PERFORMANCE IMPROVEMENTS FOR ISO TRANSPORT

RICHARD COLELLA - ROBERT ARONOFF - KEVIN MILLS

NATIONAL BUREAU OF STANDARDS
Building 225, Room B217
Gaithersburg, Maryland 20899

ABSTRACT

The NBS Protocol Performance Laboratory is developing enhanced protocol mechanisms for OSI class 4 transport that will improve the throughput efficiency achieved on a satellite channel. A selective acknowledgement mechanism has been shown to improve throughput efficiency by as much as 34%. Several alternative expedited data mechanisms have demonstrated throughput efficiency improvements as great as 38%. Most of the protocol mechanism enhancements considered require only minor changes to the international standard OSI transport protocol.

INTRODUCTION

The ISO transport protocols[3] were designed for effective and efficient communications over a set of network services available from public data networks (PDNs), local area networks (LANs), and combinations of PDNs and LANs; however, satellite networks have not been completely accomodated. This paper proposes and evaluates two minor, optional changes to the ISO class 4 transport protocol that provide substantial performance improvements over transmission paths exhibiting long propagation delay, high bandwidth, and unstable error properties.

First, an option permitting more flexible transmission of expedited data is considered. Use of this variant of expedited data is shown to provide throughput efficiency improvements of up to 38% and delay improvements of up to 61% when compared with the standard ISO expedited mechanism.

Second, an optional selective acknowledgement capability is examined. Use of this optional mechanism is demonstrated to yield throughput efficiency increases of up to 34% when compared with a standard one-for-one acknowledgement scheme; however, the improvement is realized at a cost of increased receive buffer utilization.

Both the optional expedited data modification and the selective acknowledgement procedures can be added to the ISO class 4 transport protocol such that implementations of the protocol that do not include the options will be completely interoperable with implementations including the options. Thus, these improvements will increase the range of network media over which the ISO transport protocols can operate efficiently without sacrificing the effectiveness of the protocols.

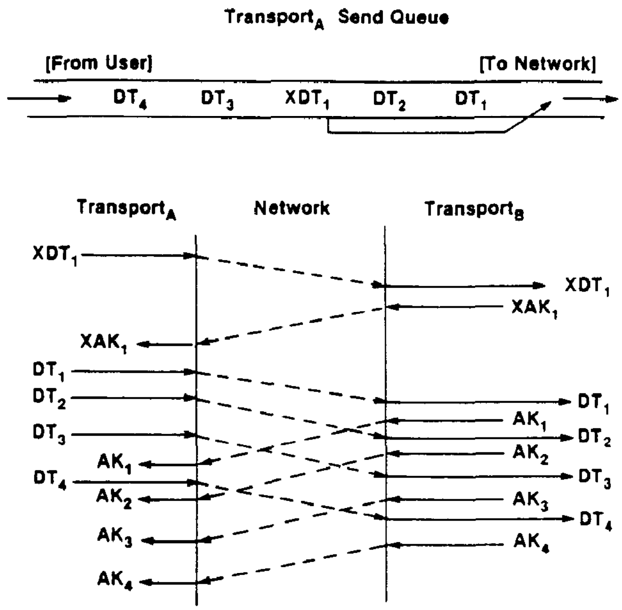
ISO EXPEDITED DATA MECHANISM

Two forms of the expedited data mechanism are discussed and contrasted here. The first mechanism is that currently specified in the ISO transport protocol specification. The ISO expedited data mechanism specifically precludes sending any new normal data after transmitting an expedited data message until the expedited acknowledgement is received (although retransmission of previously sent data is permitted)[3].

Figure 1 shows a time graph representation of how the ISO expedited data mechanism operates. As expedited data is received from the user it is promoted to the front of the send queue ahead of all normal data. It is then sent by $Transport_A$ to $Transport_B$ which acknowledges the expedited data message, XDT_1, when received satisfactorily. Not until this acknowledgement, XAK_1, is received by $Transport_A$ can $Transport_A$ resume sending normal data.

Thus, as can be seen in Figure 1 where the time dimension progresses down the page, the transport connection is idle between the time of initial sending of the expedited data message, XDT_1, and the time of receipt of the expedited acknowledgement, XAK_1, by $Transport_A$. It is only at that point that normal data transmission can resume. Furthermore, were there additional expedited data messages to be sent, there would be additional delay in the resumption of normal data transmission directly proportional to the number of expedited data messages and the propagation delay on the communications network.

From this brief analysis of the ISO expedited data mechanism, there appears to be an undesirable throttling effect which occurs because all normal data transmissions are suspended until the expedited data message is acknowledged. Thus, the transport connection is idle for the period of time required to send the expedited data and receive its acknowledgement.



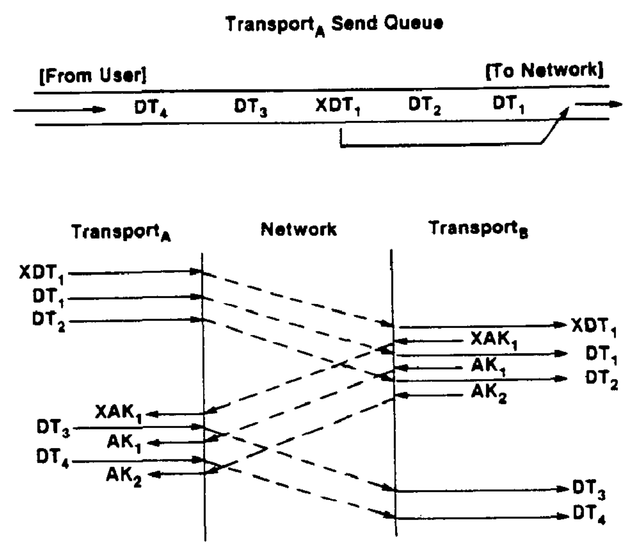
Message Sequencing for ISO Expedited Data Mechanism.
Figure 1

NBS EXPEDITED DATA MECHANISM

The second expedited data mechanism is proposed as part of the NBS transport specification[1], which also satisfies the ISO service specification[4]. The NBS expedited data mechanism may reduce or eliminate the transport connection idle time incurred by the ISO scheme during expedited data transfer. The reduction is accomplished by allowing some normal data to be transmitted while the expedited data acknowledgement is awaited. One restriction applies: the normal data must have been received from the sending user before the expedited data. This satisfies the NBS and ISO transport service specifications [2,4] which guarantee that expedited data received from the user by a transport entity will be delivered to the receiving entity's user before any subsequent normal data received from the sending user. This is achieved because the sending transport entity does not transmit any normal data received from the user after the expedited data until the expedited data has been acknowledged, although data already sent may be retransmitted.

The distinction between the ISO and NBS schemes lies in the signficance attached to the order of data received from the user. With the NBS mechanism, normal data received from the user before the expedited data may be transmitted while awaiting the expedited data acknowledgement (this assumes that expedited data will be placed ahead of normal data in the send queue). With the ISO scheme, the order of receipt from the user is ignored; no data (except retransmissions) may be sent while awaiting an expedited data acknowledgement. The projected improvement in performance for the NBS expedited data mechanism derives from allowing transmission of normal data which had been received from the user up to the time of receipt of the expedited data from the user. By so doing, the normal data does not suffer the full stop-and-wait delay to which the ISO mechanism is subject and which results in idle time on the transport connection. As noted above, this idle time is equal to the difference between the time expedited data is transmitted and the time of receipt of the expedited acknowlegement.

Figure 2 presents a time graph representation of how the NBS expedited data mechanism operates. From this figure we can see the transport connection is being used to send normal data during the time between sending expedited data, XDT_1, and receiving the related expedited acknowledgement XAK_1. More specifically, we can postulate the ideal situation for this mechanism to be one in which there is sufficient normal data in the send queue prior to arrival of the expedited data such that the sending of normal data fills the entire time between sending the expedited data message and receiving the acknowledgement. With such a situation, the NBS expedited data mechanism allows no idle transport connection time.



Message Sequencing for NBS Expedited Data Mechanism
Figure 2

This expedited mechanism is easily implemented, requiring no changes in message formats. Normal data are marked with the most recent expedited sequence number, preserving their order of receipt from the user with respect to the expedited data. The transport entity can then determine which normal data is allowed to be transmitted based upon the status of expedited data transmission.

This mechanism, furthermore, can be implemented unilaterally, placing no restriction on interoperation with implementations based on the ISO mechanism. The receiving transport entity assumes that all normal data received in sequence can be delivered. Thus, the receiver relies on the sender to transmit only what is deliverable. It is transparent to the receiver which of the two mechanisms is being used by the sender.

ONE-FOR-ONE ACKNOWLEDGEMENT

The second protocol mechanism studied is the acknowledgement mechanism. In general, an acknowledgement and retransmission mechanism determines the scheme for confirming data that has been received so that the sender can determine which normal data messages, if any, must be retransmitted. Two such schemes are discussed here.

First, consider the acknowledgement and retransmission mechanism scheme currently specified in the NBS transport protocol specification. This is a one-for-one acknowledgement scheme. In a one-for-one acknowledgement and retransmission mechanism, an acknowledgement message is sent for every normal data message received.

According to both the ISO and NBS protocol specifications, the sequence number in the acknowledgement confirms all normal data messages with lower sequence number. If there is a gap in the sequence numbers of received data messages, only those before the gap can be acknowledged. Thus, when a data message is acknowledged, all other messages with lower sequence numbers are also acknowledged. Normal messages with sequence numbers after a gap will be kept by the receiver, if buffer space is available, but cannot be acknowledged until the missing messages are received. Retransmission of the missing data occurs when the retransmission timer at the sender expires.

SELECTIVE ACKNOWLEDGEMENT AND RETRANSMISSION MECHANISM

In order to try to obtain improved performance, we have proposed a variation of the one-for-one acknowledgement mechanism, based upon the selective acknowledgement concept. The selective acknowledgement messages contain one additional optional field not found in the one-for-one acknowledgement messages (Figure 3). When present, this field contains a sequence number for a normal data message which cannot be represented by the regular sequence number field in the acknowledgement message because of an intervening gap in the received data message stream.

For example, assume Transport_A transmits normal data messages with sequence numbers one through five. When message number one arrives at Transport_B, an acknowledgement is returned confirming all messages up to and including message one using the normal format shown in Figure 3. Similarly, when message two arrives the acknowledgement confirms message two and all previous messages. If message three is lost in the network and message number four now arrives at Transport_B, a selective acknowledgement is returned (Figure 3) where the sequence number confirms messages numbered two and below, as before, and the optional sequence number is set to four, selectively acknowledging that individual message. A similar acknowledgement is returned when message number five is received. Eventually, the retransmission timer for message three will expire and message three will be resent by Transport_A. Correct receipt at Transport_B will allow a normal acknowledgement message to be returned for all messages up to five, inclusive.

Using selective acknowledgement, improved performance of the transport protocol may be realized in two ways. First, data messages correctly received, but after a gap, are now acknowledged immediately, allowing the sending transport to release, for other uses, the buffer space occupied by those data messages. Thus, this mechanism may promote more efficient buffer utilization at the sending transport entity.

The second way that selective acknowledgement may improve performance over the one-for-one mechanism is by eliminating unnecessary retransmissions. If a normal data message is lost or damaged in transmission, the retransmission timer for that message will eventually expire causing it to be sent again. However, using the one-for-one scheme, before the retransmitted data message can be received and all received messages confirmed, some number of timers for messages subsequent to the one in error will expire, causing retransmission of correctly received (but unacknowledged) data. These unnecessary retransmissions are eliminated using selective acknowledgement; after a gap at the receiver, data received correctly are confirmed individually using the optional field in the selective acknowledgement message. Transmission bandwidth, therefore, should be used more efficiently.

EXPERIMENTAL APPROACH

The experiments reported in this paper investigate class 4 transport performance for bulk data transfer with particular emphasis on comparing alternative internal mechanisms of the protocol. In order to evaluate and obtain a quantitative comparison of the relative performance characteristics of the contrasted transport mechanisms described above, we used a simulation model of class 4 transport[5], developed at the NBS, to run experiments and generate data for the performance metrics described in Table 1[6].

Acknowledgement Formats

Normal (One-For-One) Acknowledgement

Octet	1	2	3 4	5 6 7 8	9 10
	Length Indication	AK	Destination Reference	Sequence Number	Credit

Selective Acknowledgement

Octet	1	2	3 4	5 6 7 8	9 10	11	12	13 14 15 16
	Length Indication	AK	Destination Reference	Sequence Number	Credit	Parameter Code (Opt. Seq. #)	Parameter Length (4)	Optional Sequence Number

Comparison of Acknowledgement Message Formats Used in the One-For-One and Selective Acknowledgement Mechanisms
Figure 3

TABLE 1

PERFORMANCE METRICS USED IN EXPERIMENTS

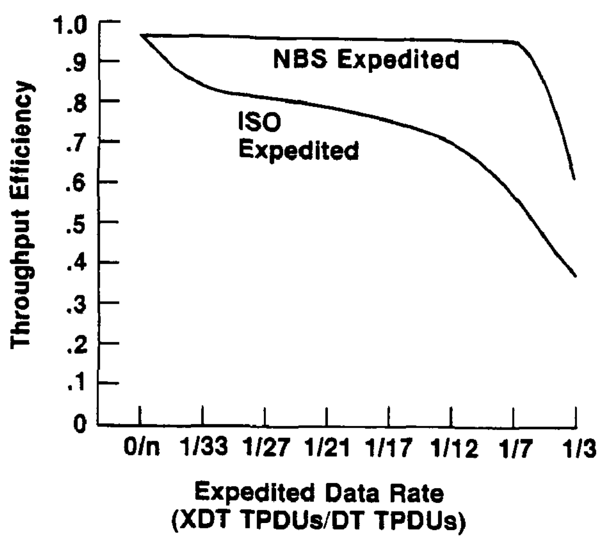
Throughput Efficiency	The ratio of user throughput to maximum theoretical throughput. User throughput is the data transfer rate, in octets/sec., as perceived by the user. It is calculated by dividing the amount of user data transferred by the transfer time. Maximum theoretical throughput is the subnetwork access link speed expressed in octets/sec.
Average Per Octet Delay	The delay, as seen by the users, between transmission and receipt of a TSDU (transport service data unit) divided by the number of octets in the TSDU, averaged over all TSDUs sent over the connection. TSDU is the logical unit for data exchange by the users of transport. A TSDU may be broken into multiple normal data messages by the transport entity.
Average Memory Utilization	The average amount of buffer memory used by a transport entity during an experiment.
Maximum Memory Utilization	The maximum amount of buffer memory used by a transport entity during an experiment.

EXPERIMENT 1 (EXPEDITED DATA) -- DESCRIPTION

Experiment 1 explores the performance differences between the two previously described expedited data mechanisms. Table 2 lists the parameters and their values which describe this experiment. The independent variable in this experiment is the expedited data rate. The expedited data rate is the rate of occurrence of expedited data requests per normal data requests at the user interface. The decimal values listed in Table 2 for expedited data rate represent the specific ratios used (i.e., the number of expedited messages divided by the number of normal messages). This variable defines the x-axis of the results graphs. The y-axis represents the particular metric (throughput efficiency or average per octet delay) being used. This experiment is run for each of the two expedited data mechanisms, ISO and NBS.

EXPERIMENT 1 (EXPEDITED DATA) -- RESULTS

The simulation results indicate that the NBS and ISO expedited data mechanisms exhibit very different performance characteristics. Figure 4 maps throughput efficiency as a function of expedited data rate for the NBS and ISO mechanisms. The reason ISO throughput efficiency is lower than NBS efficiency is due to the restriction placed on normal data transmission while awaiting the expedited data acknowledgement.



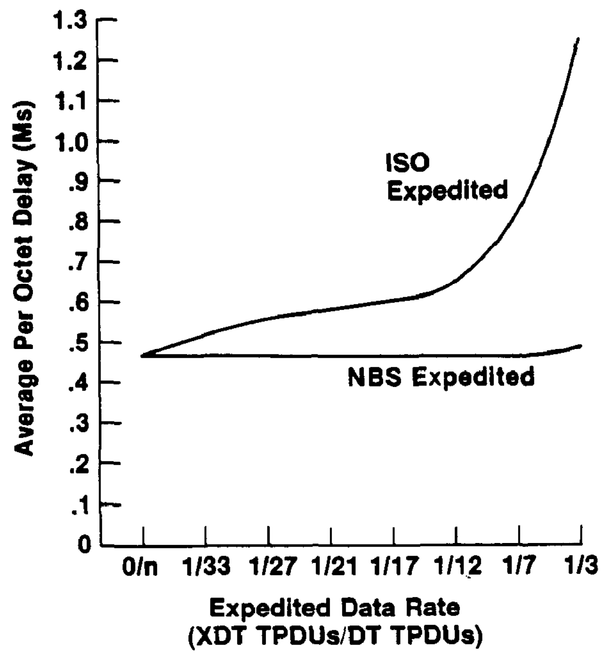
Throughput Efficiency as a Function of Expedited Data Rate for NBS and ISO Expedited Mechanisms
Figure 4

Regardless of the order in which data was received from the user, no normal data may be transmitted until an outstanding expedited message is acknowledged. For each expedited message transmitted, this forces a fixed amount of idle time on the transport connection equal to the round trip time.

The NBS scheme allows normal data received from the user before expedited data to be transmitted after the expedited message is sent, yet before it is confirmed. The idle time on the link is thereby greatly reduced or eliminated. At rates between 0/n and 1/7, any time on the transport connection between transmission and confirmation of expedited data is consumed with normal data transmission. The result is an expeditious transfer of all data.

The throughput efficiency curve for NBS expedited drops significantly as the expedited data rate increases from 1/7 to 1/3. In this region, transport is not receiving sufficient normal data before each expedited message to keep the transport connection fully utilized, resulting in idle time on the connection and lowered throughput efficiency. In effect, the expedited mechanism is imposing an artificial window on normal data flow.

The cost of the ISO mechanism in throughput efficiency is also reflected in delay (Figure 5). Increasing the rate of expedited data causes a concomitant increase in interference in the normal data; the data transfer is stretched out over a longer period of time.



Average Per Octet Delay as a Function of Expedited Data Rate for NBS and ISO Expedited Mechanisms
Figure 5

EXPERIMENT 2 (ACKNOWLEDGEMENT MECHANISMS) -- DESCRIPTION

Table 3 lists the parameters and values describing this experiment, which compares the performance of two acknowledgement schemes. The independent variable in this experiment is the bit error rate. The bit error rate is the ratio of the number of

TABLE 2

EXPERIMENT 1: EXPEDITED DATA MECHANISMS

Experiment Parameters

TRAFFIC CHARACTERISTICS	
No. of Transport Connections	1
Distribution of Arrrivals	Continuously Queued
Directionality	Simplex
TSDU Size	10240 Octets
No. of TSDUS to Send	60
NETWORK CHARACTERISTICS	
Network Window (MSGS)	4
Satellite Link Speed	64 Kbps
Bit Error Rate	10E-9
TRANSPORT CHARACTERISTICS	
Send Buffer Size (Octets)	32K
RCV Buffer Size (Octets)	32K
Window Size (Octets)	12K
Retransmission Timer	2000 ms
Window Timer	6000 ms
Maximum Ack. Delay Timer	Infinite
DT TDPU Size	1024
Expedited Data Rate (expedited/normal)	.33, .14, .08, .0588, .0455, .037, .03, None
Acknowledgement Mechanism	One-For-One
Expedited Mechanism	ISO; NBS

TABLE 3

EXPERIMENT 2: NBS ACK. VS. SELECTIVE ACK.

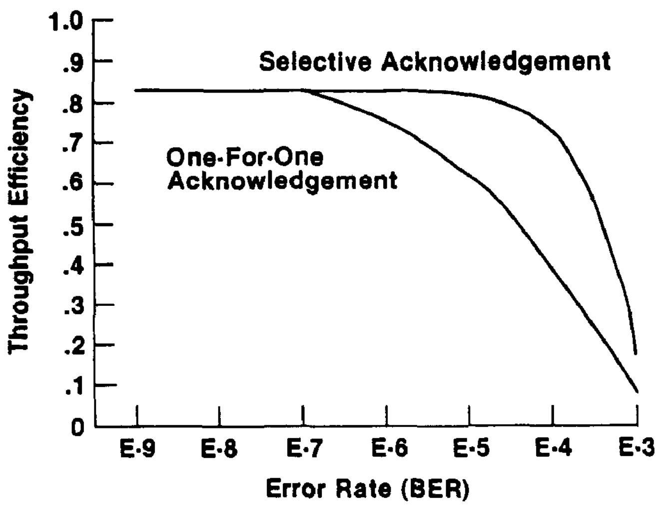
Experiment Parameters

Traffic Characteristics	
No. of Transport Connections	1
Distribution of Arrivals	Continuously Queued
Directionality	Simplex
TSDU Size	10240 Octets
No. of TSDUS to Send	20 (5 at 10E-3 Error Rate)
Network Characteristics	
Network Window (MSGS)	4
Satellite Link Speed	64 Kpbs
Bit Error Rate	10E-3, 10E-4, 10E-5, 10E-6,
Transport Characteristics	
Send Buffer Size (Octets)	32K
RCV Buffer Size (Octets)	32K
Window Size (Octets)	20K
Retransmission Timer	650 ms
Window Timer	2000 ms
Maximum Ack. Delay Timer	Infinite
DT TPDU Size	128 Octets
Expedited Data Rate	0
Acknowledgement Mechanism	One-for-One; Selective
Expedited Mechanism	NBS

bits received incorrectly to the total number of bits received. For example, a bit error rate of 10E-6 indicates that, on average, a bit will be received incorrectly once in one million bits. This variable defines the x-axis of the results graphs. The y-axis represents the particular metric (throughput efficiency, average and maximum receive memory utilization) being used. This experiment is run for each of the two acknowledgement schemes, one-for-one and selective.

EXPERIMENT 2 (ACKNOWLEDGEMENT MECHANISMS) -- RESULTS

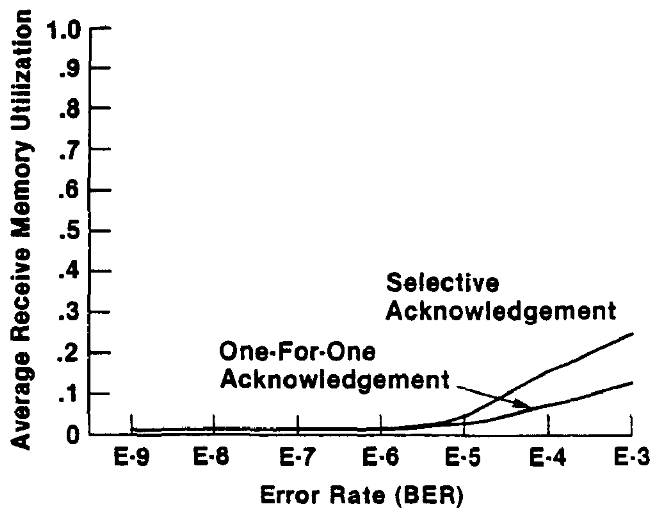
This experiment examines the performance of transport using two different acknowledgement schemes as described earlier, selective and one-for-one. As shown in Figure 6, the benefit of using selective acknowledgement is increased throughput efficiency for the medium to high error rates (E-7 to E-3). The increase is due to a reduction in the number of retransmissions as was anticipated.



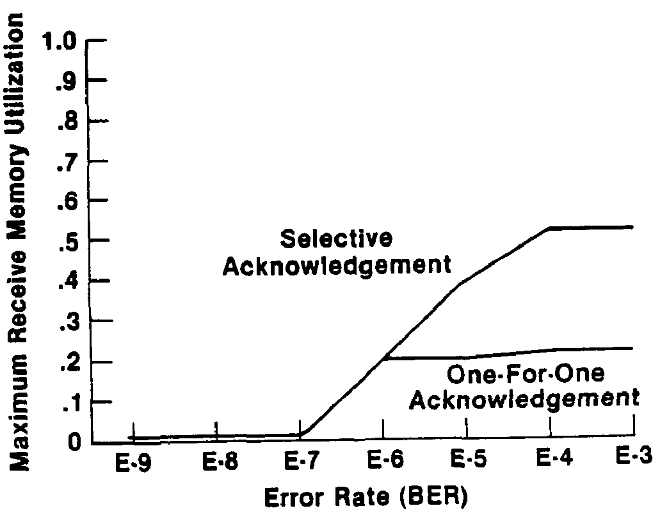
Throughput Efficiency as a Function of Error Rate for One-For-One and Selective Acknowledgement
Figure 6

As one should expect, there is a cost associated with the performance improvement discussed above. In order to maintain a continuous flow of original data messages using selective acknowledgement, the transmit window must be set to twice that required by the one-for-one mechanism. Under the one-for-one scheme, once one retransmission timer expires, a large number will subsequently expire due to the long delay experienced with satellite communications; hence, a major portion of the satellite link bandwidth is consumed by the retransmitted data. Selective acknowledgement, however, eliminates most retransmissions, allowing further original data messages to be sent while awaiting the acknowledgement. The amount of additional data sent is approximately equal to the window size used in the one-for-one mechanism. Therefore, selective acknowledgement requires twice the window size to maintain continuous data flow.

The effect on average and maximum receive memory utilization is shown in Figures 7 and 8. The increase in memory utilization is due to the fact that fewer retransmissions (and therefore more original transmissions) arrive at the receiver per unit time. The received data may not be forwarded to the user except as a sequenced stream of data. Therefore, the receiving transport must buffer the correctly received data, pending the arrival of missing messages. Many more original messages can be sent in a specified period of time when selective acknowledgement is used; thus, many more messages must be buffered by the receiver.



Average Receive Memory Utilization as a Function of Error Rate for One-For-One and Selective Acknowledgement
Figure 7



Maximum Receive Memory Utilization as a Function of Error Rate for One-For-One and Selective Acknowledgement
Figure 8

CONCLUSIONS

This paper has shown that two minor, optional changes to the ISO class 4 transport protocol can extend the range of network media over which the ISO transport protocols can achieve efficient operation. A simple modification in the discipline for sending transport expedited data can provide throughput efficiency improvements of up to 38% and delay improvements of up to 61%. The modification in the sending procedures requires no corresponding change in the receiving procedures; therefore, an implementation of the class 4 transport protocol can include the enhanced expedited transfer mechanism with no loss in interoperability with those implementations that adhere to the ISO standard.

Inclusion of an optional selective acknowledgement procedure can yield throughput efficiency increases of up to 34%, but only when both the sender and receiver are able to use the mechanism. The option can be included such that transport implementations using selective acknowledgement can fully interoperate with those implementations which do not use the selective acknowledgement discipline. This interoperability can be achieved in the short-term by careful selection and publication of the parameter code used to identify the optional selective sequence number in acknowledgement messages.

The changes to the ISO class 4 transport protocol proposed in this paper will be submitted to appropriate standards bodies for possible inclusion in subsequent revisions to the ISO transport protocol standard. Adoption of these optional procedures by the standards bodies is necessary to ensure that any short-term agreements to implement the procedures will be protected in the long run by codification within international standards.

REFERENCES

[1] "Specification of a Transport Protocol for Computer Communications, Volume 3: Class 4 Protocol," National Bureau of Standards, February 1983.

[2] "Specification of a Transport Protocol for Computer Communications, Volume 4: Formal Service Specification," National Bureau of Standards, January 1983.

[3] International Organization for Standardization, "Transport Protocol Specification," ISO/TC 97/SC 16, International Standard ISO/IS 8073 rev., June 1984.

[4] International Organization for Standardization, "Transport Service Definition," ISO/TC 97/SC 16, International Standard ISO/IS 8072 rev., June 1984.

[5] Mills, Kevin L., Marnie Wheatley, and Richard Colella, "Simulation of an International Standard Transport Protocol," National Bureau of Standards (pending publication).

[6] Colella, Richard, Marnie Wheatley, and Kevin Mills, "COMSAT/NBS Experiment Plan for Transport Protocol," National Bureau of Standards (pending publication)