Feb 79 N 1/5

A QUEUEING MODEL FOR HDLC-CONTROLLED DATA LINKS

W. Bux and H.L. Truong

Institute of Switching
and Data Technics

University of Stuttgart

Stuttgart , Fed. Rep. of Germany

In this paper it is shown that the performance of the Balanced Class of HDLC Procedures can be evaluated for a wide range of applications by using a rather simple queueing model. The key idea to our approach is the use of the so-called "virtual transmission time" which comprises the real transmission time of a message as well as the duration of possible recovery actions performed in case of transmission errors.

In order to derive analytic results for the delay and the throughput of a data link, we take all the mechanisms into account which through a detailed simulation of the protocol turned out to have the major impact on performance. In this way explicit and simply computable expressions for the maximum throughput as well as the mean transfer time of messages in the presence of transmission errors are found and validated by comparison with results from a simulation study of HDLC-controlled data links where the information exchange phase of HDLC Balanced Class of Procedures was implemented in full detail.

The developed model provides not only a deeper insight into the performance determining mechanisms of the Balanced Class of HDLC Procedures but it can also serve as an efficient and reliable tool, if it is used as a submodel in performance studies of higher protocol levels.

1. INTRODUCTION

In the protocol hierarchy of computer communication networks the data link control level represents an important part. The main functions of this level are: (i) to achieve an error-protected transport channel, and (ii) to provide means for flow-control.

In this investigation we concentrate on the ISO link contol procedure HDLC,

---

Paper presented at the International Symposium on Flow Control in Computer Networks, February 12-14, 1979 Paris, France.

especially on its Asynchronous Balanced Mode of Operation (ABM) [1,2,3]. The Balanced Class of operation is intended for point-to-point links carrying large traffic volumes, where transfer times and link efficiency are important factors. The stations at both sides of the link are so-called Combined Stations: each station can send both commands and responses. The type of transmission response is asynchronous, which means that each station can transmit information (I-) or supervisory (S-) frames without having to wait for an explicit permission by the other station.

The objective of the investigation is to study the performance of HDLC,ABM with a special emphasis on its behavior in the presence of transmission errors. All presented results are based on an implementation of this protocol in the framework of a simulation study of HDLC controlled data links [4] .
In addition to this simulation study we develop here an analytic approach to the performance evaluation of the Balanced Class of Procedures of HDLC in order to serve two purposes: first, to provide a tool for a fast and reliable performance evaluation of HDLC controlled data links, and second, to provide more insight into how the various parameters of an HDLC implementation and the data channel characteristics influence the performance.

2. DATA LINK MODEL

Figure 1 shows the structure of our data link model. It consists of two data stations connected by a full-duplex data circuit. Messages to be transmitted from station A to station B, or vice versa, are stored in the send buffer of the sending station, where they have to wait for transmission. Messages are transmitted according to a first-come-first-served discipline, one message per I-frame. Throughout this paper we assume that the message buffers are unlimited, and that processing time can be neglected.

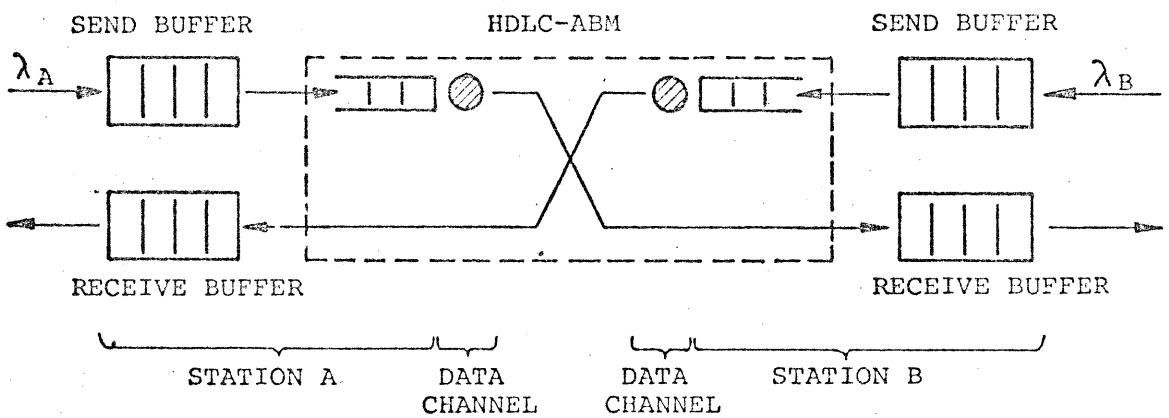


Fig. 1: Structure of the data link model

The transmission channels are characterized by their transmission rate V and the bit-error probability p_{bit}. We assume independent bit-errors. Propagation delay is neglected, which is justified for most terrestrial links.

We assume furthermore that the messages to be transmitted to Station B arrive at Station A according to a Poisson process with rate λ_A, and that the messages to be transmitted to Station A arrive at Station B according to a Poisson process with rate λ_B. The messages are of constant length ℓ .

The transfer of information across the data link is controlled according to HDLC, Balanced, Asynchronous Response Mode, Class of Procedures with the optional function REJ (reject). The REJ-command/response is used for a more timely reporting of I-frame sequence errors.

3. PERFORMANCE CONSIDERATIONS

3.1 Performance measures

The performance of a data link control procedure can be defined in various ways. For reasons of clarity we distinguish between two different kinds of traffic situations of a data link, the saturated and the non-saturated case.

The saturated case assumes the idealistic condition that a station has always information to be sent. The appropriate performance measure for this kind of operation is the throughput, i.e. the number of successfully transmitted information bits per unit of time. It is obvious that the throughput in the saturated case has the maximum achievable value.

The non-saturated case corresponds to the more realistic condition where the traffic on the data link varies statistically and the channels are only loaded to a certain percentage of their maximum capacity. In this case we are mainly interested in the various delay times, for instance in the transfer time T_f of messages from their arrival at one station until their successfull receipt at the other station or the buffer holding time T_b of messages from their arrival at one station until their deletion due to a positive acknowledgment received from the other station.

Whereas related studies on the performance of HDLC procedures have dealt with the saturated case only [5-7] , our approach attempts to handle both cases (saturated and non-saturated) in a uniform way.

3.2 Transfer time analysis

3.2.1 Outline of analysis

The following considerations are related to one direction of transmission, namely from Station A to Station B. Due to the Balanced Mode of Operation the results for the other direction can be obtained in the same way.

In case of zero error-probability our problem is nearly equivalent to the M/D/1-queue, where M stands for the Poisson (Markovian) input process with rate λ_A, D is the constant I-frame transmission time, and the figure 1 indicates one server (the transmission channel from A to B). For non-zero error probability, however, the problem becomes very complex due to the various error recovery actions following a detected I-frame error.

In Figure 2 a simple example of operation is shown, using symbols of references [2,3]. Considering the direction A to B we assume that I-frame no. 1 is received in error. A recovery action, however, is not started before the next error-free I-frame is received at Station B. Upon receipt of I-frame no. 2 in our example, Station B reports a sequence error to Station A using a REJ-frame. Having received the REJ-frame Station A retransmits I-frame no.1 and subsequently I-frames no. 2,3,...

Our approach to the evaluation of the transfer time is described as follows. As indicated in Fig. 2 we conceive that I-frame no.i occupies the transmission channel for a fictitious time $T_v(i)$, which we denote "virtual transmission time". The time $T_v(i)$ begins with the start of transmission of I-frame no.i under the condition that the preceding frame no. i-1 has been successfully transmitted. The time $T_v(i)$ ends when I-frame no.i has also been successfully

transmitted, i.e. without FCS and sequence error.[1] The virtual transmission time $T_v(i)$ is equal to the real transmission time of I-frame no.i, if the frame is transmitted without FCS and sequence error; in case of an FCS error of I-frame no.i it is prolonged (i) by the time until the receiver detects a sequence error or a timer expires, (ii) by the recovery action, and (iii) by the retransmission of I-frame no. i.

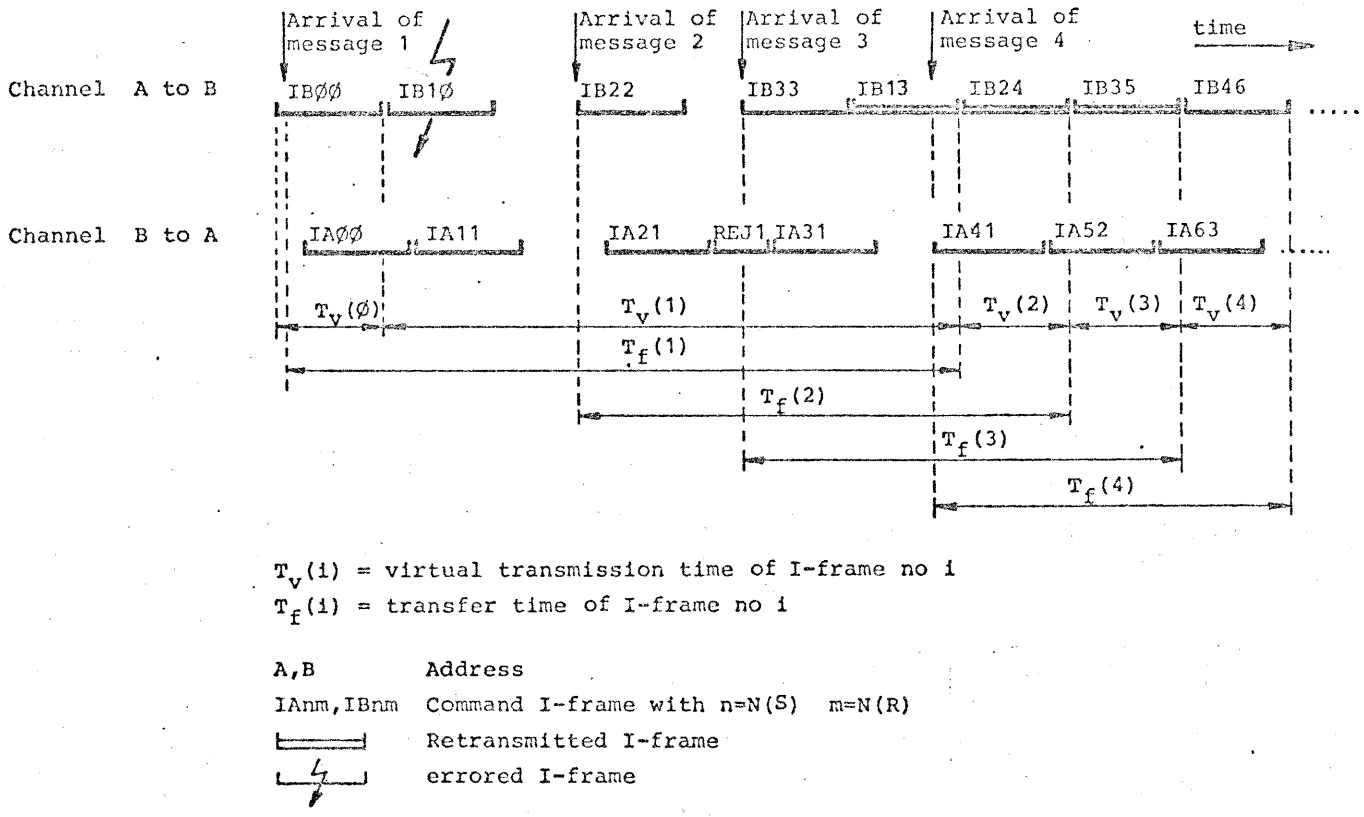


Fig. 2: Example of HDLC, Balanced Mode of Operation

It should be noted that our definition of the virtual transmission time is not equivalent to the commonly used "effective transmission time", which is usually defined from the first transmission of a message until its successfull receipt at the other station. Due to the specific kind of the error recovery mechanisms of HDLC, ABM, the latter definition cannot be used for our purposes.

By the introduction of the virtual transmission time we have obtained a model wherein the I-frames consecutively ("virtually") occupy the transmission channel. Therefore, we are now able to approximately determine the transfer time by using well-known results from the theory of M/G/1 queueing systems (Markovian input, Generally distributed service times, 1 server), if we succeed in determining the distribution function (or a sufficient number of moments) of the virtual transmission time.

---

[1] FCS error: transmission error within an HDLC frame detected with the aid of the 16-bit Frame Checking Sequence (FCS).

Sequence error: condition at the receiver when an I-frame is received error-free (no FCS error) but contains a send sequence count N (S) that is not equal to the receive state variable.

The mean transfer time t_f of messages from Station A to Station B can thus be determined with the aid of the well-known Pollaczek-Khintchine formula [8]:

$$t_f = E[T_v] + \frac{\lambda_A E[T_v^2]}{2(1 - \lambda_A E[T_v])} \qquad (1)$$

Therefore, the main problem we are faced with is to determine the first two moments of the virtual transmission time, if we are interested in the mean transfer time of messages.

3.2.2 Virtual transmission time

Since the HDLC-controlled information flow over the link is a very complex process we are only able to derive approximate results for the virtual transmission time. Our goal is to take into account all those effects which through our simulation results turned out to have the major impact. Besides the traffic conditions on the link there exist mainly two effects which may influence the virtual transmission time: transmission errors and the modulo value determining the maximum number of simultaneously outstanding (unacknowledged) I-frames. A detailed simulation study of HDLC controlled data links, where the information exchange phase of HDLC, ABM was implemented in full detail showed, however, that under our assumptions (constant I-frame length, no propagation delay) the modulo value (8 or 128) has no significant impact [4]. Therefore, we concentrate on the effect of transmission errors in determining the virtual transmission time.

The following abbreviations will be used (c.f. Fig. 3):
(Let I-frame no.i be the frame for which the virtual transmission time is derived.)

t_I : transmission time of an I-frame ($=\{\ell + 48 \text{ bit}\} / V$ in case of modulo 8).

t_S : transmission time of an S-frame ($= 48 \text{ bit}/ V$ in case of modulo 8).

t_{out} : duration of time-out.

T_e : random time between end of transmission of the disturbed I-frame no.i and either end of transmission of the first I-frame without FCS error following the disturbed I-frame no.i, or end of transmission of an RR command with P-bit=1, following a time-out.

T_r : random time between either end of transmission of the first I-frame without FCS error following the disturbed I-frame no.i, or end of transmission of an RR command with P-bit=1 following a time-out and start of retransmission of I-frame no.i.

$T_{res}(2)$: random time interval between receipt of either the first I-frame without FCS error following the disturbed I-frame no.i, or of an RR command with P-bit=1 at Station B and start of transmission of REJi or RRi,F.

$T_{res}(1)$: random time interval between receipt of REJ i or RRi,F at Station A and retransmission of I-frame no.i.

p_B : block error probability, probability that an I-frame is disturbed

($p_B = 1 - (1-p_{bit})^{\{\ell+48 \text{ bit}\}}$ in case of modulo 8).

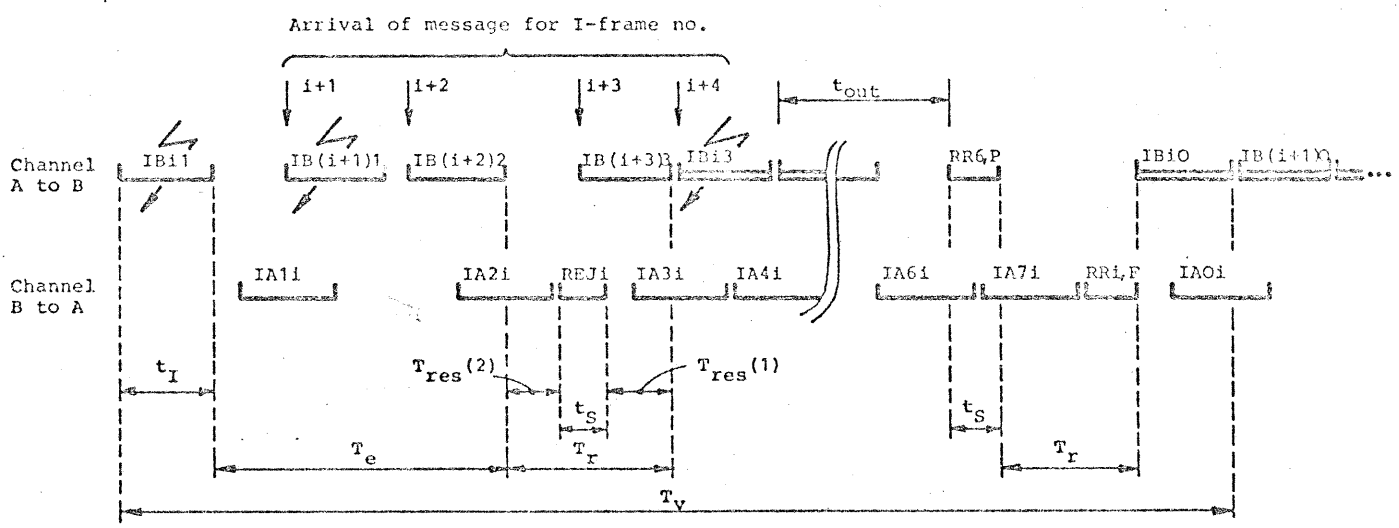


Fig. 3: Components of the virtual transmission time

Figure 3 shows an example of the various time components which make up the virtual transmission time if the considered I-frame is disturbed two successive times.

The time until the first retransmission of I-frame no.i is equal to $t_I+T_e+T_r$.

If the retransmitted I-frame no.i is again received with FCS error then no REJ recovery is possible according to the rules of HDLC, but the error situation has to be resolved by time-out recovery or P/F bit-recovery. For our analysis we assume that in this case recovery is initiated by the expiration of a timer. After Station A has timed out it transmits an RR-command with the P-bit set to 1. In response to the received P-bit, Station B transmits an RR-response with the F-bit set to 1. Upon receipt of this RR-frame retransmission of I-frame no.i is initiated according to the rules of P/F-bit recovery.

Therefore, if I-frame no.i is disturbed N successive times ($N \geqq 1$) its virtual transmission time is composed as follows:

$$T_v = t_I+T_e+T_r+(N-1)(t_I+t_{out}+t_s+T_r) + t_I \qquad (2)$$

N is distributed according to:

$$P\ (N=n) = (1-p_B)\ p_B^{\ n} \qquad (3)$$

From this it follows:

$$E[T_v] = \frac{1}{1-p_B}t_I + p_B E[T_e] + \frac{p_B}{1-p_B}E[T_r] + \frac{p_B^2}{1-p_B}(t_{out}+t_S) \qquad (4)$$

$$E[T_v^2] = (1+3p_B)t_I^2 + p_B E[T_e^2] + \frac{p_B^2(1+p_B)}{1-p_B}t^2 + \frac{p_B}{1-p_B}E[T_r^2]$$
$$+ 4t_I \left(p_B E[T_e] + \frac{p_B^2}{1-p_B}t + \frac{p_B}{1-p_B}E[T_r] \right)$$
$$+ 2E[T_e] \left(\frac{p_B^2}{1-p_B}t + \frac{p_B}{1-p_B}E[T_r] \right)$$
$$+ 4\frac{p_B^2}{(1-p_B)^2}t\,E[T_r] + \frac{2p_B^2}{(1-p_B)^2}E[T_r]^2 \qquad (5)$$

with $t = t_{out} + t_I + t_S$

Although the time components T_e and T_r are not independent in a strict sense, eq. (5) can serve as a good approximation, because it is intuitively clear and confirmed by our simulation results that their correlation is very small.

Herewith, it remains to determine the first two moments of T_e and T_r.

Time component T_e

As indicated in Fig. 3 the time component T_e begins at the end of transmission of the disturbed I-frame no. i. The end of T_e is determined either by the end of transmission of the first I-frame without FCS error following I-frame no. i or the expiration of a timer. We assume for simplicity that a timer is started every time an I-frame is transmitted.

If we assume, that

$$M\,t_I \leq t_{out} < (M+1)\,t_I \qquad (6)$$

then the distribution and the first two moments of T_e can be written as

$$P(T_e \leq t) = \sum_{x=1}^{\infty} (1-p_B)p_B^{x-1}\,P(T_e(x) \leq t) \qquad (7)$$

$$E[T_e] = \sum_{x=1}^{M+1} (1-p_B)p_B^{x-1}\,E[T_e(x)] + p_B^{M+1}(t_{out}+t_S) \qquad (8)$$

$$E[T_e^2] = \sum_{x=1}^{M+1} (1-p_B)p_B^{x-1}\,E[T_e(x)^2] + p_B^{M+1}(t_{out}+t_S)^2 \qquad (9)$$

The time $T_e(x)$ is equal to T_e, given that I-frame no. i+x is the first I-frame without FCS error following the disturbed I-frame no. i.

The time between end of transmission of I-frame no.i and end of transmission of I-frame no. i+x can be conceived as the time ξ_x until the xth departure of a customer from an M/D/1 queue measured from a departure epoch, namely the end

of transmission of I-frame no.i.

The distribution of ξ_x has been derived by PACK [9] :

$$F_x(t)=P(\xi_x\leq t)=\begin{cases}1-\sum_{m=0}^{x-1}\left[1-Q_{x-m}(t)\right]\pi_m & t\geq xt_I\ ;\ x\geq 0\\ 0 & \text{otherwise}\end{cases} \tag{10}$$

where $Q_n(t)$ is defined as :

$$Q_n(t)=\begin{cases}1-\sum_{k=1}^{n}\lambda_A^{n-k}\,e^{-\lambda_A(t-kt_I)}\,(t-nt_I)(t-kt_I)^{n-k-1}/(n-k)! & t\geq nt_I\ ;\ n\geq 0\\ 0 & \text{otherwise}\end{cases} \tag{11}$$

The π_m in (10) are the steady-state probabilities of the queue length at the departure instants. They are given by

$$\pi_m=\begin{cases}\overline{\pi}_m-\overline{\pi}_{m-1} & m\geq 1\\ \overline{\pi}_0 & m=0\end{cases} \tag{12}$$

with

$$\overline{\pi}_m=(1-\lambda_A t_I)\sum_{i=0}^{m}e^{i\lambda_A t_I}\,(-i\lambda_A t_I)^{m-i}/(m-i)! \qquad \lambda_A t_I<1$$

In order to determine $T_e(x)$, we have to distinguish between two cases, as indicated in Fig. 4, namely :

$$T_e(x)=\begin{cases}\xi_x & \xi_x\leq t_{out}+t_I\\ t_{out}+t_S & \xi_x>t_{out}+t_I\end{cases} \tag{13}$$

Using the definitions

$$P(\xi_x\leq t)=F_x(t)=\int_0^t f_x(t)\,dt \tag{14}$$

$$P(\,T_e(x)\leq t\,)=G_x(t)=\int_0^t g_x(t)\,dt \tag{15}$$

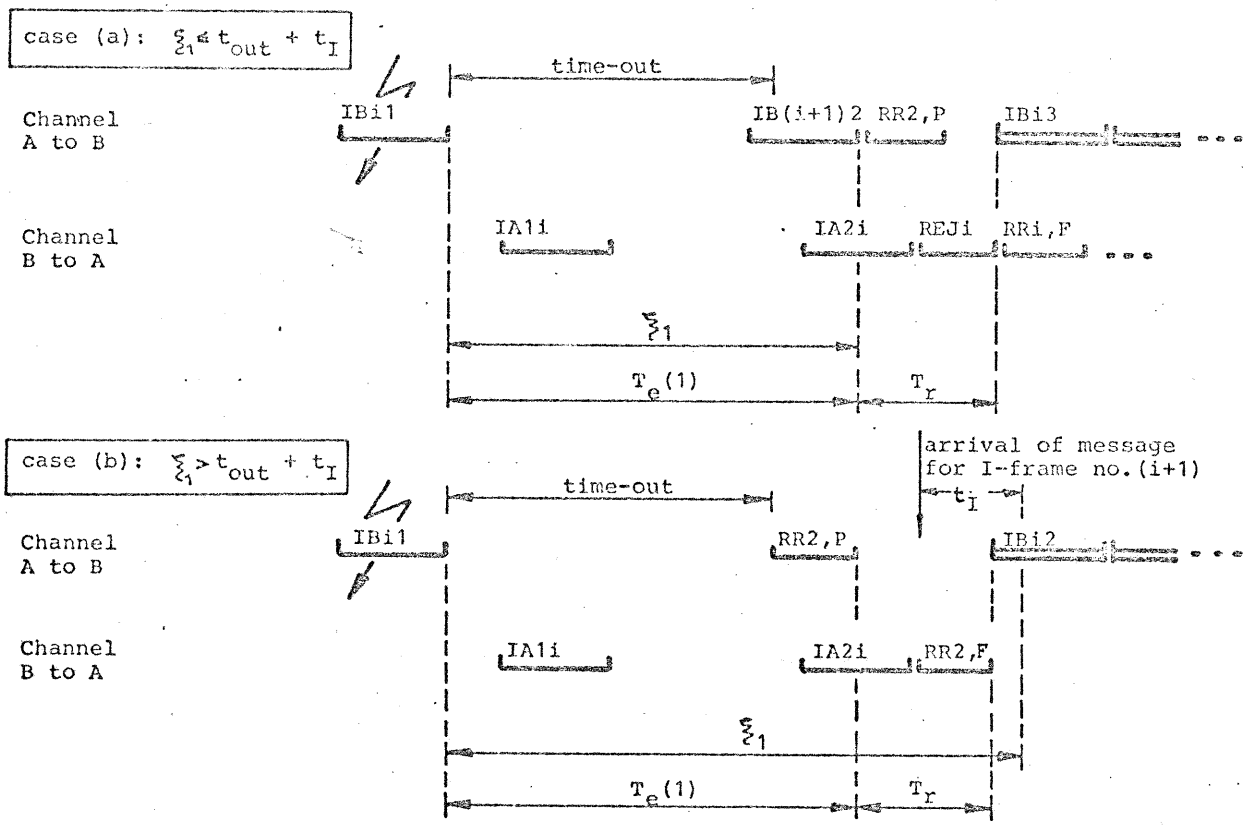


Fig. 4 : Duration of $T_e(x)$ (here : x=1)

the following expressions for the first two moments of $T_e(x)$ can be derived :

$$E[T_e(x)] = \int_0^\infty t\, g_x(t)\, dt$$

$$= t_{out} + t_S + (t_I - t_S)\, F_x(t_{out}+t_I) + F_x^*(xt_I) - F_x^*(t_{out}+t_I) \qquad (16)$$

$$E[T_e(x)^2] = \int_0^\infty t^2\, g_x(t)\, dt$$

$$= (t_{out}+t_S)^2 + \{(t_{out}+t_I)^2 - (t_{out}+t_S)^2\}\, F_x(t_{out}+t_I) - 2(t_{out}+t_I)\, F_x^*(t_{out}+t_I) + 2F_x^{**}(t_{out}+t_I) + 2\, xt_I\, F_x^*(xt_I) - 2\, F_x^{**}(xt_I) \qquad (17)$$

with $F_x^*(t) = \int F_x(t)\, dt$; $F_x^{**}(t) = \int F_x^*(t)\, dt$

The functions $F_x^*(t)$ and $F_x^{**}(t)$ in (16) and (17) can be determined by integrating $F_x(t)$ two times.

If we define $Q_n^*(t) = \int Q_n(t)dt$ and $Q_n^{**}(t) = \int Q_n^*(t)dt$ we obtain

$$F_x^*(t) = \int F_x(t)\ dt = \begin{cases} t(1-\sum_{m=o}^{x-1} \pi_m) + \sum_{m=o}^{x-1} Q_{x-m}^*(t)\ \pi_m & t \geq xt_I \\ o & \text{otherwise} \end{cases} \tag{18}$$

$$F_x^{**}(t) = \int F_x^*(t)\ dt = \begin{cases} \frac{t^2}{2}(1-\sum_{m=o}^{x-1} \pi_m) + \sum_{m=o}^{x-1} Q_{x-m}^{**}(t)\ \pi_m & t \geq xt_I \\ o & \text{otherwise} \end{cases} \tag{19}$$

Using the definitions

$$\begin{aligned} A_{n,k}(t) &= e^{-\lambda_A(t-kt_I)}\ (t-kt_I)^{n-k} \\ B_{n,k}(t) &= (n-k)t_I\ e^{-\lambda_A(t-kt_I)}\ (t-kt_I)^{n-k-1} \\ A_{n,k}^*(t) &= \int A_{n,k}(t)\ dt \\ A_{n,k}^{**}(t) &= \int A_{n,k}^*(t)\ dt \\ B_{n,k}^*(t) &= \int B_{n,k}(t)\ dt \\ B_{n,k}^{**}(t) &= \int B_{n,k}^*(t)\ dt \end{aligned} \tag{20}$$

we can write :

$$Q_n(t) = \begin{cases} 1 - \sum_{k=1}^{n} \frac{\lambda_A^{n-k}}{(n-k)!} [\ A_{n,k}(t) - B_{n,k}(t)\ ] & t \geq nt_I \\ o & \text{otherwise} \end{cases} \tag{21}$$

$$Q_n^*(t) = \begin{cases} t - \sum_{k=1}^{n} \frac{\lambda_A^{n-k}}{(n-k)!} [\ A_{n,k}^*(t) - B_{n,k}^*(t)\ ] & t \geq nt_I \\ o & \text{otherwise} \end{cases} \tag{22}$$

$$Q_n^{**}(t) = \begin{cases} \frac{t^2}{2} - \sum_{k=1}^{n} \frac{\lambda_A^{n-k}}{(n-k)!} [\ A_{n,k}^{**}(t) - B_{n,k}^{**}(t)\ ] & t \geq nt_I \\ o & \text{otherwise} \end{cases} \tag{23}$$

The functions $A^*_{n,k}$, $A^{**}_{n,k}$, $B^*_{n,k}$, $B^{**}_{n,k}$ can be most efficiently computed with the aid of the following recursive relations :

$$A^*_{n,k}(t) = -\frac{1}{\lambda_A} A_{n,k}(t) + \frac{n-k}{\lambda_A} A^*_{n-1,k}(t) \tag{24}$$

$$A^{**}_{n,k}(t) = -\frac{1}{\lambda_A} A^*_{n,k}(t) + \frac{n-k}{\lambda_A} A^{**}_{n-1,k}(t) \qquad k,n\in\{1,2,\ldots\} \tag{25}$$

$$B^*_{n,k}(t) = t_I \left[A_{n,k}(t) + \lambda_A A^*_{n,k}(t)\right] \qquad n\geq k \tag{26}$$

$$B^{**}_{n,k}(t) = t_I \left[A^*_{n,k}(t) + \lambda_A A^{**}_{n,k}(t)\right] \tag{27}$$

By substituting eqs. (24) to (27) into eqs. (22) and (23) we can directly evaluate the functions F^*_x and F^{**}_x according to eqs. (18) and (19). This enables us to compute $E[T_e(x)]$ and $E[T_e(x)^2]$ in eqs. (16) and (17), and hence $E[T_e]$ and $E[T_e^2]$ according to eqs. (8) and (9).

Time component T_r

If we neglect the block error probability of S-frames, which is justified if I-frames are considerably longer than S-frames, then the time component T_r becomes (c.f. Fig. 3) :

$$T_r = T_{res}(2) + t_S + T_{res}(1) \tag{28}$$

From this, the first and second moment of T_r are derived :

$$E[T_r] = E[T_{res}(2)] + t_S + E[T_{res}(1)] \tag{29}$$

$$\begin{aligned} E[T_r^2] = {} & E[T^2_{res}(2)] + t_S^2 + E[T^2_{res}(1)] + \\ & + 2E[T_{res}(2)](t_S + E[T_{res}(1)]) + 2t_S E[T_{res}(1)] \end{aligned} \tag{30}$$

In order to estimate the first two moments of $T_{res}(2)$ and $T_{res}(1)$, we assume that the arrival times at station B of the first correct I-frame following I-frame no.i or the arrival times of RR-frames with P-bit = 1 following a time-out as well as the arrival times at station A of the REJ-frames or RR-frames with F-bit = 1 are purely random. Then $T_{res}(1)$ and $T_{res}(2)$ can be considered as residual life times of a renewal process with constant interevent times (the transmission times of I-frames), if the channel is occupied. This consideration leads to [8] :

$$E[T_{res}(1)] = Y(1)\ t_I/2 \tag{31}$$

$$E[T^2_{res}(1)] = Y(1)\ t_I^2/3 \tag{32}$$

The quantity Y(1) represents the utilization of the channel from station A to station B :

$$Y(1) = \lambda_A t_I \tag{33}$$

In case of extremely high block-error rates the accuracy of the values of $E[T_r]$ and $E[T_r^2]$ obtained from eqs. (29) and (30) can be improved by taking into account that the channel utilization grows due to retransmitted I-frames. In these extreme cases it can be useful to increase Y(1) by adding the following term $Y_e(1)$, which estimates the additional channel load due to errors :

$$Y_e(1) = p_B \lambda_A t_I (1+z_e+z_r) \tag{34}$$

with

$$z_e = \sum_{i=1}^{M} i(1-p_B)p_B^{i-1} \{F_i(t_{out}+t_I)+p_B[1-F_{i+1}(t_{out}+t_I)]\} + (M+1)(1-p_B)p_B^M F_{M+1}(t_{out}+t_I) \tag{35}$$

$$z_r = \lambda_A t_I + \lambda_B t_I - \lambda_A \lambda_B t_I^2 - (1-\lambda_A t_I)\frac{\lambda_B}{\lambda_A}(1-e^{-\lambda_A t_I}) \tag{36}$$

The quantities z_e and z_r represent estimates of the mean numbers of I-frames with sequence errors transmitted during T_e and T_r, respectively.
The first two moments of $T_{res}(2)$ are derived in the same manner.

We have now derived all quantities necessary to determine the first two moments of the virtual transmission time. Therefore, we are now in a position to compute the mean transfer time of the messages, using eq. (1).

In Section 4 of the paper we present numerical results demonstrating the range of validity of our approximate solution by a comparison with simulation results.

3.3 Buffer holding time analysis

In addition to the transfer time derived above, a second time is of some interest : the buffer holding time, namely the time that messages remain in the buffer of a sending station. The knowledge of this time is necessary for a proper dimensioning of the send buffer.

As indicated in Fig. 5 the buffer holding time T_b comprises the transfer time of messages plus an additional delay T_{ack} necessary for sending back the positive acknowledgment :

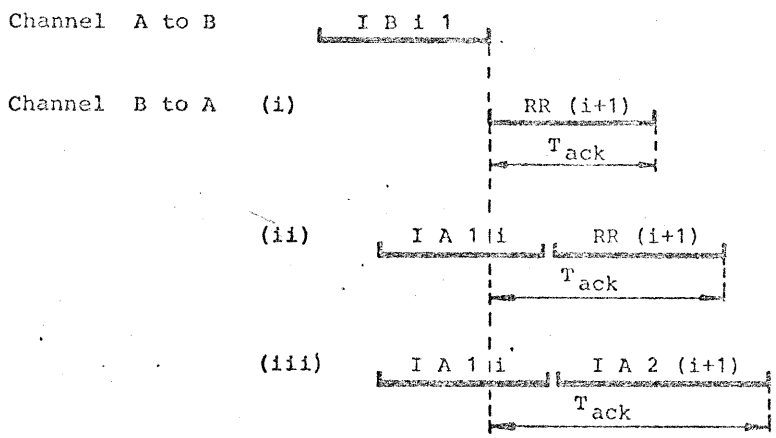
$$T_b = T_f + T_{ack} \tag{37}$$

If the expectation of T_b

$$E[T_b] = t_f + E[T_{ack}] \tag{38}$$

were known then the send buffer load Y_b (mean number of messages in the send buffer) could be easily determined :

$$Y_b = \lambda_A E[T_b] \tag{39}$$



Case (i) : Channel B to A is idle

Case (ii) : Channel B to A is occupied, but station B has no additional I-frame to send afterwards

Case (iii) : Channel B to A is occupied, and station B has additional I-frame to send

Fig. 5 : Derivation of the acknowledgment time

The expectation of the acknowledgment time T_{ack} can be determined as follows (c.f. Fig. 5) :

When an I-frame has been correctly received at station B (i.e. without FCS and sequence error), then three different situations must be distinguished (see the definitions of the probabilities π_m in eq. (12)) :

(i) With probability $\pi_0 = 1-Y(2)$ the channel B to A is idle. Then the received I-frame is acknowledged by using an S-frame (e.g. RR-command/response).

(ii) With probability $\pi_1 = [e^{Y(2)}-1][1-Y(2)]$ the channel B to A is occupied, but station B has no additional I-frames to send afterwards. In this case, the received I-frame will also be acknowledged with the help of an S-frame.

(iii) With probability $1-\pi_0-\pi_1 = 1-e^{Y(2)}[1-Y(2)]$ the channel B to A is occupied and station B has additional I-frames to send. Then the received I-frame is acknowledged by setting N(R) of the next I-frame sent by station B to one plus the value N(S) of the received I-frame.

Therefore, the mean acknowledgment time can be estimated by :

$$E[T_{ack}] = [1-Y(2)]\ e^{Y(2)} t_S + [1+ \frac{Y(2)}{2} + Y(2) e^{Y(2)} - e^{Y(2)}] t_I \qquad (40)$$

This expression neglects a possible small prolongation of the acknowledgment time in case of transmission errors. The comparisons with simulation results in Section 4 of the paper demonstrate that this simplification causes no sensible error.

3.4 Throughput analysis

In this section we briefly describe the analysis of the saturated case, where we assume that the message queues in both directions are never empty. The interesting performance value in this case is the throughput, i.e. the number of information bits being successfully transmitted over the link per unit of time. Again the following results are valid for one direction of transmission.

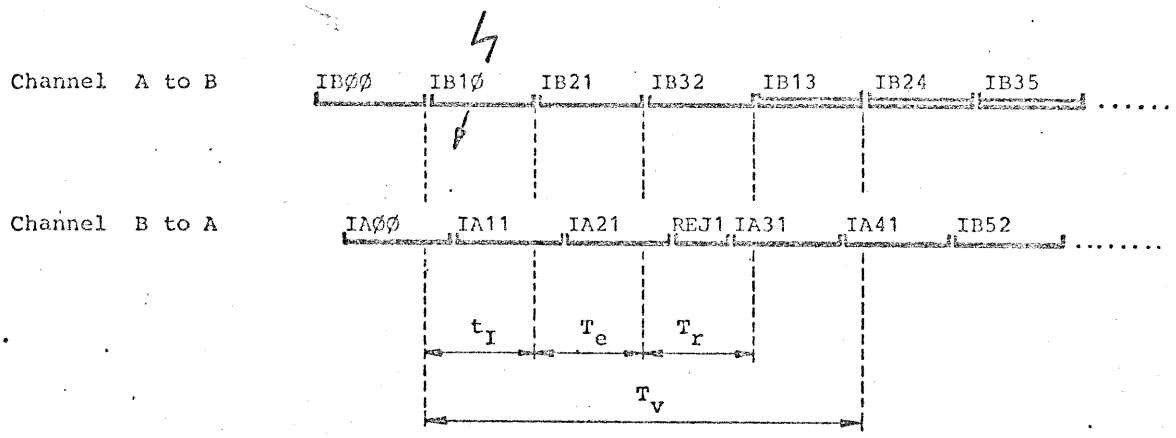


Fig. 6 : Components of the virtual transmission time for the saturated case.

Using the same techniques as in the preceding sections we arrive at the following formulae for the various components of the virtual transmission time in the saturated case (c.f. Fig. 6) :

$$E[T_e] = (1-p_B^{M+1})(1-p_B)^{-1} t_I + p_B^{M+1} t_S \tag{41}$$

$$E[T_r] = t_I + t_S \tag{42}$$

$$E[T_v] = \frac{1+Mp_B^2}{1-p_B} t_I + \frac{p_B^2}{1-p_B} t_S + p_B E[T_e] + \frac{p_B}{1-p_B} E[T_r] \tag{43}$$

Herewith, the throughput T of information bits per unit of time is given by :

$$T = \ell \,/\, E[T_v] \tag{44}$$

4. CALCULATION AND SIMULATION RESULTS

Here we present typical numerical results obtained by the analysis described in Section 3. Since the results are approximative we compare them to values obtained by the simulation of a data link model according to Fig. 1. Within the simulation program the information exchange phase of HDLC, Asynchronous Balanced Mode with the optional function REJECT was implemented in full detail [4]. All simulation results are shown with their 95% confidence intervals.

4.1 Saturated case

Figure 7 shows the throughput of information bits in case of saturated queues relative to the transmission rate as a function of the message length, i.e. the length of the I-field of the frames. In case of no errors the normalized throughput follows, of course, the ratio of ℓ and Vt_I. For non-zero error rate the throughput curves show distinct maxima, which are typical for link control procedures employing error detection and retransmission strategies [10]. The explanation of this behavior is obvious : For short message lengths the block-error probability is low but the relative overhead due to the flag, address, control, and frame checking sequence bits is high. For longer messages the relative overhead decreases but the block-error probability and thus retransmission activity increases.

A comparison of calculation and simulation results shows that our approximation yields very accurate results for short message lengths, whereas it tends to slightly overestimate the throughput for long messages. In spite of this inaccuracy it can be generally observed that the location of the maximum throughput value is rather precisely given by our approximation.

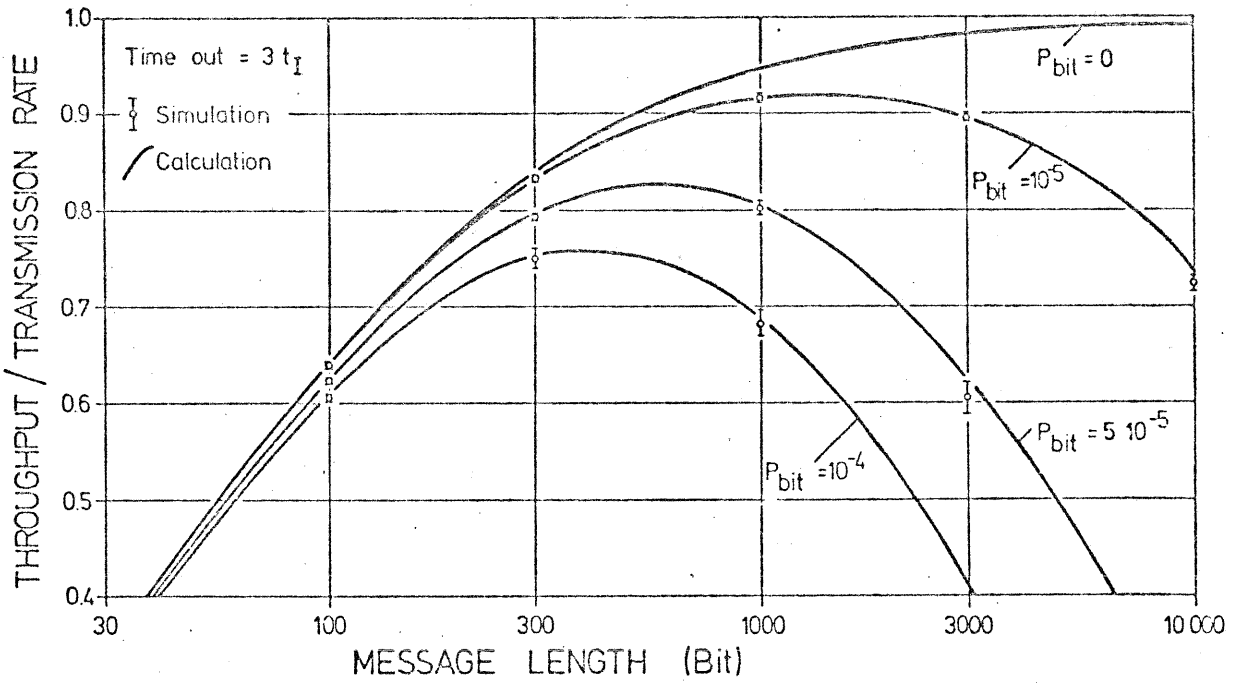


Fig. 7 : Throughput/Transmission rate vs message length.

4.2 Non-saturated case

Here we consider the more realistic case, that the transmission channels are only loaded corresponding to a fraction of their full capacity. Figures 8 and 9 represent the mean transfer time relative to the transmission time of an I-frame as a function of the useful channel load for two different values of the message length, 1000 bit and 5000 bit, respectively.

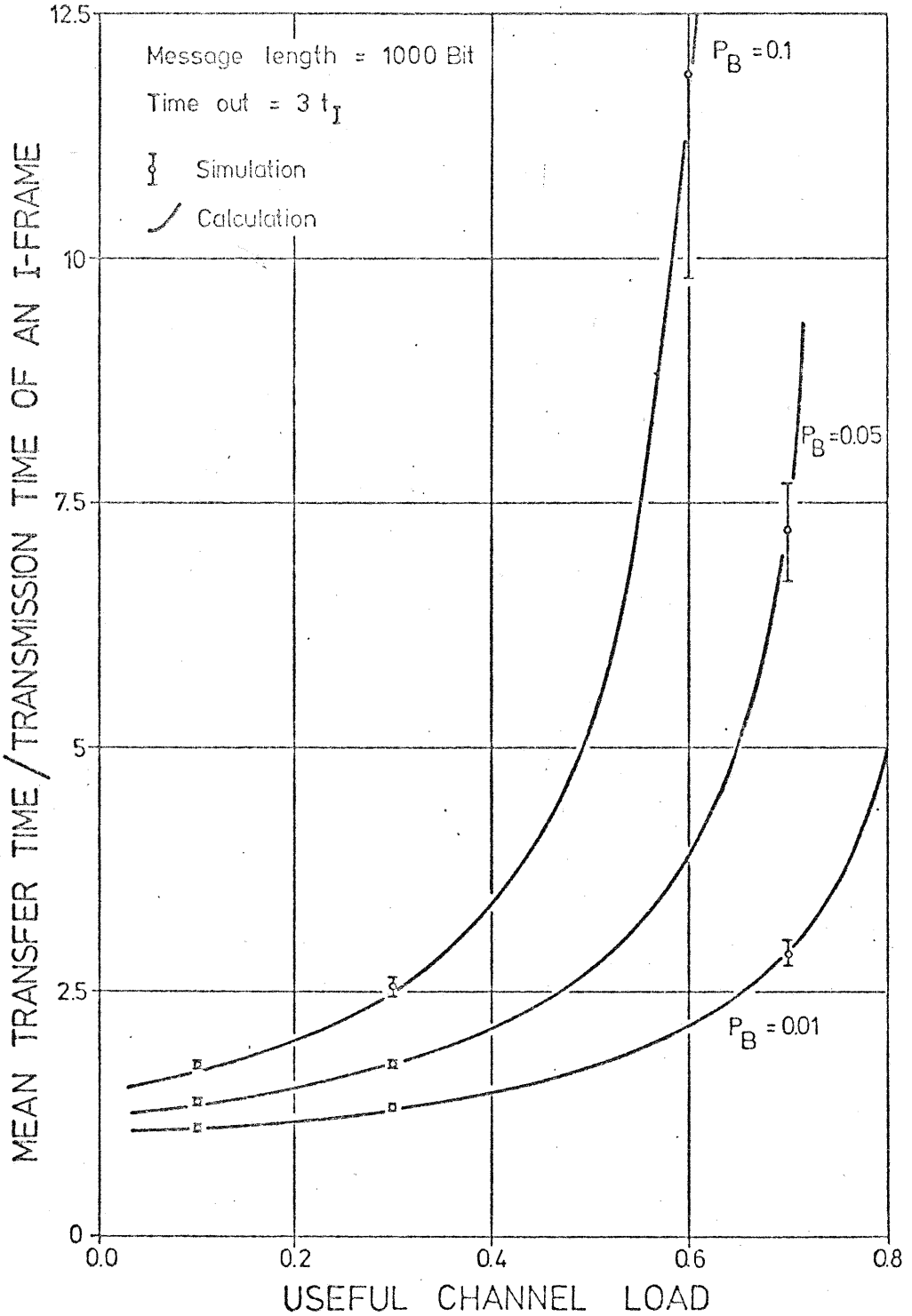


Fig. 8 : Mean transfer time/Transmission time of an I-frame vs useful channel load (message length = 1000 bit)

The useful channel load Y_u corresponds to that portion of the total channel load which is caused by the successfull transmission of information bits :

$$Y_u = \lambda_A \frac{\ell}{V}$$

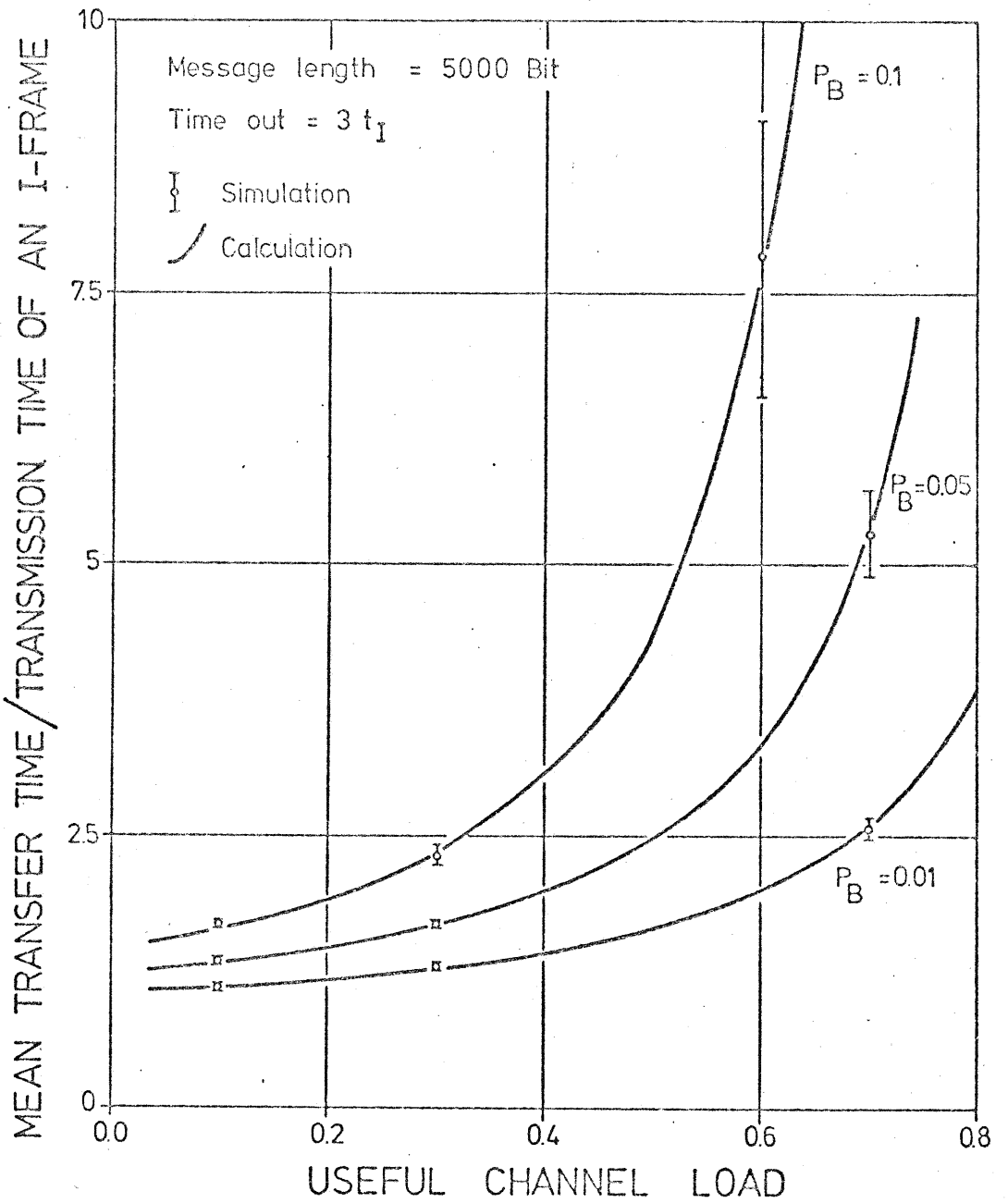


Fig. 9 : Mean transfer time/Transmission time of an I-frame vs useful channel load (message length = 5000 bit)

The duration of the time-out was chosen as the transmission time of three I-frames.

The block-error probability p_B, i.e. the probability that an I-frame is disturbed varies between 0.01 and 0.1.

It should be noted, that even in case of an abnormally high block-error probability of 0.1 our analysis yields rather precise results.

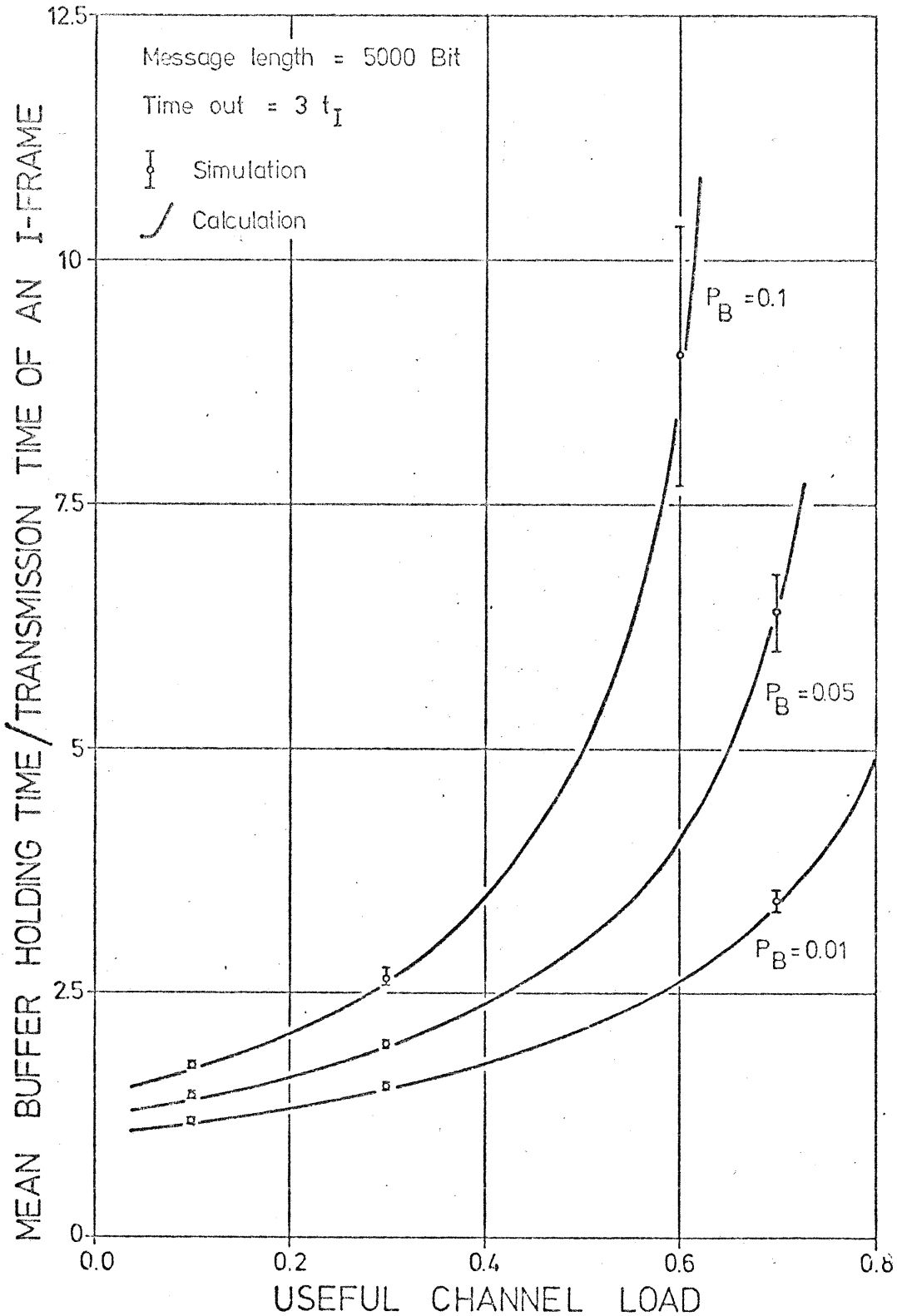


Fig. 10 : Mean buffer holding time/Transmission time of an I-frame vs useful channel load (message length = 5000 bit)

Figure 10 shows the mean buffer holding time relative to the transmission time of an I-frame as a function of the useful channel utilization for a message length of 5000 bits and three different values of the block-error probability.

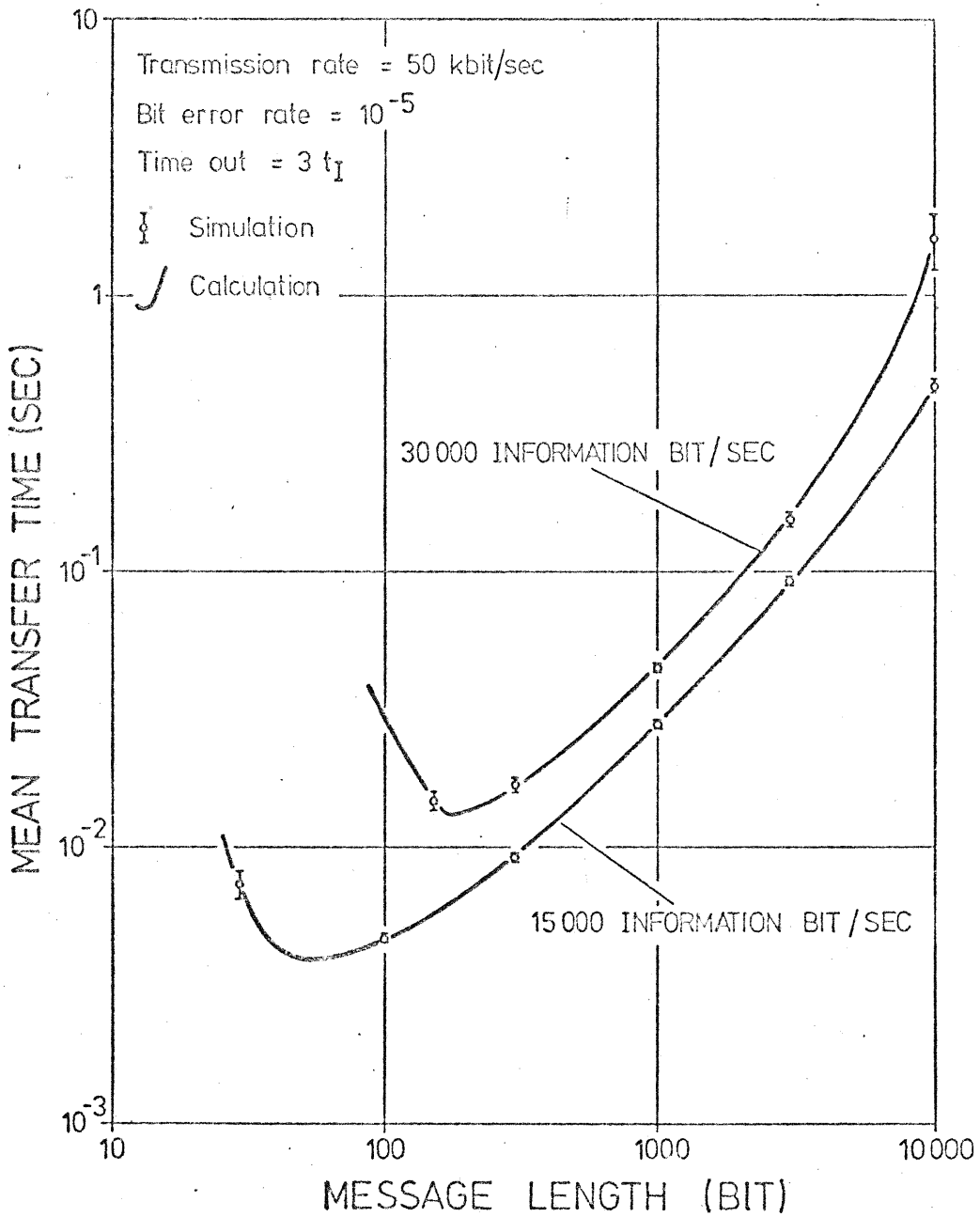


Fig. 11 : Mean transfer time vs message length.

Finally, Figure 11 represents the mean transfer time of messages vs the message length for two fixed values of the throughput : 30,000 and 15,000 information bit/sec, respectively. The parameters of the transmission channel have been chosen as follows : transmission rate 50 kbit/sec, bit-error rate 10^{-5}, $t_{out}=3t_I$. The curves show distinct minima which can be globally explained by the large relative overhead per I-frame (flag, address, control, FCS field) for short messages and the increase of the virtual transmission time for long messages. The latter effect is caused by three reasons : (i) the real transmission time increases, (ii) the block-error probability increases, and (iii) the error recovery takes longer. A comparison of Figures 7 and 11 shows that the minimum transfer times are obtained for significantly shorter messages than the maximum throughput values. This should be taken into account when choosing the parameters of a specific application.

5. CONCLUSIONS

The major conclusions drawn from this investigation are:
It is possible to evaluate the performance of the Balanced Class of Procedures of HDLC for a wide range of applications by using a rather simple queueing model. The key to the analysis is the introduction of the "virtual transmission time", defined in Section 3.2.1.

Comparisons of our analytic findings with results obtained from the simulation of a data link, where the protocol was fully implemented, indicate that our approach comprises the major effects which determine the protocol performance.

Two findings are particularly interesting: First, the virtual transmission time of the I-frames depends not only on their real transmission time and the error-rates of the channels but also on the traffic intensities in both directions of the link. This fact has not been taken into consideration in former studies of related link control procedures, although it may have a significant impact on performance. Second, throughput as well as mean transfer time show distinct optimum values, depending on the I-frame length, however, the optima are usually located at different values of the I-frame length. This should be taken into account when designing actual protocols.

Acknowledgment

The authors would like to thank Prof. Dr. A. Lotze and Dr. K. Kümmerle for many helpful discussions.

References

[1] Data communication - High level data link control procedures - Frame structure. International Standard ISO 3309.

[2] Data communication - High level data link control procedures - Elements of procedures (independent numbering). International Standard ISO 4335.

[3] HDLC - Proposed Balanced Class of Procedures. Document ISO/TC 97/SC6-No 1444.

[4] Bux, W., Kümmerle, K., Truong, H.L.: Results on the Performance of Balanced HDLC Procedures. National Telecommunications Conference, Birmingham (Alabama), Dec. 4-6, 1978.

[5] Traynham, K.C. and Steen, R.F.: SDLC and BSC on satellite links: a performance comparison. ACM Computer Communication Review, Oct. 1977, Vol. 7, No.4, pp. 3-14.

[6] Gelenbe, E., Labetoulle, J., Pujolle, G.: Performance Evaluation of the protocol HDLC. Proceedings of the Symposium on Computer Network Protocols, Liège (Belgium), Febr. 13-15, 1978, pp G3.1-G3.5.

[7] Masunaga, Y.: Optimal HDLC I-frame structure in a two-way file transfer and IIASA data communication network. IIASA Research Memorandum RM-76-31, May 1976.

[8] Kleinrock,L.: Queueing systems. Vol.I: Theory.J. Wiley&Sons,New York,1975.

[9] Pack, C.D.: The output of an M/D/1 Queue. Operations Research, Vol.23, No. 4, July-August 1975, pp. 750-760.

[10] Chu, W.W.: Optimal message block size for computer communications with error detection and retransmission strategies. IEEE Trans. on Comm., Vol. COM-22, No. 10, Oct. 1974, pp 1516-1525.