FAIRNESS AND CONGESTION CONTROL IN LAN/WAN PROTOCOL RELAYS

Edmundo Monteiro*, Fernando Boavida* and Vasco Freitas**
edmundo@uc.pt boavida@uc.pt vf@uminho.pt

* Universidade de Coimbra
Laboratório de Informática e Sistemas
Quinta da Boavista, Lote 1, 1º
P-3000 COIMBRA
Tel.: +351 39 701775
Fax: +351 39 701266

** Universidade do Minho
Centro de Ciências e Engenharia de Sistemas
Largo do Paço
P-4719 BRAGA Codex
Tel: +351 53 612234

Abstract: This paper addresses the problem of congestion control and fairness in relay systems for Connection Mode LAN/WAN interworking. To ensure congestion collapse avoidance and fairness against QoS parameters, a channel service algorithm and an acknowledgment algorithm are proposed and evaluated by simulations. Results show that, in a situation of overload and for different traffic patterns with different QoS requirements, the relay's channel service and acknowledgment algorithms can guarantee justice in relay resource utilization and the expected QoS among the active channels.

1. Introduction

In connection mode (CO) networks the issue of congestion control is not as urgent as in ATM or in connectionless mode (CL) networks. The main reasons are the existence of the connection establishment phase that enables resource reservation for the data transfer phase, (preventing congestion due to resource starvation), and the fact that connection mode protocols (such as X.25) have more built-in functions for congestion control than CL protocols (such as IP) do. There are, however, some important issues, related to congestion control in CO networks, that need to be studied, such as those brought up by network interconnection.

Network interconnection became a very important field in past few years. At the present moment, several functional standardization activities address network layer relaying, covering CLNS relaying, CONS relaying and X.25 Packet Level Protocol relaying [Stallings 93]. These functional specifications, or profiles, are being developed by regional workshops (e.g. EWOS) and are at different development stages. Current profile work addresses the interconnection of different types of sub-networks, e.g., CSMA/CD, Token Ring, PSDN, and FDDI, in various combinations.

This paper addresses the problem of congestion control in relay systems for Connection Mode LAN/WAN interworking. For this particular interworking scenario two different relay approaches are possible: *service relays* and the *protocol relays*, classified, according to the ISO taxonomy, as RB5x.yyyy and RC5x.1yyy, respectively [ISO 91, EWOS 90].

The RB5x.1yyy profile family specifies *connection-mode network service relays* (RB). The RC5x.1yyy profile family specify an X.25 *protocol relay* (RC). The 5x stands for LAN sub-network identifiers 51 (CSMA/CD), 52 (Token bus), 53 (Token Ring) and 54 (FDDI). The 1yyy stands for PSDN sub-network identifiers from 1111 to 1231, which include permanent and switched access to the PSDN, and leased line, digital data circuit or ISDN B-channel access methods.

Although the RB relay family represents a more orthodox approach to network layer relaying, because their operation is based on a standardized layer, the RC family have, in despite of their non orthodoxy, some interesting characteristics and functionality. They support both kinds of Network layer services (with appropriate convergence sub-layers) and they are capable of relaying non-OSI traffic, namely X.29 (triple X) PAD traffic or traffic from proprietary communication architectures.

This paper addresses the congestion control and fairness issues in RC5x.1yyy relay family, focusing mechanisms that act in the relaying module and in the X.25 protocol built-in functions.

Some may argue that the study of RC relays is no longer important because X.25 and CONS belongs to the past and now is time for CLNS or IP over high speed sub-networks such as ATM or Frame Relay. This is not true. X.25 and CONS are going to be included in the (almost ready to appear) 4.4 BSD Unix release, side-by-side with the CLNS and TCP/IP [Husemann 92]. X.25 networks have a strong implantation in Europe, with a continuous growth over the last years, and can operate at high speeds [Holleczec 92] providing a communication infrastructure for LAN interconnection, directly supporting the CONS, and also the CLNS and IP.

2. RC5x.1yyy Relays

The internal architecture of an RC5x.1yyy relay is shown in figure 1. In the PSDN side the relay has a pure X.25 [ISO 90] stack: X.21, X.21 bis or ISDN-B channel in the Physical Layer (depending of the PSDN access method), LAP B or I440 in the Link Layer and X.25 PLP as the SNAcP (Sub-network Access Protocol) of the Network Layer.

In the LAN side, the X.25 PLP (also playing the SNAcP role) is used above the MAC (Physical Layer). The LLC Type 2 is used in-between to provide the X.25 PLP a link service with the error free characteristics needed by the X.25 and not provided by the MAC protocol or the LLC type 1. This protocol stack is proposed in the ISO standard IS8881-2 [ISO 87].

The relaying functions between the PSDN and the LAN sub-networks are accomplished on the basis of the X.25 PDUs (X.25 packets).

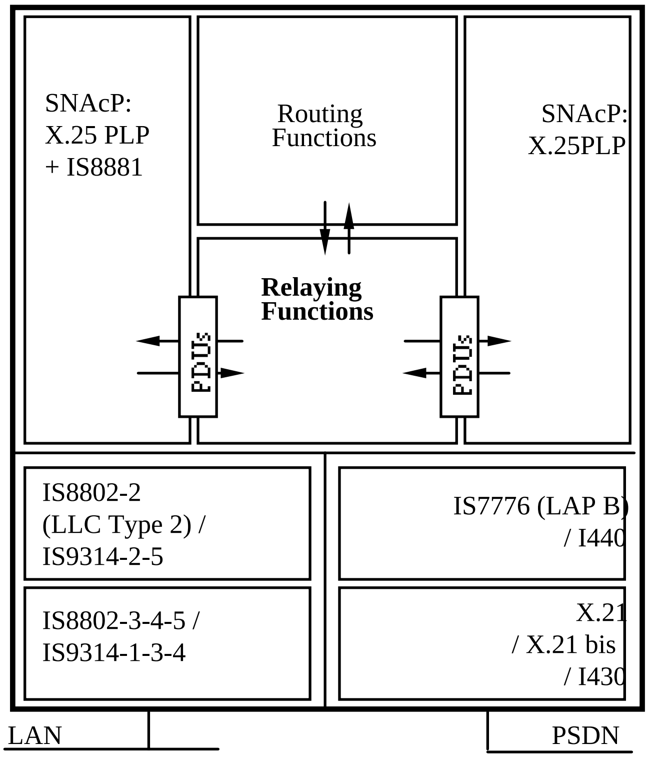

Figure 1 - An RC5x.1yyy relay

Because RC relays operate at the SNAcP sub-layer of the Network Layer, it is possible, with appropriate convergence functions, (SNDCP sub layer) to use them to support the two different types of Network Layer Services (even at the same time), and non-OSI traffic like PAD traffic or traffic from non-OSI architecture (TCP/IP, SNA, DNA, etc.) implanted over X.25 sub-networks.

Some functions of the relaying module are covered by the ISO Technical Report 10029 [ISO 89a]. This Technical Report describes the functions of an X.25 IWU (Interworking Unit) which is a module that provides the way to the interconnection of X.25 PLP based sub-networks. TR 10029 does not regard sub-networks lower layers and thus, refers to all RC5x.1yyy relays.

3. Congestion control in RC5x.1yyy relays

Performance issues are very important in intermediate system design. In RC5x.1yyy relays the issue is of great importance because of the bandwidth difference between the interconnected sub-networks (typically 2 to 3 orders of magnitude).
Several factors affect performance. They can be split into two major groups according to their origin: *processing factors* and *communication factors*. Processing factors deal with the hardware that supports the intermediate system, like its architecture, processing speed, amount of memory, etc. Communication factors deal with the communication itself. They include link speed and error characteristics, communication protocol performance, and issues related to the performance of the intermediate system routing and relaying modules.

The relaying module is responsible for the forwarding of the information between the interconnected sub-networks. Thence its performance has a major influence in the overall intermediate system performance.
When the amount of traffic arriving at the relay exceeds its relaying capacity (due to processing speed limitation, to buffer limitations or to bandwidth limitations) it is said that the relay is *congested*. A severe congestion state is characterized by *throughput* approaching to zero and *transit delay* approaching to infinite (this state is also known as the *congestion collapse* state [Jain 87]).

To avoid performance degradation due to congestion, special functions need to be added to the communication systems (end systems and intermediate systems). These functionalities are known as *congestion control* or *congestion management* [Jain 87] functions.

In relay systems, congestion control must avoid performance degradation due to overload and guarantee *fairness* in resource utilization among active users. Fairness is not a clear concept. It can have many definitions and measurement criteria.
In this work, fairness is evaluated against the user *Quality of Service* (QoS) parameters. If QoS parameters are respected the system is said to be *fair*, otherwise it's *unfair*.

Congestion control functions can be accomplished acting at several levels: using mechanisms built in protocol layers (e.g., window flow control), acting upon the relaying module and acting upon the routing module.

Some routing functions can be used by congestion control. For example, new paths can be selected according to their load or to the delay they introduce. Routing functions must be used in conjunction with other congestion control mechanisms because, by themselves, they are not capable of resolving congestion due to a load larger then the total available bandwidth (in all the available links).

Routing functions are more efficient in connectionless networks than in connection mode ones. In CL networks, routes can be selected in a per packet basis, reflecting load dynamics. In CO networks, routing functions act only during the connection establishment phase and, thence, can't usually be adjusted to load variations during the information transfer phase.

The most sensitive module concerning congestion control is the *relaying module*. It is the only point where congestion control functions can ensure fairness in resource utilization among users. This work is focused in congestion control mechanisms acting in the relaying module and in the X.25 protocol built in functions.

In order to analyze the congestion control mechanisms, a relay model was developed. The model includes all the relaying and X.25 protocol functions needed for congestion control as well as all the end system and sub-network aspects that have influence in throughput and transit delay.

In the relaying module two levels of congestion control policies must be applied: buffer allocation policies and channel service policies.

Buffer allocation policies have great impact in transit delay. In connection mode networks, with window flow control, it's possible to predict the buffer need for each channel, during call setup phase. To avoid packet loss due to buffer starvation, the buffer space allocated to each active channel is determined by the product of the window size, *W* by the maximum packet length, *P:*

$$NeedBuffer_{channel} = W_{channel} \times P_{channel} \qquad (3.1)$$

Channel service policies must guarantee fairness in bandwidth utilization among users, according to their QoS parameter set.

In [Nag 87] and later, in [Demers 89] a channel service algorithm called "*fair queuing*" is proposed and analyzed for gateways in connectionless mode networks. This algorithm ensures that available bandwidth is *equally* shared among source-destination pairs, and ensures protection from ill-behaved users. The fair queuing algorithm doesn't distinguish between users and considers all them equally.

In connection mode networks each channel has its own QoS parameter set to be respected. To ensure that the negotiated QoS minimum throughput is guaranteed by relay intermediate systems a more sophisticated algorithm is needed. Let us define the *Sharing Index* of a channel for the instant *ti*, *SI(ti)*, as:

$$SI(t_i)_{channel} = \frac{EfectiveThroughput}{QoSThroughput} = \frac{\frac{1}{t_i}\sum SizePacketsTransmitted}{QoSThroughput} \qquad (3.2)$$

The next channel to be served at instant *ti* is scheduled by ascending order of the channels *SI(ti)*. This algorithm guarantees that QoS throughput parameters are respected. The evaluation of *SI(ti)* doesn't require much calculation and can be made in parallel with packet transmission to avoid transmission delays. When the *k* packet is transmitted by the relay, the packet finishing transmission time, t_k, is calculated and the *SI* vector is updated for the instant t_k, using the following expressions:

$$SI(t_0) = 1 \qquad (3.3)$$

$$SI(t_k) = \frac{t_{k-1}}{t_k} SI(t_{k-1}) \qquad (3.4)$$

$$SI(t_k) = \frac{1}{t_k}\left(t_{k-1} SI(t_{k-1}) + \frac{PacketSize_k}{QoSThroughput} \right) \qquad (3.5)$$

Expression (3.3) defines de initial conditions for *SI(ti)* evaluation, (3.4) is the *SI(ti)* update expression for the idle channels (with or without packets pending) and (3.5) is the update expression for the channel currently busy with the transmission of packet *k.*

QoS transit delay parameters aren't explicitly controlled by the relay. Once QoS throughput is satisfied QoS maximum delay is automatically bounded by expression (3.6) below. This value can be imposed during the QoS negotiation phase.

$$QoSMaxDelay = \frac{W \times MaxPacketSize}{QoSThroughput} \qquad (3.6)$$

Table 3.1 summarizes the relaying module, X25 protocol, CSMA/CD sub-network, PSDN sub-network and end system characteristics that are to be included in the model.

module	characteristics
Relaying	- Static buffer allocation policy: buffers are allocated according to window size and packet length - Dynamic QoS channel service policy - Segmentation and reassembling capabilities, to match different window and packet sizes
X.25 protocol	- Negotiable window size - Negotiable packet size - Acknowledgment policy: ACKs are sent as soon as buffers are available to receive a new packet
CSMA/CD sub-network lower layers	- Bit rate of 10 M bps - Error free (the effect of error control is neglected) - Bit stuffing (to guarantee minimum frame size) - Half-duplex medium - Effect of traffic not destined (originated) to (in) the relay in bandwidth consumption - Propagation time of ACKs
PSDN sub-network lower layers	- Variable bit rate (set to 64 K bps) - Error free (the effect of error control is neglected) - Full-duplex medium - Propagation time of ACKs
End system (or next relay system)	- Generated data block size (e.g., size of the N-Data primitives in the case of network service users) - Data block inter-arrival time - Number of data blocks generated - QoS parameters: - throughput - transit delay

Tab. 3.1 Summary of characteristics included in the model.

Based on the aspects summarized on table 3.1 a simulation model was built up using Activity Cycle Diagrams (ACDs) [Clementson 82]. In ACDs two kinds of elements exist: *entities* and *activity cycles.* Entities represent real elements such as packets, buffers or available window. The dynamic behavior of entities is represented by activity cycles. Activity cycles are closed sequences (cycles) of alternating activities and waiting states through which entities circulate. ACDs are supported by the ECSL (Extended Control and Simulation Language) simulation language, to which the description can be easily rewritten into [Clementson 82].

4. Simulation results

Without congestion control mechanisms, communication systems exhibit the behavior illustrated in figure 4.1 [Jain 87].

In the *linear zone* transit delays are low and throughput increases with load. If the load continues to increase the communication system approaches its limit capacity, queues start to build up and transit delay increases rapidly -- the system is said to be in the *congestion zone*. Due to the overflow of internal queues and retransmissions when the communication system is overloaded, throughput decreases exponentially and transit delay increases also exponentially; this region is know as the *congestion collapse zone*.

Ideally, congestion control functions should keep the communication system within the linear zone and they must recover if, by any reason, the system enters the congestion or, further more, the collapse region.

In order to evaluate the congestion collapse avoidance and the fairness guarantee behavior of RC5x.1yyy relays, two different situations have been simulated (Situation A and Situation B).

In both situations six active channels were considered in the simulation setup (this number was imposed by the simulation hardware environment, but it was found to be enough to the analysis made). The load is considered to be balanced in the two traffic flow directions (LAN -> RELAY -> PSDN and PSDN -> RELAY -> LAN). Data for channels is generated in blocks of variable size with a Weibull probability distribution (corresponding to the generation of *N-DATA.request* primitives by the Network service users) and are then splited (if necessary) into PSDN or LAN packets. The data inter-arrival time is ruled by a Poisson distribution. As said before the Relay performs segmentation / reassembling functions to fit different packet sizes in the PSDN and LAN sides.

In order to obtain load/throughput and load/delay charts the total load is continuously increased from a small value (corresponding to a large inter-arrival time) to a high value (small inter-arrival time). There are two kinds of channel users: *“well-behaved”* users that don’t increase their demand above their QoS throughput; *“ill-behaved”* ones that don’t respect this limit. The goal is to see if the ill-behaved channel users get more than their fair share of bandwidth and if they harm the well-behaved channels with their misconduct.

In situation A all the channel users are ill-behaved. The QoS throughput and delay are asymmetrically distributed among channels (see table 1). To evaluate the bandwidth redistribution capabilities of the Relay, channel 1 stops transmitting for some time and then restarts his activity.

In situation B, interactive, well-behaved traffic is mixed with ill-behaved file-transfer-like-traffic. Two channels (2 and 5) carry terminal generated traffic with small data blocks and medium inter-arrival times. The demand of these channels is constant and small (2% of the available bandwidth each). Ill-behaved channels (1, 3, 4 and 6) use 96% of available bandwidth (24 % each) but they don’t respect this limit and their demand is continuously increased.

In both situations the QoS Max. Delay parameter is calculated by expression (3.6) using LAN window and packet sizes. This is because the most relevant relay delays are those introduced in the LAN -> PSDN traffic.

Tables 4.1 and 4.2 show a summary of the most relevant parameters of the simulations made in Situation A and Situation B respectively.

Parameter	Chan. 1	Chan. 2	Chan. 3	Chan. 4	Chan. 5	Chan. 6
QoS Throughput [bps]	26800	19200	9600	4800	2400	1200
QoS Max. Delay [ms]	305	425	850	1700	3410	6820
PSDN Window [packets]	4	4	4	4	4	4
PSDN Max. Packet Size [bytes]	128	128	128	128	128	128
PSDN N-DATA med. size [bytes]	512	512	512	512	512	512
PSDN N-DATA inter-arrival [ms]	8000-200	8000-250	8000-300	8000-400	8000-500	8000-600
LAN Window [packets]	2	2	2	2	2	2
LAN Max. Packet Size [bytes]	1024	1024	1024	1024	1024	1024
LAN N-DATA med. size [bytes]	512	512	512	512	512	512
LAN N-DATA inter-arrival [ms]	8000-200	8000-250	8000-300	8000-400	8000-500	8000-600
Cannel user behavior	ill	ill	ill	ill	ill	ill

Tab. 4.1 Summary of the simulation parameters for Situation A

Figures 4.2 and 4.3 show the simulation results relative to RELAY -> PSDN throughput and RELAY queues delay respectively. Throughputs are plotted normalized to the communication system bottleneck (PSDN link speed of 64000 bps). PSDN -> LAN traffic is not analyzed in detail because it doesn’t congest the relay. For this traffic direction, queues never build up in the relay, and the delay introduced by it is neglectable, compared to the total delay.

As can be observed from the average plots the congestion collapse zone is never reached, even under severe load situations (2 times the bottleneck capacity). For load under around 80 % of the bottleneck capacity (PSDN link speed), the relay operates in the linear zone with low transit delay and throughput proportional to load.

As load increases above 80%, the relay moves from the linear zone into the congestion zone. Queues built up in the relay and transit delays increases exponentially with load. When the load reaches the bottleneck capacity, the relay queues saturate. Flow control avoids queue overflow and thence avoids the entrance of the relay in the congestion collapse zone.

Looking at the throughput/load plots it can be seen that, when overload arises, bandwidth is shared according to QoS throughput parameters.

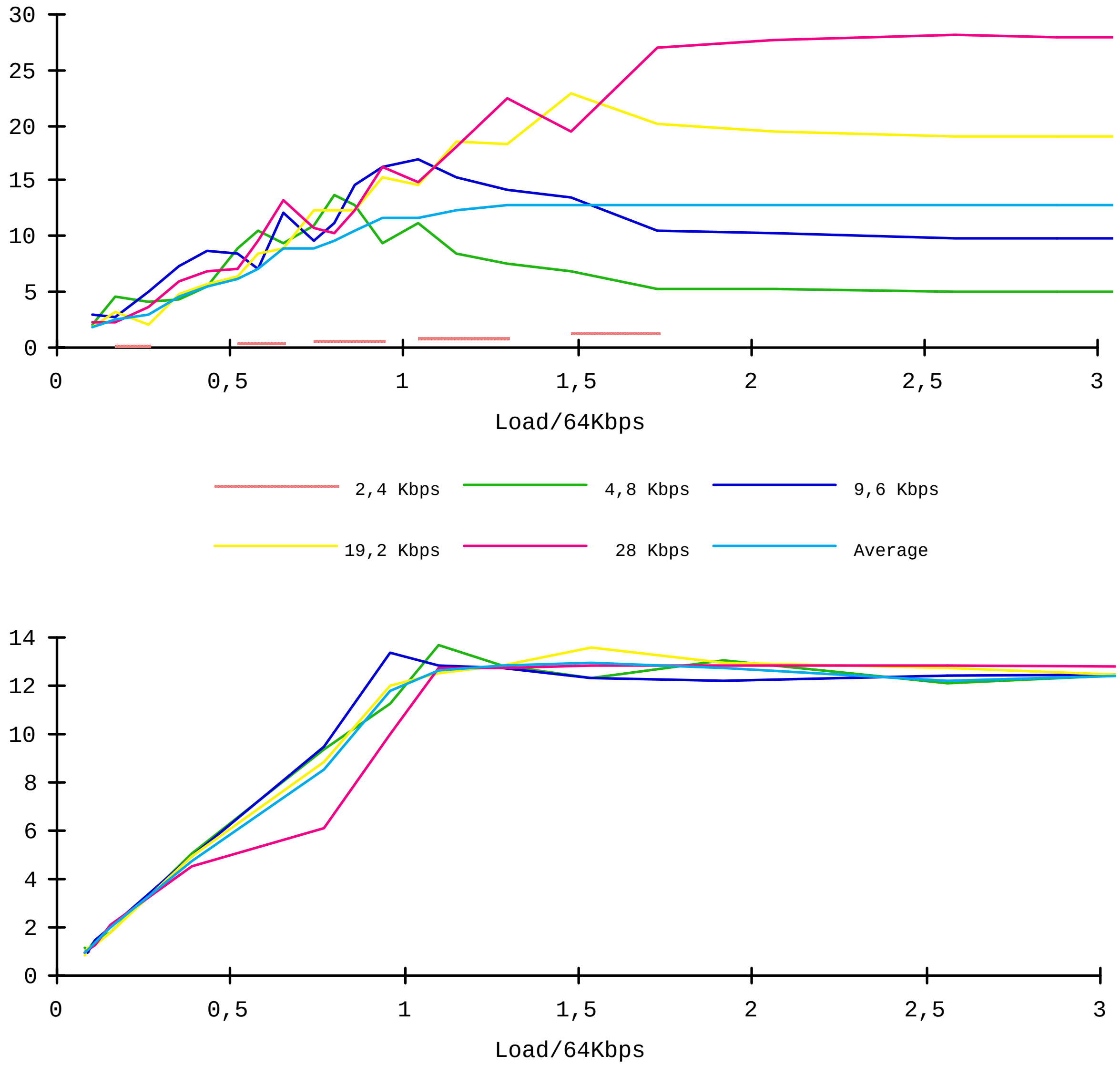

Fig. 4.2 Simulation results relative to RELAY -> PSDN throughput

In situation A throughput increases with load and, when the bottleneck is reached, channel throughput is regulated by QoS throughput values. Also in situation A, when channel 1 becomes idle it's bandwidth is splited between the active channels, proportionally to QoS throughput values. When it becomes busy again its bandwidth is recovered from the other channels.

Situation B shows that well-behaved users are protected from ill-behaved ones. Ill-behaved file-transfer-like channels (1, 3, 4, 6) have a share of 24 % of the bandwidth each, but their demand is continuously increasing is spite of this limit; well-behaved channels have a share of 2% and a constant demand of 1% each. When the system reaches the congestion zone the channel service algorithm ensures protection to the well-behaved channels.

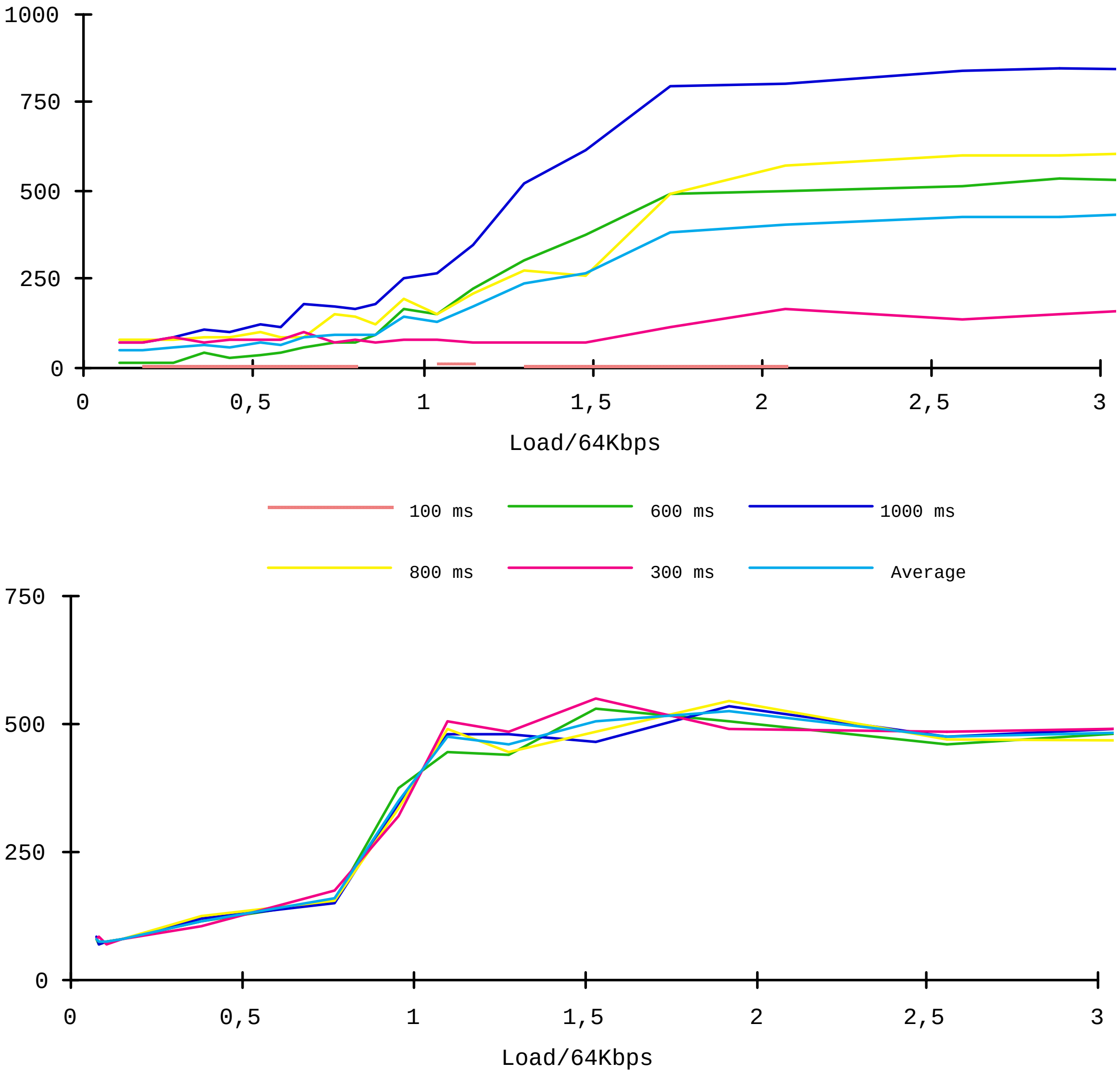

Fig. 4.3 Simulation results relative to RELAY queues

Looking now at the transit delay QoS parameters Figure 4.3 shows that, in both situations A and B, QoS Max. Delay is always respected. Another interesting observation is that for demands below 80% of the QoS throughput class, delays are very small compared to QoS delay parameter. When the demand exceeds QoS throughput (ill-behavior) channel users are penalized with delays near de QoS Max. Delay parameter.

5. Conclusion

The simulation study of the X.25 protocol relay for the interconnection of CSMA/CD LANs and X.25 PSDNs highlighted some important aspects of the relay congestion control behavior.

A channel service algorithm was proposed to avoid unfairness in respect to bandwidth and to internal resources utilization. Fairness is measured against the QoS throughput and delay parameters.

The simulations made revealed that the proposed channel service algorithm in conjunction with X.25 built-in flow control mechanisms, can avoid the congestion collapse in X.25 protocol relays, even under severe load situations.

Simulation also showed that fairness can be achieved. User throughput and transit delay QoS parameters are respected even when the demand increases above the system bottleneck capacity (PSDN link). The proposed algorithm creates “fire-walls” between well and ill behaved channel users protecting the first ones from the misbehavior of the seconds.

The queuing delay analysis revealed that the optimal operating point for channels is around 80% of channel's QoS throughput. Below this point, delays are low compared to the QoS delay parameter. When the load exceeds that point queues build-up exponentially and the delay reaches the QoS delay. This effect acts like a punishment to ill-behaved users.

These simulation results are being substantiated by experiments with a prototype implementation of an RC51.1111 relay, in order to effectively evaluate the feasibility and usefulness of the proposed algorithms.

Acknowledgment: This paper has been partially funded by JNICT (Junta Nacional de Investigação Científica e Tecnológica), under contract PCMT/C/TIT/454/90.

References

[Boyer 92] - Boyer, P., Guillemin, F., Servel, M. and Coudreuse, J. "Spacing Cells Protects and Enhances Utilization of ATM Network Links", *IEEE Network*, **6**, (5), pp. 38-49, September 1992.

[Clementson 82]. - Clementson, A., *ECSL-Users Manual*, GLE.COM, 1982.

[Demers 89] - Demers, A., "Analysis and Simulation of a Fair Queuing Algorithm", *Computer Communication Review,* **19,** (4), pp. 1-12, September 1989.

[Eckberg 92] - Eckberg, A., "B-ISDN/ATM Traffic and Congestion Control", *IEEE Network*, **6**, (5), pp. 28-37, September 1992.

[EWOS 90] - European Workshop for Open Systems, "Lower Layer Relays", *EWOS Technical Guide 006*, August 1990.

[Gerla 88] - Gerla, M. and Kleinrock, M., "Congestion Control in Interconnected LANs", *IEEE Network*, **2**, (1), pp. 72-76, January 1988.

[Holleczec 92] - Holleczek, P. and Baumgarten, T., "Throughput measurements in a 2 Mbps X.25 network", *Computer Networks and ISDN Systems*, **25**, (4-5), pp. 351-356, November 1992.

[Husemann 92] - Husemann, D., "ISO CONS in LANs - making it all work: A European contribution to 4.4 BSD Unix", *Computer Networks and ISDN Systems*, **25**, (4-5), pp. 411-419, November 1992.

[ISO 87] - International Organization for Standardization, "Use of the X.25 packet level protocol in local area networks - Part 2: Use with LLC Type 2 procedures". *ISO IS 8881-2*, 1987.

[ISO 89] - International Organization for Standardization, "Operation of an X.25 Interworking Unit", *ISO/IEC TR 10029*, March 1989.

[ISO 90] - International Organization for Standardization, "Information Technology - Data Communications - X.25 Packet Level Protocol for Data Terminal Equipment", *ISO/IEC IS 8208*, March 1990.

[ISO 91] - International Organization for Standardization, "Information Technology - Framework and taxonomy of International Standardized Profiles - Part 2: Taxonomy of profiles", *ISO/IEC TR 10000-2*, 1991.

[Jacobson 88] - Jacobson, V., "Congestion Avoidance and Control", *Computer Communication Review,* **18,** (4), pp. 314-329, September 1988.

[Jain 87] - Jain, R., Ramakrishnan, K. and Chi, D., "Congestion Avoidance in Computer Networks With a Connectionless Network Layer", *DEC-TR-506*, Digital Equipment Corporation, August 1987.

[Jain 90] - Jain, R., "Congestion Control in Computer Networks: Issues and Trends", *IEEE Network*, **4**, (3), pp. 24-30, May 1990.

[Jain 92] - Jain, R., "Myths about Congestion Management in High-speed Networks", *Internetworking Research and Experience,* **3**, (3), pp. 101-113, September 1992.

[Keshav 91] - Keshav, S., "A Control-Theoretic Approach to Flow Control", *Computer Communication Review,* **21,** (4), pp. 3-17, September 1991.

[Lea 92] - Lea, C., "What Should be the Goal for ATM", *IEEE Network*, **6**, (5), pp. 60-66, September 1992.

[Mankin 90] - Mankin, A., "Random Drop Congestion Control", *Computer Communication Review,* **20,** (4), pp. 1-9, September 1990.

[Mishra 92] - Mishra, P. and Kanakia, H., "A Hop by Hop Rate-based Congestion Control Scheme", *Computer Communication Review,* **22,** (4), pp. 112-123, Oct. 1992.

[Nagle 87] - Nagle, J., "On Packet Switches with Infinite Storage", *IEEE Transactions on Communications*, **35**, pp. 435-438, 1987.

[Stallings 93] - Stallings, W., "Components of OSI: International Standardized Profiles", *Connexions*, **7**, (1), pp. 2-15, January 1993.

[Trajkovic 92] - Trajkovic, L. and Golestani S., "Congestion Control for Multimedia Services", *IEEE Network*, **6**, (5), pp. 20-27, September 1992.

[Turner 92] - Turner, J., "Managing Bandwidth in ATM Networks with Bursty Traffic", *IEEE Network*, **6**, (5), pp. 50-59, September 1992.