ATNP/WG2/
WP/
06 July 1995

AERONAUTICAL TELECOMMUNICATIONS NETWORK PANEL

WORKING GROUP TWO

Rome 17.7.95-21.7.95

Minimal solution for Congestion Management

Presented By Henk Hof

Prepared by Leen Goossens

SUMMARY

This document presents a proposal for requirements and recommendations for congestion management in the transport entity. It should be contained in Appendix 8 of the ATN Manual. The document describes a minimal solution for congestion management. This solution exists in the backoff of the transport entity that is sending TPDUs when it presumes that the network is experiencing congestion. Congestion management is exercised by the sending transport entity by limiting the amount of data that may be send before an acknowledgement is received and by adjusting the time the transport entity will wait for acknowledgement before retransmitting a TPDU.

MINIMAL SOLUTION FOR CONGESTION MANAGEMENT

1. TRANSPORT ENTITY CONGESTION AVOIDANCE

The transport entity shall implement congestion avoidance algorithms that interact with the transport protocol for the purpose of reducing load on network resources.

1.1 Advertised window

A transport entity that is receiving TPDUs shall provide the transport entity that is sending the TPDUs with the lower window edge and the size of the *advertised window* by using the explicit flow control mechanism explained in ISO/IEC 8073.

Note.- The ***advertised window*** *is the window advertised by the receiver of the data. It indicates the amount of data that the receiver is willing to accept.*

1.2 Congestion window

A transport entity that is sending TPDUs shall maintain and control the size of the *congestion window.* The congestion window shall always be a subwindow of the advertised window and the lower window edge of the congestion window shall always equal the lower window edge of the advertised window.

Note.- The ***congestion window*** *is chosen by the sender of the data. It indicates the amount of data the sender will send before an acknowledgement is received.*

1.3 Retransmission timer

An ATN implementation shall use a retransmission timer on a per transport connection basis.

1.4 Sending Transport Entity Backoff

1.4.1 Adjustment of the congestion window

1. The size of the congestion window shall be equal to one.

- when the connection is established,
- when a packet needs to be retransmitted because the retransmission timer of the packet expired,
- when the network entity has signalled congestion.

Note 1.- When a packet needs to be retransmitted, it is assumed that the packet got discarded due to congestion. This congestion management strategy is based on the assumption, that packet loss is mostly due to congestion and not due to fact that the packet is damaged. A second assumption is that a packet is lost when no acknowledgement is received for it before the retransmission timer expires. If the retransmission time is too small, the timer will expire before the acknowledgement is received. Therefore it is important that the right value is chosen for the retransmission time.

Note 2.- When the network entity receives an Error NPDU indicating congestion, it can signal this to the transport entity. The manner by which the transport entity is informed is a local matter.

2. If the size of the congestion window is smaller than the size of the advertised window, the size of the congestion window shall be gradually increased until it equals the size of the advertised window. The way the size of the congestion window is increased shall depend on a certain threshold value. This threshold value shall equal half of the size of the congestion window at the end of the previous phase.

- If the size of the congestion window is smaller than the threshold value, the size of the congestion window shall be increased each time an acknowledgement arrives. The size of the congestion window shall be increased with the number of packets that are acknowledged, but it shall not be increased to a value greater than the threshold value or greater than the size of the advertised window.

- If the size of the congestion window equals or exceeds the threshold value, the size of the congestion window shall be increased with one each time “size(congestion window)” acknowledgements are received. The size of the congestion window shall not be increased to a value greater than the size of the advertised window.

Note 1.- Schematically this becomes:

```
size(cong_win) = 2 * α;
while ( transport connection exists )
{
    /* Beginning of a phase */
    threshold = max( size(cong_win ) / 2, 1 );
    size(cong_win ) = 1;
    ackrcvd = 0;
    while ( no retransmission timer expires )
    {
        if ( (acknowl. arrives ) &&
             ( size(cong_win) < size(advertised_win) ) )
        then
            if ( size(cong_win) < threshold )
            then
                size(cong_win) =
                    min(size(cong_win) + #packets_ackn,
                        threshold,
                        size(advertised_win));
            else
            {
                ackrcvd = ackrcvd + #packet_ackn;
                if ( ackrcvd ≥ size(cong_win) )
                then
                {
                    ackrcvd = ackrcvd - size(cong_win);
                    size(cong_win) =
                        min(size(cong_win)+1,
                            size(advertised_win));
                }
            }
    }
    /* End of a phase */
}
```

α is the initial value for the threshold.

Note 2.- The initial value of α is a local issue.

1.4.2 Dynamic adjustment of the retransmission timer

Recommendation.- *The value of the retransmission timer should be dynamically adjusted when the round trip time of packets significantly changes. The round trip time should be estimated by measuring the delay between the sending of a packet and the receiving of the corresponding acknowledgement. This delay should be measured frequently. The delay may not be measured for retransmitted packets.*

The algorithm used to compute the value of the retransmission time should be:

```
estimated_round_trip_time =
    ( 1 - α )  * estimated_round_trip_time + α *
measured_round_trip_time;
retransmission time = β * estimated_round_trip_time;

with 0 < α < 1 and 1 < β
```

β should be a value greater than 1.

Note 1.- The round trip time is the time passed between the moment a packet is sent and the moment the acknowledgement for the packet is received. An approximate value for the maximum round trip time is given by

$$E_{LR}+E_{RL}+A_R+x$$

where *E_{LR} is the expected maximum transit delay local-to-remote,*

E_{RL} is the expected maximum transit delay remote-to-local

A_R is the remote acknowledgement time,

x is a small quantity to allow for additional internal delays, ...

*Note 2.- ((1 - α) * 100)% of the new estimated round trip time is taken from the previous estimated round trip time and (α * 100)% is taken from the last measured round trip time. The recommended value for α is 0.1, so that a temporary change in the round trip time doesn't influence the retransmission time too much.*

Note 3.- estimated_round_trip_time is multiplied by β to take sudden variances in round trip time into account.

Note 4.- The value of β is TBD.

Note 5.- It is not necessary to measure the round trip delay for each packet. One timer could suffice to do the measurements. If the timer is already in use when a packet is transmitted, the round trip delay for that packet will not be measured.

Under TP4 every TPDU has to be retained after transmission until it is acknowledged. An alternative way to measure the round trip time is to timestamp each retained TPDU against a local clock. Then, when an acknowledgement arrives, the current local time less the timestamp on the least recently transmitted and acknowledged TPDU provides the measured round trip delay without needing a single timer. Only a clock is necessary.

How the round trip time is measured is a local matter.

Note 6.- Because the transport entity makes use of the CLNS, there can be great variances in the transit delay experienced by packets. Therefore it is important that the sliding window flow control strategy is coupled with the dynamic adjustment of the retransmission time.

Note 7.- The delay between the sending of a packet and the receiving of the corresponding acknowledgement is not measured for retransmitted packets because there is no way of telling that the received acknowledgement corresponds to the first transmission of a packet or to one of the retransmissions of the packet.

Note 8.- If the remote acknowledgement time is added to the round trip time measured for the connection request packet or for the connection confirm packet, then this sum can be taken as a start value for the estimated round trip delay. If the acknowledgement times are not exchanged during connection establishment, an estimated value for the remote acknowledgement time can be used.