            Profile for Datagram Congestion Control Protocol (DCCP)
     Congestion ID 4: TCP-Friendly Rate Control for Small Packets (TFRC-SP)
                       draft-ietf-dccp-ccid4-05.txt

Status of This Memo

   This Internet-Draft will expire on 2 January 2010.

Copyright Notice

Abstract

   This document specifies a profile for Congestion Control Identifier
   4, the Small-Packet variant of TCP-Friendly Rate Control (TFRC), in
   the Datagram Congestion Control Protocol (DCCP). CCID 4 is for
   experimental use, and uses TFRC-SP [RFC4828], a variant of TFRC
   designed for applications that send small packets. CCID 4 is
   considered experimental because TFRC-SP is itself experimental, and
   is not proposed for widespread deployment in the global Internet at
   this time. The goal for TFRC-SP is to achieve roughly the same
   bandwidth in bits per second (bps) as a TCP flow using packets of up
   to 1500 bytes but experiencing the same level of congestion. CCID 4
   is for use for senders that send small packets and would like a TCP-
   friendly sending rate, possibly with Explicit Congestion Notification
   (ECN), while minimizing abrupt rate changes.

Table of Contents

TO BE DELETED BY THE RFC EDITOR UPON PUBLICATION:

Changes from draft-ietf-dccp-ccid4-04.txt:

* Added captions to figures. IESG General Area review
  by Miguel Garcia.

* Added a sentence about the need for a ccid codepoint
  for experimentation across the Internet with different
  codebases. Feedback from Miguel Garcia and Adrian Farrel
  from IESG review.

Changes from draft-ietf-dccp-ccid4-03.txt:

* Minor editing from Working Group Last Call. This includes
clarification of the use of RFC 5348 instead of RFC 3448. Feedback
from Arjuna Sathiaseelan and Gorry Fairhurst.

Changes from draft-ietf-dccp-ccid4-02.txt:

* Updated boilerplates.

* Updated to refer to RFC 5348 instead of RFC 3448. RFC 5348 is now
required, instead of RFC 3448.

* Added a section on "Experimental Status of this Document."
Feedback from Gorry Fairhurst.

* Minor editing changes. Feedback from Gorry Fairhurst.

* Minor editing, from feedback from Alfred Hoenes.

Changes from draft-ietf-dccp-ccid4-01.txt:

* Added a Design Considerations section with a discussion on the
length of the Drop Count field in the Dropped Packets option, and on
the lengths of the fields in the Loss Intervals option. From
feedback in the Working Group.

Changes from draft-ietf-dccp-ccid4-00.txt:

* Added that the RFC 4342 errata applies to CCID 4 as well. From
email from Leandro Sales.

* Added the phrase "If the sender is calculating the loss event rate
itself" to a non-normative description in Section 5. Feedback from
Gerrit Renker.

* Deleted the Send Dropped Packets feature, since it is not used in
CCID 4. In CCID 4, the Dropped Packets option is mandatory.

Changes from draft-floyd-dccp-ccid4-01.txt:

* Title changed to draft-ietf-dccp-ccid4-00.txt.

* Incorporated material from draft-kohler-dccp-ccid3-drops-01.txt.

* Added a reference to RFC3448bis.

* Added a sentence saying that this is Experimental because TFRC-SP
is Experimental.

* General editing in response to feedback from Gorry.

Changes from draft-floyd-dccp-ccid4-00.txt:

* Added a subsection describing calculation of the average loss
interval in TFRC-SP.

* Changed the assumed DCCP-Data header size from 12 bytes to 16
bytes, for 48-bit sequence numbers. Feedback from Ian McDonald.

* Added that the CCID 4 sender can send two packets in a burst, if
limited by OS granularity. From Ian McDonald.

* Added that the implementor may track Faster Restart and implement
it before an explicit update to the CCID4 RFC. From Ian McDonald.

* Added an example to Section 8.4 of when errors can occur in using
the Window Counter to detect loss intervals of at most two round-trip
times.

Changes from draft-floyd-ccid4-00.txt:

* Added the Dropped Packets option for reporting the number of
packets dropped in a loss interval.

* Added examples to Section 8.4 of the receiver incorrectly inferring
whether a loss interval was short or not.

END OF SECTION TO BE DELETED.

1. Introduction

    This document specifies an experimental profile for Congestion
    Control Identifier 4, TCP-Friendly Rate Control for Small Packets
    (TFRC-SP), in the Datagram Congestion Control Protocol (DCCP)
    [RFC4340]. CCID 4 is a modified version of Congestion Control
    Identifier 3, CCID 3, which has been specified in [RFC4342]. This
    document assumes that the reader is familiar with CCID 3, instead of
    repeating from that document unnecessarily.

    CCID 3 uses TCP Friendly Rate Control (TFRC), which is now specified
    in RFC 5348 [RFC5348]. CCID 4 differs from CCID 3 in that CCID 4
    uses TFRC-SP [RFC4828], an experimental, small-packet variant of
    TFRC. The original specification of TFRC, RFC 3448 [RFC3448], has
    been obsoleted by RFC 5348. The CCID 3 and TFRC-SP documents both
    predate RFC 5348 and refer instead to RFC 3448 for the specification
    of TFRC. However, this document assumes that RFC 5348 will be used
    instead of RFC 3448 for the specification of TFRC.

    CCID 4 differs from CCID 3 only in the following respects:

    o Header size: For TFRC-SP, the allowed transmit rate in bytes per
       second is reduced by a factor that accounts for packet header
       size. This is specified for TFRC-SP in Section 4.2 of [RFC4828],
       and described for CCID 4 in Section 5 below.

    o Maximum sending rate: TFRC-SP enforces a minimum interval of
       10 milliseconds between data packets. This is specified for TFRC-
       SP in Section 4.3 of [RFC4828], and described for CCID 4 in
       Section 5 below.

    o Loss rates for short loss intervals: For short loss intervals of
       at most two round-trip times, the loss rate is computed by
       counting the actual number of packets lost or marked. For such a
       short loss interval with N data packets, including K lost or
       marked data packets, the loss interval length is calculated as
       N/K, instead of as N. This is specified for TFRC-SP in Section
       4.4 of [RFC4828]. If the sender is computing the loss event rate,
       the Dropped Packets option specified in Section 8.7 is required,
       in addition to the default CCID 3's Loss Intervals option.
       Section 8.7 describes the use of the Dropped Packets option in
       calculating the loss event rate. The computation of the loss rate
       by the receiver for the Loss Event Rate option is described for
       CCID 4 in Section 8.4 below.

    o The nominal segment size: In TFRC-SP, the nominal segment size
       used by the TCP throughput equation is set to 1460 bytes. This is
       specified for TFRC-SP in Section 4.5 of [RFC4828], and described

for CCID 4 in Section 5 below.

2. Conventions

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].

Additional terminology is described in Section 2 of [RFC4342].

3. Usage

Like CCID 3, CCID 4's congestion control is appropriate for flows
that would prefer to minimize abrupt changes in the sending rate,
including streaming media applications with small or moderate
receiver buffering before playback.

CCID 4 is designed to be used either by applications that use a small
fixed segment size, or by applications that change their sending rate
by varying the segment size. If CCID 4 is used by an application
that varies its segment size in response to changes in the allowed
sending rate in bps, we note that CCID 4 doesn't dictate the segment
size to be used by the application; this is done by the application
itself. The CCID 4 sender determines the allowed sending rate in
bps, in response to on-going feedback from the CCID 4 receiver, and
the application can use information about the current allowed sending
rate to decide whether to change the current segment size.

We note that in some environments there will be a feedback loop, with
changes in the packet size or in the sending rate in bps affecting
congestion along the path, therefore affecting the allowed sending
rate in the future.

3.1. Relationship with TFRC and TFRC-SP

The congestion control mechanisms described here follow the TFRC-SP
mechanism specified in [RFC4828]. As with CCID 3, conformant CCID 4
implementations MAY track updates to the TCP throughput equation
directly, as updates are standardized in the IETF, rather than
waiting for revisions of this document. This document is based on
CCID 3 [RFC4342], TFRC, and TFRC-SP. For TFRC, RFC 3448 [RFC3448]
has been obsoleted by RFC 5348 [RFC5348].

3.2. Example Half-Connection

This example shows the typical progress of a half-connection using
CCID 4's TFRC Congestion Control, not including connection initiation
and termination. The example is informative, not normative. This

example differs from that for CCID 3 in [RFC4342] only in one respect; with CCID 4 the allowed transmit rate is determined by [RFC4828] as well as by [RFC5348].

1. The sender transmits DCCP-Data packets, where the sending rate is governed by the allowed transmit rate as specified in [RFC4828]. Each DCCP-Data packet has a sequence number, and the DCCP header's CCVal field contains the window counter value, used by the receiver in determining when multiple losses belong in a single loss event.

   In the typical case of an ECN-capable half-connection, each DCCP-Data and DCCP-DataAck packet is sent as ECN-Capable, with either the ECT(0) or the ECT(1) codepoint set. The use of the ECN Nonce with TFRC is described in Section 9.

2. The receiver sends DCCP-Ack packets acknowledging the data packets at least once per round-trip time, unless the sender is sending at a rate of less than one packet per round-trip time [RFC5348] (Section 6). Each DCCP-Ack packet uses a sequence number, identifying the most recent packet received from the sender, and includes feedback about the recent loss intervals experienced by the receiver.

3. The sender continues sending DCCP-Data packets as controlled by the allowed transmit rate. Upon receiving DCCP-Ack packets, the sender updates its allowed transmit rate as specified in [RFC5348] (Section 4.3) and [RFC4828]. This update is based upon a loss event rate calculated by the sender, based on the receiver's loss intervals feedback. If it prefers, the sender can also use a loss event rate calculated and reported by the receiver.

4. The sender estimates round-trip times and calculates a nofeedback time, as specified in [RFC5348] (Section 4.4). If no feedback is received from the receiver in that time (at least four round-trip times), the sender halves its sending rate.

4. Connection Establishment

   The connection establishment is as specified in Section 4 of [RFC4342].

5. Congestion Control on Data Packets

   CCID 4 uses the congestion control mechanisms of TFRC [RFC5348] and TFRC-SP [RFC4828]. [RFC4828] MUST be considered normative except where specifically indicated.

Loss Event Rate

As with CCID 3, the basic operation of CCID 4 centers around the
calculation of a loss event rate: the number of loss events as a
fraction of the number of packets transmitted, weighted over the last
several loss intervals. For CCID 4, this loss event rate, a round-
trip time estimate, and a nominal packet size of 1460 bytes are
plugged into the TCP throughput equation, as specified in RFC 5348
(Section 3.1) and [RFC4828].

Because CCID 4 is intended for applications that send small packets,
the allowed transmit rate derived from the TCP throughput equation is
reduced by a factor that accounts for packet header size, as
specified in Section 4.2 of [RFC4828]. The header size on data
packets is estimated as 36 bytes (20 bytes for the IPv4 header, and
16 bytes for the DCCP-Data header with 48-bit sequence numbers). If
the DCCP sender is sending N-byte data packets, the allowed transmit
rate is reduced by N/(N+36). CCID 4 senders are limited to this fair
rate. The header size would be 32 bytes instead of 36 bytes when
24-bit sequence numbers were used in the DCCP-Data header.

As explained in Section 4.2 of [RFC4828], the actual header could be
larger or smaller than the assumed value, due to IP or DCCP options,
IPv6, IP tunnels, header compression, and the like. Because we are
only aiming at rough fairness, and at a rough incentive for
applications, the default use of a 32-byte or 36-byte header in the
calculations of the header bandwidth is sufficient for both IPv4 and
IPv6.

If the sender is calculating the loss event rate itself, the loss
event rate can be calculated using recent loss interval lengths
reported by the receiver. Loss intervals are precisely defined in
Section 6.1 of [RFC4342], with the modification in [RFC4828]
(Section 3) for loss intervals of at most two round-trip times. In
summary, a loss interval is up to 1 RTT of possibly lost or ECN-
marked data packets, followed by an arbitrary number of non-dropped,
non-marked data packets. The CCID 3 Loss Intervals option is used to
report loss interval lengths; see Section 8.6.

For loss intervals of at most two round-trip times, CCID 4 calculates
the loss event rate for that interval by counting the number of
packets lost or marked, as described in Section 4.4 of [RFC4828].
Thus, for such a short loss interval with N data packets, including K
lost or marked data packets, the loss interval length is calculated
as N/K, instead as N. The Dropped Packets option is used to report
K, the count of lost or marked data packets.

Unlike CCID 3, the CCID 4 sender enforces a minimum interval of 10 ms
between data packets, regardless of the allowed transmit rate. If
operating system scheduling granularity makes this impractical, up to
one additional packet MAY be sent per timeslice, providing that no
more than three packets are sent in any 30 ms interval.

Other Congestion Control Mechanisms

The other congestion control mechanisms such as slow-start and
feedback packets are exactly as in CCID 3, and are described in the
subsection on "Other Congestion Control Mechanisms" of Section 5 in
[RFC4342].

5.1. Response to Idle and Application-limited Periods

This is described in Section 5.1 of [RFC4342]. If Faster Restart is
standardized in the IETF for TFRC [KFS07], then Faster Restart MAY be
implemented in CCID4 without having to wait for an explicit update to
this document.

5.2. Response to Data Dropped and Slow Receiver

This is described in Section 5.2 of [RFC4342].

5.3. Packet Sizes

CCID 4 is intended for applications that use a fixed small segment
size, or that vary their segment size in response to congestion.

The CCID 4 sender uses a segment size of 1460 bytes in the TCP
throughput equation. This gives the CCID 4 sender roughly the same
sending rate in bytes per second as a TFRC flow using 1460-byte
segments but experiencing the same packet drop rate.

6. Acknowledgements

The acknowledgements are as specified in Section 6 of [RFC4342] with
the exception of the Loss Interval lengths specified below.

6.1. Loss Interval Definition

The loss interval definition is as defined in Section 6.1 of
[RFC4342], except as specified below. Section 6.1.1 of RFC 4342
specifies that for all loss intervals except the first one, the data
length equals the sequence length minus the number of non-data
packets the sender transmitted during the loss interval, with a
minimum data length of one packet. For short loss intervals of at
most two round-trip times, TFRC-SP computes the loss interval length

as the data length divided by the number of dropped or marked data
packets (rather than as the data length of the loss interval).

Section 5.4 of RFC 4342 describes when to use the most recent loss
interval in the calculation of the average loss interval. [RFC4828]
adds to this procedure the restriction that the most recent loss
interval is only used in the calculation of the average loss interval
if the most recent loss interval is greater than two round-trip
times. The pseudocode is given in Section 3 of [RFC4828].

6.2. Congestion Control on Acknowledgements

The congestion control on acknowledgements is as specified in Section
6.2 of [RFC4342].

6.3. Acknowledgements of Acknowledgements

Procedures for the acknowledgement of acknowledgements are as
specified in Section 6.3 of [RFC4342].

6.4. Quiescence

The procedure for detecting that the sender has gone quiescent is as
specified in Section 6.4 of [RFC4342].

7. Explicit Congestion Notification

Procedures for the use of Explicit Congestion Notification (ECN) are
as specified in Section 7 of [RFC4342].

8. Options and Features

CCID 4 can make use of DCCP's Ack Vector, Timestamp, Timestamp Echo,
and Elapsed Time options, and its Send Ack Vector and ECN Incapable
features. CCID 4 also imports the currently defined CCID-3-specific
options and features [RFC4342], augmented by the Dropped Packets
option specified in this document. Each CCID-4-specific option and
feature contains the same data as the corresponding CCID 3 option or
feature, and is interpreted in the same way, except as specified
elsewhere in this document (or in a subsequent IETF standards-track
RFC that updates or obsoletes this specification).

	Option		DCCP-	Section
Type	Length	Meaning	Data?	Reference
-----	------	-------	-----	---------
128-191		Reserved		
192	6	Loss Event Rate	N	8.5
193	variable	Loss Intervals	N	8.6
194	6	Receive Rate	N	8.3
195	variable	Dropped Packets	N	8.7
196-255		Reserved		

Table 1: DCCP CCID 4 Options

The "DCCP-Data?" column indicates that all currently defined
CCID-4-specific options MUST be ignored when they occur on DCCP-Data
packets.

As with CCID 3, the following CCID-specific features are also
defined.

		Rec'n	Initial		Section
Number	Meaning	Rule	Value	Req'd	Reference
------	-------	-----	-----	-----	---------
128-191	Reserved				
192	Send Loss Event Rate	SP	0	N	8.4
193-255	Reserved				

Table 2: DCCP CCID 4 Feature Numbers

More information is available in Section 8 of [RFC4342].

8.1. Window Counter Value

The use of the Window Counter Value in the DCCP generic header's
CCVal field is as specified in Section 8.1 of [RFC4342]. In addition
to their use described in CCID 3, the CCVal counters are used by the
receiver in CCID 4 to determine when the length of a loss interval is
at most two round-trip times. None of these procedures require the
receiver to maintain an explicit estimate of the round-trip time.
However, Section 8.1 of [RFC4342] gives a procedure that implementors
may use if they wish to keep such an RTT estimate using CCVal.

8.2. Elapsed Time Options

The use of the Elapsed Time option is defined in Section 8.2 of
[RFC4342].

8.3. Receive Rate Option

   The Receive Rate option is as specified in Section 8.3 of [RFC4342].

8.4. Send Loss Event Rate Feature

   The Send Loss Event Rate feature is as defined in Section 8.4 of
   [RFC4342].

   See [RFC5348], Section 5 and [RFC4828], Section 4.4 for a normative
   calculation of the loss event rate. Section 4.4 of [RFC4828]
   modifies the calculation of the loss interval size for loss intervals
   of at most two round-trip times.

   If the CCID 4 receiver is using the Loss Event Rate option, the
   receiver needs to be able to determine if a loss interval is short,
   of at most two round-trip times. The receiver can heuristically
   detect a short loss interval by using the Window Counter in arriving
   data packets. The sender increases the Window Counter by 1 every
   quarter of a round-trip time, with the caveat that the Window Counter
   is never increased by more than five, modulo 16, from one data packet
   to the next. Using the Window Counter to detect loss intervals of at
   most two round-trip times could result in some false positives, with
   some longer loss intervals incorrectly identified as short ones. For
   example, if the loss interval contained data packets with only two
   Window Counter values, say, k and k+5, then the receiver could not
   tell if the loss interval was at most two round-trip times long or
   not. Similarly, if the sender sent data packets with Window Counter
   values of 4, 8, 12, 0, 5, but the packets with Window Counter values
   of 8, 12, and 0 were lost in the network, then the receiver would
   only receive data packets with Window Counter values of 4 and 5, and
   would incorrectly infer that the loss interval was at most two round-
   trip times.

8.5. Loss Event Rate Option

   The Loss Event Rate option is as specified in Section 8.5 of
   [RFC4342].

   See [RFC5348] (Section 5) and [RFC4828] for a normative calculation
   of the loss event rate.

8.6. Loss Intervals Option

   The Loss Intervals option is as specified in Section 8.6 of
   [RFC4342].

8.7. Dropped Packets Option

   This section describes the Dropped Packets option, a mechanism for
   reporting the number of lost and marked packets per loss interval.
   By reporting both the Loss Intervals and Dropped Packets options on
   the feedback packets, the receiver gives the sender sufficient
   information to calculate the loss event rate, or to verify the
   calculation of the reported loss event rate, if the sender so
   desires.

   The core information reported by CCID 4 receivers is a list of recent
   loss intervals, where a loss interval begins with a lost or ECN-
   marked data packet; continues with at most one round-trip time's
   worth of packets that may or may not be lost or marked; and completes
   with an arbitrarily long series of non-dropped, non-marked data
   packets. Loss intervals model the congestion behavior of TCP NewReno
   senders, which reduce their sending rate at most once per window of
   data packets. Consequently, the number of packets lost in a loss
   interval is not important for either TCP's or TFRC's congestion
   response. CCID 3's Loss Intervals option reports the length of each
   loss interval's lossy part, not the number of packets that were
   actually lost or marked in that lossy part.

   However, for computing the loss event rate for periods that include
   short loss intervals the TFRC-SP sender needs to know the number of
   packets lost or marked in a loss interval, over and above the length
   of the loss interval in packets. The Dropped Packets option, a
   CCID-4-specific option, reports this information. Together with the
   existing Loss Intervals option, the Dropped Packets option allows the
   CCID 4 sender to discover exactly how many packets were dropped from
   each loss interval. The receiver reports the number of lost or
   marked packets in its recently observed loss intervals using the
   Dropped Packets option.

   The Dropped Packets Option is specified as follows:

```
   +--------+--------+-------...-------+--------+-------
   |11000011| Length |   Drop Count    | More Drop Counts...
   +--------+--------+-------...-------+--------+-------
    Type=195                  3 bytes
```

         Figure 1: Dropped Packets Option

   The Dropped Packets option contains information about one to 84
   consecutive loss intervals, always including the most recent loss
   interval. As with the Loss Intervals option, intervals are listed in
   reverse chronological order. Should more than 84 loss intervals need
   to be reported, multiple Dropped Packets options can be sent; the

second option begins where the first left off, and so forth.

One Drop Count is specified per loss interval. Drop Count is a
24-bit number that equals the number of packets lost or received ECN-
marked during the corresponding loss interval. By definition, this
number MUST NOT exceed the corresponding loss interval's Loss Length.

CCID 4 receivers MUST report Dropped Packets options with every
feedback packet. Any packet containing a Loss Intervals option MUST
also contain a Dropped Packets option covering the same loss
intervals. If a feedback packet does not include a relevant Dropped
Packets option, and the CCID 4 sender is computing the loss event
rate itself, the sender MUST treat the relevant loss intervals' Drop
Counts as equal to the corresponding Loss Lengths, as specified
below.

Consider a CCID 4 receiver. As specified in Section 8.6.1 of
RFC 4342, the receiver sends the Loss Intervals option for all
intervals that have not been acknowledged by the sender. When this
receiver sends a feedback packet containing information about the N
most recent loss intervals (packaged in one or more Loss Intervals
options), then the receiver includes on the same feedback packet one
or more Dropped Packets options covering exactly those N loss
intervals. CCID 4 senders MUST ignore Drop Counts information for
loss intervals not covered by a Loss Intervals option on the same
feedback packet. Conversely, a CCID 4 sender might want to
interpolate Drop Counts information for a loss interval not covered
by any Dropped Packets options; such a sender MUST use the
corresponding loss interval's Loss Length as its Drop Count.

Each loss interval's Drop Count MUST by definition be less than or
equal to its Loss Length. A Drop Count that exceeds the
corresponding Loss Length MUST be treated as equal to the Loss
Length.

8.7.1. Example

Consider the following sequence of packets, where "-" represents a
safely delivered packet and "*" represents a lost or marked packet.
This sequence is repeated from [RFC4342].

```
Sequence
 Numbers: 0          10        20        30        40  44
          |          |         |         |         |   |
          ----------*--------***-*--------*----------*-
```

     Figure 2: Sequence of Delivered (-) and Lost (*) Packets

Assuming that packet 43 was lost, not marked, this sequence might be
divided into loss intervals as follows:

```
        0          10         20         30        40  44
        |          |          |          |         |   |
        ----------*--------***-*--------*----------*-
        _____/_____/_____/_____/
            L0        L1          L2         L3
```
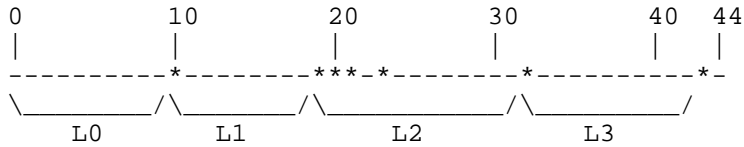
Figure 3: Loss Intervals for the Packet Sequence

A Loss Intervals option sent on a packet with Acknowledgement Number
44 to acknowledge this set of loss intervals might contain the bytes
193,39,2, 0,0,10, 128,0,1, 0,0,10, 0,0,8, 0,0,5, 0,0,10, 0,0,8,
0,0,1, 0,0,8, 0,0,10, 128,0,0, 0,0,15; for interpretation of this
option, see [RFC4342]. A Dropped Packets option sent in tandem on
this packet would contain the bytes 195,14, 0,0,1, 0,0,4, 0,0,1,
0,0,0. This is interpreted as follows.

195 The Dropped Packets option number.

14 The length of the option, including option type and length bytes.
    This option contains information about (14 - 2)/3 = 4 loss
    intervals. Note that the two most recent sequence numbers are
    not yet part of any loss interval -- the Loss Intervals option
    includes them in its Skip Length -- and are thus not included in
    the Dropped Packets option.

0,0,1
    These bytes define the Drop Count for L3, which is 1. As
    required, the Drop Count is less than or equal to L3's Loss
    Length, which is also 1.

0,0,4
    The Drop Count for L2 is 4.

0,0,1
    The Drop Count for L1 is 1.

0,0,0
    Finally, the Drop Count for L0 is 0.

9. Verifying Congestion Control Compliance With ECN

   Verifying congestion control compliance with ECN is as discussed in
   Section 9 of [RFC4342].

9.1. Verifying the ECN Nonce Echo

   Procedures for verifying the ECN Nonce Echo are as specified in
   Section 9.1 of [RFC4342].

9.2. Verifying the Reported Loss Intervals and Loss Event Rate

   Section 9.2 of [RFC4342] discusses the sender's possible verification
   of loss intervals and loss event rate information reported by the
   receiver.

10. Implementation Issues

10.1. Timestamp Usage

   The use of the Timestamp option is as discussed in Section 10.1 of
   [RFC4342].

10.2. Determining Loss Events at the Receiver

   The use of the window counter by the receiver to determine if
   multiple lost packets belong to the same loss event is as described
   in Section 10.2 of [RFC4342].

10.3. Sending Feedback Packets

   The procedure for sending feedback packets is as described in Section
   10.3 of [RFC4342].

11. Design Considerations

   This section discusses design considerations for the field sizes in
   the Loss Intervals and Dropped Packets Options.

11.1. The Field Size in the Loss Intervals Option

   Section 8.6 of RFC 4342 specifies a Loss Intervals Option with three
   fields for each loss interval, for reporting the Lossless Length,
   Loss Length, and Data Length. Each field is specified to be three
   bytes. Section 8.6 of this document specifies that CCID 4 use the
   same Loss Intervals Option as CCID 3, with the same field sizes.
   This has the significant advantage of minimizing the implementation
   differences between CCID 3 and CCID 4. However, it has been
   suggested that CCID 4 *could* use a Loss Intervals Option with
   smaller field sizes, since a CCID 4 sender enforces a minimum
   interval of 10 ms between data packets. This section explains the
   reason for CCID 4 to use the same Loss Intervals option as specified
   for CCID 3.

The Lossless Length field reports the number of packets in the loss
intervals' lossless part, and the Loss Length field reports the
number of packets in the loss interval's lossy part. The Data Length
field reports the number of packets in the loss interval's data
length (excluding non-data packets). A two-byte Data Length field
can report a data length of 65,536 packets, corresponding to a loss
event rate of 0.00002; this is enough to give the CCID 4 sender an
allowed sending rate of roughly 250 packets per RTT, which is enough
for a connection with a round-trip time of at most 2.5 seconds. For
a CCID 4 connection with a larger round-trip time, the three-byte
Lossless Length and Data Length fields would be needed.

For the Loss Length field in the Loss Intervals Option, reporting the
number of packets in the one-RTT lossy part of the loss interval, a
one-byte field would not be sufficient for a CCID 4 connection with a
long RTT (three seconds or longer). For the Loss Length field, a
two-byte field should be sufficient for CCID 4. However, our
judgement is that the advantages of using the same Loss Intervals
Option as in CCID 3 outweigh any advantages of using a CCID 4 Loss
Intervals Option that uses eight bytes instead of nine bytes for
reporting the fields for each loss interval.

11.2. The Field Size in the Dropped Packets Option

Section 8.7 specifies the Dropped Packets Option for reporting the
number of lost or marked packets per loss interval, allocating three
bytes for the drop count field for each loss interval reported. The
three-byte field is partly for simplicity, to give the same field
size as the fields in the Loss Intervals option specified in
RFC 4342. It has been suggested that CCID 4 *could* use a smaller
field size for the Dropped Packets option. This section discusses
the issue of the size of the drop count field in the Dropped Packets
Option.

It is not necessary to specify a three-byte field for the Dropped
Packets Option. A one-byte field would allow a reported drop count
of 255, and a two-byte field would allow a reported drop count of
65,535. A two-byte field would clearly be sufficient for the drop
count field for CCID 4.

In fact, a one-byte field would *probably* be adequate for reporting
the drop count for a loss interval in a CCID 4 connection. Because a
CCID 4 sender enforces a minimum interval of 10 ms between data
packets, a sender would need a round-trip time of over 2.55 seconds
to have more than 255 packets lost or marked in a single loss
interval; round-trip times of greater than three seconds are not
unusual for some flows traversing satellite links. The drop count
field is used in CCID 4 to compute the actual loss rate for short

loss intervals, rather than using the loss event rate that is used
for longer loss intervals. If a loss interval of at most two round-
trip times included N packets sent, with more than 255 of those
packets lost or marked, a drop count field of one byte would allow a
drop count of at most 255 to be reported, resulting in a computed
loss rate for that interval of 255/N. This loss rate might be less
than the actual loss rate, but it is significantly higher than the
loss event rate of 1/N, and should be sufficient to prevent a steady-
state condition of a CCID 4 connection with multiple packets dropped
each round-trip time. Thus, a one-byte field would probably be
adequate for reporting the drop count for a loss interval in a CCID 4
connection. However, at the moment this document specifies a three-
byte field, for consistency with the field size in the Loss Intervals
Option.

12. Experimental Status of this Document

   TFRC-SP is a congestion control mechanism defined in RFC 4828.
   Section 10 of [RFC4828] describes why TFRC-SP is currently specified
   as experimental and why it is not intended for widespread deployment
   at this time in the global Internet. Since TFRC-SP is experimental,
   CCID 4 is therefore also considered experimental. If the IETF
   publishes a standards-track RFC that changes the status of TFRC-SP,
   then CCID 4 should then be updated to reflect the change of status.

13. Security Considerations

   Security considerations include those discussed in Section 11 of
   [RFC4342]. There are no new security considerations introduced by
   CCID 4.

14. IANA Considerations

   This specification defines the value 4 in the DCCP CCID namespace
   managed by IANA. This is a permanent codepoint, as is needed for
   experimentation across the Internet using different codebases.

   CCID 4 also uses three sets of numbers whose values should be
   allocated by IANA, namely CCID-4-specific Reset Codes, option types,
   and feature numbers. This document makes no particular allocations
   from the Reset Code range, except for experimental and testing use
   [RFC3692]. We refer to the Standards Action policy outlined in
   [RFC5226].

14.1. Reset Codes

   Each entry in the DCCP CCID 4 Reset Code registry contains a
   CCID-4-specific Reset Code, which is a number in the range 128-255; a
   short description of the Reset Code; and a reference to the RFC
   defining the Reset Code. Reset Codes 184-190 and 248-254 are
   permanently reserved for experimental and testing use. The remaining
   Reset Codes -- 128-183, 191-247, and 255 -- are currently reserved,
   and should be allocated with the Standards Action policy, which
   requires IESG review and approval and standards-track IETF RFC
   publication.

14.2. Option Types

   Each entry in the DCCP CCID 4 option type registry contains a
   CCID-4-specific option type, which is a number in the range 128-255;
   the name of the option, such as "Loss Intervals"; and a reference to
   the RFC defining the option type. The registry is initially
   populated using the values in Table 1, in Section 8. This includes
   the value 195 allocated for the Dropped Packets option. This
   document allocates option types 192-195, and option types 184-190 and
   248-254 are permanently reserved for experimental and testing use.
   The remaining option types -- 128-183, 191, 196-247, and 255 -- are
   currently reserved, and should be allocated with the Standards Action
   policy, which requires IESG review and approval and standards-track
   IETF RFC publication.

14.3. Feature Numbers

   Each entry in the DCCP CCID 4 feature number registry contains a
   CCID-4-specific feature number, which is a number in the range
   128-255; the name of the feature, such as "Send Loss Event Rate"; and
   a reference to the RFC defining the feature number. The registry is
   initially populated using the values in Table 2, in Section 8. This
   document allocates feature number 192, and feature numbers 184-190
   and 248-254 are permanently reserved for experimental and testing
   use. The remaining feature numbers -- 128-183, 191, 193-247, and 255
   -- are currently reserved, and should be allocated with the Standards
   Action policy, which requires IESG review and approval and standards-
   track IETF RFC publication.

15. Thanks

   We thank Gorry Fairhurst, Alfred Hoenes, Ian McDonald, Gerrit Renker,
   and Leandro Sales for feedback on this document.

Normative References

   [RFC2119]   S. Bradner. Key Words For Use in RFCs to Indicate
                   Requirement Levels. RFC 2119.

   [RFC5226]   T. Narten and H. Alvestrand. Guidelines for Writing
                   an IANA Considerations Section in RFCs. RFC 5226.

   [RFC3448]   M. Handley, S. Floyd, J. Padhye, and J. Widmer, TCP
                   Friendly Rate Control (TFRC): Protocol Specification,
                   RFC 3448, Proposed Standard, January 2003. Obsoleted
                   by RFC 5348.

   [RFC3692]   T. Narten. Assigning Experimental and Testing Numbers
                   Considered Useful. RFC 3692.

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

   [RFC4342]   Floyd, S., Kohler, E., and J. Padhye. Profile for
                   Datagram Congestion Control Protocol (DCCP) Congestion
                   Control ID 3: TCP-Friendly Rate Control (TFRC),
                   RFC 4342, March 2006.

   [RFC4828]   S. Floyd and E. Kohler. TCP Friendly Rate Control
                   (TFRC): the Small-Packet (SP) Variant. RFC 4828,
                   April 2007.

   [RFC5348]   S. Floyd, M. Handley, J. Padhye, and J. Widmer, TCP
                   Friendly Rate Control (TFRC): Protocol Specification,
                   RFC 5348, Proposed Standard, September 2008.

Informative References

   [KFS07]   Kohler, E., S. Floyd, and A. Sathiaseelan, Faster
                   Restart for TCP Friendly Rate Control (TFRC),
                   Internet-draft draft-ietf-dccp-tfrc-faster-
                   restart-06.txt, work-in-progress, July 2008.

   [RFC3448]   M. Handley, S. Floyd, J. Padhye, and J. Widmer, TCP
                   Friendly Rate Control (TFRC): Protocol Specification,
                   RFC 3448, Proposed Standard, January 2003. Obsoleted
                   by RFC 5348.

Authors' Addresses

    Sally Floyd
    ICSI Center for Internet Research
    1947 Center Street, Suite 600
    Berkeley, CA 94704
    USA

    Email: floyd@icir.org

    Eddie Kohler
    4531C Boelter Hall
    UCLA Computer Science Department
    Los Angeles, CA 90095
    USA

    Email: kohler@cs.ucla.edu