INTERNAL MODEL HOP-BY-HOP CONGESTION CONTROL FOR HIGH-SPEED NETWORKS

Antonio Pietrabissa
Università di Roma "La Sapienza" – Dipartimento di Informatica e Sistemistica (DIS)
via Buonarroti 18, 00184 Roma, Italy; Tel: +39 06 44585975; Fax: +39 06 32501463; e-mail: pietrabissa@dis.uniroma1.it

Keywords: congestion control, internal model control, Smith's principle.

Abstract
This paper presents a hop-by-hop congestion control for high-speed networks. The control policy relies on the data exchange between adjacent nodes of the network (nearest-neighbour interaction). The novelty of this paper consists in the extensive use of Internal Model Control (IMC) to set the rates of the traffic flows. As a result, the proposed congestion control provides upper-bounds of the queue lengths in all the network buffers (*overflow avoidance*), avoids wasting the assigned capacity (*full link utilisation*) and guarantees the congestion recovery. Numerical simulations prove the effectiveness of the scheme.

1 Introduction
Congestion control schemes regulate the traffic rate of the flows entering the network nodes. In this paper, two kinds of traffic classes are considered: i) High Priority (HP) traffic, which is used to transport real-time traffic, has bandwidth guarantees and is not subject to congestions; ii) Low Priority (LP) traffic, which is used to transport data traffic (e.g. IP best-effort traffic), has no bandwidth guarantees and is subject to congestions. In particular, the capacity that is not used by the HP traffic has to be dynamically shared among the LP traffic flows. Thus, this paper deals with the LP flows and considers the HP flows as background traffic.
The control objective is twofold: on one hand, the queue lengths in the network buffers must be kept low in order to minimize the occurrence of buffer overflows (i.e., packet losses); on the other hand, it is desirable to keep some packets in the queues, in order to take advantage of sudden availability of bandwidth. As detailed in Section 2, these objectives can be expressed as proposed in [9]:
1) *Stability*: the queue lengths in the network buffers should be upper-bounded by the buffer sizes, so that buffer overflows are avoided;
2) *Full link utilization*: when the nodes are allowed to transmit at a certain rate, the buffers should always have some packets to send.

In high-speed networks, congestion control is a crucial issue; consequently, it has been the subject of widespread researches. Most proposed congestion control schemes are based on end-to-end (ETE) feedback exchange: the traffic sources receive the feedback from the nodes of the network and adjust their transmission rate. By far, the most used congestion control protocol is the TCP ([14]), but ETE schemes has been proposed by high number of studies, based on classical control theory ([9]), optimal control ([15]), neural networks ([16]) and many other approaches.
In contrast, in hop-by-hop (HBH) congestion control also the nodes of the network regulate the transmission rates, on the basis of feedback from the neighbouring nodes. The main advantages of HBH schemes are the following ([2]):
i) *Reactiveness*: the ETE feedback delay can be much larger than the feedback delay of a single hop, thus HBH algorithms can react faster to traffic changes;
ii) *Resource utilisation*: in ETE congestion controls, the packets are accumulated in a single node of the network (bottle-neck node), while HBH ones are able to utilise also the preceding nodes along the path.
However, currently used congestion control schemes are ETE: the main reason is that HBH schemes introduce complexity in the switches and lead to scalability problems, since they require the nodes to keep per-flow information.
Recently, there is a growing interest in hop-by-hop schemes ([1]-[4]), since next generation switches and routers will have more processing capabilities in order to enhance the Quality-of-Service (QoS) perceived by the users. In [1], the congestion control relies on the prediction of the HP traffic; in [2], the control equation is based on a model of the system, whose inaccuracies are corrected via the feedback; [3] regulates the transmission rates on the basis of the queue length in the down-stream node; [4] models the network as a linear parameter varying (LPV) system.
The algorithms [1]-[4] have some drawbacks: [1] assumes that the traffic sources are persistent (i.e., they have infinite backlogged traffic) and that the destination are capable of absorbing the whole received traffic; [1] and [2] assume that the transmission delay between each couple of adjacent nodes is the same; [4] does not require per-flow buffering, but the stability of the scheme is not exponential, so that the robustness is affected; none of them [1]-[4] guarantees the avoidance of buffer overflows and the *full link utilization*.
The proposed scheme overcomes the above-mentioned problems, and is based on the following assumptions:
1) *Per*-flow buffering policy: the switches store each flow in a separate buffer, as shown by Fig. 1.
2) The buffers are served in a round-robin fashion (which, in high-speed networks, is a good approximation of fair queuing [7]), but the nodes are capable of ensuring that the depletion rate of the buffers is bounded by a certain rate ([2]).
3) The link delays are known in advance; this assumption is straightforward if the *control traffic* has strict priority over the *data traffic* ([9]).
Assumption 3) can be neglected if the feedback delays are

estimated on-line. However, for the sake of simplicity, this assumption is considered in this paper.
Since network traffic is hardly predictable ([13]), no assumption is made on both HP and LP traffic.
The proposed scheme is based on Internal Model Control (IMC) ([10]) and utilizes a Smith Predictor (SP) controller; since the SP can be put in an IMC-form, the scheme is named *Internal Model Congestion Control* (IMCC). Model-based control utilises the model of the system to compute the control law, while feedback messages are utilised to evaluate and correct the model inaccuracies. The effectiveness of using model-based control for congestion control has been examined in several ETE schemes ([5], [6]) and also in the HBH scheme presented in [2].

The paper is organized as follows: in Section 2, the network model is developed; in Section 3 and 4 the node model and controller are developed; Section 5 presents the numerical simulation; in Section 5, the conclusions are drawn.

2 Network Model

The traffic sources transmit packets towards a certain destination through the network elements: links, switches and buffers (see Figure 1). The data transmission between a source and a destination is defined as a flow. The switches are capable of directing the different flows to the proper output buffers. Let the node be the switch altogether with its output buffers. The links connect the different nodes; each link is characterized by its capacity, which is shared by the different flows coming from the same node and directed to the same node.
The congestion control scheme of the present paper is based on the exchange of control messages between adjacent nodes of the network (nearest-neighbor interaction). For the sake of simplicity, we will refer to the buffer fed by the j^{th} flow as the j^{th} buffer of the node/source/destination.

Fig. 2 shows the path followed by data packets among adjacent nodes (*data path*), as well as the exchange of control packets (*control path*). The figure refers to 3 consecutive nodes traversed by the j^{th} flow, and the control and data paths are relevant to the j^{th} flow. The forward and the backward loops are characterised by two (generally different) feedback delays: $T^{i,j}$. and $T^{i-1,j}$.
A brief description of the data and control messages of Fig. 2 follows:
For each flow j, node i transmits the *allowed rate*, $r_{ALL}^{i,j}(t)$, to node (i-1) and the *requested rate*, $r_{REQ}^{i,j}(t)$, to node (i+1):

- The allowed rate is computed in order to avoid the overflows of the j^{th} buffer, and represents the desired input rate;
- The requested rate is computed in order to avoid the underruns of the j^{th} buffer, and represents the desired output rate.

The dynamic of the j^{th} buffer of the i^{th} node is driven by the following data rates:

- The buffer transmits towards the j^{th} buffer of the $(i+1)^{th}$ node with the output data rate $r_{OUT}^{i,j}(t)$, which is regulated by the bandwidth availability of the i^{th} link and by the allowed rate received from the $(i+1)^{th}$ node, $r_{ALL}^{i+1,j}(t)$, which, in turn, depends on the requested rate, $r_{REQ}^{i,j}(t)$;
- The buffer receives packets from the j^{th} buffer of the $(i-1)^{th}$ node with the input data rate $r_{IN}^{i,j}(t)$, which is regulated by the bandwidth availability of the $(i-1)^{th}$ link and on the allowed rate communicated by the i^{th} node, $r_{ALL}^{i,j}(t)$.

Flow j in node i can be subject to *forward* and *backward congestions*:
Forward congestions lead to a reduction of the buffer output rate, and can be caused by two conditions:
i) by bandwidth constraint of the i^{th} link, due to the HP traffic and by the concurrent LP traffic;
ii) by the allowed rate of the controller of node i+1.
Similarly, *backward congestions* lead to a reduction of the buffer input rate, and can be caused by two conditions:
i) by bandwidth constraint of the $(i-1)^{th}$ link (i.e., when node i-1 cannot transmit at the allowed rate);
ii) by the requested rate of the controller of node i-1.

Hereafter, all the variables refer to the j^{th} flow; thus, without ambiguity, the index j will be omitted.
As previously mentioned, the control objective is to maintain the queue length $q^i(t)$ in the interval [0, B^i], where B^i is the buffer size of the i^{th} node. It is known that the required buffer size is tightly related to the product transmission rate-feedback delay. The objective of the following Sections is the determination of such a relation. In other words, given the buffer size B^i and the feedback delays T^{i-1} and T^i, the maximum transmission rate which allows to meet the control objectives will be established. This rate will be referred to as *target rate* r_{MAX}^i.
In view of these considerations, the control objectives can be expressed as follows:
1) The control actions $r_{ALL}^i(t)$ and $r_{REQ}^i(t)$ must guarantee that $q^i(t)$ is bounded between 0 (*full link utilization*) and B^i (*stability*);
2) In congestion-less case, the flow transmission rate must tend to the *target rate* (*congestion recovery*)[1].

3 Node Model

Referring to the j^{th} flow, the model of the i^{th} node consists of 3 elements:
1) The *buffer j*, in which the packets of the j^{th} flow wait for transmission towards the $(i+1)^{th}$ node, is modelled by an integrator. Let $q^i(t)$ denote the queue length in this buffer; its variation is given by the input rate minus the output rate:

$$\dot{q}^i(t) = r_{IN}^i(t) - r_{OUT}^i(t) \qquad (3.1)$$

Note that (3.1) is a linearized buffer model. In the reality, $q^i(t)$ can not be lower than zero ($q^i(t) = 0$ means an empty buffer) and greater than the buffer size B^i. As shown in the following Section, the proposed scheme manages to

[1] The beginning of the flow transmission is regarded as a *backward congestion*.

keep $q^i(t)$ in the linear dynamic, since it meets the *stability* and *full link utilization* objectives.

2) The *controller* of buffer *j* of node *i*, $G^{i,j}(t)$, is in charge of computing the control actions on the basis of the available measures (*queue length* and *input rate*). The controller $G^{i,j}(t)$ will be defined in Section 4.

3) Finally, the *switch* of node *i*, has the task of limiting the *allowed rate* by the *desired rate* and by the *requested rate*, as described by the following equation:

$$r^i_{ALL}(t) = \min\{r^i_{DES}(t), r^{i-1}_{REQ}(t - T^{i-1}/2)\} \quad (3.2)$$

To model the *backward congestion* due to the requested rate of node *i*–1, let define the *request disturbance*:

$$d^i_{REQ}(t) = \begin{cases} 0 \text{ if } r^i_{DES}(t) \le r^{i-1}_{REQ}(t - T^{i-1}/2) \\ r^i_{DES}(t) - r^{i-1}_{REQ}(t - T^{i-1}/2) \text{ else} \end{cases} \quad (3.3)$$

Thus, equation (3.2) can be rewritten as follows:

$$r^i_{ALL}(t) = r^i_{DES}(t) - d^i_{REQ}(t) \quad (3.4)$$

and $d^i_{REQ}(t)$ meets the following inequality:

$$0 \le d^i_{REQ}(t) \le r^i_{DES}(t) \quad (3.5)$$

The *forward* and *backward loops* are modelled as follows:

4) The *forward loop* is modelled by a delay equal to T^i followed by an additive disturbance, $d^i_{FW}(t)$, named *forward disturbance*, which models the *forward congestions*. Considering that, when no *forward congestion* is occurring, $r^i_{OUT}(t) = r^i_{REQ}(t - T^i)$, while, in the other case, $r^i_{OUT}(t) \le r^i_{REQ}(t - T^i)$, the disturbance can be defined as follows:

$$d^i_{FW}(t) = \begin{cases} 0 \text{ if } r^i_{OUT}(t) \le r^i_{REQ}(t - T^i) \\ r^i_{REQ}(t - T^i) - r^i_{OUT}(t) \text{ else} \end{cases} \quad (3.6)$$

Thus, $d^i_{FW}(t)$ meets the following inequality:

$$0 \le d^i_{FW}(t) \le r^i_{REQ}(t - T^i) \quad (3.7)$$

5) Similarly, the *backward loop* is modelled by a delay equal to T^{i-1} followed by an additive disturbance, $d^i_{BW}(t)$, named *backward disturbance*. $d^i_{BW}(t)$, which models the *backward congestiosn* due to the concurrent traffic of node *i*-1, is defined as follows:

$$d^i_{BW}(t) = \begin{cases} 0 \text{ if } r^i_{IN}(t) \le r^i_{ALL}(t - T^{i-1}) \\ r^i_{ALL}(t - T^{i-1}) - r^i_{IN}(t) \text{ else} \end{cases} \quad (3.8)$$

Thus, $d^i_{BW}(t)$ meets the following inequality:

$$0 \le d^i_{BW}(t) \le r^i_{ALL}(t - T^{i-1}) \quad (3.9)$$

Fig. 3 shows the proposed linear node model. By taking into account equations (3.6), (3.8) and (3.4), the following equations hold:

$$\begin{aligned} r^i_{IN}(t) &= r^i_{ALL}(t - T^{i-1}) - d^i_{BW}(t) = \\ &= r^i_{DES}(t - T^{i-1}) - d^i_{REQ}(t - T^{i-1}) - d^i_{BW}(t) \end{aligned} \quad (3.10)$$

$$r^i_{OUT}(t) = r^i_{REQ}(t - T^i) - d^i_{FW}(t) \quad (3.11)$$

Remark 1:

Equations (3.5), (3.7) and (3.9) express the physical constraint stating that the packet transmission rates cannot be negative, and imply that also equations (3.10) and (3.11) are non-negative. □

4 Node Controller

Let the *adjustment rate* $r^i_{ADJ}(t)$ be defined as follows:

$$r^i_{ADJ}(t) = K^i \cdot \left(q^i(t) - \int_{t-T^i}^{t} r^i_{REQ}(t) \cdot dt - \int_{t-T^{i-1}}^{t} r^i_{ADJ}(t) \cdot dt \right) \quad (4.1)$$

where K^i is a constant to be properly set.

The control actions of the controller $G^{i,j}(t)$ are the following:

$$r^i_{DES}(t) = r^i_{MAX} - r^i_{ADJ}(t) \quad (4.2)$$

$$r^i_{REQ}(t) = r^i_{IN}(t) + r^i_{ADJ}(t - T^{i-1}) \quad (4.3)$$

where r^i_{MAX} is a constant representing the *target rate*.

The initial conditions at the start of the transmission, for $t = 0$, are $q^i(0) = 0$ and $r^i_{REQ}(t) = r^i_{ADJ}(t) = 0$ for $t \le 0$.

The proposed node controller scheme, named Internal Model Congestion Control (IMCC), is shown in Fig. 4 a).

The idea beneath the controller (4.1), (4.2) and (4.3) is to let the transmission rate tend to the *target rate*; the congestion recovery is achieved by using the two measured variables disjointedly:

- The *input rate* is used to regulate the *requested rate* with the aim of recovering from *backward congestions*.
  In fact, considering equations (3.10), (4.2) and (4.3), it follows that the *requested rate*, computed on the basis of $r^i_{IN}(t)$, does not depend on $q^i(t)$:

$$r^i_{REQ}(t) = r^i_{MAX} - d^i_{REQ}(t - T^{i-1}) - d^i_{BW}(t) \quad (4.4)$$

- On the contrary, the *desired rate* is independent of the *input rate*: the queue length is used to compute the *adjustment rate*, which, in turn, is used to regulate the *desired rate* (see equation (4.2)) with the aim of recovering from *forward congestions*.

As a consequence, the *forward* and *backward* disturbance are uncoupled: as a matter of fact, by inspection of Fig. 4 a), it can be noted that the following theorem holds:

Theorem 1:

If $d^i_{FW}(t) \equiv 0$, the IMCC scheme (4.1), (4.2) and (4.3) is equivalent to the *backward congestion scheme* shown in Fig. 4 b), while, if $r^i_{MAX} \equiv d^i_{REQ}(t) \equiv d^i_{BW}(t) \equiv 0$, the IMCC scheme is equivalent to the *forward congestion scheme* shown in Fig. 4 c), where the $K^i(t)$ block is the SPC block of Fig. 4 a). □

Remark 2:

As already mentioned, the aim of the controller $G^{i,j}(t)$ is to use the measures of the queue length $q^i(t)$ in order to deal with the unmeasured disturbance $d^i_{FW}(t)$. For this purpose, the scheme utilises the Internal Model (highlighted as IM in Fig. 4 a)) of the *rate-based request scheme* highlighted in Fig. 4 b). Given that the initial condition $q^i(0) = 0$ is known, the *forward*

disturbance $d^i_{FW}(t)$ is the only difference between the *rate-based request scheme* and the IM block, which produces an estimate of the queue length:

$$q^i_r(t) = \int_{t-T^i}^{t} r^i_{REQ}(t) \cdot dt \qquad (4.5)$$

Thus, the effect of $d^i_{FW}(t)$ can be evaluated by measuring the queue error:

$$q^i_e(t) = q^i(t) - q^i_r(t) \qquad (4.6)$$

and $r^i_{ADJ}(t)$ depends on $d^i_{FW}(t)$ only. □

With reference to the schemes of Fig. 4 b) and c), the following lemmi hold[2].

Lemma 1:
The queue length $q^i_r(t)$ of the *backward congestion scheme* (shown in Fig. 4 b)), is bounded between 0 and $r^i_{MAX} \cdot T^i$. □

Lemma 2:
Assuming that $K^i > 0$ and $0 \le d^i_{FW}(t) \le r^i_{MAX}$, the queue length $q^i_e(t)$ of the *forward congestion scheme* (shown in Fig. 4 c)), is bounded between 0 and $r^i_{MAX} \cdot (T^{i-1} + 1/K^i)$. □

Remark 3:
The controller $K^i(t)$, shown in the SPC block of Fig. 4 a), is based on the Smith's principle ([11]).
The feedback delay, known in literature as Dead Time (DT), might cause stability problems. In order to handle the DT, the proposed scheme avails of a Smith Predictor (SP) controller. The SP has the capability of improving the control of loops with DT, thanks also to the fact that the feedback delay of the control traffic is constant and, therefore, that the DT is exactly known in advance.
The plant of the forward congestion scheme of Fig. 4 c) is $P^i(s) = (1/s)e^{-s \cdot T^{i-1}}$, thus the SP controller is derived as follows ([12]):

$$K^i(s) = \frac{r^i_{ADJ}(s)}{q^i_e(s)} = \frac{K^i}{1 + K^i \cdot \left(\frac{1}{s} - \frac{1}{s} \cdot e^{-s \cdot T^{i-1}} \right)} \qquad (4.7)$$

where K^i is constant and must be properly set.
As a result of the application of the SP controller, the delay disappears from the denominator of the transfer function of the *forward disturbance scheme*:

$$\frac{q^i_e(s)}{d^i_{FW}(s)} = \frac{s + K^i \cdot (1 - e^{-sT^{i-1}})}{s \cdot (s + K^i)} \qquad (4.8)$$

In conclusion, the *adjustment rate* given by equation (4.1) is obtained by substituting equations (4.5) and (4.6) in the inverse Laplace transform of equation (4.7). □

[2] Due to space reasons, the paper does not include the proofs of the *Lemmi* and *Theorems* (included in the reviewed version); please contact the author if interested.

The following theorem proves that, by properly setting the SP controller gain K^i and the *target rate* r^i_{MAX}, the proposed control scheme matches the control objectives.

Theorem 2:
By setting $K^i > 0$ and r^i_{MAX} such that:

$$r^i_{MAX} \le \frac{B^i}{T^i + T^{i-1} + 1/K^i} \qquad (4.9)$$

the IMCC scheme (4.1), (4.2) and (4.3) meets the control objectives:

(i) *Stability* and *full link utilization*: $0 \le q^i(t) \le B^i$.

(ii) *Congestion recovery*: after a congestion, $r^i_{IN}(t)$ and $r^i_{OUT}(t)$ tend to r^i_{MAX}, and $q^i(t)$ tends to $r^i_{MAX} \cdot T^{i-1}$. In particular, *backward congestions* are recovered, in the worst case, in $(T^{i-1} + T^i)$ sec., while *forward congestion* recovery is exponential, with time constant $\tau^i = 1/K^i$, and has no oscillation and overshoots. □

5 Numerical Simulations

In this Section, numerical simulations of the proposed congestion control scheme are provided. A single node has been simulated, while the behaviour of the preceding and successive nodes is simulated by the *forward*, *backward* and *request* disturbances. Simulating a single node is meaningful, since no assumption is made on the disturbances.
The simulation parameters are the following: B^i = 400 packets, T^i = 200 ms, T^{i-1} = 100 ms, K^i = 0.01 ms^{-1}. From equation (4.9), the target rate has been selected as follows:

$$r^i_{MAX} = \frac{B^i}{T^i + T^{i-1} + 1/K^i} = 1 \frac{pkt}{ms}$$

The first simulation has been performed in order to test the worst-case backward congestion and the backward congestion recovery. This situation is represented by the following *backward disturbance*:

$$d^i_{BW}(t) = r^i_{MAX} \cdot [u_{-1}(t - 1s) - u_{-1}(t - 2s)]$$

where $u_{-1}(t)$ is the step function. Obviously, the same situation can be modelled using the *request disturbance*. The simulation results, presented in Fig. 5 a), show that the queue length is non-negative. Note that the actual and the estimated queue lengths are identical during the simulation.
The second simulation has been performed in order to test the worst-case forward congestion and the forward congestion recovery. This situation is represented by the following *forward disturbance*:

$$d^i_{FW}(t) = r^i_{MAX} \cdot [u_{-1}(t - 1s) - u_{-1}(t - 2s)]$$

The simulation results, presented in Fig. 5 b), show that the queue length is always less than 200 pkts.
The third simulation has been performed in order to test the protocol behaviour with a white-noise random *backward disturbance*, subject to the physical limitations of equation (3.8). The obtained actual and the estimated queue lengths are shown in Fig. 6 a): the figure shows that the queue length is always bounded between 0 (*full link utilisation*) and

$r^i_{MAX} \cdot T^i = 200$ pkts, and that $q^i(t)$ matches the estimated one.

The fourth simulation has been performed in order to test the protocol behaviour with a white-noise random *forward disturbance*, subject to the physical limitations of equation (3.6). The obtained actual and the estimated queue lengths are shown in Fig. 6 b): the figure shows that the queue length is always bounded between $r^i_{MAX} \cdot T^i = 200$ pkts and $B^i = 400$ pkts (*stability*).

The fifth simulation has been performed in order to test the protocol behaviour d with white-noise random *forward* and *backward disturbances*, subject to the limitations of equation (3.6) and (3.8), respectively. The obtained actual and the estimated queue lengths are shown in Fig. 6 c): the figure shows that the queue length is always bounded between 0 (*full link utilisation*) and $B^i = 400$ pkts (*stability*).

Remark 4:

In the actual implementation the continuous-time scheme must be sampled, assuming that the exchange of control messages between adjacent nodes is regular. In this paper, the discretization is not explicitly given, since it relies on the same considerations given in [9] and [17]. In particular, if the system is sampled with sampling time T_C, the SP controller gain K^i of eq. (4.8) has not to be higher than $1 / (2 \cdot T_C)$. □

6 Conclusions

In this paper, a hop-by-hop congestion control protocol for high-speed networks with per-flow buffering has been presented. The novelty of this paper consists in the extensive use of Internal Model Control to regulate the rates of the active traffic flows in the network.

Hop-by-hop (HBH) congestion control schemes are based on the data exchange between adjacent nodes of the flow path (nearest-neighbour interaction). In the proposed scheme, the node i of the flow path communicates with the successive node i+1 (*forward control loop*) and with the preceding node i-1 (*backward control loop*), and copes with the problem of having different feedback delays associated to the two loops, T^i and T^{i+1}, respectively.

In particular, at time t, the controller associated to the i^{th} node of a flow path communicates the *requested rate* to the $(i+1)^{th}$ node and the *allowed rate* to the $(i-1)^{th}$ node, which represent the maximum desired output rate at time $t + T^i$, and the maximum desired input rate at time $t + T^{i-1}$, respectively. The control is based on measurement of the queue length in the buffers and of the rate of the traffic feeding the buffers.

In contrast with other proposed HBH protocols, the presented scheme does not rely on unrealistic hypotheses, such as considering persistent sources or assuming that the links are characterized by identical delays. Nevertheless, the proposed scheme meets the control objectives:

i) The buffer queue lengths are upper-bounded by the buffer size (*stability*), thus the congestion control scheme guarantees the overflow avoidance;
ii) When the nodes are allowed to transmit at a certain rate, the buffers always have some packets to send, thus the allowed transmission rate is always fully utilised (*full link utilisation*);
iii) The recovery from a *backward congestion*, caused by a reduction of the *requested rate* of the $(i-1)^{th}$ node or by the competing traffic of the $(i-1)^{th}$ node, is achieved, in the worst case, in $(T^{i-1} + T^i)$ sec;
iv) The recovery from a *forward congestion*, caused by a reduction of the *allowed rate* of the $(i+1)^{th}$ node or by the competing traffic of the i^{th} node, is exponential and has no oscillations and overshoots;
v) Finally, the protocol is independent of the generally unknown statistical characteristics of the network traffic, since it does not rely on traffic models.

Numerical simulations have been provided, which validated the effectiveness of the proposed scheme.

Bibliography

[1] G. Ramamurthy, B. Sengupta, "*A Predictive Hop-by-Hop Congestion Control Policy for High Speed Networks*", in Proc. of INFOCOM '93, pp. 1033-1041, 1993.

[2] P. Mishra, H. Kanakia, S. Tripathi, "*On Hop-by-Hop Rate Based Congestion Control*", IEEE ACM Transactions on Networking, 4(2), 224-239, 1996.

[3] H. Zhang, O. W. Yang, "*The hop-by-hop* flow *controller for high-speed networks_single VC case*", in Proc. of IEEE GLOBECOM '97, Vol. 2, pp. 785-789, 1997.

[4] S. Bohacek, "*Stability of hop-by-hop congestion control*" Proc. of the 39th IEEE Conference on Decision and Control, Vol. 1, pp. 67-72, 2000.

[5] S. Keshav, "*A control-theoretic approach to* flow *control*", in Proc. of ACM SIGCOMM'91, Zurich, Switzerland, pp. 3-15, September 1991.

[6] K. Ko, P. P. Mishra, S. K. Tripathi, "*Interaction Among Virtual Circuits Using Predictive Congestion Control*", Comp. Networks and ISDN Systems, 25, pp. 681-699, 1993.

[7] A. Demers, S. Keshav, and S. Shenker, "*Analysis and simulation of a fair queueing algorthm*", Proc. of ACM SIGCOMM, pp. 3-12, 1989.

[8] F. Delli Priscoli, A. Pietrabissa, "*Resource management for ATM-based geostationary satellite networks with on-board processing*", Computer Networks, 39, pp. 43-60, 2002.

[9] S. Mascolo, "*Congestion control in high-speed communication networks using the Smith principle*", Automatica, vol. 35, n. 12, pp. 1921-1935, Dec. 1999.

[10] R. D. Braatz, "*Internal Model Control*," in The Control Handbook (S. Levine, ed.), CRC Press, pp. 215-224,1996.

[11] O. Smith, "*A Controller to Overcome Dead Time*," ISA Journal, Vol. 6, No. 2, Feb. 1959.

[12] Z. J. Palmor, "*Time-delay compensation - Smith predictor and its modifications*," in The Control Handbook (S. Levine, ed.), CRC Press, pp. 224-237, 1996.

[13] V. Paxson and S. Floyd, "*Wide-area Traffic: The Failure of Poisson Modeling*," IEEE/ACM Transactions on Networking, pp.226-244, June 1995.

[14] W. Stalling, "*High Speed Networks: TCP/IP and ATM Design Principles*", Prentice Hall, San Francisco, 1998.

[15] E. Altman, T. Basar, R. Srikant, "*Congestion control as a stochastic problem with action delays*", Automatica, Dec., 1999.

[16] S. Jagannathan, J. Talluri, "*Congestion control of ATM networks using multilayer neural network approach: multiple source/single switch scenario*" in Proc. of ACC, Vol. 5, pp. 3789-3794, 2001.

[17] F. Delli Priscoli, A. Pietrabissa, "*Control-theoretic bandwidth-on-demand protocol for satellite networks*", in Proc. of CCA, vol. 1, pp. 530-535, Sept. 2002.

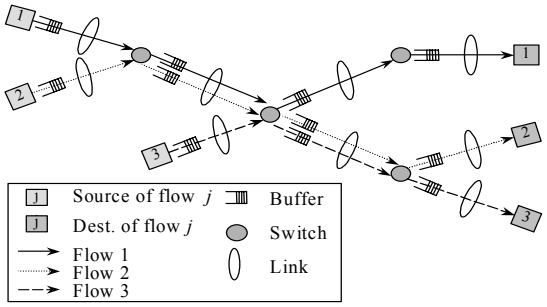

Figure 1: Network topology with per-flow buffering

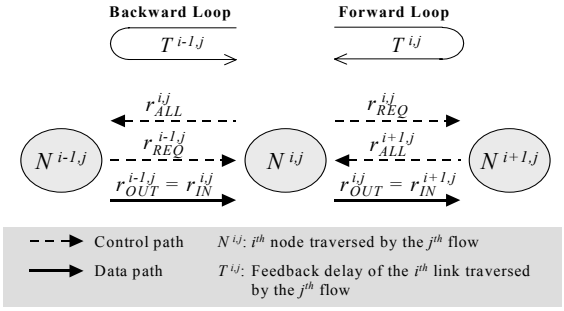

Figure 2: Node i - control and data paths of flow j

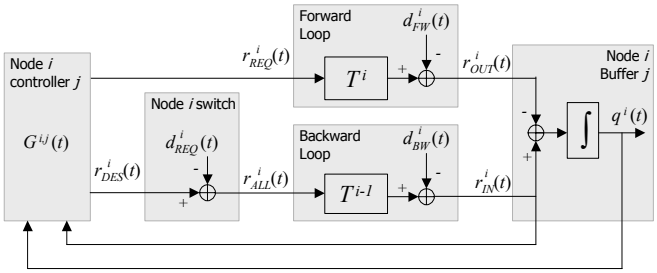

Figure 3: Node model – T^i = feedback delay of the i^{th} link

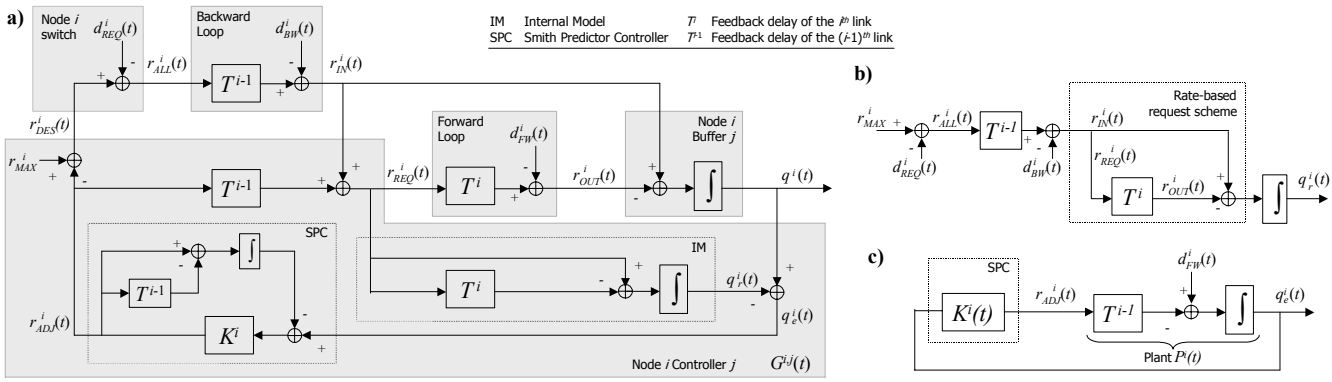

Figure 4: a) Internal Model Congestion Control (IMCC) scheme; b) Backward cong. scheme; c) Forward cong. scheme

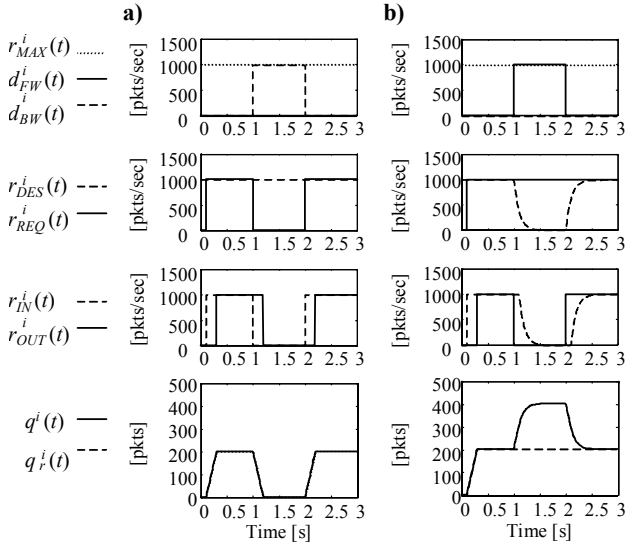

Figure 5: Disturbances, maximum, desired, requested, input and output rates, queue length and estimated queue length in worst-case backward, a), and forward, b), congestion

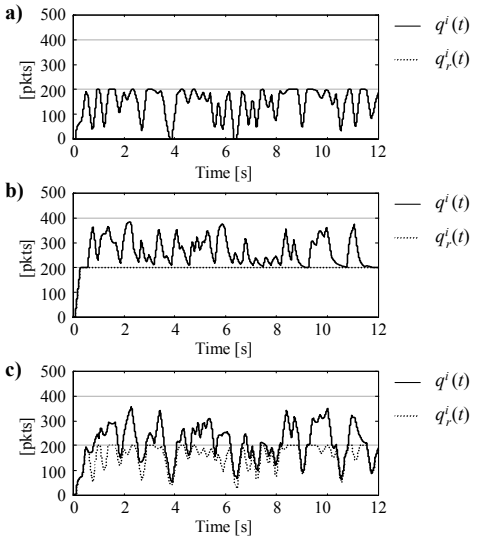

Figure 6: Queue length and estimated queue length in three situations: a) backward congestion; b) forward congestion; c) forward and backward congestion