DCCP Simultaneous-Open Technique:

A journey through middleboxes

Gerrit Renker

Godred Fairhurst

draft-ietf-dccp-simul-open-00

Outline

1. Background Information (brief)

2. Changes Since Last Revision

3. Further Work

Background Information

- existing DCCP:
  - NAT traversal similar to TCP
  - but no simultaneous-open
  - hence has to rely on relay or tunnel
- this draft:
  - produces *effect* of simultaneous-open
  - without changes to standard DCCP operation
  - *indicator packet* to establish middlebox state
  - may help with *traversal of other middleboxes*

Server initialisation

- server initially needs to know ...
  - client IP-address:port and
  - client's chosen service code
- possible ways: SDP, RTSP, SIP, etc.
- to populate the 5-tuple ...
  <src-IP, src-port, Service Code, dst-IP, dst-port>
- ... which is then communicated to NAT

DCCP-Listen Packet

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Source Port          |           Dest Port           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Data Offset=5 | CCVal | CsCov |           Checksum            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Res | Type  |1|   Reserved    |  Sequence Number High Bits    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Sequence Number Low Bits                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Service Code                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

DCCP-Listen Packet

- intended for NAT, ignored by endpoints
- reuses DCCP-Request layout
- differerent Type value, because ...
  - avoid *semantic overloading* of packet types
  - RFC4340-clients may reply with Reset or Sync
  - would require *complicated processing rules*
- Format:
  - no payload / options carried on DCCP-Listen
  - *reuses 48-bit generic-header variant* (X=1)
  - sequence number values are ignored

Standard DCCP passive-open

```
unspecified remote            +---------+
      +-----------------------| CLOSED  |
      |                       +---------+
      |
      |
      v
+---------+
| LISTEN  |
+---------+
```

Standard DCCP passive-open

```
unspecified remote         +--------+
     +---------------------| CLOSED |
     |                     +--------+
     |
     |
     v
+--------+
| LISTEN |
+--------+
     |
     |
     |
     |
     |
     |
     |  receive Request   +---------+
     +------------------->| RESPOND |
        send Response     +---------+
```

Fully-specified passive-open

```
unspecified remote         +---------+   fully specified remote
     +---------------------| CLOSED |---------------------+
                           +---------+
     |
     |
     |
     v
+---------+
| LISTEN |
+---------+
     |
     |
     |
     |
     |
     |
     |  receive Request   +----------+
     +------------------->| RESPOND |
        send Response     +----------+
```

Fully-specified passive-open

```
unspecified remote         +---------+   fully specified remote
     +---------------------| CLOSED  |-------------------------+
     |                     +---------+   send DCCP-Listen      |
     |                                                         |
     |                                                         |
     v                                                         v
+---------+
| LISTEN  |
+---------+
     |
     |
     |
     |
     |
     |
     |  receive Request   +----------+
     +------------------->| RESPOND  |
        send Response     +----------+
```

Fully-specified passive-open

```
unspecified remote         +--------+   fully specified remote
     +---------------------| CLOSED |---------------------+
     |                     +--------+   send DCCP-Listen  |
     |                                                    |
     |                                                    |
     v                                                    v
+--------+                                           +---------+
| LISTEN |                                           | INVITED |
+--------+                                           +---------+
     |
     |
     |
     |
     |
     |
     |  receive Request   +---------+
     +------------------->| RESPOND |
        send Response     +---------+
```

Fully-specified passive-open

```
unspecified remote          +---------+  fully specified remote
     +----------------------| CLOSED  |--------------------+
     |                      +---------+  send DCCP-Listen  |
     |                                                     |
     |                                                     |
     v                                                     v
+---------+                                           +----------+
| LISTEN  |                                           | INVITED  |
+---------+                                           +----------+
     |                                                     |
     |                                                     |
     |                                                     |
     |                                                     |
     |                                                     |
     |                                                     |
     |  receive Request        +----------+  receive Request  |
     +--------------------->| RESPOND  |<-------------------+
        send Response          +----------+  send Response
```

What if DCCP-Listen is lost?

```
unspecified remote          +---------+   fully specified remote
     +----------------------| CLOSED  |----------------------+
     |                      +---------+   send DCCP-Listen   |
     |                                                        |
     |                                                        |
     |                                                        |
     v                                                        v
+---------+                                            +----------+
| LISTEN  |                                            | INVITED  |
+---------+                                            +----------+
     |                                                        |
     |                                                        |
     |                                                        |
     |                                                        |
     |                                                        |
     |                                                        |
     |  receive Request    +----------+    receive Request   |
     +-------------------->| RESPOND  |<---------------------+
        send Response      +----------+    send Response
```

What if DCCP-Listen is lost?

```
unspecified remote         +--------+   fully specified remote
     +---------------------| CLOSED |---------------------+
     |                     +--------+   send DCCP-Listen  |
     |                                                    |
     |                                                    |
     v                                                    v
+---------+                                  timeout +---------+
| LISTEN  |                              +-----------| INVITED |
+---------+                              |           +---------+
     |                                   |             ^  |
     |                                   |             |  |
     |                                   +-------------+  |
     |                                     resend Listen  |
     |                                                    |
     |                                                    |
     |  receive Request   +---------+    receive Request  |
     +-------------------->| RESPOND |<--------------------+
        send Response     +---------+    send Response
```

... or lost again?

```
unspecified remote         +---------+  fully specified remote
     +---------------------| CLOSED  |--------------------+
     |                     +---------+  send DCCP-Listen  |
     |                                                    |
     |                                                    |
     v                                                    v
+--------+                                  timeout  +---------+
| LISTEN |                               +-----------| INVITED |
+--------+                               |           +---------+
     |                                   |  1st / 2nd  ^  |
     |                                   |  retransm.  |  |
     |                                   +-------------+  |
     |                                    resend Listen   |
     |                                                    |
     |                                                    |
     |  receive Request   +---------+    receive Request  |
     +-------------------->| RESPOND |<--------------------+
        send Response     +---------+    send Response
```

Fallback after 2 retransmissions

```
unspecified remote         +--------+   fully specified remote
     +---------------------| CLOSED |---------------------+
     |                     +--------+   send DCCP-Listen  |
     |                                                    |
     |                                                    |
     |                                                    |
     v                                                    v
+--------+                                  timeout  +---------+
| LISTEN |<------------------------------+-----------| INVITED |
+--------+  more than 2 retransmissions  |           +---------+
     |                                   |  1st / 2nd  ^  |
     |                                   |  retransm.  |  |
     |                                   +-------------+  |
     |                                    resend Listen   |
     |                                                    |
     |                                                    |
     |  receive Request   +---------+    receive Request  |
     +------------------->| RESPOND |<--------------------+
        send Response     +---------+    send Response
```

Fallback after 2 retransmissions

```
unspecified remote         +--------+   fully specified remote
     +---------------------| CLOSED |---------------------+
     |                     +--------+   send DCCP-Listen  |
     |                                                    |
     |                                                    |
     v                                                    v
+--------+                                  timeout  +---------+
| LISTEN |<------------------------------+-----------| INVITED |
+--------+  more than 2 retransmissions  |           +---------+
     |                                   |  1st / 2nd  ^  |
     |                                   |  retransm.  |  |
     |                                   +-------------+  |
     |                                    resend Listen   |
     |                                                    |
     |                                                    |
     |  receive Request   +---------+    receive Request  |
     +------------------->| RESPOND |<--------------------+
        send Response     +---------+    send Response
```

The extension is conditional

```
unspecified remote         +--------+   fully specified remote
     +---------------------| CLOSED |---------------------+
     |                     +--------+   send DCCP-Listen  |
     |                                                    |
     |                                                    |
     v                                                    v
+--------+                                  timeout  +---------+
| LISTEN |<------------------------------+-----------| INVITED |
+--------+  more than 2 retransmissions  |           +---------+
     |                                   |  1st / 2nd  ^  |
     |                                   |  retransm.  |  |
     |                                   +-------------+  |
     |                                    resend Listen   |
     |                                                    |
     |                                                    |
     |  receive Request   +---------+    receive Request  |
     +------------------->| RESPOND |<--------------------+
        send Response     +---------+    send Response
```

The extension is conditional

```
unspecified remote         +---------+   fully specified remote
     +---------------------| CLOSED  |----------------------+
     |                     +---------+   send DCCP-Listen   |
     |                                                      |
     |                                                      |
     v                                                      v
+---------+                                 timeout   +----------+
| LISTEN |<-------------------------------+-----------| INVITED  |
+---------+ more than 2 retransmissions   |           +----------+
     |                                    | 1st / 2nd   ^  |
     |                                    | retransm.   |  |
     |                                    +-------------+  |
     |                                     resend Listen   |
     |                                                     |
     |                                                     |
     |  receive Request  +----------+   receive Request    |
     +------------------>| RESPOND  |<---------------------+
        send Response    +----------+   send Response
```

The extension is conditional

```
unspecified remote         +--------+   fully specified remote
     +---------------------| CLOSED |---------------------+
     |                     +--------+   send DCCP-Listen  |
     |                                                    |
     |                                                    |
     v                                                    v
+--------+                                  timeout  +---------+
| LISTEN |<------------------------------+-----------| INVITED |
+--------+  more than 2 retransmissions  |           +---------+
     |                                   |  1st / 2nd  ^  |
     |                                   |  retransm.  |  |
     |                                   +-------------+  |
     |                                    resend Listen   |
     |                                                    |
     |                                                    |
     |  receive Request   +---------+    receive Request  |
     +------------------->| RESPOND |<--------------------+
        send Response     +---------+    send Response
```

Retransmission of DCCP-Listen

- Requirement: *establish state before client retransmits DCCP-Request (1 second)*
- Goals:

  1) keep it simple (soft-state only)

  2) avoid correlated loss

  3) longer RTT: middlebox may not be next hop

  4) provide fall-back if everything else fails

Retransmission of DCCP-Listen

- Requirement: *establish state before client retransmits DCCP-Request (1 second)*
- Goals:

  1) keep it simple (soft-state only)

  2) avoid correlated loss

  3) longer RTT: middlebox may not be next hop

  4) provide fall-back if everything else fails

- Chosen solution:
  - 2 retransmissions, 200 msec apart
  - after 600 msec, fall back to standard LISTEN

Traversal Implementation

- Linux DCCP NAT module
  - current work-in-progress (Patrick McHardy)
  - quite likely available in upcoming 2.6.26
  - UDP-Lite NAT traversal already works!
- Linux kernel support for DCCP-Listen
  - seems straightforward
  - testing as soon as NAT module available

Traversal Implementation

- Linux DCCP NAT module
  - current work-in-progress (Patrick McHardy)
  - quite likely available in upcoming 2.6.26
  - UDP-Lite NAT traversal already works!
- Linux kernel support for DCCP-Listen
  - seems straightforward
  - testing as soon as NAT module available
- We are interested in other / further implementations!

Conclusions

- no change to standard DCCP operation
  - conditional extension
  - communicate all state to middlebox
  - simple retransmission scheme
- may work for other middleboxes, too:
  - interested in implementation
  - solicit feedback regarding wider applicability