TCP Maintenance and Minor Extensions (tcpm)
Internet-Draft
Intended status: Informational
Expires: August 2, 2009

F. Gont
UTN/FRH

January 29, 2009

On the problem of long delays between connection-establishment attempts in TCP

draft-gont-tcpm-connection-delays-00.txt

Status of this Memo

Copyright Notice

Abstract

This document discusses a number of solutions to the problem of long delays between connection establishment attempts in TCP.

Table of Contents

1. Introduction

This document discusses a number of alternative solutions to the one proposed in [I-D.ietf-tcpm-tcp-soft-errors] for the problem of long delays between connection establishment attempts in TCP.

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].

2. A more conservative approach

A more conservative approach would be to abort a connection in the SYN-SENT or SYN-RECEIVED states only after an ICMP Destination Unreachable has been received a specified number of times, and the SYN segment has been retransmitted more than some specified number of times.

Two new parameters would have to be introduced to TCP, to be used only during the connection-establishment phase: MAXSYNREXMIT and MAXSOFTERROR. MAXSYNREXMIT would specify the number of times the SYN segment would have to be retransmitted before a connection is aborted. MAXSOFTERROR would specify the number of ICMP messages indicating soft errors that would have to be received before a connection is aborted.

Two additional variables would need to be introduced to store additional state information during the connection-establishment phase: "nsynrexmit" and "nsofterror". Both would be initialized to zero. "nsynrexmit" would be incremented by one every time the SYN segment is retransmitted. "nsofterror" would be incremented by one every time an ICMP message that indicates a soft error is received.

A connection in the SYN-SENT or SYN-RECEIVED states would be aborted if nsynrexmit was greater than MAXSYNREXMIT and "nsofterror" was simultaneously greater than MAXSOFTERROR.

This approach would give the network more time to solve the connectivity problem. However, it should be noted that depending on the values chosen for the MAXSYNREXMIT and MAXSOFTERROR parameters, this approach could still lead to long delays between connection establishment attempts, thus not solving the problem. For example, BSD systems abort connections in the SYN-SENT or the SYN-RECEIVED state when a second ICMP error is received, and the SYN segment has been retransmitted more than three times. They also set up a "connection-establishment timer" that imposes an upper limit on the time the connection establishment attempt has to succeed, which

expires after 75 seconds [Stevens2]. Even when this policy may be better than the three-minutes timeout policy specified in [RFC1122], it may still be inappropriate for handling the potential problems described in this document. This more conservative approach has been implemented in BSD systems since, at least, 1994 [Stevens2].

3. Asynchronous Application Notification

In section 4.2.4.1, [RFC1122] states that there MUST be a mechanism for reporting soft TCP error conditions to the application. Such a mechanism (assuming one is implemented) could be used by applications to cycle through the destination IP addresses. However, this approach would increase application complexity, and would take a long time to kick in, as it would require all existing applications to be modified.

4. Issuing several connection requests in parallel

For those scenarios in which a domain name maps to several IP addresses, several connection requests could be issued in parallel, each one to a different destination IP address. The host would then use the first connection attempt to succeed, eliminating the potential delay in establishing a connection with the destination host. However, this would mean that every attempt to connect to a multihomed host would imply sending several SYN segments, making it hard for network operators to distinguish valid connection attempts from those performing Denial of Service (DoS) attacks.

An alternative approach would be as follows. A host would issue a connection request to the first IP address in the list returned by the name-to-address mapping function. If this connection request didn't succeed in some time, a connection request to the second IP address in the list would be issued in parallel. If none of these connection requests succeeded in some time, and there were still more addresses left in the list, they would be tried in the same way. While this approach would, in principle, avoid the problems of the previous approach, it might be hard to define the time interval to wait before issuing each parallel connection request. A short time interval would lead to the problems caused by the previous approach, while a long time interval would likely still lead to long delays in establishing a connection with the destination host. In any case, it must be noted that both approaches have the same drawbacks as the solution described in Section 2: they would increase application complexity, and would take too long to begin to be used by applications.

5. Security Considerations

To be included in future revisions of this document.

6. IANA Considerations

This document has no actions for IANA.

7. Acknowledgements

8. References

8.1. Normative References

[RFC1122] Braden, R., "Requirements for Internet Hosts - Communication Layers", STD 3, RFC 1122, October 1989.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

8.2. Informative References

[I-D.ietf-tcpm-tcp-soft-errors] Gont, F., "TCP's Reaction to Soft Errors", draft-ietf-tcpm-tcp-soft-errors-09 (work in progress), November 2008.

[Stevens2] Wright, G. and W. Stevens, "TCP/IP Illustrated, Volume 2: The Implementation", Addison-Wesley , 1994.

Author's Address

Fernando Gont
Universidad Tecnologica Nacional / Facultad Regional Haedo
Evaristo Carriego 2644
Haedo, Provincia de Buenos Aires 1706
Argentina

Phone: +54 11 4650 8472
Email: fernando@gont.com.ar
URI: http://www.gont.com.ar