Taps Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 4, 2015

J. You
Huawei

June 2, 2015

3RED Model for TAPS

draft-you-taps-3red-model-00

Abstract

This document defines a 3RED model to describe transport service features. Applications can make use of the 3RED model to request transport services from the TAPS by sending a request which contains the explicit 3RED requirement parameters. The purpose of this document is to enable applications to make use of various transport services without customization or re-implementation.

Requirements Language

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].

Table of Contents

1. Introduction

Application programmers face difficulty when they use the transport services provided by the transport protocols other than TCP or UDP, such as SCTP, DCCP, WebSockets, MPTCP, UDP-Lite, etc. The goal of the TAPS working group is to help application and network stack programmers by describing an (abstract) interface for applications to make use of transport services.

This document is to define a 3RED model to describe transport service features. Applications can make use of the 3RED model to request transport services from the TAPS by sending a request which contains the explicit 3RED requirement parameters. The purpose of this document is to enable applications to make use of various transport services without customization or re-implementation.

2. Terminology

This section contains definitions of terms used in this document.

3RED: Rate, Response, Reliability, Efficiency and Differentiation

BDP: Bandwidth-Delay Product

DCCP: Datagram Congestion Control Protocol

MPTCP: Multipath TCP

RTO: Retransmission Timeout

RTT: Round-Trip Time

SCTP: Stream Control Transmission Protocol

TCP: Transmission Control Protocol

UDP: User Datagram Protocol

UDP-Lite: Lightweight User Datagram Protocol

3. 3RED Model

```
                 +------------+
                 |            |
                 |    APPs    |
                 |            |
                 +-----+------+
                       |
                       |
                   3RED|Model
                       |
                       |
                 +-----v------+
                 |            |
                 |    TAPS    |
                 |            |
                 +-----^------+
                       |
                       |
                       |
```

```
              |
              |
        +-----v------+
        | Transport  |
        | Protocols  |
        |            |
        +------------+
```

Figure 1: 3RED Model

The APP extracts or derives the service requirement parameters (i.e. 3RED model) of the requested service, and then requests transport

services from the TAPS by sending a request which contains the explicit 3RED requirement parameters. The usage of 3RED model is depicted in Figure 1.

3.1. Rate

Rate, also known as bandwidth utilization rate, is the percentage of the access bandwidth (in bit/s) that goes to the actually achieved throughput when transmitting enough data.

For example, TCP Reno throughput is measured as:

```
Throughput <= min ( BW, WindowSize/RTT, MSS/(RTT*sqrt(p)) )
```

BW: maximum bandwidth

WindowSize: congestion window size

RTT: round trip time

MSS: maximum segment size (fixed for each Internet path, typically 1460Bytes)

p: packet loss probability

For basic 4K streaming, it requires TCP throughput no less than 45Mbps. Assuming 100M network with RTT=50ms; p=0.01%; MSS=1460Bytes; WindowSize=256KBytes. Then,

```
            TCP Throughput <= min ( 100, 40.96, 23.36 )
                           <= 23.36Mbps
```

As we can see TCP responds to loss of packets and RTT, this results in poor throughput. TCP is inefficiency in high bandwidth-delay product (BDP) networks. It has an inherent throughput bottleneck that becomes obvious and severe with increased packet loss and latency.

Table 1 shows the throughput and bandwidth utilization rate of different congestion algorithms under different packet loss conditions on a 100Mbps link with RTT=100ms when downloading a file. The simulation topology is shown in Figure 2.


```
    +----------+    100M +----------+             +----------+
    | Sender   +---------+ Switch   |     +---------+ Receiver|
    +----------+         +-----+----+     |         +----------+
                               |100M      |
                               |          |
                               |          |
                               |      +---+-----+
                               +-------+ Network |
                                      | Emulator|
                                      +---------+
```

Figure 2: Simulation Topology

Table 1: Throughput and Rate

	p=0.1%		p=0.01%		p=0.005%	
	Throughput	Rate	Throughput	Rate	Throughput	Rate
Cubic	641 KB/s	5.26%	3.14 MB/s	26.38%	4.93 MB/s	41.41%
Westwood	1.19 MB/s	10.00%	3.97 MB/s	33.25%	5.41 MB/s	45.44%

Veno	637 KB/s	5.23%	1.99 MB/s	16.72%	2.80 MB/s	23.52%
Illinois	3.12 MB/s	26.21%	10.2 MB/s	85.68%	9.82 MB/s	82.49%

So Rate can be graded into 5 levels, as shown in Table 2.

Table 2: Rate

Rate Level	Utilization Rate
5	80%-100%
4	60%-80%
3	40%-60%
2	20%-40%
1	0%-20%

3.2. Response

Response (i.e. response time) is the time elapsed between the emission of the first bit of a data and the accumulated data that is able to make a service usable.

For web browsing, the response time is the time elapsed between the emission of the first bit of a web page and the reception of the last bit of that web page. For example, with an initial window of 3

segments (IW3), a transfer of 25 segments of web page will require 5 round trips to complete. While with the larger initial window (e.g. IW=10 [RFC6928]), it will require only 2 round trips. An increase of the initial window from 3 segments to 10 segments improves the response latency.

For 4K streaming video, the response time is the time elapsed between the emission of the first bit of a data and the accumulated data that is ready to play 4K video.

For most non real time applications, such as web browsing, email, FTP, a response time of 2-5 seconds is suggested. For most real time applications, such as video conferencing, videophony, VoIP, a response time of less than 150ms is suggested [QoS-Requirements].

Table 3 shows the average response time of different congestion algorithms when downloading files with different size on a 100Mbps link with IW=10, meanwhile assuming a packet loss occurs during the first RTT. The simulation topology is the same as shown in Figure 2.

Table 3: Response Time

	Object Size 10K Bytes	Object Size 40K Bytes	Object Size 100K Bytes	Object Size 256K Bytes	Object Size 3.75M Bytes
Cubic	3RTT	4RTT	8RTT	16RTT	85RTT
Westwood	3RTT	5RTT	10RTT	18RTT	110RTT
Veno	3RTT	4RTT	8RTT	13RTT	98RTT
Illinois	3RTT	5RTT	9RTT	15RTT	50RTT

Response can be graded into 5 levels, as shown in Table 4.

Table 4: Response

Response Level	Response Time
5	<10ms
4	10-100ms
3	100-400ms
2	400-5000ms
1	>5000ms

3.3. Reliability

Reliability is measured by a maximum application-level packet loss rate that is tolerable for the applications.

For web-browsing using HTTP over TCP, the expected data loss of HTTP objects is zero since TCP is a reliable transfer protocol in which erroneous packets are sent again using a retransmission policy. For

HDTV quality MPEG-2 video streams, the tolerable loss rate should be less than 0.0001%. For broadcast audio streams, a tolerable loss rate of no more than 0.1% is recommended. For VoIP, the tolerable loss rate should be less than 1% in the G.729 codec [QoS-Requirements]. When addressing the QoS needs of streaming video traffic, loss should be no more than 5%.

So Reliability can be graded into 5 levels, as shown in Table 5.

Table 5: Reliability

Reliability Level	Tolerable Loss Rate
5	0%
4	0%-0.1%
3	0.1%-1%
2	1%-5%
1	>5%

3.4. Efficiency

Efficiency, also known as bandwidth utilization efficiency, is the percentage of network bottleneck bandwidth (in bit/s) that goes to the actually achieved throughput when a number of competing connections traverse the same network bottleneck. [Performance-Evaluation] evaluates a set of TCP variants over a single bottleneck network.

Table 6 shows the efficiency of different congestion algorithms under different packet loss conditions when downloading files using different number of flows. The simulation topology is shown in Figure 3, Access Bandwidth= 1Gbps, Bottleneck Bandwidth=100Mbps, RTT=100ms.


```
  +----------+      1G +----------+                           +----------+
```

```
 | Sender  +---------+ Switch  |       +---------+ Receiver|
 +---------+         +----+----+       |         +---------+
                          |100M        |
                          |            |
                          |            |
                          |        +---+-----+
                          +-------+ Network |
                                  | Emulator|
                                  +---------+
```

Figure 3: Simulation Topology

Table 6: Efficiency

	p=0.1%:			p=0.01%		p=0.005%	
	1 flow	5 flows	10 flows	1 flow	5 flows	1 flow	5 flows
Cubic	5%	23%	43%	25%	100%	43%	100%
Westwood	9%	67%	100%	33%	100%	40%	100%
Veno	5%	26%	53%	17%	85%	23%	100%
Illinois	26%	100%	100%	75%	100%	96%	100%

Efficiency can be graded into 5 levels, as shown in Table 7.

Table 7: Efficiency

Efficiency Level	Efficiency
5	90%-100%
4	60%-80%
3	30%-60%
2	10%-30%
1	0%-10%

3.5. Differentiation

Differentiation means that a number of competing connections are provided with differential opportunity to transfer data when they sharing a single bottleneck link.

To provide a numerical measure reflecting the differential share distribution across the various connections, the differentiation index for one specified connection is defined as:

```
   (maximum throughput from all competing connections
             - the throughput of the specified connection)
---------------------------------------------------------------
     maximum throughput from all competing connections
```

For example, for three competing connections, their throughputs are 5Mbps, 4Mbps and 1Mbps respectively; then their differentiation indexes are 0, 20% and 80% respectively.

Differentiation determines whether users or applications are receiving a differential share of network resources when network is congested. Application may specify differentiation to individual media, in order to facilitate appropriate priority handling; for example, when the congestion is experienced, it informs the downstream nodes to downgrade their data transmission rate to some degree and vice versa.

Assuming beta is a constant multiplication decrease factor applied for window reduction at the time of loss in TCP Reno, the throughput is measured as:

```
Throughput <= min ( BW, WindowSize/RTT, sqrt(1/beta-1/2)*MSS/(RTT*sqrt(p)) )
```

So through adjusting beta, the throughput could be controlled. Accordingly, the differentiation level could be implemented.

Differentiation can be graded into 5 levels, as shown in Table 8.

Table 8: Differentiation

Differentiation Level	Differentiation
5	0%-10%
4	10%-30%
3	30%-50%
2	50%-70%
1	>70%

4. 3RED Implementation

3RED (i.e. Rate, Response, Reliability, Efficiency and Differentiation) usually cannot be satisfied at the same time. For example, if Reliability would be preferred, then Rate and Response have to be downgraded. When several flows share a single bottleneck link, if Response would be preferred, then Efficiency may be sacrificed.

The application could belong to end-user or service provider. For end-user, it may care more about Rate, Response and Reliability;

however for service provider, it may care more about Efficiency and Differentiation.

5. IANA Considerations

This document has no actions for IANA.

6. Security considerations

This document does not introduce any new security considerations.

7. Acknowledgement

TBD.

8. Normative References

[Performance-Evaluation]
Alrshah, M. and M. Othman, "Performance Evaluation of Parallel TCP, and its Impact on Bandwidth Utilization and Fairness in High-BDP Networks Based on Test-bed", 2013 IEEE 11th Malaysia International Conference on Communications, November 2013.

[QoS-Requirements]
Chen, R., Farley, T., and N. Ye, "QoS Requirements of Network Applications on the Internet", Information Knowledge Systems Management, January 2004.

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

[RFC6928] Chu, J., Dukkipati, N., Cheng, Y., and M. Mathis, "Increasing TCP's Initial Window", RFC 6928, April 2013.

Author's Address

Jianjie You
Huawei

101 Software Avenue, Yuhuatai District
Nanjing, 210012
China

Email: youjianjie@huawei.com