Performance of SCTP-controlled Failovers in M3UA-based SIGTRAN Networks[†]

Karl-Johan Grinnemo TietoEnator AB Lagergrens gata 2 SE-651 15 Karlstad, SWEDEN Email: Karl-Johan.Grinnemo@tietoenator.com Anna Brunstrom Dept. of Computer Science Karlstad University SE-651 88 Karlstad, SWEDEN Email: Anna.Brunstrom@kau.se

Keywords: failover, SCTP, multi-homing, SIGTRAN, M3UA

Abstract

There are some large economic, operational, and, to some extent, technical incentives to replace the traditional telecom network with IP. However, such a large transition will not happen overnight – maybe never. Meanwhile, IP-based and traditional TDM-based telephony will have to co-exist. To address this situation, the IETF SIGTRAN working group has developed an architecture for transportation of Signaling System No. 7 (SS7) traffic over IP. Still, it remains to be shown that the introduction of the SIGTRAN architecture will not significantly deteriorate the performance of SS7. To this end, this paper evaluates the failover performance in SIGTRAN networks. Specifically, the paper evaluates the performance of SCTP-controlled failovers in M3UA-based SIGTRAN networks. The paper suggests that in order to obtain a failover performance with SCTP comparable to that obtained in traditional TDM-based SS7 systems, SCTP has to abandon many of the configuration recommendations of RFC 2960 and become much more aggressive in its failover behavior. Furthermore, the paper suggests that the SCTP parameter Path.Max.Retrans has a major impact on the SCTP failover performance. Our evaluation also indicates that for those path propagation delays envisioned in future SIGTRAN networks, the impact of the path propagation delay on the failover performance is marginal.

1. INTRODUCTION

Unlike a datacom network, a telecom network logically comprises two networks: a transport and a signaling network. The transport network carries the voice traffic, while the signaling network carries the control information that is needed for the administration and supervision of calls, and the management of the telecom network itself.

Traditionally, signaling traffic and voice traffic are both carried over TDM-based, circuit-switched connections. However, this is about to change. Using IP networks and protocols,

[†]Parts of the results in this paper were presented as work in progress at the First Swedish National Computer Networking Workshop (SNCNW) 2003. telecom operators are seeing ways to improve resource utilization and reduce the operational, maintenance, and network infrastructure costs. Still, the transition from TDM to IP will not happen overnight – maybe never. The traditional telecom network represents a huge capital investment¹ and is still unsurpassed in terms of reliability and QoS [12]. To address the situation of two different, co-existing, networks, one TDM based and one IP based, the IETF SIGTRAN working group has developed an architecture for signaling traffic over IP. In particular, they have developed an architecture for running Signaling System No. 7 (SS7), the predominant signaling system in traditional TDM-based telecom networks, over IP. Together with the so-called SoftSwitch architecture, the SIGTRAN architecture [14] constitutes a complete solution for the integration of the two networks.

The interoperability between the traditional TDM-based telecom network and its IP counterpart requires that the signaling performance in the IP network is comparable to that of TDM. Although some time has passed since the SIGTRAN architecture was first published, it is still unclear if it will perform comparable to the traditional telecom network [5], or if it will lead to unacceptable performance degradations [6].

The SIGTRAN architecture specifies a common transport protocol for all SS7 signaling traffic – SCTP [17], and a number of adaptation layers that run on top of SCTP. Although several adaptation layers have been specified, it seems as if a majority of telecom companies have embraced the MTP-L3 User Adaptation Layer (M3UA) [16]. This adaptation layer mimics the functionality of MTP-L3, the SS7 transport layer, and makes it possible to run all layers of the SS7 stack above MTP-L3 without modification on top of SCTP.

The Message Transfer Part (MTP) of the SS7 stack, of which MTP-L3 is the topmost layer, is not only responsible for the reliable transmission of signaling traffic, but also for network redundancy. In particular, link failures in traditional TDM-based SS7 networks are primarily managed by MTP. When a link failure occurs, this is detected by layer 2 in MTP (MTP-L2). MTP-L2 informs MTP-L3 about the failed link, and a so-called changeover is performed by MTP-L3. The

¹There is more than \$350 billion of legacy equipment installed in the current telecom network [3].



Figure 1. Evaluated network scenario.

changeover procedure diverts the signaling traffic carried by the unavailable link to alternate links as quickly as possible while avoiding message loss, duplication, or reordering.

To obtain a corresponding network redundancy in a SIGTRAN network as in a traditional SS7 network, SCTP supports so-called multi-homed associations. Multi-homed associations make it possible to manage several TCP-like connections, 'paths' in SCTP, as one redundant logical connection. When one path goes down, SCTP performs a failover and switches all traffic to an alternative path. A similar failover mechanism as the one in SCTP is also provided by M3UA, therefore we henceforth call failovers in SCTP, SCTP-controlled failovers.

This paper evaluates the performance of SCTP-controlled failovers in M3UA-based SIGTRAN networks: both in terms of SCTP failover times, and in terms of the maximum Message Signal Unit (MSU) transfer times experienced by M3UA users during failovers. Moreover, the paper studies to what extent the performance of SCTP-controlled failovers correlates with the path propagation delay, and with the SCTP parameter Path.Max.Retrans, the upper bound on the SCTP path error counter.

Our main contribution is to show that in order to have performance similar to the changeover procedure in a traditional SS7 network, SCTP has to be configured much more aggressively than what is recommended in RFC 2960. It is also shown that for the envisioned path propagation delays in future SIGTRAN networks, the effect of the path propagation delay on the SCTP failover performance is minor. However, there seems to be a strong correlation between failover performance and the value of the Path.Max.Retrans parameter. Specifically, we observe that in order to comply with the SS7 performance requirements, SCTP should not have Path.Max.Retrans set to a value larger than 3.

A similar experiment as the one presented in this paper has been carried out by Jungmaier et al. [10]. However, their experiment considered the MTP-L2 Peer-to-Peer adaptation layer (M2PA) [13]. Furthermore, Caro Jr. et al. at the University of Delaware have made extensive simulation studies of issues related to SCTP multi-homed associations. They have, among other things, suggested a two-level threshold mechanism [4] as an improvement to the existing SCTP failover mechanism.

The remainder of the paper is organized as follows. Section 2 describes the experimental procedure and setup. The results of the experiment are presented and analyzed in Section 3. Finally, Section 4 concludes the paper and makes some comments on future work.



Figure 2. Experiment setup.

2. METHODOLOGY

The purpose with our experiment was to evaluate the performance of SCTP-controlled failovers in the typical network scenario depicted in Figure 1. SEP1 and SEP2 are two SIGTRAN signaling endpoints, each one running an M3UA application. The two M3UA applications are engaged in a signaling session in which the SEP1 application acts as the source of the signaling traffic and the SEP2 application acts as the sink. During the signaling session, SEP2 becomes unreachable via its primary path; SCTP at SEP1 detects the failed primary path and performs a failover to the alternate path. When the failover has completed, the signaling session continues on the alternate path, and ends before the primary path has been recovered.

To evaluate the failover performance of SCTP in the network scenario of Figure 1, we used the experiment setup illustrated in Figure 2. The flow of events in the test runs of the experiment mimicked closely the flow of events in the evaluated network scenario. The source application at SEP1 continuously sent MSUs to the sink application at SEP2. When 30s of a test run had elapsed, i.e., more than enough time for SCTP to enter its stationary transmission behavior, the primary path was broken. A failover occurred, and the source application resumed its transmission on the alternate path. The test run ended 90s after the primary path was taken down, which was enough time for SCTP to conclude the failover and regain its stationary transmission behavior.

The two paths in between SEP1 and SEP2 consisted of links of bandwidth 100Mbps. Both paths included link emulators (L1 and L2 in Figure 2) that enabled us to vary the propagation delays of the two paths. In addition, L1 enabled us to introduce path breaks on the primary path. The link emulators were PCs running FreeBSD 5.0 and dummynet [15].

All tests were run automatically by the SEP1 test manager program with assistance from the SEP2 test manager and the L1 and L2 path managers. The SEP1 test manager directly administered the execution of the source application and the

Parameter	SCTP Configuration				
	RFC2960	Telecom(p)			
RTO_{init}	3000ms	80ms			
RTO_{min}	1000ms	80ms			
RTO_{max}	60000ms	150ms			
Path.Max.Retrans (p)	5	2	3	4	5
Heartbeat Interval	30000ms	30000ms			
SACK Timer	200ms	40ms			

Table 1. Evaluated SCTP configurations.

Table 2. Executed tests.

SCTP Configuration	Path Propagation Delay (ms)
RFC2960	5, 10, 20
Telecom(2)	5, 10, 20
Telecom(3)	10
Telecom(4)	10
Telecom(5)	5, 10, 20

SEP1 SIGTRAN stack. Furthermore, via commands, the SEP1 test manager controlled the execution of the SEP2 test manager and the L1 and L2 path managers. The SEP2 test manager and the L1 and L2 path managers, in their turn, acted as proxies to the SEP1 test manager. That is, on behalf of the SEP1 test manager, they administered the execution of the sink application and the SEP2 SIGTRAN stack, as well as performed the configuration of dummynet at L1 and L2.

In all test runs, event logging took place at both SEP1 and SEP2. Therefore, it was important that the local clocks of SEP1 and SEP2 were synchronized. To this end, NTP was used which kept the clocks of SEP1 and SEP2 differ with about 10ms in our experiment.

Six SCTP configurations were evaluated. The six evaluated SCTP configurations are shown in Table 1. The configuration denoted RFC2960 is the configuration of SCTP recommended in RFC 2960 [17]. A special notation is used for the remaining five SCTP configurations, Telecom(p), where 'p' is the value of the SCTP parameter Path.Max.Retrans.The notation alludes to the fact that these configurations are all variations of Telecom(2), which is the configuration recommended by some large telecom companies. In particular, the other four Telecom configurations included in the experiment are all examples of SCTP configurations which, in terms of failover, are more conservative than Telecom(2).

Tests were performed with three different path propagation delays: 5ms, 10ms, and 20 ms. These delays are believed to represent typical path propagation delays in future dedicated SIGTRAN networks.

Only a subset of the possible combinations of path propagation delay and SCTP configuration were tested. Specifically, our experiment comprised the 11 tests listed in Table 2. Each test was run 10 times giving a total of 110 test runs.

As follows from Table 2, RFC2960, Telecom(2), and Telecom(5) were tested with all three path propagation delays. This made it possible for us to study the correlation between failover performance and path propagation delay for, on one hand, the SCTP configuration recommended by IETF, and, on the other hand, for the, in terms of failover conservativeness, extremes of the Telecom configurations. The SCTP configurations Telecom(3), and Telecom(4) were only tested with a path propagation delay of 10ms. However, combined with the corresponding tests for Telecom(2) and Telecom(5), these tests enabled us to study the correlation between the SCTP failover performance and the SCTP parameter Path.Max.Retrans.

3. RESULTS

As briefly mentioned in Section 2, event logging at SEP1 and SEP2 took place in all test runs. Specifically, the time the primary path was broken and the time the path failure was detected by SCTP at SEP1 were logged. The failover time in a test run was then calculated as the difference between the SCTP detection time and the actual time of the path failure.

Also the sending times of the MSUs by the source application, and the reception times of the MSUs by the sink application were logged during each test run. (Note that the timing of the MSUs occurred at the level of the M3UA application, and not at the SCTP level.) Based on these values, the MSU transfer times were calculated as the difference between the reception and the sending times of the MSUs.

Figure 3 and Table 3 summarize the results of the measurements of the failover times and the MSU transfer times for the three SCTP configurations: RFC2960, Telecom(2), and Telecom(5). Recall from Section 2 that RFC2960 is the configuration of SCTP recommended in RFC 2960 [17]; that Telecom(2) is an SCTP configuration with strong proponents in the telecom sector; and that Telecom(5) is a conservative version of Telecom(2). In particular, Telecom(5) is a merge of Telecom(2) and RFC2960: The RTO-parameters of Telecom(5) are the same as for Telecom(2), i.e., are set with respect to the envisioned delays in future SIGTRAN networks, while the failover behavior of Telecom(5) is as conservative as for RFC2960.

The lin-log graphs in Figure 3(a) plot the sample means of the measured failover times in the tests as a function of the path propagation delay. The sample means are also listed in Table 3. Specifically, Table 3 lists the sample means and their corresponding 99% confidence intervals.

It follows from Table 3 that the mean failover times for RFC2960 were of magnitude 63s for all three path propagation delays considered. This is not surprising since with five retransmissions until a path is abandoned (i.e., Path.Max.Retrans = 5), the theoretical failover time for RFC2960 (assuming that $RTO = RTO_{min}$, which was the case is all our tests) becomes exactly 63s: 1s + 2s + 4s + 8s +16s + 32s = 63s.

As shown in Figure 3(a), the failover times for the Telecom configurations were several orders of magnitude less than for RFC2960. In particular, it follows from Table 3 that the failover times of Telecom(2) were mostly in the range of 435ms - 505ms, while Telecom(5) had roughly twice the failover times of Telecom(2).



Figure 3. Failover performance vs. path propagation delay.

Table 3. 99% confidence intervals for failover performance vs. path propagation delay.

	Failover Time (ms)			Max. MS	U Transfer Tii	ne (ms)
Path Propagation Delay (ms)	5	10	20	5	10	20
RFC2960	63086 ± 44	63147 ± 46	63244 ± 28	61809 ± 1403	62612 ± 31	62735 ± 45
Telecom(2)	458 ± 23	484 ± 20	480 ± 16	457 ± 41	457 ± 29	495 ± 46
Telecom(5)	975 ± 17	1008 ± 16	1093 ± 29	592 ± 24	620 ± 19	718 ± 23

As mentioned in Section 1, the corresponding path failure scenario to the one studied in our experiment is managed by the MTP-L3 changeover procedure in a traditional SS7 network. According to ITU-T recommendation Q.706 [9], the changeover time in an SS7 network must be less than or equal to 800ms. Since basically the same applications will be used in future SIGTRAN networks that is used in current SS7 networks, it is reasonable to assume that the requirements are roughly the same. Thus, it follows from our experiment that RFC2960 most likely will fail to meet the Q.706 requirement on changeover. In fact, the failover times of RFC2960 were almost 80 times the changeover limit of Q.706. This is, of course, to be expected, and is in agreement with the results reported in [6] and [10]. More interestingly, we observe that while the failover times of Telecom(2) were well below the changeover limit of Q.706, this were not the case for Telecom(5). Thus, it seems that if SCTP is to be used for transfer of signaling traffic, it not only has to abandon the conservative RTO settings of RFC 2960, but also has to switch from a failed path less conservatively than recommended by RFC 2960.

Figure 3(a) and Table 3 also suggest that the path propagation delay only had a minor impact on the SCTP failover time – at least for propagation delays no greater than 20ms, i.e., for those path propagation delays considered typical in future dedicated SIGTRAN networks. Specifically, the increase in mean failover time for RFC2960 when the path propagation delay was increased from 5ms to 20ms was much less than 1%; for Telecom(2) the increase was about 5%; and for Telecom(5) the increase was close to 12%.

Still, there was indeed a correlation between failover time and path propagation delay. The correlation could, as follows from Table 3, only be established for RFC2960 and Telecom(5). However, for these two SCTP configurations there was, with a 99% confidence, an increase of the failover time when the path propagation delay increased from 5ms to 20ms.

In the same way as for the failover times, Figure 3(b) and Table 3 give the results of the measurements of the maximum MSU transfer times. To avoid having the SCTP slow start and the transient behavior of SCTP during the termination of a test run interfere with the results, the first and last seconds of a test run were excluded from the calculation.

The graphs show that the maximum MSU transfer times for RFC2960 and Telecom(2) were almost the same as their failover times, while Telecom(5) had maximum MSU transfer times about 380ms less than its failover times. Contrary to the failover times, there is no ITU-T recommendation that explicitly governs the MSU transfer times. Instead, the upper bound of the MSU transfer times are determined by the application layers atop MTP-L3, i.e., the MTP-L3 stakeholders.

The primary stakeholders of MTP-L3 in terms of MSU transfer time are the ISUP (ISDN User Part) [8] and TCAP (Transaction Capabilities Application Part) [7] application protocols. The basic function of ISUP is to control setup,



Figure 4. Failover performance vs. Path.Max.Retrans for Telecom(p) (10ms path propagation delay).

connection, and teardown of telephone calls, while TCAP is an application protocol that is used by a large number of distributed SS7 applications. Examples of applications using TCAP include various Intelligent Networking (IN) applications and mobility support applications in mobile networks (i.e., GSM and IS-41).

Although, neither ISUP nor TCAP imposes any explicit requirements on MSU transfer times, analyses have been made [1], [2], [11] suggesting that the maximum permissible MSU transfer times with respect to these application protocols are in the range of 600ms - 1000ms, with 1000ms being barely acceptable. With these figures in mind, it is obvious that RFC2960, with maximum MSU transfer times of about 63s, did not comply with the ISUP/TCAP requirements. Again, as with the RFC2960 failover times, this was to be expected. Less expected was that also Telecom(5) had some difficulties passing the ISUP/TCAP requirements. As follows from Table 3, the mean maximum MSU transfer time for Telecom(5) at a path propagation delay of 20ms was 718ms. Considering that the ISUP/TCAP requirements are worst case values, and that the measurements took place in a scenario with no competing traffic, Telecom(5) may not give adequate MSU transfer times during a failover in a real SIGTRAN network. Thus, the outcome of the maximum MSU transfer time measurements only reinforces the outcome of the failover times: If SCTP is to be used for signaling traffic, then it has to be much less conservative than recommended by RFC 2960.

Figure 3(b) and Table 3 also show that the path propagation delay only had a minor impact on the maximum MSU transfer times experienced during a failover. Furthermore, the correlation between maximum MSU transfer time and path propagation time was weak, and could only be established for Telecom(5).

We also performed a more detailed study of the impact of the SCTP parameter Path.Max.Retrans on the SCTP failover performance. The outcome of this study is compiled in the graphs in Figure 4². The graphs plot the sample means of the measured failover times and maximum MSU transfer times together with their 99% confidence intervals.

It follows from the graphs that the value of Path.Max.Retrans had indeed a major impact on the failover time. An increase of Path.Max.Retrans from 2 to 3 resulted in a relative increase of the mean failover time by 40%. And, when Path.Max.Retrans was increased from 3 to 4, or from 4 to 5, the relative increase of the mean failover time was about 20% in both cases. Even more important is to note that already with a Path.Max.Retrans of 4, SCTP failed to meet the failover requirement of Q.706. Thus again reinforcing the need for SCTP to be much more aggressive than what is recommended by RFC 2960 if it is to be used for SS7 signaling transport.

The graphs also show that the value of Path.Max.Retrans had some influence on the maximum MSU transfer time. Specifically, the maximum MSU transfer time increased with approximately 35% when the value of Path.Max.Retrans was changed from 2 to 5. However, the maximum MSU transfer times were below ISUP/TCAP requirements for all values of Path.Max.Retrans. Thus, in terms of MSU transfer time there was no problem having Path.Max.Retrans configured as conservatively as recommended by RFC 2960 and still meet the SS7 signaling transport needs.

4. CONCLUSIONS

This paper presents an evaluation of the performance of SCTP-controlled failovers in future M3UA-based SIGTRAN networks. The evaluation suggests that in order to meet the failover performance objectives of a traditional SS7 network,

²The dotted lines in the graphs are only provided to make the trends more clear, and do not suggest that Path.Max.Retrans is continuous.

SCTP has to abandon the conservative failover behavior recommended by RFC 2960. Specifically, it has to set the parameter Path.Max.Retrans to a value no larger than 3. In addition, it has to change from the RTO-parameter configuration recommended by RFC 2960 to a parameter configuration far more in line with the actual path propagation delays in the SIGTRAN network.

The evaluation also suggests that the configuration of the SCTP parameter Path.Max.Retrans has a major impact on the failover performance: Especially in terms of failover time, but also to some extent in terms of the maximum MSU transfer time experienced by an M3UA application during failover.

In contrast, the evaluation indicates that for path propagation delays in the range of 5ms to 20ms, i.e., for path propagation delays believed to be representative for dedicated SIGTRAN networks, the path propagation delay has only a minor impact on the failover performance.

Our future work includes studying the effects of introducing competing signaling traffic on the performance of SCTP-controlled failovers. In particular, to study the tradeoff between shorter failover times and spurious failovers. However, we also want to study to what extent the SCTP failover performance degrades with different levels and mixtures of competing traffic. Furthermore, it remains to find out how other configurable SCTP parameters, e. g., RTO_{min} and RTO_{max} , affect the failover performance.

5. ACKNOWLEDGEMENTS

We would like to acknowledge all persons at TietoEnator Telsys in Karlstad who have made this paper possible. In particular Patrik Forslund, Mikael Blom and Nils Böjeryd for their assistance during the specification of the experiment in this paper. Thanks also to the many people at TietoEnator who gave generously of their time and expertise during the work with this paper. This includes Ulf Melin who helped us with the setup and configuration of the SIGTRAN stacks; Ulf Asplund and Lennart Wålfors who assisted us during the development of the test programs; and Robin Thunell and Johan Öberg for their generous support on SCTP and M3UA implementation details. Last but not least, our sincere thanks to the management of TietoEnator Telsys who made this paper possible in the first place.

REFERENCES

- Broscius, T. S. A., Huitema, C., and Lin, H.-A. P. "Performance requirements for signaling in Internet telephony,". Internet draft, IETF, November 1998. Expired May 16, 1999.
- [2] Broscius, T. S. A., Huitema, C., and Lin, H.-A. P. "Performance requirements for TCAP signaling in Internet telephony,". Internet draft, IETF, November 1998. Expired May 16, 1999.
- [3] Buga, W. "The evolution of softswitch architecture,". Annual Review of Communications 55 (1997), 73–76.
- [4] Caro Jr., A. L., Iyengar, J. R., Amer, P. D., Heinz, G. J., and Stewart, R. R. "Using SCTP multihoming for fault tolerance and load balancing,". In *SIGCOMM 2002 Poster Session* (Pittsburg, Pennsylvania, USA, August 2002).
- [5] Coene, L., and Pastor, J. "Telephony signalling transport over SCTP applicability statement,". Internet draft, IETF, August 2003. Work in Progress.

- [6] Gradischning, K. D., and Tuexen, M. "Signaling transport over IPbased networks using IETF standards,". In *Proceedings of the Third International Workshop on the Design of Reliable Communication Networks (DRCN)* (Budapest, Hungary, October 2001), pp. 168–174.
- [7] ITU-T. "Signalling system no. 7 functional description of transaction capabilities,". ITU-T Recommendation 771, ITU, June 1997.
- [8] ITU-T. "Signalling system no. 7 ISDN user part functional description,". ITU-T Recommendation 761, ITU, December 1999.
- [9] ITU-T. "Signalling system no. 7 message transfer part signalling performance,". ITU-T Recommendation 706, ITU, March 1999.
- [10] Jungmaier, A., Rathgeb, E. P., and Tuexen, M. "On the use of SCTP in failover scenarios,". In *Proceedings of the 6th World Multiconference* on Systemics, Cybernetics and Informatics (Orlando, Florida, USA, July 2002).
- [11] Lin, H.-A. P., Yang, K.-M., Seth, T., and Huitema, C. "VoIP signaling performance requirements and expectations,". Internet draft, IETF, October 1999. Expired March 22, 2000.
- [12] Molinero-Fernandez, P., McKeown, N., and Zhang, H. "Is IP going to take over the world (of communications)?,". ACM Computer Communication Review 33 (January 2003), 113–118.
- [13] Morneault, K., Dantu, R., Sidebottom, G., Bidulock, B., and Heitz, J. "Signaling system 7 (SS7) message transfer part 2 (MTP2) — user adaptation layer,". RFC 3331, IETF, September 2002.
- [14] Ong, L., Rytina, I., Garcia, M., Schwarzbauer, H., Coene, L., Lin, H., Juhasz, I., Holdrege, M., and Sharp, C. "Framework architecture for signaling transport,". RFC 2719, IETF, October 1999.
- [15] Rizzo, L. "Dummynet: a simple approach to the evaluation of network protocols,". ACM Computer Communication Review 27, 1 (January 1997), 31–34.
- [16] Sidebottom, G., Morneault, K., and Pastor-Balbas, J. "Signaling system 7 (SS7) message transfer part 3 (MTP3) — user adaptation layer (M3UA),". RFC 3332, IETF, September 2002.
- [17] Stewart, R., Xie, Q., Morneault, K., Sharp, C., Schwarzbauer, H., Taylor, T., Rytina, I., Kalla, M., Zhang, L., and Paxson, V. "Stream control transmission protocol,". RFC 2960, IETF, October 2000.