself
for nothing.  Then the old BDR will send prune message to upstream
routers.

When an existed router receives hello messages from a new router, and
the existed router is neither DR nor BDR on the LAN, the existed
router will compare the existed BDR and the new router.  If the new
router is superior to all the other routers, the existed DR router
will treat the new router as new BDR.

## 5.  Deployment suggestion

If there are two and more routers on a share-media LAN, and the
multicast services is sensitive due to the lost of multicast packets,
the LAN should deploy the function of DR and BDR in this document.

## 6.  Security Considerations

For general PIM Security Considerations.

## 7.  IANA Considerations

IANA is requested to allocate OptionTypes in TLVs of hello message.
Include DR and BDR.

## 8.  Normative References

   [I-D.ietf-pim-rfc4601bis]
              Fenner, B., Handley, M., Holbrook, H., Kouvelas, I.,
              Parekh, R., Zhang, J., and L. Zheng, "Protocol Independent
              Multicast - Sparse Mode (PIM-SM): Protocol Specification
              (Revised)", draft-ietf-pim-rfc4601bis-06 (work in
              progress), August 2015.

   [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>.

Authors' Addresses

   Zheng(Sandy) Zhang
   ZTE Corporation
   No. 50 Software Ave, Yuhuatai Distinct
   Nanjing
   China

   Email: zhang.zheng@zte.com.cn


   Fangwei Hu
   ZTE Corporation
   No.889 Bibo Rd
   Shanghai
   China

   Email: hu.fangwei@zte.com.cn


   BenChong Xu
   ZTE Corporation
   No. 68 Zijinghua Road, Yuhuatai Distinct
   Nanjing
   China

   Email: xu.benchong@zte.com.cn