Study and performance evaluation of TCP modifications and tuning over satellite links.

Mario Marchese
CNIT - Italian National Consortium for Telecommunications, DIST - Univ. of Genoa Research Unit
Department of Communications, Computer and System Science (DIST)
University of Genoa, Italy
Via Opera Pia 13, 16145 Genova (Italy)
Phone: +39-010-3532985
Fax: +39-010-3532154
e-mail: mario.marchese@cnit.it

Abstract - **The paper presents a study and an analysis of the performance offered by TCP over a GEO (Geostationary Orbit) satellite link. The characteristics of satellite channels (notably the large round trip delay) heavily influence the TCP flow control, which is essentially based on the acknowledgments. The effect is a delay in the acknowledgement reception and, consequently, in the delivery of messages, with respect to cabled networks. The drawbacks of using TCP in a satellite environment may be mitigated by a proper tuning of some TCP parameters. The behavior of the protocol as a consequence of variations in the buffer length of both transmitter and receiver and in the initial congestion window is investigated in the paper and a proper configuration that drastically improves performance (measured by the throughput in bytes/s and by the overall transmission time) is proposed. Two test environments have been used to evaluate the proposed modifications: a real testbed, composed of two remote hosts connected through a satellite channel, and a satellite network emulator, composed of three PCs (two of them representing two hosts, the third one reproducing the behavior of the satellite link). In both environments, a single ftp-like application designed for the aim has represented the reference application; three different file sizes have been used and the different effect of the tuning depending on the transfer length has been evidenced.**

I. INTRODUCTION

Due to the inherent broadcast capability of satellites, which can connect remote sites when there is no terrestrial infrastructure, providing, at the same time, high-speed links, there is a growing interest in providing interconnection and multimedia services by using satellite links. Several projects have been dedicated to this issue; we mention here just a few. Most of the research reported in the paper has been developed in the framework of [1], which is a project aimed at analyzing the problems related to a satellite or terrestrial/satellite interconnection, from the point of view of both transmission and network protocols. The project is funded by the Italian Space Agency (ASI), and managed by the Italian Consortium for Telecommunications (CNIT), a research center composed by several Italian universities. It is divided into two integrated lines: an experimental activity of multimedia services over a terrestrial-satellite network and a study activity for the system evolution concerning network protocols, integration of satellite and terrestrial networks, medium access techniques, resource allocation, development of terminal equipment and user interfaces. The project in [1] uses the ITALSAT satellite and works at 2 Mbit/s with an antenna of 1.8 m in the Ka-band (20÷30 GHz). It is strictly related to other European activities planned by the European Space Agency (ESA), as, for instance, DICE (Direct Inter-establishment Communications in Europe) [2], proposed as a multi-point videoconference system via satellite, and CODE (Co-operative Data Exchange), which is a VSAT system dedicated to LAN inter-connection (see [2] for an overview and [3] for more details). The European program about the development of 'Switched VSAT' [4] is also worth mentioning. In this system, ATM will be the reference technology (see [5] for a list of related projects) for the provision of different services, ranging from the interconnection of LANs to videoconference services. Moreover, the ARTES-3 program should be mentioned, where a research concerning novel pioneering systems, strictly related to the research activity of the project in [1], will be carried out. Another important project to be highlighted is the NASA ACTS [6], whose activity is, at least from the internetworking viewpoint, very similar to the topics addressed by [1] and by this paper.

The characteristics of a satellite network are very different from wired or wireless networks for which some protocols have been designed. As a consequence, the performance of protocols may be very different. For instance, the delay of a remote login or file retrieval in a satellite environment may be unacceptable for the user. So, when constructing a network containing a satellite link, or, as in this paper, when using a single satellite link to connect remote sites, there are many considerations to take into account and many points where to act [7]. An important issue is the performance of the transport protocol. If the transport protocol does not offer high performance, the network throughput may become really low and, as a consequence, the quality perceived by the users may be poor. The object of the paper is the investigation of the transport protocol and the proposal of some modifications to improve

This work has been supported by the Italian Space Agency (ASI) under the CNIT contract ASI-ARS-99-205 "Integration of Multimedia Services on Heterogeneous Satellite Networks".

0-7803-6283-7/00/$10.00 © 2000 IEEE

performance. TCP [8] has been chosen as transport protocol. The problem is currently investigated [9-14]. A review of the problems due to the utilization of TCP in a satellite environment is presented in [7], while an overview of the research performed in the field may be found in [9, 15]. This paper, which contains an extension of the activity presented in [16] and new results, studies TCP performance by tuning two TCP parameters, namely the buffer dimension and the initial congestion window. The behavior has been investigated by using an ftp-like application suited to transfer files of various dimensions. A single connection case is investigated. The experimental environment is based on a real testbed, where two hosts are connected through a Geostationary Orbit (GEO) [17] satellite link by means of routers, and on a satellite network emulator.

The paper is structured as follows. Section II contains the description of the main problems related to TCP in a satellite environment. Section III describes the test network and the application used to get the results. The proposals about the modified TCP and the performance analysis are contained in Section IV. Section V presents the conclusions and some possible future activities.

II. TCP IN A SATELLITE ENVIRONMENT

As stated in the Introduction, the use of the TCP/IP stack in satellite environments is really important: many applications are based on this protocol family and any connection with the Internet of the future would imply the use of TCP/IP. The network used heavily affects the behavior of the transport protocol; even in a satellite environment, the problems are different if a LEO (Low Earth Orbit), a MEO (Medium Earth Orbit) or a GEO satellite system is used [17]. Issues related to each environment are listed in [7]. A Geostationary system is approached in this work and the Round Trip Time (RTT) is above 500 ms. TCP has not been designed for these network characteristics: the high delay to receive acknowledgements decreases performance and makes the quality perceived by the users really poor. Improvements can be obtained by properly tuning some parameters and by modifying algorithms. The decision of taking into account a single link and a single connection may seem a strong limitation. It is surely true that the presence of a long heterogeneous path or of multiple high loaded connections may heavily influence the protocol behavior and, perhaps, vanify the improvement obtained with the satellite TCP version, but it is also true that the satellite part may receive a different treatment and attention with respect to the cabled parts of the networks [14, 18] and that the single link and single application case represents both a valid benchmark for working and a configuration used in real environments [6, 19]. The problem of enhancing TCP performance in satellite environment is approached in two ways in the literature. Either the application level is modified to better adapt the network characteristics [19], or the transport protocol itself is revised, which is the approach followed in this paper. A short summary of the TCP characteristics related to the paper topic may be of help before introducing the testing environment and the protocol modifications presented. A NewReno TCP is used under the 2.2.1 version of the Linux kernel. The parameters are substantially set following the standard in [20] and [21]. The notation used herein has been introduced in [20]. The TCP parameters of interest are reported in Table I; a C-like language is used for the description. The acronym cwnd is the congestion window, smss is the sender maximum segment size, ssthresh is the slow start threshold. FlightSize is the measure (in bytes) of the amount of data sent but not yet acknowledged, i.e., the packets still in flight. The real transmission window (TW) is set, in any case, to the minimum between cwnd and the receiver's advertised window (rwnd), if the buffer space of the transmitter does not represent a bottleneck. In this last case, the transmission buffer governs the transmission speed. The receiver window rwnd has been measured to be 32 kbytes at the beginning of the transmission. The receiver buffer space is automatically set by TCP to 64 kbytes. The Selective Acknowledgment (SACK) mechanism is utilized [21]. The paper proposes a study and a tuning of the receiver/transmitter buffer space and of the initial congestion window. The other algorithms are untouched. The performance metrics considered are the throughput and the overall time required for the transmission.

TABLE I
TCP PARAMETERS

TW=min{cwnd, rwnd}	
SLOW START [cwnd<ssthr]	cwnd=1·smss ssthr=∞ ACK → cwnd=cwnd+1·smss
CONGESTION AVOIDANCE [cwnd≥ssthr]	<cwnd> ACK → cwnd=cwnd+1
FAST RETRANSMIT / RECOVERY	ssthr=max{FlightSize/2, 2·smss} cwnd=ssthr+3·smss Delayed ACK → cwnd=cwnd+1·smss cwnd=ssthr

II. TEST NETWORK

The results presented in the following have been obtained by using two different tools: a real testbed and a satellite network emulator. The application used to perform the tests is described in Sub-section 3.3.

3.1. Real testbed

The real testbed is shown in Fig. 1: two remote hosts are connected through a satellite link by using IP routers. The TCP/IP protocol stack is used. The data link level of the router uses HDLC encapsulation on the satellite side, where a serial interface is utilized, and Ethernet on the LAN side. The main characteristics of the Radio Frequency (RF) devices used for the tests are contained in Table II; more details are described in [1]. A raw Bit Error Rate - BER (i.e., BER with no channel coding) approximately of 10^{-2} has been measured; the utilization of a sequential channel coding with a code rate of 1/2, to correct transmission errors, has allowed to reach a BER of about 10^{-8}. As a consequence, the higher layer protocol, the data link protocol, 'sees' a reliable channel. The system offers the possibility of selecting the transmission bit rate over the satellite link and a bit rate of 2048 kbit/s has been used for the tests.

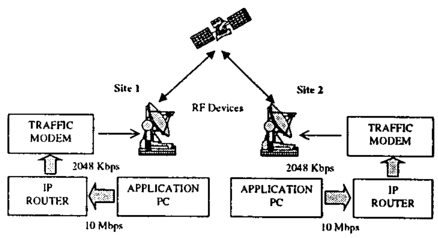

Fig. 1. Testbed.

TABLE II.
RF DEVICE CHARACTERISTICS (SSPA STANDS FOR SOLID STATE POWER AMPLIFIER)

Antenna diameter	1.8 m
TX Gain	52 dB
RX Gain	48.5 dB
SSPA Power	37 dBm (5 W)
Polarization	V
TX Center Frequency	29.75 GHz
RX Center Frequency	19.95 GHz

3.2. Network emulator

The satellite environment described in the previous Sub-section has been emulated by using three PCs connected as in Fig. 2: two of them (PC1 and PC3) represent the remote hosts; PC2 emulates the satellite channel and imposes a delay corresponding to the real measures on the satellite link (about 250 ms each hop) and a bandwidth of 2048 kbit/s. The application program NistNet is installed in PC 2. NistNet allows to impose the mentioned delay and the bandwidth bottleneck to adapt to the real situation. The emulator provides accurate results and may be succesfully used to substitute the real testbed when it is not available or when the tests would be too long and expensive to be performed in the real environment. It has actually been used to tune parameters and to complete the analysis in the case of a 11 Mbytes file transfer (see the Sub-section 3.3 and the Section 4 below).

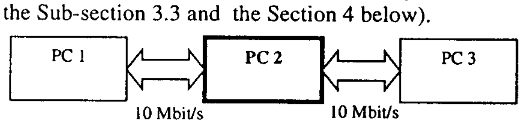

Fig. 2. Emulation testbed.

3.3. Test application

The application used to get the results is a simple ftp-like one, i.e. a file transfer application located just above the TCP. It allows to transfer data of variable dimension (H [bytes] in the following) between the two remote sites. A single connection case is taken into account; i.e. a single file transfer is performed a time. It should be highlighted that most of the Internet applications, as browsing, are mainly based on file transfer. The work has taken this application as a reference because it is thought as fundamental for most of the applications of interest.

IV. RESULTS

The analysis is dedicated to investigate the behavior of TCP by varying the value of the initial window (IW, measured in bytes, as indicated in Section II; i.e. the notation IW=1 means IW=1·smss) and of the buffer dimension, intended as the memory availability in bytes, for source and destination, which is kept equal; i.e., the buffer has the same length both for the source and the destination. It is identified with the variable 'buf' in the following. Concerning the initial congestion window, the issue has been treated in the literature. Simulation studies, though not for the specific satellite environment ([22]), show the positive effect of increased IW for a single connection. [23] clarifies the strict dependence of the performance on the application environment and suggests that "larger initial windows should not dramatically increase the burstiness of TCP traffic in the Internet today". IW is set to 1 in classical TCP, as stated in Section II. The performance improvement (i.e., the reduction of the time required for the whole transmission and the higher throughput) provided by varying the value of IW is shown in Fig. 3 for a buffer (buf) of 64 kbytes, which is the value set by TCP if no modification is imposed. A file transfer of 2.8 Mbytes is performed in this case.

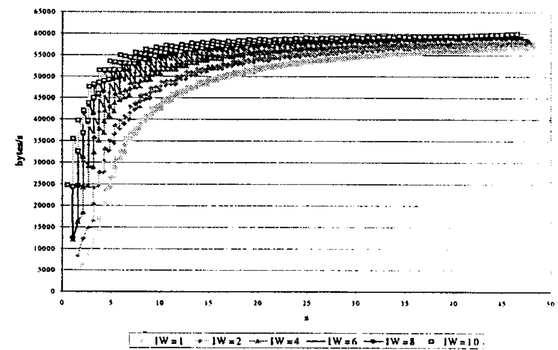

Fig. 3. Throughput (bytes/s) versus time for different values of the initial congestion window (IW), buf=64 kbytes, H=2.8 Mbytes.

The advantage of using an increased IW is more evident for shorter transfers. The transmission time necessary to perform a 100 kbytes transfer is reported in Table III, for different values of IW, along with the gain in percentage obtained with respect to the reference (buf=64 kbytes, IW=1).

The gain is computed as follows: if T_{REF} is the reference transmission time (T_{REF}=4.42 s, in Table III) and T is a generic transmission time, the percentage gain is defined as

$$\%\,\text{Gain} = \begin{cases} \dfrac{T_{REF} - T}{T_{REF}} \cdot 100, & \text{if } T < T_{REF} \\ 0, & \text{otherwise} \end{cases}$$

. If $T_{REF} \leq T$, there is no gain. For example, referring to Table III, if T=2.70 s, %Gain=(4.42-2.70)/4.42=38.9 %. The improvement obtained is noticeable. The best results for shorter files are due to the initial window behavior in the TCP: IW, as will be remarked after the analysis concerning the buffer dimension, is mainly responsible of the behavior in the first part of the transmission. As a consequence, short file transfers receive a real advantage from an increased IW. In both cases, Fig. 3 and Table III, there is no congestion and TCP is more aggressive, depending on the initial window; if cwnd, as presented in Section II, starts from a higher value, more data can be sent without waiting a too long time.

TABLE III.
OVERALL TRANSMISSION TIME FOR DIFFERENT VALUES OF THE INITIAL CONGESTION WINDOW (IW), BUF=64 KBYTES, H=100 KBYTES

IW	Transmission Time [s]	% Gain
1	4.42 s	
2	3.76 s	14.9 %
4	3.21 s	27.4 %
6	2.70 s	38.9 %
8	2.66 s	39.8 %
10	2.62 s	40.7 %

The tuning of the initial congestion window allows to mitigate the problem introduced by the channel delay. As regards the second parameter, Fig. 4 shows the behavior of the throughput (bytes/s) versus time for different values of the buffer length for IW=1. The classical TCP, currently used in cabled networks, is labeled as buf=64 kbytes. It is clear that, in this situation, the increase in speed is very slow, the TCP is drastically blocked by the satellite delay, the transmission window cannot increase its length because the buffer dimension represents an actual bottleneck for the system. The time required to transfer data is shown in Table IV, along with the percentage gain with respect to the reference (IW=1, buf=64 kbytes). If the buffer dimension is higher, the number of packets in flight is increased, as well as the system performance. The improvement is outstanding, as shown by the transfer time reduced below 16 s in the best case, corresponding to an improvement of about 68%.

The buffer length is very important for the performance of the system; it rules the congestion window by imposing a bottleneck to its increment. A short buffer drastically limits performance, but an excessively long buffer makes the system congested, as shown by the line labeled as 640 kbytes in Fig. 4.

Fig. 5 allows to detect the 'ideal' combination of IW and buf values for the configurations taken into account in the paper.

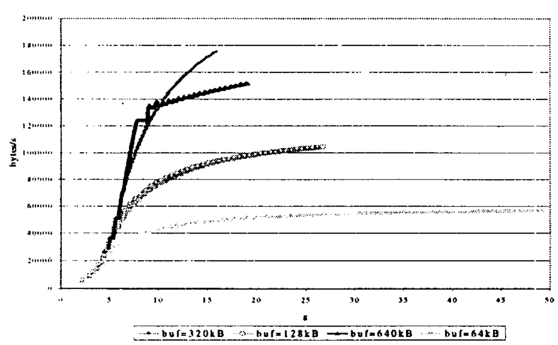

Fig. 4. Throughput (bytes/s) versus time for different values of the buffer length, IW=1, H=2.8 Mbytes.

TABLE IV
OVERALL TRANSMISSION TIME FOR DIFFERENT VALUES OF THE BUFFER, IW=1, H=2.8 MBYTES

buf [kbytes]	Transmission Time [s]	% Gain
64	49.21 s	-
128	26.67 s	45.80 %
320	15.87 s	67.75 %
640	26.67 s	61.03 %

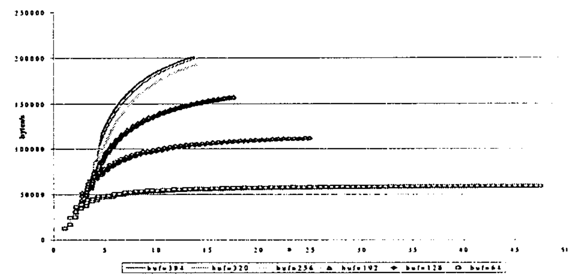

Fig. 5. Throughput (bytes/s) versus time for different values of the buffer length, IW=6, H=2.8 Mbytes.

Values of buffer over 384 kbytes, not shown in the graph, saturate the intermediate router; a value of 320 kbytes seems to guarantee a very high gain with respect to the classical solution of 64 kbytes and, at the same time, a certain margin against saturation. The percentage gain of the overall transfer time is reported in Table V versus the buffer length, up to the value of 384 kbytes, with respect to the (IW=6, buf=64 kbytes) configuration. The same quantities, up to 320 kbytes, with the same reference, are reported in Table VI and Table VII for H=100 kbytes and H=11 Mbytes, respectively. As is evident from the tables, the gain is light for short files, since the buffer does not represent a bottleneck for the system and IW rules the efficiency, while it is meaningful for larger transfers.

TABLE V
OVERALL TRANSMISSION TIME FOR DIFFERENT VALUES OF THE BUFFER, IW=6, H=2.8 MBYTES

buf [kbytes]	Transmission Time [s]	% Gain
64	47.55 s	-
128	24.97 s	47.5 %
192	17.61 s	63.0 %
256	14.24 s	70.1 %
320	13.96 s	70.6 %
384	13.76 s	71.1 %

TABLE VI
OVERALL TRANSMISSION TIME FOR DIFFERENT VALUES OF THE BUFFER, IW=6, H=100 KBYTES

buf [kbytes]	Transmission Time [s]	% Gain
64	2.7 s	-
128	2.69 s	0.4 %
192	2.67 s	1.1 %
256	2.66 s	1.5 %
320	2.65 s	1.9 %

TABLE VII

OVERALL TRANSMISSION TIME FOR DIFFERENT VALUES OF THE BUFFER, IW=6, H=11 MBYTES

buf [kbytes]	Transmission Time [s]	% Gain
64	191.1 s	-
128	95.4 s	50.1 %
192	64.8 s	66.1 %
256	56.1 s	70.6 %
320	55.5 s	71 %

Table VIII allows to evidence the effect of the parameter tuning performed in the paper. The table contains the combination of the two parameters analyzed (IW and buf), the time required for the overall transmission and the gain in percentage obtained with respect to the basic configuration (IW=1, buf=64 kbytes). The measures obtained with the real testbed and H=2.8 Mbytes have been chosen to summarize the results of the paper. The gain in the overall transmission time (up to 71.63 %) is mainly due to the buffer length, which may represent a real bottleneck for the system. On the other hand, the throughput in the initial phase is governed by IW. It is sufficient to observe Fig. 6, where the throughput versus time is shown for the same configurations of Table VIII. If the configurations with the same buffer length are analyzed, the difference between the increase in speed for different values of IW is outstanding. The behavior after 5 s may be taken as an example: (IW=1, buf=64 kbytes) has a throughput of about 26 kbytes/s, (IW=6, buf=64 kbytes) of 48 kbytes/s. The throughput for (IW=1, buf=320 kbytes) is about 35 kbytes, whereas it is 118 kbytes/s for (IW=6, buf=320 kbytes).

TABLE VIII

COMPARISON OF TCP CONFIGURATIONS BY VARYING THE INITIAL CONGESTION WINDOW AND THE BUFFER LENGTH, H=2.8 MBYTES

IW , buf [kbytes]	Transmission Time [s]	% Gain
1 , 64	49.21 s	-
6 , 64	47.55 s	3.37 %
1 , 320	15.87 s	67.75 %
6 , 320	13.96 s	71.63 %

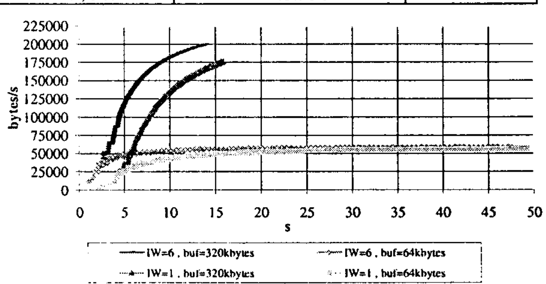

Fig. 6. Throughput (bytes/s) versus time for different values of the buffer length and of the initial congestion window, H=2.8 Mbytes.

V. CONCLUSIONS

The paper has presented an analysis of TCP behavior by tuning parameters as the initial congestion window and the buffer length. The TCP version which is commonly used in cabled networks, taken as a reference in the paper, results inefficient. The delay imposed by the satellite link heavily affects the acknowledgement mechanism and drastically reduces the overall throughput. The improvement due to an increased dimension of the transmitter/receiver buffer and of the initial congestion window is outstanding. A gain above 71% is reached with respect to the reference configuration. The case of one hop and one connection taken into account in the paper is used extensively in the literature and it is very useful to analyze some peculiarities of TCP. The multiple connection case will be considered in the future along with the behavioral analysis of slow start and congestion avoidance schemes, as well as of fast retransmit/recovery algorithms.

REFERENCES

[1] D. Adami, A. Conti, D. Dardari M. Marchese, L. S. Ronga, "CNIT-ASI Project 'Integration of Multimedia Services on Heterogeneous Satellite Networks': An Overview", Proc. Fifth Ka Band Utilization Conference, Oct. 1999, Taormina, Italy.

[2] The DICE/CODE Project Page, http://www-inw.tu-graz.ac.at/dice/welcome.htm.

[3] G. Fairhurst, C.T. Spracklen, N. Samaraweera, K. Ord, "A Report on the Performance of the IP Networking Software Used by CODE", Contract Report, January 1993.

[4] ESA Switched VSAT (XVSAT), http://www.erg.abdn.ac.uk/users/silas/systems/xvsat.html

[5] Summary Information about ATM Projects, http://www.erg.abdn.ac.uk/users/silas/systems/index.html

[6] NASA ACTS (University of Maryland), http://www.erg.abdn.ac.uk/users/silas/systems/UoMACTS.html

[7] M. Allman, D. Glover, L. Sanchez, "Enhancing TCP Over Satellite Channels using Standard Mechanism", IETF, RFC 2488, January 1999.

[8] Information Sciences Institute, University of Southern California, "Transmission Control Protocol - Darpa Internet Program - Protocol Specification", IETF, RFC 793, September 1981.

[9] C. Partridge, T. J. Shepard, "TCP/IP Performance over Satellite Links", IEEE Network, September/October 1997, pp. 44-49.

[10] T.V. Lakshman, 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, June 1997

[11] M. Allman, "On the Generation and Use of TCP Acknowledgments", ACM Computer Communication Review, 28(5), October 1998.

[12] R. Durst, G.J. Miller, E.J. Travis, "TCP Extensions for Space Communications", MOBICOM'96, November 1996, New York, USA, pp. 15-26.

[13] Department of Engineering, Aberdeen University, "Monitoring and Optimisation of CODE Protocols", Estec Contract 9980/92/NL/LC, March 1994.

[14] N. Ghani, S. Dixit, "TCP/IP Enhancement for Satellite Networks", IEEE Comm. Mag., July 1999, pp. 64-72.

[15] M. Allman, D. Glover, J. Griner, J. Heidemann, S. Ostermann, K. Scott, J. Semke, J. Touch, D. Tran, "Ongoing TCP Research Related to Satellites", IETF, Internet Draft, draft-ietf-tcpsat-res-issues-05.txt, November 1998.

[16] M. Marchese, "Modifications to TCP's Parameters to Enhance Performance over Satellite Channels", Fifth Ka-band Utilization Conference, Taormina, Italy, October 1999, pp. 343-350.

[17] G Maral, M. Bousquet, "Satellite Communications Systems - Systems, Techniques and Technology, IIIed", John Wiley & Sons, Chichester, 1998.

[18] Y. Zhang, D. De Lucia, B. Ryu, Son K. Dao, "Satellite Communication in the Global Internet", http://www.wins.hrl.com/people/ygz/papers/inet97/index.html.

[19] M. Allman, "Improving TCP Performance Over Satellite Channels", Master Thesis, Ohio University, June 1997.

[20] M. Allman, V. Paxson, W. S. Stevens, "TCP Congestion Control", IETF, RFC 2581, April 1999.

[21] S. Floyd, T. Henderson, "The NewReno Modification to TCP's Fast Recovery Algorithm", IETF, RFC 2582, April 1999.

[22] K. Poduri, K. Nichols, "Simulation Studies of Increased Initial TCP Window Size ", IETF, RFC 2415, September 1998.

[23] M. Allman, S. Floyd, C. Partridge, "Increasing TCP's Initial Window", IETF, RFC 2414, September 1998.