Network Working Group
Request for Comments: 5597
BCP: 150
Category: Best Current Practice
R. Denis-Courmont
VideoLAN project
September 2009

Network Address Translation (NAT) Behavioral Requirements for the Datagram Congestion Control Protocol

Abstract

This document defines a set of requirements for NATs handling the Datagram Congestion Control Protocol (DCCP). These requirements allow DCCP applications, such as streaming applications, to operate consistently, and they are very similar to the TCP requirements for NATs, which have already been published by the IETF. Ensuring that NATs meet this set of requirements will greatly increase the likelihood that applications using DCCP will function properly.

Status of This Memo

Copyright and License Notice

Table of Contents

1. Introduction

For historical reasons, NAT devices are not typically capable of handling datagrams and flows for applications that use the Datagram Congestion Control Protocol (DCCP) [RFC4340].

This memo discusses the technical issues involved and proposes a set of requirements for NAT devices to handle DCCP in a way that enables communications when either or both of the DCCP endpoints are located behind one or more NAT devices. All definitions and requirements in [RFC4787] are inherited here. The requirements are otherwise designed similarly to those in [RFC5382], from which this memo borrows its structure and much of its content.

Note however that, if both endpoints are hindered by NAT devices, the normal model for DCCP of asymmetric connection will not work. A simultaneous-open must be performed, as in [RFC5596]. Also, a separate, unspecified mechanism may be needed, such as Unilateral Self Address Fixing (UNSAF) [RFC3424] protocols, if an endpoint needs to learn its own external NAT mappings.

2. Definitions

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].

This document uses the term "DCCP connection" to refer to individual DCCP flows, as uniquely identified by the quadruple (source and destination IP addresses and DCCP ports) at a given time.

This document uses the term "NAT mapping" to refer to a state at the NAT that is necessary for network address and port translation of DCCP connections. This document also uses the terms "endpoint-independent mapping", "address-dependent mapping", "address and port-dependent mapping", "filtering behavior", "endpoint-independent filtering", "address-dependent filtering", "address and port-dependent filtering", "port assignment", "port overloading", "hairpinning", and "external source IP address and port" as defined in [RFC4787].

3. Applicability Statement

This document applies to NAT devices that want to handle DCCP datagrams. It is not the intent of this document to deprecate the overwhelming majority of deployed NAT devices. These NATs are simply not expected to handle DCCP, so this memo is not applicable to them.

Expected NAT behaviors applicable to DCCP connections are very similar to those applicable to TCP connections (with the exception of REQ-6 below). The following requirements are discussed and justified extensively in [RFC5382]. These justifications are not reproduced here for the sake of brevity.

In addition to the usual changes to the IP header (in particular, the IP addresses), NAT devices need to mangle:

- the DCCP source port for outgoing packets, depending on the NAT mapping,

- the DCCP destination port for incoming packets, depending on the NAT mapping, and

- the DCCP checksum, to compensate for IP address and port number modifications.

Because changing the source or destination IP address of a DCCP packet will normally invalidate the DCCP checksum, it is not possible to use DCCP through a NAT without dedicated support. Some NAT devices are known to provide "generic" transport-protocol support, whereby only the IP header is mangled. That scheme is not sufficient to support DCCP.

4. DCCP Connection Initiation

4.1. Address and Port Mapping Behavior

A NAT uses a mapping to translate packets for each DCCP connection. A mapping is dynamically allocated for connections initiated from the internal side, and is potentially reused for certain subsequent connections. NAT behavior regarding when a mapping can be reused differs for different NATs, as described in [RFC4787].

REQ-1: A NAT MUST have an "Endpoint-Independent Mapping" behavior for DCCP.

4.2. Established Connections

REQ-2: A NAT MUST support all valid sequences of DCCP packets (defined in [RFC4340] and its updates) for connections initiated both internally as well as externally when the connection is permitted by the NAT. In particular, in addition to handling the DCCP 3-way handshake mode of connection initiation, A NAT MUST handle the DCCP simultaneous-open mode of connection initiation, defined in [RFC5596]. That mode updates DCCP by adding a new packet type: DCCP-Listen. The DCCP-Listen packet communicates the information necessary to uniquely identify a DCCP session. NATs may utilise the connection information (address, port, Service Code) to establish local forwarding state.

4.3. Externally Initiated Connections

REQ-3: If application transparency is most important, it is RECOMMENDED that a NAT have an "Endpoint-independent filtering" behavior for DCCP. If a more stringent filtering behavior is most important, it is RECOMMENDED that a NAT have an "Address-dependent filtering" behavior for DCCP.

- o The filtering behavior MAY be an option configurable by the administrator of the NAT.

- o The filtering behavior for DCCP MAY be independent of the filtering behavior for any other transport-layer protocol, such as UDP, UDP-Lite, TCP, and SCTP (Stream Control Transmission Protocol).

REQ-4: A NAT MUST wait for at least 6 seconds from the reception of an unsolicited, inbound DCCP-Listen or DCCP-Sync packet before it may respond with an ICMP Port Unreachable error, an ICMP Protocol Unreachable error, or a DCCP-Reset. If, during this interval, the NAT receives and translates an outbound DCCP-Request packet for the connection, the NAT MUST silently drop the original unsolicited, inbound DCCP-Listen packet. Otherwise, the NAT SHOULD send an ICMP Port Unreachable error (Type 3, Code 3) for the original DCCP-Listen unless the security policy forbids it.

5. NAT Session Refresh

The "established connection idle-timeout" for a NAT is defined as the minimum time a DCCP connection in the established phase must remain idle before the NAT considers the associated session a candidate for removal. The "transitory connection idle-timeout" for a NAT is defined as the minimum time a DCCP connection in the CLOSEREQ or CLOSING phases must remain idle before the NAT considers the associated session a candidate for removal. DCCP connections in the TIMEWAIT state are not affected by the "transitory connection idle-timeout".

REQ-5: If a NAT cannot determine whether the endpoints of a DCCP connection are active, it MAY abandon the session if it has been idle for some time. Where a NAT implements session timeouts, the default value of the "established connection idle-timeout" MUST be of 124 minutes or longer, and the default value of the "transitory connection idle-timeout" MUST be of 4 minutes or longer. A NAT that implements session timeouts may be configurable to use smaller values for the NAT idle-timeouts.

NAT behavior for handling DCCP-Reset packets or connections in the TIMEWAIT state is left unspecified.

6. Application-Level Gateways

Contrary to TCP, DCCP is a loss-tolerant protocol. Therefore, modifying the payload of DCCP packets may present a significant additional challenge in maintaining any application-layer state needed for an Application Level Gateway (ALG) to function properly. Additionally, there are no known DCCP-capable ALGs at the time of writing this document.

REQ-6: If a NAT includes ALGs, these ALGs MUST NOT affect DCCP.

NOTE: This is not consistent with REQ-6 of [RFC5382].

7. Other Requirements Applicable to DCCP

A list of general and UDP-specific NAT behavioral requirements are described in [RFC4787]. A list of ICMP-specific NAT behavioral requirements are described in [RFC5508]. The requirements listed below reiterate the requirements from these two documents that directly affect DCCP. The following requirements do not relax any requirements in [RFC4787] or [RFC5508].

7.1. Port Assignment

REQ-7: A NAT MUST NOT have a "Port assignment" behavior of "Port overloading" for DCCP.

7.2. Hairpinning Behavior

REQ-8: A NAT MUST support "hairpinning" for DCCP. Furthermore, a NAT's hairpinning behavior MUST be of type "External source IP address and port".

7.3. ICMP Responses to DCCP Packets

REQ-9: If a NAT translates DCCP, it SHOULD translate ICMP Destination Unreachable (Type 3) messages.

REQ-10: Receipt of any sort of ICMP message MUST NOT terminate the NAT mapping or DCCP connection for which the ICMP was generated.

8. Requirements Specific to DCCP

8.1. Partial Checksum Coverage

DCCP supports partial checksum coverage. A NAT will usually need to perform incremental changes to the packet Checksum field, as for other IETF-defined protocols. However, if it needs to recalculate a correct checksum value, it must take the checksum coverage into account, as described in Section 9.2 of [RFC4340].

REQ-11: If a NAT translates a DCCP packet with a valid DCCP checksum, it MUST ensure that the DCCP checksum is translated such that it is valid after the translation.

REQ-12: A NAT MUST NOT modify the value of the DCCP Checksum Coverage.

The Checksum Coverage field in the DCCP header determines the parts of the packet that are covered by the Checksum field. This always includes the DCCP header and options, but some or all of the application data may be excluded as determined on a packet-by-packet basis by the application. Changing the Checksum Coverage in the network violates the integrity assumptions at the receiver and may result in unpredictable or incorrect application behaviour.

8.2. Services Codes

DCCP specifies a Service Code as a 4-byte value (32 bits) that describes the application-level service to which a client application wishes to connect [RFC4340].

REQ-13: If a NAT translates a DCCP packet, it MUST NOT modify its DCCP Service Code value.

Further guidance on the use of Service Codes by middleboxes, including NATs, can be found in [RFC5595].

9. DCCP without NAT Support

If the NAT device cannot be updated to support DCCP, DCCP datagrams can be encapsulated within a UDP transport header. Indeed, most NAT devices are already capable of handling UDP. This is however beyond the scope of this document.

10. Security Considerations

[RFC4787] discusses security considerations for NATs that handle IP and unicast (UDP) traffic, all of which apply equally to this document. Security concerns specific to handling DCCP packets are discussed in this section.

REQ-1 and REQ-6 through REQ-13 do not introduce any new known security concerns.

REQ-2 does not introduce any new known security concerns. While a NAT may elect to keep track of some DCCP-specific, per-flow state (compared to UDP), it has no obligations to do so.

REQ-3 allows a NAT to adopt either a more secure or a more application-transparent filtering policy. This is already addressed in [RFC4787] and [RFC5382].

Similar to [RFC5382], REQ-4 of this document recommends that a NAT respond to unsolicited, inbound Listen and Sync packets with an ICMP error delayed by a few seconds. Doing so may reveal the presence of a NAT to an external attacker. Silently dropping the Listen makes it harder to diagnose network problems and forces applications to wait for the DCCP stack to finish several retransmissions before reporting an error. An implementer must therefore understand and carefully weigh the effects of not sending an ICMP error or rate-limiting such ICMP errors to a very small number.

REQ-5 recommends that a NAT that passively monitors DCCP state keep idle sessions alive for at least 124 minutes or 4 minutes, depending on the state of the connection. To protect against denial-of-service attacks filling its state storage capacity, a NAT may attempt to actively determine the liveliness of a DCCP connection, or the NAT administrator could configure more conservative timeouts.

11. Acknowledgments

The author would like to thank Gorry Fairhurst, Eddie Kohler, Dan Wing, Alfred Hoenes, Magnus Westerlund, Miguel Garcia, Catherine Meadows, Tim Polk, Lars Eggert, and Christian Vogt for their comments and help on this document.

This memo borrows heavily from [RFC5382] by S. Guha (editor), K. Biswas, B. Ford, S. Sivakumar, and P. Srisuresh.

12. References

12.1. Normative References

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

[RFC4340] Kohler, E., Handley, M., and S. Floyd, "Datagram Congestion Control Protocol (DCCP)", RFC 4340, March 2006.

[RFC4787] Audet, F. and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP", BCP 127, RFC 4787, January 2007.

[RFC5508] Srisuresh, P., Ford, B., Sivakumar, S., and S. Guha, "NAT Behavioral Requirements for ICMP", BCP 148, RFC 5508, April 2009.

[RFC5596] Fairhurst, G., "Datagram Congestion Control Protocol (DCCP) Simultaneous-Open Technique to Facilitate NAT/ Middlebox Traversal", RFC 5596, September 2009.

12.2. Informative References

[RFC3424] Daigle, L. and IAB, "IAB Considerations for UNilateral Self-Address Fixing (UNSAF) Across Network Address Translation", RFC 3424, November 2002.

[RFC5382] Guha, S., Biswas, K., Ford, B., Sivakumar, S., and P. Srisuresh, "NAT Behavioral Requirements for TCP", BCP 142, RFC 5382, October 2008.

[RFC5595] Fairhurst, G., "The Datagram Congestion Control Protocol (DCCP) Service Codes", RFC 5595, September 2009.

Author's Address

Remi Denis-Courmont
VideoLAN project

EMail: rem@videolan.org
URI: http://www.videolan.org/