Forcing Fragmentation to Network MTU

Status of This Memo

Abstract

There exists a class of applications which cannot accept Path MTU discovery [RFC-1981]. This is recommended to be turned on by default for IPv6 [IPV6-SPEC-V2, 4.5, 5]. This document describes an extension to the BSD API [RFC-2133] to inform the kernel when it should be fragmenting at the network MTU rather than trying to discover the path MTU.

Path MTU discovery works well when there is a stream of data. However for distributed servers sending small answers larger than the network MTU to many clients the lack of fragmentation in intermediate nodes is anathema.

RFC-1981 Section 5.6 recommends that a system utility be provided to:

- Specify that Path MTU Discovery not be done on a given path.

- Change the PMTU value associated with a given path.

This documents specifies how to do the same at the application level to a individual socket.

1 - No Path MTU Discovery

Disabling path MTU. This would be a binary option, on a per socket basis. If set the kernel would fragment all IP packets generated by this socket at the network MTU (1280) [IPV6-SPEC-V2, 5] unless it has a priori knowledge of the path MTU which it could then use.

```
int discover = 0;

if (setsockopt(s, IPPROTO_IPV6, IPV6_MTU_DISCOVER,
        (char*) &discover, sizeof(discover)) == -1)
        perror("setsockopt IPV6_MTU_DISCOVER");
```

2 - Setting IP Fragmentation Size

A socket option to explicitly set the maximum IP fragment size for messages derived from this socket. This should be used when it is known that the path MTU is greater than the network MTU.

When combined with disabling path MTU discovery the resultant behaviour should be the same as if the network MTU was set to this value.

```
size_t mtu = 1280;

if (setsockopt(s, IPPROTO_IPV6, IPV6_MTU,
         (char*) &mtu, sizeof(mtu)) == -1)
         perror("setsockopt IPV6_MTU");
```

3 - Usage Consideration

Applications sending IPv6 datagrams larger than the network MTU should ensure that the reviever has buffer space large enough to receive them. How this is done is left upto the application.

4 - Security Considerations

This document is believed to introduce no security problems.

Existing security problems with using IPv6 datagrams larger than the network MTU include:

- Fragment overlay attacks.

- Reassembly buffer exhaustion.

References

[RFC-1981]
McCann, J., Deering, S., and Mogul, J. "Path MTU Discovery for IP version 6", RFC 1981, August 1996.

[RFC-2133]
Gilligan, R. E., Thomson, S., and Bound, J., "Basic Socket Interface Extensions for IPv6", RFC 2133, April 1996.

[IPV6-SPEC-V2] work in progress
Deering, S., and Hinden, R., "Internet Protocol, Version 6 (IPv6) Specification", draft-ietf-ipngwg-ipv6-spec-v2-01.txt, November 1997.

Author' Address

Mark Andrews
CSIRO Mathematical and Information Sciences
Locked Bag 17
North Ryde NSW 2113
AUSTRALIA
+61 2 9325 3148
<Mark.Andrews@cmis.csiro.au>