An Analysis of the Performance of an MIH Based Switch Strategy for Heterogeneous Network Mobility

Enda FALLON[1,2], Liam MURPHY[1] , John MURPHY[1]
[1] *Performance Engineering Laboratory, University College Dublin, Ireland*
Tel: +353 1 716 2914 Email: liam.murphy@ucd.ie, j.murphy@ucd.ie
[2] *Software Research Centre, Athlone Institute of Technology, Ireland*
Tel: + 353 90 6471877, Email: efallon@ait.ie

Abstract: The increasingly ubiquitous deployment of wireless networks has created significant opportunities for mobile application developers. The availability of high-capacity, low financial cost networks has the potential to enable mobile application developers to offer feature-rich end-user-oriented product offerings. Standardisation efforts such as Media Independent Handover (MIH) will support this trend by providing developers with a detailed view of the financial cost and performance capabilities of available wireless and mobile networks. Even with the increasingly pervasive deployment of Wireless LAN (WLAN), it is unlikely those mobile applications will exclusively utilise WLAN networks, as many are deployed without consideration of seamless network mobility. In the scenario where WLAN coverage gaps exist, it will be necessary to utilise mobile networks such as UMTS as a communication bridge. In such an environment, the mobility support provided by multi-homed protocols such as the Stream Control Transmission Protocol (SCTP) has significant potential. This paper investigates how MIH performance characteristics - in particular the Received Signal Strength (RSS) from a WLAN Access Point (AP) - can be used to optimise SCTP switch performance. Results indicate that in certain situations there is a performance problem with utilising a "retransmit on same path" approach to lost packet retransmission when the primary path is explicitly set, as would be the case with MIH. Careful configuration of the RSS threshold for switchover is shown to alleviate the problem.

Keywords: SCTP, WLAN, Network Migration

1. Introduction

The availability of multiple underlying heterogeneous networks with diverse capabilities is driving the need for seamless session continuation during network migration. Originally introduced in 2002 SCTP [1] was intended for use in fixed line telecommunications networks. The primary focus of the protocol was to implement path redundancy for telecommunications signalling data. While the multi-homing features of SCTP are suitable for supporting path redundancy they also have an applicability to network mobility. In a heterogeneous wireless/mobile network environment the Mobile SCTP [2] extension of the SCTP protocol can be used to detect network paths, add an additional network path to an existing association and delete a path from an association. While the multi-homed features of SCTP originally designed for path redundancy can also be used to implement path handover during network migration, a number of issues have been raised in relation to the use of SCTP for network mobility.

The proposed IEEE Media Independent Handover protocol [3], also known as 802.21, defines an approach to facilitating service transition based on triggering and providing network detection and selection assistance. This paper investigates how MIH performance characteristics, in particular RSS, can be used to optimise SCTP switch performance for heterogeneous WLAN and UMTS networks.

This paper is organized as follows. Section 2 details related work in the area. Section 3 provides an overview of MIH and SCTP. The test results are presented in Section 4. The business implications of the work undertaken are discussed in Section 5. Conclusions are discussed in Section 6.

2. Related Work

In [4] the effect of paths with diverse transmission capabilities within a single association is investigated. The results presented indicate that the default SCTP retransmission strategy degrades acutely when the secondary path delay is less than the primary path at a certain level. The coupling of the logic for data acknowledgment and path monitoring are investigated in [5] and two SCTP stall scenarios are presented. In [6] the authors identify that a finite receiver buffer will block CMT-SCTP transmission when the quality of one path is lower than others. Several retransmission policies are studied which can alleviate receiver buffer blocking. In [7] an SCTP performance limitation in WLAN environments is presented. The results indicate that SCTP behaves in a counter intuitive manner allowing more time for switchover as network conditions degrade.

There are also a number of efforts ongoing which investigate how the performance of low layer switch strategies can be enhanced through the use of MIH type functionality. In [8] an integrated MIH/SCTP based switch strategy is adopted. The performance of this strategy is assessed in a number of wired and wireless network environments. In [9] and [10] an enhanced handover mechanism utilising MIH type primitives are proposed for MIPv6 networks. These studies indicate that switch performance can be enhanced by removing router discovery time. In [11] an MIH based Universal Information service is presented which enables seamless network migration without the need for human intervention.

3. Technology Description

3.1 Media Independent Handover

802.21, also known as Media Independent Handover is an emerging IEEE standard which supports seamless handover between homogenous and heterogeneous networks. 802.21 does not in itself implement network handover, rather it provides information to allow handover to and from a range of networks including cellular, GSM, GPRS, WiFi, Bluetooth. The network handover enabling function within the protocol is implemented through the MIH function. The MIH function consists of 3 elements:

- Event Service - typically used to detect the need for handovers
- Command Service - enables higher layers to control the physical, data link and logical link layers
- Information Service - provides a framework by which a MIHF (Media Independent Handover Function) entity may discover and obtain network information existing within a geographical area to facilitate handovers

3.2 Stream Control Transmission Protocol Path Management

SCTP identifies a path by the IP address of the destination. During the protocol startup stage, each SCTP endpoint selects one of its peer's IP addresses as a primary path for data transmission. When the primary path fails, a backup path will be selected as primary path. SCTP sends heartbeat packets periodically to an idle address to detect the reachability of the address. SCTP defines "idle address" as: no data chunks are sent within the current heartbeat interval. An address is considered active if the sender received the expected acknowledgement from its peer within a designated period. Otherwise, if the number of consecutive transmission timeouts exceeds the protocol parameter Path.Max.Retrans, it means the address is inactive. If this inactive address is current primary path, a handover will occur. In SCTP, the selection of primary address is based on local policy, whereas in the Dynamic Address Reconfiguration extension of SCTP, an SCTP endpoint can recommend a primary address for its peer.

Calculation of Retransmission Time-Out (RTO) in SCTP

If an SCTP sender can not receive a response for an SCTP data chunk from its receiver within the time of RTO, the sender will consider this data chunk lost. If the number of data loss exceeds the SCTP parameter Path.Max.Retrans, the address will be marked as inactive by the sender. Therefore, RTO is a very important factor for handover and the stability of the protocol. RTO is calculated for each destination address separately based on the Smoothed Round-Trip Time (SRTT) and Round-Trip Time Variation (RTTVAR) of the path. It is initialized with RTO.Initial which is an SCTP parameter and can be configured by the user:

$$RTO = RTO.Initial \quad (3.1)$$

SRTT and RTTVAR of a path are calculated by the measurement of Round-Trip Time (RTT) of the path. The RTT measurement for a path is made for every round trip. When SCTP gets the first measurement of RTT: RTT.1st, SRTT and RTTVAR are initialized as:

$$SRTT = RTT.1^{st} \quad (3.2)$$

$$RTTVAR = {}^{RTT.1st}\!/_{2} \quad (3.3)$$

And RTO is updated to

$$RTO = SRTT + 4\ x RTTVAR \quad (3.4)$$

For each time SCTP gets a new measurement of RTT: RTT.new, SRTT and RTTVAR will be updated as follows:

$$RTTVAR.new = (1-\beta)\ x\ RTTVAR.old + \beta\ x\ (SRTT.old - RTT.new) \quad (3.5)$$

$$SRTT.new = (1-\alpha)\ x\ SRTT.old + \alpha\ x\ RTT.new \quad (3.6)$$

Where β and α are constants and their recommended values are 1/4 and 1/8 respectively. Then the new RTO is:

$$RTO = SRTT.new + 4\ x\ RTTVAR.new \quad (3.7)$$

If the new RTO is less than RTO.Min, it will be set to RTO.Min. If the new RTO is greater than RTO.Max, it will be set to RTO.Max. Every time a transmission timeout occurs for an address, the RTO for this address will be doubled:

$$RTO = RTO\ x2 \quad (3.8)$$

And if the new RTO is greater than RTO.Max, RTO.Max will be used for the new RTO. If the sender gets a response from the receiver and a new RTT is measured, SCTP will use this new RTT to calculate RTTVAR, SRTT and finally RTO by the equations (3.5) to (3.7).

4. Results

4.1 Experimental Results

In order to dimension WLAN network behaviour for use in further simulated studies 25 baseline tests were executed. Each tests involved pinging an AP with 65500 bytes of data every second. The tests were initiated adjacent to the access point. The mobile client then moved at slow walking pace, approximately 1m/sec, a distance of 100M. Figure 1 illustrates the average link delay for the 25 tests. These average values are then used in further simulated studies.

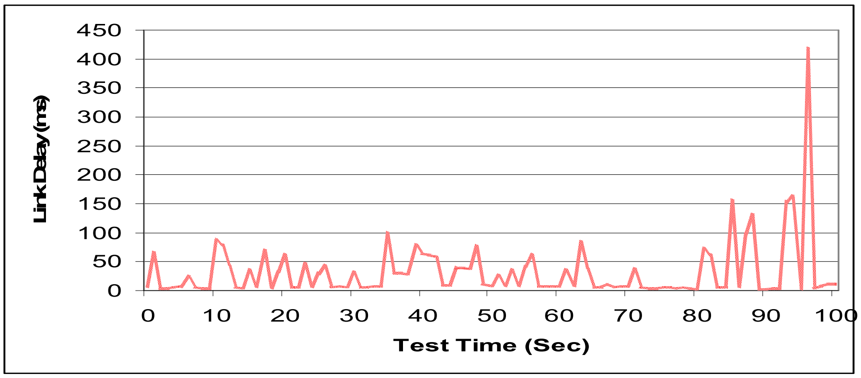

Figure 1: Average WLAN Link Delay

As the mobile node moves from the coverage of the AP the signal strength degrades. Figure 2 illustrates the average signal degradation.

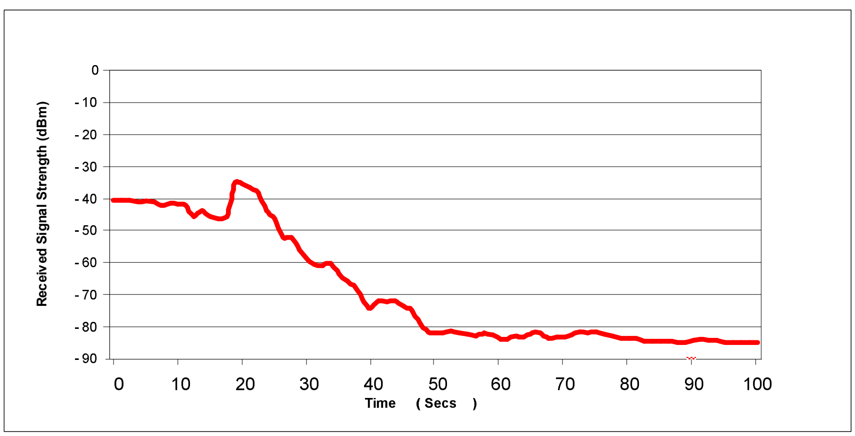

Figure 2: Received Signal Strength of Mobile Node

Table 1 details the percentage loss rate for 20 second periods. In the period 1-20 seconds the average loss rate on the link of 0%. As the mobile node moved from the coverage of the AP the signal degradation causes increased loss rates. In the time periods 21-40, 41-60 and 61-80 the average % loss rates were 2%, 3% and 8% respectively. Finally in the period 81-100 seconds the loss rate increases substantially to 15%.

Time Period (secs)	**1-20**	**21-40**	**41-60**	**61-80**	**81-100**
%Loss Rate	0	2	3	8	15

Table 1: Average % Loss Rates

4.2 Simulated Results

In order to accurately analyse the performance of a MIH based approach to heterogeneous network mobility an NS2 simulation was created which utilised the University of Delewares [12] SCTP module. The simulation used the results of the experimental study described in section 4.1. The simulation topology is detailed in Figure 3.

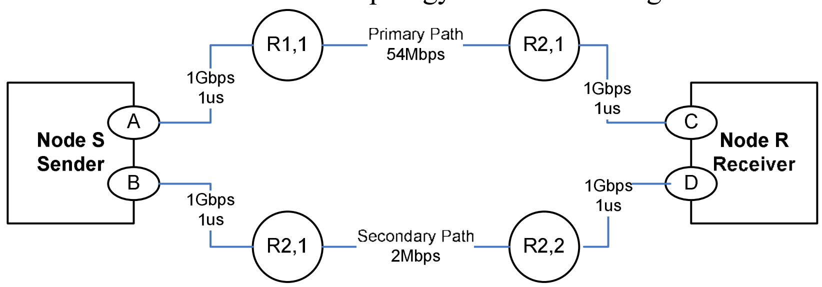

Figure 3: Simulation Configuration

Node S and Node R are SCTP sender and receiver respectively. Both SCTP endpoints have two addresses. R1,1, R1,2, R2,1 and R2,2 are routers. The implementation is configured with no overlap between the two paths. Node S begins to send FTP data to Node R after 0.5 seconds.

Previous investigations [4] have established that the default SCTP fast retransmission algorithm decreases performance when the primary and secondary path delay have a significant differential in transmission capability, for example, when a mobile node utilises both WLAN and UMTS networks. Results indicated that, contrary to the default mechanism specified in [1], fast retransmission on the primary path was encouraged when path bandwidths are high. For this reason the simulation was configured with the SCTP NS2 *rtxToAlt_* parameter set to retransmit lost packets on the same/primary (WLAN) path. For each test the default SCTP based switch functionality was disabled by setting an arbitrarily high PMR value. 10 tests were executed in which 200 seconds of FTP data was transmitted. For each test an explicit switchover was requested 10 seconds later than the previous test, using the SCTP NS2 *set-primary-destination* method. The last switchover was requested at 100 seconds. Figure 3 illustrates the total amount of data transmitted for each of the tests.

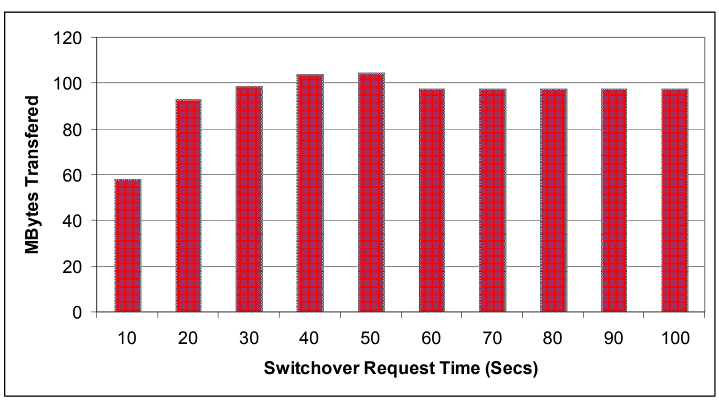

Figure 4: Total Data Transferred (Retransmission on Same Path)

The first test which requested an explicit switchover from the primary WLAN path to the secondary UMTS path after 10 seconds transmitted 58.15 Mbytes, the least amount of data transmitted of any of the tests. The subsequent tests with switchovers at 20, 30, 40 and 50 seconds transmitted 92.77, 98.83, 104.05 and 104.49 Mbytes respectively. The increase in total data transmission is logical as the tests which delay switchover use the higher capacity WLAN path for longer durations before switching to the secondary, lower capacity UMTS path.

The behaviour for the tests with switchover times from 60 to 100 seconds is more complex. Figure 5 illustrates the accumulated amount of data transmitted for the tests with switchover times from 10 to 60 seconds. The performance of tests with switchover times at 70, 80, 90 and 100 are identical to that of the test which switches over at 60 seconds. Figure 5 illustrates that the tests with switchover times from 60 to 100 seconds have each transmitted 97.92 Mbytes of data after 56 seconds. There is no subsequent successful transmission of data.

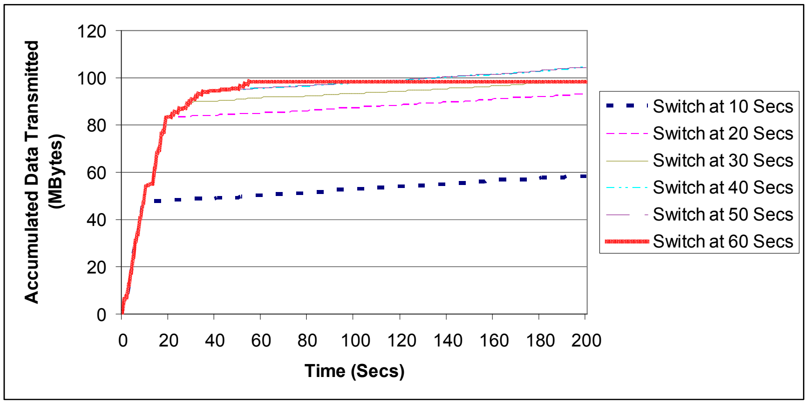

Figure 5: Accumulated Total Data Transferred (Retransmission on Same Path)

In [7] it was identified that increased Round Trip Times (RTT) in WLAN environments cause a significant increase in RTO calculations as a result of (3.7) and (3.8). The RTO calculation for the test where the retransmissions were on the same path and switchover was requested after 60 seconds is illustrated in figure 6. After 56 seconds a packet loss is detected and all buffered packets for the primary path are discarded. Using (3.8) the RTO is doubled from 1 to 2. The lost packet is retransmitted on the primary path rather than the alternate path according to [4]. As the network conditions for the WLAN have degraded this retransmitted packet is also lost, (3.8) is again applied doubling the RTO. At 60 seconds the primary path is explicitly changed to the UMTS path, as would be the case with an MIH based switch. However, since retransmitted packets are required to be transmitted on the same path as they were lost the explicit switchover has no effect. New packet transmission must await the successful transmission of the lost packets on the WLAN path.

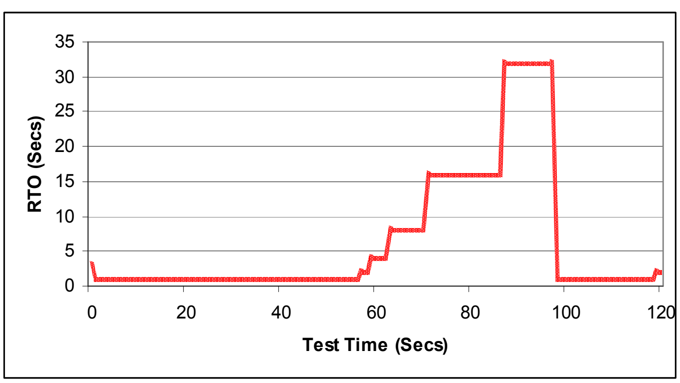

Figure 6: RTO for Test with Switchover at 60 Seconds (Retransmission on Same Path)

In order to overcome the path blocking caused by the retransmission of packets on the same path the default SCTP retransmission behaviour, *retransmit on alternate*, was configured and the tests were repeated. Figure 7 illustrates the results for the retransmit on alternate strategy and compares them to the retransmit on same strategy.

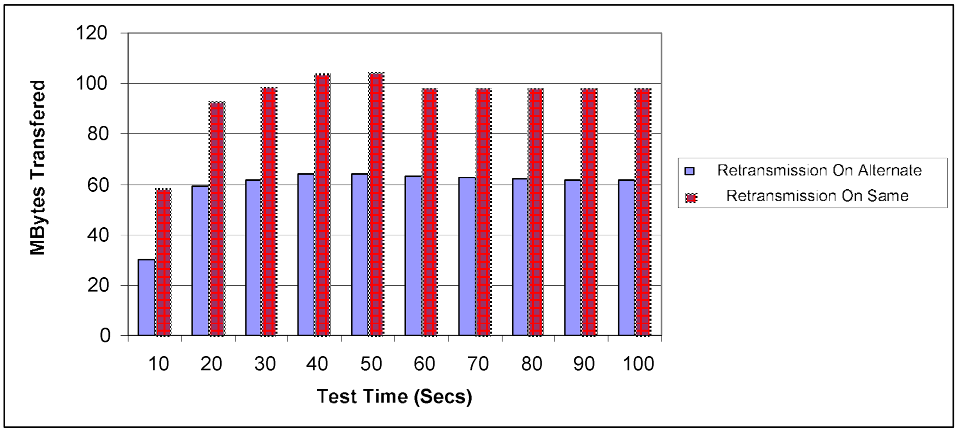

Figure 7: Total Data Transmitted for Switchover Times

Figure 7 illustrates that in general the retransmit on same strategy is more effective than the retransmit on alternate strategy. Figure 7 however, is over simplistic as only the final totals are detailed. It does not take account of the fact that after 57 seconds the test which utilise a transmit on same approach have effectively ceased transmitting data. Figure 8, which details the accumulated data transmitted for the retransmit on same strategy, illustrates that the tests which retransmit on alternate paths continue to transmit data, albeit at a poor performance rate for the entire duration of the test. The reason for this poor performance relates to the findings of [7].

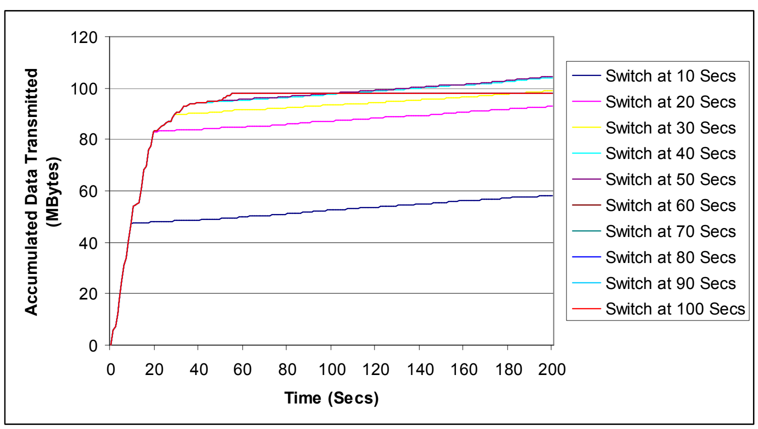

Figure 8: Accumulated Data Transmitted (Retransmission on Alternate Path)

5. Business Benefits

The implementation of mobile applications is a complex task. Developers must be network capability aware, device capability aware as well as having the necessary software development skills. It is as if traditional layered protocol models no longer apply to mobile computing. What is required is a scenario where mobile application developers can concentrate on the development of feature rich, end user oriented applications rather than taking the technical capabilities of the proposed deployment environment as the starting point. The work undertaken seeks to enable an application transparent approach to network utilisation. In this way application developers can take more of a network independent approach to application development.

6. Conclusions

This paper investigated how MIH performance characteristics, in particular the RSS from a WLAN AP could be used to optimise SCTP switch performance. The results indicate that there is a performance implication of utilising a *retransmit on same* approach to lost packet retransmission in a WLAN environment. The performance issue arises when an SCTP client experiences a period of increased RTO and increased loss rate. If a *retransmit on same* approach is employed the explicit setting of a primary path is not effective. This retransmission policy continues to resend lost packets on a degraded path while effectively delaying the transmission of first transmission packets on the newly selected primary path.

The results also indicate that a MIH oriented approach to switchover which uses received signal strength as the decision to switch path can improve performance when a retransmit on same strategy is used. Using Figure 2 in conjunction with the results detailed in Figure 4, the optimal time to switch path would be when the received signal strength reaches approximately -80dBm. By implementing a MIH based switch management strategy the issues of ineffective explicit primary path setting and performance degradation as a result of SCTP path differential can be avoided.

References

[1] R. Stewart, et al: Stream Control Transmission Protocol, RFC2960

[2] R. Stewart, et al: Stream Control Transmission Protocol, Dynamic Address Reconfiguration, proposed extension May 2006

[3] IEEE 802.21, Media Independent Handover Services, IEEE Standard under development. Available: http://www.ieee802.org/21/

[4] Qiao, Y, et al: "SCTP Performance Issue on Path Delay Differential" International Conference on Wireless/Wired Internet Communications (WWIC), Coimbra, Portugal 2007

[5] J. Noonan, et al: Stall and Path Monitoring Issues in SCTP, Proc. Of IEEE Infocom, Conference on Computer Communications, Barcelona, April 2006.

[6] J. Iyengar, et al: Receive buffer blocking in concurrent multipath transfer, IEEE Globecom 2005, St. Louis, Nov. 2005.

[7] Fallon, S, et al: "SCTP Switchover Performance Issues in WLAN Environments" , 5th IEEE Consumer Communications & Networking Conference (CCNC) January 2008 Las Vegas, Nevada

[8] Yung-Mu, , et al: "SCTP-based handoff based on MIH triggers information in campus networks", (2006) 8th International Advanced Communication Technology Conference

[9] Yoon, Y, et al: "Reduction of Handover Latency Using MIH Services in MIPv6", (2006) 20th International Conference on Advanced Information Networking and Applications, 2006

[10] Mussabbir, Q, et al: "Optimized FMIPv6 Handover using IEEE802.21 MIH Services" (2006) Proceedings of first ACM/IEEE international workshop on Mobility in the evolving internet architecture

[11] Dannewitz , et al: "An IEEE 802.21-based Universal Information Service" in Proc. of the Wireless World Research Forum Meeting 2006

[12] A. Caro, et al : ns-2 SCTP module, Version 3.5, www.armandocaro.net/software/ns2sctp/.