Network Working Group   M. Tuexen
Internet-Draft   I. Ruengeler
Intended status: Standards Track   Muenster Univ. of Appl. Sciences
Expires: November 1, 2012   R. Stewart
                                                        Adara Networks
                                                         April 30, 2012

SACK-IMMEDIATELY Extension for the Stream Control Transmission Protocol
         draft-tuexen-tsvwg-sctp-sack-immediately-09.txt

Abstract

   This document defines a method for the sender of a DATA chunk to
   indicate that the corresponding SACK chunk should be sent back
   immediately and not be delayed.

Status of this Memo

   This Internet-Draft is submitted in full conformance with the
   provisions of BCP 78 and BCP 79.

   Internet-Drafts are working documents of the Internet Engineering
   Task Force (IETF). Note that other groups may also distribute
   working documents as Internet-Drafts. The list of current Internet-
   Drafts is at http://datatracker.ietf.org/drafts/current/.

   Internet-Drafts are draft documents valid for a maximum of six months
   and may be updated, replaced, or obsoleted by other documents at any
   time. It is inappropriate to use Internet-Drafts as reference
   material or to cite them other than as "work in progress."

   This Internet-Draft will expire on November 1, 2012.

Table of Contents

1. Introduction

According to [RFC4960] the receiver of a DATA chunk should use
delayed SACKs. This delaying is completely controlled by the
receiver of the DATA chunk.

In specific situations the delaying of SACKs results in reduced
performance of the protocol. If such a situation can be detected by
the receiver, the corresponding SACK can be sent immediately. For
example, [RFC4960] recommends the immediate sending if the receiver
has detected message loss or message duplication. However, if the
situation can only be detected by the sender of the DATA chunk,
[RFC4960] provides no method of avoiding the delaying of the SACK.
Thus the protocol performance might be reduced.

This document overcomes this limitation and describes a simple
extension of the SCTP DATA chunk by defining a new flag, the I-bit.
The sender of a DATA chunk indicates by setting this bit that the
corresponding SACK chunk should not be delayed.

Upper layers of SCTP using the socket API as defined in [RFC6458] may
subscribe to the SCTP_SENDER_DRY_EVENT for getting a notification as
soon as no user data is outstanding anymore. To avoid an unnecessary
delay while waiting for such an event, the application might set the
I-Bit on the last DATA chunk sent before waiting for the event. This
enabling is possible using the extension of the socket API described
in Section 6.

There are also situations in which the SCTP implementation can set
the I-bit without interacting with the upper layer. If the
association is in the SHUTDOWN-PENDING state, the I-bit should be
set. This reduces the number of simultaneous associations in case of
a busy server handling short living associations. Another case is
where the sending of a DATA chunk fills the congestion or receiver
window. Setting the I-bit in these cases improves the throughput of
the transfer.

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].

3. The I-bit in the DATA Chunk Header

The following Figure 1 shows the extended DATA chunk.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type = 0    |  Res  |I|U|B|E|            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              TSN                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Stream Identifier       |    Stream Sequence Number     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Payload Protocol Identifier                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
\                                                               \
/                           User Data                           /
\                                                               \
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
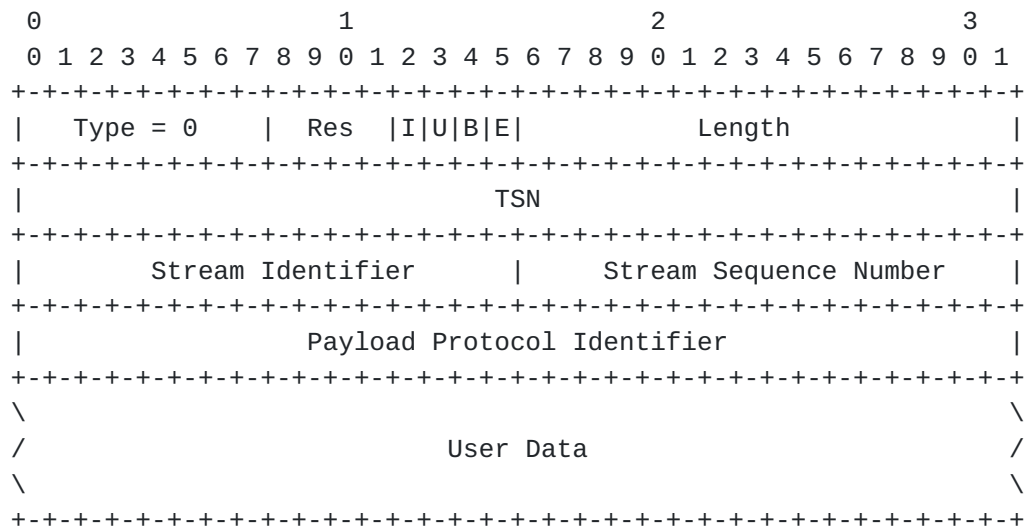
Figure 1: Extended DATA chunk format

The only difference between the DATA chunk in Figure 1 and the DATA
chunk defined in [RFC4960] is the addition of the I-bit in the flags
field of the chunk header.

4. Procedures

4.1. Sender Side Considerations

Whenever the sender of a DATA chunk can benefit from the
corresponding SACK chunk being sent back without delay, the sender
MAY set the I-bit in the DATA chunk header. Please note that it is
irrelevant to the receiver why the sender has set the I-bit.

Reasons for setting the I-bit include but are not limited to the
following:

o   The application requests to set the I-bit of the last DATA chunk
    of a user message when providing the user message to the SCTP
    implementation (see Section 6).

o   The sender is in the SHUTDOWN-PENDING state.

o   The sending of a DATA chunk fills the congestion or receiver
    window.

4.2. Receiver Side Considerations

   On reception of an SCTP packet containing a DATA chunk with the I-bit
   set, the receiver SHOULD NOT delay the sending of the corresponding
   SACK chunk and SHOULD send it back immediately.

5. Interoperability Considerations

   According to [RFC4960] the receiver of a DATA chunk with the I-bit
   set should ignore this bit when it does not support the extension
   described in this document. Since the sender of the DATA chunk is
   able to handle this case, there is no requirement for negotiating the
   support of the feature described in this document.

6. Socket API Considerations

   This section describes how the socket API defined in [RFC6458] is
   extended to provide a way for the application to set the I-bit.

   Please note that this section is informational only.

   A socket API implementation based on [RFC6458] is extended by
   supporting a flag called SCTP_SACK_IMMEDIATELY, which can be set in
   the snd_flags field of the struct sctp_sndinfo structure or the
   sinfo_flags field of the struct sctp_sndrcvinfo structure, which is
   deprecated.

   If the SCTP_SACK_IMMEDIATELY flag is set when sending a user message,
   the I-bit of the last DATA chunk of the corresponding user message is
   set.

7. IANA Considerations

   [NOTE to RFC-Editor:

      "RFCXXXX" is to be replaced by the RFC number you assign this
      document.

   ]

   Following the chunk flag registration procedure defined in [RFC6096]
   IANA should register a new bit, the I-bit, for the DATA chunk. The
   suggested value is 0x08. The reference for the new chunk flag in the
   chunk flags table for the DATA chunk available at sctp-parameters [1]
   should be RFCXXXX.

8. Security Considerations

   This document does not add any additional security considerations in
   addition to the ones given in [RFC4960].

9. Acknowledgments

   The authors wish to thank Mark Allmann, Brian Bidulock, Janardhan
   Iyengar, and Kacheong Poon for their invaluable comments.

10. References

10.1. Normative References

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

   [RFC4960] Stewart, R., "Stream Control Transmission Protocol",
              RFC 4960, September 2007.

   [RFC6096] Tuexen, M. and R. Stewart, "Stream Control Transmission
              Protocol (SCTP) Chunk Flags Registration", RFC 6096,
              January 2011.

10.2. Informative References

   [RFC6458] Stewart, R., Tuexen, M., Poon, K., Lei, P., and V.
              Yasevich, "Sockets API Extensions for the Stream Control
              Transmission Protocol (SCTP)", RFC 6458, December 2011.

URIs

   [1] <http://www.iana.org/assignments/sctp-parameters>

Authors' Addresses

   Michael Tuexen
   Muenster University of Applied Sciences
   Stegerwaldstr. 39
   48565 Steinfurt
   DE

   Email: tuexen@fh-muenster.de

Irene Ruengeler
Muenster University of Applied Sciences
Stegerwaldstr. 39
48565 Steinfurt
DE

Email: i.ruengeler@fh-muenster.de

Randall R. Stewart
Adara Networks
Chapin, SC 29036
US

Email: randall@lakerest.net