     Services provided by IETF transport protocols and congestion control
                                mechanisms
                     draft-fairhurst-taps-transports-00

Abstract

   This document describes services provided by existing IETF protocols
   and congestion control mechanisms.  It is designed to help
   application and network stack programmers and to inform the work of
   the IETF TAPS Working Group.

Status of This Memo

Copyright Notice

1.  Introduction

   Most Internet applications make use of the Transport Services
   provided by TCP (a reliable, in-order stream protocol) or UDP (an
   unreliable datagram protocol).  We use the term "Transport Service"
   to mean an end-to-end facility provided by the transport layer.  That
   service can only be provided correctly if information is supplied
   from the application.  The application may determine the information
   to be supplied at design time, compile time, or run time and may
   include guidance on whether an aspect of the service is required, a
   preference by the application, or something in between.  Examples of
   Transport service facilities are reliable delivery, ordered delivery,
   content privacy to in-path devices, integrity protection, and minimal
   latency.

   Transport protocols such as SCTP, DCCP, MPTCP, UDP and UDP-Lite have
   been defined at the transport layer.

   In addition, a transport service may be built on top of these
   transport protocols, using a fraemwork such as WebSockets, or RTP.
   Service built on top of UDP or UDP-Lite typically also need to
   specify a congestion control mechanism, such as TFRC or the LEDBAT
   congestion control mechanism.  This extends the set of available
   Transport Services beyond those provided to applications by TCP and
   UDP.

   Transport services can aslo be differentiated by the services they
   provide: for instance, SCTP offers a message-based service that does
   not suffer head-of-line blocking when used with multiple stream,
   because it can accept blocks of data out of order, UDP-Lite provides
   partial integrity protection when used over link-layer services that
   can support this, and LEDBAT can provide low-priority "scavenger"
   communication.

2.  Terminology

   This section presents the terminology used in this document.

   [EDITOR'S NOTE: Terminology to be discussed in Honolulu.  We need to
   determine what a "service" as used by the IETF, as opposed to a
   "service component", "property", an "aspect", "dimension", etc.]

3.  Transport Protocols

   This section provides a list of known IETF transport protocol and
   transport protocol frameworks.

   [EDITOR'S NOTE: combine these tables into one?  Also, reorder them to
   match ths sections below.]

```
+---------+-----------------------------------+------+--------------------+
| Section | Benefit                           | Setup| Mode               |
+---------+-----------------------------------+------+--------------------+
|   3.1   | Transmission Control Protocol (TCP) | CO | Unicast            |
|   3.1.1 | Multipath-TCP (MPTCP)             | CO   | Unicast            |
|   3.2   | SCTP                              | CO   | Unicast            |
|   3.2.1 | SCTP-PR                           | CO   | Unicast            |
|   3.3   | User Datagram Protocol (UDP)      | DG   | Unicast/Multicst   |
|   3.4   | UDP-Lite                          | DG   | Unicast/Multicst   |
|   3.5   | DCCP                              | CO   | Unicast            |
|   3.X   | More as needed                    |      |                    |
+---------+-----------------------------------+------+--------------------+
```

Table 1: Key IETF Transport Protocol - by cmmunication mode

```
+---------+-----------------------------------+------+--------------------+
| Section | Benefit                           | Style| Reliability        |
+---------+-----------------------------------+------+--------------------+
|   3.1   | Transmission Control Protocol (TCP) | Str | Ordered Byte Stream |
|   3.1.1 | Multipath-TCP (MPTCP)             | Str  | Ordered Byte Stream |
|   3.2   | SCTP                              | Mess | Message Streams    |
|   3.2.1 | SCTP-PR                           | Mess | Partial M Streams  |
|   3.3   | User Datagram Protocol (UDP)      | Mess | Datagram Message   |
|   3.4   | UDP-Lite                          | Mess | Error Tolerant DG  |
|   3.5   | DCCP                              | Mess | Unrel Message Stream|
|   3.X   | More as needed                    |      |                    |
+---------+-----------------------------------+------+--------------------+
```

Table 2: Key IETF Transport Protocol - by reliability

   "Setup" defines whether the protocol performs a connection-oriented
   protocol handshake prior o communication or is datagram based.  This
   provides reliable negotiation of options, including negotiation of a
   suitable congestion control mechanism.This property can impact the
   ability of the protocol to traverse firewalls.

```
+---------+-----------------------------------+--------------------------+
| Section | Benefit                           | Congestion Control       |
+---------+-----------------------------------+--------------------------+
|   3.1   | Transmission Control Protocol (TCP)| Yes                     |
|   3.1.1 | Multipath-TCP (MPTCP)             | Yes   (Multipath)        |
|   3.2   | SCTP                              | Yes                      |
|   3.2.1 | SCTP-PR                           | Yes                      |
|   3.3   | User Datagram Protocol (UDP)      | At application layer     |
|   3.4   | UDP-Lite                          | At application layer     |
|   3.5   | DCCP                              | Yes, Various CCIDs defined |
|   3.X   | More as needed                    |                          |
+---------+-----------------------------------+--------------------------+
```

Table 3: Key IETF Transport Protocol - by congestion control

   Some other protocol frameworks that may potentially be considered for
   inclusion in future versions of this document.  Examples are:

   o  Multicast - RMT

   o  RTP-based methods

   o  HTTP-based methods

   o  TLS

   o  DTLS

   The following subsections describes each of these transports.

3.1.  Transport Control Protocol (TCP)

   TCP provides a bidirectional byte-oriented stream over a connection-
   oriented protocol.  The protocol and API use the byte-stream model.

   [EDITOR'S NOTE: Describe the aspects(?) of TCP: reliable, connection-
   oriented, congestion-controlled, single-stream-oriented, non-
   boundary-preserving... Note that we want to describe the
   characteristics of the SOCK_STREAM API as well as just the wire
   protocol.]

3.1.1.  Multipath TCP (MPTCP)

   [EDITOR'S NOTE: aspects of MPTCP beyond TCP.]

3.2.  Stream Control Transmission Protocol (SCTP)

   This section will describe SCTP.

   SCTP provides a bidirectional set of logical unicast streams over one
   a connection-oriented protocol.  The protocol and API use messages,
   rather than a byte-stream.  Each stream of messages is independently
   managed, therefore retransmission does not hold back data sent using
   other logical streams

3.2.1.  Partial Reliability SCTP (PR-CTP)

   SCTP-PR [RFC3758] is a variant of SCTP that provides partial
   reliability.

3.3.  User Datagram Protocol (UDP)

   The User Datagram Protocol (UDP) provides a unidirectional minimal
   message-passing transport that has no inherent congestion control
   mechanisms.  The service may be multicast and/or unicast.

   [EDITOR'S NOTE: Describe the aspects(?) of UDP: unreliable,
   congestion control to be applied above the transport, datagram-
   oriented, connectionless, boundary-preserving... Note that we want to
   describe the characteristics of the SOCK_DGRAM API as well as just
   the wire protocol.]

   Using UDP robustly requires each application to implement a raft of
   functions (mostly re-inventing or adaptng mechansism already found in
   TCP, SCTP and DCCP).  [EDITOR'S NOTE: reference RFC 5405/bis ]

3.4.  UDP-Lite

   A special class of applications can derive benefit from having
   partially-damaged payloads delivered, rather than discarded, when
   using paths that include error-prone links.  Such applications can
   tolerate payload corruption and may choose to use the Lightweight
   User Datagram Protocol (UDP-Lite) The service may be multicast and/or
   unicast

   [EDITOR'S NOTE: compare to UDP]

   [RFC3828] and [RFC 5405/bis]

3.5.  Datagram Congestion Control Protocl (DCCP)

   The Datagram Congestion Control Protocol (DCCP) [RFC4340] is a
   bidirectional transport protocol that provides unicast connections of
   congestion-controlled unreliable messages.  DCCP is suitable for
   applications that transfer fairly large amounts of data and that can
   benefit from control over the tradeoff between timeliness and
   reliability.

   [EDITOR'S NOTE: Describe the aspects(?) of DCCP...]

   [FC4340 et al]

3.6.  Realtime Transport Protocol (RTP)

   RTP provides an end-to-end network transport service, suitable for
   applications transmitting real-time data, such as audio, video or
   data, over multicast or unicast network services, including TCP, UDP,
   UDP-Lite, DCCP.

   [EDITOR'S NOTE: Describe the aspects(?) of RTP...]

3.7.  Hypertext Transport Protocol (HTTP) as a pseudotransport

   HTTP provides end-to-end network unicast transport service.

   [EDITOR'S NOTE: Reference BCP 56, note that this implies TCP but also
   brings with it object semantics you may not want.]

3.7.1.  WebSockets

   [EDITOR'S NOTE: point out how websockets kind of fixes this.]

4.  Transport service components

   Aspects as derived from the subsections above.

   This section is blank for now.

5.  Acknowledgements

   The authors were part-funded by the European Community under its
   Seventh Framework Programme.  The views expressed are solely those of
   the authors.

   Comments are welcome to the authors or via the IETF TAPS mailing
   lists.

6.  IANA Considerations

   XXX RFC ED - PLEASE REMOVE THIS SECTION XXX

   This memo includes no request to IANA.

7.  Security Considerations

   This document introduces no new security considerations.  Each RFC
   listed in this document discusses the security considerations of the
   specification it contains.

8.  References

8.1.  Normative References

   [RFC0791]  Postel, J., "Internet Protocol", STD 5, RFC 791, September
              1981.

   [RFC2460]  Deering, S. and R. Hinden, "Internet Protocol, Version 6
              (IPv6) Specification", RFC 2460, December 1998.

8.2.  Informative References

   [RFC0768]  Postel, J., "User Datagram Protocol", STD 6, RFC 768,
              August 1980.

   [RFC0793]  Postel, J., "Transmission Control Protocol", STD 7, RFC
              793, September 1981.

   [RFC0896]  Nagle, J., "Congestion control in IP/TCP internetworks",
              RFC 896, January 1984.

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

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

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

   [RFC5405]  Eggert, L. and G. Fairhurst, "Unicast UDP Usage Guidelines
              for Application Designers", BCP 145, RFC 5405, November
              2008.

Authors' Addresses

   Godred Fairhurst
   University of Aberdeen
   School of Engineering, Fraser Noble Building
   Aberdeen  AB24 3UE
   UK


   Email: gorry@erg.abdn.ac.uk


   Brian Trammell
   ETH Zurich
   Gloriastrasse 35
   Zurich  8092
   CH


   Email: ietf@trammell.ch