Time-Out Management in Packet Switching Networks

H.R. van As

Dept. of Communications, University of Siegen
Siegen, Federal Republic of Germany

ABSTRACT

The paper addresses design and operational guidelines for packet retransmission strategies necessary to support recovering from lost or damaged packets, a scheme typically applied in packet switching networks. In simulating the so-called time-out mechanism, it has been observed that traffic performance results obtained from analytical models which are based merely on a distribution function, are far too optimistic. Motivated by this observation an extensive simulation study has been carried out. Performance impacts of several time-out management strategies are presented and relevant parameters are discussed. Moreover, it will be shown that it is very important to control the packet retransmissions in an adaptive way. Otherwise, network congestion may render the recovery function into an additional source flooding the network with blind traffic and thus amplifying network congestion up to a final crash.

1. INTRODUCTION

Interprocess communication in packet switching networks has been based on reliable and efficient protocols. To overcome occasional loss or damage to packets, protocols make use of an error-recovery function which is typically implemented at various protocol levels. In particular, transmission errors and input buffer overflows are recovered node-by-node according to a link level protocol (e.g. HDLC), whereas packets lost along the path from source to destination due to malfunctioning are recovered globally, either at the user transport level (end-to-end) or at the network transport level (entry-to-exit). Some proposed congestion control mechanisms even drop packets from the network and thus rely on an adequate error-recovery function [1,2].

A common feature of the procedures implemented at the different protocol levels is that they all make use of the so-called time-out (TMO) mechanism. The time-out mechanism retransmits automatically those packets whose acknowledgements (ACK) do not return to the source within the specified period of time and thus are presumed to be lost. The choice of the time-out value may become a critical performance factor when acknowledgement delays are prolonged due to network congestion. A long interval will retard the recovery action resulting in a considerable packet delivery delay and long buffer occupancy time for packet copies.

A short interval, however, implies many unnecessary retransmissions, which cause an additional network load. Load amplification even may result in an unavoidable crash. Due to this potential deterioration of network performance, an efficient time-out management plays a fundamental role in the specification of an overall congestion control plan for packet switching networks.

In the past the time-out mechanism has been considered in switching node models in that a probabilistic decision is made for an exponentially delayed retransmission [3,4,5]. The influence of time-outs on the Send-and-Wait protocol has been analyzed in [6]. Time-out retransmissions appear in [7] as one of the protocol parameters which have an impact on the efficiency of interprocess communication protocols. In [8,9] the time-out interval has been a basic parameter in the analysis of network entry-to-exit window flow control. A general treatment on the subject has been given in [10], in which also the need for an adaptive scheme has been identified and propositions have been made. In [11] the instability phenomenon has been noticed, too.

2. MODELING

2.1 The packet switching network

Traffic performance of a packet switching network is mainly determined by communication channel capacity, processing time, nodal storage, and operating rules which allocate the hardware resources to the packets (e.g. communication protocols, congestion control).

Fig. 1: Model for a virtual circuit connection in a packet switching network
In modelling such a network, e.g. [12, 13] communication channels are represented by queueing systems whereas nodal processors often can be described by switching points only. This is due to the fact that packet processing delay is typically substantially less than the delay incurred at a communication channel. This results in a queueing network model as shown in Fig. 1 in which one source-destination relation - according to the virtual circuit concept with all packets routed along the same path - has been depicted in detail with a forward path for packets and a backward path for their acknowledgements. On top of this basic model other refinements are possible [14].

Since we are particularly interested in global time-out management, we assume that all transmission errors and input buffer overflows are recovered by the link level protocol. Moreover, no packets are dropped from the network due to congestion control. In that case, packets can be lost only due to malfunctioning (e.g. a switching node crashes or a packet header has been damaged). Under normal conditions the probability of loss is very small. Therefore, the packet switching network has been modelled without loss, so that the pure effects of unnecessary time-out traffic on network performance can be studied.

2.2 Time-out management model
In Fig. 2, a functional diagram of the time-out management model has been shown. Packets arrive with rate $\lambda$ at a black box system representing the packet switching network. Before entering, a time-out event is set. If it is not reset by the corresponding acknowledgement this timing expires, and an additional packet - its copy - is submitted to the black box system. At the same time the next time-out event is set. All packet copies may arrive with rate $\lambda_{TMO}$, so that a total traffic rate $\lambda = \lambda + \lambda_{TMO}$ is offered to the black box system.

The considered network configurations have been put together in Fig. 3. Configuration Q10 reflects a network consisting of only one communication channel and an acknowledgement path without delay. In configuration Q11 acknowledgements are returned over the backward channel of a duplex communication connection. They are either piggybacked in the header of packets travelling in the reverse direction or if no data packets are available special ACK-packets are sent. Therefore, it is assumed that packets arriving at their destination are returned to the source as acknowledgements.

Packet copies, however, will make use of the forward path only. Analogous considerations hold for the configurations Q22 and Q33 consisting of two resp. three communication channels between source and destination. Additional network traffic interfering with the considered source-destination pair traffic has been taken into account by a separate traffic stream at each communication channel.

In Fig. 4 time-out handling has been illustrated. Timing is set at the arrival instant of a packet and when a packet copy has been generated. In case that a timer expires, time-out handling may be delayed depending on the state of a so-called time-out window (see section 2.3.3). A timer reset, initiated by the reception of an acknowledgement, can be made both in the timing part and in the time-out handling queue.

2.3 Time-out management strategies
For the definition of an adequate time-out management strategy the following issues can be identified:
- choice of the time-out interval
- number of time-outs allowed per packet
- time-out handling control
2.3.1 Choice of the time-out interval

With respect to the determination of the time-out interval a distinction can be made between a fixed scheme and an adaptive scheme.

Fixed time-out strategies turned out to be very sensitive to load variations. A temporary increase of the acknowledgement delay may trigger an avalanche of time-outs and without control the data traffic breaks down. However, if the time-out interval is adapted to recent information about the network state, a good traffic performance can be achieved. Realization variants differ in:

- the kind of information used
- estimation method for the next time-out interval
- update frequency, both for the state information and for the interval estimation.

Useful network state information can be obtained by measuring acknowledgement delays (also referred to as roundtrip delays). If the arrival time of a packet has been stored, the delay measurement can be made every time an acknowledgement is returned. Otherwise, test packets must be sent out periodically. The test packet method must be used in any case if no data packets have been sent away during a specified period of time.

Based on these measurements the next optimal time-out interval must be estimated. In case of traffic stationarity, this can be done by averaging over recent measurements. As will be shown, there exists an optimal range for the size of the measurement set. In an environment with time-dependent load variations a more sophisticated estimation method is certainly advantageous.

Due to the overhead required to evaluate the acknowledgement delay measurements, it may be preferred to make time-out interval estimates periodically instead of each time a new time-out must be set. However, whereas processing overhead is improved by a periodic update, traffic performance decreases.

In addition to acknowledgement delay measurements, congestion indicators provided by flow and congestion control mechanisms may serve as valuable information to prolong time-out intervals. Furthermore, the number of time-outs in a given interval may indicate either congestion or a hardware failure.

2.3.2 Number of time-outs per packet

A limitation of the number of time-outs per packet is a good policy. Then, if a packet times out more than once, it can be assumed that packets either are delayed by congestion somewhere in the network or a faulty part of the system (e.g. node, trunk, receiving host) makes packet delivery impossible. Whereas in the former case an acknowledgement can be expected soon, it will fail to come in case of a fault. Therefore, the status of the connection path used must be determined by inquiry packets, which are handled with priority.

To overcome the overhead of sending inquiry packets when acknowledgements have been delayed, subsequent time-out intervals for the same packet can successively be prolonged. For example, binary prolongation \(-2, 4, 8, \ldots\) times the current time-out interval - can be applied.

2.3.3 Time-out handling control

Since the network state may have changed completely between the time that the time-out has been set and the time it becomes active, time-out handling control is advantageous. The parameter of control is the number of packets - belonging to the considered source-destination pair - which are currently managed by the time-out function. Therefore, it is an easily accessible quantity. Purpose of the control, as already shown in Fig. 3, is to delay the time-out handling if the number of packets waiting for an acknowledgement has exceeded a given limit. In analogy with window flow control, this mechanism will be denoted by time-out window control and is operated independently from flow control mechanisms.

By delaying time-out handling - which turned out to be quite short - two positive effects are achieved. First the network, which is presumably overloaded in that case, is freed from additional packets and secondly, a part of the delayed time-outs can be reset before being handled. In order to prevent a deadlock in case that none of the packets of the considered connection will be acknowledged, an additional time supervision is necessary each time the time-out window control becomes active.

3. ANALYSIS AND SIMULATION

3.1 Analysis based on a distribution function

In [15] it is shown that for an open queueing network with product-form solution and with strictly forward routing, the probability density function of the End-to-End delay can be expressed by a rational Laplace transform. Furthermore, if the assumption is made that time-outs occur independently of the actual load situation in the network, the time-out traffic \(\lambda_{TOO}\) can be derived from a distribution function. This approach has already been practised in previous papers on the time-out subject. By a combination of both achievements, network traffic interference can be included into analytical studies on the choice of the time-out interval.

For this, we consider a packet switching network - like that of Fig. 1 - consisting of \(N\) communication channels. Queueing systems representing the communication channels are of the infinite Markovian type and a First-In First-Out queue discipline is applied. A given traffic matrix is offered to the network, but as we are particularly interested in the data communication from source A to destination B, network topology and related traffic matrix are projected on one of the configurations of Fig. 3.

Therefore, all traffic which belongs not to the considered connection is assumed to include already the time-out traffic and only for the source A a fresh traffic rate \(\lambda_{OA}\) is specified. This rate is used as a starting value for an iterative procedure, so that the final traffic rate offered by source A to the network will increase to \(\lambda_{OA} + \lambda_{TOO}\).

The main point of the procedure is to determine the acknowledgement distribution \(\Phi(T_{ACK})\). In the underlying open queueing network, flow times at individual queueing systems are independent of each other and obey a negative exponential distribution.
Thus, the Laplace-Stieltjes transform (LST) of the acknowledgement distribution - splitted up in forward path terms and backward path terms - is determined by

$$F_{ACK}(s) = \prod_{i \in AB} \frac{1/tfi}{s+1/tfi} \cdot \prod_{j \in BA} \frac{1/tfj}{s+1/tfj}$$

The mean flow time $t_{fi}$ of the communication channel $i$ in the forward path (the same expression holds for the communication channel $j$ in the backward direction) can be obtained by a well-known relation of a M/M/1 queueing system:

$$t_{fi} = \frac{h_i}{1-\rho_i}, \quad i \in \text{forward path AB}$$

with mean exponential service time $h_i = 1/\mu i [\text{sec}]$ communication channel capacity $C_i [\text{bits/sec}]$ mean exponential packet length $17\mu [\text{bits}]$ traffic load $\rho_i = \frac{h_i}{C_i}$

For a vector of external traffic rates $\lambda_Q$ and the traffic routing matrix $Q$, the arrival rate $\lambda_1$ is obtained by the rule of flow conservation:

$$\lambda_1 = \lambda_Q + \sum_{k=1}^{N} \lambda_k Q_k$$

The acknowledgement distribution itself is found by integration of $F_{ACK}(s)$ in the Laplace domain and subsequent inversion:

$$F(T_{ACK} < t) = \frac{1}{s} F_{ACK}(s)$$

In a next step the time-out traffic is determined. For a strategy with a fixed time-out interval $T$ and no restriction on the number of time-outs per packet, the mean number of retransmissions can be obtained by the infinite summation of equidistant points of the complementary acknowledgement distribution:

$$E[N_{TMO}^+] = \sum_{i=1}^{N} F(T_{ACK} > iT)$$

So that the increased arrival rate at which the source $A$ submits its packets to the network is given by

$$\lambda_{TMO}^* = \lambda_{TMO} (1+E[N_{TMO}^+])$$

In summary, $\lambda_{TMO}$ is chosen and its corresponding $\lambda_{TMO}^*$ is calculated from $F(T_{ACK} > iT)$. All traffic rates in the forward path are increased by this amount and in repeating the procedure until convergence the final value for $\lambda_{TMO}$ is obtained.

The End-to-End delay distribution $F(T_{d} < t)$ can be obtained similary if in equation (1) the terms of the backward path are omitted. The mean acknowledgement time $t_{ACK}$, resp. mean End-to-End delay $t_d$ is given by

$$t_{ACK} = \sum_{i \in AB} \frac{h_i}{1-\rho_i} + \sum_{j \in BA} \frac{h_j}{1-\rho_j}$$

$$t_d = \sum_{i \in AS} \frac{h_i}{1-\rho_i}$$
4.2 Comparison between analysis and simulation

The comparison between results obtained by analysis based on a distribution function and those found by simulation shows striking differences. Fig. 5 gives the additional time-out traffic $\Delta_{T_{\text{out}}}$ generated by the offered traffic $\lambda$ if a strategy with a fixed time-out interval $T=16$ and no restriction on the number of time-outs per packet is applied. Characteristic for the results of all the network configurations is the premature avalanche of time-outs in the simulation experiments, whereas the instability predicted by the analysis is found at much higher traffic values. This discrepancy is caused by the inherently positive feedback of the time-out mechanism: if long delays are encountered, many time-outs occur and the increased network load causes in turn an increase in the acknowledgement delays. A distribution function, averaging the time-out traffic, cannot describe this effect properly. As can be expected, the discrepancies - though still significant - decrease if the number of communication channels between source and destination increases. This is a consequence of the smaller variance in the acknowledgement delay. Furthermore, one should notice the sudden increase of time-out traffic in the configurations $Q_{10}$ and $Q_{11}$, whereas for $Q_{22}$ and $Q_{33}$ an initially gradual increase can be observed. In this part, results for analysis and simulation coincide.

4.3 Adaptive time-out strategies

The inherent instability of a fixed time-out strategy can be prevented by an adaptive choice of the retransmission interval. In Fig. 6 the additionally generated time-out traffic is depicted for all 4 network configurations. For one set of curves (AD2) the time-out interval is adaptively set twice the acknowledgement delay which is estimated by an average of 10 recent measurements. For the other set (AD3) the current estimate is multiplied by 3. From the picture several conclusions can be drawn. First, adaptation performs well both for small and for high traffic, because for these traffic values a good estimate for the choice of the next retransmission interval can be made. In the low traffic range packets incur no or only small delays and only in a few cases the time-out interval is exceeded. In the high traffic range a fairly continuous stream of packets is available, which also allow to make a good estimate. However, due to larger fluctuations for intermediate traffic values, it happens more often that the time-out events related to a cluster of arrivals are all based on the choice of a small retransmission interval, and thus more timings will expire before a correction of the time-out interval becomes active. As a second conclusion the positive influence of taking a larger multiplication factor (AD3 instead of AD2) can be noticed. And finally, a large number of communication channels in the acknowledgement path makes the network less sensitive to the time-out mechanism.

Fig. 7 shows the performance improvement achieved by a modification of the basic adaptive time-out strategy:
- a maximum of 1 resp. 2 time-outs per packet
- binary prolonged intervals for successive time-outs of the same packet
- adaptation only above a minimum level (4, 8, 16)
In Fig. 8 the influence of the size of the measurement set A has been demonstrated for the configuration Q10 and adaptive strategy AD3. In case that the interval estimate is based only on the last measurement, a high time-out traffic is generated. Performance becomes better if more measurements are taken into account. However, if the measurement set is too large—i.e., the measurement average contains old and thus irrelevant information—, again more time-out traffic will be offered to the network. Obviously, there exists an optimal value, which is about 10 for the considered configuration. For the limiting case that all measurements are taken into account, the adaptive strategy will become a fixed time-out strategy with unstable behaviour.

Fig. 9 illustrates for two network configurations (Q10,Q11) the case that the choice of the retransmission interval is based on a periodic update instead of an update on-demand. With increasing length of the period, traffic performance degrades. This is a consequence of the fact that a short time-out interval cannot be corrected before the next update. Therefore, a cluster of packet arrivals does much more harm.

In Fig. 10 the network transfer delay $t_d$ has been depicted for network configuration Q11. The curves show the increase in delay—compared to the $M/M/1$-delay—for different adaptive strategies (AD1,AD2,AD3). Furthermore, curves for the strategy AD2 with increasing update period have been given. It should be noticed that due to the adaptive strategy all curves will approximate the $M/M/1$-curve for high traffic values.

4.4 Time-out handling control

Fig. 11 demonstrates that with the time-out window control—here $W=40$—, even the unstable fixed time-out strategy can be managed, however, at the expense of much overhead in resetting time-outs in the TMO-handling queue. The slope of the sudden increase of traffic can arbitrarily be reduced by choosing a smaller window. For comparison dashed curves show the instability limits in case of no control. Especially in relation with adaptive strategies, time-out window control is an efficient completion. This is illustrated in the next two diagrams.

**Adaptive time-out strategies**

Fig. 8: influence of the measurement set
Fig. 8: influence of the time-out update period
Fig. 10: network transfer delay $t_d$
Fig. 12 shows –for the network configuration and the strategies AD2 and AD3 – that the time-out traffic can be reduced by a control window W.

4.5 Interfering network traffic

The interfering network traffic described in the sequel has been made. This demonstrates that the amount of generated time-out traffic decreased for higher values of CV. It turned out for all network configurations, that adaptive time-out management strategy must be taken into account in an overall congestion control plan. In the paper, it has been shown that adaptive time-out schemes completed by the proposed time-out window control will prevent potential network crashes.

REFERENCES


Q.1 (O.G. Soto)

In your paper, the finite source effect of flow central and the global network routing and controls are not considered. Which is the validity of this type of modeling for an actual packet switching network (ie. based on X.25) in which that mechanisms produce a proved and efficient traffic self-regulation before the Time-outs are reached.

A.1 (H.R. Van As)

The different catch words you have used in your single question could be material enough to fill the program of another congress of this size. So, you made it very difficult for me to give a short answer.

First, if we speak about X.25 we must keep in mind that this recommendation defines only the interface between network users and network suppliers and very often offers flow control concepts are used within the network itself. Second, we should keep in mind that there exist two types of packet switching networks: datagram based networks (DATAPAC, DN1) and Virtual Circuit type networks (eg. Transpac, Sna).

Now coming to performance.

Global network renting which makes sense only in a datagram based packet switching network, has been known to give poor performance during overload situations.

Furthermore, individual flow control between users or to a lesser extent individual flow control between network nodes, is very well self-regulating at an individual basis, but since we have hundreds, thousands or more of them, the self-regulating effect is by no means achieved for the total network.

Taking into account all these aspects, it is important to recognize that neither global network routing nor individual flow control alone can prevent network performance degradation, and the problem of time-out management is therefore a practical problem and not at all an academic exercise. As I have shown we get just then a lot of time-outs when we don't need them, ie. at the time we have overload situations in the network, so traffic situations become worse and finally we might have a network collapse.