IDEA GROUP PUBLISHING
701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com

ITB12853

This paper appears in the book, *Emerging Trends and Challenges in Information Technology Management, Volume 1 and Volume 2* edited by Mehdi Khosrow-Pour © 2006, Idea Group Inc.

Improving TCP/IP Performance Over Geosynchronous Satellite Links: A Comparative Analysis

Joseph M. Lawson & Michael R. Grimaila, Air Force Institute of Technology, Wright-Patterson AFB, OH 45433-7765,
P 937-255-3636 x4800, F 937-656-4699Joseph.Lawson@AFIT.EDU, Michael.Grimaila@AFIT.EDU

ABSTRACT

The need for global information interchange in military applications has dramatically increased the need for geosynchronous (GEO) satellite data links beyond their existing capacity. In this paper, we identify shortcomings inherent in the configuration and utilization of TCP protocol over GEO satellite data links that prevent efficient utilization of the data link bandwidth. To date, both military and commercial organizations have relied upon use of Performance Enhancing Proxies (PEPs) to maximize bandwidth usage. However, new capabilities in existing routing equipment software may provide the capability to attain similar improvements in transmission efficiency with no additional cost. Our work focuses on the modeling and simulation of PEPs versus acceleration using existing end point routers in order to compare their effectiveness. We discuss our work to date and the importance of taking advantage of the new capabilities of end point routers to reduce costs and improve mission effectiveness.

1. INTRODUCTION

The advent of the Internet has revolutionized the way information is shared between geographically separated entities. From E-business sites, such as Amazon.com, to United States military Command and Control systems, the Internet has become the medium of choice for the push and pull of data and information. Although this new reality seemingly simplifies network topology and design, the characteristics of the underlying transport protocols of the Internet need to be considered when used in tandem with a satellite communications architecture.

The types of orbits utilized by constellations of satellite systems orbiting the earth include low earth orbit (LEO), highly elliptical orbit (HEO), and geosynchronous orbit (GEO). GEO satellites are the most commonly utilized communications asset for deployed US military forces and are the focus of this paper. GEO satellites orbit at an elevation of approximately 36,000 kilometers above the surface of the earth (Roddy, 2001). At this elevation over the equator, satellites are able to achieve a speed that matches the rotational velocity of earth. This condition enables GEO satellites to remain stationary relative to a location on the equator, known as a sub-satellite point. Due to a GEO satellite's high elevation and relatively stationary position, it is able to attain a coverage area of approximately ±75° latitude. Three geostationary satellites, spaced apart by 120° longitude along the equator, could provide whole earth coverage. For these reasons, GEO satellites have become a high demand asset during US military operations in areas with little or no terrestrial telecommunications infrastructure.

As a result of a GEO satellite's high elevation, lengthy propagation times occur for radio signals traveling the slant distance between an earth station and the satellite. Propagation delays can vary from 239.6 to 279 milliseconds (ms) for one "hop" (ground station to satellite to ground station), depending on the location of ground stations in a satellite's spot beam. From these figures it can be calculated that the Round Trip Time (RTT) for a message and a corresponding response between two ground stations communicating via a GEO satellite could take between 479.2 and 588 ms (Allman, 1999). It will be shown later that this characteristic negatively impacts the throughput capability of a TCP-based connection.

2. TCP/IP OVERVIEW

IP represents the de facto standard in the logical mapping of devices connected via the Internet. The basic characteristics of IP will not be discussed here as they are well known throughout the community. TCP also utilizes four algorithms to mitigate network congestion that include Slow Start, Congestion Avoidance, Fast Restransmit, and Fast Recovery (Kamm, 2000). These algorithms are used to detect network congestion and to reduce the transmission rate of a network device to a level that can be supported by the available network resources.

2.1 TCP Congestion Control Algorithms

Congestion control algorithms utilize two state variables, Congestion Window (CWND) and Slow Start Threshold (SSTHRESH). CWND represents the amount of segments a device can inject into a network before receiving an ACK. CWND will also be limited to the size of the advertised window of a receiver. The CWND value can be increased or decreased based on the perceived amount of network congestion. SSTHRESH is used to determine what algorithm will be used to increase CWND. If CWND is less than SSTHRESH, the Slow Start algorithm is used. If CWND is greater than or equal to SSTHRESH, the congestion avoidance algorithm is used (Allman, 1999). The initial SSTHRESH value will be the receiver advertised window size. The SSTHRESH value can also be set at the level in which network congestion was detected (Allman, 1999). While these algorithms are useful in preventing the congestive failure of a network, they also have a negative impact on the performance of TCP over a large delay or high bit-error rate network commonly encountered in a satellite-based transmission environment (Allman, 1997).

Slow Start is the mechanism TCP utilizes to establish a network connection or restart a connection after a RTO has occurred (Kamm, 2000). The purpose of Slow Start is to ensure a network device does not transmit too large of a burst of data segments that could overwhelm the available network resources. The Slow Start algorithm begins by initially setting the value for CWND to one segment and SSTHRESH to the receiver's advertised window size. This limits the network device to the transmission of one segment and to wait for an ACK of receipt of that segment (Allman, 1999). For each ACK the transmitting device receives, the CWND is increased by one segment (Kamm, 2000). For example, after receipt of the initial ACK, the transmitting device will be able to send two segments (CWND equals two). After receipt of an ACK for each of the two segments, the sender will be able to send four segments (CWND equals four). This behavior represents an exponential growth pattern and will continue until CWND equals or exceeds the value of SSTHRESH or a segment loss is detected. It is important to note that if the RTO expires for a transmitted segment, TCP will initiate the retransmission of the segment and will perceive the RTO expiration as a sign of network congestion. As a response to the perceived network congestion, TCP will reduce the transmission rate of a device by cutting the SSTHRESH value to half of the current CWND value and reset the CWND value to one, starting the Slow Start mechanism anew (Allman, 1997). When CWND equals or exceeds SSTHRESH, however, the Congestion Avoidance algorithm is used to further increase the size of CWND (Allman, 1999).

The Congestion Avoidance algorithm is a more conservative measure to increase CWND than Slow Start and is primarily used to slowly probe the network for additional capacity (Allman, 1999). Congestion Avoidance only allows an increase in value of CWND if all segments transmitted in a window have a corresponding ACK, making the growth rate of CWND linear in nature (Broyles, 1999). Mathematically, the congestion window will be increased by 1/CWND for each segment that is ACKed during use of the Congestion Avoidance algorithm. From this property, CWND will be increased by roughly one segment for every round-trip time (Allman, 1997).

The Fast Retransmit and Fast Recovery Algorithms work hand in hand to limit the time it takes a sender to detect a dropped segment (Kamm, 2000). Under normal circumstances during a TCP connection, segments are assumed lost and are retransmitted when RTO occurs. Unfortunately, needless retransmissions of segments can occur despite successful transmission of a segment because the corresponding ACK is still traveling through the network or the segment awaits processing in a receiver's buffer when RTO expires (Broyles, 1999). To counteract this, Fast Retransmit provides a way to retransmit a packet prior to RTO occurring by assuming that three duplicate ACKs corresponds with a lost segment (Allman, 1997).

When a segment is retransmitted by the Fast Retransmit protocol, the Fast Recovery Algorithm is activated in response to a perceived congestion of the network. The Fast Recovery Algorithm reduces the CWND to half the current value and resets the SSTHRESH value to the new value of CWND (Allman, 1997). The CWND is then artificially increased to match the number of duplicate ACKs, under the assumption that duplicate ACKs indicate a lost segment that is no longer in the network and therefore additional network capacity exists (Broyles, 1999). Depending on the size of CWND, additional segments may be transmitted, however, the receipt of a non-duplicate ACK will reduce CWND to the value of SSTHRESH and will cause the start of the Congestion Avoidance algorithm (Allman, 1997).

2.2 TCP Limitations in a Satellite Environment

TCP's limitations in a satellite environment can be generalized into two areas: increasing the sliding window size and the limitations of TCP throughput in a high delay environment. The Slow Start and Congestion Avoidance algorithms, as outlined previously, determine the rate of size increase or decrease of the sliding window. The steady-state behavior of TCP determines the maximum throughput capability based on factors of window size and RTT.

The time the Slow Start algorithm takes to reach a window size of W segments on a network with a RTT of R can be calculated as shown in Equation 1 (Allman, 1997):

$$\text{SStime} = R^*\log_2(W) \quad (1)$$

Assuming a GEO satellite link with a RTT of 570 ms, 256 byte segments, and a maximum advertised window size (64 Kilobytes that would result in 256 segments), it would take 4.56 seconds to increase CWND to the advertised window size. In contrast, a terrestrial link with the same characteristics, but a RTT of 70 ms, would take 560 ms. This illustrates the potential waste in available bandwidth by TCP in a high delay environment.

As mentioned previously, Congestion Avoidance is utilized by TCP to slowly probe a network for additional capacity. The linear nature in which Congestion Avoidance increases CWND takes an exorbitant amount of time in a high loss or large delay transmission system. For example, if a loss (such as the loss of a segment) occurs on a network, the value of CWND is reduced to half of the original value. If 256 byte segments and a maximum window size were in use prior to the loss, the resulting new value for CWND would be 32 Kilobytes. Under these conditions, Congestion Avoidance would take 72.96 seconds to return CWND to the maximum window size. In contrast, a terrestrial link with similar properties, but a RTT of 80 ms, would take only 10.24 seconds to reach the maximum window size under Congestion Avoidance. This discrepancy in recovery time again highlights the potential waste of bandwidth utilizing TCP over high delay paths.

TCP also experiences a bottleneck in throughput capability in high delay transmission systems. Assuming a loss- and congestion-free network, the maximum throughput of TCP can be calculated as shown in Equation 2 (Allman, 1997):

$$\text{MT} = \text{Receive Window Size/RTT} \quad (2)$$

With a maximum window size of 64 Kilobytes and a satellite link with a RTT of 570 ms, the maximum throughput of a TCP connection would approximately be 114,977 bytes per second. If the satellite link operated at a T1 (1536 Mbits/second with removal of 8 Mbits for overhead) rate on a transponder, the link would only be 60% utilized (114,977 bytes/192,000 bytes). TCP would again fail to fully utilize the available bandwidth of a network.

3. IMPROVING TCP PERFORMANCE OVER SATELLITE LINKS

There are multiple strategies to enhance the performance of TCP over satellite links. The window scale option is one strategy in which a modification is added to the TCP header. During the initial synchronization of a TCP session, "SYN" segments are transmitted between sender and receiver. The SYN segment purpose is to establish the initial parameters (window sizes, segment lengths, etc.) for a TCP session. Adding the window scale option to the SYN segment would allow a sender to ascertain the available buffer size for a receiver. Based on the buffer size available and "agreement" between sender and receiver, a scale factor can be added to the TCP session parameters that would increase the maximum window size up to twice the value found in a standard TCP session (Mathis, 1996). This relatively simple solution would allow increased bandwidth utilization for high delay paths based on the maximum throughput equation discussed in the previous section. Fortunately, most routing equipment now incorporates the capability of using windows scaling, but requires that it must be manually enabled (Cisco, 2005).

Another option to improve TCP performance would be the use of selective acknowledgements (SACKs). SACKs allow receivers in a TCP session to communicate the success of transmission of every received segment. By bypassing the cumulative ACK system of standard TCP, senders would no longer need to wait for multiple RTT durations to determine what segments have been lost and will instead be able to retransmit specific lost segments in one or two RTT cycles (Jacobson, 1992 and Mathis, 1996). The SACK strategy mitigates the occurrence of RTO and results in the avoidance of activating TCP congestion control mechanisms that hinder throughput (Stewart, 2000). Once again, many router manufacturers have added the capability of using SACKs, but it must be manually enabled.

It should be noted that there are network devices can also be added to satellite transmission links that significantly enhance the performance of router networks utilizing TCP. These devices are commonly referred to as packet accelerators, or Performance Enhancing Proxies (PEPs), and utilize protocols such the Space Communication Protocol Standard-Transport Protocol (SCPS-TP) and Xpress Transport Protocol (XTP) as a replacement for TCP in long delay, high bit error, and asymmetric bandwidth conditions. Both SCPS-TP and XTP enabled devices use both dynamic window sizing and SACKs in an effort to improve network throughput. These devices require no modification of the TCP protocol and instead encapsulate TCP segments with their own specialized protocols for transmission.

4. MILITARY CONSIDERATIONS

Leasing bandwidth on commercial geostationary (GEO) satellites is a costly endeavor, with prices ranging up to $1.5M for an 8 Mbit/sec allocation per year. Military supported constellations, particularly those spacecraft that are part of the Defense Satellite Communications System (DSCS) and Wideband Gapfiller programs (which currently or will support a majority of the United States military's X-band communication requirements), require significant Department of Defense (DoD) funding to maintain required operational and maintenance levels. As a result, military communication support entities, as a consumer of both commercial and military satellite systems, are continually seeking ways to efficiently utilize space-based communication bandwidth allocations in order to maximize return on investment.

PEPs have a relatively short history of use in US military operations. The first large scale employment of these devices by the DoD occurred during the initial stages of Operation IRAQI FREEDOM at the Landstuhl Standardized Tactical Entry Point (STEP) in Landstuhl, Germany. STEP sites represent communication elements that provide deployed combat units access (typically via satellite or terrestrial mediums) to the Defense Information Systems Network (DISN) and associated services (secure/non-secure Internet circuits, video teleconference circuits, messaging systems, secure/non-secure voice circuits, etc.). Though the initial planning and installation phases of the supplied PEPs were somewhat problematic, the subsequent performance benefits of these devices provided a great bit of incentive by the STEP program office for the procurement of additional PEPs for installation at all 18 active STEP sites.

Procurement and installation of PEPs at STEP sites worldwide was rolled into the ongoing Enhanced STEP upgrade program. Though PEPs were viewed to be effective in mitigating TCP link degradations associated with satellite communications, the additional burden of support (training, maintenance, and operation) and the cost of acquisition for these devices proved to be a significant undertaking. The average cost of a PEP procured by the STEP program was $3,000, with training (train the trainer program) costs varying between $10,000 and $15,000 per person. Taking in consideration that 18 STEP sites were to be upgraded, with a minimum of 32 devices, as well as the training of at least 18 personnel, the cost for the PEP implementation of this magnitude could easily exceed $2M. In addition, STEP sites were also expected to allocate a large amount of manpower in the reorganization of their sites to accommodate the rack space and power requirements of these devices. Some STEP sites, such as Ramstein STEP (located at Ramstein Air Force Base, Germany) had severe site limitation both with power availability and physical space available. In addition, STEP sites also needed to reevaluate their respective facility back-up power mechanisms; assessments would provide insight as to whether or not facility back-ups could have handled the additional load presented by the PEPs in the event of a power outage.

Due to the acquisition, installation, and training costs for PEPs it is not surprising that many organizations are researching techniques to incorporate in commonly employed technologies (routers, satellite modems, etc.) to combat link performance issues and negate the need for another device to acquire, maintain, and operate. Of particular interest is the use of selective acknowledgements (SACKs) and increases in the TCP window size, that have been incorporated in already released Cisco router Internetwork Operating System versions. Since Cisco products represent the largest population of network devices used by military communication units, it is reasonable to assume that a significant cost savings (in terms of dollars and physical requirements) could be realized if SACKs and window scale techniques rival the performance benefits of a PEP system.

5. MODELING AND SIMULATION

Our current research effort focuses on the modeling, simulation, and comparison of both PEPs and existing Cisco routers using OPNET V10.5 Modeler, an environment for network modeling and simulation, running on a standard personal computer (OPNET, 2005). OPNET provides the capability for extensive, accurate modeling of the environment being simulated. For this reason, care is being taken to model the environment accurately by examining existing STEP sites and their network bandwidth utilization. The results we originally proposed to be included in this paper were too simplified and thus flawed. For this reason, we have chosen not to include our intermediate results and instead refocused our paper as a research in progress paper.

Our experimentation plan is to model a TCP session over satellite link with Cisco routers at each end of the link. We will conduct multiple simulation runs were completed using different router configuration parameters to show a comparison between unmodified, or standard TCP protocol implementation, and TCP with enhancements (use of SACKs, larger initial windows, and both) using models of existing routers in use by the Air Force. In addition, we will model environmental/transmission system factors, such as variable bit error rates, link/propagation delays, asymmetric bandwidth limitations, and traffic loads (individual client sessions to high density (multi-client) sessions.

6. CONCLUSIONS

Leasing bandwidth on commercial GEO satellites is a costly endeavor so it is essential to maximize the efficiency of satellite communications. Unfortunately, a majority of these units have not been given the capability (packet accelerators) or the knowledge (TCP modifications) to fully exploit use of the bandwidth they were allocated. Work is currently underway to model and simulate the effectiveness of using two simple modifications on existing Cisco end point routers to eliminate the need to purchase dedicated packet accelerators for satellite links. If the DoD is to maximize satellite program investments and information sharing of combat units, recognition and use of advanced capabilities of existing equipment needs to be exploited instead of needlessly purchasing commercially available equipment that yields the same performance improvements in satellite links.

7. DISCLAIMER

The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

8. REFERENCES

Allman, et. al. "Request for Comments 2488: Enhancing TCP Over Satellite Channels Using Standard Mechanisms." The Internet Society. 1999 p. 1-19.

Allman, Mark. "Improving TCP Performance Over Satellite Channels." Ohio University. 1997 p. 1-64.

Broyles, Ren H. "Performance Analysis of TCP Enhancements in Satellite Data Networks." Masters Thesis, Air Force Institute of Technology. 1999.

Cisco Systems, Inc., "TCP Windows Scaling," Available at http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122newft/122t/122t8/tcpwslfn.pdf, 2005.

V. Jacobson, R. Braden, and D. Borman, "RFC 1323: TCP Extensions for High Performance," International Engineering Task Force Request for Comments, May 1992, Available at www.faqs.org/rfcs/rfc1323.html.

Kammermann, Maik and Hugeus Latour. "Enhancing TCP Performance Over Satellite Channels." Defense Research & Development Canada. Ottawa, Canada: DREO TR 2000-079, November 2000 p. 1-37.

Lammle, Todd. CCNA: Cisco Certified Network Associate Study Guide. San Francisco: Sybex Inc., 2000.

Mathis, M., Mahdavi, J. , Floyd, S., and Romanow, A., "RFC 2018: TCP Selective Acknowledgment Options," International Engineering Task Force Request for Comments, October 1996, Available at www.faqs.org/rfcs/rfc2018.html.

OPNET Technologies, Inc., OPNET Network Simulator, Available at http://www.opnet.com, 2005.

Stewart, R., Xie, Q, Morneault, K., Sharp, C., Schwarzbauer, H., Taylor, T., Rytina, I., Kalla, M., Zhang, L. and Paxson, V., "RFC 2960 Stream Control Transmission Protocol," International Engineering Task Force Request for Comments, October 2000, Available at www.faqs.org/rfcs/rfc2960.html.

Roddy, Dennis. Satellite Communications: Third Edition. New York: McGraw-Hill Companies, Inc., 2001.