Impact of Traffic Load on SCTP Failovers in SIGTRAN

Karl-Johan Grinnemo[1] and Anna Brunstrom[2]

[1] TietoEnator AB, Lagergrens gata 2, S-651 15 Karlstad, SWEDEN,
karl-Johan.grinnemo@tietoenator.com

[2] Karlstad University, Dept. of Computer Science, S-651 88 Karlstad, SWEDEN,
anna.brunstrom@kau.se

Abstract. With Voice over IP (VoIP) emerging as a viable alternative to the traditional circuit-switched telephony, it is vital that the two are able to intercommunicate. To this end, the IETF Signaling Transport (SIGTRAN) group has defined an architecture for seamless transportation of SS7 signaling traffic between a VoIP network and a traditional telecom network. However, at present, it is unclear if the SIGTRAN architecture will, in reality, meet the SS7 requirements, especially the stringent availability requirements. The SCTP transport protocol is one of the core components of the SIGTRAN architecture, and its failover mechanism is one of the most important availability mechanisms of SIGTRAN. This paper studies the impact of traffic load on the SCTP failover performance in an M3UA-based SIGTRAN network. The paper shows that cross traffic, especially bursty cross traffic such as SS7 signaling traffic, could indeed significantly deteriorate the SCTP failover performance. Furthermore the paper stresses the importance of configuring routers in a SIGTRAN network with relatively small queues. For example, in tests with bursty cross traffic, and with router queues twice the bandwidth-delay product, failover times were measured which were more than 50% longer than what was measured with no cross traffic at all. Furthermore, the paper also identifies some properties of the SCTP failover mechanism that could, in some cases, significantly degrade its performance.

1 Introduction

Since Voice over IP (VoIP) roared into prominence during the latter part of the 1990s, the idea of a converged network based on IP technology for voice, video, and data has gained strong momentum. However, in spite of all prospective advantages with IP it would be naive to think that the transition from the traditional circuit-switched network to IP would happen overnight.

In light of this, the IETF Signaling Transport (SIGTRAN) working group has defined an architecture, the SIGTRAN architecture [1], for seamless Signaling System #7 (SS7) signaling between VoIP and the traditional telecom network. The SIGTRAN architecture essentially comprises two components: a new IP transport protocol, the Stream Control Transmission Protocol (SCTP) [2], specifically designed for signaling traffic; and an adaptation sublayer. The adaptation sublayer shields SS7 from SCTP and IP, and depending on how much of the SS7 stack is run atop SCTP, different adaptation protocols are used. Examples of adaptation protocols include: M2PA [3] for adaptation of the SS7 MTP-L3 [4] protocol to IP, and M3UA [5] for adaptation of SCCP [6] and user part protocols such as ISUP [7].

It is widely recognized that to gain user acceptance, the SIGTRAN architecture has to perform comparable to the traditional circuit-switched telecom network [8]. In particular, it has to provide the same level of availability as a traditional SS7 network. Considering that ITU-T prescribes an availability level of 99.9988% [9], i.e., no more than 10 minutes downtime per year, and that many telecom networks have an even higher availability level [10], this is indeed a great challenge.

To meet the stringent requirements of SS7, several availability mechanisms have been included in the SIGTRAN architecture of which the SCTP failover mechanism is one of the more important ones – if not the most important one. It corresponds with the MTP-L3 changeover procedure, and enables rapid re-routing of traffic from a failed signaling route within a SIGTRAN network. In particular, the SCTP failover mechanism constitutes part of SCTPs multi-homing support.

Although, the SCTP failover mechanism plays a key role in the availability support of the SIGTRAN architecture, very few results are available on its actual performance in this context. Jungmaier et al. [11] have studied the SCTP failover performance in an M2PA-based network, and showed that it only meets ITU-T requirements provided it is configured very aggressively, and provided the network path propagation delays are very short. A similar result was also obtained by Grinnemo et al. [12] when they performed measurements on SCTP failover performance in an M3UA-based network.

Both the study in [11] and in [12] took place in unloaded networks, i.e., under quite unrealistic conditions. This paper advances the work in [12], and partly the work in [11], by studying the impact of traffic load on the SCTP failover performance in an M3UA-based SIGTRAN network. The main contribution of the paper is that it demonstrates that cross traffic, especially bursty cross traffic such as SS7 signaling traffic, could indeed significantly deteriorate the SCTP failover performance. Furthermore, the paper stresses the importance to keep the router queues in a SIGTRAN network relatively small. In fact, the paper shows that bursty traffic in combination with ill-dimensioned router queues may well cause the SCTP failover mechanism to not comply with the ITU-T requirement on the MTP-L3 changeover procedure [9]. Furthermore, the paper identifies some issues regarding the design of the SCTP failover mechanism which in some cases negatively affect the failover performance.

The remainder of the paper is organized as follows. Section 2 gives a brief description of the SCTP failover mechanism. Then, in Section 3 follows a description of the design and execution of the experiment that underlies our study. Next, in Section 4, we elaborate on the results of the experiment. Finally, in Section 5, the paper ends with some concluding remarks and words on future work.

2 Failovers in SCTP

While IP networks have many virtues, high availability and reliability have traditionally not been seen as two of them. Unlike circuit-switched paths, which exhibit changeover and failover times on the order of milliseconds, measurements show that it may take well over ten seconds before the routers in the Internet reach a consistent view after a path failure [13] – in other words, too long for delay-sensitive SS7 signaling traffic.

In the SIGTRAN architecture, the unsuitability of IP for high-availability routing of SS7 signaling messages is addressed through various redundancy mechanisms at the transport and adaptation layers. As previously mentioned, one of the most important network redundancy mechanisms in SIGTRAN is the SCTP failover mechanism.

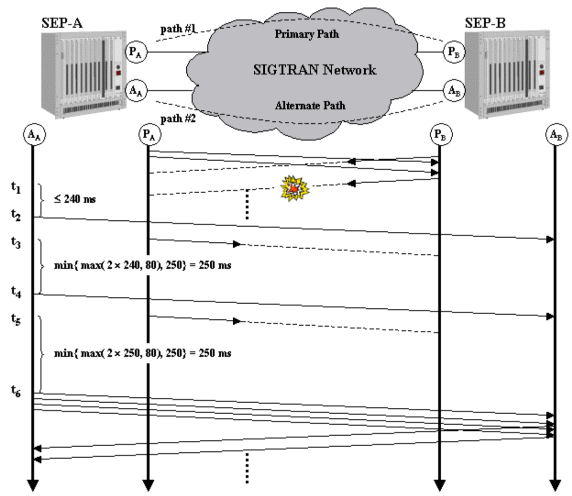


Fig. 1. Failover scenario between two dual-homed signaling endpoints.

An example of how the SCTP failover mechanism works is illustrated in Figure 1. In this example, we have an SCTP connection, a so-called association, between two signaling endpoints: SEP-A and SEP-B. The association comprises two routing paths: path #1 and path #2. Since SCTP does not support load-sharing, one path in an association is always designated the primary path and is the path on which signaling traffic is normally sent. The remaining paths, if any, become backup or alternate paths. In our example, path #1 is the primary path and path #2 an alternate path.

SCTP continuously monitors reachability on the primary and alternate paths – on an active primary path SCTP probes for reachability using the transferred data packets themselves, and on idle alternate paths specific heartbeat packets are used. Furthermore, for each path (actually network destination), SCTP keeps an error counter that counts the number of consecutively missed acknowledgements to data or heartbeat packets. A path is considered unreachable when the error counter of the path exceeds the value of the SCTP parameter `Path.Max.Retrans`. In the remaining discussion, it is assumed that the SCTP stacks at SEP-A and SEP-B are configured with `Path.Max.Retrans` set to 2.

As follows from the time line in Figure 1, a failure occurs on the primary path at time t_1. At that time, the SCTP retransmission timeout (RTO) variable is assumed to be

240 ms, and it is assumed that there are outstanding traffic. Thus, at $t_2 \leq t_1 + 240\,ms$, the SCTP retransmission timer, T3-rtx, expires and a timeout occurs; an SCTP packet worth of outstanding data is retransmitted on the alternate path, and the error counter of the primary path is incremented by one. Furthermore, the RTO variable is backed off, or more precisely

$$RTO \leftarrow min\left\{max\left(2 \times RTO_{cur}, RTO_{min}\right), RTO_{max}\right\}, \tag{1}$$

where RTO_{cur} denotes the current value of the RTO variable, and RTO_{min} and RTO_{max} are SCTP parameters that limit the range of the RTO variable. Here, it is assumed that RTO_{min} is set to 80 ms and RTO_{max} to 250 ms.

At time t_3, new data is sent out on the primary path, and the T3-rtx timer is restarted with the value of the updated RTO variable. The flow of events that occurred at times t_2 and t_3 are repeated at times t_4 and t_5. When time t_6 is reached, the error counter of the primary path becomes 3, i.e., greater than `Path.Max.Retrans`, and SCTP considers the path failed and starts sending new data onto the alternate path. In other words, the failover concludes.

3 Methodology

To be able to study the impact of traffic load on the SCTP failover performance, we considered the network scenario depicted in Figure 2.

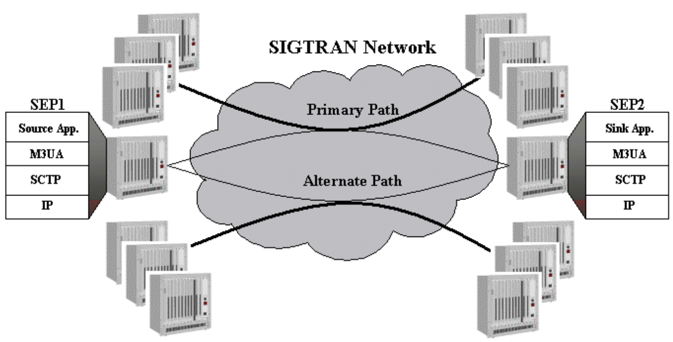


Fig. 2. Studied network scenario.

In this scenario, two M3UA users at signaling endpoints SEP1 and SEP2 were engaged in a signaling session over a SIGTRAN network with varying degrees of traffic load. The session took place over a multi-path association with one primary and one alternate path. Initially, all signaling traffic in the M3UA session was routed on the primary path. However, 30 s into the signaling session a failure occurred on the primary path. As a result, the signaling traffic was re-routed from the primary to the alternate path. The network scenario ended when 90 s had elapsed from the time of the path failure.

The network scenario in Figure 2 was modeled using the experiment setup illustrated in Figure 3. The M3UA session between SEP1 and SEP2 was modeled as a constant bit rate flow of 200 Kbps. Although it could be argued that a constant bit rate flow is not a particularly realistic model of actual SS7 traffic [14], a more realistic model would make it much more difficult to measure the failover times. Particularly, introducing randomness in the traffic generation at SEP1 would render it difficult to establish the start times of the failovers.

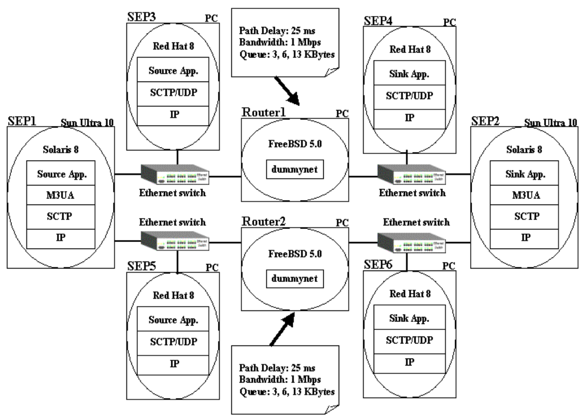


Fig. 3. Experiment setup.

The cross traffic comprised single SCTP flows between SEP3 and SEP4, and SEP5 and SEP6. Since the SS7 traffic in future dedicated SIGTRAN networks will presumably be bursty [14, 15], the cross traffic was generated as bursty flows. Tests were run for a range of cross traffic flows representing a spectrum of traffic loads with different degrees of burstiness. Specifically, tests were run with cross traffic flows having burst sizes and inter-burst gaps as listed in Table 1. It should be noted that CT-NONE denotes no cross traffic at all, and that the CT-HIGH cross traffic case actually meant that the SEP1 source application did not impose any limits on the SCTP transmission rate.

To be able to study the impact of queueing delay on the SCTP failover performance, tests were run with three different router queue sizes: 3 Kbytes (approximately half the bandwidth-delay product), 6 KBytes (approximately the same as the bandwidth-delay product), and 13 KBytes (approximately twice the bandwidth-delay product). These queue sizes were selected with the intent to model the router configurations found in both controlled, delay-sensitive, networks, and uncontrolled networks.

The SCTP stacks at SEP1 and SEP2 were configured to meet the ITU-T requirements on the MTP-L3 changeover procedure [9], i.e., according to the findings in [11, 12]. More precisely, they were configured as shown in Table 2, with the remaining pa-

Table 1. Cross Traffic Characteristics.

Name	Burst Size (KBytes)	Inter-Burst Gap (ms)
CT-NONE	0	0
CT-LOW	4	200
CT-MEDIUM	8	100
CT-HIGH	16	50

Table 2. SCTP configuration.

Parameter	Setting
RTO_{init}	250 ms
RTO_{min}	80 ms
RTO_{max}	250 ms
Path.Max.Retrans	2
SACK timer	40 ms

rameters set as recommended in RFC 2960 [2]. The SCTP stacks at the remaining SEPs were configured in accordance with RFC 2960.

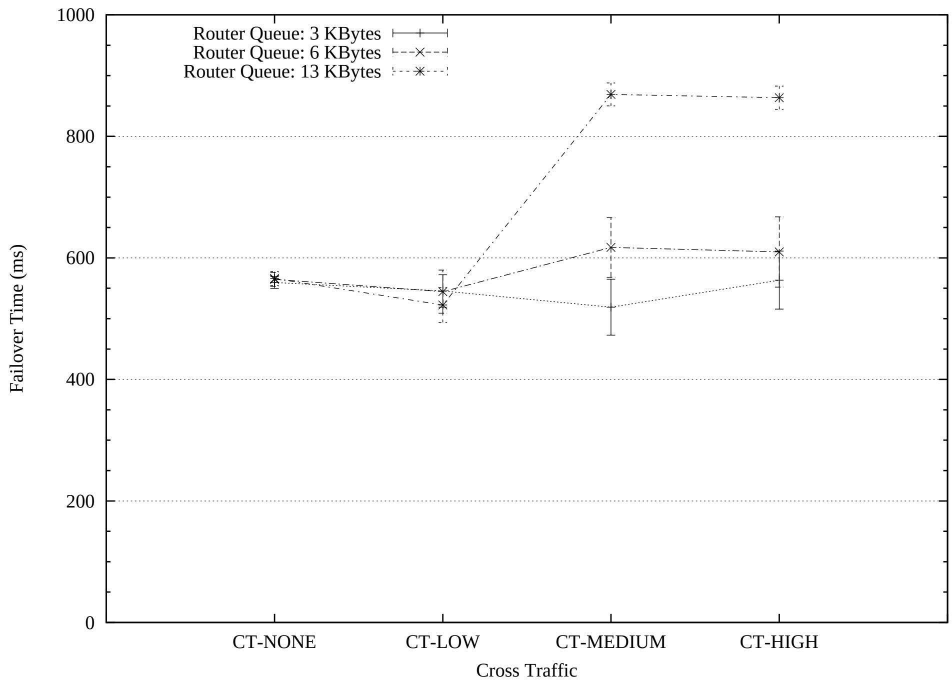
Tests were run for all combinations of cross traffic and router queue sizes, giving a total of 12 tests. Furthermore, to obtain statistical validity each test was repeated 40 times.

4 Results

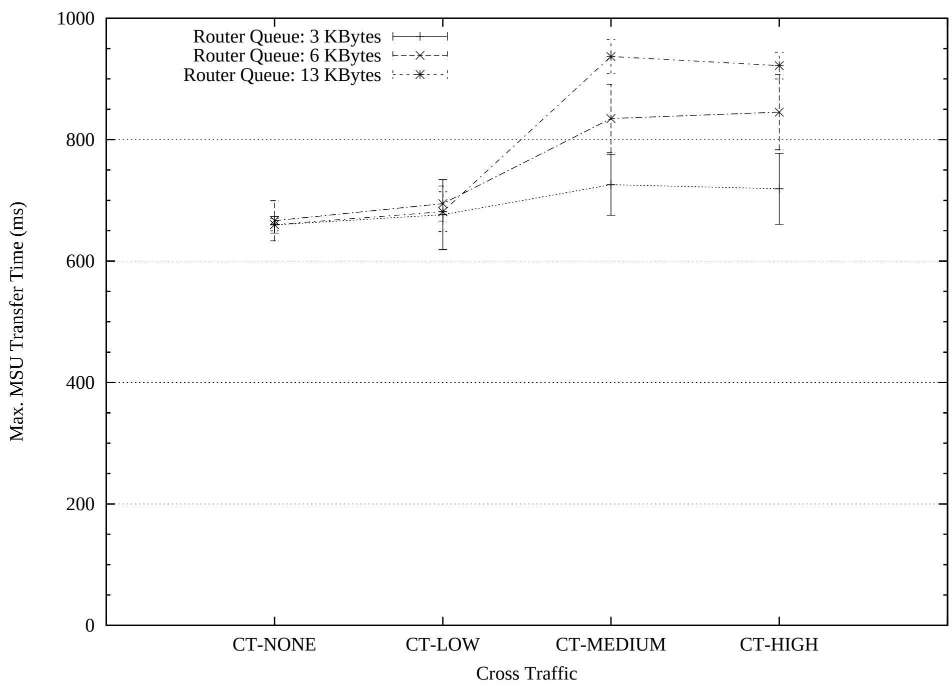
The SCTP failover performance was evaluated in terms of two metrics: the failover time experienced by the SEP1 source application, and the maximum Message Signal Unit (MSU) transfer time measured during failover in the M3UA session between SEP1 and SEP2. As estimates of the failover times and the max. MSU transfer times in the tests, the sample means were used.

Figure 4 summarizes the result of our experiment. In Figure 4 (a), it is shown how the SCTP failover time was affected by increasing traffic load at different router queue sizes, while Figure 4 (b) shows the same relationship for the max. MSU transfer time. The error bars depict the 99% confidence intervals, and the lines connecting the mean failover times and max. MSU transfer times are only supplied as a visualization aid. Specifically, these lines are only included to help visualize the trends.

As follows from Figure 4, the deteriorating effect of the cross traffic on the failover performance increased with increased traffic load and router queue size. When the Router1 queue was only 3 KBytes, the cross traffic did not inflict significantly on the failover and max. MSU transfer times. However, as the queue size was increased, the effect of the cross traffic became more and more apparent. Thus, when the Router1 queue was 13 KBytes, the CT-HIGH cross traffic increased the failover time with more



(a) Failover Time vs. Cross Traffic.



(b) Max. MSU Transfer Time vs. Cross Traffic.

Fig. 4. Impact of traffic load on SCTP failover performance.

than 50% and the max. MSU transfer time with almost 40% as compared with no cross traffic at all.

The reason to the increased failover and max. MSU transfer times was the queueing delays that arose at Router1 when the router queue was fairly large, and when the cross traffic was bursty (i.e., when the short-term bandwidth requirement of the cross traffic sometimes exceeded the bandwidth capacity of the primary path). As a matter of fact, in previous tests with the same test flow, but with constant bit rate cross traffic, it was found that the traffic load had no significant impact on the failover performance provided it was less than the path capacity.

Another observation worth making concerns the SCTP failover times with regards to the requirement of ITU-T on the MTP-L3 changeover procedure [9]. To comply with this requirement, the SCTP failover times should be no more than 800 ms. However, as follows from Figure 4, this requirement was only fulfilled in those cases the Router1 queue was relatively small (3 KBytes or 6 Kbytes). In the tests with a router queue of 13 KBytes or twice the bandwidth-delay product (to our knowledge a quite common configuration [16]), the failover times averaged well above 850 ms at medium (CT-MEDIUM) and high (CT-HIGH) traffic loads.

Interestingly, in all tests, the measured failover times were significantly larger than what could be expected given the RTOs. However, the discrepancy was larger with larger traffic loads and router queues. Consider, for example, the test with a 13 KByte Router1 queue and the CT-HIGH cross traffic. When this test was re-ran with tracing on the RTO, the RTO at the time of the path failure, RTO_t, was measured to 240 ms. Only considering the timeout periods, this gives us a theoretical failover time of $240\,ms + 250\,ms + 250\,ms = 740\,ms$ (see Section 2). However, the measured failover time was 920 ms, or 180 ms larger than our estimate.

The reason to this discrepancy turned out to be substantial delays between the expiration of the T3-rtx timer and its restart during the failover (see Figure 5). When a timeout occurred, the SCTP congestion window at SEP1 was reduced to 1 Maximum Transmission Unit (MTU). As a result, no packets were sent out on the primary path, and the T3-rtx timer was not restarted, until the amount of outstanding data went below 1 MTU. This meant, as shown in Figure 5, an extra delay (apart from the timeout delay) of about 80 ms at each timeout event.

Although, an extra delay of 80 ms at each timeout during a failover has to be considered as a quite large delay in this context (SS7 signaling), even larger delays could be expected in real-world SIGTRAN networks. Specifically, it could take several transmission rounds before the T3-rtx timer of the primary path is restarted again after a timeout in cases with large amounts of outstanding data at the time of a path failure.

Finally, as an aside, we would like to mention the significant penalty in terms of failover performance that could be the result of setting RTO_{init}, the initial value of RTO, too low. Specifically, a too low value on RTO_{init} with respect to the round-trip time of the alternate path[3] could result in one extra retransmission, and thus one

[3] Note that the first transmission round on the alternate path within a timeout period only comprises a single SCTP packet. Consequently, the SACK timer delay adds to the initial round-trip time in a timeout period on the alternate path, something that is easily overlooked when RTO_{init} is configured.

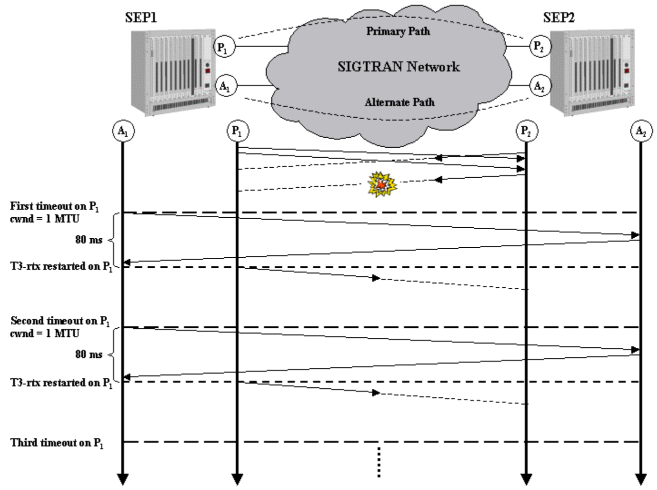


Fig. 5. Management of the T3-rtx timer during failover.

extra timeout period, before SCTP considers the primary path failed. To gain some appreciation of the extent to which this could in fact impede on the failover performance in a SIGTRAN network, we re-ran the test with the Router1 queue set to 13 KBytes and with no cross traffic (CT-NONE), but this time with RTO_{init} at SEP1 and SEP2 configured to 80 ms instead of 250 ms. The result of this test was that we observed an increase in failover time with about 180 ms, or 32%, compared with the original test (cf. Figure 4 (a)).

5 Conclusions

This paper studies the impact of traffic load on the SCTP failover performance in an M3UA-based SIGTRAN network. Two performance metrics are considered: the SCTP failover time, and the maximum transfer time experienced by an M3UA user during failover. The paper shows that cross traffic, especially bursty cross traffic such as SS7 signaling traffic, could indeed significantly deteriorate the SCTP failover performance. Furthermore, the paper demonstrates how important it is to configure the routers in a SIGTRAN network with relatively small queues. For example, in tests with bursty cross traffic and with router queues twice the bandwidth-delay product (to our knowledge a quite common configuration), failover times were measured which on the average were more than 50% longer than what was measured with no cross traffic at all. In fact, in these situations, our study suggests the SCTP failover performance may not even meet the requirement of ITU-T on MTP-L3 changeovers.

Two important observations are also made in the paper which concern the SCTP failover behavior. First, it is shown that the delays which occur in between the expiration of the SCTP retransmission timer (T3-rtx) and its restart during a failover could contribute significantly to the failover and max. MSU transfer times. Second, the paper comments on the extent to which a too low initial retransmission timeout (RTO) value, i.e., a too low value on the SCTP parameter RTO_{init}, could deteriorate the failover performance.

While cross traffic, T3-rtx restart delays, and low values on RTO_{init} could have a significant negative effect on the SCTP failover performance, it still remains that a major factor is the length of the timeout periods. Thus, in our future work, we intend to study ways of shortening these periods without threatening network stability. Specifically, we intend to study the effect of introducing a more relaxed RTO backoff mechanism.

References

1. Ong, L., Rytina, I., Garcia, M., Schwarzbauer, H., Coene, L., Lin, H., Juhasz, I., Holdrege, M., Sharp, C.: Framework architecture for signaling transport. RFC 2719, IETF (1999)
2. Stewart, R., Xie, Q., Morneault, K., Sharp, C., Schwarzbauer, H., Taylor, T., Rytina, I., Kalla, M., Zhang, L., Paxson, V.: Stream control transmission protocol. RFC 2960, IETF (2000)
3. George, T., Bidulock, B., Dantu, R., Schwarzbauer, H.J., Morneault, K.: Signaling system 7 (SS7) message transfer part 2 (MTP2) - user peer-to-peer adaptation layer (M2PA). Internet draft, IETF (2004) draft-ietf-sigtran-m2pa-12.txt.
4. ITU-T: Specifications of signalling system no. 7 - message transfer part: Signalling network functions and messages. ITU-T Recommendation 704, ITU (1996)
5. Sidebottom, G., Morneault, K., Pastor-Balbas, J.: Signaling system 7 (SS7) message transfer part 3 (MTP3) - user adaptation layer (M3UA). RFC 3332, IETF (2002)
6. ITU-T: Specifications of signalling system no. 7 - signalling connection control part: Signalling connection control part procedures. ITU-T Recommendation 714, ITU (1996)
7. ITU-T: Specifications of signalling system no. 7 - ISDN user part: ISDN user part signalling procedures. ITU-T Recommendation 764, ITU (1997)
8. Fu, S., Atiquzzaman, M.: SCTP: State of the art in research, products, and technical challenges. IEEE Communications Magazine **42** (2004) 64–76
9. ITU-T: Specifications of signalling system no. 7 - message transfer part signalling performance. ITU-T Recommendation 706, ITU (1993)
10. Kuhn, R.: Sources of failure in the public switched telephone network. IEEE Computer **30** (1997) 31–36
11. Jungmaier, A., Rathgeb, E.P., Tuexen, M.: On the use of SCTP in failover scenarios. In: 6th World Multiconference on Systemics, Cybernetics and Informatics (SCI), Orlando, Florida, USA (2002)
12. Grinnemo, K.J., Brunstrom, A.: Performance of SCTP-controlled failovers in M3UA-based SIGTRAN networks. In: Advanced Simulation Technologies Conference (ASTC), Applied Telecommunication Symposium (ATS), Arlington, Virginia, USA (2004) 86–91
13. Labovitz, C., Ahuja, A., Bose, A., Jahanian, F.: Delayed Internet routing convergence. IEEE/ACM Transactions on Networking **9** (2001) 293–306
14. Scholtz, F.J.: Statistical analysis of common channel signaling system no. 7 traffic. In: 15th Internet Traffic Engineering and Traffic Management (ITC) Specialist Seminar, Wurzburg, Germany (2002)
15. Andersen, A.T.: Modelling of packet traffic with matrix analytic methods. PhD thesis, Informatics and Mathematical Modelling, Technical University of Denmark, DTU (1995)
16. Allman, M.: TCP byte counting refinements. ACM Computer Communications Review **3** (1999)