Protocol Overhead in IP/ATM Networks

John David Cavanaugh*
Minnesota Supercomputer Center, Inc.

This paper discusses the sources of protocol overhead in an IP/ATM protocol stack. It quantifies the amount of overhead and makes some observations about its effects. It shows that an application is limited to 87% or less of the line speed in an IP-over-ATM environment, that most of the protocol overhead is due to the lower levels of the protocol stack, and that MTU size can have a significant effect on the amount of protocol overhead.

1.0 Introduction

An application which transfers data across a network can not expect to use the full bandwidth of the communications medium. Along with the application's data, the communications protocols transmit information of their own, which we refer to as *overhead*.

Each protocol layer imposes some overhead in the form of a header and/or trailer which are added to the information received from the next higher protocol layer. Fields in these headers and trailers are used mainly for improving protocol reliability (e.g. HEC; AAL 5 CRC; IP, UDP and TCP checksum) and for multiplexing (e.g. ATM VPI and VCI; IP source and destination addresses; UDP and TCP port numbers). They also carry information such as the length of the data in the protocol data unit (PDU), connection state information, and so forth.

The remainder of this paper discusses the sources and quantities of protocol overhead in an IP/ATM protocol stack over a variety of physical media. We assume that the reader understands the basics of SONET, ATM, and IP, and make no effort to provide a tutorial on these subjects.

This paper was written as part of network research for the MAGIC Gigabit Testbed [5].

2.0 Sources of Protocol Overhead

In this section, we examine each of the layers of an IP/ATM protocol stack and discuss the contribution each makes to protocol overhead.

2.1 Physical Layer

We consider these physical layers:

- SONET OC-3c (155.520 Mbps)
- SONET OC-12c (622.080 Mbps)

* This work was supported by ARPA under Contract Number F19628-92-C-0072.

- Multi-mode fiber at 100 Mbps
- Multi-mode fiber at 140 Mbps
- DS3 (44.736 Mbps)

2.1.1 SONET

SONET is based on the transmission of frames of data. One frame is sent every 125 microseconds, for a total of 8,000 frames per second. The size of the frame depends on the data rate. An OC-1 frame is 810 bytes in length, for a data rate of 51.840 Mbps. Frames for higher data rates are multiples of this basic frame size. An OC-3 or OC-3c frame is 2430 bytes long (9 rows of 270 bytes; see Figure 1), for a data rate of 155.520 Mbps. An OC-12 or OC-12c* frame is 9720 bytes long (9 rows of 1080 bytes), for a data rate of 622.080 Mbps. The OC-12c frame is similar in structure to the OC-3c frame; the replicated overhead bytes (A1, A2, C1, H1, H2, H3, B2) are each repeated twelve times instead of three and each row in the payload is 1044 bytes long instead of 261.

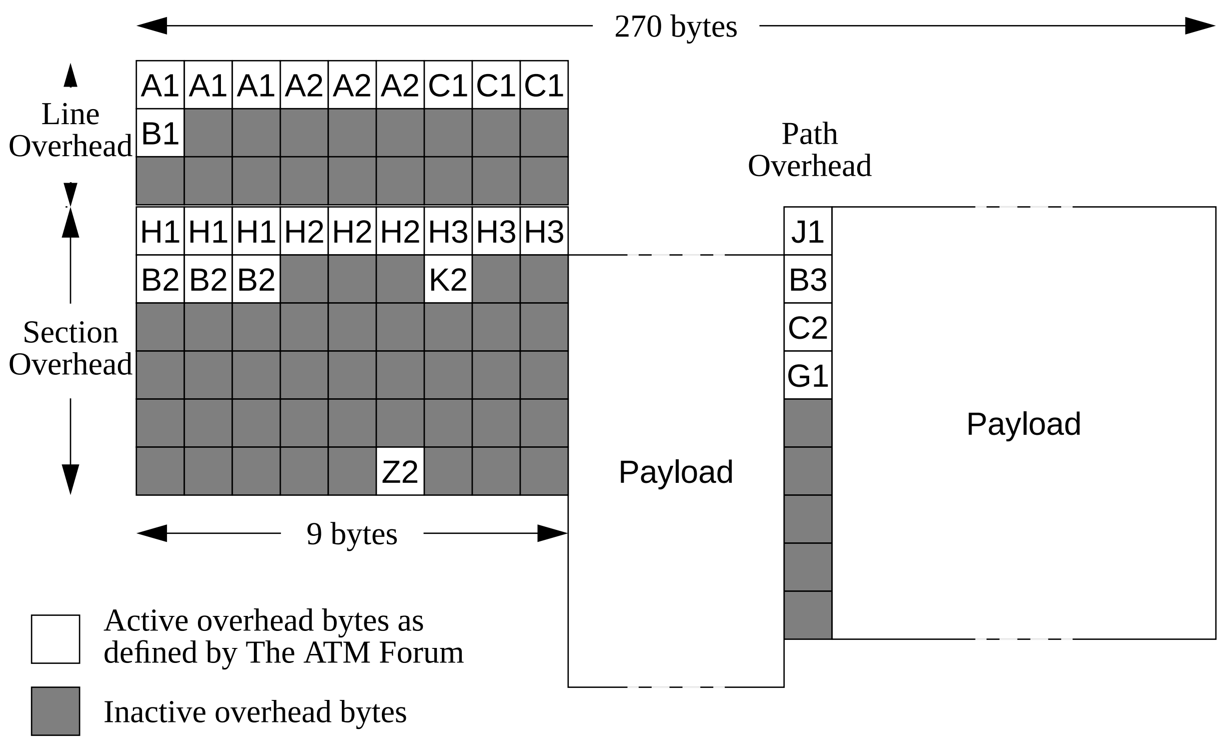

FIGURE 1. SONET OC-3c Frame Format

SONET has three types of overhead: section overhead (SOH), line overhead (LOH) and path overhead (POH). SOH, LOH and POH consume 90 bytes of the 2430-byte OC-3c frame, or 5.760 Mbps of an OC-3c stream, leaving 149.760 Mbps available to the layer above SONET. They consume 333 bytes of an OC-

* [2] defines a physical layer interface for SONET OC-3c, but not for SONET OC-12c. We assume that an OC-12c stream will be used in a manner similar to an OC-3c.

12c frame, or 21.312 Mbps of an OC-12c stream, leaving 600.768 Mbps for the next higher layer. The meanings of the SONET overhead bytes are given in Table 1.

Byte	Coding	Replicated?	Use	Description
A1	1111 0110	Yes	Framing	Flags the beginning of the SONET frame.
A2	0010 1000	Yes	Framing	Flags the beginning of the SONET frame.
B1	BIP-8	No	Section error monitoring	
C1	0000 0001—0000 1100	Yes	SONET identifier byte	Gives the position in the concatenated stream of the corresponding OC-1 stream.
B2	BIP-96	Yes	Line error monitoring	
H1, H2		Yes	Pointer value	H1 and H2 point to the first byte (J1) of the path overhead. Path AIS is indicated by a value of all ones.
H3		Yes	Pointer action	
K2 (bits 6-8)	111	No	Line Alarm Indication Signal (AIS)	
Z2		No	Line Far End Block Error (FEBE)	Contains the B2 error count.
B3	BIP-8	N/A	Path error monitoring	
C2	0001 0011	N/A	Path signal level indicator	
G1 (bits 1-4)		N/A	Path FEBE	Contains the B3 error count.
G1 (bit 5)	0 or 1	N/A	Path yellow alarm	Set when loss of cell delineation is detected.
J1		N/A	STS path trace	Repetitively transmits a 64-byte string for verification by a Path receiving terminal.

TABLE 1. SONET Overhead Bytes

2.1.2 Multi-mode Fiber

The multi-mode fiber interface is based on the physical medium used for FDDI. It uses 62.5 micron optical fiber with a 125 or 175 Mbps line rate and a 4B/5B encoding, yielding bit rates of 100 or 140 Mbps*. The multi-mode fiber interface is sometimes referred to as the TAXI interface, after a chipset used in FDDI (and multi-mode fiber) interface implementations.

* The 140 Mbps version of the multimode interface is not considered a standard interface—it is not mentioned in [2]. We include it here for the sake of completeness.

The 4B/5B encoding of the multi-mode interface provides 32 5-bit codes. Sixteen of these are used to indicate 4-bit nybbles of data. Some of the remaining sixteen are used as command symbols. Three pairs of these command symbols have special significance in the multi-mode interface. They are JK, the *sync* (or *idle*), code, TT, the *start of cell* code, and QQ, the *loss of signal* code.

The overhead of the multimode fiber interface consists of a single *start of cell* code inserted at the start of each cell and a single *idle* code between each pair of cells*. At 100 Mbps, this overhead amounts to 3.636 Mbps, leaving 96.364 Mbps for the next higher layer. At 140 Mbps, it is 5.091 Mbps, leaving 134.909 Mbps for the next higher layer. The format of an ATM cell stream on multi-mode fiber is shown in Figure 2.

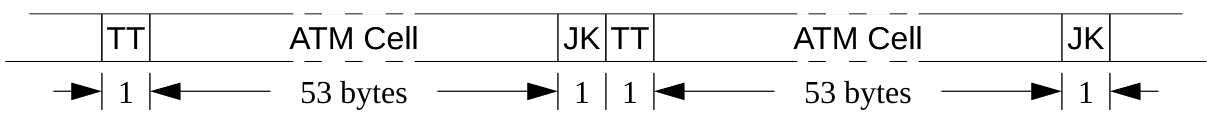

FIGURE 2. Multi-mode Fiber Transmission Format

2.1.3 DS3

The DS3 service is a wide-area digital data service provided by telephony carriers. To adapt ATM cells to the DS3 service, [2] defines a Physical Layer Convergence Protocol (PLCP) for mapping ATM cells into the payload of a DS3 circuit. The format of the PLCP frame is shown in Figure 3.

* There can be fewer idle codes—[2] allows as few as 1 sync every 0.5 second. We assume 1 sync following each cell to be conservative in our calculations.

The line rate of a DS3 circuit is 44.736 Mbps. A frame of data is transmitted every 125 microseconds (8,000 frames per second). A PLCP frame carries 12 ATM cells, providing an effective data rate of 40.704 Mbps to the next higher layer.

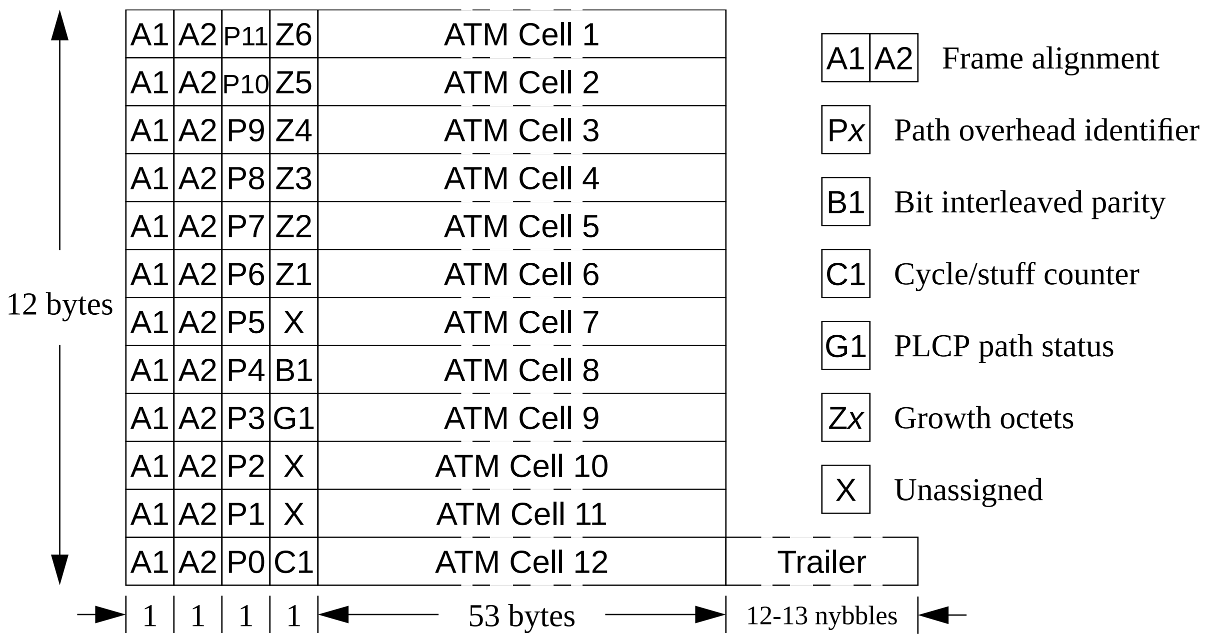

FIGURE 3. DS3 PLCP Frame Format

2.2 ATM

ATM contributes quite a bit to protocol overhead. Each ATM cell is 53 bytes long, comprising a 5-byte header and a 48-byte payload, so nearly 10% of each cell is overhead. Fields in the header provide for multiplexing (virtual paths and channels), traffic type indicators, and a header checksum.

2.3 AAL 5

There are two sources of overhead in AAL 5: protocol overhead in the form of a trailer at the end of the data unit, and the fact that an AAL 5 PDU occupies an integral number of ATM cells.

AAL 5 has an 8-byte trailer which is appended to each PDU. The trailer gives the length of the PDU and contains a checksum which covers the entire PDU. AAL 5 uses an integral number of cells for each PDU. The AAL 5 trailer occupies the last 8 bytes of the last cell. Since AAL 5 PDUs are variable-length (with a maximum of 65,535 bytes), there may be up to 47 bytes of padding. For example, a 576-byte PDU exactly fills 12 ATM cells, leaving no room for the AAL 5 trailer. A 13th cell would therefore be required; it would contain 40 bytes of padding and the 8-byte AAL 5 trailer.

2.4 Protocol Encapsulation

[4] provides for two distinct types of protocol encapsulation for IP over AAL 5. They are *LLC encapsulation*, which calls for LLC, OUI, and Ethertype headers to be precede the IP datagram, and *VC-based multiplexing*, which defines a null encapsulation (an AAL 5 PDU contains only an IP datagram).

We assume the use of LLC encapsulation. The overhead of the LLC, OUI, and Ethertype headers amounts to 8 bytes per datagram.

2.5 IP

Overhead at the IP layer consists of the IP header. Although the header is variable in length, it is almost always 20 bytes long (its minimum length). For the purpose of this paper, we assume that IP always contributes 20 bytes of overhead per datagram.

2.6 TCP

Like IP, TCP's overhead consists of the TCP header. Again, the header is variable-length, but is usually 20 bytes in length. For the purpose of this paper, we assume that TCP always contributes 20 bytes of overhead per segment, and that each TCP segment corresponds to an IP datagram.

2.7 UDP

UDP's overhead consists of an 8-byte header. Fields in the header provided for multiplexing and data integrity.

2.8 RTP

Like UDP, RTP has an 8-byte header. However, RTP can be used in two ways: as a client of IP or as a client of UDP. Since the latter is the case in MAGIC, this paper treats RTP's overhead as 16 bytes, the sum of the UDP and RTP headers.

3.0 Results

Tables 2 through 4 show the amount of bandwidth that remains after each successive layer has claimed its share of the overhead. Each row in the tables shows how much bandwidth is available to the indicated protocol layer. For example, the second row, labeled "To ATM layer," shows how much bandwidth is available to the ATM layer (this number is the line rate less the physical layer overhead). There are columns for SONET OC-3c and OC-12c, multi-mode fiber at 100 and 140 Mbps, and for DS3. Where the maximum transmission unit (MTU*) size is significant, there are subcolumns for the different MTUs. We consider three MTUs: 576 (the Internet inter-network default), 9,180 (the proposed default for IP over ATM), and 65,527 (the maximum for IP over AAL 5).

---

* We consider MTU to be the size of the IP datagram, including the IP header, but not including the LLC/SNAP header or the AAL 5 overhead. The largest IP MTU considered is 65,527 bytes; this is the size of the largest AAL 5 payload less 8 bytes for the LLC/SNAP header.

For an OC-3c stream, the bandwidth available to the application is in the range of 116.504 to 135.498 Mbps. For an OC-12c stream, the range is 467.361 to 543.519 Mbps. In both cases, the range is 75% to 87% of the raw SONET bandwidth.

	OC-3c			**OC-12c**		
Line rate		155.520			622.080	
To ATM		149.760			600.768	
To AAL		135.632			544.092	
	MTU=576	MTU=9,180	MTU=65,527	MTU=576	MTU=9,180	MTU=65,527
To LLC/SNAP	126.937	135.220	135.563	509.214	542.439	543.818
To IP	125.198	135.102	135.547	502.239	541.966	543.752
To transport	120.851	134.808	135.506	484.800	540.786	543.586
To appl. via UDP	119.112	134.690	135.489	477.824	540.313	543.519
To appl. via RTP	117.374	134.572	135.472	470.849	539.841	543.453
To appl. via TCP	116.504	134.513	135.464	467.361	539.605	543.420

TABLE 2. Bandwidth Available after Protocol Overhead (SONET)

For 100 Mbps multi-mode fiber, the bandwidth available to the application is in the range 74.965 to 87.181 Mbps. For 140 Mbps multi-mode fiber, the range is 104.951 to 122.053 Mbps, or 75% to 87% of the raw bandwidth.

	100Mbps			**140Mbps**		
Line rate		100.000			140.000	
To ATM		96.364			134.909	
To AAL		87.273			122.182	
	MTU=576	MTU=9,180	MTU=65,527	MTU=576	MTU=9,180	MTU=65,527
To LLC/SNAP	81.679	87.008	87.229	114.350	121.811	122.120
To IP	80.560	86.932	87.218	112.783	121.704	122.105
To transport	77.763	86.743	87.192	108.867	121.439	122.068
To appl. via UDP	76.644	86.667	87.181	107.301	121.333	122.053
To appl. via RTP	75.525	86.591	87.171	105.734	121.227	122.038
To appl. via TCP	74.965	86.553	87.165	104.951	121.174	122.031

TABLE 3. Bandwidth Available after Protocol Overhead (Multi-mode Fiber)

	DS3		
Line rate		44.736	
To ATM		40.704	
To AAL		36.864	
	MTU=576	MTU=9,180	MTU=65,527
To LLC/SNAP	34.501	36.752	36.845
To IP	34.028	36.720	36.841
To transport	32.847	36.640	36.830
To appl. via UDP	32.374	36.608	36.825
To appl. via RTP	31.902	36.576	36.821
To appl. via TCP	31.665	36.560	36.818

TABLE 4. Bandwidth Available after Protocol Overhead (DS3)

For a DS3 circuit the bandwidth available to the application is in the range 31.665 to 36.825 Mbps. This is 71% to 82% of the raw DS3 bandwidth, which is slightly lower than for either SONET or multi-mode fiber.

Figure 4 shows the effect of MTU size on bandwidth available to the application (via TCP). MTU sizes from 40 to 65,527 are shown. The “notches” in the graph are due to the padding imposed by AAL 5.

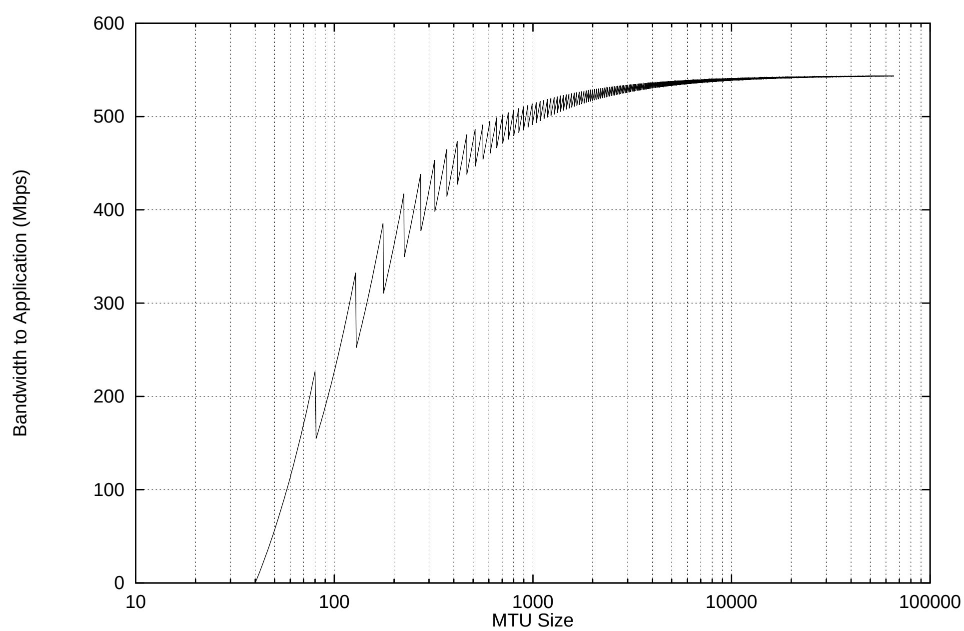

FIGURE 4. Effect of MTU Size on Available Bandwidth

Table 5 shows the percentage of bandwidth which is consumed by overhead when the MTU is 576 (the smallest defensible MTU). The first column for each of the physical media shows the bandwidth available at the various layers as a percentage of the raw bandwidth. The second shows the bandwidth required at each layer in order to provide 100% of the bandwidth required by the application. Note that the figures for SONET and multi-mode fiber are quite similar.

	SONET		Multi-mode fiber		DS3	
	% of line rate	**% of appl. bandwidth**	**%of line rate**	**% of appl. bandwidth**	**% of line rate**	**% of appl. bandwidth**
Physical	100.0	135.1	100.0	133.3	100.0	141.3
ATM	96.6	130.5	96.4	128.5	91.0	128.5
AAL	87.5	118.2	87.3	116.4	82.4	116.4
LLC/SNAP	80.7	109.1	81.7	108.9	77.1	108.9
IP	79.6	107.6	80.6	107.5	76.1	107.5
TCP	76.8	103.8	77.8	103.7	73.4	103.7
Application	74.0	100.0	75.0	100.0	70.8	100.0

TABLE 5. Bandwidth as a Percentage of Line and Application Rates

Table 6 shows what part of the bandwidth available to it is consumed by each protocol layer. Again, the MTU for this table is 576; the numbers for layers above ATM would be considerably lower at a higher MTU.

Protocol Layer	% Overhead
SONET OC-3c	3.70%
SONET OC-12c	3.43%
100 Mbps multi-mode fiber	3.64%
140 Mbps multi-mode fiber	3.64%
DS3 (including PLCP)	9.01%
ATM	9.43%
AAL 5	6.41%
LLC/SNAP	1.37%
IP	3.47%
UDP	1.44%
RTP	2.88%
TCP	3.60%

TABLE 6. Percent Overhead by Protocol Layer

4.0 Observations

- At best, only 87% of the line rate is available to an application in an IP-over-ATM environment.
- Most of the protocol overhead is consumed by the lower layers (see Table 6). The bandwidth available to the AAL layer is only a little over 87% of the line rate.
- The size of the data unit transmitted has more impact than the choice of transport protocol on the degree of protocol overhead.

Acronyms

AAL	ATM Adaptation Layer
AIS	Alarm Indication Signal
ATM	Asynchronous Transfer Mode
BIP	Bit Interleaved Parity
CRC	Cyclic Redundancy Check
DS3	Digital Signal 3
DoD	Department of Defense
FDDI	Fiber Distributed Data Interface
FEBE	Far End Block Error
HEC	Header Error Control
IP	Internet Protocol
Kbps	Kilobits per second
LLC	Logical Link Control
LOH	Line Overhead
Mbps	Megabits per second
MTU	Maximum Transmission Unit
OC-*n*	Optical Carrier Level *n*
OC-3c	Optical Carrier Level 3 Concatenated
OC-12c	Optical Carrier Level 12 Concatenated
OUI	Organizationally Unique Identifier
PDU	Protocol Data Unit
PLCP	Physical Layer Convergence Protocol
POH	Path Overhead
RFC	Request for Comments
RTP	Real-time Transport Protocol
SOH	Section Overhead
SONET	Synchronous Optical Network
STS	Synchronous Transport Signal
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
UNI	User–Network Interface
VC	Virtual Channel
VCI	Virtual Channel Identifier
VPI	Virtual Path Identifier

References

[1] Atkinson, Randall J. *Default IP MTU for use over ATM AAL*. RFC 1626. May 1994.

[2] The ATM Forum. *ATM User–Network Interface Specification, Version 3.0*. Englewood Cliffs, NJ: Prentice Hall, 1993.

[3] Cavanaugh, J. D. and T. J. Salo. “Internetworking with ATM WANs,” Minnesota Supercomputer Center Technical Report, Minneapolis, MN, December 14, 1992. Reprinted in Stallings, William, *Advances in Local and Metropolitan Area Networks*, Los Alamitos, California: IEEE Computer Society Press, 1994.

[4] FA-NWT-001109. *Broadband ISDN Transport Network Elements Framework Generic Criteria*. Bellcore, Issue 1, December 1990.

[5] Fuller, B. B., and I. Richer, “An Overview of the MAGIC Project,” M93B0000173, The MITRE Corp., Bedford, MA, December 1, 1993.

[6] Heinanen, Juha. *Multiprotocol Encapsulation over ATM Adaptation Layer 5*. RFC 1483, July 1993.

[7] Postel, J., Ed. *DoD Standard Internet Protocol*. RFC 760, January 1980.

[8] Postel, J., Ed. *DoD Standard Transmission Control Protocol*. RFC 761, January 1980.

[9] Postel, J. *User Datagram Protocol*. RFC 768, August 1980.

[10] Schulzrinne, H., and S. Casner. *RTP: A Real-Time Transport Protocol*. Internet Draft, July 30, 1993.

[11] TR-NWT-000253. *Synchronous Optical Network (SONET) Transport Systems: Common Generic Criteria.* Bellcore, Issue 2, December 1991.