Congestion Control using Multilevel Explicit Congestion Notification in Satellite Networks[1]

Arjan Durresi, Mukundan Sridharan,
Chunlei Liu, Mukul Goyal
Department of Computer and Information Science
The Ohio State University, Columbus, OH 43210, USA
Telephone: (614) 688-5610, Fax: (614) 292-2911
durresi@cis.ohio-state.edu

Raj Jain

Raj Jain is now at
Washington University in Saint Louis
Jain@cse.wustl.edu
http://www.cse.wustl.edu/~jain/

***Abstract*-Exponential growth of Internet traffic and the proliferation of new user applications warrant the development of new Internet infrastructure. Due to the fundamental satellite system characteristics such as global coverage, broadcast nature, and bandwidth on demand, satellite systems are excellent candidates for providing high data rate Internet access and global connectivity accommodating multimedia applications. However, to meet this goal, provisioning of quality-of-service (QoS) within the advanced satellite network systems is the critical requirement. Congestion remains the main obstacle to Quality of Service (QoS) on the Internet. In today's TCP networks, ECN is the only explicit mechanism, which delivers congestion signals to the source.**

In this paper we present a new traffic management scheme based on an enhanced Explicit Congestion Notification (ECN) mechanism. In particular we used multilevel ECN, which conveys more accurate feedback information about the network congestion status than the current ECN scheme. We have designed a TCP source reaction that takes advantage of the extra feedback information and tunes better its response to the congestion than the current schemes. Our analysis and simulations results show that our scheme performed better than the current ECN, having less losses, better network utilization, less delays, and the solution is scalable.

I. INTRODUCTION

In recent years, significant investments have been made in the planning and development of broadband satellite networks. The exponential growth of the Internet provides a good opportunity for satellites to service the increasing new applications, such as Web surfing, desktop and video conferencing. Interest in Ka-band satellite systems has dramatically increased, with over 450 satellite applications filed with the ITU. In the U.S., there are currently 13 Geostationary Satellite Orbit (GSO) civilian Ka-band systems licensed by the Federal Communications Commission (FCC), comprising a total of 73 satellites. Two Non-Geostationary Orbit (NGSO) Ka-band systems, compromising another 351 satellites, have also been licensed. Eleven additional GSO, four NGSO, and one hybrid system Ka-band application for license and 16 Q/V-band applications have been filed with the FCC [1]. The main advantages and disadvantages of GSO versus non-GSO architectures have been discussed in [2].

The delays in GSO systems and delay variations in NGSO systems affect both real-time and non-real-time applications. In an acknowledgement and time-out-based congestion control mechanism, e.g., TCP, performance is inherently related to the delay-bandwidth product of the connection. Moreover, TCP round-trip time (RTT) measurements are sensitive to delay variations, which may cause false timeouts and retransmissions. As a result, the congestion control issues for broadband satellite networks are somewhat different from those of lower-latency terrestrial networks [23-26].

Despite the fact that a number of schemes have been proposed for network congestion control, the search for new schemes continues [4-18]. The research in this area has been going on for at least two decades. There are two reasons for this. First, there are requirements for congestion control schemes that make it difficult to get a satisfactory solution. Second, there are several network policies that affect the design of a congestion scheme. Thus, a scheme developed for one network, traffic pattern, or service requirement may not work on another network, traffic pattern, or service requirements. For example, many of the schemes developed in the past for best-effort data networks will not work satisfactorily for multi-class IP networks.

Recognizing the need for a more direct feedback of congestion information, the Internet Engineering Task Force (IETF) has come up with Explicit Congestion Notification (ECN) method for IP routers [3, 4]. ECN is much more powerful than the simple packet drop indication, used by existing routers and is more suitable for high distance-bandwidth networks. In this study we present some enhancement to ECN based on multilevel ECN and apply them to a satellite network scenario. Our results show that Multilevel ECN (MECN) improves considerably the congestion control on satellite links.

II. SATELLITE – BASED INTERNET ARCHITECTURES

The satellite-based Internet has several architectural options, due to the diverse design of satellite systems. In general a satellite network can serve as part of the Internet backbone, a broadband Internet access or both of them. While the use of satellites in Internet backbone dates back around 25 years, their use in access network is relatively new.

Traditionally satellite networks used satellite that simply broadcast whatever they receive, also known as "bent-pipes" [21]. In Fig. 1 it is shown a satellite network with bent-pipes satellites. The satellites used can be GSO, MEO or LEO. This architecture provides Internet services through gateways on Earth or directly to some users. Because there is no onboard processing (OBP) capability and no possibility to use

[1] This work was supported in part by grants from TRW, Honeywell, OAI, Cleveland, Ohio and NSF CISE grant #9980637

inter satellite links (ISL), the architecture suffers low spectrum efficiency, lack of flexibility and high delays.

Satellite with OBP capabilities can be used to build satellite ISL networks, and to implement sophisticated protocols, able to offer flexible and high quality service. But, on the other hand OBP increases the complexity. Fig. 2 shows a satellite network with OBP and ISL. Onboard processing involves demodulation and demulti-plexing of the received signal. The payload performs decoding and encoding, processing the header information, and routing the data, pointing the antennas, buffering, multiplexing, and retransmitting the data on downlink or inter-satellite link. The major reasons for OBP include separation of the uplink from the downlink, a gain of approximately 3 dB in performance, and provision of resources on demand. The impact of OBP and switching are discussed in [22].

Another architecture uses the satellite links only as downlinks via direct broadcast satellites (DBS), employed for television broadcast. In this architecture the users need only satellite receiver. The reverse link to server is provided by terrestrial links. Fig. 3 illustrates the DBS architecture. The advantage of DBS is the lack of uplink transmitter and suites the Internet traffic asymmetry between server and users.

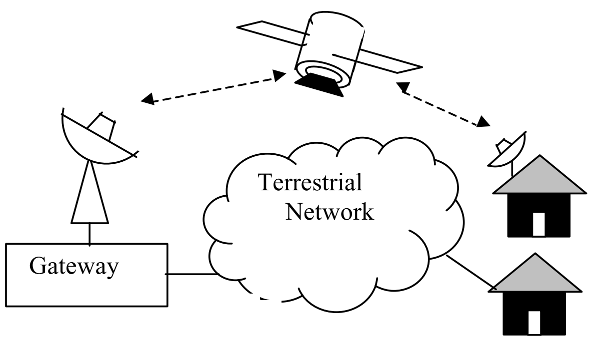

Fig. 1.Bent-pipe architecture.

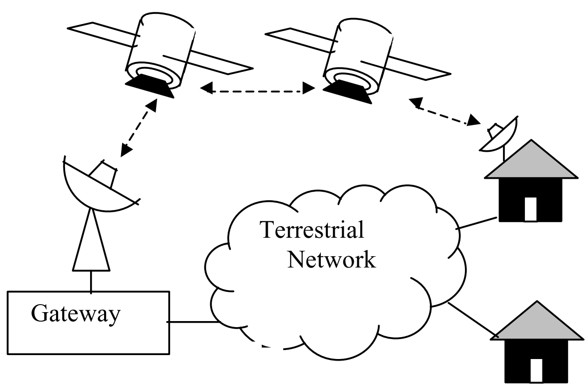

Fig. 2. Inter Satellite Link Architecture with OBP

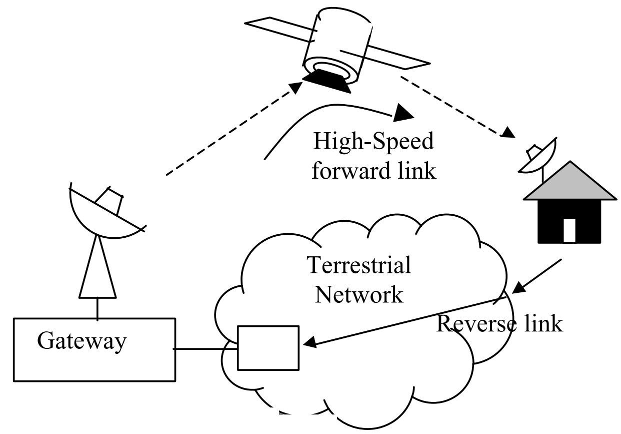

Fig. 3. DBS architecture.

III. END-TO-END CONGESTION CONTROL WITH ECN

In order to manage the traffic, IP routers need to inform the sources about their load levels, so that the sources can increase or decrease their traffic to match the available capacity. In the simplest case, the feedback can consist of dropping packets, as is currently done in the IP routers. The next step is to include explicit feedback in the network layer header. The Internet Engineering Task Force (IETF) has recently introduced, with RFC 2481, the "Explicit Congestion Notification (ECN)." Two bits, CE and ECT in the IP header have been reserved for this purpose. CE bit is used to indicate congestion in the router. Routers start marking this bit using a RED-like algorithm based on the average queue length. ECT bit is used by the sources to indicate whether the flow is ECN capable. The receivers echo the ECN bit back to the source through TCP ACKs. Sources respond to ECN once per round trip time (RTT). To guard against loss of ACKs, receiver continues to set ECN-Echo bit in subsequent ACKs (even if further packets do not have CE bit set) until it receives a packet with Congestion Window Reduced (CWR) bit set. A source, after responding to congestion indication by halving the congestion window (CWND), sets CWR bit in next packet sent (in order to inform receiver about action taken in response to congestion). After receiving a packet with CWR bit set, receiver does not set ECN-Echo bit in ACKs until it gets another packet with CE bit set.

The two major advantages of ECN scheme are: in case of not very high level congestion, the packets are not dropped and second, it could provide a more detailed information about congestion as will be seen in study later. Both these ECN features are very suited to satellite networks, where it is very important not to drop packets that already have consumed critical resources of satellite long links. Satellite links can be considered as feedback control system with long feedback time, so it is very important to have the most accurate information about congestion situation, in order to tune faster the source reaction without waiting for several long RTTs.

IV. MULTILEVEL ECN – MECN

A. *Marking the bits at routers*

The current proposal for ECN uses two bits in the IP header (bits 6 and 7 in the TOS octet in Ipv4, or the Traffic class octet in Ipv6) to indicate congestion. These two bits can be used more efficiently to indicate congestion. With two bits we can indicate 4 different congestion levels. If non ECN-capable packets are identified by the bit combination of '00', we have three other combinations to indicate three levels of congestion. In our scheme the bit combination '01'-indicates no congestion,'10'-indicates incipient congestion and '11'- indicates moderate congestion. Packet drop occurs only if there is severe congestion in the router and buffers overflow. Including packet-drop we can indicate four different levels of congestion. Then appropriate action could be taken by TCP sources, depending on the level of congestion. The four levels of congestion are summarized in Table 1.

The marking of CE, ECT bits is done using a multilevel RED scheme. The RED scheme has been modified to include another threshold called the mid_thresh, in addition to the min_threshold and max_threshold. If the size of the average queue is in between min_th and mid_th, there is incipient congestion and the CE, ECT bits are marked as '10' with a maximum probability of P1max. If the average queue is in between mid_th and max_thresh, there is moderate congestion and the CE, ECT bits are marked as '11' with a maximum probability P2max. If the average queue is above the max_thresh all packets are dropped. The marking policy is shown in Fig. 4.

B. *Feedback from Receiver to Sender*

The receiver reflects the bit marking in the IP header, through TCP ACKs. Since we have three levels of marking instead of 2-level marking in the traditional ECN, we make use of 3 combination of the 2 bits 8, 9 (CWR, ECE) in the reserved field of the TCP header. In the current ECN, the bit combination '00' indicates no congestion and '01' indicates congestion. Again, these 2 bits are just going to reflect the 2 bits in the IP header. The packet drop is recognized using traditional ways, by timeouts or duplicate ACKs.

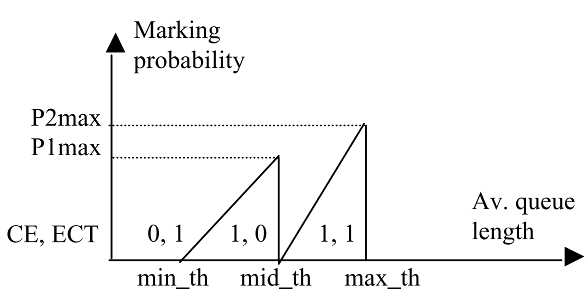

Fig. 4. Marking at routers.

TABLE I
ROUTER RESPONSE TO CONGESTION BY MARKING OF CE AND ECT BITS AND DROPPING PACKETS

CE bit	ECT bit	Congestion State
0	1	No Congestion
1	0	Incipient congestion
1	1	Moderate congestion
Packet drop		Severe congestion

In our scheme, the receiver marks the CWR, ECE bits in ACKs as "01', if the received packet has CE, ECT bits marked by the router as '10'. When a packet with CE, ECT bits marked as '11' is received, the receiver marks CWR, ECE bits in ACKs as '11'. If the received packet has CE. ECT bits marked as '00' or '01', the receiver marks CWR, ECE bits of the ACKs as '00'. The marking of the ACKs CWR and ECE bits is shown in Table 2. In the current ECN standard, the CWR bit has the possibility of being set only in packets from source to the receiver and the receiver stops reflecting the ECN bits if it receives a packet with CWR set. But in our scheme the CWR is used in both directions.

C. *Response of TCP Source*

We believe that ECN marking should not be treated the same way as packet drop, since ECN indicates just the beginning of congestion and the buffers still have space. For this reason, using multiple levels of congestion feedback, the TCP's response needed to be refined as follows.

When there is a packet-drop, the 'cwnd' is reduced by α_3 = 50%. This done for two reasons: First, a packet-drop means severe congestion and buffer overflow, so some severe actions need to be taken. Second, to maintain backward compatibility with routers, which don't implement ECN.

When there is no congestion, the 'cwnd' is allowed to grow additively as usual. When the marking is '10'(incipient congestion), 'cwnd' is decreased by α_1 = 20 %. When the marking is '11'(moderate congestion) the 'cwnd' is decreased multiplicatively by a factor α_2 % less than 50% but more than α_1. In Table 3 there are shown the TCP source responses and the values of parameters α_x we have implemented. In future work we will study the influence of parameter values α_x.

V. SIMULATION RESULTS

In order to compare current ECN with our multi-level ECN scheme, we carried out a set of simulations using the ns simulator [27]. The RED queue in the ns has been modified to include the mid_thresh, in addition to the min_threshold and max_threshold. The marking policy is shown in Fig. 4 and is explained earlier.

TABLE II
RECEIVER MARKING OF CWR AND ECE BITS AND DROPPING PACKETS

CWR bit	ECE bit	Congestion
0	0	No Congestion
0	1	Incipient congestion
1	1	Moderate congestion
Packet Drop		Severe congestion

TABLE III
TCP SOURCE RESPONCE

Congestion State	cwnd change
No congestion	Increase ‘cwnd’ additively
Incipient congestion	Decrease multiplicatively by $\alpha_1 = 20\%$
Moderate congestion	Decrease multiplicatively by $\alpha_2 = 40\%$
Severe congestion	Decrease multiplicatively by $\alpha_3 = 50\%$

The TCP in the ns simulator is also modified according to our algorithm. The receiver reflects the markings in the IP header, in the experimental field of the TCP header. The sender reduces its congestion window by 20% if it gets a mild congestion marking and reduces the window by 40%, if it gets a heavy congestion marking. If there is any timeout or duplicate acks (packet loss) the TCP reduces the window by 50%. When the TCP sender sends the congestion window reduced (CWR) signal, the receiver stops echoing the level of ECN, which it marked first. For example, suppose if there is congestion in a given router, which starts marking packets in the next level. The receiver gets packets and starts echoing that particular level of ECN in all ACKs. If the congestion makes into next level, before the receiver gets a congestion window reduced (CWR) signal, the receiver remember, which level was marked first and stops echoing that level and starts echoing the next level of ECN. The connection establishment phase and the ECN negotiation are not modified.

For simplicity, the max Probability of dropping, for both levels of ECN are kept the same, P1max = P2max. Also for the same reason, we have applied for MECN max_th = 2 mid_th, and mid_th = 2 min_th and for simple ECN max_th = 2 min_th. The aim of the simulation is not to fix the best parameter of the RED queue, but to illustrate the advantage of multi-level ECN. Further study is needed to optimize these parameters.

A Simulation Configuration

For all our simulations, we used the following configuration, shown in Fig. 5. A Number of sources S1, S2, S3.., Sn are connected to a router R1 through 10Mbps, 2ms delay links. Router R1 is connected to R2 through a 1.5Mbps, 65ms delay link. R2 is connected to R3 through a 1.5Mbps, 65ms delay link and a number of destinations D1, D2, D3.., Dn are connected to the router R3 via 10Mbps 4ms delay links. The link speeds are chosen so that the congestion will happen only between routers R1 and R2 where our scheme is tested. This configuration can simulate the case of MEO and LEO satellite networks.

An FTP application runs on each source. Reno TCP is used as the transport agent. (The modifications were made to the Reno TCP). The packet size is 1000 bytes and the acknowledgement size is 40 bytes.

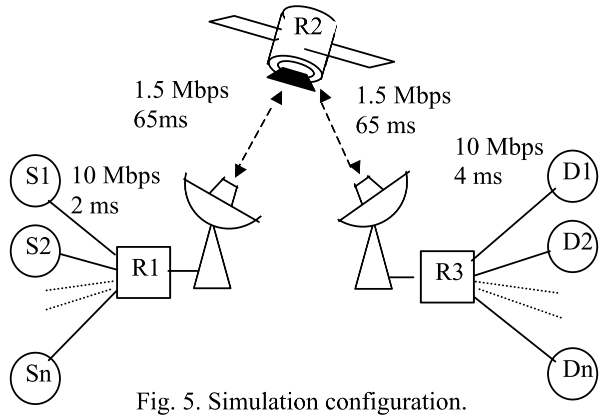

Fig. 5. Simulation configuration.

B Simulation Scenarios

With the basic configuration described above, the following simulation scenarios were used to test our scheme.

1. Ten overlapping connections with same RTT, each connection starting 0.3 seconds after the previous one. The RTT for all the connections are fixed at 272 ms
2. Ten connections with different RTT. The minimum RTT is 272 ms and for each connection the RTT increases by 10 ms, so RTTs are 272, 282,…up to 362 ms.

In Fig. 6a there are shown the instantaneous and average queue lengths, in number of packets, for the case of ten sources with different RTT, applying simple ECN. In Fig. 6b there are the results for the same configuration, applying our scheme MECN. It is clear that in case of MECN, the queue length converges faster and with less oscillations compared to the simple ECN case.

Then we compared the link efficiencies versus min_threshold, in number of packets, obtained with simple ECN and MECN. Fig. 7 shows the results for the configuration with ten sources with the different RTT. As shown, in both experiments the use of MECN not only has improved the link efficiency for all values of min_threshold, but the best link efficiency in case of MECN is reached for a lower value of min_threshold, that means with less delays introduced to the packets.

In Fig. 8a we compared the losses, in number of packets, between simple ECN and MECN, for the configuration with ten overlapping sources with different RTTs. As shown, in case of MECN there are less losses than with simple ECN. Furthermore, MECN reaches the point of zero loss with less min_threshold than ECN. So there two important advantages of MECN compared ECN: MECN enables less losses and reaches the point of zero losses with much less min_threshold, that means with mach less delay introduced to packets. In Fig. 8b there are compared the losses between ECN and MECN for the configuration with ten sources with different RTT. Again MECN performs much better than simple ECN. Again MECN has less losses and reaches the point of zero losses with much less min_threshold, that means with much less delay introduced to the traffic.

All our simulation experiments show that Multi-level ECN performs better than simple ECN as congestion control scheme for TCP. These results are explained by the fact that in case of MECN, the feedback control system uses more accurate feedback information and consequently is able to react better to the congestion. We plan to study in the future the influence of different parameters involved, such as min, mid, max thresholds, P1max, P2max, α_x. Also we plan to simulate more complex configurations to study better the advantages of MECN.

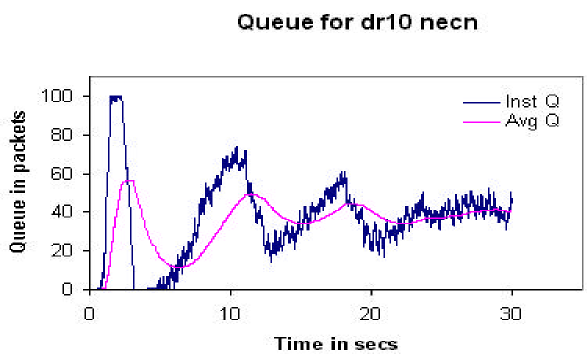

Fig. 6a. Queue length with simple ECN.

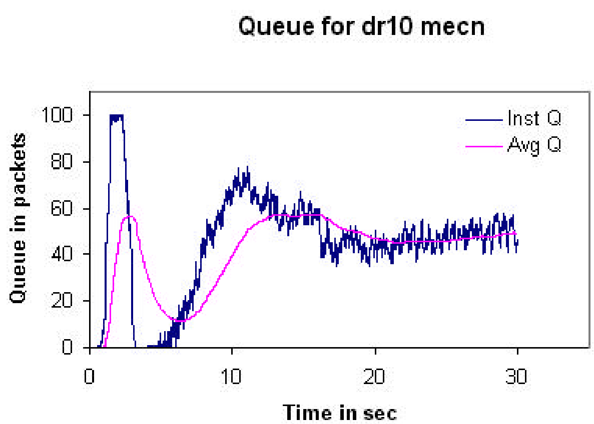

Fig. 6b. Queue length with simple MECN.

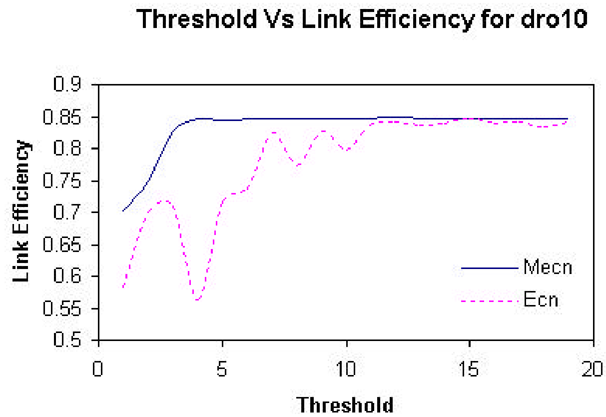

Fig. 7. Link efficiency with ten overlapping sources of different RTTs

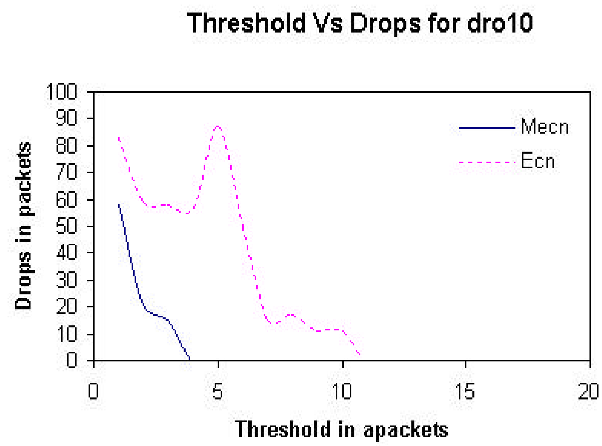

Fig. 8a. Packet drops with ten overlapping sources of different RTTs

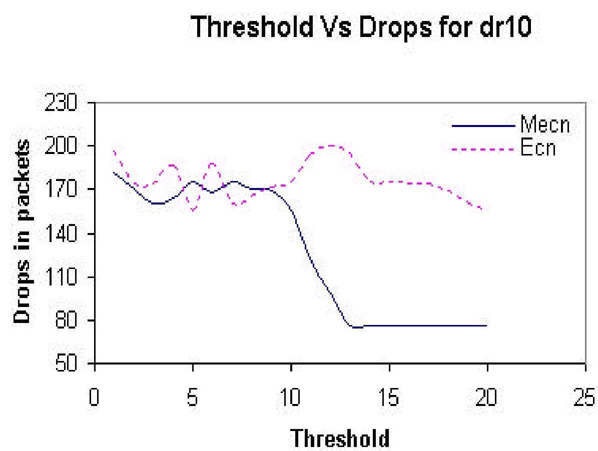

Fig. 8b. Packet drops with ten sources of different RTTs

V. Conclusions

In this study we used Multi-level ECN (MECN), a new and enhanced ECN congestion control scheme, in a satellite network scenario. With MECN, routers are able to send more accurate feedback information about the congestion to TCP source through destinations. More information about the congestion enables TCP sources to have a better tuned response to the congestion. Consequently the MECN congestion control scheme converges faster, with less losses than simple ECN and improves other important QoS parameters such as link utilization and delay. MECN uses of the same bits in IP and TCP header already used by ECN, so it is compatible with the accepted standards.

All simulation results in the satellite network show that MECN improves the QoS parameters such as throughput, link utilization, delay, losses, and queue oscillation compared to ECN scheme. These improvements are very important especially in case of more expensive satellite links compared to terrestrial ones. We believe that MECN is a step forward in the right direction to deal with Internet congestion in general, including satellite networks. We will analyze in future the values of the parameters involved and will optimize the MECN scheme to different types of satellite networks.

References

[1] Sastri Kota, "Multimedia Satellite Networks: Issues and Challenges," *Proc. SPIE International Symposium on Voice, Video, and Data Communications, Boston*, Nov 1-5, 1998.

[2] Arjan Durresi, Sastri Kota, Raj Jain, Mukul Goyal, "Achieving QoS for TCP Traffic in Satellite Networks with Differentiated Services", Accepted for publication in *Space Communications Journal*.

[3] K. Ramakrishnan and S. Floyd, ``A Proposal to add Explicit Congestion Notification (ECN) to IP," *RFC 2481*, January 1999

[4] Sally Floyd, ``TCP and Explicit Congestion Notification," *Computer Communications Review*, Vol. 24, No. 5, October 1994, pp. 10-23.

[5] Sally Floyd and Van Jacobson, ``Random Early Detection gateways for Congestion Avoidance," *IEEE/ACM Transactions on Networking*, Vol. 1, No. 4, August 1993, pp. 397-413.

[6] Sally Floyd and Van Jacobson, ``On Traffic Phase Effects in Packet-Switched Gateways," *Internetworking: Research and Experience*, Vol. 3, No. 3, September 1992, pp. 115-156.

[7] Sally Floyd, ``Connections with Multiple Congested Gateways in Packet-Switched Networks Part 1: One-way Traffic," *Computer Communications Review*, Vol. 21, No. 5, October 1991, pp. 30-47.

[8] David D. Clark and Wenjia Fang, ``Explicit allocation of best-effort packet delivery service," *IEEE/ACM Trans. on Networking*, Vol. 6, No. 4, August 1998, pp. 362-373.

[9] W. Feng, D. Kandlur, D. Saha and K. Shin, ``Blue: A New Class of Active Queue Management Algorithms," University of Michigan, *Technical Report UM-CSE-TR-387-99*, 1999.

[10] S. Kalyanaraman, R. Jain, S. Fahmy R. Goyal and B. Vandalore, ``The ERICA Switch Algorithm for ABR Traffic Management in ATM Networks," Submitted to *IEEE/ACM Trans. on Networking*, November 1997

[11] Sally Floyd and Kevin Fall, ``Promoting the Use of End-to-End Congestion Control in the Internet," *IEEE/ACM Trans. on Networking*, August 1999.

[12] Sally Floyd and Kevin Fall, ``Router Mechanisms to Support End-to-End Congestion Control," *Unpublished Manuscript*, http://www-nrg.ee.lbl.gov/floyd/papers.html

[13] M. Mathis, J. Semke, J. Mahdavi and T. Ott, ``The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm," *Computer Communications Review*, Vol. 27, No. 3, July 1997, pp. 67-82.

[14] D. Chiu and R. Jain, ``Analysis of the Increase/Decrease Algorithms for Congestion Avoidance in Computer Networks," Journal of Computer Networks and ISDN Systems, Vol. 17, No. 1, June 1989, pp. 1-14.

[15] Sally Floyd and Tom Henderson, ``The NewReno Modification to TCP's Fast Recovery Algorithm,", *Internet Draft - Work in Progress*, Feburary 1999.

[16] M. Mathis, J. Mahdavi, S. Floyd and A. Romanow, ``TCP Selective Acknowledgment Options, " *RFC 2018*, October 1996.

[17] K. K. Ramakrishnan and R. Jain, ``A binary feedback scheme for congestion avoidance in computer networks," *ACM Transactions on Computer Systems*, Vol. 8, No. 2, May 1990, pp. 158-181.

[18] R. Jain, S. Kalyanaraman, R. Viswanathan, ``Rate Based Schemes: Mistakes to Avoid." *AF-TM 94-0882*, September 1994.

[19] R. Jain, S. Kalyanaraman, and R. Viswanathan, ``Ordered BECN: Why we need a timestamp or sequence number in the RM Cell," *ATM Forum/94-0987*, October 1994.

[20] W. Stevens, ``TCP Slow Start, Congestion Avoidance, Fast Retransmit, and Fast Recovery Algorithms," *RFC 2001*, January 1997.

[21] Yurong H, Victor O. K. Li, "Satellite-based Internet: A tutorial*," IEEE Communications Magazine*, pp.154-162, Vol. 39, No 3, March 2001.

[22] H. J. Morgan et. al., "Throughput Analysis for Satellite Network Architectures*," Proc. of AIAA 18th International Communication Satellite Systems Conference*, Oakland, 2000, pp. 602-612.

[23] Mark Allman, Chris Hayes, Hans Kruse, Shawn Osterman, "TCP Performance over satellite links," *5th International Conference on Telecommunication Systems*, 1997

[24] Thomas R. Henderson, Randy H. Katz, "TCP performance over satellite channels*," UCB Computer Science Technical Report 99-1083*, December 1999.

[25] Thomas R. Henderson, "Networking over next-generation satellite systems," *Ph.D. dissertation, University of California at Berkeley*, 1999

[26] Nasir Ghani, Sudhir Dixit, "TCP/IP Enhancements for satellite networks," *IEEE Communications Magazine*, pp.64-72, Vol. 37, No 7, July 1999

[27] NS Simulator, Available from http://www-mash.cs.berkeley.edu/ns