ATM QoS Experiments Using TCP Applications: Performance of TCP/IP Over ATM in a Variety of Errored Links

Brian D. Frantz
GTE Technology Organization, Cleveland, Ohio

William D. Ivancic
Glenn Research Center, Cleveland, Ohio

January 2001

The NASA STI Program Office . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the Lead Center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA's counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or cosponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized data bases, organizing and publishing research results . . . even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at *http://www.sti.nasa.gov*

- E-mail your question via the Internet to help@sti.nasa.gov

- Fax your question to the NASA Access Help Desk at 301–621–0134

- Telephone the NASA Access Help Desk at 301–621–0390

- Write to:
  NASA Access Help Desk
  NASA Center for AeroSpace Information
  7121 Standard Drive
  Hanover, MD 21076

ATM QoS Experiments Using TCP Applications: Performance of TCP/IP Over ATM in a Variety of Errored Links

Brian D. Frantz
GTE Technology Organization, Cleveland, Ohio

William D. Ivancic
Glenn Research Center, Cleveland, Ohio

Available from

NASA Center for Aerospace Information
7121 Standard Drive
Hanover, MD 21076
Price Code: A03

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22100
Price Code: A03

Available electronically at http://gltrs.grc.nasa.gov/GLTRS

ATM QoS Experiments Using TCP Applications:
Performance of TCP/IP Over ATM in a Variety of Errored Links

Brian D. Frantz
GTE Internetworking
Cleveland, Ohio

William D. Ivancic
National Aeronautics and Space Administration
Glenn Research Center
Cleveland, Ohio

ABSTRACT

Asynchronous Transfer Mode (ATM) Quality of Service (QoS) experiments using the Transmission Control Protocol / Internet Protocol (TCP/IP) were performed for various link delays. The link delay was set to emulate a Wide Area Network (WAN) and a Satellite Link. The purpose of these experiments was to evaluate the ATM QoS requirements for applications that utilize advance TCP/IP protocols implemented with large windows and Selective ACKnowledgements (SACK). The effects of cell error, cell loss, and random bit errors on throughput were reported. The detailed test plan and test results are presented herein.

INTRODUCTION

There is a great amount of interest in understanding the ATM QoS requirements for services such as large data transfers that utilize TCP/IP, particularly over Long Fat Networks (LFNs). The Bandwidth Delay Product (BDP) is defined as transmission rate times the Round Trip Time (RTT). RTT is also called link delay in the paper. Networks with a BDP of larger than 65535 bytes are considered LFNs.

The TCP protocol was designed to be a reliable transport protocol. Acknowledgements are sent from the receiver to the sender upon reception of packets to indicate the transmission was successful. The sender retransmits packets for lost packets due to congestion and or errors. Modifications to the TCP protocol such as the fast retransmit and fast recovery algorithms were implemented to improve the efficiency of TCP in congested and errored environments.[1] Originally the TCP protocol was not designed for LFNs. As a result several extensions have been implemented to improve the efficiency of the protocol. The extensions are the window scale option, the timestamp option, and the protection against wrapped sequence numbers and are described in detail in RFC 1323[2]. The SACK protocol (RFC2018)[3] was designed to handle multiple dropped packets within one window of data.

This paper quantitatively reports the effects of cell error, cell loss, and random bit errors on throughput for advanced TCP/IP over ATM protocols implemented with large windows and SACK.

EXPERIMENT SETUP

The physical setup for the experiment is shown in figure 1. The equipment is listed below:

- Hewlett Packard HP-E4210B Broadband Series Test Equipment
  - OC3 Line Card E1697A
  - Network Impairment Module E4219A
- Adtech AX4000 ATM Traffic Generator and Analyzer
- FORE Systems ASX-200Bx ATM switch
  - One 4 Port Multimode OC3c Netmod (Product Number NM-4/155MMSCC)
  - One 4 Port Singlemode Netmod (Product Number NM-4/155SMSRC)
- Two SUN Ultra II workstations
  - 200 MHz Sparcv9 processor
  - Operating System: Solaris 7 (SunOS 5.7) Generic_106541-02
  - FORE Sun Bus Adapter (SBA-200E) ATM Network Interface Card (NIC)
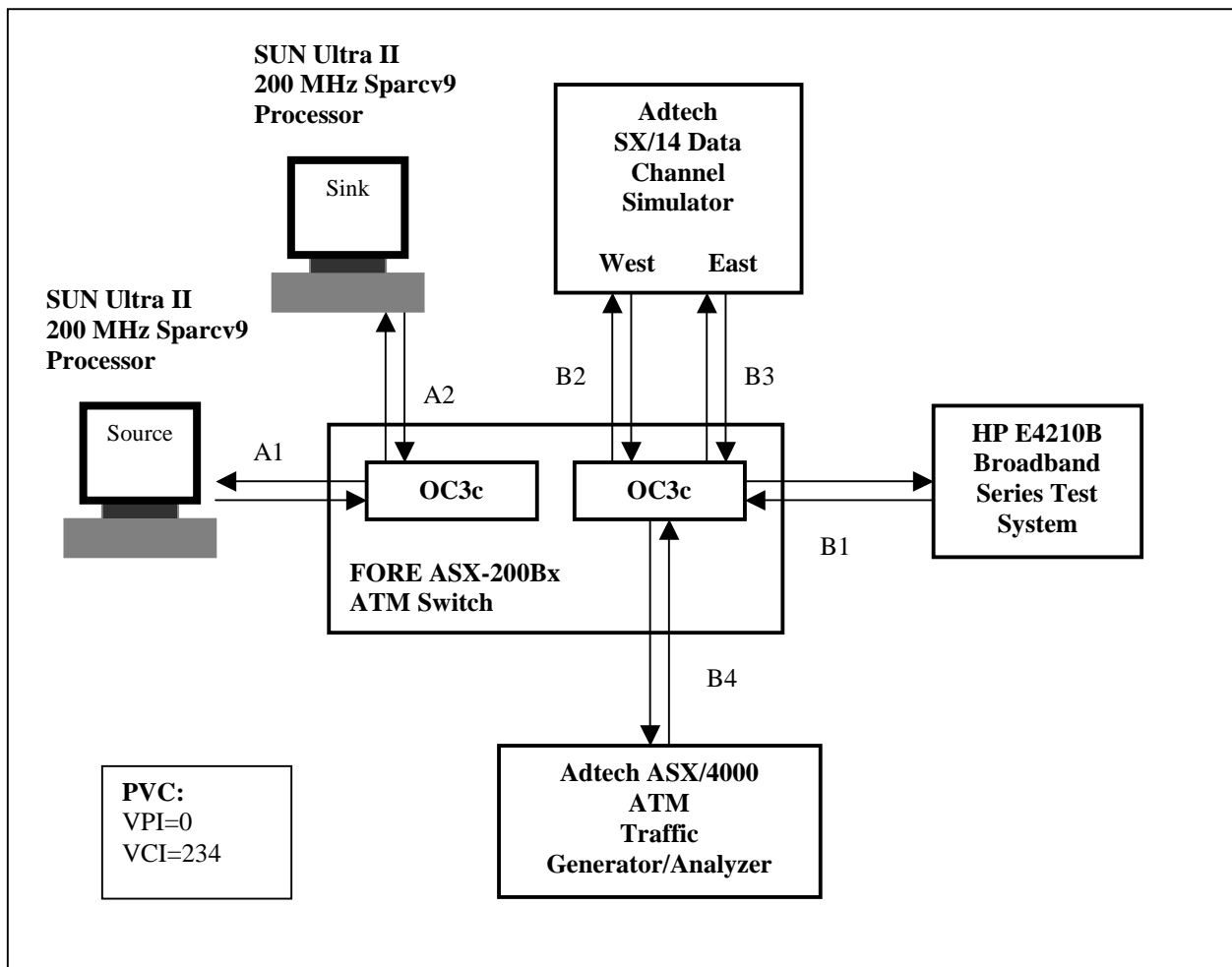   - Driver version: ForeThought_5.0.0.7 (36118)

Figure 1.—Physical Experimental Setup.

EQUIPMENT FUNCTIONS/SETUP

The two SUN ULTRA based workstations were used to generate the large high speed data transfers. The Test TCP (TTCP) application was used to generate the data stream and obtain the resulting throughput. The following shows the TTCP syntax and options:

(NOTE: We have added the –S (Enable SACK, OSF1 Only!) option, used in the initial search for a viable SACK TCP stack, and the –T,-P,-z, and –y options for other experiments. Our modified TTCP source code can be requested by sending email to bfrantz@grc.nasa.gov.

```
ttcp: either -r or -t must be set

Usage: ttcp -t [-options] host [ < in ]
           ttcp -r [-options > out]
Common options:
       -l##     length of bufs read from or written to network (default 8192)
       -u       use UDP instead of TCP
       -p##     port number to send to or listen at (default 5001)
       -s       -t: source a pattern to network
                -r: sink (discard) all data from network
       -A##     align the start of buffers to this modulus (default 16384)
       -O       start buffers at this offset from the modulus (default 0)
       -v       verbose: print more statistics
       -d       set SO_DEBUG socket option
       -b##     set socket buffer size (if supported)
       -w##     set TCP window shift (if supported)
       -S       Enable Selective Acknowledgement (SACK) OSF1 Only!
       -T##     Set TOS field
       -P##     Set Precedence bit (0-7)
       -z##     Set UDP start/stop packet size
       -y##     Set the number of UDP stop packets
Options specific to -t:
       -n##     number of source bufs written to network (default 2048)
       -D       don't buffer TCP writes (sets TCP_NODELAY socket option)
       -N       do not print dots while doing I/O
Options specific to -r:
       -B       for -s, only output full blocks as specified by -l (for TAR)
```

For our testing, the following key options were utilized: window size (-b## buffer size), TCP write packet size (-l## length of bufs read or written to network), and the number of packets generated (-n## number of source bufs written to network).

The Solaris 7 operating system was selected to perform the set of experiments because of the availability of SUN ULTRA workstations and Sun Bus (Sbus) ATM interface cards. A comparison of experimental results for other TCP stacks would have proven valuable, but could not be obtained due to a lack of ATM NIC drivers and stable SACK implementations.

Four TCP parameters were modified so that the TCP stack would support large windows and SACK. The following script was used to setup both workstations prior to testing:

/usr/sbin/ndd -set /dev/tcp tcp_sack_permitted 2
/usr/sbin/ndd -set /dev/tcp tcp_wscale_always 1
/usr/sbin/ndd -set /dev/tcp tcp_max_buf 9100000
/usr/sbin/ndd -set /dev/tcp tcp_cwnd_max 9100000

The tcp_sack_permitted parameter was set to two to initiate and accept connections with SACK options. The tcp_wscale_always parameter was set to one to enable the window scale option. The maximum allowable window size is determined by the minimum value of tcp_max_buf and tcp_cwnd_max. Both of these parameters were set to 910000 bytes, the largest window size needed during the experiments.

The CLassical IP (CLIP) over ATM protocol (RFC1577) was used because of its ability to utilize Permanent Virtual Circuits (PVCs). The PVCs allowed the data and acknowledgements to be explicitly routed through each piece of test equipment. The following commands were used to setup up the CLIP interfaces:

Source Workstation:

```
/etc/fore/clipconfig add -if ci0
/usr/sbin/ifconfig ci0 10.1.1.1 netmask 255.255.255.0 up
/etc/fore/cliparp add -pvc 0 304 llc_routed ci0 10.1.1.2 -reval 0
```

Sink Workstation:

```
/etc/fore/clipconfig add -if ci0
/usr/sbin/ifconfig ci0 10.1.1.2 netmask 255.255.255.0 up
/etc/fore/cliparp add -pvc 0 304 llc_routed ci0 10.1.1.1 -reval 0
```

The clipconfig command created the interface "ci0". The ifconfig command applied the IP address to the new interface. The cliparp command created an IP address (destination host) to PVC mapping.

The HP E4210B is capable of inserting errors at the ATM cell level. Thus, it was used to insert cell loss and cell errors with deterministic and binomial distributions. The Adtech SX/14 is unaware of any higher level protocols like ATM and operated only on the physical network layer. This enabled us to apply a Binomial logic error distribution at the OC3c level. A logic error changes ones to zeroes and zeros to ones. The Adtech AX/4000 was used to verify that all error and delay parameters of the link were properly setup prior to performing an experiment. The FORE ATM switch provided the PVCs used to route the ATM cells through the test equipment. Figure 2 shows the logical flow of data and acknowledgements through each piece of equipment described above. The following PVCs were manually set in the ATM switch:

Data Path PVC descriptions:
- Source workstation input port to HP E4210B output port
- HP E4210B input port to Adtech SX/14 West channel output port
- SX/14 East channel input port to ASX/4000 output port
- ASX/4000 input port to the sink workstation output port

Correspondingly, the FORE switch setup (See figure 1):

```
Configuration vcc new A1 0 234 B1 0 234
Configuration vcc new B1 0 234 B2 0 234
Configuration vcc new B3 0 234 B4 0 234
Configuration vcc new B4 0 234 A2 0 234
```

Acknowledgment Path PVCs

- Sink workstation output port to SX/14 East channel output port
- SX/14 West channel input port to source workstation input port

Correspondingly, the FORE switch setup (See figure 1):

```
Configuration vcc new A2 0 234 B3 0 234
Configuration vcc new B2 0 234 A1 0 234
```
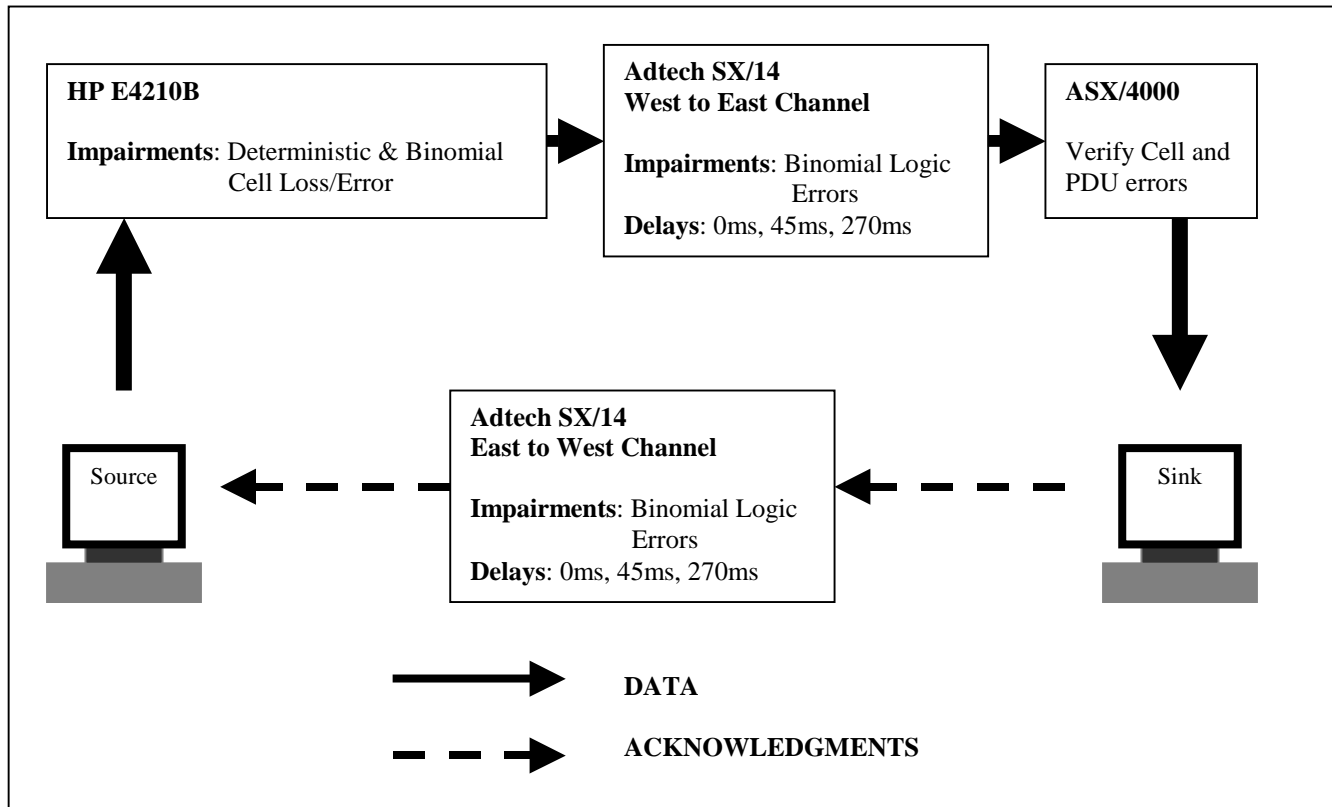

Figure 2.—Logical Experimental Setup.

EXPERIMENT PROCEDURES

Every experiment was configured by the following steps:

1) The SX/14 East and West channel delays were set to one half the required RTT. For a LAN, WAN, or Satellite link, the RTT selected was 0ms, 70ms, and 540ms respectively.
2) Depending on the type of error distribution required, the HP-E4210B or SX/14 error generators were enabled or disabled appropriately.
3) The sink workstation started the TTCP application as follows:

    ttcp –s –r –b[buffer size] –l 9100 –n [#packets]

The buffer size for each link delay was determined by setting the buffer size to a value less than the BDP and increasing the value until maximum throughput was obtained. The optimum buffer size for the LAN, WAN, and Satellite links were 80,000, 1,300,000, and 9,100,100 bytes respectively. A packet size of 9100 was chosen so that fragmentation of packets would not occur, reducing the overall throughput. The packet size derivation is below:

```
Packet Size (Bytes) = Maximum Transmission Unit (MTU) of the interface
                    – TCP header length – IP header length
                    – the maximum size of TCP header options

Packet Size = 9180 bytes (For Classical IP over ATM) – 20bytes – 20bytes – 40bytes
            = 9100 bytes
```

The number of packets were selected so that the packet error ratio converged to a stable value, indicating any major transients were removed.

4) The Adtech AX/4000 statistics were reset.
5) Source workstation initializes a TCP snoop.
6) The source workstation starts the TTCP application as follows:

   ttcp –s –t –b[buffer size] –l 9100 –n [#packets] 10.1.1.2

7) After TTCP data transfer completed, throughput results were recorded.
8) Repeat steps 1 through 7 for various delays and error distributions.

Experiment 1:
Deterministic cell error and cell loss distribution for a LAN, WAN and Satellite link.

Purpose:
To verify congestion control and SACK algorithms were operating correctly and to determine the performance of TCP over ATM over a variety of link delays for lost and errored ATM cells with a deterministic distribution.

Results:
The congestion control and SACK algorithms performed as expected. The throughput results for this experiment are shown in figure 3. The cell loss and cell error curves where nearly identical without any correlation. Even though a cell loss on the last cell of one protocol data unit (PDU)[1] would cause a loss of two PDUs the effect on the throughput was not noticeable due to the interaction of the SACK protocol.
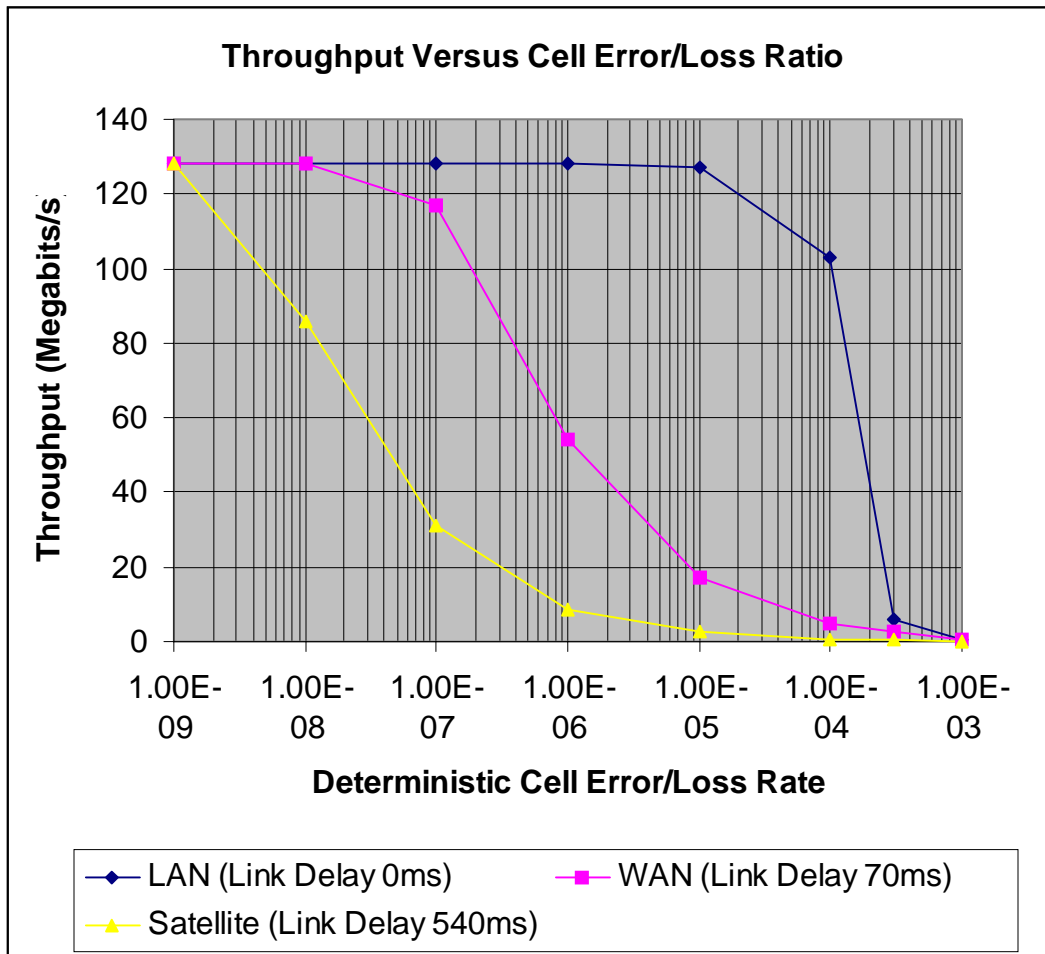

Figure 3

---

[1] An ATM AAL5 PDU is equivalent to one TCP packet. The end of a PDU is determined by a single bit in the last ATM cell for that PDU.

Experiment 2:
Binomial cell error and cell loss distribution for a LAN, WAN and Satellite link.

Purpose:
To determine the performance of TCP over ATM for a variety of link delays for lost and errored ATM cells with a Binomial distribution.

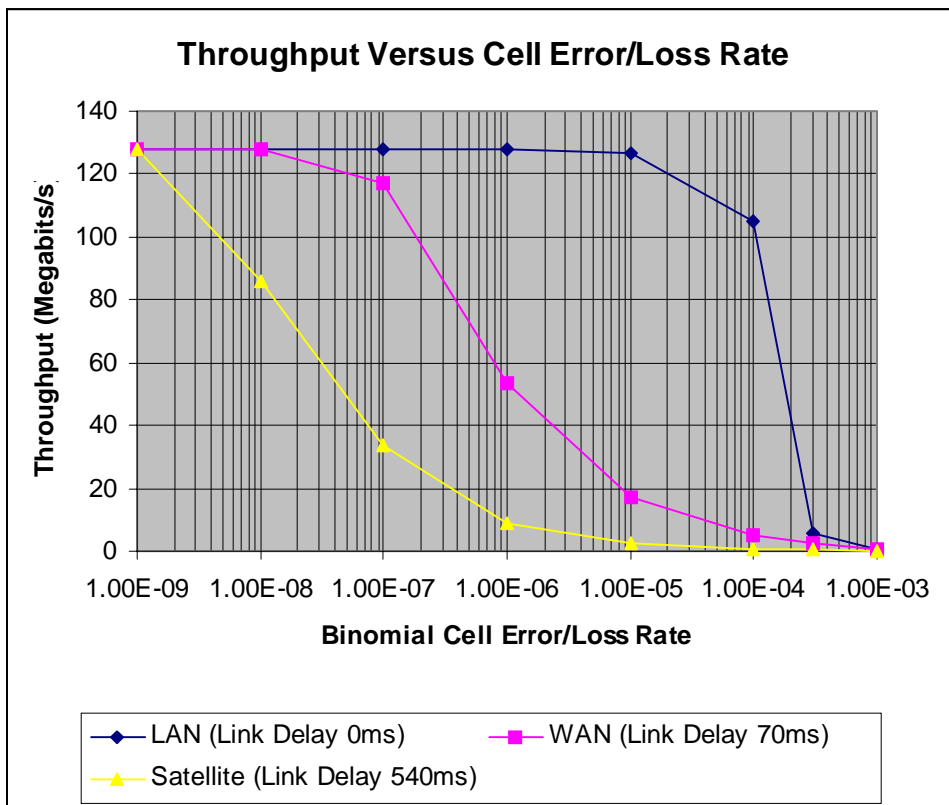## Results:
The throughput results are shown in figure 4.

Figure 4

Experiment 3:
Binomial bit error distribution for a LAN, WAN and Satellite link.

Purpose:
To determine the performance of TCP over ATM for an errored physical OC3c link for a variety of link delays.

Results:
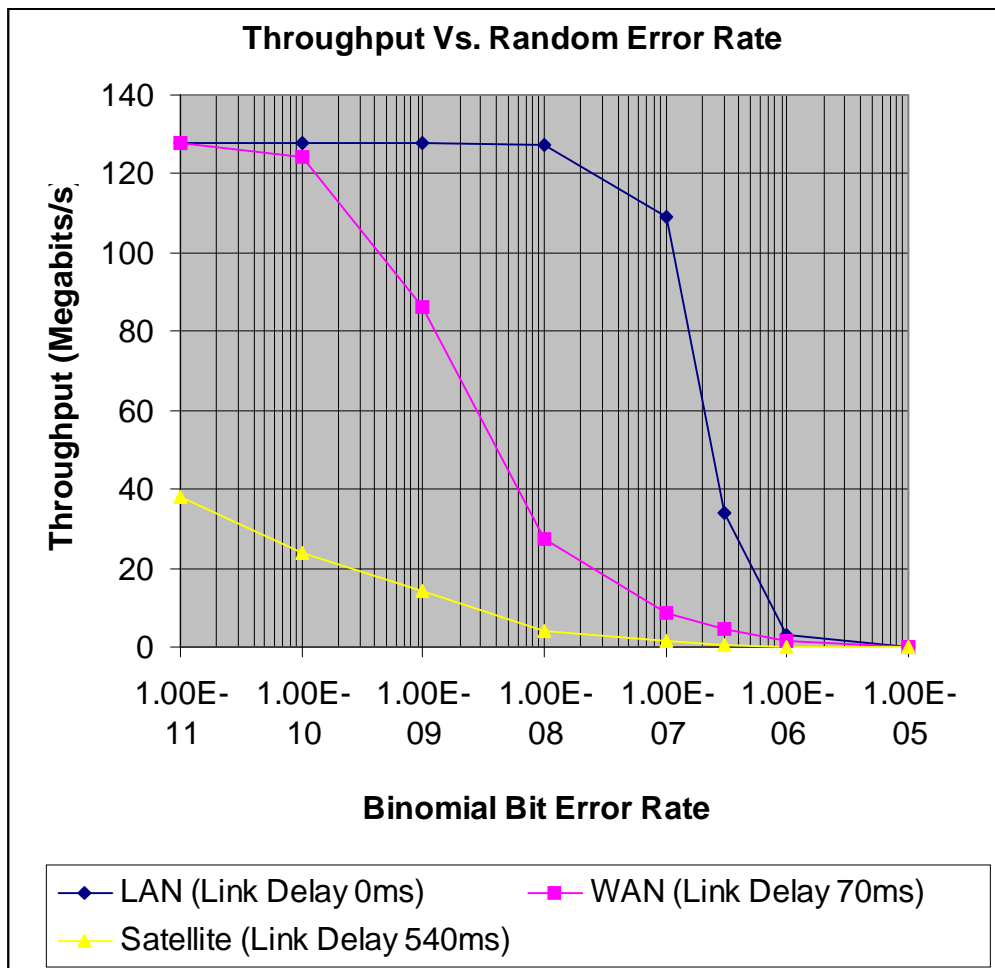The results are shown in figure 5.

Figure 5

CONCLUDING REMARKS

This paper presented the performance of an advanced TCP/IP stack in the presence of several error distributions and link delays. The paper provided a detailed experiment setup and procedure so that the experiments could be duplicated and verified. Parameters such as the packet size and buffer length were chosen to be fixed to reduce the variables to a reasonable amount. Research on the effects of these variables would be valuable. In our experiments an OC3c physical link was utilized. It would be useful to extend this work to include DS3 LFNs. The errors were applied at the ATM cell and network layers due the capabilities of the test equipment. Extending the errors to the physical layer could be accomplished by utilizing commercial satellite modems and injecting guassian noise at the RF level. Finally, the results of these experiment can be used as a foundation to compare how other advanced TCP/IP stacks perform in errored links.

REFERENCES

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

[2] Jacobson, V., Braden, R., and Borman, D., "TCP Extensions for High Performance", RFC 1323, May 1992.

[3] Mathis, M., Mahdavi, J., Floyd, S., and Romanow, A., " TCP Selective Acknowledgement Options", RFC 2018, October 1996.

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	January 2001	Technical Memorandum

4. TITLE AND SUBTITLE

ATM QoS Experiments Using TCP Applications: Performance of TCP/IP Over ATM in a Variety of Errored Links

5. FUNDING NUMBERS

WU–322–20–2A–00

6. AUTHOR(S)

Brian D. Frantz and William D. Ivancic

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

National Aeronautics and Space Administration
John H. Glenn Research Center at Lewis Field
Cleveland, Ohio 44135–3191

8. PERFORMING ORGANIZATION REPORT NUMBER

E–11990

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

National Aeronautics and Space Administration
Washington, DC 20546–0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

NASA TM—2001-209644

11. SUPPLEMENTARY NOTES

Brian D. Frantz, GTE Technology Organization, 21000 Brookpark Rd., M.S. 142–1, Cleveland, Ohio 44135; William D. Ivancic, NASA Glenn Research Center. Responsible person, William D. Ivancic, organization code, 5610 216–433–3494.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Unclassified - Unlimited
Subject Category: 17   Distribution: Nonstandard

Available electronically at http://gltrs.grc.nasa.gov/GLTRS
This publication is available from the NASA Center for AeroSpace Information, 301–621–0390.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Asynchronous Transfer Mode (ATM) Quality of Service (QoS) experiments using the Transmission Control Protocol/Internet Protocol (TCP/IP) were performed for various link delays. The link delay was set to emulate a Wide Area Network (WAN) and a Satellite Link. The purpose of these experiments was to evaluate the ATM QoS requirements for applications that utilize advance TCP/IP protocols implemented with large windows and Selective ACKnowledgements (SACK). The effects of cell error, cell loss, and random bit errors on throughput were reported. The detailed test plan and test results are presented herein.

14. SUBJECT TERMS

Satellite; ATM; QoS; TCP/IP SACK

15. NUMBER OF PAGES

16

16. PRICE CODE

A03

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102