ISIS: Path MTU calculation in ISIS

Status of This Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on February 5, 2009.

Abstract

This draft defines a method for calculating the PMTU for each IPv6 destinations using ISIS routing protocol. By doing so the overhead incurred in the existing path maximum transferable unit discovery mechanism is reduced and the same solution can be extended to other link state routing protocols as well.

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 RFC 2119 [RFC2119].

Abstract ........................................................... 1
Specification of Requirements ................................... 1
1. Introduction .................................................. 2
2. Deploying scenarios ......................................... 3
3. New fields .................................................. 3
   3.1 Extended IS Reachability TLV & MT Intermediate Systems TLV .. 3
   3.2 IPv6 Reachability TLV & MT IPv6 Reachability TLV .......... 4
      3.2.1 MTU representation in IPv6 Reachability TLV & MT IPv6 Reachability TLV ............................ 4
      3.2.1.1 Intra area route information ..................... 5
      3.2.1.2 Inter level redistribute route information ...... 5
4. Computation .................................................. 6
1. Introduction

The current IPv6 PMTU discovery has the following drawbacks,

1. The IPv6 PMTU discovery is done by trial and error method, which can result in inefficient forwarding such as described below and this in turn can result in delay in packet transmission.
   - Packets may be dropped because of packet too big reason by any intermediate router.
   - Packets that are very small in size may be forwarded for considerable amount of time resulting in inefficient usage of available bandwidth.

2. The source comes to know about the packet drop only by ICMPv6 packet too big error. But this error packet will have to travel from the problem occurred router to the source of the packet, which consumes considerable amount of bandwidth on all the intermediate links between the originator and the problem occurred node.

This document describes a method to dynamically compute an optimal PMTU for each IPv6 destinations using IS-IS and [IPv6PMTU] describes the modification needed in IPv6 to convey the optimal PMTU to the hosts.

2. Deploying scenarios

This paper suggests a solution to calculate IPv6 PMTU for each IPv6 destinations using the IGP ISIS. The solution proposed can work independently in a single domain as shown below,

```
|                |
H-|  IGP domain    |-H

If a router is originating a packet then it can have the PMTU for the destined route and if a host is originating a packet then it can get the PMTU value from the nearest gateway router as described in [IPv6PMTU].

```

```
|             |    |         |    |             |
H-| IGP domain  |-ER-|   EGP   |-ER-| IGP domain  |-H

In the above topology R represents router and ER represents edge router.
For the solution to work across routing domain as shown above protocols such as BGP should carry out similar kind of PMTU calculation for its route else the scope of the PMTU so computed will be confined to the ISIS IGP domain.

As other routing protocols such as BGP run over backbone which have higher MTU links the BGP domain generally won’t constitute for PMTU bottleneck and hence it is just enough if other routing protocols such as BGP convey the PMTU computed from one ISIS domain to another.

The method by which the PMTU computation or PMTU information conveying is done in other protocols is out of scope of this document.

3. New fields

New sub-TLVs are added to the following existing TLV for calculating PMTU to all the ISIS IPv6 destinations.

3.1 Extended IS Reachability TLV & MT Intermediate Systems TLV

A new sub TLV is added to the Extended IS Reachability TLV(22) and MT Intermediate Systems TLV(222). The value field should indicate the MTU of the link represented by respective TLVs.

The format of the sub-TLV is as shown below and this sub-TLV can appear at most once in each Extended IS Reachability TLV and MT Intermediate Systems TLV.

- TYPE - 0x88
- LENGTH - Total length of the value field, it should be 4
- VALUE - 4-byte MTU value of the link

```
+-----------------+
|     MTU value   |
|                |  4
+-----------------+
```

As each NBR TLV indicates a link between a pair of nodes in the domain, this link can be mapped to an interface. The MTU associated with that interface is advertised in the sub-TLV.

Whenever there is a change in MTU value of a physical or logical interface represented by the Extended IS Reachability TLV or MT Intermediate Systems TLV, ISIS should re-originate the respective TLV with the new MTU value.

3.2 IPv6 Reachability TLV & MT IPv6 Reachability TLV

A new sub TLV is added to the IPv6 Reachability TLV(236) and MT IPv6 Reachability TLV(237). The value field should indicate the MTU of the interface represented by respective TLVs.

The format of the sub-TLV is as shown below and this sub-TLV can appear at most once in each IPv6 Reachability TLV and MT IPv6 Reachability TLV.

- TYPE - 0x88
- LENGTH - Total length of the value field, it should be 4
3.2.1 MTU representation in IPv6 Reachability TLV & MT IPv6 Reachability TLV

IPv6 Reachability TLV and MT IPv6 Reachability TLV can represent these kinds of routes and the MTU value advertised in each of these should convey the below described information.

3.2.1.1 Intra area route information

The intra area route information advertised by IPv6 Reachability TLV and MT IPv6 Reachability TLV should carry the MTU value of the physical/logical interface. If the MTU is not applicable to an interface or if MTU value is not available for an interface, then this sub TLV need not be advertised in IPv6 Reachability TLV and MT IPv6 Reachability TLV at which the router carries out the best effort PMTU computation as described in sec 4.

3.2.1.2 Inter level redistribute route information

When a router receives the inter-level redistributed route (L1 to L2 or L2 to L1) it will not have the full path to the destination. The received router will know the path only till the L12 router that redistributed the route. In order that all routers calculate the PMTU till the destination the inter level redistribute router should convey additional information.

Thus inter level redistributing router should advertise the calculated PMTU from itself to the destination for each destination while redistributing routes from one level to other.

3.2.1.3 External domain route information

While external domain routes are redistributed into ISIS if external domain has PMTU information for any of its routes then this information should be propagated into ISIS domain and ISIS should advertise the above received PMTU information for each of its route. If PMTU information is not available then this sub TLV need not be advertised in the IPv6 Reachability TLV and MT IPv6 Reachability TLV at which the router carries out the best effort PMTU computation till the inter-domain redistributing router as described in sec 4.

3.2.1.4 Summary route information

The maximum value of the computed PMTU for any of the individual route that has been summarized should be the PMTU for the summarized route.

4. Computation
4.1 PMTU for a destination node

The least value of the MTU, which is advertised in Extended IS Reachability TLV or MT Intermediate Systems TLV, from the root to destination node along the SPF path should be considered as the PMTU to reach that particular destination node.
If a router has not advertised MTU value for its link and if that link is on SPF tree then the link MTU non-advertisement should be ignored and the PMTU to reach that node is same as the PMTU to reach its parent node.

If the MTU un-available link has the least MTU to a destination then the originator of the IPv6 packet will come to know about this bottleneck through the ICMPv6 packet too big error else PMTU computation is optimal. Thus the above mechanism ensures the best effort PMTU computation.

As described in [ISO10589], sec ANNEX F.2.1 each path can be described as \(<N, d(N), \{\text{Adj}(N)\}>\). Here N denotes the destination node, \(d(N)\) represents the cost to the destination and \(\{\text{Adj}(N)\}\) represents the set of Adjacency from which the destination node N can be reached. An additional PMTU parameter is added to the PATH attribute. Thus an PATH in SPF tree can be represented by \(<N, d(N), \{\text{Adj}(N)\}, \{\text{Pmtu}(N)\}>\) where \(\{\text{Pmtu}(N)\}\) is the set of PMTU calculated from the source node to reach the destination node N. In case of equicost multiple paths the highest PMTU among the various paths should be considered as the PMTU to reach a particular destination.

4.2 PMTU for a destination address

The least value of the following should be considered as the PMTU to reach that particular destination address,

1. The MTU advertised for a destination address in IPv6 Reachability TLV or MT IPv6 Reachability TLV.
2. The PMTU computed till the destination node, which advertised the destination prefix, as described in sec 4.1.

If either of this information is not available then the remaining one should be considered as the PMTU to reach a particular destination address and if none of this information is available then the PMTU computation to that particular destination address can be ignored.

This computed PMTU should be advertised by L12 router in the IPv6 Reachability TLV or MT IPv6 Reachability TLV while redistributing the routes across levels.

4.3 PMTU for a default route due to attach bit

The L1 router that receives ATT bit should calculate the PMTU to reach nearest L12 router, which advertises ATT bit and associate this PMTU value to the default route.

5. Backward compatibility

The routers incapable of recognizing the MTU sub-TLV will ignore MTU sub-TLV and parse the rest of the information in Extended IS Reachability TLV, MT Intermediate Systems TLV, IPv6 Reachability TLV & MT IPv6 Reachability TLV.

If a PMTU capable router receives an Extended IS Reachability TLV, MT Intermediate Systems TLV, IPv6 Reachability TLV & MT IPv6 Reachability TLV without the MTU information then the router can carry out the best effort PMTU computation as described in sec 4.1.
6. Future extension

The PMTU computation for each route can be extended to IPv4 destination in ISIS and other IPv6/IPv4 link state routing protocols.

7. Acknowledgments

The author would like to thank Saravana Kumar and K.L.Srini

8. Security Consideration

ISIS security applies to the work presented. No specific security issues with the proposed solutions are known. The authentication procedure for ISIS PDUs is the same regardless of MTU information inside the ISIS PDUs.

9. IANA Considerations

IANA is requested to create a new registry, "IS-IS PMTU ID values" with the assignment listed in this document and registration policies for future assignments.

IANA is also requested to update the IS-IS codepoint registry (http://www.iana.org/assignments/isis-tlv-codepoints) so that sub TLV code 136 refers to this document.

[[ note to the RFC-editor: the above paragraph may be removed or reworded prior to RFC publication ]]

10. References

10.1. Normative References


[IPv6PMTU] Vijay Kumar Vasantha "draft-kumar-ipv6-pmtu-using-routing/proto-00.txt".

10.2. Informative References

[RFC2463] Internet Control Message Protocol (ICMPv6) for the Internet protocol Version 6 (IPv6) Specification


11. Authors’ Addresses

Vijay Kumar Vasantha
Huawei Technologies India Private Limited
Bangalore, India - 560008
vijaykumar@huawei.com

Full Copyright Statement

Copyright (C) The IETF Trust (2008).

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.