Enhancing End-to-End Performance of Information Services Over Ka-Band Global Satellite Networks

Kul B. Bhasin, Daniel R. Glover, William D. Ivancic, and Thomas C. vonDeak
Lewis Research Center, Cleveland, Ohio

Prepared for the
Third Ka Band Utilization Conference
sponsored by the IIC-Istituto Internazionale delle Comunicazioni
Sorrento, Italy, September 15–18, 1997

National Aeronautics and Space Administration
Lewis Research Center

December 1997
ENHANCING END-TO-END PERFORMANCE OF INFORMATION SERVICES OVER KA-BAND GLOBAL SATELLITE NETWORKS

Kul B. Bhasin, Daniel R. Glover, William D. Ivancic, and Thomas C. von Deak
National Aeronautics and Space Administration
Lewis Research Center
21000 Brookpark Road
Cleveland, Ohio 44135 U.S.A
(216) 433-3676, (216) 433-2847, (216) 433-3494, (216) 433-3277
(216) 433-8705 (FAX)
bhasin@lerc.nasa.gov, dglover@lerc.nasa.gov, wivancic@lerc.nasa.gov, tvondeak@lerc.nasa.gov

1. ABSTRACT
The Internet has been growing at a rapid rate as the key medium to provide information services such as e-mail, WWW and multimedia etc., however its global reach is limited. Ka-band communication satellite networks are being developed to increase the accessibility of information services via the Internet at global scale. There is need to assess satellite networks in their ability to provide these services and interconnect seamlessly with existing and proposed terrestrial telecommunication networks.

In this paper the significant issues and requirements in providing end-to-end high performance for the delivery of information services over satellite networks based on various layers in the OSI reference model are identified. Key experiments have been performed to evaluate the performance of digital video and Internet over satellite-like testbeds. The results of the early developments in ATM and TCP protocols over satellite networks are summarized.

2. INTRODUCTION
NASA has been at the forefront of telecommunications technology and Ka-Band experimentation through the Advanced Communications Technology Satellite Program (ACTS) conducted at the NASA Lewis Research Center in Cleveland, Ohio. In addition to this activity NASA has now partnered with the satellite industry to address the issues concerning seamless interoperability between satellite and terrestrial networks in the emerging GII. Interoperability becomes an issue whenever the standards used on one side of a network interface are incompatible with network conditions on the other side of the interface or become significantly inefficient in their operation. This is the case with many aspects of the ATM and Internet protocols. This paper outlines the issues that are currently recognized and being addressed by NASA with its industry partners.

The depiction in Figure 1 outlines the complexity and scope of the issues between satellite and terrestrial communications in a top level format. Some of the key features that come to light are that:

![Diagram of network topologies](image-url)

FIGURE 1 - Top Level Reference Model
1. Low-Earth-Orbit, Middle-Earth-Orbit, and Geosynchronous-Earth-Orbit (LEO/MEO/GEO) satellites will provide a wide range of data rates to a variety of implementation scenarios ranging from low data rate mobile users served by a LEO satellite constellation to SONET supporting fixed earth stations served by a single GEO satellite.

2. The LEO/MEO/GEO satellite configurations may directly interconnect with one another utilizing intelligent on-board switching electronics.

3. The various implementation scenarios will have to seamlessly interconnect with each other and the terrestrial networks in order to achieve the goals of the GII.

The methodology of addressing the large scope of interoperability has been to develop reference models and identify the key interoperability issues at each of the layers and its respective interfaces (Figure 2). In recognition of the two driving protocols of the GII, ATM and TCP/IP, the protocol stack models are ATM, TCP/IP over ATM, and TCP/IP residing directly on the physical layer. The remainder of the paper is therefore divided into ATM and TCP/IP issues.

3. ATM ISSUES

Asynchronous Transfer Mode (ATM) is a protocol originally conceived by the telecommunication industry to handle multimedia traffic over wide area networks (WANs). The protocol is a connection oriented cell switched protocol developed for fiber optic systems which have near-error-free performance characteristics. The protocol encompasses features of both the data link layer and the routing layer, layers 2 and 3 of the Open Systems Interface (OSI) model.

The main features of ATM are guaranteed quality of service (QoS), ease of switching and multimedia compliant. The ATM cell structure of 53 bytes - a 5 byte header and 48 byte payload - was designed for high bandwidth implementations with low switching and routing overhead. The 5 byte header of each cell contains all the information necessary for the end-to-end delivery of the cell. If the header of the cell is corrupted beyond the corrective abilities of the included 1 byte CRC check then the cell is discarded. The small header and limited routing rules allow the switching to be performed in high speed hardware. The small size of the ATM cell - 53 bytes - allows small internal buffers to be used in the switches keeping cell delay and jitter to a minimum.

The main requirements of ATM are: a near-error-free link, fixed cell size and order of cells must be maintained. These requirements directly affect the design of satellite and wireless networks that wish to utilize ATM.

The following interrelated issues must be addressed for ATM to be utilized over satellite (and other wireless) links:

- Quality of Service
- Traffic Management and Congestion Control
- Routing
- Handover
- Signaling
- Multicasting

---

1 We defined cell switching as a subset of packet switching where all packets are of one fixed size.
Quality of Service

Quality of service is application dependent. Voice quality may be acceptable with bit error rates of 10E-6 or higher. Medical imaging may find 10E-7 acceptable as file transfers are not delay sensitive and the high-level protocols will resolve any problems resulting from non-error free transmission. Thus, the question to be resolved is: "What link quality will the satellites have to provide in order to be globally interoperable with terrestrial systems?"

The Satellite Networks and Architectures Branch (SN&AB) performed QoS measurements using an EF-Data SDM-9000 IDR Modem [1]. The results presented here indicate that the requirements for all classes of ATM service as proposed in ITU-T Draft Recommendation I.356 may not be accommodated by transmission services designed to just comply with ITU-T Recommendation G.826, such as the IDR service as specified in Tables 2 and 3, IDR Performance, p. 36A, Intelsat Earth Station Standard IESS-308 (Rev. 7A). For certain classes of ATM it will be necessary to augment the error performance of the basic IDR service through the use of the optionally specified Reed-Solomon outer codec or other link enhancer schemes. The results presented here do not include such performance enhancing techniques. Figure 3 also illustrates the satellite link availability requirements as contained in ITU-R S.1062 with the ATM current and proposed QoS requirements of ITU-T I.356, B-ISDN ATM Layer Cell Transfer Performance. Notice that there is potential for an order of magnitude or more improvement required for the QoS requirements for stringent telecommunication applications such as compressed data and video. Also note that as the QoS requirements increase, the link availability decreases.

![Figure 3](image)

Recent experimental result using MPEG-2 compressed digital video indicate that the MPEG-2 decoders temporarily loose synchronization at CLRs and CERs in the order of 1.0E-7 and 4.0E-7 respectively [2]. As an example, the movie, "Apollo 13," was video taped to get some idea of the long term Quality of Service requirements. An EF-Data SDM 9000 45 Mbps QPSK modem was configured to utilize 3/4 convolutional coding (without the Reed-Solomon code activated) at an Eb/No of 8.0 dB which corresponds to a 2.0E-8 BER, 5.0E-7 Cell Error Ratio, and 4.0E-8 Cell Loss Ratio. In 2 hours and 20 minutes, there were at least 12 identifiable errors, block errors or loss of synchronization. Video was also recorded using combined 3/4 convolutional coding and Reed-Solomon block coding resulting in a BER, CLR, and CER of approximately 1.0E-8, 4.5E-7 and 5.5E-7 respectively. Block errors and synchronization losses averaged approximately four to five per hour.

* As agreed to at the ITU-T Study Group 13 Rapporteur’s meeting held in Lannion, France, 13-17 November 1995.
six per half hour. Thus, one would expect the CLR and CER requirement to be at least 1.0E-8 and 1.0E-7 or better depending on the customer’s acceptance level.

Traffic Management and Congestion Control

The SN&AB is funding Sterling Software Inc. to investigate ATM traffic management and congestion control issues related to satellite systems. The principle investigator is Dr. Raj Jain at Ohio State University [3]. Recent work has addressed Optimization of performance of TCP/IP over satellite ATM networks and switch algorithms. Of particular interest has been Dr. Jain and Seong-Cheol Kim’s work regarding segment-by-segment congestion control known as “virtual source/virtual destination” (VS/VD). This control scheme is advantageous to long delay links as - in the limit - each segment may be independently controlled [Figure 4].

![Segment-by-Segment Control Diagram](image)

**FIGURE 4 - Segment-by-Segment Control**

Additional work is being performed by Dr. Jain and his colleagues regarding Internet protocols over ATM over satellite including TCP (Transmission Control Protocol) over UBR (Unavailable Bit Rate) and TCP over ABR (Available Bit Rate). The key parameters being studied are the performance with limited buffers, and buffer requirements for zero loss. Various algorithms are also under investigation such as Early Packet Discard, Fast Retransmit Recovery and fair buffer allocation.

Routing

Route selection in the emerging ATM networks is based on the capabilities of the “Private Network-Network Interface Specification Version 1.0” (PNNI 1.0) [4] and its forthcoming antecedents. The protocol defines the distribution of topology information between switches and clusters of switches. This information is used to compute paths through the network. A second protocol within PNNI is defined for signalling, that is message flows, used to establish point-to-point and point-to-multipoint connections across the ATM network. This signalling protocol is based on the ATM Forum UNI signalling, with mechanisms added to support source routing, crankback, and alternate routing of call setup requests in case of connection setup failure.

The interconnection of satellite networks with ATM networks will require that the satellite networks participate in the routing protocols. At this time NASA Lewis Research Center is planning a set of experiments [5] to be completed in the summer of 1997 that will:

1. Establish a baseline network for studying the issues related to the signalling operation of a hybrid satellite/terrestrial network.
2. Determine signalling parameters issues related to supporting "normal" network and user services that are required in a hybrid satellite/terrestrial network.
3. Collect data from the baseline network under various operating conditions including that of when a terrestrial or satellite link is abnormally disabled (network survivability).

Handover

Handover is the mechanism that supports the mobility of a network device, its network links, and its active connections. The ATM specification for handover is currently in development in the ATM Forum Wireless ATM working group [6]. There are three types of handover that are required to be designed for in satellite networks:
1. End-user to satellite
2. satellite to satellite
3. satellite to network access point

It is anticipated that the three types of handovers will operate independently of one another in order to be compatible with the handover protocol envisioned in the ATM Forum. A key element of the handover mechanism is the rerouting of active circuits. This element also depends on the PNNI protocol for implementation. Within the ATM Forum other mechanisms such as Fault Tolerance [7][8], and Connection Modify [9] will also rely on the reroute mechanism. All of these mechanisms are critical to the optimal operation of a hybrid satellite-terrestrial network.

A contribution to the ATM Forum has been made by NASA Lewis Research Center to define a generic reroute protocol that is applicable to the mechanisms of Handover, Fault Tolerance, and Connection Modify [10]. Additional contributions on the subject will be submitted by NASA to the ATM Forum.

Signaling
NASA foresees some issues regarding network management and control signaling for some satellite architectures including LEO and MEO networks, and TMDA and Multi-beam systems. Signaling must be maintained in order to maintain routing status and congestion notification. This must be designed into the networks and satellites – particularly for onboard processing satellites.

Multicasting
NASA foresees multicasting as an area that satellites have a tremendous advantage in over terrestrial networks in that the routing tree structure should be significantly smaller utilizing a satellite network for broadcast and multicast services such as the “push services,” applications that push information to integrated users - also known as netcasting applications. This area must be exploited for satellites to be a viable competitor in the Internet environment. Much work is being done by the terrestrial telecommunication industry regarding IP multicasting, ATM multicasting and integrating IP multicasting with ATM. This must be followed closely to ensure satellite friendly implementations.

Security
Security is always an issue for networking and more so for ATM because of its structure. Since ATM was developed for easy of switching via the header, it is easy to filter off messages by monitoring the header. For free-space communications, this is a difficult problem. The header cannot be scrambled or modified for security as the switching, routing, and header error control will be compromised. Therefore, ATM security must take place in the payload. A variety of groups are working this problem. The US military is a major funding source for such activity as they wish to utilize ATM technology within their networks.

Common Air Interface
A new generation of satellite systems are being designed, developed and deployed for a range of services — fixed, mobile and broadcast. Examples of such systems include geosynchronous, low earth orbit and medium earth orbit mobile systems, as well as Ka-band geosynchronous and low earth orbit systems for high speed services. A number of these systems are envisioned to be regional and global. A set of common air interface standards for these systems will be beneficial from the point of view of interoperability.

A standards subcommittee project, Common Air Interface for Satellite Systems, has been approved by the Communications and Interoperability Section of the Telecommunications Industry Association’s (TIA) Satellite Communications Division [11]. The work will be carried out in a new working group to be formed under standards subcommittee TR-34.1. The satellite systems to be considered will range from satellite personal communications systems to broadband satellite systems.

TR-34.1 is also involved in other areas of communications and interoperability. Its specific working groups are Wireless Asynchronous Transfer Mode (ATM), Internet over Satellite, ATM Speech, ATM Traffic Management and Congestion Control, ATM Quality of Service and Hybrid Reference Models.
4. TRANSPORT LAYER ISSUES

Satellite research in the Internet protocol stack known as TCP/IP has tended to focus on the Transmission Control Protocol (TCP) since it is the basis for World Wide Web traffic and is the standard protocol for reliable transport. There are several approaches to dealing with TCP over satellite links including spoofing, protocol conversion, caching, and TCP extensions [12]. We have concentrated on TCP extensions as being the most general solution to interoperability. There are three main issues which need to be addressed in using TCP over satellite: 1) Use of Long, Fat Network (LFN) extensions, and 2) Inefficiencies in the slow start and congestion avoidance algorithms, and 3) Loss recovery.

The issue of the use of LFN extensions is basically one of implementation. The LFN extensions were designed to allow TCP steady-state operation over high bandwidth-delay product connections and have existed (and been discussed and slightly modified) for many years [13, 14]. They are an optional TCP extension and are not necessarily available in a particular TCP implementation. The product of the round trip time for a connection and its data rate is known as the bandwidth-delay product. A high bandwidth-delay product connection can result from a very high data rate just as well as from a high delay, thus LFN extensions are needed for high data rate terrestrial connections as well as for satellite connections.

A key part of the LFN extensions is the window scaling option or “large windows.” TCP is a sliding window protocol and the size of the window determines the maximum throughput of a TCP connection. The size of the window should be equal to the bandwidth-delay product for full utilization of a channel in steady-state. Without the LFN extensions, the maximum window size for TCP is 64 KB. The typical default value for the window size in a particular implementation might be 8 KB. Window scaling allows as much as a 1 GB window if both ends of the connection support the option.

Table 1 shows a comparison of bandwidth*delay products for examples of terrestrial and satellite delays. The LFN extensions are not even needed for a typical low rate satellite connection and it is not until an individual satellite connection gets up to around 1 Mbps that they are needed. Table 2 shows that even for short terrestrial links, LFN extensions will be necessary as data rates increase. The 10 Mbps entry of 52 ms means that a terrestrial cross-country link of 10 Mbps with a round trip time of 80 ms (>52 ms) will require LFN extensions to utilize the full bandwidth of the link. Thus, as data rates increase everywhere, LFN extensions will have to become more prevalent in TCP implementations.

### TABLE 1 - Comparison of 80 ms and 560 ms bandwidth-delay products

<table>
<thead>
<tr>
<th>RATE (bps)</th>
<th>RTT (sec)</th>
<th>B*D (bytes)</th>
<th>LFN (Y/N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>33600</td>
<td>0.08</td>
<td>336</td>
<td>N</td>
</tr>
<tr>
<td>33600</td>
<td>0.56</td>
<td>2352</td>
<td>N</td>
</tr>
<tr>
<td>128000</td>
<td>0.08</td>
<td>1280</td>
<td>N</td>
</tr>
<tr>
<td>128000</td>
<td>0.56</td>
<td>8960</td>
<td>N</td>
</tr>
<tr>
<td>1550000</td>
<td>0.08</td>
<td>155000</td>
<td>N</td>
</tr>
<tr>
<td>1550000</td>
<td>0.56</td>
<td>108500</td>
<td>Y</td>
</tr>
<tr>
<td>10000000</td>
<td>0.08</td>
<td>1000000</td>
<td>Y</td>
</tr>
<tr>
<td>10000000</td>
<td>0.56</td>
<td>700000</td>
<td>Y</td>
</tr>
<tr>
<td>1.55E+08</td>
<td>0.08</td>
<td>1550000</td>
<td>Y</td>
</tr>
<tr>
<td>1.55E+08</td>
<td>0.56</td>
<td>1.1E+07</td>
<td>Y</td>
</tr>
</tbody>
</table>
TABLE 2 - Round trip times to match the maximum TCP window size without scaling option

<table>
<thead>
<tr>
<th>RATE (bps)</th>
<th>RTT for B*D=64KB (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>33600</td>
<td>15.604</td>
</tr>
<tr>
<td>128000</td>
<td>4.096</td>
</tr>
<tr>
<td>1550000</td>
<td>0.338</td>
</tr>
<tr>
<td>10000000</td>
<td>0.052</td>
</tr>
<tr>
<td>1.55E+08</td>
<td>0.003</td>
</tr>
<tr>
<td>6.22E+08</td>
<td>0.001</td>
</tr>
</tbody>
</table>

In addition to scaled windows, LFN extensions provide a timestamps option which is used for the Round Trip Time Measurement and for Protect Against Wrapped Sequences algorithms. A selective acknowledgment (SACK) option was defined as part of the original LFN extensions, but was split out to be technically addressed as a separate topic and has only recently been published [15]. SACK could be used to improve TCP performance over noisy links [16]. Although the SACK RFC describes how to implement selective acknowledgments, it does not describe a retransmit strategy beyond existing congestion avoidance, so further research is needed for SACK implementation.

The second issue is a research problem as opposed to an implementation problem and is concerned with the dynamic behavior of TCP as it starts up a connection or deals with lost packets. With a long round trip time, the slow start and congestion avoidance algorithms built into TCP to combat network congestion operate very slowly. Slow start is used to ramp a TCP connection up onto the network when it is first established, and it is also sometimes used (followed by an even slower algorithm called congestion avoidance) when a packet is lost indicating congestion.

It can be shown [17] that the time spent in slow start from the beginning of a connection to steady-state in the absence of losses is

\[ t = R \log_2 W \] (1)

where \( t \) is the time spent in slow start, \( R \) is the round trip time, and \( W \) is the size of the window in packets (with delayed acknowledgments, which are commonly used, the time will be longer than the value given by equation 1). The key thing to recognize from equation 1 is that the time spent in slow start is proportional to the round trip time. For a geosynchronous satellite connection of, say, 560 ms round trip time, the slow start time compared to a terrestrial cross-country connection of, say, 80 ms will be seven times as long. Similarly, time spent in congestion avoidance over that same geosynchronous satellite link will be around seven times longer than that terrestrial cross-country connection (starting from similar initial conditions).

For small file transfers, such as are now typical on the World Wide Web (WWW) [18], a connection will not get out of the slow start phase before all the data has been sent. With the move to HTTP 1.1, the Web will move to larger file transfers which may allow connections to get through slow start or at least use more of the available bandwidth. Efforts are also underway to speed up slow start. One proposal being put before the IETF is for a larger initial window of up to 4KB. The results of testing and simulation thus far indicates that this modification will not hurt the network [19]. This will shave a couple of round trip times off a transfer.

Studies have shown that multiple connections (in the range of 4 to 8) allow an application to utilize more of the available bandwidth [20], but this is not looked upon by the networking community in general as a "fair" solution. Multiple connections can be thought of as allowing more aggressive start-up and less aggressive congestion control [21]. Some WWW browsers currently allow multiple connections to be established.

A problem for noisy links with TCP's congestion handling algorithms is that TCP uses packet loss as an indication of congestion and throttles back on the transmission rate when a packet is lost. The best way to avoid the inefficiencies of congestion avoidance is not to needlessly invoke it. SACK could help in lost packet cases,
but further research is needed to develop ways of discriminating between packet loss and congestion. TCP works best over low error rate links like fiber. A case can even be made that error rates on high data rate/long delay links should be even lower than those on fiber for TCP to work efficiently [22].

5. CONCLUDING REMARKS

Although this paper has outlined the major issues concerning seamless interoperability between satellite and terrestrial networks in the emerging GII it in fact only touches the first layer of complexity that needs to be addressed. Every year additional telecommunication services are envisioned that are by and large incompatible with the unique transmission characteristics of satellite. The role of satellites in the emerging global information infrastructure is critical to provide information services to anyone, anywhere, at anytime which makes the issue of interoperability between satellite and terrestrial networks critical. To address the interoperability issue it is important that satellite and terrestrial telecommunication standards for the GI evolve simultaneously. The product of this involvement will lead to enhanced end-to-end global information services in the developing global information market.

6. REFERENCES

**Enhancing End-to-End Performance of Information Services Over Ka-Band Global Satellite Networks**

Kul B. Bhasin, Daniel R. Glover, William D. Ivancic, and Thomas C. vonDeak

National Aeronautics and Space Administration
Lewis Research Center
Cleveland, Ohio 44135-3191


The Internet has been growing at a rapid rate as the key medium to provide information services such as e-mail, WWW and multimedia etc., however its global reach is limited. Ka-band communication satellite networks are being developed to increase the accessibility of information services via the Internet at global scale. There is need to assess satellite networks in their ability to provide these services and interconnect seamlessly with existing and proposed terrestrial telecommunication networks. In this paper the significant issues and requirements in providing end-to-end high performance for the delivery of information services over satellite networks based on various layers in the OSI reference model are identified. Key experiments have been performed to evaluate the performance of digital video and Internet over satellite-like testbeds. The results of the early developments in ATM and TCP protocols over satellite networks are summarized.