Performance Modeling of SCTP Multihoming

Shaojian Fu and **Mohammed Atiquzzaman**
Telecommunications and Networks Research Lab
School of Computer Science
University of Oklahoma,
Norman, OK 73019-6151, USA.
Email: {sfu,atiq}@ou.edu

Abstract— **Stream Control Transmission Protocol (SCTP) is a new transport layer protocol which is being standardized by IETF. Multi-homing, one of the most attractive features of SCTP, makes it competitive in high-availability and mobile environments. In this paper, we propose an analytical model for evaluating the performance of SCTP multi-homing. The proposed model has been validated against and found to agree well with simulation results.**

I. INTRODUCTION

Stream Control Transmission Protocol (SCTP) is a new transport protocol over IP networks [1]. Multihoming is one of its new built-in features that is not available in TCP. Multi-homed endpoints can utilize the redundancy in network, and allow high-availability applications to perform uninterrupted switch-overs during link failures.

SCTP's multihoming has received much attention from the research community [2], resulting in a number of papers on using multihoming to improve data transmission. For example, Jungmaier et al. [3] investigated the effect of SCTP multihoming on the recovery of SS7 network linkset failures; Iyengar et al. [4] studied possible packet reordering from simultaneous data transfers over multiple destination addresses; Fu et al. [5] proposed a Seamless IP diversity based Generalized Mobility Architecture (SIGMA), which also utilizes SCTP multihoming feature. However, the authors are not aware of any analytical framework to predict SCTP's throughput over a multihomed association. The *objective* of this paper is to propose an analytical model to fill in this research gap.

SCTP is based on the congestion control principles of TCP. Recently, several papers have reported analytical models to predict the throughput of TCP [6]–[9]. Since TCP does not support multihoming, the models did not consider the effect of multihoming on transport layer throughput, and thus cannot be readily applied for SCTP. *This paper differs from previous research in the fact that the model proposed in the paper explicitly takes multihoming into account* in the analysis of SCTP's throughput.

In this paper, we model the throughput of SCTP multiohoming using the fixed-point method [9]. Our model is split into two parts: source model and network model. The advantage of this methodology is its ability to isolate the analysis of

SCTP's congestion control algorithms from network dynamics, rendering the model clear and accurate.

The rest of the paper is organized as follows. The modeling approach is described in Sec. II, followed by details of our proposed SCTP source and network models in Secs. III and IV, respectively. We then validate the accuracy of the proposed model against simulations in Sec. V. The accuracy of our proposed model is presented in Sec. VI, followed by concluding remarks in Sec. VII.

II. OVERVIEW OF ANALYTICAL MODEL

In this section, we describe our modeling approach, including modeling assumptions (Sec. II-A), architecture of the model (Sec. II-B), and notations used in the model (Sec. II-C).

We consider an SCTP multihoming association over the typical network topology of Fig. 1. The source node (SRC) is attached to N FTP flows which send data to the destination node (DST); DST connects to the network through two access routers AR1 and AR2. Tuples $(B_i, K_i), 1 \leq i \leq 5$, are the bandwidths and queue sizes of of the links in the topology.
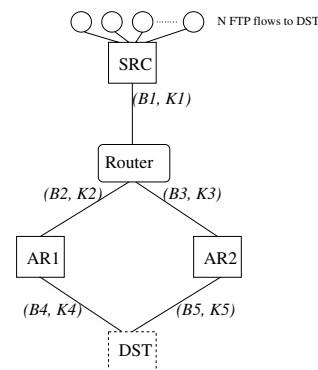

Fig. 1. Network topology.

A. Modelling Assumptions

We make the following assumptions (also used in [9]–[11]) for developing our analytical model.

- Aggregation of a large number of SCTP traffic sources results in the overall traffic arrival to the network being Poisson;
- Loss between subsequent segments in the network are independent;
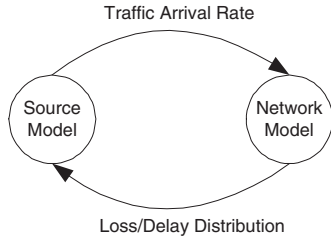- Round Trip Time (RTT) has an exponential distribution.

Fig. 2. Overall modeling architecture.

- SCTP associations carry long-lived FTP traffic.

B. Overall architecture

In addition to packets being lost due to buffer overflows in the network shown in Fig. 1, they can also be delayed due to queueing in the buffers. The overall modeling architecture, taking into account these loss and delay effects, is shown in Fig. 2. The traffic rate from the source model is fed into the network model as the arrival traffic, and the packet loss/delay distribution from the network model is fed into the source model to compute the new traffic arrival pattern. This process is iterated until the traffic arrival rate into the network model reaches a stable value, indicating that the overall model has achieved an equilibrium point.

C. Notations

The notations used in this paper are given below.

p_q, d_q Segment loss probability and mean delay obtained from the queueing network model.

d_{pt} Propagation and transmission delay between source and destination.

θ Round Trip Time (RTT) between source and destination; $\theta = d_{pt} + d_q$.

$cwnd$ Congestion window size (segments).

W_t Slow start threshold (segments).

$wmax$ Maximum value of $cwnd$.

N Number of SCTP sources.

T Value of Retransmission Time Out (RTO) (seconds).

$ccwnd, pcwnd$ Value of $cwnd$ size after and before a state transition.

π Steady state distribution of tuple $(cwnd, W_t, l)$.

$P_w(j)$ Probability of j segments lost in a window of size w.

P_w^{TO} Probability that a Time Out (TO) occurs when $cwnd = w$.

P_w^{FR} Probability that a Fast Retransmit (FR) occurs when $cwnd = w$.

$P\left(loss^{(k)}\right)$ Probability that k segments were lost during the last state transition.

$P\left(pcwnd^{(i)}, ccwnd^{(j)}\right)$ Probability that $pcwnd = i$ and $ccwnd = j$.

G Expected number of total segments generated by source model per RTT.

$E[L]$ Expected number of total losses per RTT.

λ_{source} Traffic rate generated by source model (segments/sec).

III. SCTP SOURCE MODEL

In this section, we develop the average traffic rate generated by an SCTP source, given a certain packet loss probability p_q and packet delay d_q in the queueing network. We first consider a single-homed SCTP association case (Sec. III-B) then a multihomed association case (Sec. III-C).

A. Difference between congestion control of TCP and SCTP

SCTP's congestion control is based on and very similar to the well proven rate-adaptive window-based congestion control of TCP. The common features include the adoption of slow start, congestion avoidance, timeout and fast retransmit algorithms. However, there are several major differences between the congestion control mechanisms of TCP and SCTP. Since our modeling approach is based on that used for TCP [9], we list below the differences between the congestion control of TCP and SCTP.

- SCTP doesn't have an explicit fast-recovery phase. SCTP achieves fast recovery implicitly through the use of SACK [1].
- SCTP begins slow start algorithm from $cwnd = 2$ instead of one in TCP.
- Mandatory use of SACK in SCTP allows more robust reaction in the case of multiple losses from a single window of data. This avoids a time-consuming slow start stage after multiple segment losses, thus saving bandwidth and increasing throughput.
- TCP begins fast retransmit after the receipt of three Duplicate Acknowledgements (DupACKs); SCTP begins after four DupACKs. However, SCTP is able to clock out new data on receipt of the first three DupACKs, and can also retransmit a lost segment by ignoring whether the flight size is less than $cwnd$.

B. Single-homed SCTP association

We show the state transition diagram of an SCTP association with one destination in Fig. 3; it is based on TCP's state transition diagram [9] and incorporated two differences between TCP and SCTP: (a) SCTP's slow start begins from two segments instead of one, (b) SCTP begins fast retransmit after four DupACKs, and therefore, the triggering of fast retransmit in SCTP requires a current congestion window of at least five, whereas it is four for TCP.

In Fig. 3, every state includes three elements $(cwnd, W_t, l)$, where l is the loss indication: 0 means no loss occurred during previous transition and 1 means one or multiple losses occurred. For ease of reading, only $cwnd$ is shown in the circles, and thick circles correspond to states with $l = 1$. Here, $wmax = 16$ is assumed to model the largest receiver window $(rwnd)$ of 16, and initial $W_t = wmax$. The rightmost column with thick circles denotes states undergoing fast retransmission. Since this column is identical for $W_t = 2, 4, 8, 16$, to keep the figure readable, only the case for $W_t = 2$ is shown.

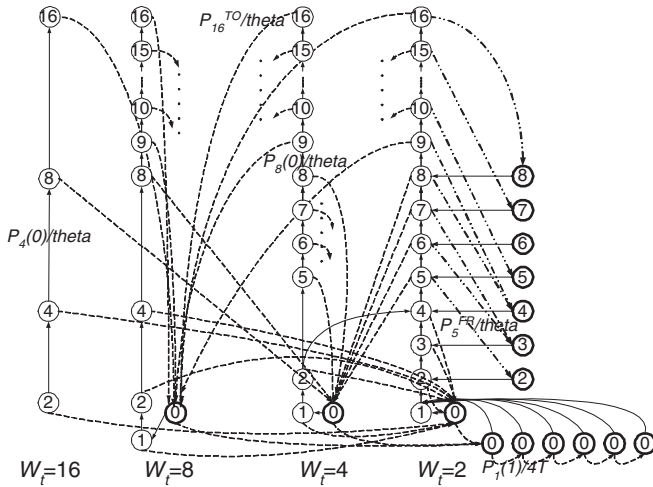The state transitions in Fig. 3 can be classified into four categories:

Fig. 3. State transition of SCTP source - single-homed case.

Fig. 4. State transition of SCTP source - multihomed case.

- **Slow Start**: state transitions from $(w, W_t, 0)$ to $(2w, W_t, 0)$ with a transition rate of $P_w(0)/\theta$. This means sender's congestion window size grows from w to $2w$ in one RTT, if there is no loss. For example, in Fig. 3, the transition probability from $cwnd = 4$ to 8 at $W_t = 16$ is $P_4(0)/\theta$.
- **Congestion Avoidance**: state transitions from $(w, W_t, 0)$ to $(w + 1, W_t, 0)$ with transition rate of $P_w(0)/\theta$. This means sender's current window size grows from w to $w + 1$ in one RTT if there is no loss. For example, in Fig. 3, the transition probability from $cwnd = 8$ to 9 at $W_t = 4$ is $P_8(0)/\theta$.
- **Timeout**: state transitions from $(w, W_t, 0)$ to $(0, \lfloor w/2 \rfloor, 1)$ with transition rate of P_w^{TO}/θ. This means sender's current window size drops from w to 0, and slow start threshold drops from W_t to $\lfloor w/2 \rfloor$, and l changes from 0 to 1 within one RTT if timeout happens.

$$P_w^{TO} = \begin{cases} \sum_{i=1}^{w-4} P_w(i)\left(1 - (1-p_q)^i\right) + \sum_{i=w-3}^{w} P_w(i) & w \geq 5 \\ 1 - P_w(0) & w < 5 \end{cases} \tag{1}$$

Although $cwnd = 1$ after a timeout in SCTP, we add the state $cwnd = 0$ as an intermediate state to model the waiting time before a timeout is detected. During this time, no segment is sent, so we count $cwnd$ as 0. For example, in Fig. 3, the transition probability from $(cwnd = 16, W_t = 4, 0)$ to $(cwnd = 0, W_t = 8, 1)$ is P_{16}^{TO}/θ.
- **Exponential Backoff**: In case of repeated timeouts, the SCTP sender will perform an exponential backoff. state transitions from $(0, W_t, 1)$ to $(0, 2, 1)$ with transition rate of $P_1(1)/(2^j T)$, $j = 1, 2, \cdots, 6$ for jth successive timeout. An example in Fig. 3 is the transition rate from the second to third timeout is $P_1(1)/4T$.
- **Fast Retransmit**: state transitions from $(w, W_t, 0)$ to $(\lfloor w/2 \rfloor, \lfloor w/2 \rfloor, 1)$ with transition rate of P_w^{FR}/θ. This means that sender's $cwnd$ drops from w to $\lfloor w/2 \rfloor$, slow start threshold drops from W_t to $W_t/2$, and l changes from 0 to 1 in one RTT if timeout happens.

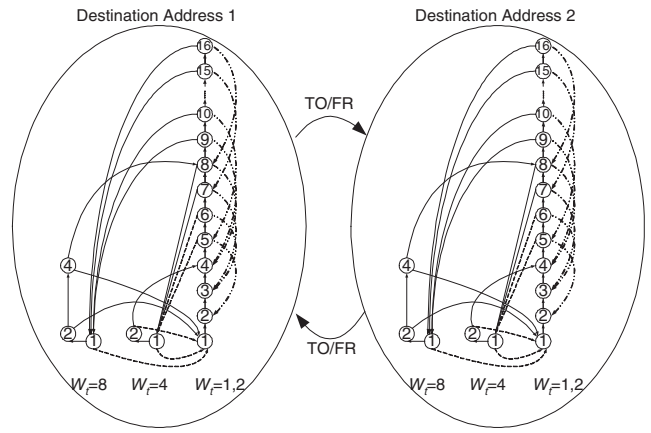$$P_w^{FR} = \begin{cases} 1 - P_w^{TO} - P_w(0) & w \geq 5 \\ 0 & w < 5 \end{cases} \tag{2}$$

For example, in Fig. 3, the transition rate from $(cwnd = 5, W_t = 2, 0)$ to $(cwnd = 2, W_t = 2, 1)$ is P_5^{FR}/θ.

If we assume packet losses to be independent from each other, $P_w(j)$ in Eqns. (1) and (2) can be determined by the Bernoulli formula: $P_w(j) = \binom{j}{w} p_q^j (1 - p_q)^{(w-j)}$.

After all transition rates in Fig. 3 are determined, the steady state distribution π of the $(cwnd, W_t, l)$ can be calculated by:

$$\pi Q = \pi \tag{3}$$

where Q is the transition probability matrix.

C. Multihomed SCTP association

We denote the expected number of segments generated by source model per RTT as:

$$G = \sum_{w=1}^{wmax} wP\left(cwnd^{(w)}\right) \tag{4}$$

By definition of π,

$$P\left(cwnd^{(w)}\right) = \sum_{W_t=2}^{wmax} \sum_{l=0}^{1} \pi(w, W_t, l) \tag{5}$$

To model an SCTP association with a multihomed destination, we next determine the traffic sent into the primary and alternative paths. We need to model SCTP's packet retransmission on the alternative path when there is a Time Out (TO) or a Fast Retransmit (FR). To do this, in Fig. 3, we strip the states where $l = 1$, and sum up all the losses when the system transits into these states (resulted from TO or FR) to obtain the total number of packets retransmitted on the alternative path, as shown in Fig. 4. Bayes method is used to compute the expected number of segment losses during these types of transitions as described in detail in Sec. III-D.

D. Bayes Loss Estimation

We separate the reason for the transition to a state with $cwnd = w$ into two cases: due to a fast retransmit and due to a timeout. Then we combine these two cases to get the expected segment losses during the transition given $ccwnd = w$.
1) Fast Retransmit case: since $ccwnd = w$, previous window size $pcwnd$ must be $2w$ or $2w + 1$. From Fig. 3,

the $ccwnd$ can only range from 2 to $wmax/2$ after a Fast Retransmit. Moreover, the number of losses during this transition can not be more than $2w - 4$, otherwise a timeout will occur. From Bayes formula:

$$P\left(loss^{(k)}|pcwnd^{(i)},ccwnd^{(w)}\right) = \frac{P\left(loss^{(k)}\right)P\left(pcwnd^{(i)},ccwnd^{(w)}|loss^{(k)}\right)}{P\left(pcwnd^{(i)},ccwnd^{(w)}\right)} \quad (6)$$

where $2 \leq w \leq wmax/2$, $i = 2w$ or $2w + 1$ and $1 \leq k \leq 2w - 4$.

Since we know that $P\left(loss^{(k)}\right) = P_i(k)$, and $P\left(pcwnd^{(i)},ccwnd^{(w)}\right) = P_i^{FR}$, Eqn. (6) becomes:

$$P\left(loss^{(k)}|pcwnd^{(i)},ccwnd^{(w)}\right) = \frac{P_i(k)P\left(pcwnd^{(i)},ccwnd^{(w)}|loss^{(k)}\right)}{P_i^{FR}} \quad (7)$$

Next, we want to find $P\left(pcwnd^{(i)},ccwnd^{(w)}|loss^{(k)}\right)$ in Eqn. (7). Since the transition to the current state has been due to a Fast Retransmit, given k segments lost from original transmission, $ccwnd$ will become w only when all the successive retransmissions for the k segments are successful. A timeout will happen if any of the k retransmissions are lost. So, the conditional probability that $pcwnd$ was i and $ccwnd$ becomes w, given k losses happened, can be estimated as:

$$P\left(pcwnd^{(i)},ccwnd^{(w)}|loss^{(k)}\right) = (1-p)^k \quad (8)$$

By substituting Eqn. (8) into Eqn. (7), we can get:

$$P(loss^{(k)}|pcwnd^{(i)},ccwnd^{(w)}) = \frac{P_i\left(k\right)\left(1-p\right)^k}{P_i^{FR}} \quad (9)$$

By summing up two cases for $i = 2w, 2w+1$ in Eqn. (9), we can get the marginal conditional distribution:

$$P(loss^{(k)}|ccwnd^{(w)}) = \sum_{i=2w}^{2w+1}\frac{P_i\left(k\right)\left(1-p\right)^k}{P_i^{FR}} \quad (10)$$

2) <u>Timeout</u> case: then $ccwnd=0$, and $pcwnd$ could be any value from 1 to $wmax$, and $k = 1, 2, \ldots pcwnd$. Similarly, by Bayes Formula:

$$P\left(loss^{(k)}|pcwnd^{(w)},ccwnd^{(0)}\right) = \frac{P\left(loss^{(k)}\right)P\left(pcwnd^{(w)},ccwnd^{(0)}|loss^{(k)}\right)}{P\left(pcwnd^{(w)},ccwnd^{(0)}\right)} \quad (11)$$

Since we know that $P\left(loss^{(k)}\right) = P_w(k)$ and $P\left(pcwnd^{(w)},ccwnd^{(0)}\right) = P_w^{TO}$, Eqn. (11) becomes:

$$P\left(loss^{(k)}|pcwnd^{(w)},ccwnd^{(0)}\right) = \frac{P_w(k)P\left(pcwnd^{(w)},ccwnd^{(0)}|loss^{(k)}\right)}{P_w^{TO}} \quad (12)$$

Next, we want to find $P\left(pcwnd^{(w)},ccwnd^{(0)}|loss^{(k)}\right)$ in Eqn. (12). Since the transition to the current state was caused by a timeout, given k segments were lost in the original transmission, if some of the retransmitted segments for the k segments failed or there are not enough DupACKs generated (in the case of $k = w - 3, w - 2, \ldots w$), $ccwnd$ will become zero; otherwise, a Fast Retransmit will happen. Also, because $pcwnd$ can be any value from 1 to $wmax$, we assume that $pcwnd$ ranges from 1 to $wmax$ with equal probability. So, the

conditional probability that $pcwnd$ was w, $cwnd$ is 0, given k losses happen, can be estimated as:

$$P\left(pcwnd^{(w)},ccwnd^{(0)}|loss^{(k)}\right) =$$
$$\begin{cases} \left[1 - (1-p)^k\right]/wmax & k = 1, 2, \ldots w - 4 \\ 1/wmax & k = w - 3, w - 2, \ldots w \end{cases} \quad (13)$$

Substituting Eqn. (13) into Eqn. (12), and summing up all the cases for $pcwnd = 1, 2, \ldots, wmax$, we get the marginal conditional distribution:

$$P\left(loss^{(k)}|ccwnd^{(0)}\right) =$$
$$\sum_{pcwnd=1}^{wmax} P\left(loss^{(k)}|pcwnd^{(w)},ccwnd^{(0)}\right)$$

$$= \begin{cases} \sum_{w=1}^{wmax}\frac{P_w(k)\left[1-(1-p)^k\right]}{wmaxP_w^{TO}} & \text{if } k = 1, 2, \ldots w - 4 \\ \\ \sum_{w=1}^{wmax}\frac{P_w(k)}{wmaxP_w^{TO}} & \text{if } k = w - 3, w - 2, \ldots w \end{cases} \quad (14)$$

3) <u>Combine FR and TO case</u>: given $ccwnd = w$. This is done by weighting the number of segment losses (k) by the conditional probabilities (Eqns. (10) and (14)):

$$E\left[L|ccwnd = w\right] = \sum_{k=1}^{wmax} kP\left(loss^{(k)}|ccwnd^{(w)}\right) \quad (15)$$

Finally, the overall expected segment losses occurring in the primary path, i.e. the traffic transferred into the alternative path can be obtained using:

$$E[L] = \sum_{w=1}^{wmax} E\left[L|ccwnd = w\right]P(ccwnd = w)$$

$$= \sum_{w=1}^{wmax}\sum_{k=1}^{wmax} kP\left(loss^{(k)}|ccwnd^{(w)}\right)P(ccwnd = w) \quad (16)$$

The above equation also represents the conditional expectation of segment losses occurring during transiting into all states with $l = 1$. We can thereby obtain the traffic on the primary path by subtracting the losses (which is also the traffic on the alternative path) (Eqn. (16)) from the total traffic generated by the source (Eqn. (4)).

IV. NETWORK MODEL.

Solution of the source model in Sec. III requires the value of RTT ($\theta = d_{pt}+d_q$) and loss probability (p_q). In this section, we derive the values of d_q and p_q. In the network model, we consider two cases: single queue case and multi-queue case. In single queue case, the whole network is modelled as an M/M/1/K queue. In the multi-queue case, we consider all the queues in the network separately. We denote λ as the arrival traffic rate at a link queue (segments/sec), and μ, B, K as the service rate (segments/sec), bandwidth (bps), and buffer size (segments) of a link, respectively.

A. Single queue case

In Fig. 1, when B_2 through B_5 are large enough, the only queue that affects packet loss and delay is the SRC-Router queue. We can model the queuing network as an M/M/1/K

queue with $K = K1$. We denote $\rho = \lambda/\mu$, where $\mu = B/8 *$ $PacketSize$ (segments/sec). From M/M/1/K queuing theory, the segment loss probability can be calculated as:

$$p_q = \begin{cases} \frac{1}{K+1} & \rho \geq 1 \\ \frac{(1-\rho)\rho^K}{1-\rho^{(K+1)}} & \rho < 1 \end{cases} \quad (17)$$

To find the queuing delay (d_q), let S be the mean number of segments in the queue:

$$S = \begin{cases} \frac{K}{2} & \rho = 1 \\ \frac{\rho}{(1-\rho)} - \frac{K+1}{1-\rho^{(K+1)}}\rho^{K+1} & \rho \neq 1 \end{cases} \quad (18)$$

Considering the current segment being transmitted in the queue, we can obtain the mean queuing delay as:

$$d_q = \frac{S+1}{\mu} \quad (19)$$

B. Multi-queue case

It is shown in [12] that in presence of greedy connections (such as FTP) that tend to overload the network, different queueing models provide similar estimates of the average loss probability. Therefore, a simple queue for each link on the topology can be used to approximate the ensemble behavior the whole network. Other approaches with significantly greater complexity (mainly based on group arrivals and services) were also tested by the authors of [12], but results do not change significantly in the case of long-lived flows. In Fig. 1, if (B_2, K_2) through (B_5, K_5) are finite, we assume that the queuing network can be modelled as a combination of M/M/1/K queues, as shown in Fig. 5. The input traffic into each queue
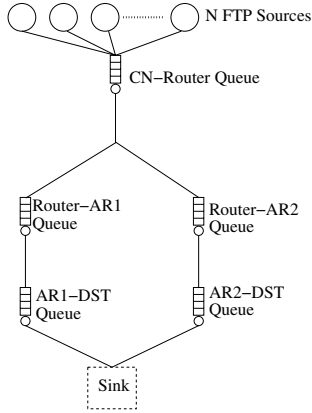

Fig. 5.   Queuing network for multi-queue case.

in Fig. 5 can be determined as: $\lambda_{SRC-Router} = \lambda_{source}$. This means the input traffic to the SRC-Router queue is the same as the traffic generated from the source model. We can also get the input traffic to the Router-AR1 queue as:

$$\lambda_{Router-AR1} = \lambda_{SRC-Router}(1-R)(1-p_{SRC-Router}) \quad (20)$$

where $p_{SRC-Router}$ denotes the loss probability at SRC-Router queue which can be determined using Eqn. (17) with $\lambda = \lambda_{SRC-Router}$, $B = B1$, and $K = K1$. R is the percentage of packets retransmitted through the alternative path (via AR2), which can be determined as:

$$R = E(L)/G \quad (21)$$

where $E(L)$ is the expected number of packet losses during one RTT (as determined by Eqn. (16)) i.e. those that will be retransmitted through the alternative path, and G (determined by Eqn. (4)) is the total traffic generated by the source model. Similarly, we can get the input traffic to the AR1-MH queue:

$$\lambda_{AR1-DST} = \lambda_{Router-AR1}(1 - p_{Router-AR1}) \quad (22)$$

Since each queue is modelled as an M/M/1/K queue, we can use Eqns. (17) and (18) to get loss probability and average queue occupancy of each individual queue. Assuming no repeated losses for traffic retransmitted into the alternative path, we can get the overall loss probability at the primary path as:

$$p_q = 1 - (1 - p_{SRC-Router})(1 - p_{Router-AR1})(1 - p_{AR1-DST}) \quad (23)$$

where $p_{Router-AR1}$ and $p_{AR1-DST}$ denote the loss probability at Router-AR1 queue and AR1-DST queue, respectively. This means that the overall loss probability is the percentage of packets that did not successfully go through all the three queues.

By Little's law, we can model the average delay in the queuing network as:

$$\begin{aligned} d_q &= \frac{S}{\lambda} \\ &= \frac{S_{SRC-Router} + S_{Router-AR1} + S_{AR1-DST}}{\lambda} \end{aligned} \quad (24)$$

where $S_{SRC-Router}$, $S_{Router-AR1}$, and $S_{AR1-DST}$ denote the average queue occupancy at SRC-Router queue, Router-AR1 queue, and AR1-DST queue, respectively; λ is the input traffic rate at SRC-Router queue. Similarly, by substituting loss probability and queue occupancy of the queues in alternative path into Eqns. (23) and (24), we can get p_q and d_q for the alternative path.

V. SIMULATION SETUP

In order to validate the accuracy of our model presented in Secs. II, III, and IV, we compare the results obtained from our proposed analytical model against simulation results obtained from the *ns-2* network simulator in Sec. VI. The simulation topology is shown in Fig. 1, where SRC is a single-homed node and DST is a multihomed one. There are fifty SCTP source agents attached to node SRC sending FTP traffic, and the fifty destination agents are attached to node DST.

Values of relevant simulation parameters are summarized in Table I. In the simulation, 50 SCTP flows share the link from $SRC \rightarrow DST$ to simulate Poisson arrival process. We vary expected RTT (θ) between 0.1 to 1 second, $wmax$ between 8 to 32 segments, and each link queue size (K) between 30 to 100 segments. For each (θ, $wmax$, K) combination, we run the simulation for a long time (500 seconds) to make sure that the results from simulation stabilize.

VI. RESULTS

In this section, we evaluate the effectiveness of our proposed model by comparing the throughput predicted by our model against the values obtained from simulation. Only the results corresponding to queue size of 50 are presented; the results for queue size of 30 and 100 are very similar but are not shown here due to the space limitations.

TABLE I
SIMULATION PARAMETERS FOR THE TOPOLOGY OF FIG. 1.

| Traffic type	FTP
Number of flows	50
Header size	52 bytes
Payload size	1448 bytes
Link queue type	drop-tail
rwnd limit	8, 16, 32 segments
Initial *ssthresh*	8, 16, 32 segments
SRC-Router link bandwidth	300Mbps
Router-AR1(AR2) link bandwidth	∞ for single queue, 100Mbps for multi-queue
AR1(AR2)-DST link bandwidth	∞ for single queue, 11Mbps for multi-queue
link propagation delay	0.005 - 0.5 s
bottleneck queue size	30 - 100

A. Single queue case

For the single queue case, we compare the primary and alternative path throughputs from our proposed model against simulation, and the results are shown in Figs. 6(a) and 6(b), respectively, for $wmax$ ranging from 8 to 32. The results show that our proposed model can predict the throughput pretty accurately. The differences between model and simulation mainly results from the assumption that the RTT between source and destination is exponential distribution. Since the propagation and transmission delay component (d_{pt}) in RTT is not random variable, this assumption will produce some errors. Also, we assume the losses within one SCTP sending window are independent from each other and follow a Bernoulli distribution. This may not be true for packet drops at the drop-tail queue.
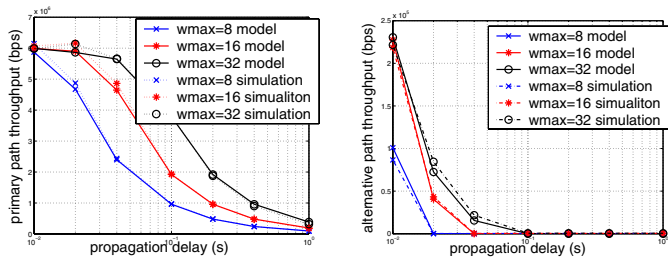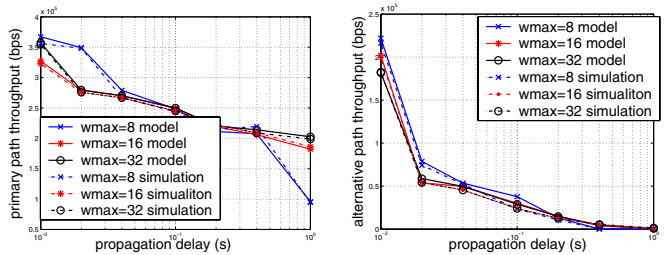
B. Multi-queue case

For the multi-queue case, we compare the primary and alternative path throughputs from our proposed model against simulation results in Figs. 6(c) and 6(d). The results show that in the case of multi-queue network, the proposed model can also predict the throughput of both primary and alternative path pretty accurately. In multi-queue case, the two assumptions that produce predicting errors in single queue still have effect. Also, when we assume that the queuing network can be modelled as a combination of M/M/1/K queues, there is also extra error introduced.

VII. CONCLUSION

Since TCP does not support multihoming, none of the previous TCP models considers the effect of multihomed node on the steady state throughput of transport protocol, therefore, they cannot be readily used to model the performance of a multihoming connection. We developed an analytical model for SCTP multihoming association, and showed that the model is accurate in estimating the steady state throughput of both primary and alternative path of multihomed SCTP associations. The model can be used by network engineers to dimension the capacity of links connecting multihomed SCTP nodes.

REFERENCES

[1] R. Stewart and Q. Xie et. al., "Stream control transmission protocol," IETF RFC 2960, October 2000.

[2] S. Fu and M. Atiquzzaman, "SCTP: State of the art in research, products, and technical challenges," *IEEE Communications Magazine*, vol. 42, no. 4, pp. 64–76, April 2004.

[3] A. Jungmaier, E.P. Rathgeb, and M. Tuxen, "On the use of SCTP in failover-scenarios," in *International Conference on Information Systems, Analysis and Synthesis*, Orlando, Florida, July 2002, pp. 363–368.

[4] J. Iyengar, P. Amer, R. Stewart, and I. Rodriguez, "Preventing SCTP congestion window overgrowth during changeover," Internet Draft, draft-iyengar-sctp-cacc-01.txt.

[5] S. Fu, L. Ma, M. Atiquzzaman, and Y. Lee, "Architecture and performance of SIGMA: A seamless mobility architecture for data networks," in *40th IEEE Internationl Conference on Communications (ICC)*, Seoul, Korea, May 2005.

[6] T.V. Lakshman and U. Madhow, "The performance of TCP/IP for networks with high bandwidth-delay products and random loss," *IEEE/ACM Transactions on Networking*, vol. 5, no. 3, pp. 336–350, June 1997.

[7] M. Mathis, J. Semke, and J. Mahdavi, "The macroscopic behavior of the TCP congestion avoidance algorithm," *Computer Communications Review*, vol. 27, no. 3, pp. 67–82, July 1997.

[8] J. Padhye, V. Firoiu, D.F. Towsley, and J.F. Kurose, "Modeling TCP Reno performance: a simple model and its empirical validation," *IEEE/ACM Transactions on Networking*, vol. 8, no. 2, pp. 133–145, April 2000.

[9] C. Casetti and M. Meo, "A new approach to model the stationary behavior of TCP connections," in *IEEE INFOCOM 2000*, Tel-Aviv, Israel, March 2000, pp. 367–375.

[10] C. Casetti and M. Meo, "Modeling the stationary behavior of TCP Reno connections," in *International Workshop on Quality of Service in Multiservice IP Networks*, Rome, Italy, January 2001, pp. 141 – 156.

[11] Takayuki Osogami Adam Wierman and Jorgen Olsn, "A unified framework for modeling TCP-Vegas, TCP-SACK, and TCP-Reno," in *11th IEEE/ACM International Symposium on Modeling, Analysis and Simulation of Computer Telecommunications Systems*, Orlando, Florida, Octobor 2003, pp. 269–278.

[12] M.Garetto, R. Cigno, M. Meo, and M.A. Marsan, "Closed queueing network models of interacting long-lived TCP flows," *IEEE/ACM Transactions on Networking*, vol. 12, no. 2, pp. 300–311, April 2004.

(a) Primary path throughput for single queue case.

(b) Alternative path throughput for single queue case.

(c) Primary path throughput for multi-queue case.

(d) Alternative path throughput for multi-queue case.

Fig. 6. Simulation validation of the analytical model