
TCP/IP over a Satellite Cell Switch

by Aaron Falk

Network Systems Engineer

TRW Space and Electronics Group

Introduction

Supporting Internet protocols (TCP, IP, and others) is a key requirement for the economic viability of any broadband satellite network. TRW has designed a satellite network architecture built around a payload cell switch. In order to provide Internet services over such a network two key, interrelated analysis are required. First, the interactions between these protocols and the underlying link mechanism, an ATM-capable cell switch, need to be explored. Second, the basic operation of today's most popular and pervasive Internet protocols will operate over such a satellite network needs to be examined.

In this paper some background on the role of satellites in providing Internet access is presented followed by a description of the TRW cell switching network architecture. The analysis is comprised of a review of the technical issues related to carrying IP and TCP over this architecture as well as a description of the standards activities currently underway. The paper will conclude with an index of web sites for those interested in monitoring publicly available activities related to Internet over satellite.

What is the Internet and How Can Satellites Fit In?

The explosion of interest in the global Internet has changed business, academic, and personal lives of millions of people. Internet connectivity has created a paradigm shift in the way information is shared for everything from large software development projects to airplane schedules to sharing baby pictures.

The Internet is defined as a network of networks each exchanging information using a shared set of loosely defined protocols referred to as TCP/IP (the acronyms of the two most popular protocols: the Transmission Control Protocol and the Internet Protocol). At a minimum, a network on the Internet must support IP. Each network can be considered autonomous lending to the Internet's somewhat anarchic flavor. In a general sense, a network can be comprised of Internet hosts, computers which can be accessed though an IP address, and routers, computers which handle communication with external networks.

From [Architectural Principals of the Internet \[RFC-1958\]](#),

"Many members of the Internet community would argue that there is no architecture, but only a tradition, which was not written down for the first 25 years (or at least not by the Internet Architecture Board). However, in very general terms, the community

believes that the goal is connectivity, the tool is the Internet Protocol, and the intelligence is end to end rather than hidden in the network.”

“Heterogeneity is inevitable and must be supported by design. Multiple types of hardware must be allowed for, e.g. transmission speeds differing by at least 7 orders of magnitude, various computer word lengths, and hosts ranging from memory-starved microprocessors up to massively parallel supercomputers. Multiple types of application protocol must be allowed for, ranging from the simplest such as remote login up to the most complex such as distributed databases.”

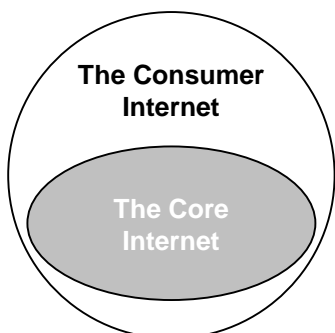


Figure 1. The Consumer Internet and Core Internet can be represented by a Venn diagram.

As illustrated in Figure 11, there are actually two Internets [Quarter]. The Core Internet [computers not behind firewalls which are capable of distributing information via ftp, telnet, and World Wide Web] can be thought of as the information supply. Computers on the core Internet are typically high powered machines with high bandwidth connectivity. They represent the set of destinations for the Consumer Internet [which includes the core plus information-consumer-capable computers]. Of course, machines in the Core Internet can be information consumers. More importantly, there is a large group of computers which are not capable of (or not efficient at) distributing data. A widely known example of these types of computers are home PCs with dial-up Internet Service Provider (ISP) accounts. We'll come back to this important distinction soon.

One reason the Internet protocols suite has grown to such heights of popularity is its ability to seamlessly interconnect networks of disparate origins, architectures, and speed. IP, the defining protocol of the Internet, has been specified in such a way that it can function on media from terabit per second fiber networks to carrier pigeons (see [IP over Avian Carriers \[RFC-1149\]](#)). The connectivity of the Internet more than any other characteristic that is its true value. Yet, there are large parts of the world which, due to constraints of the telecommunications infrastructure, have no access to the Internet.

Why Internet over Satellite?

Satellites may be used to provide connectivity to the Internet where there is none and may be used to enhance the bandwidth of existing connectivity through bypassing bottlenecks in the terrestrial infrastructure. There are two elements of networking infrastructure which satellites can address. First, satellites can provide long haul services allowing distant locales to be connected directly into the developed Internet. This can improve performance experienced by the user by eliminating intermediate bottlenecks which may inject queuing delays. Second, satellites can be used in lieu of more expensive, local infrastructure by providing wireless Internet access to large geographical areas. Even in cases where satellites don't provide the permanent solution they can be an excellent way to provide temporary infrastructure while the business case necessary to support the expense of installing a permanent terrestrial infrastructure is demonstrated.

Current Internet Services over Satellite

IP has been run over satellites since the earliest days of the Internet.

Satellites are currently used to provide two general types of Internet connectivity. Trunking is the most predominant satellite service--interconnecting routers which have performed traffic aggregation terrestrially. These services are similar to terrestrial long haul fiber backbone services. They are sold either to businesses to connect corporate LANs to the Internet or to Internet Service Providers, who sell access to end users using terrestrial infrastructure (e.g., telephone lines and modems) aggregates user traffic before routing over the satellite channel. Examples of these services include Panamsat, Karewak Inc. (an Alaska ISP), Intelsat, Astra, and Orion.

The second type of currently available Internet service uses the satellite for hybrid network access. The service provides Internet access directly to the end user. Users dial into a local modem as they would for terrestrial ISP service but when files are to be transferred from the Core Internet to the user, such as a response to a World Wide Web request, the data is routed via satellite. In this way the satellite uplink terminal can provide aggregation, such as in the satellite trunking, and provide a higher burst rate than usually found in modems. An example of this service would be DirecPC.

Proposed Internet Services over Satellite

Twenty (TBR) orbital slots and transmission licenses have been awarded by the FCC for use in the Ka-band (30-40GHz). This frequency band will be the highest used for commercial satellite communications and offers the opportunity for satellites to provide high bandwidth network services to large numbers of users. The higher frequency allows for smaller beam footprints. This will enable the same frequency to be reused in different geographical areas and this spectrum reuse will increase the total effective bandwidth of the satellite payload.

A stated goal from the FCC filings of most of the companies is to use the bandwidth to provide high rate data services to consumers and small business. However, since the filing most of the Ka-band filers have kept their strategies private. Nevertheless, publicly available information shows a range of satellite architectures—from traditional “bent pipe” designs to full on-board cell switches—revealing that there is no clear consensus on the best way to use satellites inherent advantages of flexible bandwidth allocation, universal coverage, simple multicasting, and reliability to address the growing telecommunications market.

It is safe to assume, however, that the Internet will play an important role in how these systems will be used. The rest of this paper will review some satellite network topologies, a tutorial on how Internet protocols interplay with the characteristics of a geosynchronous satellite, and an overview of a network architecture designed around a digital satellite cell switch developed by TRW which is capable of providing high speed Internet services.

Carrying IP over a Satellite Cell Switch

The Internet Protocol is the common ground upon which the global Internet is built. IP is necessary and sufficient to support a vast number of data applications and is rapidly evolving to support multimedia applications as well. While native ATM applications are only beginning deployment, IP based applications, including multimedia, are proliferating and, in many cases, are freely available.

Clearly, a satellite network needs to be able to support the Internet Protocol.

[Comer] provides some useful definitions when discussing Internet protocols:

- IP Address: an address assigned to each host that participates in the TCP/IP internet
- Router: a special purpose, dedicated computer that attaches to two or more networks and forwards packets from one to the other
- ARP or Address Resolution Protocol: the TCP/IP protocol used to dynamically bind a high-level IP Address to a low-level physical layer address

IP over Asynchronous Transfer Mode (ATM)

IP is a connectionless protocol utilizing datagrams which are routed through the network based on the destination address found in the header. In contrast, ATM is connection oriented requiring pre-connection signaling to assign connection identifiers and establish a path through the network. Traditional IP networks maintain permanent connections between the routers in the network. When routing IP traffic through an ATM network the ATM virtual circuits (VCs) may also be permanent. If so, each router needs to maintain a fixed mapping between the IP addresses of each host it is directly connected to through the ATM network and the ATM VCs which can be used to reach them. Since the connections are permanent this could be done by hand. However, the problem gets more difficult when non-permanent ATM connections, or switched virtual circuits (SVCs) are used. In this case, a mechanism needs to be developed which allows a host added to the network to obtain the address mappings.

Several standardized schemes exist for carrying IP over ATM networks. They are known as Classical IP over ATM (CLIP), LAN emulation (LANE), and Multi-Protocol over ATM (MPoA). What they all have in common is that ATM virtual circuit connections are established to carry IP datagrams which have been fragmented into ATM cells. Originally developed for use over LANs, these schemes utilize mechanisms which don't scale well over GEO satellite networks.

Satellite networks are particularly sensitive to the overhead bandwidth involved in call setup signaling and idle connections. In general, standardized IP over ATM schemes all require routing tables containing the mapping between the IP addresses of the adjacent IP hosts and their ATM addresses. Accessing this centralized table requires establishing an ATM VC. Once the table information is retrieved, another ATM VC must be established to transmit the packet. The protocols contain activity timers which tear down virtual circuits after an idle period. A balance must be struck between the delay experienced by the user before sending data and the resources wasted in maintaining idle connections.

All these schemes were designed to support LANs with relatively small numbers of nodes. Each virtual circuit in a satellite network uses up valuable resources, whether there is traffic on it or not. Schemes like MPOA which require up to three virtual circuits per host do not scale well to support very large numbers of users.

However, the impact of the delay and signaling penalties for the various ARP schemes depends on whether the system will be used to support satellite dialup services or provide mesh connectivity. In the dialup configuration, most users will be routing all data to an Internet gateway and will not need to perform the ARP functions except when they log into the system. Nevertheless, most of today's ATM switches can not accommodate very large numbers of ATM VCs for IP over ATM use and gateways handling connections from a hundreds or thousands will require development.

Connectionless IP over ATM Schemes

Several companies have developed solutions for connectionless routing of IP over ATM. Most of these schemes involve a modification of the ATM header to make the ATM cell self-routing by allowing use of a global header value to reach a specific destination. Network ATM switches which support the connectionless header format can transfer cells to the correct switch output port without per connection signaling and pre-configuration. Advantages of a connectionless ATM scheme are that connection setup time is reduced, especially for networks where mesh connectivity is desired. Disadvantages of such a scheme are that the size of the ATM header can limit the number of users that such a scheme can support. Additionally, all the switches in the network need to support handling of special format headers.

Early Packet Discard

An effect of sending IP packets in ATM cells is that if cell loss occurs due to insufficient buffers or link errors in the ATM network partial packets may be transmitted. Upon exiting the ATM network, packet reassembly will fail, the transmitted fragments will be discarded, and, if TCP is being used, a retransmission will be requested. Because of the sensitivity to bandwidth utilization, the transmission of data which is destined to be discarded is especially costly in satellite networks. If buffering is limited or links may be corrupted, Early Packet Discard (EPD) algorithms [Romanow] should be used to reduce the number of packets affected by a cell congestion event.

There are several different forms of EPD and the simplest is to implement is partial packet discard. In partial packet discard, once a cell has been discarded, all subsequent cells carrying fragments of the same packet are also discarded. If, during a buffer overflow, multiple cells would be discarded the use of partial packet discard would confine the cell discards to a smaller number of packets and, hence, will allow more useful data to be sent on the link as well as reducing the number of packet retransmissions.

TCP

If the Internet Protocol is what defines the Internet, the Transmission Control Protocol (TCP) is what gives the Internet it's current usefulness. With respect to performance over satellites, there are two fundamental characteristics of TCP. First, TCP provides reliable data transmission over an essentially unreliable network. This is accomplished through a selective retransmission protocol which uses acknowledgments from the receiver to tell the sender that data has been received successfully. There are several variants of retransmission protocols, the current recommended standard is known as Selective ACKnowledgements or SACK [RFC-2018].

Second, TCP utilizes an adaptive window algorithm to adjust to congestion in the network. This mechanism is designed to protect the network from collapse. If a network's capacity is exceeded, the buffer occupancy in the network will grow eventually leading to packet arrivals with no available storage space. In this situation packets are discarded, which leads to packet retransmission. Without a congestion control function, packet retransmissions could inject more data into the network just at the time when it has the least available capacity, leading to more packet drops and more retransmissions and, eventually, network collapse as no useful work is done by the network. To alleviate this concern, packet losses in TCP are interpreted as indicators of congestion and, when the sender becomes aware of a packet loss, he reduces the amount of data being transmitted.

The retransmission and congestion control mechanisms work to provide efficient, reliable transmission of data while maintaining a stable Internet with literally millions (billions? trillions?) of simultaneous connections.

While the retransmission and congestion control create a very reliable, stable network design, the mechanisms developed tend to sacrifice connection bandwidth and link utilization for overall network stability. This effect is exacerbated by the delays imposed by transmission over a GEO satellite network. These issues will be explored in the following sections. While TCP is used to support a wide variety of applications, this section will focus on evaluating large multi-megabit transfers, i.e. ftp file transfers, and small transfers, i.e. web pages.

Introduction to Congestion Control Mechanisms

Congestion control is an umbrella which covers several mechanisms including Slow Start, Congestion Avoidance, Fast Retransmit, and Fast Recovery. These mechanisms were developed expressly to reduce the risk of network collapse in the Internet and are critical to the smooth growth it is experiencing. [Stevens] has an excellent overview of TCP congestion control. All these mechanisms are implemented at the data sender and are implemented by manipulating the sender's window—the amount of data the sender is allowed to have unacknowledged—during a TCP connection. The maximum sender's window is determined by taking the minimum of the receiver's advertised window and the sender's congestion window.

In order to use all the available network bandwidth, the sender's window should be able to grow to be equal to the delay-bandwidth product of the connection. The delay-bandwidth product is a metric estimating the capacity of the network to hold data in transit between the sender and receiver. It is usually calculated as the product of the bandwidth of the network bottleneck link and the round trip time. In order to make maximum use of the bottleneck link, a user should be capable of "filling the pipe" and sending the maximum amount of unacknowledged data that the network can handle. If the receiver has advertised a receive window which is smaller than the delay-bandwidth product the sender will never be able to inject enough data into the network to fill the pipe.

If a TCP connection has experienced no loss it is in one of two phases: Slow Start or Congestion Avoidance. The two phases have been designed so that during Slow Start the congestion window starts small and opens very quickly. Once the estimated capacity of the network has been reached, indicated either by a packet loss—i.e. congestion—or by crossing the Slow Start threshold—usually initialized to the value of the receiver window—the connection moves into the Congestion Avoidance phase where congestion window opens much more slowly and the network is probed for additional capacity.

Slow Start has often been called a misnomer because, in fact, Slow Start opens the window very rapidly. The initial congestion window size in a TCP connection is set to around one TCP packet so only a single packet can be sent. The acknowledgement of the first packet triggers two packets to be sent and the ACKs from those two packets each trigger two packets (i.e. four packets) to be sent. So, it can be seen that after each round trip, twice as many packets have been injected into the network resulting in an exponential increase of unacknowledged packets.

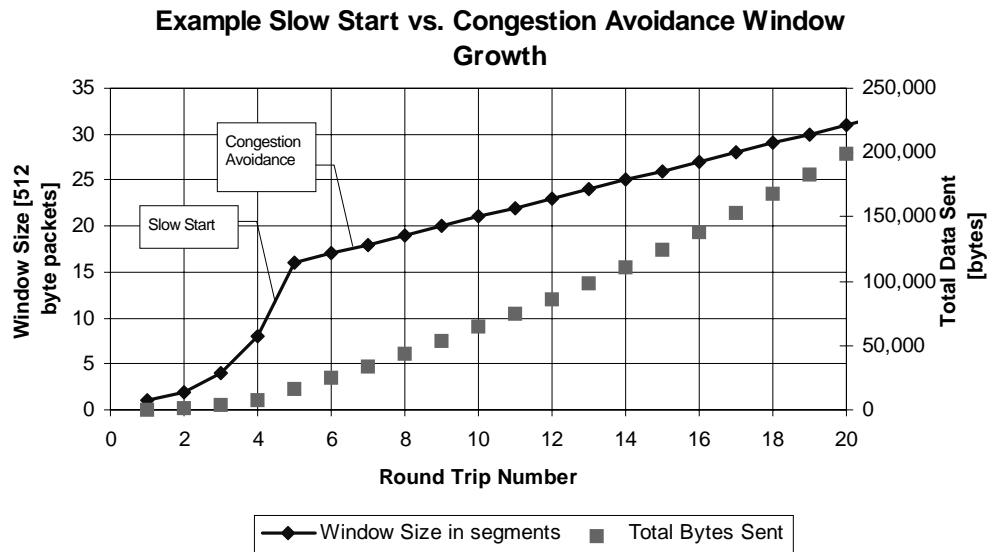


Figure 2. Window size grows exponentially while a TCP connection is in Slow Start and linearly while in Congestion Avoidance.

When the congestion window crosses the Slow Start threshold, or *ssthresh*, the Slow Start phase ends and the Congestion Avoidance phase begins. Congestion Avoidance specifies that the congestion window grow by one packet per round trip time—changing the rate of window growth from exponential to linear.

Both Slow Start and Congestion Avoidance increase the window monotonically and, eventually, the connection will reach the limits of the network capacity and a packet will be dropped. If there is still data flowing through the connection, the sender learns of a packet drop through a duplicate acknowledgement. Each ACK identifies the sequence number of the last byte in the last packet successfully received in consecutive order. For example, if the sender sends out four 512 byte packets it can expect to receive ACKs showing successful reception of bytes 511, 1023, 1535, and 2047. If the third packet is lost, the sender would receive 3 ACKs acknowledging sequence numbers 511, 1023, and 1023. The fourth packet generated an ACK but because there was a gap in the sequence number, a duplicate ACK was generated. If the last packet had been lost, there would have been no duplicate ACK and, eventually, a timer would have expired causing the connection to reenter Slow Start.

The Fast Retransmission mechanism takes advantage of the duplicate ACKs to re-send lost data before the retransmission timer can expire. Upon the receipt of three duplicate ACKs the sender knows two things: first, the sequence number of the first lost packet and, second, data is still flowing through the network (since new packets are arriving at the destination causing duplicate ACKs to be generated). Fast Retransmission specifies that the sender re-send the packet that was known to be lost.

Fast Recovery specifies that after the Fast Retransmission has resent a lost packet, the sender should pause briefly allowing the network to drain, then continue transmitting one packet for each duplicate ACK received (i.e. maintaining a fixed congestion window) and, when the retransmitted packet is acknowledged, set *ssthresh* to half the congestion window value at the time of retransmission and enter Congestion Avoidance.

A shortcoming of Fast Retransmission is that the mechanism can only reliably identify one missing packet per window. Since the duplicate ACKs show only the last successfully received, in-order packet, if more than one packet is lost the sender won't know until the round trip timer expires and the connection returns to the Slow Start. The TCP option Selective Acknowledgements (SACK) addresses this by adding a field in the ACK packets containing the sequence number ranges of successfully received packets. As long as the network can pass data, congestion control and SACK used together will eliminate the Go-Back-N behavior sometimes attributed to TCP. (If the network cannot pass data, the correct behavior is to allow the connection to time out then behave in the same fashion as a new connection, i.e. return to Slow Start.)

GEO Satellites and TCP

TCP performance over an uncongested GEO satellite can be characterized based on the link bit error rate, delay, and delay-bandwidth product. It is advantageous from the satellite system providers' perspective to maximize bandwidth utilization and the capacity of a satellite in order to maximize revenue. Additionally, when competing with terrestrial networks, response time becomes important. The speed of light places a 500ms "tax" on each round trip made over a GEO satellite. This delay tax can have deleterious effects on TCP performance. Not all of the inefficiencies will always apply and the specific issues will be explored in the following sections. For further reading see [Partridge] for more a more detailed discussion of the performance issues.

Window Size and GEO Satellite Networks

As noted in the discussion on congestion control, it is necessary for the receiver's window to be set at least to the delay-bandwidth product of the connection. The default window size supported in the TCP header is 64k bytes. Using this as the largest delay-bandwidth a TCP can support and, given a round trip time of 500ms, the maximum bandwidth a single TCP connection can support would be around 1Mbps. While this may be considerably more than a typical home user can get today over a modem connection to the Internet, it represents an artificial limitation of the bandwidth a satellite network can provide.

Fortunately, the satellite environment shares this problem with any other network that contains a large delay-bandwidth product. This includes networks with very large bandwidths, such as terrestrial gigabit networks, as well as those with large delays such as transoceanic fiber as well as GEO satellites. There exists a suite of TCP options, known as TCP for Large Windows [RFC-1323] which not only support large windows but also addresses issues of ambiguity in sequence number space which can occur when multiple packets with the same sequence number are unacknowledged. TCP for large windows extends the maximum window size to support a delay-bandwidth product of over 1Gbyte and will allow a maximum connection rate of over 16Gbps over a GEO satellite.

Packet Error Rate and TCP Throughput

In the case where the network is lightly loaded, bit errors will dominate the performance of TCP. The congestion control algorithm, by design, interprets packet loss as an indicator of network congestion and reduces the sender's window to allow the network to drain. Uncorrected bit errors caused by satellite links will have the same effect on connection performance even though there is no congestion. Specifically, if a packet loss occurs during Slow Start, the connection will move into Congestion Avoidance and rate that the window opens will slow dramatically. This can be seen in Figure 2. For a window size of N, it takes a single round trip for the window to double in Slow Start but N round trips for

the window to double in Congestion Avoidance. This slows the rate the data is transferred (i.e. the user's response time for interactive applications) and reduces the utilization of satellite network bandwidth.

If a transfer is large enough to move from Slow Start to Congestion Avoidance, the throughput can be estimated as a function of the round trip time, packet size, and packet error rate [Mathis]. For example, for a 500ms round trip time, typical packet size of 512 bytes, and uniform bit error rate of $1E-7$, the throughput can be estimated as 140 kbps. For $1E-9$, the throughput can be estimated as 1.4 Mbps.

Other factors will affect TCP throughput and the results presented above should be considered as representative of ideal performance of an arbitrarily large transfer on an unloaded network with uniform distribution of bit errors. Bit errors on a real satellite link, especially in Ka-band, are not uniform and greater correlation in bit errors will reduce the packet error rate because bit errors may be concentrated into fewer packets. On the other hand, most Internet connections will traverse networks other than just the satellite link and packet loss due to bit errors or congestion on those networks as well as congestion will work to constrain TCP throughput.

How Congestion Control Affects GEOs

Most transfers over today's Internet, if run over a GEO satellite network, would be too small to allow Congestion control to reach the steady state described above. At T1 rate (1.544 Mbps), it takes about 9 round trips to before the pipe is full and over 250 kbytes

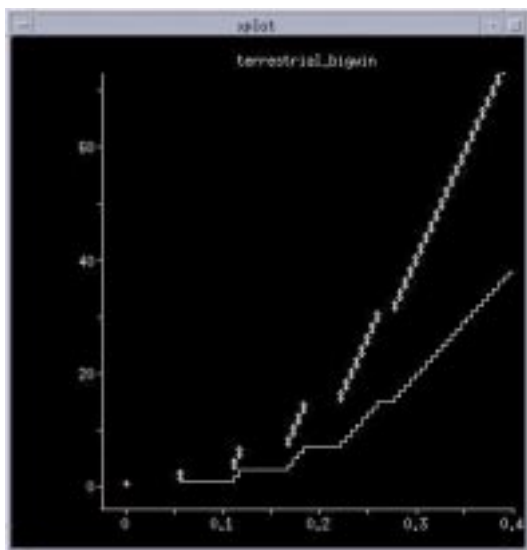


Figure 3. This plot shows a simulation of a connection in Slow Start. The stop-and-wait behavior of a TCP connection not yet transmitting at full rate can be seen. The packet numbers are on the y-axis and time, in seconds, on the x-axis. The dots show a packet transmission and the solid lines show the receipt of an acknowledgment. [plot will be updated to a connection over satellite for publication]

have been sent by then. So, World Wide Web transfers which average about 20kbytes will never reach the Congestion Avoidance phase. However, large transfers may grow more frequent as multimedia applications and software distribution become more pervasive. If a fixed bandwidth connection is allocated for a TCP connection, much of the bandwidth will be wasted. If a satellite network is designed with fixed rate circuits on the uplinks, a single TCP connection will not make efficient use the bandwidth.

The behavior during Slow Start is that some data is transmitted, then a delay, when the ACKs come back more data is transmitted (with a larger window), then another delay, etc. until the pipe is full and the sender can send continuously (see Figure 33). The underlying reason for this is that the congestion control algorithms open the congestion window once per round trip time. In networks with small delay-

bandwidth products, the window doesn't need to open as far to get to the point where the pipe is full and Congestion Avoidance kicks in to establish a steady state throughput for the connection. Large bandwidth-delay connections need additional round trips to get the window open farther. Long delay connections such as GEO satellite networks are additionally penalized because the delay for each round trip is somewhat longer than terrestrially. For long delay, high delay-bandwidth networks, the idle period is a greater percentage of the total transfer time, increasing the user's perceived transfer delay and leaving satellite bandwidth unused.

So, to summarize, there are two areas of inefficiency when running TCP over satellite: first, for small transfers Slow Start sends only a few packets per round trip time even if the bandwidth is available, and second, for large transfers Congestion Avoidance takes many round trip times to recover from a packet loss.

Questions About How TCP Performance Will Interact with ATM Rate Control and DAMA

Lower layer rate control mechanisms may have unpredictable results on TCP's congestion control and are areas which require further research. The ATM Available Bit Rate (ABR) service provides for rate management where the available rate on an individual link may vary according to the load presented. Researchers are just beginning to examine how the added feedback delay of a GEO satellite will affect the interactions between TCP Congestion control and ATM Rate Management.

Additionally, allowing multiple users to share a single uplink channel in a DAMA scheme will vary the delay and available bandwidth a TCP connection experiences. This also is an area for further research.

How Can TCP Performance Be Improved?

With all these issues, what can be done now to optimize TCP performance over satellite? The question has three components:

- a) What is available in today's standard TCP that helps performance over satellite?
- b) What magic can satellite system developers hide in their network to improve TCP performance?
- c) What satellite friendly changes should be advocated as TCP evolves?

Any changes to the TCP protocol face the constraint that they must have no negative impact to terrestrial traffic. Because TCP resides on the source and destination machines, those advocating implementation changes are at the mercy of the operating system developers for distribution of these changes. There is the possibility, however, of modifying TCP's behavior through the use of agents in the middle of the Internet known as proxies, sometimes called "spoofing." Use of these proxies is not without risk but may allow performance improvements without changing the end user's protocol stack. Also, while TCP is a standardized protocol, the philosophy behind the protocol has been to optimize interoperability by specifying only minimum functionality to allow differing systems to interoperate. Because of this, the distinction between changes to the TCP standard and changes to the implementation is important. Two TCP implementations may both be compliant with the standard and may interoperate with each other but may have drastically different performance.

IETF TCP over Satellite Recommendations

The Internet Engineering Task Force (IETF), which produces Internet standards, has created the TCP over Satellite working group (TCPSAT). TCPSAT is developing a recommended configuration for TCP stacks which will be used over satellite. At the time of this writing the document is still in working paper form but is expected to become an Informational Standard before the end of 1998. This configuration incorporates the recommended TCP options referenced above. Specifically, TCP large windows and SACK as well as path MTU discovery should be used. Path MTU discovery allows the largest IP packet possible to be used in transmission. Since the TCP congestion window is measured in packets, larger packets will allow more data to be transmitted sooner. Most modern TCP stacks have these options available.

Non-standardized Performance Enhancements

Forward error correction and buffer tuning are two implementation optimizations which are recommended by TCPSAT. The use of link forward error correction should make noisy satellite links as error free as possible. Buffer tuning allows efficient use of buffer space in senders, such as World Wide Web servers, which are likely to be supporting simultaneous connections with both large and small windows.

There are many different ways in which proxies may also be used to enhance TCP performance and research is currently ongoing to establish the best mechanism. Traditional proxying (used by web caches and corporate firewalls) involves a user routing requests to a proxy which either responds directly with cached data or fetches the data, forwards it to the user, and caches it for the next query. This mechanism breaks the end-to-end TCP connection into two separate connections. If one of the connections is isolated to the satellite network, then TCP can be modified to run over the satellite network more efficiently.

There are risks involved in the use of proxies which use a modified TCP although research is underway to minimize them. For example, if TCP congestion control is modified to make more efficient use of the satellite link, it will likely be made more aggressive. To protect the Internet from congestion collapse aggressive behavior must be contained within the network doing the proxying. One of the downsides of traditional proxies is that they require users to configure their software to take advantage of them. Hidden proxies have been proposed which modify the connection by looking inside the IP packet for the TCP header information. A variety of performance enhancing mechanisms are possible using this technique but they all are at risk because IP packet encryption will render many mechanisms unworkable by encrypting the TCP header information. There is much more that needs to be explored in the use of hidden proxies and many people in the satellite industry are working on the problem.

Finally, as the TCP protocol evolves the opportunity exists to propose satellite-friendly changes to the protocol. For example, it has been shown earlier that the rate the congestion window opens is based on round trip time. This makes TCP inherently unfair to longer delay connections which take much longer to open their window. Consideration should be given to congestion control mechanisms which are based either on a fixed rate clock or some other mechanism that will remain consistently fair from one environment to another.

Research and Standards Organizations

There remain many research issues to optimize TCP performance over a GEO satellite. As already mentioned, TCPSAT is developing recommendations on standard options and implementations to look for in a TCP for use over satellite. A second document is in development summarizing much of the relevant TCP research going on in the satellite community. However, working groups in the IETF are short lived. In accordance with IETF policy, TCPSAT may dissolve after the completion of the two documents specified above. However, another working group exists to continue as a focal point for exchange of information on research in the area of Internet protocols over satellite.

The Telecommunications Industry Association, Satellite Communications Division, has a working group addressing Internet Protocols over Satellite. In this group satellite specific issues are addressed which are not appropriate for development in the IETF. For example, a TCP/IP Common Air Interface requirements document is to be developed over the next year. This document will specify the characteristics a satellite system should provide to optimize TCP/IP performance. A requirements document on the use of proxies to enhance TCP performance over satellite will also be developed.

TRW's ToBeNamed Digital Satellite Architecture

In order to thoroughly evaluate the issues involved when TCP/IP is used over a satellite system, it is helpful to have a specific system in mind. Many issues which are not apparent when considering a generalized architecture may become critical when the specifics of creating a launchable hardware design are considered. In this section we are describing the TRW ToBeNamed architecture—a satellite network utilizing digital on-board cell switching, forward error correction, and uplink DAMA.

Historically, satellites have been based on what is known as a “bent pipe” architecture. A bent pipe satellite is a simple analog transponder re-broadcasting all the signals received with no changes. During the 1980's satellite designs were developed for NASA and DoD incorporating on-board signal processing giving birth to the first digital payloads. (The payload is the part of the satellite which performs mission specific functions, such as communication, and is distinct from the satellite bus which provides generic, mission independent, functions, such as supplying power and attitude control.)

Digital payloads allow satellite processing to be performed on the binary data—the bits—rather than being limited to the analog waveform. Access to the bits enables many, previously impossible functions to be performed on-board. Processed payload technology has evolved to a state where many of the functions found in terrestrial computers and networks can now be placed on the satellite.

There is tremendous benefit to performing switching on-board a satellite. Bandwidth is a precious resource on satellites and maximizing bandwidth utilization is paramount in an efficient design. Data transmission, Internet data in particular, is bursty often with extremely low duty cycles. For example, the average time between mouse clicks for World Wide Web users is about 15 seconds. At 56kbps this yields a transmission duty cycle of about 200:1 and is worse at higher rates. To make the most efficient use of satellite resources, statistical multiplexing should be used wherever possible. On-board packet switching allows for statistical multiplexing on the downlink. Bursty data from many sources can be aggregated on the payload to increase downlink utilization. Additionally, Demand Assigned Multiple Access (DAMA) allows for statistical multiplexing on the uplink

to the satellite by allowing many users to share a single channel. For client server applications such as web browsing, uplink DAMA can increase uplink utilization from 2% to over 40% (TBR) by sharing the uplink channel among multiple users.

Building a digital payload is not as simple as placing a Pentium computer or a Cisco router on board a satellite. ASICs and RAMs designed for terrestrial use will not survive in the harsh environment found outside the earth's atmosphere. Background radiation, single-event upsets, and latch-up phenomena wreck havoc with commercial off-the-shelf devices. Chips manufactured for use in satellites require special design rules, materials, manufacturing processes, and shielding to operate reliably in the space environment. Satellite designs are constrained by size, weight, and power (SWAP). Radiation shielding consumes significant size and weight. Additionally, terrestrial electronics are often not power efficient enough for use on satellites.

The key way in which a satellite payload design can enhance TCP/IP performance is to provide maximum link bandwidth and utilization while minimizing packet loss. The ToBeNamed digital payload uses:

- Forward error correction to reduce packet retransmission due to corrupted data
- Many uplink channels and uplink DAMA to allow many users to share access to the satellite network
- On-board Early Packet Discard so that partial packets are not clogging up the downlink
- Connectionless IP over cell switching to minimize connection establishment delay
- Sufficient on-board buffering to allow TCP congestion control to maintain stability within the network

The ToBeNamed Architecture

The TRW architecture allows the use of a custom interface between the ground terminals and the satellite while still supporting standard protocols and network interoperability from outside the satellite network. This allows a wide variety of protocols including ATM and TCP/IP to be carried over the network.

ToBeNamed consists of three primary elements: a User Terminal, a geosynchronous Satellite Payload, and a Network Operations Center (see Figure 44 and Figure 55). Other elements include the Satellite Bus and Satellite Management Center but those are not relevant to the networking architecture. Each of the primary elements will be described briefly below.

The ToBeNamed payload receives data from many time and frequency division multiplexed (TDM/FDM) uplinks. Between 32 and 128 beams can be used to cover an area the size of the continental United States. Partitioning the uplink in time and frequency into many channels allows fine granularity of uplink bandwidth allocation. This is important because uplink channels are a finite resource and need to be managed carefully to maximize system capacity. They can either be allocated in a fixed manner, where a single user is assigned a channel for the duration of his call, or in a random access fashion, where multiple users share a single TDM/FDM allocation using DAMA.

While a large number of channels are used to carry uplink data, downlinks are single broadcast data streams to the entire antenna coverage area (300-400 miles in diameter). This efficiently supports user terminals which may transmit low rate data and receive high rate data. Transmission rates as low as 16kbps can be used to keep user terminal costs

low by reducing the size of the transmission power amplifier. More expensive terminals can be designed which transmit at higher, multi-megabit rates.

Both uplink and downlink channels make use of forward error correction (FEC) to protect the data from corruption during transmission. The Ka-band frequencies, for which ToBeNamed was designed, are susceptible to atmospheric interference which can severely corrupt data without adequate FEC.

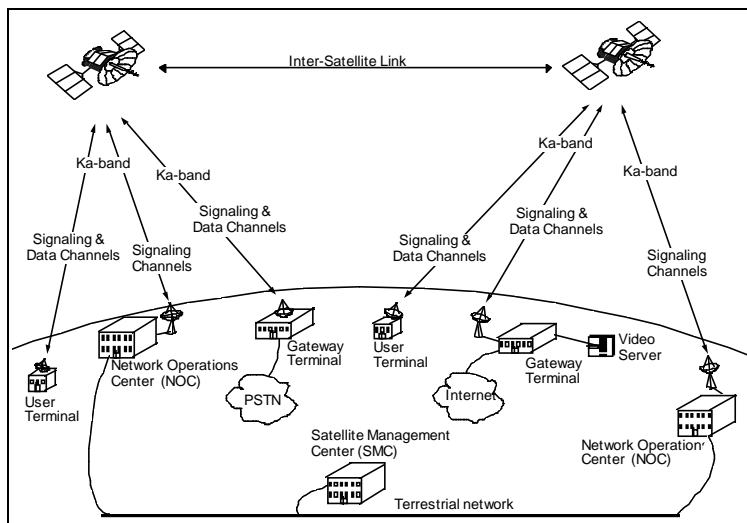


Figure 4. The ToBeNamed architecture allows User Terminals to connect to public networks such as the PSTN or the Internet.

Payload

At its heart the ToBeNamed payload contains a cell switch capable of routing fixed size cells from any uplink beam to any downlink beam. The on-board processing functions include the following:

- **Demodulation:** processing the uplink waveform to recover baseband data/cells
- **Full on-board decode/encode:** enhance TCP throughput by protecting data during transmission using forward error correction on the uplink and downlink
- **Cell switching:** route cells to correct downlink or crosslink based on information in header
- **Early packet discard:** optimize downlink efficiency by discarding partial packets
- **Connectionless IP:** carry IP traffic without requiring connection establishment signaling
- **Multicast:** replicate cells for transmission to multiple downlinks
- **Multiple priorities:** support multiple service classes
- **Statistics collection and signaling:** support switch management functions on ground

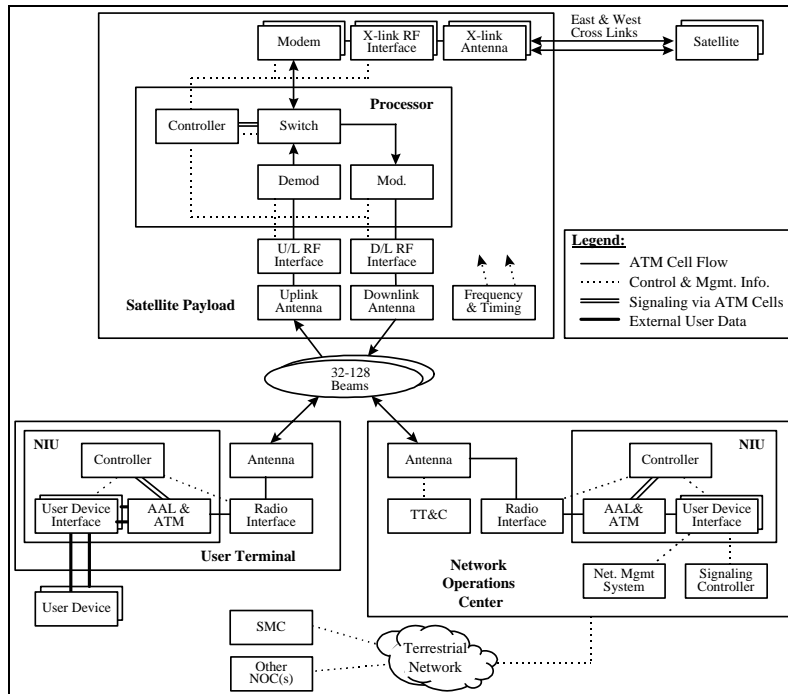


Figure 5. The ToBeNamed architecture has three functional elements: the User Terminal, Satellite Payload, and Network Operations Center. The User Terminals contain a Network Interface Unit providing signaling and adaptation from standard protocols to those consistent with the cell relay architecture. The Network Operations Center manages the satellite payload and performs call admission control when needed. The Satellite Payload routes cells from the source's uplink channel to destination downlink channel.

User Earth Terminal (UET)

The UET combines the ground antenna, modem, and digital processing required to support the air interface to the satellite on the RF side of the terminal and standard network interfaces on the digital side. Proxies may be used to enhance Internet traffic performance. The RF side of the UET may support a distributed DAMA algorithm which will allow multiple users to contend for a single TDM/FDM uplink channel. The digital portion of the UET could be a "set-top box," a plug in card for a home computer; an interface card for a router; or a stand alone box with an ATM, Ethernet, or PPP port. User terminals will likely be produced in a wide range of cost, complexity, and data rates depending on the services offered by the satellite network.

Network Operations Center (NOC)

The NOC provides a central location from which resources can be allocated in the satellite network. The NOC is responsible for management of the on-board switch including collecting connection statistics and configuring and updating routing and scheduling tables as part of connection admission control. The NOC receives all connection related signaling messages from the UETs and generates responses to them as a proxy for the switch. The NOC also supplies usage information for customer billing.

As the elements forming the satellite network, the Satellite Payload, UET and NOC work together to provide a full featured network interface to the user. As the network protocols used over this architecture are analyzed, it is important to understand the ways in which user data will traverse the network.

Conclusion

Satellites offer unique capabilities to extend the Internet. Satellites offer reliable, flexible bandwidth allocation and can be used to extend the Internet into areas of the world where high-bandwidth connectivity would otherwise be unfeasible using other infrastructure.

There are many ways to carry IP over ATM, each with tradeoffs. While appropriate for some limited network topologies, many existing schemes do not scale well in an environment where thousands or millions of network nodes may be connected to the same switch. Connectionless IP over ATM shows promise by reducing call setup time and overhead associated with maintaining idle connections.

TCP performance over satellite is dominated by congestion control. Small file transfers are impacted by Slow Start. Large file transfers are impacted by Congestion Avoidance. Geosynchronous satellite delays aggravate connection performance by adding delay but bandwidths of up to 1Gbps can be achieved with today's TCP.

TCPSAT is making recommendations on several TCP options and existing implementations which can improve TCP performance but some inefficiencies still remain. Research issues include TCP over ATM, impacts of DAMA, how best to use proxies, and how to make the next generation of TCP more fair to long delay connections. TCPSAT and TIA-SCD-CIS-IPoS are the focal points for information on these issues. Information on working group activities can be found at the web sites listed below.

TRW's ToBeNamed architecture can flexibly carry TCP/IP traffic over a space based cell switch. In spite of the challenges involved in designing space-qualified, high performance digital electronics, The ToBeNamed digital payload provides sophisticated functions to enhance system bandwidth and connection performance.

References

[Comer] D. Comer, "Internetworking with TCP/IP; Volume I: Principles, Protocols, and Architecture," Prentice-Hall, 1995

[Quarter] Quarterman, John and Smoot Carl-Mitchell, "What is the Internet, Anyway?," Matrix News, 4(8), August 1994

[Mathis] M. Mathis, J. Semke, J. Mahdavi, T. Ott, "The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm," Computer Communication Review, vol. 27, no. 3, July 1997

[Partridge] C. Partridge, T. Shepard, "TCP/IP Performance over Satellite Links," IEEE Network, September/October 1997

[Romanow] A. Romanow, and S. Floyd, "Dynamics of TCP Traffic over ATM Networks." IEEE JSAC, V. 13 N. 4, May 1995

[RFC-1149] D. Waitzman, "A Standard for the Transmission of IP Datagrams on Avian Carriers; RFC-1149," Internet Requests for Comments, no 1149, April 1990

[RFC-1323] D. Borman, R. Braden, and V. Jacobson, "TCP Extensions for High Performance; RFC-1323," Internet Requests for Comments, no. 1323, May 1992

[RFC-1958] B. Carpenter, "Architectural Principles of the Internet; RFC-1958," Internet Requests for Comments, no. 1958, June 1996

[RFC-2018] M. Mathis, J. Mahdavi, S. Floyd, A. Romanow, "TCP Selective Acknowledgement Options; RFC-2018," Internet Requests for Comments, no. 2018, October 1996.

[Stevens] W. R. Stevens, "TCP/IP Illustrated, Vol. I," Addison-Wesley, 1994.

Interesting Web Sites

TCP over Satellite working group (TCPSAT): Contains pointers to research on TCP and TCP over satellite; working drafts of the group recommendations; and the working group email archive. <http://tcpsat.lerc.nasa.gov/tcpsat>

TCP Performance Enhancing Proxies (TCPPEP): A fledgling site with brief summary of information relating to the use of proxies to enhance TCP performance over satellite. <http://tcppep.lerc.nasa.gov/tcppep>

Pittsburgh Supercomputing Center (PSC): Contains information on research on TCP over high bandwidth and high delay-bandwidth links. Good page on which operating systems are supporting satellite-friendly TCP options. <http://www.psc.edu/networking>

Internet Engineering Task Force (IETF): Contains information on Internet standards development and has pointers to working groups, standards (RFCs), and Internet Research Task Force (IRTF) where next generation protocols and other research issues are discussed. <http://www.ietf.org>

Telecommunications Industry Association, Satellite Communications Division (TIA-SCD): Contains information about standards activities to enhance interoperability between satellite and terrestrial networks. <http://www.tiaonline.org/about/satellite.html>