Behavior Engineering for Hindrance Avoidance (if taken)
Internet-Draft
Intended status: Standards Track
Expires: August 20, 2008

R. Denis-Courmont
VideoLAN project

February 17, 2008

Network Address Translation (NAT) Behavioral Requirements for DCCP
draft-denis-behave-nat-dccp-01.txt

Status of This Memo

Copyright Notice

Abstract

This document defines a set of requirements for NATs that handle DCCP.

Table of Contents

1. Introduction

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

This draft discusses the technical issues involved, and proposes a set of requirements for NAT devices to handle DCCP in a way that enables 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 [I-D.ietf-behave-tcp], 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 of asymmetric connection model of DCCP will not work. A simultaneous open must be performed, as in [I-D.fairhurst-dccp-behave-update]. Also, a separate unspecified mechanism may be needed, such as UNSAF 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 documentation uses the term "DCCP connection" to refer to invidual DCCP flows, as uniquely identified by the the 4-tuple (source and destination IP addresses and DCCP ports) at a given time.

This document uses the term "NAT mapping" to refer to state at the NAT 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 or REQ-6 below). The following requirements are discussed and justified extensively in [I-D.ietf-behave-tcp]. 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
- the DCCP checksum, to compensate for IP address and port number modifications.

Because changing the 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 a "generic" transport protocol support, whereby only the IP header is mangled. That scheme is not sufficient to support DCCP in any case.

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 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. Internally Initiated Connections

FIXME/TBD: may change as DCCP simultaneous open progresses.

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 (FIXME: currently work-in-progress in the DCCP working group).

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.

- The filtering behavior MAY be an option configurable by the administrator of the NAT.
- The filtering behavior for TCP MAY be independent of thefiltering behavior for UDP.

REQ-4: A NAT MUST NOT respond to an unsolicited inbound DCCP-Request (TBD: add stuff for DCCP simultaneous open) packet for at least 6 seconds after the packet is received. If during this interval the NAT receives and translates an outbound DCCP packet (TBD: DCCP packet type) for the connection the NAT MUST silently drop the original unsolicited inbound DCCP-Request packet. Otherwise the NAT SHOULD send an ICMP Port Unreachable error (Type 3, Code 3) for the original DCCP-Request, 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. In such cases, the value of the "established connection idle-timeout" MUST NOT be less than 2 hours 4 minutes. The value of the "transitory connection idle-timeout" MUST NOT be less than 4 minutes. The value of the NAT idle-timeouts MAY be configurable.

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

6. Application Level Gateways

Contraty to TCP, DCCP is a loss-tolerant protocol. Therefore, modifying the payload of DCCP packets present a significant

additionnal challenge in maintaining sane DCCP sequence numbers, if the size of the payload were altered. Also, there are no known DCCP-capable Application Level Gateways (ALGs) at the time of writing this document.

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

NOTE: This is not consistent with REQ-6 of [I-D.ietf-behave-tcp].

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 [I-D.ietf-behave-nat-icmp]. The requirements listed below reiterate the requirements from these two documents that directly affect DCCP. The following requirements do not relax anyrequirements in [RFC4787] or [I-D.ietf-behave-nat-icmp].

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. Futhermore, 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. DCCP without NAT support

If the NAT device cannot be updated to support DCCP, DCCP datagram could be encapsulated within an additionnal UDP transport header. Indeed, most NAT devices are already capable of handling UDP.

There are significant disadvantages to this approach:

- Both sides of the DCCP session need must be updated to use tunnelling, even though only one side might be hindered with a NAT.

- o A method MUST be defined to negociate when to use tunnelling.
- o The per-packet overhead is increased.

A DCCP transport-specific solution is specified by [I-D.phelan-dccp-natencap]. Alternatively, existing IP tunneling protocols, such as ESP-in-UDP[RFC3948] (especially with the NULL cipher suite) or Teredo[RFC4380], could be used.

9. DCCP simultaneous open

When both parties to an intended DCCP session are located behind either a NAT device or a stateful firewall, neither can act as the paassive endpoint in the connection establishment.

Unfortunately, at the time of writing, the DCCP connection state machine does not allow both peers to behave as active endpoint, as is the case in TCP simultaneous open. It is expected that this issue will be tackled in the DCCP working group shortly (TODO: reference relevant I-D).

10. Security Considerations

TBD.

11. IANA Considerations

This document raises no IANA considerations.

12. Acknowledgments

The author would like to thank ... for their comments on this document.

This memo borrows heavily from draft-ietf-behave-tcp-07, by S. Guha (editor), K. Biswas, B. Ford, S. Sivakumar and P. Srisuresh.

13. References

13.1. Normative References

[I-D.ietf-behave-nat-icmp] Srisuresh, P., Ford, B., Sivakumar, S., and S. Guha, "NAT Behavioral Requirements for ICMP protocol", draft-ietf-behave-nat-icmp-07 (work in progress), February 2008.

[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.

13.2. Informative References

[I-D.fairhurst-dccp-behave-update] Fairhurst, G. and G. Renker, "An Update for DCCP Connection Establishment to Assist NAT & Firewall Traversal", draft-fairhurst-dccp-behave-update-01 (work in progress), November 2007.

[I-D.ietf-behave-tcp] Guha, S., "NAT Behavioral Requirements for TCP", draft-ietf-behave-tcp-07 (work in progress), April 2007.

[I-D.phelan-dccp-natencap] Phelan, T., "Datagram Congestion Control Protocol (DCCP) Encapsulation for NAT Traversal (DCCP-NAT)", draft-phelan-dccp-natencap-00 (work in progress), February 2008.

[RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M. Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC 3948, January 2005.

[RFC4380] Huitema, C., "Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs)", RFC 4380, February 2006.

Author's Address

Remi Denis-Courmont
VideoLAN project

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

Full Copyright Statement

Intellectual Property

Acknowledgement

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).