# PREVENTING OVERLOAD AND VERIFYING LINK STABILITY IN AN IMPLEMENTATION OF THE DPNSS LAYER 2 PROTOCOL

Arthur W. Berger,[1] James S. Kent,[2] and Michael V. Stein [2]

AT&T Bell Laboratories
1 - Crawfords Corner Rd., Holmdel, NJ 07733-1988
2 - 2000 N. Naperville Rd., Naperville, IL 60566-7033

*ABSTRACT*

The Digital Private Network Switching System (DPNSS) layer 2 protocol specification does not explicitly indicate methods of layer 2 overload control or link stability verification. This paper describes how AT&T's 5ESS - PRX Switch has implemented these features. The key to preventing and controlling overload is to explicitly ignore some received retransmissions of incoming frames. This strategy reliably handles the range of frame transmission strategies that are consistent with the DPNSS protocol. Periodic transmission of specialized layer 2 test frames over the DPNSS link assures link stability.

## 1. INTRODUCTION

This paper discusses preventing and controlling layer 2 overload, and verifying layer 2 link stability, for a Private Branch Exchange or telephone switching exchange using the Digital Private Network Switching System (DPNSS) protocol.

AT&T's 5ESS -PRX Switch provides a DPNSS-compatible interface for use in the Digital Derived Services Network (DDSN), an Intelligent Network for the United Kingdom designed by AT&T for British Telecom.

Some exchanges using the DPNSS protocol retransmit frames at layer 2 before the initial transmission has been acknowledged by the adjacent exchange. Processing these retransmissions increases processor load, and increases delay for subsequent frames. This paper proposes optimizing processor real-time performance by intentionally discarding some transmissions. Our method of doing this leaves the probability of completely missing a frame negligible (less than $10^{-8}$) over a broad range of scenarios. We also describe a manner of delaying frames during processor overload.

Additionally, we present a way to assess the link stability at layer 2. The 5ESS-PRX Switch periodically sends out test frames which should be acknowledged and then discarded. If such frames are not acknowledged, recovery actions are undertaken.

In the 5ESS-PRX Switch, different processors are used for layer 2 and layer 3 protocol processing. Except when we explicitly state otherwise, the subject of our discussion is the layer 2 processor.

## 2. RELEVANT FEATURES OF THE DPNSS PROTOCOL

The DPNSS layer 2 protocol is specified in the British Telecom Network Requirements.[1]

Features of the protocol relevant to this paper are:

| | |
|---|---|
| Logical Channels: | The protocol recognizes up to 60 logical channels, called Link Access Protocols (LAPs), per 2 Mbps link. Timeslot 16 carries signaling messages for all LAPs. |
| Retransmissions: | Frames must be retransmitted until an acknowledgement is received, or for a minimum of 500 ms and a minimum of 64 transmissions, after which the LAP will be reset.<br><br>The minimum interval between retransmissions is determined by the requirement that there be at least one flag between frames. In practice, frames have been observed to arrive separated by as little as 1.0 ms. |
| Window Size: | The protocol has a window size of 2 at layer 2; each frame has a sequence number of either 0 or 1. |
| Link Stability: | The protocol does not explicitly mention checking link stability at layer 2. |
| Test Frames: | There are no test frames specified at layer 2. The protocol requires that null frames should be acknowledged and we have exploited this feature to provide a link assurance mechanism.[2] |
| Overload Control: | The protocol provides no mechanism for an exchange to indicate a busy condition to other exchanges, and no explicit overload controls.[3] |



CH2682-3/89/0000-0175 $1.00 © 1989 IEEE

## 3. OBJECTIVES FOR THE DPNSS LAYER 2 IMPLEMENTATION

Our objectives in implementing DPNSS layer 2 are listed below. They assume that frames are transmitted 64 times, and that at least 70% of transmitted frames will arrive without error at the far end of the link.

1. Have a secure layer 2 interface. The DPNSS protocol specification contains no explicit reliability criteria; we chose the reliability objective of one lost frame in $10^8$, which is our objective for ISDN applications.
2. Regularly exercise all links and LAPs to quickly discover those which are not operational at layer 2.
3. Maximize performance of the layer 2 interface by:
    a. Retransmitting slowly enough so that a frame can be received at the far end, and its acknowledgement received by the sender, before the sender retransmits.
    b. Under overload conditions, throttling the adjacent exchange by acknowledging frames as slowly as possible without causing reset.
    c. Creating a short FIFO buffer of incoming frames, and dropping without processing or acknowledging those frames which encounter a full buffer. This buffer is sized to keep the probability of frame loss to below $10^{-8}$. Subject to this constraint, maximum buffer size is chosen to minimize delay.

## 4. CHOOSING AN APPROPRIATE TRANSMISSION RATE

The DPNSS protocol specification permits wide latitude in choosing the frame retransmission rate. The optimal retransmission strategy has been shown to be:[4] the time between successive retransmissions should be just long enough to allow a non-errored frame to be sent to the far end, and its acknowledgement returned to the sender.

Observations of the DDSN Telecom network and experience with ISDN applications have led us to estimate an appropriate time between retransmissions to be 50 ms.

Moreover, when the 5ESS-PRX Switch is in overload, it automatically reverts to 200 ms retransmission to conserve resources.

## 5. HANDLING DIFFERENT RATES OF FRAME ARRIVALS

The transmitting exchange may send copies of frames contiguously; thus the link may be 100% utilized for significant periods of time if the transmitter has a non-empty queue of frames. To avoid congestion by possible waves of arrivals, the processor uses a limited service, cyclic schedule so that other tasks such as transmission of frames to the adjacent exchange are not affected.

To avoid needlessly resetting the link, an initial idea would be to buffer all copies of a frame by designing the receive buffer to hold hundreds of frames. However, it is not necessary that all 64 frame copies be received to meet the reset criterion of $10^{-8}$, even for frame transmission error probabilities as high as .3. Moreover, having such long frame queues would lead to high delay for particular copies that are processed. A more viable idea is to choose a buffer size large enough to meet the reset criterion over projected traffic patterns, yet small enough so that the delay is not needlessly increased.

### 5.1 Traffic Scenarios

As described above, from the viewpoint of the receive buffer, frames from some exchanges arrive at a deterministic rate that is faster than rate at which frames are taken from the FIFO buffer and processed. Viewing the buffer as a deterministic queueing system with arrival rate greater than the processing rate, then, modulo an initial transient, the buffer either will be filled or will have just one empty space. This applies whether the buffer holds 10 or 100 frames. Thus, modulo an initial transient, for the deterministic case, the buffer size does not influence the probability that a frame will we blocked by encountering a full buffer, and discarded. Adequate buffer space is still needed, however, for periods when the processor interrupts its clock schedule for other tasks.

In addition to traffic patterns seen to date, we considered other potentially stressful situations possible under the DPNSS protocol.

We have based this analysis on a transmitter that uses a round robin schedule across 30 to 60 active channels, and where frame interarrival times vary stochastically according to a hyperexponential distribution.

### 5.2 Probability of Reset

We calculate the probability of reset as the probability that no (non-errored) acknowledgement (ACK) is processed by the PBX, given that $N$ copies of the frame are sent. $N$ is less than 64 and is chosen to capture the effect of ACKs that arrive at the PBX after the 500 ms limit and the effect of periodic audits by the processor. To obtain an analytic expression for the Prob(reset), we make the independence assumptions:

- The probability of a transmission error in a frame or ACK is independent of the probability of a transmission error in any other frame or ACK.
- The probability that a frame or ACK is blocked at the 5ESS-PRX Switch or at the PBX is independent of the probability that any other frame or ACK is blocked.

- The probability of a transmission error is independent of the probability of blocking.

Define the notation:

$N$ - number of copies sent by PBX, $N < 64$.

$b_{5e}$ - probability a frame is blocked at the incoming FIFO buffer to the processor.

$b_{pbx}$ - probability an ACK is blocked at the input buffer of the PBX.

$e$ - probability a frame or ACK is in error, due to transmission noise.

Conditioning on the first good copy processed by the processor, we obtain:

$$P(reset) = [\,1 - (1 - b_{5e})(1 - e)\,]^N + \qquad (1)$$

$$[\,1 - (1 - e)(1 - b_{pbx})\,] \cdot$$

$$\frac{[1 - (1 - b_{5e})(1 - e)(1 - b_{pbx})]^N - [1 - (1 - b_{5e})(1 - e)]^N}{b_{pbx}}$$

In equation (1), $e$ and $b_{pbx}$ are exogenous inputs; however, $b_{5e}$ is influenced by the choice of the size of the waiting area for the FIFO buffer, as well as by other factors. To approximate $b_{5e}$ we use an $H_2/M/1/K$ model. That is, we assume frames arrive according to a renewal process with a 2-stage hyperexponential interarrival time distribution, the service times are exponential and independent of one another and of the arrival process, and the system (buffer plus server) has finite capacity equal to K.

The actual service times are closer to a 2 point mass distribution. The exponential assumption yields service times with greater variance than the real service times; this causes greater delays and a greater number of frames in the system, in turn causing higher blocking and yielding a high estimate for the probability of reset.

Figure 1 shows reset probability versus receive buffer size for the four cases given in Table 1. The parameter values were selected to give a range of behavior. The parameters are exogenous inputs to the model; rows 2, 3, and 4 of Table 1 determine the $H_2$ distribution of the frame interarrival times, row 5 most directly influences the delay described in Section 5.3, row 6 is input to equations (1) and (2), and row 7 is input to equation (1). For these cases, a buffer size of 11 or more will meet the reset criterion. In total, we considered over 300 cases, including cases where the PBX adopts the same strategy of a limited receive buffer and where the blocking probability at the PBX is assumed to be equal to that at the 5ESS-PRX. (This is a pessimistic assumption since the 5ESS-PRX retransmission rate is chosen to be 1 per 50 ms, which is slower than that of many DPNSS PBXs.) Although one can pick parameter values such that no choice of buffer size meets the reset criterion, we found that a buffer size of 15 is sufficient over the practical range of scenarios.

**Figure 1.** PROBABILITY OF RESET

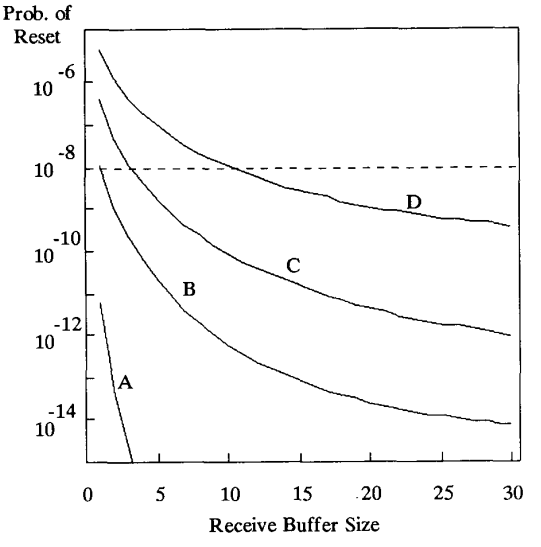



**TABLE 1.** REPRESENTATIVE SCENARIOS USED IN RESET AND DELAY TABLES

| Case | A | B | C | D |
|---|---|---|---|---|
| mean interarrival time [ms] | 3. | 3. | 3. | 3. |
| coefficient of variation of interarrival times | 1. | 3. | 3. | 3. |
| skewness of interarrival times | 2. | 10. | 10. | 10. |
| Number of active logical channels | 1 | 60 | 30 | 30 |
| Probability of frame in error | .01 | .01 | .30 | .01 |
| Probability of blocking at PBX | .3 | .3 | .3 | .6 |

### 5.3 Calculation of the Delay Through the (Layer 2) Processor, $d$

Let $d$ denote the mean delay through the processor, which we define to be the expected time from the arrival of the first copy through the processing of the first non-blocked, non-errored copy. $d$ can be partitioned into two components:

1. mean delay from the arrival of the first copy to the arrival of the first non-blocked, non-errored copy.
2. mean queueing delay in the processor's FIFO buffer of the first non-blocked, non-errored copy plus its service time by the processor.

We assume the number of copies to arrive before the first non-blocked, non-errored copy has a geometric distribution, and the queueing delay is obtained from the $H_2/M/1/K$ model above. For $n$ active channels operating in round robin with mean interarrival times of frames of $\lambda^{-1}$ ms, then:

$$d = \frac{[1 - (1 - e)(1 - b_{5e})]}{(1 - e)(1 - b_{5e})} \cdot n\lambda^{-1} + \qquad (2)$$

*mean ( queueing time plus service time )* [ms]

The choice of the FIFO buffer size has two, counteracting effects on $d$. An increase in the buffer size increases the mean queueing time, but also decreases $b_{5e}$, which decreases the first component of $d$. In general, for small list sizes the first component dominates, while for large list sizes the second dominates, particularly if the arrival rate is greater than the service rate.

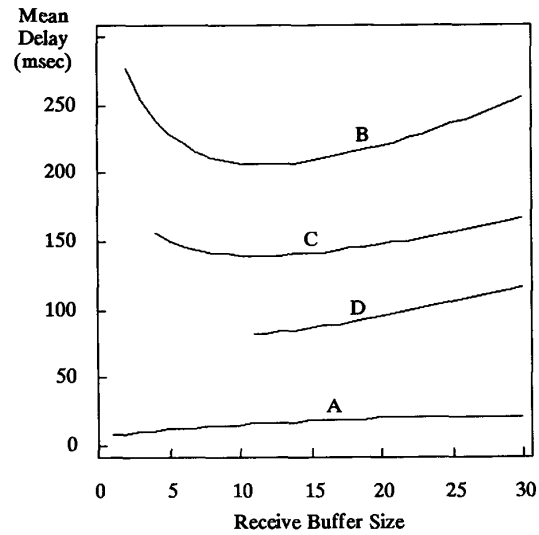
For the same four cases as in Figure 1, the delay $d$ is shown in Figure 2 for those buffer sizes that meet the reset criterion. Note that the high delay shown in Figure 1 for Case B is not from queueing in the processor, but from the delay between arrivals of copies of a given frame. With 60 active channels operating in round robin, and with 3 ms between frames, if a given copy is in error or is blocked, then 180 ms will elapse on average before the next copy arrives. Figure 2 suggests, and we find in general, that a buffer of 15 to 20 frames allows the reset criterion to be met for a broad range of traffic parameters without needlessly increasing the delay.

## 6. HANDLING PROCESSOR OVERLOAD

When an overload situation occurs, the layer 2 processor must:

1. Continue to process frames from all exchanges,
2. Throttle incoming frames in accordance with the activity on the link (high-usage links get throttled more than low-usage ones), and

**Figure 2.** MEAN DELAY FROM ARRIVAL OF THE FIRST COPY THROUGH THE ARRIVAL AND SUBSEQUENT QUEUEING AND PROCESSING OF THE FIRST NON-ERRORED, NON-BLOCKED COPY




3. Optimize processor efficiency.

The overload strategy addresses the throttling of incoming frames, the throttling of outgoing frames, and the "support" of a layer 3 processor that is in overload (recall that layer 2 and layer 3 use different processors; one may be in overload without the other being overloaded).

In all cases, the layer 2 processor will continue to service frames incoming to the 5ESS-PRX Switch and outgoing from the switch at an equal rate. Since each frame in one direction will generate an acknowledgement in the other direction, this strategy has proved efficient for servicing other layer 2 protocols.

The overload strategy for incoming frames is provided by the use of a short incoming FIFO buffer, as described above.

The Retransmission rate for outgoing frames is slowed from 50 ms to 200 ms during layer 2 processor overload.

The layer 2 processor takes actions to reduce the load on the layer 3 processor when it is in overload. In the event of layer 3 overload, we throttle incoming frames at layer 2 so that only one frame every 400 ms is supported per LAP. We do this by alternating every 400 ms between acknowledging only frames with sequence number 0 and acknowledging only frames with sequence number 1.

## 7. IDLE LINK ASSURANCE (ILA)

The DPNSS LAP is responsible for the reliable delivery of client messages at layer 2 of the OSI model. It contains protocol-specific procedures to detect fault conditions (e.g., missing or corrupted frames) and respond accordingly to initiate corrective actions. This recovery may be a retransmission of an outstanding but unacknowledged frame, a complete initialization and re-synchronization of the LAP involved or, in extreme cases, the reset of the entire link (i.e., all LAP instances reinitialized). Most of these protocol-specific error detection methods rely on message traffic to detect failures. If the far end protocol driver is in an abnormal state, and the near end protocol has no currently outstanding unacknowledged frames or new frames to transmit, then this pending error condition will go undetected until message traffic is present.

The Idle Link Assurance (ILA) functionality periodically scans the outgoing message queues of the in-service LAPs sending a special test frame over each LAP it finds with no current frame activity. This transmittal stimulates the currently idle LAP, using the normal transmit/receive protocol on both ends of the link. Failure of an ILA frame transmission provides more timely detection of off-normal states than waiting for actual call traffic to fail.

The ILA test frame is a normal information frame, but with a zero information field. The protocol will receive this frame, respond with the specified acknowledgement frame, update its protocol variables, and discard the frame. No error is attributed to this layer 2 exchange.

The ILA process is invoked periodically; it is inhibited during overload, so as not to add significant traffic or delay to the link. ILA test frames are sent as part of the normal protocol retransmission process. Frames are transmitted for all LAPs on a link with a separation of 50 ms between adjacent LAPs. This allows rapid examination of link stability without creating a burst of traffic on the link which could interfere with normal frame processing.

Idle Link Assurance is also used to verify that the link is still operational when the switch detects a high percentage of out-of-service LAPs. In that case, ILA is triggered to determine if any of the remaining in-service LAPs are still alive. If none of these "in-service" LAPs respond within the normal protocol timeout limit, the entire link is reinitialized.

## 8. CONCLUSION

We have designed an efficient strategy for dealing with DPNSS-compatible exchanges that rapidly retransmit frames; we have implemented this strategy in the 5ESS-PRX Switch to provide high performance and reliability in the DDSN network for the United Kingdom. The strategy is to optimize processor efficiency by ignoring as many frames as possible while keeping the reset probability below $10^{-8}$, and to periodically verify link stability.

*REFERENCES*


[1] Digital Private Network Signalling System No. 1 (DPNSS 1), Section 3: Link Access Protocol", BTNR 188, Issue 4, March 1986.

[2] Ibid, sec. 4.2.3.2.

[3] Ibid, sec. 3.6.

[4] Swedan, S. E., D. G. Smith, and J. L. Smith, "Effect of a Gap Period Mechanism on the DASS1 Protocol Performance", 5th ITC Seminar on Traffic Engineering for ISDN Design and Planning, Lake Como, Italy, May 4-8, 1987.