

Taxonomy and Survey of Retransmission Policies for Multipath Transport SCTP Protocol

Abdul Ridha H. Abass, Azizol Abdullah, Mohamed Othman, and Zurina Mohd Hanapi

Abstract—In this paper we present a taxonomy for retransmission policies that perform error recovery to offer completely reliable or partial reliable services for Multi Path Transport SCTP and quasi-SCTP. The taxonomy is drawn from a survey and an extensive analysis of the state of the art research works on retransmission policies for Multi Path Transport SCTP protocol. The taxonomy comprises two classification schemes: one that classifies retransmission policies with respect to retransmission path designation and one that classifies them with respect to their decision base scheme to select retransmission path which is the most salient feature of the retransmission policies for protocols that supports Multi Homing at the Transport layer. The sole purpose of our taxonomy is to gain valuable insights into these retransmission policies and thereby suggest baseline for future research. Based on our taxonomy, a survey is made of existing retransmission policies. The survey shows how retransmission policies are categorized according to our taxonomy, and exemplifies the majority of retransmission policies detailed in our taxonomy.

Index Terms—Taxonomy, survey, retransmission policy, multipath, SCTP.

I. INTRODUCTION

Reliable Transport protocols (e.g., Control Transmission Protocol (TCP) [1] and Stream Control Transmission Protocol (SCTP) [2]) was primarily designed to offer a completely reliable services appropriate for such applications as email (SMTP), file transfer (FTP), remote login (TELNET) and World Wide Web (HTTP). Whereas, Selective Retransmission Protocol (SRP) [3] and Partial Reliable Stream Control Transmission Protocol (PR-SCTP) [4] are instances of Partial Reliable Transport protocols that offer partial reliable services suitable for such applications as audio and video stream (i.e., its content state is highly sensitive to timeliness). To ensure reliable delivery a variety of error control mechanisms (e.g., sequence and acknowledgment numbers, and Error Correction mechanism) are used. Backward Error Correction (BEC) and Forward Error Correction (FEC) are typical Error correction mechanisms. In addition, BEC and FEC may be combined, such that lost packet and major errors are

recovered via a BEC and minor errors are corrected without retransmission [5].

FEC mechanism offers the sender encodes the payload using an Error Correcting Code (ECC) (e.g.,)

Parity bits, Checksums, Cyclic Redundancy Check (CRC) and Longitudinal Redundancy Check (LRC)) prior to transmission. The additional information added by the ECC is used by the receiver to recover the "most probably" original payload. FEC mechanism imposes less bandwidth overhead than BEC and require more processing and overhead to recover from errors. FEC is a better solution when bandwidth should be optimized and errors are rare [5].

Conversely, BEC mechanism also denoted as Automatic Repeat request (ARQ) uses retransmission mechanism explicitly designed to recover from loss or corrupted packet. The sending host maintains a copy of every transmitted packet while it is waiting for an Acknowledgement (ACK) for that packet. At the same time, it also starts a T3-rtx timer if an ACK is received before the timer expires the stored copy of the transmitted packet is discarded. Next time packet transmitted the T3-rtx timer restarted, in case that the receiver does not receive the transmitted packet it will not send ACK for that particular lost packet. In this case, the timer T3-rtx at the transmitter will go off after the expiry of Retransmission Time Out (RTO). On timeout, the transmitter assumes that the packet was lost and retransmits the copy of the packet. Every time the retransmission is sent The RTO is exponentially back off (i.e., $RTO_{new} = \min(RTO_{max}, RTO_{old})$) to space out repeated retransmissions of the same lost packet. Retransmission will continue until either successful transmission has occurred or the sender reaches the maximum number of retransmission attempt (e.g., most Linux hosts configured to send a maximum of 15 attempts [6]).

On the contrary, Fast retransmit algorithm is first appeared in TCP Tahoe [7] to avoid inefficiency caused by RTO. If a sender receives a specified number of acknowledgements; the default is three duplicate acknowledgements; with the same sequence number. In consequence, the sender can assume that particular packet was lost. The sender will then retransmit the packet that was presumed lost before waiting for its timeout.

Packet loss can be detected in one of two ways, either by the reception of triple-duplicate acknowledgements with the same sequence number or via timeouts. Triple-duplicate acknowledgements are interpreted as a sign of network congestion, whereas timeouts implies either severe congestion or failure en route.

The Stream Control Transmission Protocol (SCTP) [2], [8] is an IETF (Internet Engineering Task Force) standards track transport layer protocol that supports Multi Homing (i.e., a host can be addressed by multiple IP addresses). The

Manuscript received September 27, 2013; revised March 3, 2014.

Abdul Ridha H. Abass is with University Putra Malaysia, Serdang 43400, Malaysia, and University of Baghdad, Baghdad 10071, Iraq (e-mail: abdulridha.abass@gmail.com).

Azizol Abdullah, Mohamed Othman, and Zurina Mohd Hanapi are with University Putra Malaysia, Serdang 43400, Malaysia (e-mail: {azizol, mothman, zurina}@upm.edu.my).

motivation for Multi Homing in SCTP is fault tolerance, therefore we use SCTP in our taxonomy because of its relative maturity [9] and our focus on fault tolerance. In addition, SCTP provides TCP-like reliability, congestion, and flow-controlled data transfer to applications [10].

This paper provides a classification and survey of existing Multi Path Transport (MPT) retransmission policies. Nevertheless, TCP’s retransmission policy is used as benchmark to other policies. The paper shows how the existing policies are classified with respect to our taxonomy. Furthermore, the paper presents a survey of the classified retransmission policies.

The remainder of this paper is organized as follows: Our taxonomy is presented in Section II. Section III gives a survey of existing retransmission policies for MPT protocols and shows how they are classified with respect to our taxonomy. Finally, Section IV concludes the paper with a brief summary of our proposed taxonomy and a discussion of the insights gained in developing the taxonomy. Furthermore, briefly considers how these insights can be used in the design of future retransmission policies.

II. THE TAXONOMY

Our taxonomy consists of two classification schemes: one that classifies retransmission policies with respect to retransmission path designation and the other classifies them with respect to their decision base to select retransmission path. The former classification scheme is presented in Subsection A and the latter in Subsection B.

A. Classification with Respect to Retransmission Path Designation

Fig. 1, depicts the classification scheme with respect to the retransmission path designation. As follows from Fig. 1, the retransmission policies are classified along three dimensions: Hardwired, Preconfigured and Adaptive retransmission path.

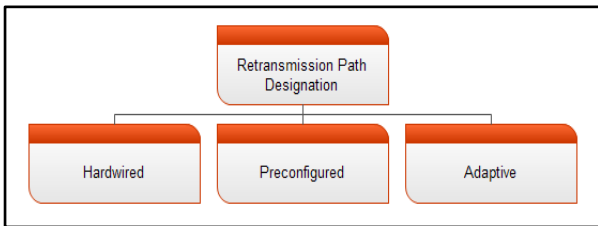


Fig. 1. Classification with respect to retransmission path designation.

- 1) *Hardwired retransmission path*: The dimension hardwired retransmission path refers to the protocol built-in retransmission path. Put differently, Single Homed reliable protocols offers single path (i.e., single peer IP address) for new payload and retransmissions. In congested situation throughput is degraded acutely. In such situation, the retransmission may also be lost, but a lost retransmission cannot be detected without timeout if there are insufficient triple-duplicate ACKs to trigger the Fast retransmit algorithm [11]. Whereas, during failure course (e.g., unreachable peer IP address)

TCP connections will timeout and abort, dictate the application to recover. Connection’s recovery overhead and associated delay can be unacceptable for such task-critical applications (e.g., e-health, e-commerce and emergency services) which should eliminate Single Points Of Failure (SPOF).

- 2) *Preconfigured retransmission path*: Multi Homed protocols offer a preconfigured retransmission path. In the course of association startup, two parties negotiate to designate only a single path denoted as primary path for new payload transmission. Whereas, the alternative paths remain for backup and just used for retransmission [2]. In other words, retransmission policy is designed to offer a constantly preconfigured retransmission path until the sender detects failure and fails over to alternative active path. Consequently, a SCTP sender cannot probe accurately alternative path’s conditions due to RTO stale of alternative paths[12].
- 3) *Adaptive retransmission path*: The last dimension refers to the situation in which, during the lifetime of a session, a retransmission policy can use any active path for retransmissions. In CMT, new data is regularly being sent into all paths. In consequence, a CMT sender maintains more timeliness and more accurate path’s conditions of all paths. Timeliness path’s conditions allow a CMT sender to use metric basis for making more informed decision where to send a retransmission.

B. Classification with Respect to Decision Base Scheme

This section explains how retransmission policies are classified with respect to their decision base scheme. The decision base comprises the metrics, rules, and/or heuristics that form the mechanism to select retransmission path. Fig. 2, illustrates the classification scheme with respect to the decision base. This dimension has three main classes: Naïve, Retransmission based and Metric based.

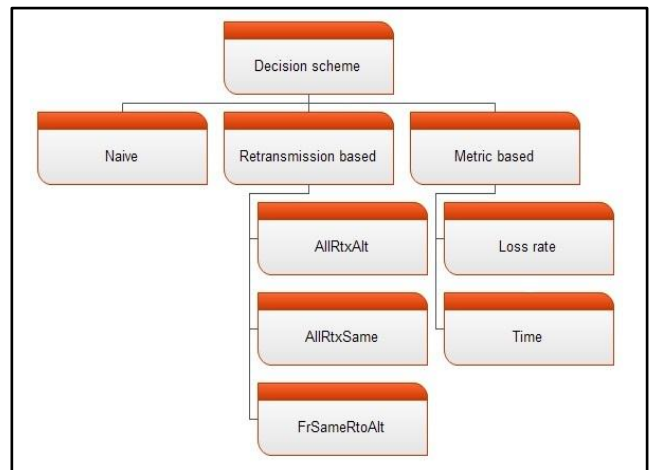


Fig. 2. Classification with respect to decision base.

- 1) *Naïve*: The dimension Naïve refers to the decision base in which the sender transmits all retransmissions on the same path used for the original transmission during the lifetime of connection. This is the case in Single Homed reliable protocol. While a Naïve decision base is simple to implement, its cost is proportional to network congestion and path failures which in practical tends to deteriorate throughput severely as the network congestion increases. However, if the path becomes unreachable, retransmission

policy will not successfully deliver any data. Naïve decision base is also useful as "baseline" method when benchmarking other decision base.

2) *Retransmission based*: Protocols supports Multi Homing at Transport layer offer more complicated decision scheme rely on retransmission type. All Rtx Alt, All Rtx Same, and Fr Same Rto Alt are main retransmission policies with respect to retransmission based [13].

a) *All Rtx Alt*: All retransmissions are sent on the alternate paths. This policy represents SCTP standard [2] to provide a very seamless resolution of transient network congestion and path failures. A drawback is that alternate paths often have undue conservative RTOs, which significantly degrades performance when retransmissions of lost packets themselves are lost.

b) *All Rtx Same*: All retransmissions are sent on the same path used for the original transmission. This policy often improves performance in non-failure scenarios (i.e., assume availability of all peer IP addresses at all times) by using the path with the most accurate RTO [13]. However, if the primary path becomes unreachable, this policy will not successfully deliver any data until the sender detects failure and fails over into to alternate paths.

c) *Fr Same Rto Alt*: Fast retransmissions are sent on the same path, but timeout retransmissions are sent on an alternate *path*. This policy is the best overall policies in non-failure scenarios. As loss rates on *primary* path are relatively lower than the alternate path, most of the losses are detected by Fast retransmit. Therefore, Fr Same RtoAlt will send most of its lost packets to the same path as All Rtx Same. More, the Fast retransmissions do not experience overly conservative RTOs for the alternate path, as they would with All Rtx Alt. Fr Same Rto Alt is able to alleviate severe loss conditions on the primary by sending timeout retransmissions to the alternate path. In contrast, All Rtx Same suffers at higher primary path loss rates by not using the alternate path. [13].

1) *Metric based*: As follows from its name, retransmission policy that select retransmission path based on direct or indirect measure of one or several such properties as packet's loss rate, congestion window (CWND), slow start threshold (SSTHRESH) or response time to retransmit are considered metric based retransmission policy. Metric based policy uses metrics to make more informed decision where to send a retransmission. Therefore, during the lifetime of a session a retransmission policy can use any active path for retransmission to attain high performance. CMT sender uses retransmission policies based on metrics. We distinguish between two classes of metric based retransmission policies:

retransmission policy that select retransmission path on some kind of estimate of the packet's loss rate or that select retransmission path based on time.

a) *Packet's loss rate*: To our best knowledge, existing retransmission policies use indirect approach to estimate the packet's loss rate. The indirect approach involves using such metrics as CWND or SSTHRESH to estimate the packet's loss rate condition. In contrast, the direct approach involves using a statistic (e.g., the arithmetic average or mean packet's loss rate) to calculate the average packet's loss rate over a fixed-length sequence of packets. Path's CWND is reflecting most recently knowledge of the network path's conditions while SSTHRESH is keeping some memory about the earlier conditions of the path. Intuitively, a larger CWND or SSTHRESH implies a less packet's loss rate. Since, the SSTHRESH is slower moving variable than CWND its value may better reflect the conditions of the respective path [14], [15].

b) *Rtx-CWND*: A retransmission of a packet is sent on the path for which the sender has the largest CWND. A tie is broken randomly.

c) *Rtx-SSTHRESH*: A retransmission of a packet is sent on the *path* for which the sender has the largest SSTHRESH. A tie is broken randomly.

d) *Rtx-LOSSRATE*: A retransmission of a packet is sent on the path with the lowest loss rate path. If multiple paths have the same loss rate, one is selected randomly. Rtx-LOSSRATE represents hypothetical ideal retransmission policy; hypothetical since in practice a sender typically does not know a priori path's loss rate; ideal since packet retransmission over the lowest loss rate path has the highest probability to deliver [14], [15].

e) *Time*: Time based retransmission policies comprise those policies that employ metric which is a function of time such as *response* time. For example, Rtx-ASAP, retransmit lost packet on path which is likely minimizes response time to retransmit in order to improve the performance.

f) *Rtx-ASAP*: A retransmission of a packet is sent on a destination for which the sender has CWND space available at the time the retransmission needs to be sent. If the sender has available CWND space for multiple paths, one is chosen randomly [14], [15].

III. A CLASSIFICATION AND SURVEY OF EXISTING RETRANSMISSION POLICIES

This section surveys existing retransmission policies and classifies them with respect to our taxonomy. Table I, shows how the retransmission policies are classified.

In [13], [16] proposed additional retransmission policies which are combinations of the three basic policies All Rtx Alt, All Rtx Same, and Fr Same Rto Alt and three algorithms Heartbeat After RTO (HAR), Time Stamps (TS), and Multiple Fast Retransmit (MFR) for transport protocols that support Multi Homing . All Rtx Alt+HAR, All Rtx Alt+TS, All Rtx Same, All Rtx Same+MFR, Fr Same Rto Alt+HAR, Fr Same

RtoA It+TS, Fr Same Rto Alt+MFR and Fr Same Rto Alt+MFR+HAR are evaluated under both failure and non-failure scenarios. Researchers concluded that the best performance achieved by policy dictates that new data transmissions and Fast retransmissions should be sent on the same path which they were originally sent, and timeout retransmissions should be sent on an alternate path. Further, policy performs best if combined with MFR (e.g., Fr Same RtoAlt+MFR) attains the best performance under the evaluated scenarios.

TABLE I: CLASSIFICATION OF RETRANSMISSION POLICIES IN OUR TAXONOMY.

Protocol	Retransmission policy	Retransmission path designation	Decision base
TCP ¹	Na ĩve	Hardwired	Na ĩve
SCTP	AllRtxAlt ²	Preconfigured	Retransmission
	AllRtxSame	Preconfigured	Retransmission
	FrSameRtoAlt	Preconfigured	Retransmission
CMT-SCTP	Rtx-LOSSRATE	Adaptive	Metric (Loss rate)
	Rtx-CWND	Adaptive	Metric (Loss rate)
	Rtx-SSTHRESH	Adaptive	Metric (Loss rate)
	Rtx-ASAP	Adaptive	Metric (Time)
CMT-PF-SCTP	CMT-PF-CWND	Adaptive	Metric (Loss rate)
	CMT-PF-SSTHRESH	Adaptive	Metric (Loss rate)
¹ Na ĩve TCP retransmission policy is a benchmark policy.			
² SCTP standard retransmission policy.			

In an investigation into the impact of path bandwidth significantly variant and constrained receive buffer on the performance of Fast retransmission policies in MultiHoming environments, [17] identifies that Fast retransmission on an alternate path may cause performance degradation due to receive buffer blocking. A theoretical model is proposed and verified using simulation for recruiting retransmission path during the Fast retransmission phase, which is based on path's conditions and receive buffer available space.

To alleviate receive CMT-Buffer Blocking effect during the congestion episode, [14] proposed five different retransmission policies and evaluated them in non-failure scenarios. Research conducted under strong assumption that receiver's buffer does not constrain the sender and a CMT use disjoint paths. Proposed retransmission policies use heuristics to reduce loss recovery durations. RTx-SAME retransmission of the packet is sent to the same destination until the destination is marked inactive due to failure. RTx-ASAP a retransmission of a packet is sent to any destination for which the sender has congestion window space attainable at the time of retransmission. RTx_CWND send a retransmission to the destination with the largest

congestion window. RTx-LOSSRATE A retransmission of a packet is sent to the destination for which the sender has the lowest loss rate. Finally, RTx-SSTHRESH a retransmission of a packet is sent to the destination for which the sender has the largest SSTHRESH. Over similar paths both RTx_CWND and RTx_SSTHRESH perform equally well during congestion-induced receiver's CMT-Buffer Blocking, and therefore, arbitrarily selected and recommended the RTx_SSTHRESH policy for CMT. However, they cannot solve the cause of the problem. In [18], they showed that CMT-SCTP with RTx_SSTHRESH policy suffers from consecutive instances of failure-induced receiver's CMT-Buffer blocking. Conclusions were vague, but the authors' perspective of retransmission policies that does not take the loss rate of each path into account will not be able to tackle the receiver's CMT-Buffer Blocking problem [19]. We argue that heuristics does not always provide a good approximation of packet's loss rate. In contrast, statistical (e.g., the arithmetic average or mean packet's loss rate) always provide a good approximation of packet's loss rate.

The [14] hypothesis that retransmission policies would best mitigate receiver's CMT-Buffer Blocking by a cutback the timeouts and the back-to-back timeouts recovery durations stimulated [15] to re-optimized retransmission policies. They confirmed conclusion drew by [14] and proposed a new policy called RTx-CSL which uses RTx_CWND, RTx_SSTHRESH and RTx-LOSSRATE to be smarter to select retransmission path for further vitiate some of the throughput degradation. Our perspective, RTx-CSL represents a tie break strategy rather than a new retransmission policy. It is basically operates as RTx_CWND retransmission policy.

The likely of path failures in practice motivates [20] to study their impact on CMT. They identified the failure-induced receiver's CMT-Buffer Blocking problem in CMT-SCTP, and proposed CMT with a Potentially-Failed destination state (CMT-PF) to alleviate receiver's CMT-Buffer Blocking during a path failure episode due to faster failure detection (i.e., reduces the path failure detection delay which is the time from a path failure until it has been marked as inactive path). [10], [18], [20] demonstrated CMT-PF's throughput gains over CMT during complete and short-term failures scenarios. More, CMT-PF's ability to avoid back-to-back timeouts on data improves its performance over CMT was demonstrated. The RTx-SSTHRESH and RTx-CWND are considered for evaluations with PF (i.e, CMT-PF-CWND and CMT-PF-SSTHRESH which employ CMT-PF with RTx-CWND and RTx_SSTHRESH retransmission policy, respectively) for Fast retransmission in realistic loss model with symmetric and asymmetric path loss and varying path RTTs [20].

IV. CONCLUSION

In summary, this paper presents ataxonomy for retransmission policies of MultiPath Transport SCTP protocol. This taxonomy provides a framework for the comparison and evaluation of these policies. In addition, the insight provided by the taxonomy and survey in this paper may be used to guide future research in this area. Studying retransmission policies is still an active area of research and requires further investigation. Making a smart decision where to send a

retransmission will help for further vitiate some of the throughput degradation. In addition, mitigate CMT-Buffer Blocking by a cutback the timeouts and the back-to-back timeouts recovery durations. Simulation results obtained by developing a new policy and a theoretical study of the MultiPath Transport SCTP protocol will be discussed and developed in the forthcoming paper.

REFERENCES

[1] J. Postel, *Transmission control protocol*, 1981.
 [2] R. R. Stewart *et al.*, *RFC 2960: Stream control transmission protocol*, Status: Standard, October 2000.
 [3] M. T. Piecuch *et al.*, "Selective retransmission protocol for multimedia on the Internet," *Information Technologies 2000*, 2001, pp. 79-90.
 [4] R. Stewart *et al.*, "Stream control transmission protocol (SCTP) partial reliability extension," RFC 3758 (Proposed Standard), 2004.
 [5] C. T. Bhunia, *Information Technology Network And Internet: New Age International (P) Limited*, Publishers, 2008.
 [6] J. Postel, *RFC 793: Transmission control protocol, September 1981*, Status: Standard, 2003.
 [7] V. Jacobson, "Congestion avoidance and control," *ACM SIGCOMM Computer Communication Review*, 1988, pp. 314-329.
 [8] J. AL Caro *et al.*, "SCTP: a proposed standard for robust internet data transport," *Computer*, vol. 36, pp. 56-63, 2003.
 [9] A. L. Caro Jr *et al.*, "Retransmission policies for multihomed transport protocols," *Computer Communications*, vol. 29, pp. 1798-1810, 2006.
 [10] P. Natarajan *et al.*, "Concurrent multipath transfer using transport layer multihoming: Performance under network failures," in *Proc. Military Communications Conference*, 2006, pp. 1-7.
 [11] R. Stewart and Q. Xie, *RFC 4960, Stream control transmission protocol*, Network Working, 2007.
 [12] C. Y. Ho *et al.*, "Fast retransmit and fast recovery schemes of transport protocols: A survey and taxonomy," *Computer Networks*, vol. 52, pp. 1308-1327, 2008.

[13] J. Noonan *et al.*, "Stall and Path Monitoring Issues in SCTP," *Infocom*, 2006.
 [14] J. AL Caro *et al.*, "Retransmission schemes for end-to-end failover with transport layer multihoming," in *Proc. Global Telecommunications Conference*, 2004, pp. 1341-1347.
 [15] J. R. Iyengar *et al.*, "Retransmission policies for concurrent multipath transfer using SCTP multihoming," in *Proc. 12th IEEE International Conference on Networks*, 2004, pp. 713-719.
 [16] J. Liu *et al.*, "Reducing receive buffer blocking in concurrent multipath transfer," in *Proc. 4th IEEE International Conference on Circuits and Systems for Communications*, 2008, pp. 367-371.
 [17] J. AL Caro *et al.*, "Retransmission policies with transport layer multihoming," in *Proc 11th IEEE International Conference on Networks*, 2003, pp. 255-260.
 [18] Y. Qiao *et al.*, "Path selection of SCTP fast retransmission in multihomed wireless environments," *Wireless and Mobile Networking*, 2008, pp. 447-458.
 [19] P. Natarajan *et al.*, "CMT using SCTP multihoming: transmission policies using a potentially-failed destination state," *Technique Report*, 2007.
 [20] T. Wallace and A. Shami, "A review of multihoming issues using the stream control transmission protocol," *Communications Surveys & Tutorials*, IEEE, vol. 14, pp. 565-578, 2012.
 [21] P. Natarajan *et al.*, "Concurrent multipath transfer using SCTP multihoming: Introducing the potentially-failed destination state," in *Proc. Networking 2008 Ad Hoc and Sensor Networks, Wireless Networks, Next Generation Internet*, 2008, pp. 727-734.



Abdul Ridha H. Abass has received his B.Sc. and M.Sc. degrees in computer science from University of Baghdad, Baghdad, Iraq, in 1995 and 2000. He had been an assistant lecturer at University of Baghdad, College of science, computer science Dept. from 2004 to 2010. He is currently pursued for Ph.D. degree of computer science at Universiti Putra Malaysia, Serdang, Malaysia. His research interests include network and intelligent computing