# Data Link-Control Performance: Results Comparing HDLC Operational Modes

W. Bux and K. Kümmerle
*IBM Zurich Research Laboratory, 8803 Rüschlikon, Switzerland*

and

H.L. Truong
*Institute of Switching and Data Techniques, University of Stuttgart, 7000 Stuttgart, F.R. Germany*

The objective of this paper is to compare the performance of the different operational modes of the standard data link-control procedure HDLC: The Normal Response Mode (NRM), an unbalanced procedure with primary and secondary stations which have different functional capabilities, and the Asynchronous Balanced Mode (ABM), a balanced procedure with functionally equivalent, so-called combined stations. Performance is measured in terms of: (1) the maximum throughput to be achieved over a point-to-point link, and (2) the mean transfer time of messages as a function of the link utilization. To obtain an accurate account of the performance of these procedures, the approach taken is to implement the two protocols in a simulation model of a data link, i.e., including all details of the sequencing mechanism and of the error-recovery procedures. The general conclusion from the study is that throughput and delay performance of ABM is better than, or at least equal to NRM. However, there exists a broad range of parameters where the performance of both modes is equivalent.

*Keywords*: Data communication, protocols, data-link control, performance evaluation, simulation, throughput, delay

**Werner Bux** received the M.S. and Ph. D. degrees in electrical engineering from Stuttgart University, Stuttgart, West Germany, in 1974 and 1980, respectively. From 1974 to 1979 he was with the Institute of Switching and Data Techniques, University of Stuttgart, where he worked primarily in the field of performance analysis of data communication networks and computer systems. He joined the IBM Zurich Research Laboratory in 1979, where he is currently working on the architecture and performance evaluation of local networks.

**Karl Kümmerle** received the M.S. and Ph. D. degrees in electrical engineering from Stuttgart University, Stuttgart, West Germany, in 1963 and 1969, respectively. From 1963 to 1969 he was Research Assistant with the Institute for Switching Techniques and Data Processing at Stuttgart University, where he was engaged primarily in investigations in the field of telephone traffic theory and mathematical statistics. From 1970 to 1972 he was Research Associate of the National Research Council at the Computation Laboratory of NASA's Marshall Space Flight Center. There he worked in the field of computer performance evaluation. In 1972 he joined the IBM Zurich Research Laboratory and has worked in the field of data communication networks. His current interest is in local area networks, where he is manager of a research project.

**Hong Linh Truong** was born in Vietnam and received the M.S. degree in electrical engineering from Stuttgart University, Stuttgart, West Germany, in 1977. Since then, he has been Research Assistant with the Institute for Switching and Data Technics, University of Stuttgart. His current research activities are in the field of computer and data communication networks with particular emphasis on performance considerations.

The authors would like to thank Prof. Dr. A. Lotze for his continued interest and support; D. Delestre, J. Grinenwald, and W. Schneider for writing the simulation models.

---

[1] Based on the article entitled "HDLC Performance: Comparison of Normal Response Mode and Asynchronous Balanced Mode of Operation" by W. Bux, K. Kümmerle, and H.L. Truong which appeared in National Telecommunications Conference, November 30–December 3, 1980, Houston, Texas. © 1980 IEEE.

North-Holland Publishing Company
Computer Networks 6 (1982) 37–51

## 1. Introduction

The increasing need for reliable and efficient data communication has forced computer manufacturers and standardization bodies to develop a number of data link-control procedures. The majority of these procedures are contained in the link-control procedures standardized by ISO under the name of HDLC (High Level Data Link Control) [11–13]. HDLC represents a data link-control architecture applicable to a wide variety of links: half-duplex, full-duplex, point-to-point, and multipoint.

In this paper, we restrict ourselves to the consideration of full-duplex, point-to-point links. For this kind of operation, we investigate the two operational modes which seem to be the most frequently applied in practice: (i) Normal Response Mode (NRM), a procedure based on the master-slave principle using two types of stations with different functional capabilities, primary and secondary, and (ii) Asynchronous Balanced Mode (ABM), a symmetrical procedure with fully equivalent, so-called combined stations.

The main differences between NRM and ABM are: In *NRM*, the primary station is responsible for link initialization and disconnection and for error recovery. The secondary station is only capable of responding to commands from the primary. In particular, a secondary is not allowed to transmit without having received explicit permission from the primary in the form of a polling command. This strict master-slave policy is mainly motivated by the operation of half-duplex links and of multipoint configurations.

For point-to-point, full-duplex links, such an unbalanced operation is possible but not necessary: *ABM* is a fully symmetrical procedure in the sense that the combined stations are equally responsible for link initialization and disconnection, and for error recovery. Furthermore, the type of transmission is asynchronous, meaning that both stations can transmit without permission from the other station.

A lot of controversial arguments are used in the discussion about the qualities of both modes. Advocates of NRM often claim the following advantages of this mode over ABM: (i) More efficient link management; (ii) simpler recovery mechanisms by putting the primary station in charge of recovery actions; and (iii) more cost-effective solution by having the major functions (and hence intelligence) in one rather than both stations [5]. On the other hand, the following arguments are used in favor of ABM: (i) More flexible solution because any two combined stations are able to communicate, whereas in NRM no communication is possible among primaries and among secondaries, respectively; and (ii) better performance in terms of higher throughput and lower delay.

This paper is intended to shed more light onto the performance aspect of both modes. To obtain accurate, unbiased, and detailed results, our approach was to implement both procedures within the simulation model of a full-duplex, point-to-point link. More precisely, we have implemented the *information transfer part* of both procedures including all details of the sequencing and error-recovery mechanisms. The transmission medium is determined through transmission rate, propagation delay, and error characteristic.

A considerable amount of effort has been expended on the performance evaluation of data link-control procedures, see e.g., refs. [1–10]. The performance analysis of the *Balanced* Classes of HDLC has been addressed in [1–3]. Furthermore, idealized protocols of the balanced type of HDLC have been analyzed in [6,7]. In Ref. [9], the throughput behavior of the *Unbalanced* HDLC procedures and of SDLC has been investigated for one-way traffic flow over a point-to-point link.

The present paper extends and complements the above analyses in the following points:

(1) Both throughput and delay performance of the Balanced and Unbalanced HDLC procedures are analyzed and compared under identical conditions.
(2) The investigation covers a broad spectrum of traffic-flow conditions, since it turns out that the relative performance of the different procedures is very sensitive to such assumptions.
(3) By having the complete protocols implemented in our simulation model, we are able to gain insight into the detailed behavior of the procedures both during error-free operation and error recovery phases. In fact, the results obtained show that throughput and delay performance is affected by rather complex interactions among the various protocol mechanisms.

In the next section, we describe our simulation model. Section 3 constitutes the main part of the paper, where we discuss and compare typical performance results for both modes. In Section 4, the general conclusions of the study are presented. For the reader not familiar with the details of both HDLC procedures, we briefly review the major characteristics of ABM and NRM in Appendix Al. Performance of the third operational mode of HDLC, the "Asynchronous Response Mode" (ARM) which uses primary and secondary stations but offers improved transmission capability of the secondary is briefly discussed in Appendix A2.

## 2. Performance Model

Figure 1 shows the data link model used for our simulation study. It consists of two data stations connected by a full-duplex circuit. The link is controlled by HDLC, operating either in NRM or in ABM. In NRM, Station-A is designated as primary, Station-B as secondary. In ABM, both stations are combined stations.

Messages to be transmitted are stored in the send buffer of the sending station where they have to wait for transmission. Messages are transmitted according to first-come, first-served, one message per I-frame. Troughout this paper, we assume that the size of the message buffers is unlimited and that the messages are either of constant length $l$ or have exponentially distributed lengths with mean $l$. The transmission channels are characterized by their transmission rate $v$, their bit-error probability $p_{\text{bit}}$ (independent bit errors), and their (one-way) propagation delay $t_{\text{prop}}$. Furthermore, we assume that processing of a received frame requires the constant time $t_{\text{proc}}$. For this investigation, we combine propagation and processing delay in a constant but arbitrarily selectable delay $t_{\text{p}} = t_{\text{proc}} + t_{\text{prop}}$. Constant processing time for the link-control functions is clearly an idealized assumption. However, since we do not consider very high-speed links in this paper and, moreover, slightly varying values of $t_{\text{proc}}$ do not significantly affect performance, this simplification appears to be justified for our purposes.

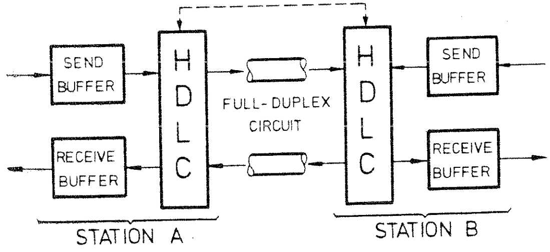

Fig. 1. Structure of the Data Link Model.

As already pointed out, the assumption for the simulation model is that the link has already been initialized. Apart from that, all mechanisms of the procedures have been implemented in full detail to obtain accurate and reliable performance results.

The following commands and responses from the basic repertoire of HDLC [12] are used in the simulation model: Information (I), Receive Ready (RR), and Receive Not Ready (RNR). In addition, we use the command/response Reject (REJ) being offered under optional functions for efficient error recovery. Details of the procedures are given in Appendix Al.

## 3. Results

This is the main section of the paper in which typical performance results for the two HDLC modes, NRM and ABM will be discussed.

We distinguish between two categories of results: maximum throughput and mean transfer time. In the first case, we determine the maximum number of information bits which can be transmitted in one direction of the link per unit time. This means, we assume that for the direction considered, information to be sent is available at any point in time. In the second case, we are interested in the transfer time of messages, defined as time interval from the arrival (or generation) of a message at one station until its successful, i.e., error-free reception at the other station. Since the transfer time includes queueing delays, the underlying assumption is that the channel load in the direction considered is only a fraction of the full channel capacity. In both cases, the traffic load on the reverse channel of the link has an impact on the performance of the direction considered; therefore, its value is treated as a parameter in the subsequent discussions.

Furthermore, two different link types are considered: A terrestrial link with processing plus propagation delay of 50 msec, and transmission rate $v = 4.8$ kbit/sec and $v = 48$ kbit/sec; and a satellite link with processing plus propagation delay of 350 msec and a transmission rate of 48 kbit/sec. The above values of $t_{\text{p}}$ are the same as in [9]. All simulation results are shown with their 95% confidence intervals.

### *3.1.* Throughput *Results*

### *3.1.1. General Throughput Behavior*

Figures 2 to 6 show typical results for the maxi-

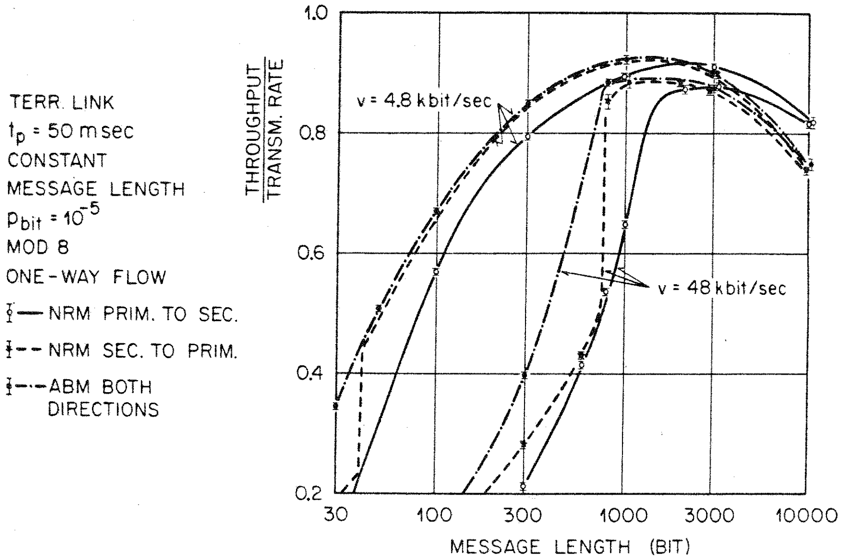

Fig. 2. Maximum Throughput Relative to Transmission Rate vs. Message Length. One-Way Flow. Impact of Transmission Rate.

mum throughput in one direction of the link as a function of message length, i.e., the length of the information field of I-frames. With respect to the traffic load on the reverse channel, we distinguish between two cases: either the reverse channel carries no I-frames at all ("one-way flow") or the reverse channel is fully loaded, too ("two-way flow"). Situations in between these extreme cases are not discussed, but obviously could easily be analyzed with the simulation model.

The throughput results are in consonance with the general behavior of link-control procedures employing an error-detection and retransmission scheme [4]: (i) The maximum throughput of information bits in any case has an upper-bound given by the ratio of I-field length to total I-frame length, where the latter contains a fixed amount of overhead bits for flags, address, control, and frame check sequence fields (see Figure 12 in Appendix A1). This explains why maximum throughput must de-

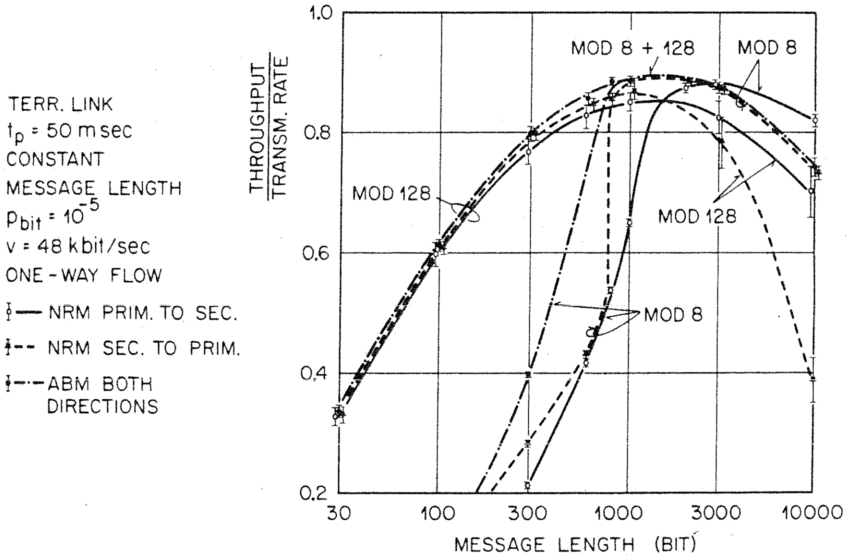

Fig. 3. Maximum Throughput Relative to Transmission Rate vs. Message Length. One-Way Flow. Impact of Modulus Value.

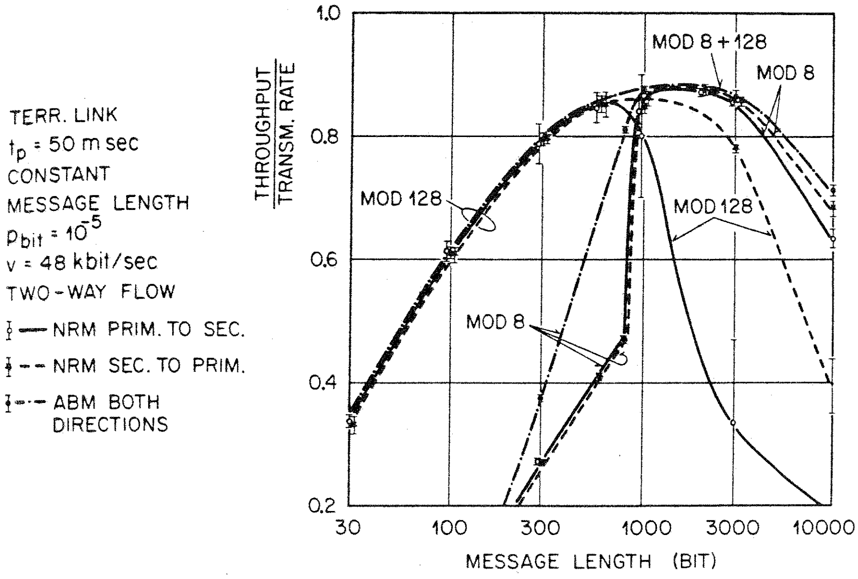

Fig. 4. Maximum Throughput Relative to Transmission Rate vs. Message Length. Two-Way Flow. Constant Message Length.

crease for very short messages. (ii) Since the block-error probability of the I-frames increases with growing frame length, the fraction of error-free, i.e., useful transmissions, decreases. These two effects explain why maximum information throughout must decrease for very short and also for very long messages. However, they are absolutely insufficient to explain the actual throughput characteristic of a particular data link-control procedure. In the case of HDLC-controlled links, maximum throughput depends on the following parameters/mechanisms:

(a) Transmission channel- and application-specific parameters: transmission rate; propagation and processing delays; error characteristic of the channels; I-field lengths; and reverse channel traffic.

(b) Protocol-specific parameters: framing overhead; modulus value M; handling of acknowledgements;

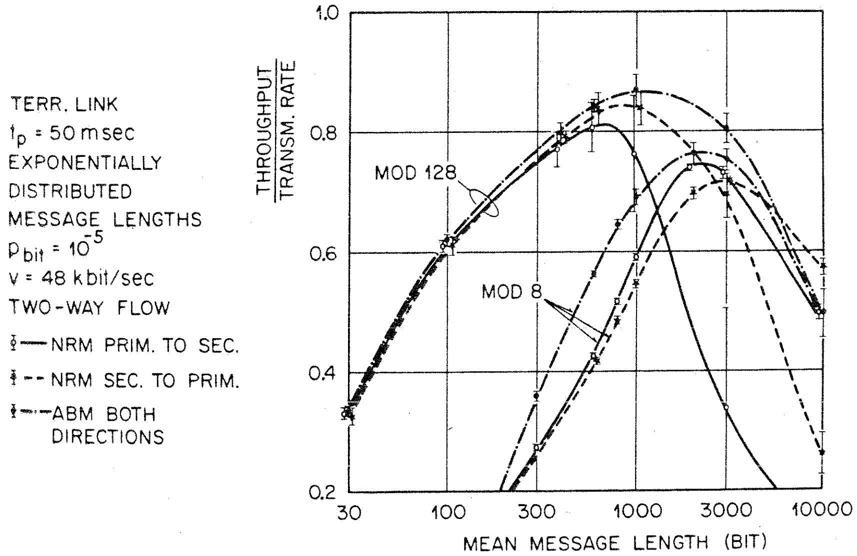

Fig. 5. Maximum Throughput Relative to Transmission Rate vs. Mean Message Length. Two-Way Flow. Exponentially Distributed Message Length.

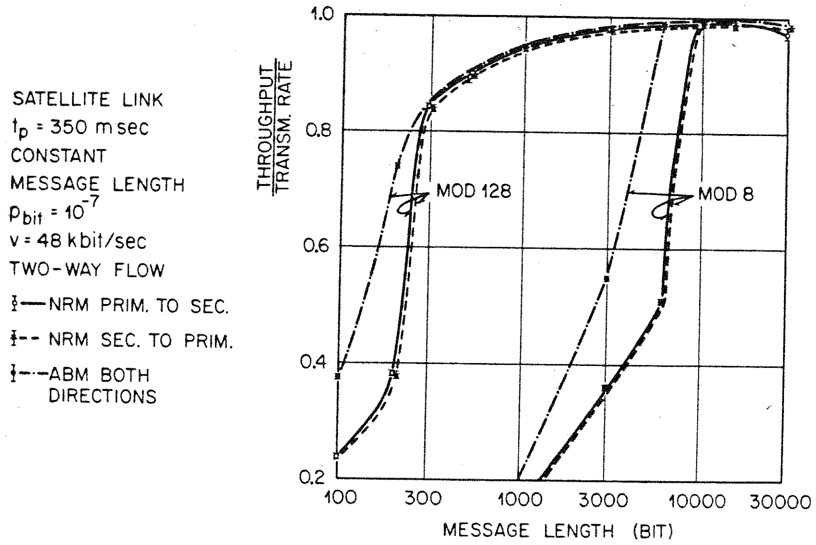

Fig. 6. Maximum Throughput Relative to Transmission Rate vs. Message Length. Two-Way Flow. Satellite Link.

error-recovery procedures; and polling mechanism.

As expected, the following results indicate that the throughput behavior is a complex function of the above parameters, and that the results can be explained only if we take into account all effects and their interaction.

### 3.1.2. *One-Way Flow Examples*

Figure 2 shows the throughput efficiency for NRM and ABM in case of one-way flow on a terrestrial link. Since the link configuration in ABM is symmetrical, we obtain the same results for both directions of the link, whereas in NRM the throughput differs in both directions. One-way flow in NRM means either that the primary station or the seondary station always has I-frames to be transmitted.

Let us first discuss the results for the region to the left of the maximum. The general difference in throughput efficiency of the 4.8 kbit/sec and the 48 kbit/sec link is due to the modulus rule of HDLC [12] (see also Appendix A1): In case of the higher transmission rate, the transmission of I-frames has more frequently to be discontinued than for the lower speed due to reaching the boundary of $M - 1 = 7$ unacknowledged I-frames. Comparing the link directions, we observe that, to the left of the maximum, the throughput efficiency from secondary to primary is constantly higher than in the other direction. The reason for this being that the primary is always able to acknowledge correctly received I-frames, whereas the secondary can do this only when it has been polled. This means that the primary reaches the boundary of seven unacknowledged I-frames more often and also over a broader range of message lengths than the secondary. The effect that in the ascending branch of the curves ABM is equal or superior to both directions of NRM has the following explanation: When the secondary reaches the number of seven unacknowledged I-frames, it stops transmission and indicates this by setting the F-bit in its last frame. Its next opportunity to resume transmission is when it receives a new P-bit. On the other hand, a combined station having reached the limit of seven unacknowledged I-frames can resume transmission at an earlier point in time, namely, when it receives an acknowledgement for one or more of the seven outstanding I-frames. Therefore, the wasted time is shorter, and hence information throughput higher for ABM than for the direction secondary-primary in NRM.

Next, we consider the throughput characteristics in the region to the right of the maximum in Fig. 2. Here, the situation changes: The direction primary to secondary shows higher throughput than the reverse direction; the ABM throughput is almost equal to the NRM throughput from secondary to primary. This result can be explained as follows: The effect limiting throughput in this region is the occurrence of transmission errors; the particular characteristics are due to the special error-recovery procedure employed in the different modes of operation, c.f. Appendix. Al. In ABM and in the case of secondary-to-primary flow of NRM, the major portion of transmission errors is recovered by REJ recovery; therefore, we observe almost equal throughput for these two cases. For one-way flow from primary to secondary, REJ recovery is possible, too. However, due to the strategy which we implemented, viz., to poll the secondary as often as possible, in most cases, checkpoint retransmission is performed before the REJ can become effective. It should be noted that the same strategy to set the P-bit as often as possible may also be applied in ABM.

From Figure 2 it can be observed that a modulus value of eight can lead to throughput degradation on higher-speed terrestrial links if the sum of propagation and processing delay is considerably longer than the I-frame transmission time. In this case, it may be appropriate to use the extended control-field format allowing a modulus value of 128.

Figure 3 shows that for short message lengths, throughput efficiency can be significantly improved by this means. However, this figure also demonstrates that, in NRM, the use of a high modulus value can lead to problems if the block-error rate is high: The striking effect of this diagram is that, for message lengths greater than 3000 bits, the throughput from secondary to primary shows a dramatic degradation for the modulo-128 case. This result is caused by the following mechanism: If a REJ frame or – more likely – a retransmitted I-frame is received in error, REJ recovery fails (c.f. section on REJ Recovery in Appendix Al). In the case of one-way flow from secondary to primary considered here, such an error is recovered when the secondary has M-1 (7 or 127) unacknowledged I-frames outstanding and therefore stops transmitting by setting the Final (F-)bit in its last frame. The next Poll (P-)bit received will cause the secondary to perform P/F-bit recovery. Since the resulting recovery delay is especially long in case of a higher modulus, we obtain the puzzling effect that – under these conditions – a small modulus value may result in a better performance than a large one. The link operated under ABM does not show this effect because time-out recovery, as described in the section on Time-Out Recovery in Appendix Al, leads to a much shorter recovery delay in this situation.

### *3.1.3. Two-Way Flow Examples*

We next discuss results for the maximum throughput under the assumption of two-way flow, i.e., that both stations always have information to be sent (Figures 4–6).

Figure 4 shows that, in the region to the left of the maximum of the throughput curves, both directions of the link under NRM are almost equivalent in the case of two-way flow. The reason for this being that, in contrast to one-way flow, acknowledgments can now be carried by I-frames. However, in the case of a modulus value of eight, we still observe a difference in throughput between ABM and NRM for the following reason: in *ABM*, the only reason why a station does not make full use of its send channel is that the limit of *M*-1 unacknowledged I-frames has been reached. As soon as a station receives an acknowledgment for one or more of these frames, it resumes I-frame transmission. In *NRM*, the same holds true for the primary. However, the mean acknowledgment delay for its I-frames is longer as compared to ABM, because despite the two-way flow there are time intervals where the secondary has sent a frame with the F-bit set to 1, and is therefore momentarily unable to respond. This longer acknowledgment delay is the source of the throughput degradation for the primary as compared to a combined station.

For the secondary, in NRM the acknowledgment delay is roughly equal to the one obtained in ABM because the primary is always capable of acknowledging a received I-frame. However, additional intervals of wasted time are caused – compared to ABM – when a secondary has received new acknowledgments, so that the modulus rule no longer prevents it from sending further I-frames but it may previously have sent a frame with the F-bit set to 1 and not yet received a new P-bit. This additional time causes the throughput degradation for the secondary as compared to a combined station. An interesting result of our simulation is that the – *per se* different – sources of performance degradation for primary and secondary result in roughly the same throughout value for both directions.

For the modulus-8 case in Fig. 4, the throughput results for both directions of the NRM link and for ABM are almost equivalent in the region of the optimum and within a wide range of longer message lengths.

As in the corresponding example of one-way flow (Figure 3), the direction secondary to primary of the NRM link exhibits rapid throughput degradation for longer message lengths if a modulus value of 128 is used. The explanation of this result is virtually the same as in the one-way flow case: If REJ recovery fails for some reason, an error is not recovered before the secondary: (i) has reached the barrier of M-1 unacknowledged I-frames; (ii) has set the F-bit; (iii) has received the next P-bit, and (iv) has performed checkpoint retransmission on the basis of this P-bit. In contrast to one-way flow, the direction from primary to secondary exhibits an even worse behavior than the other direction. This performance degradation is also caused by an inefficient recovery procedure in case REJ recovery fails. This can occur either when a retransmitted I-frame is garbled or because a REJ has been received in error. In either situation, the primary continues to transmit I-frames until it has $M$-1 I-frames outstanding. This, however, does not yet recover the error situation because the secondary may continue to transmit for a very long time until it also reaches the limit of $M$-1 unacknowledged I-frames, and therefore sets the F-bit to 1. The primary station cannot recover from this error situation before this F-bit has been received. To shorten the recovery delay in this situation, we implemented the strategy that the primary stops acknowledging received I-frames when it has $M$-1 I-frames outstanding.

A possibility to overcome this unfortunate situation is that the primary through sending an RNR, forces the secondary to stop transmitting I-frames and to set the F-bit so that P/F-bit recovery would occur much earlier. The problem with this solution, however, is that it is not generally applicable: if conditions are such that the primary frequently has $M$-1 unacknowledged I-frames outstanding, then the throughput from secondary to primary will be drastically reduced if this solution is employed.

Again, the ABM link does not exhibit the effect described because if REJ recovery fails, time-out recovery becomes active after a relatively short delay.

All results shown so far have been obtained under the assumption of constant message length. To ensure that our general results are not restricted to this particular case, we also investigated the throughput efficiency for different message-length distributions. As an example, Figure 5 shows the throughput efficiency of the same configuration as Figure 4, however, with exponentially distributed message lengths. As long as neither modulus nor transmission errors have a significant impact, the same throughput results are obtained independent of the message-length distribution. However, both effects – modulus value and transmission errors – for the following reasons lead to stronger performance degradation if message lengths are not constant: (i) The modulus rule affects the throughput over a broader range of parameters in the case of variable message lengths, because the limit of M-1 unacknowledged I-frames is more frequently reached than in the constant-length case. (ii) Since long I-frames are more error-prone than short ones, the retransmission frequency of long frames is higher. This explains the lower throughput efficiency for the exponentially distributed message lengths in the case of longer frames. It should be noted, however, that in Figures 4 and 5, the general relation among the results for the different HDLC modes is nearly the same, independent of the message-length distribution.

Figure 6 shows throughput results for two-way flow over a satellite link. Because of the better error characteristic, $p_{\text{bit}} = 10^{-7}$ as opposed to $10^{-5}$ in the previous examples, we observe over the whole range considered only the effects described above for the region to the left of the maximum. Obviously, the use of a modulo-128 numbering scheme drastically improves the throughput efficiency for short I-frames.

### 3.2. Transfer-Time Results

Here, we consider the case where the channels are only loaded to a fraction of their full capacity to guarantee finite delay. Interesting performance measures in this case are the mean transfer time of the messages and the mean buffer-holding time. Transfer time is defined as the time interval from the arrival (or generation) of a message at one station until its successful reception at the other station. Buffer-holding time is the time interval from the arrival of a message at a station until the reception of an acknowledgement for this message. This means that the transfer time includes queueing, transmission, processing and propagation times, as well as possible additional delays caused by retransmissions; the buffer-holding time comprises the transfer plus the acknowledgment time of a message.

In the following examples, we consider terrestrial links, and assume symmetrical traffic conditions, i.e., that the traffic offered to both channels is equal.

Figure 7 shows for a constant message length of 5000 bit, the mean transfer time relative to the transmission time of an I-frame as a function of the useful channel load, which is defined as the ratio of throughput of information bits per unit time and the transmission rate. In this example, we deliberately

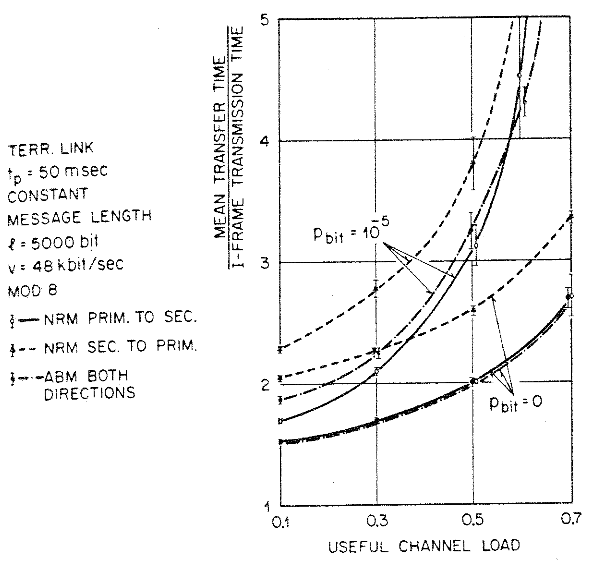

Fig. 7. Mean Transfer Time Relative to I-Frame Transmission Time vs. Useful Channel Load. 5000 Bit Constant Message Length.

chose relatively long messages to exclude any impact of the modulus rule. For clarity reasons, the latter effect is discussed below in the context of Figs. 9 and 10. Two different cases are considered: error-free transmission, $p_{\text{bit}} = 0$, and channels with a bit-error rate of $p_{\text{bit}} = 10^{-5}$. In the error-free case, ABM and the direction from primary to secondary of the NRM link show the same delay values. The messages transmitted from secondary to primary suffer an additional delay due to the polling mechanism. The mean value of this additional delay is approximately equal to the sum $t_p$ of processing and propagation delay. This result can be explained as follows: Under the ideal conditions assumed, the queueing delay for the ABM link and for the primary-secondary direction of the NRM link is that of a simple M/D/1 queue, whereas for the reverse direction of the NRM link, the following modification of an M/D/1 queueing model is necessary: As long as the server is busy, the model operates like an M/D/1 queue. When the server becomes idle, the next opportunity to obtain service is after a fixed time interval $\tau$ has elapsed. If no customer arrives during this time, a new interval $\tau$ is started, etc. It can be easily shown that the mean queueing times of the modified and of the original M/D/1 queue differ by $\tau/2$. Applied to the channel from secondary to primary, the interval $\tau$ corresponds to the time interval when the secondary has sent the F-bit but has not yet received a new P-bit. This time is approximately equal to $2t_p$, which explains our simulation results.

For both links, we obtain, of course, longer transfer times if transmission errors occur on the channels. Apart from very high loads, the direction from primary to secondary of the NRM link performs slightly better than the ABM link. We have already observed this effect for the throughput in the case of one-way flow from primary to secondary (c.f. Figure 2): In ABM, the majority of transmission errors is recovered via REJ recovery, whereas errors occurring on the primary-to-secondary channel of the NRM link are mostly recovered via P/F-bit recovery. Because of our strategy to poll the secondary as often as possible, the latter recovery mechanism works faster under light and medium traffic. This causes the difference between both delay curves.

Whereas in Figure 7 constant length messages have been assumed, Figure 8 shows the corresponding transfer-time results in case of exponentially distributed message lengths. Obviously, this leads to longer transfer times because the queueing delay increases with the variance of the transmission times. Like in the throughput case, however, we can also observe from Figures 7 and 8 that the general relation between the performance of the two different modes is robust with respect to the message-length distribution.

Figure 9 shows the same delay characteristic as Figure 7, however, with an I-field length of 800 instead of 5000 bits. This assumption changes the situation considerably. Comparing the delay of the error-free ABM link with the transfer time of an M/D/1 queue (the ideal case), it can be concluded that for loads greater than 0.4, the interarrival times of the messages are such that occasionally more than seven I-frames are ready for transmission.

For the direction primary to secondary of the NRM link, the transfer delay deviates from the ideal case already at a channel utilization of 0.2, and increases rather steeply beyond channel loads of 0.3. The reason for this being that the secondary is not able to acknowledge I-frames when it has not been polled. Thus, acknowledgment times are prolonged and the primary reaches the limit of seven unacknowledged I-frames more frequently and at lower loads than a combined station. Again, messages transmitted from secondary to primary suffer a higher transfer delay caused by the constraint that the secondary has to be polled before transmitting.

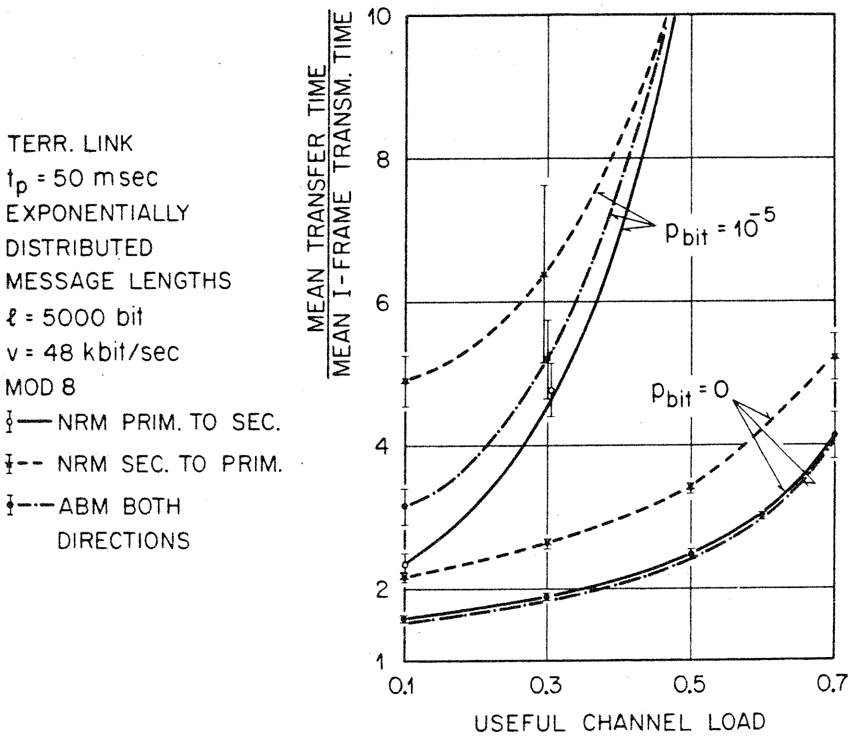

Fig. 8. Mean Transfer Time Relative to Mean I-Frame Transmission Time vs. Useful Channel Load. Exponentially Distributed Message Lengths (Mean 5000 Bit).

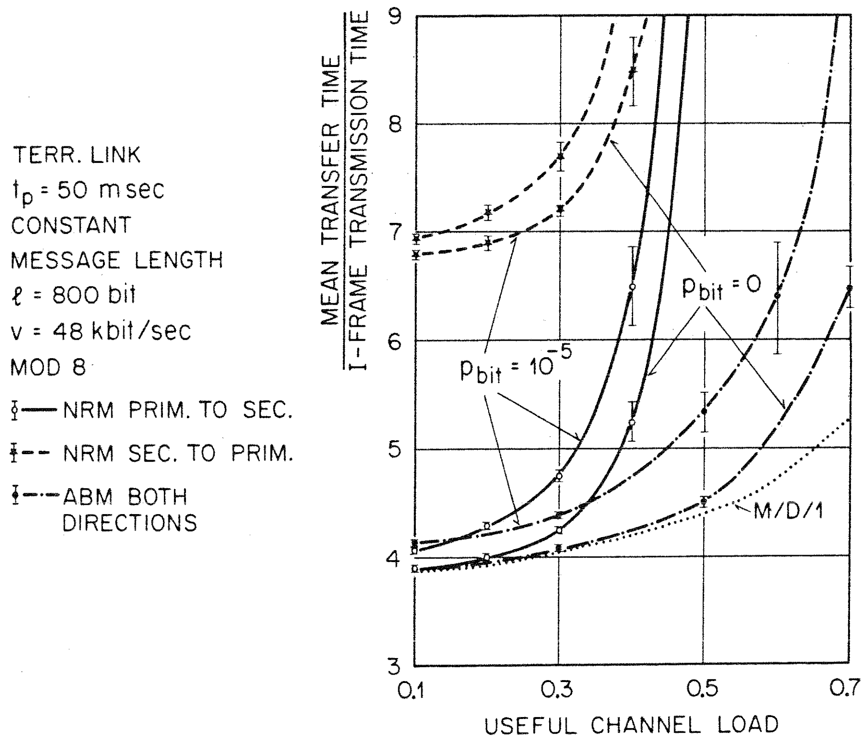

Fig. 9. Mean Transfer Time Relative to I-Frame Transmission Time vs. Useful Channel Load. 800 Bit Constant Message Length.

Figure 10 represents the transfer-time characteristic of NRM and ABM links in a more general manner. It shows the mean transfer time relative to the I-frame transmission times as a function of the sum of processing and propagation delay. The useful channel load $Y_u$ is held constant at 0.5. Although obtained for particular values of I-field length $l$ and transmission rate $v$, this diagram is also typical of other I-frame lengths and transmission rates, provided the I-frame block-error probability is 0 or $10^{-2}$.

For small values of the processing plus propagation delay, the mean transfer time in Fig. 10 grows linearly as subsequently explained: (i) In ABM and for the primary-to-secondary direction of the NRM link, the queueing delay remains constant because the useful channel load, $Y_u$, is held constant. For the reverse direction of the NRM link, the mean queueing time grows linearly with $t_p$ as described above. (ii) For ABM, it has been shown analytically that the mean recovery time in case of small processing and propagation delays grows linearly with this quantity [2,3]. Due to the similar recovery mechanisms, it is not surprising that this result also applies to NRM. Deviation from linearity indicates the impact of the modulus. Characteristically, this impact is experienced in NRM at considerably smaller ratios of processing plus propagation delay and I-frame transmission times than in ABM.

Finally, Figure 11 shows an example for the mean buffer-holding time normalized to the I-frame transmission time as a function of the useful channel load. As defined above, the buffer-holding time is the time interval for which a message has to be stored in the buffer of a sending station until a positive acknowledgment is received. Hence, it represents a measure for the buffering load imposed on a station. As can be seen from this example, buffers have to be held considerably longer within primary and secondary stations than within combined stations. Furthermore, a comparison of Figures 11 and 9 shows that the differences between the directions of the NRM link observed for transfer time do not exist with respect to buffer-holding time: Shorter transfer delays of the messages transmitted from primary to secondary are almost totally outweighed by their longer acknowledgment delays.

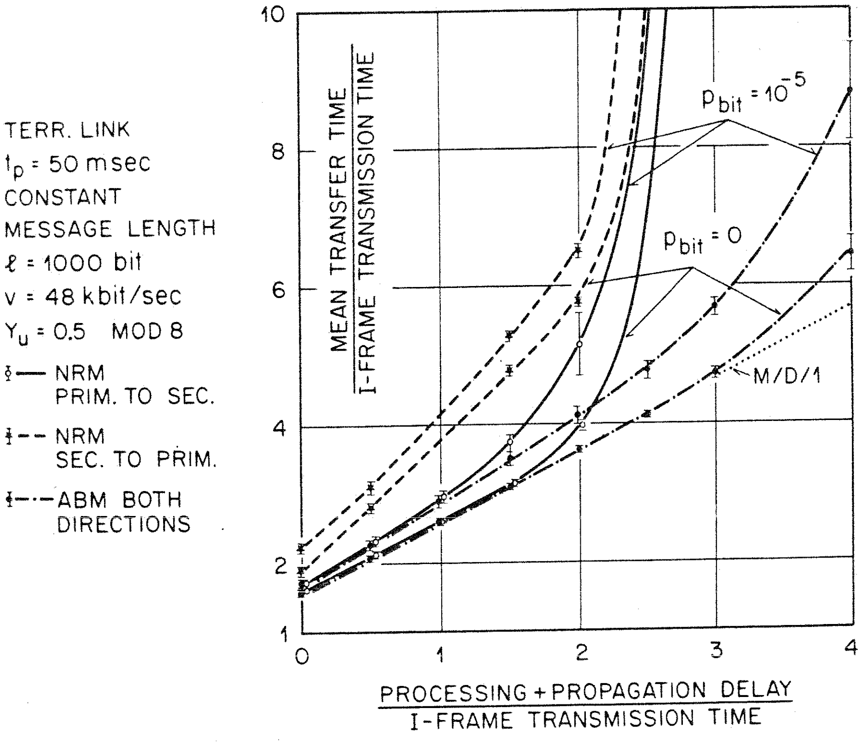

Fig. 10. Mean Transfer Time Relative to I-Frame Transmission Time vs. Processing plus Propagation Delay Relative to I-Frame Transmission Time.

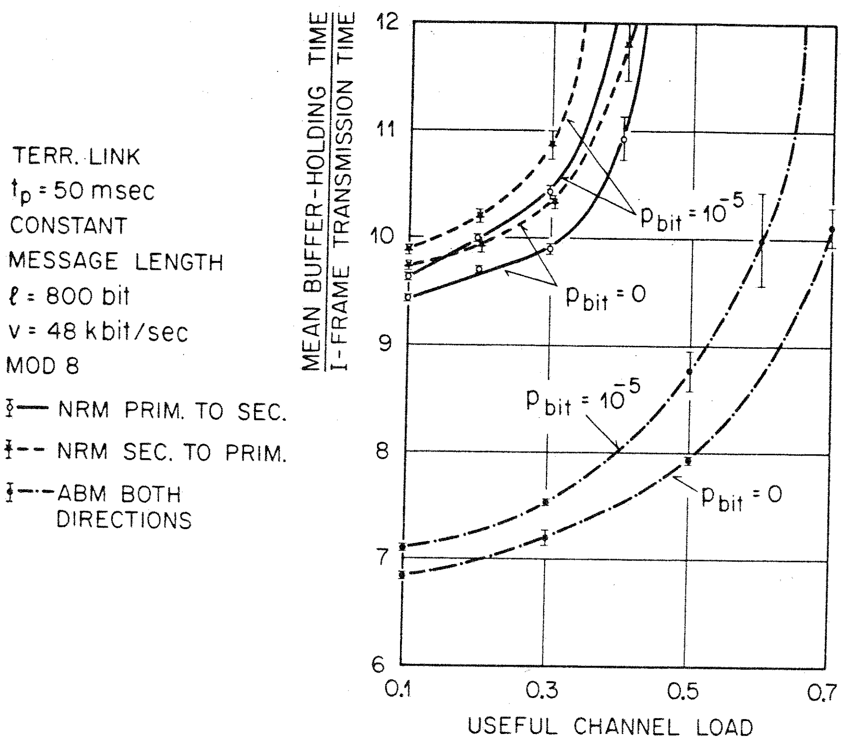

Fig. 11. Mean Buffer-Holding Time Relative to I-Frame Transmission Time vs. Useful Channel Load.

## 4. Conclusions

The performance results presented in this paper allow the following conclusions:

(1) The link operated in ABM performs better than, or at least as well as the link under NRM. However, there is a relatively broad range of parameters where the performance of both modes is equivalent.

(2) In NRM, the modulus of the sequence numbers affects performance at significantly smaller values of the processing and propagation delay than in ABM. A practical consequence of this is that a modulus value of eight might be sufficient for higher-speed terrestrial links if ABM is used, but can be too small in the case of NRM.

(3) Whereas ABM yields equal performance results in both directions of the link, throughput and delay of the NRM link can be significantly different for both directions.

(4) The situations where the throughput of both modes differs significantly are: (i) In the case of long propagation and processing delays where the modulus rule in conjunction with the polling operation results in a throughput degradation of NRM due to longer and more frequent intervals of wasted time. (ii) For high block-error probabilities where in NRM a long recovery delay can occur if REJ recovery fails or is not allowed. In ABM, this error situation is efficiently resolved by time-out recovery.

(5) With respect to message transfer time, both directions of an ABM link and the direction from primary to secondary in NRM show comparable results as long as the modulus value has no impact. The direction secondary-to-primary suffers from an additional delay due to the polling operation. The mean of this additional delay is in the order of the processing plus propagation delay. Depending on the ratio of processing plus propagation delay to I-frame transmission time, the modulus value can have a much stronger delay-increasing effect in NRM than in ABM.

(6) Buffer-holding time and buffer load in NRM are almost equal for primary and secondary stations. In the case of relatively long processing and propagation delays, they can be considerably higher than in ABM.

## Appendix

### *A1. HDLC – Normal Response Mode and Asynchronous Balanced Mode*

In this Appendix, we briefly review the main features of HDLC and discuss the commonalities and differences of the elements of procedure for the two modes of operation investigated in this paper.

### *A1.1. HDLC Frames, Commands and Responses*

HDLC defines three types of frames for the transmission of data and control information (c.f., Figures 12a and b): (a) Information frames (Information Format): They contain an arbitrary sequence of user bits within the information field. Information (I) frames are sequentially numbered with a send sequence number *N*(*S*) and allow piggybacking of acknowledgments through the receive sequence number *N*(*R*). Furthermore, they can carry a Poll (P) or Final (F) bit. (b) Supervisory frames (Supervisory Format); They are generally used to acknowledge correctly received I-frames by their receive sequence number *N*(*R*), to request retransmission in case of transmission errors, or to request a temporary suspension of the transmission of I-frames. Supervisory (*S*) frames can also carry a P/F bit. (c) Unnumbered frames (Information Format or Supervisory Format): These frames are used during link initialization or disconnection. Since we consider the performance of HDLC during its information transfer phase, unnumbered frames are irrelevant for our investigations.

### *A1.2. Error-Free Operation*

There are two properties of the HDLC procedure which mainly influence the performance in the error-free case: the modulus value of the sequence numbers and the usage of the Poll/Final (P/F) bit which is different in the two modes considered:

*1. Modulus of Sequence Numbers:* The sequence numbers of the I-frames cycle through the set of numbers 0, 1, 2, ..., *M*-1, where *M* is the modulus value. The modulus equals eight for the unextended and 128 for the extended format. Under certain conditions, e.g., long propagation delays, it may influence the data flow over the link, because a station must stop transmitting I-frames if it has *M*-1 unacknowledged I-frames simultaneously outstanding.

(a)

| FLAG | ADDRESS | CONTROL | INFORMATION | FRAME-CHECKING SEQUENCE | FLAG |
|---|---|---|---|---|---|
| 8 BIT | 8 BIT | 8 BIT | $\ell$ BIT | 16 BIT | 8 BIT |

Fig. 12a. HDLC Information Format (Unextended Control Field).

(b)

| FLAG | ADDRESS | CONTROL | FRAME-CHECKING SEQUENCE | FLAG |
|---|---|---|---|---|
| 8 BIT | 8 BIT | 8 BIT | 16 BIT | 8 BIT |

Fig. 12b. HDLC Supervisory Format (Unextended Control Field).

*2. P/F-Bit Usage*: The P-bit is generally used to solicit a response from the other station. Apart from time-out recovery situations, a station may have only one P-bit outstanding at a given time. Before it can issue another frame with the P-bit set to 1, it must receive a response from the other station with the F-bit set to 1. The particular meaning of the P- and F-bit is different in NRM and ABM.

In *NRM* the secondary cannot transmit until a command with the P-bit set to 1 is received. The secondary must set the F-bit to 1 in the last frame of its response. Following transmission of the frame with the F-bit set to 1, the secondary must stop transmitting until another frame with the P-bit set to 1 is received. The strategy in our implementation is to poll the secondary station as often as possible to give it the opportunity to transmit at the earliest point in time.

In ABM, each combined station can send both commands with or without P-bit and responses with or without F-bit. In contrast to NRM, each station transmits on an asynchronous basis, i.e., without having to wait for any permission from the other station. The P-bit is used to solicit a response at the earliest opportunity with the F-bit set to 1. For example, if a station wants to get positive acknowledgment that a particular command has been received, it may set the P-bit in that command to 1.

### *A1.3. Error Recovery*

A frame received in error is simply discarded by the receiver without any further action. If the frame is an I-frame, the error will manifest itself later in the form of a sequence error or it will be detected by means of time-out or P/F-bit recovery.

*1. P/F-Bit Recovery:* As P- and F-bits are always exchanged as a pair, they can be used for checkpointing purposes.

If in *NRM* a station receives a frame with the P/F-bit set to 1, it initiates retransmission of unacknowledged I-frames with sequence numbers less than or equal to the *N*(*S*) number of the last frame transmitted with the P/F-bit set to 1.

In *ABM*, checkpoint retransmission is only initiated based on frames received with the F-bit set to 1, to avoid possible interference with other recovery possibilities. HDLC does not specify for ABM under which conditions the P-bit has to be set. We adopted the strategy of setting the P-bit only in those cases where it is necessary to query the status of the other station, e.g., after the time-out has expired (see section on Time-Out Recovery below).

*2. REJ Recovery*: The REJ command/response is used for a more timely reporting of sequence errors. Its use is explained in the following example: Assume that the I-frame with *N*(*S*) = 1 of Station-A is received in error and, therefore, discarded by Station-B. When Station-B receives the next error-free I-frame, e.g., the frame with *N*(*S*) = 2, it informs Station-A of a sequence error by issuing a REJ frame with *N*(*R*) = 1. Upon receipt of this frame, Station-A retransmits the requested I-frame with *N*(*S*) = 1 plus all additional I-frames which have been subsequently transmitted. To avoid retransmission of correct I-frames which followed a disturbed I-frame, HDLC provides the optional Selective-Reject function (SREJ), where only the disturbed I-frame has to be retransmitted. Due to the greater complexity, the option SREJ has been implemented only rarely in current systems.

If a retransmitted I-frame is again received in error, then the REJ recovery cannot be repeated for reasons of possible ambiguities. The error situation has then to be resolved either by checkpointing or by time-out recovery. (A case where this property of the procedure is important is discussed in Section 3.1.2.)

*3. Time-Out Recovery*: A single I-frame or the last I-frame in a sequence of I-frames cannot be recovered by REJ. Also a frame with the P-bit set to 1 may be lost. To recover from such error situations, HDLC provides a time-out function. Use of the time-out also differs in NRM and ABM.

In *NRM* only the primary station has a time-out function. Its use is not specified in the Standard [12,13]. In Annex B of [12], however, labeled "Timer Considerations" the following recommendations are given: The timer is started upon transmission of a command with the P-bit set to 1. It is restarted when an error-free frame with the F-bit set to 0 is received, and it is stopped when an error-free frame with the F-bit set to 1 is received. Upon expiration of the timer, it is recommended that the primary queries status with a supervisory frame.

In *ABM*, each combined station has a time-out function. Again, HDLC does not specify how the timer should be handled. For our implementation, we adopted the following rules: The timer is started (provided it is not running already) every time an I-frame or a frame with the P-bit set to 1 is sent. It is restarted when a frame is received which acknowledges a not-yet acknowledged I-frame. The timer is stopped when no unacknowledged I-frame is outstanding.

### *A2. Performance of the Asynchronous Response Mode (ARM)*

In addition to the two modes NRM and ABM, investigated in detail in this paper, HDLC defines a third mode of operation, the "Asynchronous Response Mode" (ARM). Although this mode has not been implemented in our simulation model, we are able to make some predictions about the performance of ARM based on the experience gained for the other two modes.

Both ARM and NRM are unbalanced procedures in the sense that one station of a link is designated primary, the other secondary. In contrast to NRM, a secondary station in ARM has the capability of transmitting on an asynchronous basis, i.e., without having to wait for a P-bit from the primary. Use of the P/F bit in ARM is identical to ABM with one exception: The secondary initiates checkpoint retransmission also based on a received P-bit.

A further difference from NRM lies in the fact that ARM requires a timer in the secondary to ensure that lost I-frames are detected under all circumstances. Handling of the *primary timer* and the action taken by the primary after timer expiration can be exactly the same as for a combined station (c.f., section on Time-Out Recovery in Appendix A1). The *secondary timer* has to be handled slightly differently: It only serves the purpose of additional I-frame protection and is in no way correlated with P/F-bit usage. Since after timer expiration, the secondary has no means to query the status of the primary, it retransmits the unacknowledged I-frame with the least $N(S)$ number.

Taking into account these major differences in the operation of an ARM link as compared to NRM and ABM, we can make the following conjectures with respect to the performance of ARM:

*a) Throughput*: The possible throughput degradation of NRM as compared to ABM to the left of the optimum is due to the polling mechanism of NRM. Since the type of response in ARM is asynchronous, its throughput behavior in this region should be equal to ABM.

In the region to the right of the optimum, NRM performs worse than ABM if, due to a high error rate, REJ recovery frequently fails and the inefficient recovery mechanism of NRM in this case causes an extremely long recovery delay (c.f. Figures 3 to 5). In ARM, time-out recovery prevents such an unfortunate situation from occurring in a similar way to ABM. As described above, the only difference between the two modes in this respect is that a secondary station in contrast to a combined station is unable to query status with a P-bit. In some cases, this may lead to a slightly less efficient recovery; however, this effect can be expected to be minor. Therefore, the conclusion from this discussion is that the overall throughput performance of ARM should be similar to ABM.

*b) Transfer Time*: The major deviations between NRM and ABM obtained with respect to mean transfer time of the messages are caused by the polling mechanism of NRM and its interaction with the modulus rule. Concerning the other differences between both procedures, no significant impact on this performance measure has been observed. Therefore, we can conclude that, due to its asynchronous type of response, ARM should lead to transfer-time results similar to ABM.

## References

[1] W. Bux, K. Kümmerle and H.L. Truong, "Results on Performance of Balanced HDLC Procedures," NTC 78 Conference Record, Birmingham, Alabama, (IEEE, 1978) 28.3.1-28.3.7.

[2] W. Bux and H.L. Truong, "A Queueing Model for HDLC-Controlled Data Links," Proceedings International Symposium on Flow Control in Computer Networks, J.L. Grangé and M. Gien, Eds. (North-Holland, Amsterdam/New York/Oxford, 1979) 287–306.

[3] W. Bux, K. Kümmerle and H.L. Truong, "Balanced HDLC Procedures: A Performance Analysis," IEEE Trans. Commun., Vol. COM-28, No. 11, 1889–1898 (Nov. 1980).

[4] W.W. Chu, "Optimal Message Block Size for Computer Communications with Error Detection and Retransmission Strategies", IEEE Trans. Commun., Vol. COM-22, No. 10, 1516–1525 (Oct. 1974).

[5] R.J. Cypser, Communications Architecture for Distributed Systems, (Addison-Wesley, Reading, Massachusetts, 1978), p. 368.

[6] E. Gelenbe, J. Labetoulle and G. Pujolle, "Performance Evaluation of the Protocol HDLC," Proceedings Symposium on Computer Network Protocols, A. Danthine, Ed. (Université de Liège, Belgium, 1978), pp. G3.1-G3.5.

[7] J. Labetoulle and G. Pujolle, "Modeling and Performance Evaluation of the Protocol HDLC," Proceed-

ings International Symposium on Flow Control in Computer Networks, J.L. Grangé and M. Gien, Eds. (North-Holland, Amsterdam/New York/Oxford, 1979) 307–320.

[8] D. Towsley and J.K. Wolf, "On the Statistical Analysis of Queue Lengths and Waiting Times for Statistical Multiplexers with ARQ Retransmission Schemes," IEEE Trans. Commun., Vol. COM-27, No. 4, 693-702 (April 1979).

[9] K.C. Traynham and R.F. Steen, "SDLC and BSC on Satellite Links: A Performance Comparison," ACM Computer Communication Review, Vol. 7, No. 4, 3–14 (1977).

[10] L.W. Yu and J.C. Majithia, "An Analysis of One Direction of Window Mechanism," IEEE Trans. Commun., Vol. COM-27, No. 5, 778–788 (May 1979).

[11] Data Communication – High-Level Data Link-Control Procedures – Frame Structure, International Standard ISO 3309 (1976).

[12] Data Communication – High-Level Data Link-Control Procedures (independent numbering). International Standard ISO 4335 (1976).

[13] HDLC – Proposed Balanced Class of Procedures, Document ISO/TC 97/5C6-No. 1444 (1977).