

Network Working Group
Request for Comments: 5634
Category: Experimental

G. Fairhurst
A. Sathiaselan
University of Aberdeen
August 2009

Quick-Start for the Datagram Congestion Control Protocol (DCCP)

Abstract

This document specifies the use of the Quick-Start mechanism by the Datagram Congestion Control Protocol (DCCP). DCCP is a transport protocol that allows the transmission of congestion-controlled, unreliable datagrams. DCCP is intended for applications such as streaming media, Internet telephony, and online games. In DCCP, an application has a choice of congestion control mechanisms, each specified by a Congestion Control Identifier (CCID). This document specifies general procedures applicable to all DCCP CCIDs and specific procedures for the use of Quick-Start with DCCP CCID 2, CCID 3, and CCID 4. Quick-Start enables a DCCP sender to cooperate with Quick-Start routers along the end-to-end path to determine an allowed sending rate at the start of a connection and, at times, in the middle of a DCCP connection (e.g., after an idle or application-limited period). The present specification is provided for use in controlled environments, and not as a mechanism that would be intended or appropriate for ubiquitous deployment in the global Internet.

Status of This Memo

This memo defines an Experimental Protocol for the Internet community. It does not specify an Internet standard of any kind. Discussion and suggestions for improvement are requested. Distribution of this memo is unlimited.

Copyright Notice

Copyright (c) 2009 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents in effect on the date of publication of this document (<http://trustee.ietf.org/license-info>). Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

1. Introduction	3
1.1. Terminology	4
2. Quick-Start for DCCP	5
2.1. Sending a Quick-Start Request for a DCCP Flow	5
2.1.1. The Quick-Start Interval	5
2.2. Receiving a Quick-Start Request for a DCCP Flow	6
2.2.1. The Quick-Start Response Option	7
2.3. Receiving a Quick-Start Response	8
2.3.1. The Quick-Start Mode	8
2.3.2. The Quick-Start Validation Phase	9
2.4. Procedure When No Response to a Quick-Start Request	10
2.5. Procedure When a Packet Is Dropped While Using Quick-Start	11
2.6. Interactions with Mobility and Signaled Path Changes	11
2.7. Interactions with Path MTU Discovery	12
2.8. Interactions with Middleboxes	12
3. Mechanisms for Specific CCIDs	13
3.1. Quick-Start for CCID 2	13
3.1.1. The Quick-Start Request for CCID 2	13
3.1.2. Sending a Quick-Start Response with CCID 2	13
3.1.3. Using the Quick-Start Response with CCID 2	13
3.1.4. Quick-Start Validation Phase for CCID 2	14
3.1.5. Reported Loss or Congestion While Using Quick-Start	14
3.1.6. CCID 2 Feedback Traffic on the Reverse Path	15
3.2. Quick-Start for CCID 3	15
3.2.1. The Quick-Start Request for CCID 3	15
3.2.2. Sending a Quick-Start Response with CCID 3	15
3.2.3. Using the Quick-Start Response with CCID 3	16
3.2.4. Quick-Start Validation Phase for CCID 3	17
3.2.5. Reported Loss or Congestion during the Quick-Start Mode or Validation Phase	17
3.2.6. CCID 3 Feedback Traffic on the Reverse Path	18

3.3. Quick-Start for CCID 4	18
3.3.1. The Quick-Start Request for CCID 4	18
3.3.2. Sending a Quick-Start Response with CCID 4	18
3.3.3. Using the Quick-Start Response with CCID 4	18
3.3.4. Reported Loss or Congestion While Using Quick-Start	19
3.3.5. CCID 4 Feedback Traffic on the Reverse Path	19
4. Discussion of Issues	19
4.1. Overrun and the Quick-Start Validation Phase	19
4.2. Experimental Status	19
5. IANA Considerations	20
6. Acknowledgments	20
7. Security Considerations	20
8. References	21
8.1. Normative References	21
8.2. Informative References	21

1. Introduction

The Datagram Congestion Control Protocol (DCCP) [RFC4340] is a transport protocol for congestion-controlled, unreliable datagrams, intended for applications such as streaming media, Internet telephony, and online games.

In DCCP, an application has a choice of congestion control mechanisms, each specified by a Congestion Control Identifier (CCID) [RFC4340]. There are general procedures applicable to all DCCP CCIDs that are described in Section 2, and details that relate to how individual CCIDs should operate, which are described in Section 3. This separation of CCID-specific and DCCP general functions is in the spirit of the modular approach adopted by DCCP.

Quick-Start [RFC4782] is an experimental mechanism for transport protocols specified for use in controlled environments. The current specification of this mechanism is not intended or appropriate for ubiquitous deployment in the global Internet.

Quick-Start is designed for use between end hosts within the same network or on Internet paths that include IP routers. It works in cooperation with routers, allowing a sender to determine an allowed sending rate at the start and at times in the middle of a data transfer (e.g., after an idle or application-limited period).

This document assumes the reader is familiar with RFC 4782 [RFC4782], which specifies the use of Quick-Start with IP and with TCP. Section 7 of RFC 4782 also provides guidelines for the use of Quick-Start

with other transport protocols, including DCCP. This document provides answers to some of the issues that were raised by RFC 4782 and provides a definition of how Quick-Start must be used with DCCP.

In using Quick-Start, the sending DCCP end host indicates the desired sending rate in bytes per second, using a Quick-Start option in the IP header of a DCCP packet. Each Quick-Start-capable router along the path could, in turn, either approve the requested rate, reduce the requested rate, or indicate that the Quick-Start Request is not approved.

If the Quick-Start Request is approved (possibly with a reduced rate) by all the routers along the path, then the DCCP receiver returns an appropriate Quick-Start Response. On receipt of this, the sending end host can send at up to the approved rate for a period determined by the method specified for each DCCP CCID, and not exceeding three round-trip times. Subsequent transmissions will be governed by the default CCID congestion control mechanisms for the connection. If the Quick-Start Request is not approved, then the sender must use the default congestion control mechanisms.

DCCP receivers are not required to acknowledge individual packets (or pairs of segments) as in TCP. CCID 2 [RFC4341] allows much less frequent feedback. Rate-based protocols (e.g., TCP-Friendly Rate Control (TFRC) [RFC5348], CCID 3 [RFC4342]) have a different feedback mechanism than that of TCP. With rate-based protocols, feedback may be sent less frequently (e.g., once per Round-Trip Time (RTT)). In such cases, a sender using Quick-Start needs to implement a different mechanism to determine whether the Quick-Start sending rate has been sustained by the network. This introduces a new mechanism called the Quick-Start Validation Phase (Section 2.3).

In addition, this document defines two more general enhancements that refine the use of Quick-Start after a flow has started (expected to be more common in applications using DCCP). These are the Quick-Start Interval (Section 2.1.2), and the reaction to mobility triggers (Section 2.6).

1.1. Terminology

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. Quick-Start for DCCP

Unless otherwise specified, DCCP end hosts follow the procedures specified in Section 4 of [RFC4782], following the use specified for Quick-Start with TCP.

2.1. Sending a Quick-Start Request for a DCCP Flow

A DCCP sender MAY use a Quick-Start Request during the start of a connection, when the sender would prefer to have a larger initial rate than allowed by standard mechanisms (e.g., [RFC5348] or [RFC3390]).

A Quick-Start Request MAY also be used once a DCCP flow is connected (i.e., in the middle of a DCCP flow). In standard operation, DCCP CCIDs can constrain the sending rate (or window) to less than that desired (e.g., when an application increases the rate at which it wishes to send). A DCCP sender that has data to send after an idle period or application-limited period (i.e., where the sender has transmitted at less than the allowed sending rate) can send a Quick-Start Request using the procedures defined in Section 3.

Quick-Start Requests will be more effective if the Quick-Start Rate is not larger than necessary. Each requested Quick-Start Rate that has been approved, but was not fully utilized, takes away from the bandwidth pool maintained by Quick-Start routers that would be otherwise available for granting successive requests [RFC4782].

In contrast to most TCP applications, many DCCP applications have the notion of a natural media rate that they wish to achieve. For example, during the initial connection, a host may request a Quick-Start Rate equal to the media rate of the application, appropriately increased to account for the size of packet headers. (Note that Quick-Start only provides a course-grain indication of the desired rate that is expected to be sent in the next RTT.)

When sending a Quick-Start Request, the DCCP sender SHOULD send the Quick-Start Request using a packet that requires an acknowledgement, such as a DCCP-Request, DCCP-Response, or DCCP-Data.

2.1.1. The Quick-Start Interval

Excessive use of the Quick-Start mechanism is undesirable. This document defines an enhancement to RFC 4782 to update the use of Quick-Start after a DCCP flow has started, by introducing the concept of the Quick-Start Interval. The Quick-Start Interval specifies a period of time during which a Quick-Start Request SHOULD NOT be sent. The Quick-Start Interval is measured from the time of transmission of

the previous Quick-Start Request (Section 2.1). The Quick-Start Interval MAY be overridden as a result of a network path change (Section 2.6).

When a connection is established, the Quick-Start Interval is initialized to the Initial_QSI. The Initial_QSI MUST be at least 6 seconds (larger values are permitted). This value was chosen so that it is sufficiently large to prevent excessive router processing over typical Internet paths. Quick-Start routers that track per-flow state MAY penalize senders by suspending Quick-Start processing of flows that make Quick-Start Requests for the same flow with an interval less than 6 seconds.

When the first Quick-Start Request is sent, the Quick-Start Interval is set to:

Quick-Start Interval = Initial_QSI;

After sending each subsequent Quick-Start Request, the Quick-Start Interval is then recalculated as:

Quick-Start Interval = max(Quick-Start Interval * 2, 4 * RTT);

Each unsuccessful Quick-Start Request therefore results in the Quick-Start Interval being doubled (resulting in an exponential back-off). The maximum time the sender can back off is 64 seconds. When the back-off calculation results in a larger value, the sender MUST NOT send any further Quick-Start Requests for the remainder of the DCCP connection (i.e., the sender ceases to use Quick-Start).

Whenever a Quick-Start Request is approved (at any rate), the Quick-Start Interval is reset to the Initial_QSI.

2.2. Receiving a Quick-Start Request for a DCCP Flow

The procedure for processing a received Quick-Start Request is normatively defined in [RFC4782] and summarized in this paragraph. An end host that receives an IP packet containing a Quick-Start Request passes the Quick-Start Request, along with the value in the IP Time to Live (TTL) field, to the receiving DCCP layer. If the receiving host is willing to permit the Quick-Start Request, it SHOULD respond immediately by sending a packet that carries the Quick-Start Response option in the DCCP header of the corresponding feedback packet (e.g., using a DCCP-Ack packet or in a DCCP-DataAck packet).

The Rate Request field in the Quick-Start Response option is set to the received value of the Rate Request in the Quick-Start option or to a lower value if the DCCP receiver is only willing to allow a lower Rate Request. Where information is available (e.g., knowledge of the local Layer 2 interface speed), a Quick-Start receiver **SHOULD** verify that the received rate does not exceed its expected receive link capacity. The TTL Diff field in the Quick-Start Response is set to the difference between the received IP TTL value (Hop Limit field in IPv6) and the Quick-Start TTL value. The Quick-Start Nonce in the Response is set to the received value of the Quick-Start Nonce in the Quick-Start option (or IPv6 Header Extension).

The Quick-Start Response **MUST NOT** be resent if it is lost in the network [RFC4782]. Packet loss could be an indication of congestion on the return path; in which case, it is better not to approve the Quick-Start Request.

If an end host receives an IP packet with a Quick-Start Request with a requested rate of zero, then this host **SHOULD NOT** send a Quick-Start Response [RFC4782].

2.2.1. The Quick-Start Response Option

The Quick-Start Response message must be carried by the transport protocol using Quick-Start. This section defines a DCCP Header option used to carry the Quick-Start Response. This header option is **REQUIRED** for end hosts to utilize the Quick-Start mechanism with DCCP flows. The format resembles that defined for TCP [RFC4782].

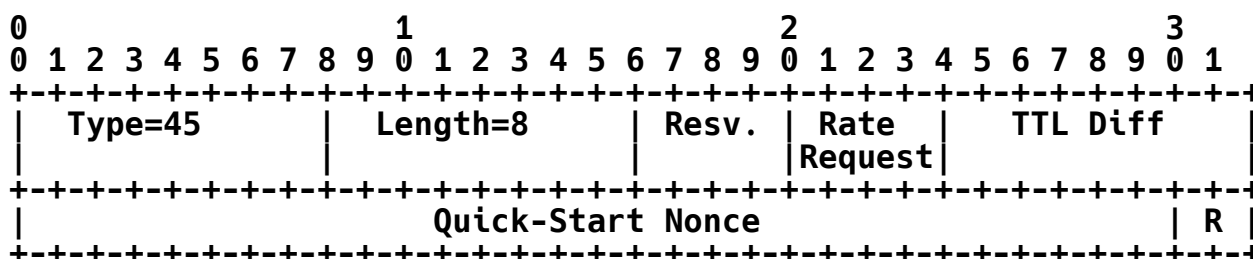


Figure 1. The Quick-Start Response Option

The first byte of the Quick-Start Response option contains the option kind, identifying the DCCP option (45).

The second byte of the Quick-Start Response option contains the option length in bytes. The length field **MUST** be set to 8 bytes.

The third byte of the Quick-Start Response option contains a four-bit Reserved field, and the four-bit allowed Rate Request, formatted as in the IP Quick-Start Rate Request option [RFC4782].

The fourth byte of the DCCP Quick-Start Response option contains the TTL Diff. The TTL Diff contains the difference between the IP TTL and Quick-Start TTL fields in the received Quick-Start Request packet, as calculated in [RFC4782].

Bytes 5-8 of the DCCP option contain the 30-bit Quick-Start Nonce and a 2-bit Reserved field [RFC4782].

2.3. Receiving a Quick-Start Response

On reception of a Quick-Start Response packet, the sender MUST report the approved rate, by sending a Quick-Start Report of Approved Rate [RFC4782]. This report includes the Rate Report field set to the Approved Rate, and the QS Nonce set to the QS Nonce value sent in the Quick-Start Request.

The Quick-Start Report of Approved Rate is sent as an IPv4 option or IPv6 header extension using the first Quick-Start Packet or sent as an option using a DCCP control packet if there are no DCCP-Data packets pending transmission.

The Quick-Start Interval is also reset (as described in Section 2.1.1).

Reception of a Quick-Start Response packet that approves a rate higher than the current rate results in the sender entering the Quick-Start Mode.

2.3.1. The Quick-Start Mode

While a sender is in the Quick-Start Mode, all sent packets are known as Quick-Start Packets [RFC4782]. The Quick-Start Packets MUST be sent at a rate not greater than the rate specified in the Quick-Start Response. The Quick-Start Mode continues for a period up to one RTT (shorter, if a feedback message arrives acknowledging the receipt of one or more Quick-Start Packets).

The procedure following exit of the Quick-Start Mode is specified in the following paragraphs. Note that this behavior is CCID-specific and the details for each current CCID are described in Section 3.

2.3.2. The Quick-Start Validation Phase

After transmitting a set of Quick-Start Packets in the Quick-Start Mode (and providing that no loss or congestion is reported), the sender enters the Quick-Start Validation Phase. This phase persists for a period during which the sender seeks to affirm that the capacity used by the Quick-Start Packets did not introduce congestion. This phase is introduced, because unlike TCP, DCCP senders do not necessarily receive frequent feedback that would indicate the congestion state of the forward path.

While in the Quick-Start Validation Phase, the sender is tentatively permitted to continue sending using the Quick-Start rate. This phase normally concludes when the sender receives feedback that includes an acknowledgment that all Quick-Start Packets were received.

However, the duration of the Quick-Start Validation Phase MUST NOT exceed the Quick-Start Validation Time (a maximum of 2 RTTs). Implementations may set a timer (initialized to the Quick-Start Validation Time) to detect the end of this phase. There may be scope for optimization of timer resources in an implementation, since the Quick-Start Validation period temporarily enforces more strict monitoring of acknowledgements than normally used in a CCID (e.g., an implementation may consider using a common timer resource for Quick-Start Validation and a nofeedback timer).

An example sequence of packet exchanges showing Quick-Start with DCCP is shown in Figure 2.

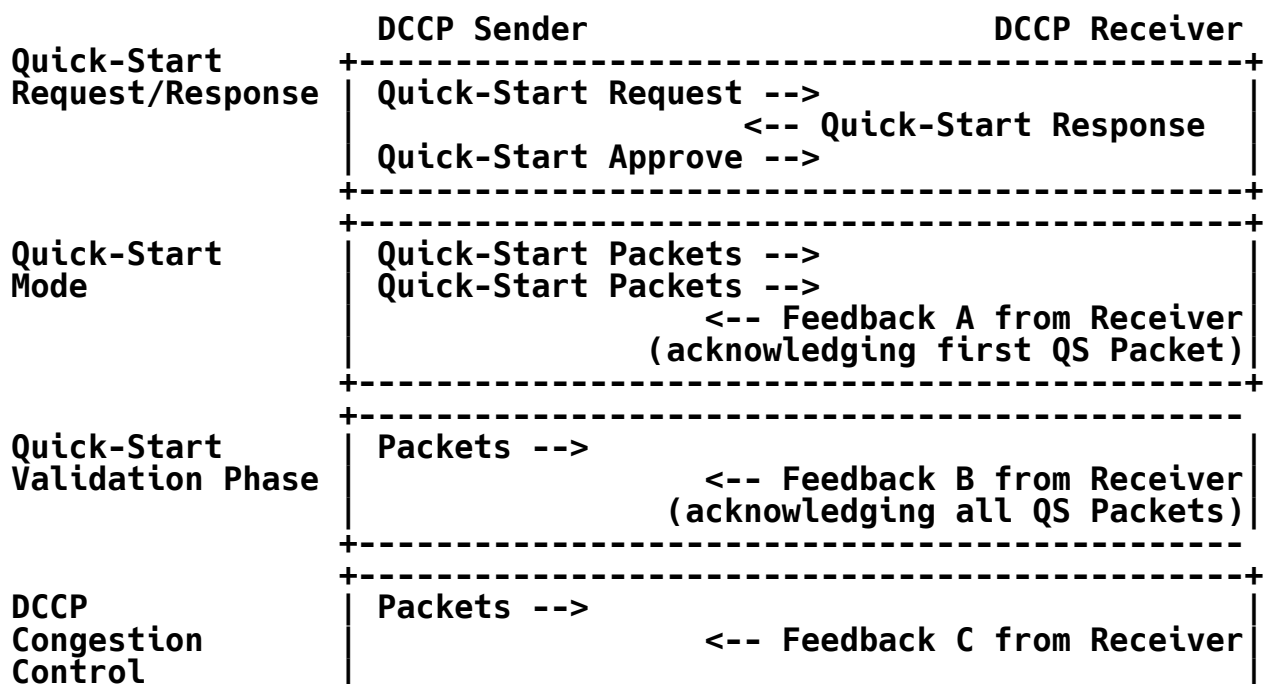


Figure 2. The Quick-Start Mode and Validation Phase

On conclusion of the Validation Phase (Feedback B in the above figure), the sender expects to receive assurance that it may safely use the current rate. A sender that completes the Quick-Start Validation Phase with no reported packet loss or congestion stops using the Quick-Start rate and continues to adjust its rate using the standard congestion control mechanisms. For example, if the DCCP sender was in slow-start prior to the Quick-Start Request, and no packets were lost or ECN-marked (Explicit Congestion Notification) since that time, then the sender continues in slow-start after exiting Quick-Start Mode until the sender sees a packet loss, or congestion is reported.

2.4. Procedure When No Response to a Quick-Start Request

As in TCP, if a Quick-Start Request is dropped (i.e., the Request or Response is not delivered by the network) the DCCP sender MUST revert to the congestion control mechanisms it would have used if the Quick-Start Request had not been approved. The connection is not permitted to send a subsequent Quick-Start Request before expiry of the current Quick-Start Interval (Section 2.1.1).

2.5. Procedure When a Packet Is Dropped While Using Quick-Start

A lost or ECN-marked packet is an indication of potential network congestion. The behavior of a DCCP sender following a lost or ECN-marked Quick-Start Packet or a lost feedback packet is specific to a particular CCID (see Section 3).

2.6. Interactions with Mobility and Signaled Path Changes

The use of Quick-Start may assist end hosts in determining when it is appropriate to increase their rate following an explicitly signaled change of the network path.

When an end host receives a signal from an upstream link/network notifying it of a path change, the change could simultaneously impact more than one flow, and may affect flows between multiple endpoints. Senders should avoid responding immediately, since this could result in unwanted synchronization of signaling messages, and control loops (e.g., a synchronized attempt to probe for a larger congestion window), which may negatively impact the performance of the network and transport sessions. In Quick-Start, this could increase the rate of Quick-Start Requests, possibly incurring additional router load, and may result in some requests not being granted. A sender must ensure this does not generate an excessive rate of Quick-Start Requests by using the method below:

A sender that has explicit information that the network path has changed (e.g., a mobile IP binding update [RFC3344], [RFC3775]) SHOULD reset the Quick-Start Interval to its initial value (specified in Section 2.1.1).

The sender MAY also send a Quick-Start Request to determine a new safe transmission rate, but must observe the following rules:

- It MUST NOT send a Quick-Start Request within a period less than the initial Quick-Start Interval (Initial_QSI) since it previously sent a Quick-Start Request. That is, it must wait for at least a period of Initial_QSI after the previous request, before sending a new Quick-Start Request.
- If it has not sent a Quick-Start Request within the previous Initial_QSI period, it SHOULD defer sending a Quick-Start Request for a randomly chosen period between 0 and the Initial_QSI value in seconds. The random period should be statistically independent between different hosts and between different connections on the same host. This delay is to mitigate the effect on router load of synchronized responses by multiple connections in response to a path change that affects multiple connections.

Hosts do not generally have sufficient information to choose an appropriate randomization interval. This value was selected to ensure randomization of requests over the Quick-Start Interval. In networks where a large number of senders may potentially be impacted by the same signal, a larger value may be desirable (or methods may be used to control this effect in the path change signaling).

2.7. Interactions with Path MTU Discovery

DCCP implementations are encouraged to support Path MTU Discovery (PMTUD) when applications are able to use a DCCP packet size that exceeds the default Path MTU [RFC4340], [RFC4821]. Quick-Start Requests SHOULD NOT be sent with packets that are used as a PMTUD Probe Packet, since these packets could be lost in the network increasing the probability of loss of the request. It may therefore be preferable to separately negotiate the PMTU and the use of Quick-Start.

The DCCP protocol is datagram-based and therefore the size of the segments that are sent is a function of application behavior as well as being constrained by the largest supported Path MTU.

2.8. Interactions with Middleboxes

A Quick-Start Request is carried in an IPv4 packet option or IPv6 extension header [RFC4782]. Interactions with network devices (middleboxes) that inspect or modify IP options could therefore lead to discard, ICMP error, or DCCP-Reset when attempting to forward packets carrying a Quick-Start Request.

If a DCCP sender sends a DCCP-Request that also carries a Quick-Start Request, and does not receive a DCCP-Response to the packet, the DCCP sender SHOULD resend the DCCP-Request packet without including a Quick-Start Request.

Similarly, if a DCCP sender receives a DCCP-Reset in response to a DCCP-Request packet that also carries a Quick-Start Request, then the DCCP sender SHOULD resend the DCCP-Request packet without the Quick-Start Request. The DCCP sender then ceases to use the Quick-Start Mechanism for the remainder of the connection.

A DCCP sender that uses a Quick-Start Request within an established connection and does not receive a response will treat this as non-approval of the request. Successive unsuccessful attempts will result in an exponential increase in the Quick-Start Interval (Section 2.1.1). If this grows to a value exceeding 64 seconds, the DCCP sender ceases to use the Quick-Start Mechanism for the remainder of the connection.

3. Mechanisms for Specific CCIDs

The following sections specify the use of Quick-Start with DCCP CCID 2, CCID 3, and for DCCP with TFRC-SP as specified in CCID 4.

3.1. Quick-Start for CCID 2

This section describes the Quick-Start mechanism to be used with DCCP CCID 2 [RFC4341]. CCID 2 uses a TCP-like congestion control mechanism.

3.1.1. The Quick-Start Request for CCID 2

A Quick-Start Request MAY be sent to allow the sender to determine if it is safe to use a larger initial cwnd. This permits a faster start-up of a new CCID 2 flow.

A Quick-Start Request MAY also be sent for an established connection to request a higher sending rate after an idle period or application-limited period (described in Section 2.1). This allows a receiver to use a larger cwnd than allowed with standard operation.

A Quick-Start Request that follows a reported loss or congestion event MUST NOT request a Quick-Start rate that exceeds the largest congestion window achieved by the CCID 2 connection since the last packet drop (translated to a sending rate).

3.1.2. Sending a Quick-Start Response with CCID 2

A receiver processing a Quick-Start Request uses the method described in Section 2.3. On receipt of a Quick-Start Request, the receiver MUST send a Quick-Start Response (even if a receiver is constrained by the ACK Ratio).

3.1.3. Using the Quick-Start Response with CCID 2

On receipt of a valid Quick-Start Response option, the sender MUST send a Quick-Start Approved option [RFC4782] (see Section 2.3).

If the approved Quick-Start rate is less than current sending rate, the sender does not enter the Quick-Start Mode, and continues using the procedure defined in CCID 2.

If the approved Quick-Start rate at the sender exceeds the current sending rate, the sender enters the Quick-Start Mode and continues in the Quick-Start Mode for a maximum period of one RTT.

The sender sets its Quick-Start cwnd (QS_cwnd) as follows:

$$\text{QS_cwnd} = (R * T) / (s + H) \quad (1)$$

where R is the Rate Request in bytes per second, T is the measured round-trip path delay (RTT), s is the packet size, and H is the estimated DCCP/IP header size in bytes (e.g., 32 bytes for DCCP layered directly over IPv4, but larger when using IPsec or IPv6).

A CCID 2 sender MAY then increase its cwnd to the QS_cwnd. The cwnd should not be reduced (i.e., a QS_cwnd lower than cwnd should be ignored, since the CCID 2 congestion control method already permits this rate). CCID 2 is not a rate-paced protocol. Therefore, if the QS_cwnd is used, the sending host MUST implement a suitable method to pace the rate at which the Quick-Start Packets are sent until it receives a DCCP-ACK for a packet sent during the Quick-Start Mode [RFC4782]. The sending host SHOULD also record the previous cwnd and note that the new cwnd has been determined by Quick-Start, rather by other means (e.g., by setting a flag to indicate that it is in Quick-Start Mode).

When the sender receives the first DCCP-ACK to a packet sent in the Quick-Start Mode, it leaves the Quick-Start Mode and enters the Validation Phase.

3.1.4. Quick-Start Validation Phase for CCID 2

A CCID 2 sender MAY continue to send at the paced Quick-Start Rate while in the Validation Phase. It leaves the Validation Phase on receipt of an ACK that acknowledges the last Quick-Start Packet, or if the validation phase persists for a period that exceeds the Quick-Start Validation Time of 1 RTT. It MUST then reduce the cwnd to the actual flight size (the current amount of unacknowledged data sent) [RFC4782] and uses the congestion control methods specified for CCID 2.

3.1.5. Reported Loss or Congestion While Using Quick-Start

A sender in the Quick-Start Mode (or Validation Phase) that detects congestion (e.g., receives a feedback packet that reports new packet loss or a packet with a congestion marking), MUST immediately leave the Quick-Start Mode (or Validation Phase). It then resets the cwnd to half the recorded previous cwnd and enters the congestion avoidance phase described in [RFC4341].

In the absence of any feedback at the end of the Validation period, the sender resets the cwnd to half the recorded previous cwnd and enters the congestion avoidance phase.

3.1.6. CCID 2 Feedback Traffic on the Reverse Path

A CCID 2 receiver sends feedback for groups of received packets [RFC4341]. Approval of a higher transmission rate using Quick-Start will increase control traffic on the reverse path. A return path that becomes congested could have a transient negative impact on other traffic flows sharing the return link. The lower rate of feedback will then limit the achievable rate in the forward direction.

3.2. Quick-Start for CCID 3

This section describes the Quick-Start mechanism to be used with DCCP CCID 3 [RFC4342]. The rate-based congestion control mechanism used by CCID 3 leads to specific issues that are addressed by Quick-Start in this section.

3.2.1. The Quick-Start Request for CCID 3

A Quick-Start Request MAY be sent to allow the sender to determine if it is safe to use a larger initial sending rate. This permits a faster start-up of a new CCID 3 flow.

A Quick-Start Request MAY also be sent to request a higher sending rate after an idle period (in which the nofeedback timer expires [RFC5348]) or an application-limited period (described in Section 2.1). This allows a receiver to increase the sending rate faster than allowed with standard operation (i.e., faster than twice the rate reported by a CCID 3 receiver in the most recent feedback message).

The requested rate specified in a Quick-Start Request MUST NOT exceed the TFRC-controlled sending rate [RFC4342] when this is bounded by the current loss event rate (if any), either from calculation at the sender or from feedback received from the receiver. CCID 3 considers this rate is a safe response in the presence of expected congestion.

3.2.2. Sending a Quick-Start Response with CCID 3

When processing a received Quick-Start Request, the receiver uses the method described in Section 2.3. In addition, if a CCID 3 receiver uses the window counter to send periodic feedback messages, then the receiver sets its local variable `last_counter` to the value of the window counter reported by the segment containing the Quick-Start Request. The next feedback message would then be sent when the window counter is greater or equal to `last_counter + 4`. If the CCID 3 receiver uses a feedback timer to send period feedback messages, then the receiver MUST reset the CCID 3 feedback timer, causing the

feedback to be sent as soon as possible. This helps to align the timing of feedback to the start and end of the period in which Quick-Start Packets are sent, and will normally result in feedback at a time that is approximately the end of the period when Quick-Start Packets are received.

3.2.3. Using the Quick-Start Response with CCID 3

On receipt of a valid Quick-Start Response option, the sender **MUST** send a Quick-Start Approved option [RFC4782] (see Section 2.3). The sender then uses one of three procedures:

- * If the approved Quick-Start rate is less than the current sending rate, the sender does not enter the Quick-Start Mode and continues using the procedure defined in CCID 3.
- * If loss or congestion is reported after sending the Quick-Start Request, the sender also does not enter the Quick-Start Mode and continues using the procedure defined in CCID 3.
- * If the approved Quick-Start rate exceeds the current sending rate, the sender enters the Quick-Start Mode and continues in the Quick-Start Mode for a maximum period of 1 RTT. The sender sets its Quick-Start sending rate (QS_sendrate) as follows:

$$\text{QS_sendrate} = R * s / (s + H); \quad (2)$$

where R the Rate Request in bytes per second, s is the packet size [RFC4342], and H the estimated DCCP/IP header size in bytes (e.g., 32 bytes for IPv4). A CCID 3 host **MAY** then increase its sending rate to the QS_sendrate. The rate should not be reduced.

CCID 3 is a rate-paced protocol. Therefore, if the QS_sendrate is used, the sending host **MUST** pace the rate at which the Quick-Start Packets are sent over the next RTT. The sending host **SHOULD** also record the previous congestion-controlled rate and note that the new rate has been determined by Quick-Start rather by other means (e.g., by setting a flag to indicate that it is in the Quick-Start Mode).

The sender exits the Quick-Start Mode after either:

- * Receipt of a feedback packet acknowledging one or more Quick-Start Packets,
- * A period of 1 RTT after receipt of a Quick-Start Response, or
- * Detection of a loss or congestion event (see Section 3.2.5).

3.2.4. Quick-Start Validation Phase for CCID 3

After transmitting a set of Quick-Start Packets in the Quick Start Mode (and providing that no loss or congestion marking is reported), the sender enters the Quick-Start Validation Phase. A sender that receives feedback that reports a loss or congestion event MUST follow the procedures described in Section 3.2.5.

The sender MUST exit the Quick-Start Validation Phase on receipt of feedback that acknowledges all packets sent in the Quick-Start Mode (i.e., all Quick-Start Packets) or if the Validation Phase persists for a period that exceeds the Quick-Start Validation Time of two RTTs.

A sender that completes the Quick-Start Validation Phase with no reported packet loss or congestion stops using the `QS_sendrate` and MUST recalculate a suitable sending rate using the standard congestion control mechanisms [RFC4342].

If no feedback is received within the Quick-Start Validation Phase, the sender MUST return to the minimum of the recorded original rate (at the start of the Quick-Start Mode) and one half of the `QS_sendrate`. The `nofeedback` timer is also reset.

3.2.5. Reported Loss or Congestion during the Quick-Start Mode or Validation Phase

A sender in the Quick-Start Mode or Validation Phase that detects congestion (e.g., receives a feedback packet that reports new packet loss or a packet with a congestion marking) MUST immediately leave the Quick-Start Mode or Validation Phase and enter the congestion avoidance phase [RFC4342]. This implies re-calculating the sending rate, `X`, as required by RFC 4342:

$$X = \max(\min(X_{\text{calc}}, 2 * X_{\text{recv}}), s / t_{\text{mbi}});$$

where `X_calc` is the transmit rate calculated by the throughput equation, `X_recv` is the reported receiver rate, `s` is the packet size and `t_mbi` is the maximum back-off interval of 64 seconds.

The current specification of TFRC [RFC5348], which obsoletes RFC 3448, uses a set of `X_recv` values and uses the maximum of the set during application-limited intervals. This calculates the sending rate, `X` as:

$$X = \max(\min(X_{\text{calc}}, \text{recv_limit}), s / t_{\text{mbi}});$$

where `recv_limit` could be `max(X_recv_set)` or `2*max(X_recv_set)` depending on whether there was a new loss event during a data-limited interval, or no loss event during an application-limited interval respectively. When the sender is not application-limited, the `recv_limit` is set to `2*max(X_recv_set)`.

A sender using RFC 4342 updated by [RFC5348], calculates the sending rate, `X`, using the above formula normatively defined in [RFC5348].

3.2.6. CCID 3 Feedback Traffic on the Reverse Path

A CCID 3 receiver sends feedback at least once each RTT [RFC4342]. Use of Quick-Start is therefore not expected to significantly increase control traffic on the reverse path.

3.3. Quick-Start for CCID 4

This section describes the Quick-Start mechanism to be used when DCCP uses TFRC-SP [RFC4828] in place of TFRC [RFC5348], as specified in CCID 4 [RFC5622]. CCID 4 is similar to CCID 3 except that a sender using CCID 4 is limited to a maximum of 100 packets/second.

The Quick-Start procedure defined below therefore resembles that for CCID 3.

3.3.1. The Quick-Start Request for CCID 4

The procedure for sending a Quick-Start Request using CCID 4 is the same as for CCID 3, defined in Section 3.2.1. In addition, the requested rate **MUST** be less than or equal to the equivalent of a sending rate of 100 packets per second [RFC4828]. CCID 4 [RFC4828] specifies that the allowed sending rate derived from the TCP throughput equation is reduced by a factor that accounts for packet header size.

3.3.2. Sending a Quick-Start Response with CCID 4

This procedure is the same as for CCID 3, defined in Section 3.2.2.

3.3.3. Using the Quick-Start Response with CCID 4

This procedure is the same as for CCID 3, defined in Sections 3.2.3, 3.2.4, and 3.2.5, except that the congestion control procedures is updated to use TFRC-SP [RFC4828].

A CCID 4 sender does not need to account for headers a second time when translating the approved Quick-Start rate into an allowed sending rate (as described in Section 5 of [RFC5622]).

3.3.4. Reported Loss or Congestion While Using Quick-Start

This procedure is the same as for CCID 3, defined in Section 3.2.5, except that the congestion control procedures is updated to use TFRC-SP [RFC4828].

3.3.5. CCID 4 Feedback Traffic on the Reverse Path

A CCID 4 receiver sends feedback at least once each RTT. Use of Quick-Start is therefore not expected to significantly increase control traffic on the reverse path.

4. Discussion of Issues

The considerations for using Quick-Start with DCCP are not significantly different to those for Quick-Start with TCP. The document does not modify the router behavior specified for Quick-Start.

4.1. Overrun and the Quick-Start Validation Phase

The less frequent feedback of DCCP raises an issue in that a sender using Quick-Start may continue to use the rate specified by a Quick-Start Response for a period that exceeds one path round trip time (i.e., that which TCP would have used). This overrun is a result of the less frequent feedback interval used by DCCP (i.e., CCID 2 may delay feedback by at most one half cwnd and CCID 3 and CCID 4 provide feedback at least once per RTT). In the method specified by this document, the Quick-Start Validation Phase bounds this overrun to be not more than an additional two RTTs.

The selected method was chosen as a compromise that reflects the need to terminate quickly following the loss of a feedback packet, and the need to allow sufficient time for end host and router processing, as well as the different perceptions of the path RTT held at the sender and receiver. Any reported loss or congestion results in immediate action without waiting for completion of the Quick-Start Validation period.

4.2. Experimental Status

There are many cases in which Quick-Start Requests would not be approved [RFC4782]. These include communication over paths containing routers, IP tunnels, MPLS paths, and the like, that do not support Quick-Start. These cases also include paths with routers or middleboxes that drop packets containing IP options (or IPv6 extensions). Quick-Start Requests could be difficult to approve over paths that include multi-access Layer-2 networks.

Transient effects could arise when the transport protocol packets associated with a connection are multiplexed over multiple parallel (sometimes known as alternative) links or network-layer paths, and Quick-Start is used, since it will be effective on only one of the paths, but could lead to increased traffic on all paths.

A CCID 2 sender using Quick-Start can increase the control traffic on the reverse path, which could have a transient negative impact on other traffic flows sharing the return link (Section 3.1.6). The lower rate of feedback will then limit the achievable rate in the forward direction.

[RFC4782] also describes environments where the Quick-Start mechanism could fail with false positives, with the sender incorrectly assuming that the Quick-Start Request had been approved by all of the routers along the path. As a result of these concerns, and as a result of the difficulties and the seeming absence of motivation for routers, such as core routers, to deploy Quick-Start, Quick-Start has been proposed as a mechanism that could be of use in controlled environments, and not as a mechanism that would be intended or appropriate for ubiquitous deployment in the global Internet.

Further experimentation would be required to confirm the deployment of Quick-Start and to investigate performance issues that may arise, prior to any recommendation for use over the general Internet.

5. IANA Considerations

IANA has assigned a DCCP Option Type (45) from the DCCP Option Types Registry. This Option is applicable to all CCIDs and is known as the "Quick-Start Response" Option and is defined in Section 2.2.1. It specifies a length value in the format used for options numbered 32-128.

6. Acknowledgments

The author gratefully acknowledges the previous work by Sally Floyd to identify issues that impact Quick-Start for DCCP, and her comments to improve this document. We also acknowledge comments and corrections from Pasi Sarolahti, Vincent Roca, Mark Allman, Michael Scharf, and others in the IETF DCCP Working Group (WG).

7. Security Considerations

Security issues are discussed in [RFC4782]. Middlebox deployment issues are also highlighted in Section 2.8. No new security issues are raised within this document.

8. References

8.1. 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.
- [RFC4341] Floyd, S. and E. Kohler, "Profile for Datagram Congestion Control Protocol (DCCP) Congestion Control ID 2: TCP-like Congestion Control", RFC 4341, 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.
- [RFC4782] Floyd, S., Allman, M., Jain, A., and P. Sarolahti, "Quick-Start for TCP and IP", RFC 4782, January 2007.
- [RFC4828] Floyd, S. and E. Kohler, "TCP Friendly Rate Control (TFRC): The Small-Packet (SP) Variant", RFC 4828, April 2007.
- [RFC5348] Floyd, S., Handley, M., Padhye, J., and J. Widmer, "TCP Friendly Rate Control (TFRC): Protocol Specification", RFC 5348, September 2008.
- [RFC5622] Floyd, S., and E. Kohler, "Profile for Datagram Congestion Control Protocol (DCCP) Congestion ID 4: TCP-Friendly Rate Control for Small Packets (TFRC-SP)", RFC 5622, August 2009.

8.2. Informative References

- [RFC3344] Perkins, C., Ed., "IP Mobility Support for IPv4", RFC 3344, August 2002.
- [RFC3775] Johnson, D., Perkins, C., and J. Arkko, "Mobility Support in IPv6", RFC 3775, June 2004.
- [RFC3390] Allman, M., Floyd, S., and C. Partridge, "Increasing TCP's Initial Window", RFC 3390, October 2002.
- [RFC4821] Mathis, M. and J. Heffner, "Packetization Layer Path MTU Discovery", RFC 4821, March 2007.

Authors' Addresses

Godred Fairhurst
School of Engineering
University of Aberdeen
Aberdeen, AB24 3UE
Scotland, UK

EMail: gorry@erg.abdn.ac.uk
URI: <http://www.erg.abdn.ac.uk/users/gorry>

Arjuna Sathiaseelan
School of Engineering
University of Aberdeen
Aberdeen, AB24 3UE
Scotland, UK

EMail: arjuna@erg.abdn.ac.uk
URI: <http://www.erg.abdn.ac.uk/users/arjuna>