Analysis of multi-homed SCTP and delay-centric transmission for VoIP traffic in lossy networks

Rudolfo A. E. Rüncos
Departamento de Engenharia Elétrica
Universidade Federal do Paraná - Curitiba, Brazil
Email: rrscar@gmail.com

Eduardo P. Ribeiro
Departamento de Engenharia Elétrica
Universidade Federal do Paraná - Curitiba, Brazil
Email: edu@eletrica.ufpr.br

***Abstract*—Nowadays a number of Internet access possibilities can be simultaneously available to fixed or mobile users. Adequate selection of the most appropriate communication path can improve transmission quality for Voice over IP (VoIP) applications by providing smaller delay and smaller packet loss. Stream Control Transmission Protocol (SCTP) is currently being studied as a good alternative for implementing this type of control due to its multi-homing capability. Different parameter configurations can affect performance of path selection mechanisms. In this paper the performance of delay-centric path selection method and SCTP standard mechanism to deal with path failure are analyzed. Communication performance of VoIP transmission is evaluated by comparing the session Mean Opinion Score (MOS) that is calculated using E-model from ITU-T. Simulated scenarios include packet losses and different levels of average delay on the available paths. Results show that delay-centric method combined with standard failure detection mechanism with non-default parameters performs well for most common cases.**

I. Introduction

Constant innovation in data transmission technologies has been increasing the different possibilities of internet access media. Devices capable of connecting simultaneously to more than one network, such as WiFi, 3G, WiMax or Ethernet, are becoming common. Multimedia services, which are particularly sensitive to network delay and losses, can benefit from this new reality. Mechanisms capable of identifying the best path among those available and perform a seamless handover are an important topic of research.

Several approaches have been proposed at different stack layers. According to Eddy [1], the most suitable layer for implementing such mechanism is the transport layer. SCTP (Stream Control Transmission Protocol) [2] provides a good framework for associations with multiple IP addresses and presents itself as an interesting alternative. Kelly [3] proposed a method for monitoring path delay and automatic handover called delay-centric. This method was also implemented and tested by Gavriloff [4]. Other methods [5] [6] estimate path MOS (Mean Opinion Score) and use it as parameter for handover decisions, which then also take packet losses into account. Crosslayer metrics are also investigated [7].

Nevertheless, SCTP has its own loss monitoring mechanism which aims at detecting path failure. Qiao [8] studied this mechanism for different retransmission policies and PMR (Path Maximum Retransmition) values in order to optimize throughput in lossy networks.

In this work, SCTP performance for different PMR values was evaluated with respect to VoIP (Voice over Internet Protocol) traffic between multi-homed hosts and considering scenarios with different average losses and delay. The influence of delay-centric path selection method in such scenarios was also studied. MOS was used as performance metric, calculated using E-model.

This paper is structured as follows. Section 2 brings a brief overview on the most relevant concepts concerning this work. Simulated scenarios are presented in Section 3. Results are exposed in Section 4. The paper ends with conclusions in section 5.

II. Concepts Overview

A. SCTP

Stream Control Transmission Protocol (SCTP), third transport protocol ratified by the Internet Engineering Task Force (IETF), is a connection oriented transport protocol that inherited several features from TCP (Transport Control Protocol), such as congestion avoidance and reliable and ordered data delivery. It allows selective acknowledgments (SACKS). Extensions to SCTP [9] allow unreliable transmissions and unordered delivery, which are features present in UDP.

One of the main features of SCTP is the concept of association, that is analogous to a TCP connection but spanning multiple IP addresses. This means that multi-homed hosts can achieve a more reliable association using SCTP, if more than one path is available.

When more than one path is available SCTP chooses one IP (path) and sets it as primary. All outgoing traffic is routed to the primary destination IP, unless the primary path becomes inactive due to failure or an upper layer requests the traffic to be sent to another destination. Thus, SCTP, through SACKs packets or probe packets called Heartbeats (HB), can monitor every path status, detect failures and switch to an alternate path when necessary in order to maintain the association.

PMR (Path.Max.Retransmitions) is the parameter that establish the tolerance for loss in each path, and is set by default to 5. A path is declared inactive if the number of consecutive retransmissions due to timeout becomes greater than the PMR value. In this work the performance of this failure detection mechanism is analyzed for different PMR values, concerning VoIP traffic.

B. Delay-centric

The delay-centric algorithm was first proposed by Kelly [3]. Its main goal is to provide better quality of service (QoS) for delay sensitive traffic, such as VoIP. Delay-centric algorithms work by constantly monitoring delay in every available path and by seamlessly setting as primary the path with the lowest delay. Delay monitoring on each path is performed by using SCTP's internal variable SRTT (Smoth Round Trip Time) values for each path. SRTT is calculated by averaging RTT (Round Trip Time) that is obtained when an acknowledge is received (from data packets on primary path or from heartbeats on alternate paths). In this work, a modified SCTP class on NS2 which implements this behavior was used [4].

C. Gilbert error model

The Gilbert error model consists of a Markov chain with two states, a "good" or lossless state, and a "bad" or lossy state, as illustrated in figure 1. This model is defined by three parameters, P1, P2 and L. P1 is the probability of going from "good" to "bad" state. P2 is the probability of going from "bad" to "good" state. L is the probability that a packet will be lost at "bad" state.

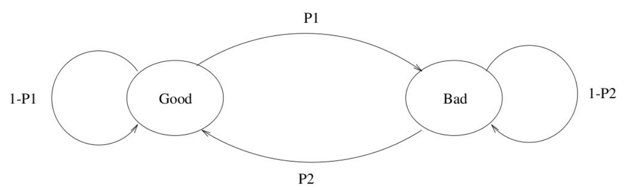

Fig. 1. A two-state Markov Chain represents the Gilbert error model.

Gilbert error model was used in this work because it represents a good compromise between realism, ease of implementation and computational cost [10] [11]. In this work, possible changes of state occur at constant time intervals. Theoretical average losses, L_{avg}, can be calculated by considering the limit when time is infinite:

$$L_{avg} = \frac{L.P1}{P1 + P2} \qquad (1)$$

D. E-model

For any voice service, the Mean Opinion Score (MOS) is an important metric. It can be evaluated after well conducted subjective tests and with a significant number of test subjects. However, these evaluations are expensive to be made and, for a large number of evaluations, they could become impracticable. E-model can provide a feasible alternative for the estimation of MOS in some cases.

The E-model is standardized by ITU-T as G.107 [12]. A scalar rating of call quality, the R factor, is the output of the E-model and is calculated based on several informations and measurements from the network:

$$R = R_0 - I_s - I_d - I_{e,eff} + A \qquad (2)$$

R_0 : Basic signal-to-noise ratio;
I_s : Impairments simultaneous to voice encoding;
I_d : Impairments due to network transmission;
$I_{e,eff}$: Equipment impairments, including packet loss;
A : Advantage factor.

The R factor can be converted to MOS values, according to the following equation:

$$MOS = \begin{cases} 1 & R < 0 \\ 1 + 0.035R + \\ +7R(R-60)(100-R) \times 10^{-6} & 0 < R < 100 \\ 4.5 & R > 100 \end{cases} \qquad (3)$$

It is important to highlight that this objective evaluation does not substitute a subjective MOS evaluation, but it can serve as a rough estimate when a more appropriate subjective test is impracticable.

In this work, MOS calculated from R factor is used as performance metric and comparison parameter at different simulated scenarios. The R factor is computed according to the formulas presented in G.107 [12] and SG12 D.106 [13]. I_d contains measured average delay from the network, $I_{e,eff}$ includes average packet loss and B_{pl} stands for the packet loss robustness factor, which is codec dependent.

III. Simulation setup

All simulations were run using NS2 (Network Simulator 2), a discrete event network simulation software widely used in computer network research. The version 2.32 with a modification for delay-centric SCTP [4] was used.

Figure 2 shows the simulated topology, which was the same for all simulations. All links are full duplex, with 1Mbps transmission rate and 30ms propagation delay. SCTP agents are attached to nodes 8 and 9. They are configured for unreliable (no retransmission), unordered delivery with no delayed SACKs. Moreover, $HBinterval$ = 0.1s, $RTOmin$ = 0.5s and $RTOinitial$ = 1s. All other parameters kept their default values. Path 0 was set as primary at the beginning of the simulation. Those parameters were the same for all simulations.

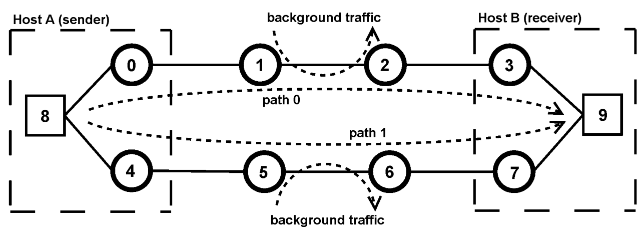

Fig. 2. Topology used in all simulations. Each link is full-duplex, with transmission rate of 1Mbps and 30ms of propagation delay. Background traffic between nodes 1 - 2 (path 0) and 5 - 6 (path 1). Losses occur in the links between nodes 2 - 3 and 6 - 7. Path 0 is set to primary at the beginning of each simulation.

Simulated VoIP traffic was considered to be G.711 [14]. In order to emulate it, a CBR (Contant Bit Rate) traffic generator was attached to SCTP agent of node 8, generating packets of 160ms each 20ms. A fixed anti-jitter buffer at the application level was considered. A playout delay of 100ms was defined. The delay threshold was set as delay of the first packet plus playout delay. Packets that did not meet this deadline were considered lost. It was also assumed the use of PLC (Packet Loss Concealing) algorithm as described in G.711 appendix 1 [15].

Background traffics were generated between nodes 1 and 2, and 5 and 6. They generated constant size UDP packets of 500 bytes at exponentially distributed random time intervals. The average time interval was calculated so that an arbitrary average delay could be achieved for this M/D/1 queue system (taking into account also the CBR traffic). Each VoIP packet then experienced a constant delay, resulting from the constant propagation and transmission delay at all three links (approximately 92ms) plus a variable queue delay at nodes 1 or 5, depending on the path.

Packet losses can occur at links between nodes 2 and 3, and 6 and 7, so background traffic packets were not discarded and losses would not affect average path delay. A Gilbert error model was used with time granularity of 5ms for all simulations. Its parameters were set in order to achieve specific desired average losses. Average error burst intervals were between 25ms and 250ms. Average lossless intervals were between 1s and 10s. L was set between 0.8 and 0.9 for error bursts.

At the beginning of each simulation, 5 parameters were set differently to generate a specific scenario: average queue delay for paths 0 and 1, average losses for paths 0 and 1, and a PMR value. Table I presents simulated parameters values. Each combination resulted in a distinct scenario. All scenarios were simulated with and without SCTP's delay-centric algorithm (DCon and DCoff). Each scenario was then simulated 7 times, each time using a different random number generator seed. VoIP traffic duration was 100s for all simulations.

TABLE I
VARIABLE SIMULATION PARAMETERS

PARAMETER	VALUES
Path 0 average queue delay (ms)	5, 10, 20, 30, 50, 100
Path 1 average queue delay (ms)	5, 10, 20, 30, 50, 100
Path 0 average losses (%)	0, 0.2, 0.5, 1, 2, 3, 5, 10, 15, 20
Path 1 average losses (%)	0, 0.2, 0.5, 1, 2, 3, 5, 10, 15, 20
PMR (unitless)	0, 1, 5

As G.711 with PLC algorithm was considered, Ie = 0 and Bpl = 25.1 were used in the E-model calculations. These values can be found on G.113 appendix 1 by ITU-T [16].

IV. RESULTS

Figure 3 is a graphical representation of all packet delay data from a single simulation. This simulation was from a scenario where average queue delay was 50ms for path 0 and 10ms for path 1, average losses for both paths were 0.5%, PMR was 1 and delay-centric algorithm was active. In this plot, each symbol (cross or circle) refer to a packet, with arrival time on the X axis and its end-to-end delay on the Y axis. The dashed line represents the playout delay threshold. Packets that fell above the dashed line were discarded and counted as additional losses for MOS calculation, which in this case was 4.08. All three path changes (0 to 1 and 1 to 0 alike) were caused by the delay-centric mechanism.

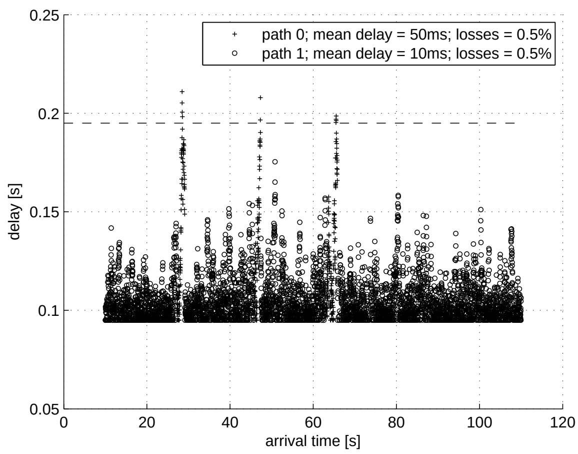

Fig. 3. Results from a single simulation. Circles are packets that arrived from path 0. Crosses are packets that arrived from path 1. X axis is the packet arrival time and Y axis is its delay. The dashed line is the jitter-buffer threshold. In this simulation, PMR = 1 and delay-centric algorithm is active.

A. *Delay-centric inactive*

As there were 5 variable parameters among all simulated scenarios, in order to evaluate results each generated graphic had values for average queue delays on both path fixed and average packet losse on path 0 fixed. X axis represents average losses on path 1 and Y axis represents MOS values (figure 4 is an example). Moreover, all three PMR values are represented by three distinct curves. Each point on the graphic is the mean from all 7 simulations run with different random number generation seeds.

For any scenario in which average losses on path 0 were less than 0.5%, there was no statistical difference in MOS for different PMR values. In such scenarios, path 0 rarely became inactive, even for low values of PMR, since average losses are too low. Thus, in most simulations where path 0 average losses were under 0.5%, all VoIP traffic flowed through path 0, and for all PMR values they presented the very same MOS.

For scenarios having average queue delay on path 0 equal or greater than that on path 1, low PMR (0 or 1) values performed better when path 0 average losses were greater than 0.5%. Figure 4 shows an example.

Low PMR values were detrimental to VoIP traffic only when average queue delay on path 0 was considerably lower than average queue delay on path 1 (difference of at least 20ms in simulated scenarios) and average losses on path 0 were between 1% and 5%. Low PMR values on these scenarios lead VoIP flow to path 1 more often than high PMR values.

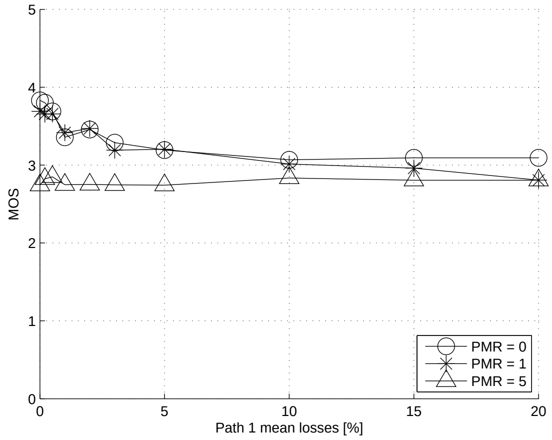

Fig. 4. Path 0: delay = 30ms and losses = 3%. Path 1: delay = 10ms.

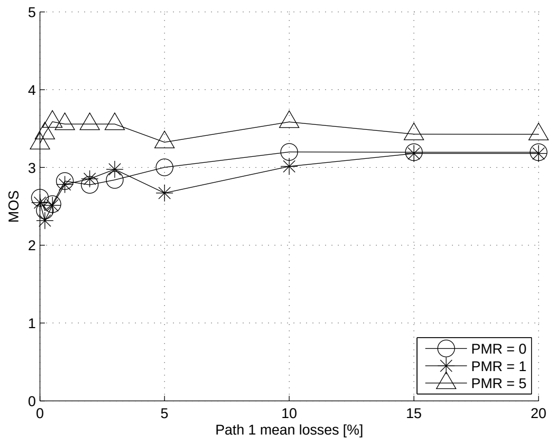

Fig. 5. Path 0: delay = 10ms and losses = 3%. Path 1: delay = 50ms.

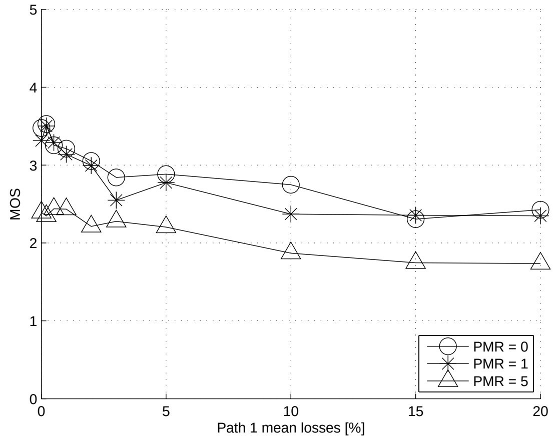

Fig. 6. Path 0: delay = 10ms and losses = 15%. Path 1: delay = 30ms.

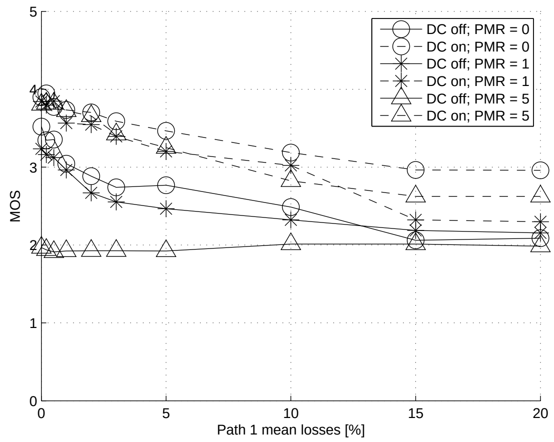

Fig. 7. Path 0: delay = 50ms and losses = 3%. Path 1: delay = 20ms.

As QoS on path 1 was worse, MOS was lower for low values of PMR. This can be observed on figure 5.

For average losses greater than 10% on path 0, low PMR values granted better MOS, independently of average queue delay, like the scenario exposed on figure 6. A fast response to loss is interesting in scenarios with high path 0 average loss ratings.

B. Delay-centric active

Analysis in this subsection focus on comparing scenarios that differ only by the status of their delay-centric algorithm (active or inactive). Therefore, from this point on all graphics will have 3 additional curves, representing all three PMR values for scenarios with active delay-centric.

The use of the delay-centric algorithm brought gain to MOS ratings in most cases. The greater was the queue delay on one or both paths, the greater the gain brought by this mechanism. Delay-centric helps in preventing moments of high delay by switching paths, which decrease the number of discarded packets due to the anti-jitter buffer threshold. Figure 7 illustrate this gain.

Path switching became a lot more frequent for scenarios with delay-centric activated, even when average queue delay on both paths were low. Also, VoIP traffic tended to flow longer through the lower average delay path, being this tendency stronger for higher differences in average queue delay between the two paths. This behavior explains the following two situations where the use of delay-centric was detrimental.

The first situation where delay-centric was not interesting happened when, on path 0, average queue delay was lower than 20ms and average losses were lower than 0.5%. If average queue delay and losses on path 1 were equally low, there was no difference in MOS. Otherwise delay-centric caused a reduction in MOS, due to more frequent path switching. Without delay-centric VoIP packets flowed almost exclusively through path 0, which had a higher QoS. Figure 8 is an example of this situation.

The second uninteresting situation happened when average queue delays on both paths were the same and equal to 30ms or lower, while average losses were lower than 3% on one path

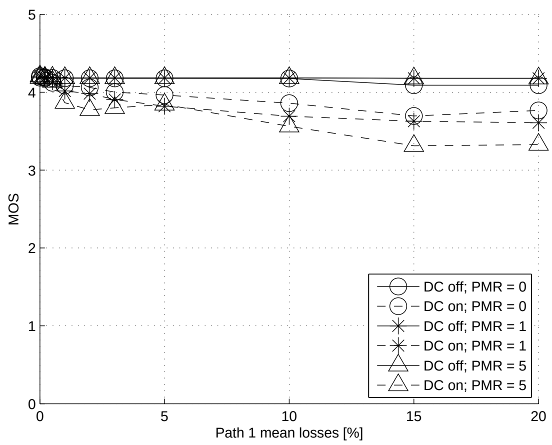

Fig. 8. Path 0: delay = 10ms and losses = 0,2%. Path 1: delay = 10ms.

and grater than 5% on the other, with at least 5% of difference (that is, there was a considerable asymmetry in average losses). In these cases delay-centric forced VoIP traffic to flow on both path, which would otherwise flow mostly through the lower average loss path for inactive delay-centric scenarios. This way, total losses are greater with delay-centric active and MOS decreases. Exceptions to this specific case, meaning that MOS increased by the use of delay-centric, happened for a PMR value of 5 and when average losses on path 0 were greater than on path 1. Figure 9 illustrate this second situation.

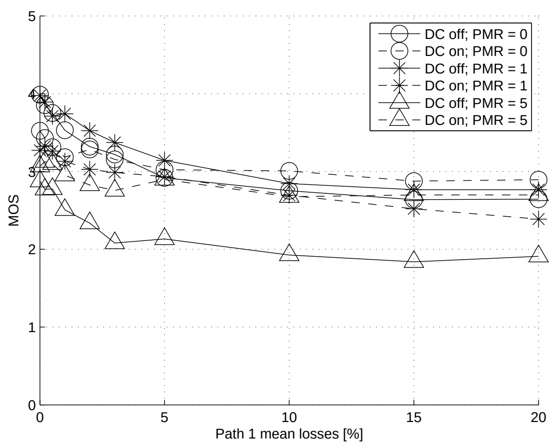

Fig. 9. Path 0: delay = 20ms and losses = 15%. Path 1: delay = 20ms.

V. Conclusion

SCTP standard failure detection mechanism was evaluated regarding VoIP traffic (G.711) through simulated scenarios with packet loss. The interaction and influence of delay-centric path selection method was also evaluated.

Results showed how failure detection mechanism reacts to packet loss. PMR value acts as a sensitivity threshold and influences reaction time (path switching) against loss bursts. Low values of PMR performed better in many simulations. However, this may not be the case when a high average delay path has low average loss ratings, because the failure detection mechanism can cause VoIP packets to flow through this higher delay path, degrading voice quality.

On the other hand, delay-centric algorithm is reactive to path delays but not to packet loss. This is advantageous in many cases, but when a lower average delay path presents high loss ratings this may reduce MOS severely. Also, frequent path switching caused by delay-centric method may degrade call quality even when there is a low average delay and loss path.

The use of lower PMR values (0 or 1) and delay-centric method simultaneously proved to be a good and robust alternative to maintain the quality of VoIP traffic. However, there are some specific situations in which this is not the best combination. Another conclusion was that delay-centric dominates the failure detection mechanism when average path delay asymmetry is intense. This is not the ideal behavior, since losses may be the main call quality degradation factor, depending on the situation.

References

[1] W. Eddy, "At what layer does mobility belong?" *Communications Magazine, IEEE*, vol. 42, no. 10, pp. 155–159, 2004.

[2] R. Stewart, I. Arias-Rodriguez, K. Poon, A. Caro, and M. Tuexen, "Stream control transmission protocol (sctp) specification errata and issues," *IETF RFC 4460*, no. 4460, April 2006. [Online]. Available: http://www.ietf.org/rfc/rfc4460.txt

[3] A. Kelly, G.-M. Muntean, P. Perry, and J. Murphy, "Delay-centric handover in sctp over wlan," *Transactions on AUTOMATIC CONTROL and COMPUTER SCIENCE*, 2004.

[4] I. Gavriloff, "Análise de aspectos envolvidos no mecanismo de seleção de caminho baseado em atraso para sistemas multiabrigados utilizando sctp," Master's thesis, UFPR, Curitiba, Brazil, March 2009.

[5] J. Fitzpatrick, S. Murphy, and J. Murphy, "An approach to transport layer handover of voip over wlan," *IEEE*, 2005.

[6] J. Fitzpatrick, S. Murphy, M. Atiquzzaman, and J. Murphy, "Evaluation of voip in a mobile environment using an sctp based handoff mechanism," *IST Mobile Summit*, 2007.

[7] J. Fitzpatrick, S. Murphy, M. Atiauzzaman, and J. Murphy, "Using cross-layer metrics to improve the performance of end-to-end handover mechanisms," *Computer Communication Preprint Online*, vol. 10.1016, p. 13 Pages, 2009.

[8] Y. Qiao, E. Fallon, J. Murphy, L. Murphy, X. Zhu, G. Hayes, A. Matthews, and A. Hanley, "Performance analysis of multi-homed transport protocols with network failure tolerance," *IET Communications*, vol. 2, pp. 336–345, 02/2008 2008.

[9] R. Stewart, M. Ramalho, Q. Xie, M. Tuexen, and P. Conrad, "Stream control transmission protocol (sctp) partial reliability extension," RFC 3758, Internet Engineering Task Force, May 2004. [Online]. Available: http://www.ietf.org/rfc/rfc3758.txt

[10] J. pierre Ebert, A. Willig, and A. Wolisz, "A gilbert-elliot bit error model and the efficient use in packet level simulation," Technical University Berlin, Tech. Rep., March 1999.

[11] P. Bhagwat, P. Bhattacharya, A. Krishma, and S. K. Tripathi, "Using channel state dependent packet scheduling to improve tcp throughput over wireless lans," *Wirel. Netw.*, vol. 3, no. 1, pp. 91–102, 1997.

[12] ITU-T, "The e-model, a computational model for use in transmission planning," *G.107*, 2000.

[13] ——, "Estimates of ie and bpl parameters for a range of codec types," ITU-T, January 2003.

[14] ITU.T, "Pulse code modulation (pcm) of voice frequencies," *G.711*, November 1988.

[15] I.-T. R. G. appendix I, "A high quality low-complexity algorithm for packet loss concealment with g.711," September 1999.

[16] ITU-T, "Provisional planning values for the equipment impairment factor ie and packet-loss robustness factor bpl," *G.113 - Appendix 1*, 2002.