Distributed Control System for Low Priority Controllable Traffic in Packet Switched Backbone Networks

Kimmo Pulakka and Jarmo Harju

Tampere University of Technology
Telecommunications Laboratory
P.O.BOX 553
FIN-33101 Tampere
FINLAND
Tel. +358-3-365 2111
Fax. +358-3-365 3808

Abstract. This paper studies the problems of congestion and optimal use of resources in packet switched backbone networks. The developed control system adjusts the data rates and selects the optimal paths through the backbone network for the controllable traffic category. Controllable traffic is low priority traffic, whose sources can be effectively controlled by the network. Data rates of the controllable traffic are calculated by the fuzzy controllers and statistical calculations. The adaptive neuro-fuzzy inference system (ANFIS) has been used to predict data rates of high priority traffic. This prediction tries to optimise the performance of the control system. All computationally intensive control operations are distributed to the edge switches of the backbone network. A special control protocol transfers the control information between the edge switches and the source stations of the controllable traffic, located in the access networks. Simulation tests have been used to verify the high performance of the control system.

1. Introduction

Packet-based backbone data networks, also called core networks, transfer data between different types of access networks connected to the core by edge switches. Usually backbone networks operate in wide geographical areas and the networks consist of a large number of switches.

Different user applications require different transmission services. Hence, data transmission services should be classified according to these various needs, and backbone networks should be capable of serving data transfer requests of different service categories [1]. Currently, the idea of integrated services in packet switched networks is defined most completely in ATM (Asynchronous transfer mode) networks. ATM networks include several service categories and they are capable of sharing their transmission resources between connections according to the QoS (Quality of Service) requirements of categorised connections.

The different service categories can be roughly divided into prioritised, controllable, and best effort services. The data flows of the prioritised services are served by giving certain privileges over the other traffic in the network. The prioritised services can also include guaranteed services. For such services, the network guarantees a certain quality of the transmission services for these traffic flows. The prioritised services require transmission contracts between the user application and the network.

The controllable and the best effort traffic can use the free capacity of the network, left over by the prioritised traffic flows. The difference between the controllable and best effort service categories is the control strategy of these traffic flows. The data flows of the controllable service categories are always somehow controlled by the network so that the free capacity would be used for these traffic flows as optimally as possible. Best effort traffic is not controlled in any way.

According to the aim of controllable traffic, the traffic control systems should ensure that the controllable traffic flows are controlled in real-time according to the abrupt load changes of the backbone network. The long transmission delays of large backbone networks require some kind of prediction of the load of the network [2]. Although the prediction of the traffic load in large networks is very difficult, it is valuable to try to produce some estimations. Good predictions increase the performance of the control system and decrease the need of control actions.

In this paper, we present our own data rate control and path selection system for generic controllable (GC) traffic in packet switched backbone networks. During the data transmission periods of the controllable traffic flows, the control system iteratively chooses the optimal path and calculates the optimal data rate for each GC traffic flow, depending on the load of the higher priority traffic flows. The control system is based on the use of 1) *mamdani fuzzy controllers*, 2) *ANFIS neuro-fuzzy system* and 3) *statistical calculations*. Fuzzy controllers and statistical calculations perform the control actions for the GC sources, and ANFIS generates the estimations of the high priority traffic load values. These estimation values are used for the control of the GC data flows. Several research groups have also used fuzzy controllers for network control problems. In these studies, fuzzy controllers have been used for policing the ATM network and to control the data rates of ABR connections [3], [4], [5]. Some research groups (e.g. [6]) have also found that fuzzy control systems are also suitable for solving routing problems of networks. The ANFIS neuro-fuzzy system [7] has been found to be effective in modelling the behaviour of highly non-linear dynamic systems [8].

Our backbone network model does not follow any existing standardised network architecture. However, the data rate control of the GC traffic closely follows the basic control principles of the ABR (Available Bit Rate) service category traffic in the ATM networks, and the modelled backbone network contains typical elements of connection oriented high speed packet switched networks.

The performance of the control system has been tested by simulations. The simulation results show that the control system can reach a high level of performance, even if the physical network environment and the load of the backbone vary greatly.

In Section 2 we explain the basic properties of the control protocol, focusing on the data rate adjustment and the path selection mechanisms of the control system. In Section 3 simulation results are given and analysed. Finally, some conclusions are drawn.

2. Description of the Control System

2.1. General Principles

The studied backbone network model consists of edge and core switches. The access networks are connected to the backbone network via edge switches. The data packets coming from the source access networks flow across the source edge switches and travel through the core switches to the destination edge switches. The destination edge switches send data packets to the destination access networks. In this model, we assume that the access networks connected to the backbone network are also some kind of packet based data networks, and they can transfer all incoming and outgoing traffic without the danger of congestion. In the backbone network, data are transferred using constant length packets. In the switches, every input and output link, connected to the switches, is handled by individual input and output ports. Output ports have a scheduling mechanism for the control traffic, prioritised traffic and GC traffic (see Fig. 1). The control traffic has the highest priority and the GC traffic has the lowest priority. Best effort traffic is not included here, but if it were, its priority would be the lowest. In this model, we assume that traffic coming from input ports could be switched to the output ports without any need of input buffering. The structure of the studied network environment is described in Fig. 1.

The control system can be divided to 1) *the transmission of the control information between the network components* and 2) *the control calculations*. The transmission of the control information can be further divided to transmission between the edge switches and GC source stations, and the collection of the network state information. The control calculations include the calculations for GC data flows and the calculations to define the state of the output links of the core switches.

The edge switches periodically control the GC source stations by the special GC control (GCC) packets. These packets include the data rate values of the GC data flows to reach the optimal use of the network resources. The frequency of sending the control packets depends on the estimated mean RTT (Round Trip Time) of the network. The data rate values are calculated by the control functions of the edge stations.

The collection of the state information of the network is done by the ERM (Edge Resource Management) packets. The ERM packets, originated by an edge switch, are broadcast by the core switches of the backbone network to all other edge switches. In this way, the ERM packets flow between the edge switches using all possible paths in the network and collect path specific information about the network, roughly in the same way as the BGP protocol is used in inter-domain routing. The state information of the core switches is saved to the ERM packets according to the max-min fairness principle [9]. In this way, the core switch whose current load of the output link is the highest, is allowed to represent the state of the path. The frequency of sending the ERM packets is constant and it depends on the mean RTT of the network. After receiving any ERM packet, the edge switches update the state information related to the path of the received ERM packet to a special database.

The core switches periodically calculate state information for each output link. This state information is calculated for the ERM packets. The variables used to define states of the output links are described in Table 1. The edge switches are responsible for most of the control calculations for the GC data flows. Using the state information

of the special database of paths, they 1) *estimate the load of the high priority traffic*, 2) *define the optimal path for each active GC data flow*, and 3) *adjust the data rate values of the GC data flows to the optimal values*. Calculations of the optimal paths and optimal data rates are done periodically together with the sending of the GCC packets to the GC source stations. The optimal data rate values are copied to the GCC packets so that the GC source stations can adjust their data rates. The optimal paths of active GC flows are saved into the edge switches. Between two control moments, the edge switches direct the incoming data packets from the GC source stations to the optimal paths. The estimations of the load of the high priority traffic are made every time when edge switches receive any ERM packet. These estimations are done individually for every alternative path, and they are done by the load estimation values of ERM packets or ANFIS estimators. The estimated loads of the high priority traffic are used to define optimal paths and optimal data rates.

Fig. 1. Structure of studied network environment.

Table 1. Control information of core switches

Variables	Meaning
Load	Current and predicted load of the output port normalised to the link capacity. Simple linear load predictor estimates the load value of the link after RTT (round Trip Time). Load values are calculated for all traffic types.
# of GCs	Number of active GC flows through the output port
Occupancy of GC buffer	Occupancy level of the GC buffer
Transmission speed of link	The total transmission capacity of the link (for all types of traffic)

$$\text{Total.curr. load} = \sum_{i=1}^N \frac{\text{Data flow of traffic class (i)}}{\text{Capacity of the link (Mbit / s)}} \quad (1)$$

N = Number of diff. traffic classes (3)
 {control traffic, high priority traffic, GC traffic}

$$\text{Occu.level} = \frac{\# \text{ of packets in GC buffer (packets)}}{\text{Optimal \# of packets in buffer}} \quad (2)$$

2.2. Prediction of the High Priority Traffic Loads

The estimation of the amount of the high priority (HIP) traffic along the paths is done in two alternative ways. If the occupancies of the GC output buffers of the heaviest loaded output links are exceeding the defined limit, the load prediction values of the latest ERM packets are used to estimate high priority loads in the paths in near future. On the other hand, if the occupancy is below this limit, the prediction is based on the ANFIS predictor. The load predictors of core switches predict the load values based on the latest load information, whereas ANFIS predictors use a little bit earlier load values for the prediction. For this reason, the prediction values of the core switches are used when congestion may occur in the path.

The ANFIS network is adaptive network based fuzzy inference system, which can perform the functions of the Sugeno or Tsukamoto fuzzy models. The important characteristic of the network is the capability to learn the behaviour of the dynamic systems by training data pairs. The network includes a set of linear and non-linear parameters. The values of the parameters are adjusted by a special hybrid learning rule, which is a combination of the gradient method and the least squares estimate (LSE). The ANFIS network can include several input variables but only one output variable [7].

The aim of the ANFIS predictors is to produce statistically reliable estimations of HIP loads for the future. Based on HIP load values measured in the past. The values of the input variables are defined by the HIP load information of several ERM packets. The input variables of ANFIS predictor are 1) *Maximum HIP load value during the observation time*, 2) *weighted average value of three latest HIP load values*, and 3) *the load change between two latest HIP load values*. The output of the ANFIS predictor is the estimation of the maximum possible HIP load value in future. The estimation time of future maximum HIP load values is defined by the operator of the control system. The ANFIS predictor is used in an on-line manner. Whenever ERM packet is received, the values of the input variables are calculated and the ANFIS network calculates the HIP load estimation for the current path. Calculated HIP load estimation values are saved into the edge switch to be used for the path selection and the data rate adjusting functions.

In our prediction case, the parameters of the ANFIS network are first adjusted in off-line mode using measured learning data pairs. The off-line mode is used, because adjusting the parameters is too time consuming to be made on-line by the edge switches. The data pairs for the parameter adjusting are collected from all possible paths in the network. In this way, the quality of the learning data pairs increases.

2.3. Selection of Active Path

The special databases of the edge switches contain the control information of alternative paths in the backbone network between the edge switches. The edge switches calculate the quality value for every alternative path. The paths with the best quality values will be chosen as the active paths of the GC sources. The calculation is based on two control variables. These variables are 1) *The predicted total load of the heaviest loaded output link in the path*, and 2) *the fuzzy control factor of the path*. The second variable indicates how much the data rates should be increased or decreased to reach the optimal load of the network. The values of the second variable are calculated by the fuzzy controllers. The fuzzy controller will be defined in the subsection 2.4.

The quality values of the paths are calculated according to Eq. 4. The quality of the path is high when the value of Eq. 3 is high. High values are reached when the predicted load of the link is low and the value of the fuzzy control variable is low. Low values of the fuzzy control variable indicate that the data rates of the GC traffic flows should be increased.

$$\text{path quality} = (1 - \text{PredTotLoad}) * (1 - \text{FCF}) \quad (3)$$

where :

PredTotLoad = The predicted total load of the
heaviest loaded output link

FCF = The fuzzy control factor [0,1]

2.4. Data Rate Adjustments of GC Traffic Flows

Fuzzy controller and load prediction based data rate adjustment strategies. The data rates of the GC traffic flows are adjusted using either the fuzzy controllers or the load predictors, depending on the states of the paths of the network. Fuzzy controllers are used, when the GC data rates can be controlled without danger of congestion in the output links of the paths. On the other hand, data rate changes are based on the load predictions when there is urgently need to adapt to the abruptly high load changes of the high priority traffic.

The selection possibility of the data rate adjusting strategy tries to ensure that the occupancy levels of the GC buffers in the heaviest loaded output links in the chosen paths would always be non-zero, yet lower than the overflow level of the GC buffers. The selection procedure estimates the occupancy changes of the GC buffers during the mean RTT of the control actions (see Eq. 4). If the estimated occupancy level changes cause a buffer overflow or underflow situations, the load prediction based strategy is chosen. Otherwise, the fuzzy controller based strategy will be chosen. The selection is made individually for every path of the network in every control moment of the paths.

$$\text{Occu.change} = \frac{(\text{PredTotLoad} - 1) * \text{LC (bit / s)}}{\text{Packet size (bit)}} * \text{RTT(s)} \quad (4)$$

where :

PredTotLoad = The predicted total load of the output link

LC = Link capacity

RTT = Round trip time

Fuzzy controller based data rate adjustments. The fuzzy controller based data rate adjusting (see Eq. 5) is based on two control factors. These factors are 1) *the fuzzy control factor*, and 2) *the physical data rate change factor*. The fuzzy control factor indicates how much the data rates should be relatively increased or decreased normalized to the value of the physical data rate change factor. The values of the physical data rate factor indicate the allowed data rate change during one control interval. The value definition of the physical data rate change factor depends on the capacity of the network components and the geographical size of the network. (see Eq. 7).

The fuzzy controllers, designed to calculate the fuzzy control factor values, include two input variables, 1) *the occupancy of the GC buffer*, and 2) *the predicted change of the occupancy level of the GC buffer*. The input variables of the fuzzy controllers are defined in Eq. 6. The second input variable indicates how the occupancy level of the GC buffer will be changed depending on the predicted total load, during some precisely defined time period. The predicted total load is the sum of the predicted load of the control, high priority and GC data flows. The precisely defined time is usually related to the RTT of the control loop.

To ensure the simplicity of the fuzzy controllers, the fuzzy variables and the rule-bases of the fuzzy controllers are defined by the operator of the network. The first input variable is defined by seven linguistic values and second input variable by nine linguistic values. The output variable is defined by seven linguistic values and the numerical range of the variable is [0,1]. The mamdani fuzzy inference system has been used for every fuzzy controller and the centroid area method has been used for defuzzification of the fuzzy output set.

$$DR(n+1) = DR(n) + FCF(n+1) * PhyDRCF(n+1) \quad (5)$$

where :

DR = Data rate

FCF = Fuzzy control factor (scaled to range [-1,1])

PhyDRCF = Physical data rate change factor

$$Input1 = \frac{\text{Occupancy of GC buffer (packet)}}{\text{Optimal GC occupancy (packet)}} \quad (6)$$

$$Input2 = \frac{(\text{Pred. total load} - \text{Link capacity}) * \text{Obs. time}}{\text{Optimal GC occupancy (packets)}}$$

Optimal GC occupancy :

= Optimal # of GC data packets in the GC buffer

Obs.time = Observation time, which is $\leq$ RTT

$$\text{PhyDRCF} = \frac{\text{Occupancy limit}}{\text{RTT}} \quad (7)$$

Occupancy limit:

Maximum allowed variance of the occupancy of the GC buffers during defined time.

RTT:

Estimated avg. round trip time of the control information

Load prediction based data rate adjustments. This data rate adjusting strategy is purely based on the predicted load values of different types of data flows in near future. The load values of the high priority traffic are predicted by the linear or ANFIS predictors, described in subsection 2.2. The load changes of the GC traffic flows are estimated based on the previous control actions. It is important to notice, that the predictions of the GC load changes cannot be exact values, because edge switches do not know the GC control actions of other edge switches. Because the frequency of the sending of the ERM packets is constant, the up-to-date load value of the control traffic can be used as an estimation of the control traffic load in future. Using these three predicted load values, the equally shared free capacity for the GC data flows are calculated as described in Eq 8.

$$\text{EQS capacity} = \frac{\text{Link capacity} - (\text{Pred.CL} + \text{Pred.HIPL} + \text{Curr.GCL} + \text{Pred. GCLC})}{\# \text{ of the active GC traffic flows}} \quad (8)$$

Where :

EQS capacity = Equally shared free capacity

Pred.CL = Predicted load of control traffic

Pred.HIPL = Predicted load of high priority traffic

Curr.GCL = Current load of GC traffic

Pred.GCLC = Predicted load change of GC traffic

3. Simulations

3.1. Description of the Simulated Network Environment

The simulated network environment consisted of one backbone network and five access networks. The capacity of the transmission links between the switches of the backbone network was 35 Mbit/s, and the transmission capacity of the access networks was 100 Mbit/s. The high transmission capacities of the access networks ensure that no congestion can occur in the access networks in any case. The backbone network was loaded by 11 continuously flowing high priority traffic flows and seven GC traffic flows with randomly alternating idle and active periods. Poisson and exponential distributions were used to define the lengths of the active and idle periods of the GC traffic flows. The feasible mean lengths of the active periods were 5,10 and

20 seconds, and the mean lengths of the idle periods were 4,8, and 12 seconds. Transmission rates of seven high priority traffic flows were generated by a first-order autoregressive process described in [10]. The other four high priority traffic flows adjusted their data rates according to the measured sample data rates of H.261 video transmission. The data rates of the H.261 sources were scaled to the same data-rate range as the data-rates generated by the first-order autoregressive process. The constant data packet size in the backbone network was 500 bytes. The control system tried to share the free capacity equally between the GC traffic flows so that the network would have been loaded optimally. Another control target was to prevent congestion in the network. However, small data packet loss ratios were allowed to reach the optimal throughput. The control frequencies of the GC source stations varied according to the estimated RTT of the network. The constant interval for sending the ERM packets and measuring the loads of the output ports was 0.01 s. The structure of the simulated network environment is described in Fig. 2.

Fig. 2. Structure of the simulated network environment.

The transmission delays of the links and the buffer sizes of the GC buffers of the output ports in the switches will have an effect to the performance of the control system. We used four tests to study the effects of these factors. The tests differed from each other by the distances between the switches and the buffer sizes of the GC traffic. In every test, the performance of our control system was tested by five different test cases. Slightly different traffic scenarios were used for different test cases. The simulation time of each test case was 50 seconds.

3.2. Description of Simulation Results

The performance variables. The performance of the control system is tested by four performance variables. These variables are: 1) *the average occupancy of the GC buffers in the output ports of the edge switches*, 2) *the standard deviation of the occupancy of these GC buffers*, 3) *the packet loss ratio*, and 4) *mean throughput*. The values of the mean throughput were obtained by calculating, for each GC traffic flow, the ratio of the achieved throughput to the theoretical throughput value given by the max-min fairness principle under the assumption that 100% of the available capacity of the network was used.

The combination of the high throughput values and zero packet loss indicate the high performance of the system. The values of the average occupancy levels and standard deviations of the GC buffers indicate how the buffers of the output ports are used. In our simulated network environment, we have measured the performance values for the output buffers of the edge switches. These output ports should always be used heavily, because all incoming GC traffic injected to the backbone network is transferred to core switches using these output ports. The single output port is used optimally, if the average value of the occupancy level of the GC buffer is between zero and the optimal size of the GC buffer, and the value of the standard deviation is low. In our control system, the low values of the standard deviation can not be reached. The control system adaptively changes the paths of the GC traffic flows, and this causes quite high variation of the occupancy levels in the output buffers of the switches.

Discussion about the performance results. According to the test results, the average of the throughput values of the tests was 0.97. The variation of the throughput values of different tests was low. However, the packet loss ratio of the GC packets varies between the tests. The packet loss ratio was zero, when the optimal sizes of the GC output buffers were high enough compared to the distances between the switches. On the other hand, the packet loss ratio increases when the distances between the switches are increased, but the optimal GC buffer sizes are kept unchanged. From this point of view, it can be noticed that the packet losses can be avoided, if the sizes of the GC buffers are high enough compared to the RTT of the network. According to the average and standard deviation values of the output buffers of the edge switches, the GC output buffers have had most of the time some data packets waiting for the transmission. The test results are described in Table 2.

It is important to notice that the performance results (e.g. throughput values, packet loss ratios) are dependent on several factors of the control system. These factors are 1) *the frequency of the sending of the ERM packets*, 2) *the frequency of the load measurements of the output ports of the core switches*, 3) *the frequency of the sending of the GCC packet to the GC source stations*, and 4) *accuracy of the predicted load values*. Increasing the frequencies of the sending of the control information in the network and other control related measurements increases the performance of the control system, but at the same time the free capacity for the data flows decrease. On the other hand, the accuracy of the load predictions increases by more complex predictor systems, but more complex prediction system requires also more calculations. The ANFIS network used for these tests included six membership functions for each input variable and total number of fuzzy rules was 216.

Analysis of the path selection procedure. A positive value of the path quality factor indicates that more traffic can be transmitted through the path, without danger of congestion in the network. The negative values are only reached when the load of the path is higher than its transmission capacity (see Eq. 3). From this point of view, if the network is not totally overloaded by prioritised data flows, the quality values of the chosen optimal paths for GC flows should be positive most of the time. Small positive quality values indicate that the variance of the load of the prioritised traffic is low and the control system can share the free capacity optimally. On the other hand, large positive values indicate that the network is heavily underloaded.

Table 2. Description of the test results

Test case {Dist. between switches (km), optimal buff. size, packet loss level}	Avg. Throughput	Packet loss ratio of GC traffic	Avg. occupancy of GC buffers of edge switches (normalized to optimal buff. size)	Avg. occupancy of GC buffers of edge switches (normalized to optimal buff. size)
{100,500,1500}	0.959	0.00003	0.07	0.19
{100,1000,3000}	0.969	0	0.10	0.22
{300,500,1500}	0.966	0.00001	0.10	0.23
{300, 1000, 3000}	0.972	0.00001	0.12	0.24

The distribution of the quality values of the chosen paths during one test case is described in Fig. 3. The distribution of the quality values indicates that the path selection procedure have found some underloaded path as an optimal path in most of the control cases. However, the highest probability density is around the quality value zero. This means that in most of the control cases even the optimal paths have been loaded close to capacities of the paths.

Fig. 3. The probability density distribution of the quality values of the optimal paths

4. Conclusions

In this paper, we have presented the distributed data rate and path selection control system for the generic controllable (GC) traffic in a generic packet switched backbone network. The key idea of the control system is to control both the data rates and paths of the GC traffic, so that the free capacity of the network, after the load of the high priority traffic, can be optimally used for the GC traffic. The control functions are based on the use of fuzzy controllers, ANFIS neuro-fuzzy inference system and statistical calculations. The control functions of the control system are distributed to the edge switches of the network.

According to tests, the mean throughputs of the GC traffic flows are about 0.97 of the theoretical maximum throughput value calculated according to the max-min principle. Although the throughput values are high, the control system can avoid congestion in most of the load situations. However, more research work is needed to study the problems of the dynamic path selection procedure.

References

1. Paul Ferguson, Geoff Huston. Quality of Service, DeliveringQoS on the Internet and in Corporate Networks. Wiley Computer Publishing, John Wiley & Sons, Inc. 1998, Chapter 9.
2. Ahmet Sekercioglu, Y.; Pitsillides A., Intelligent control techniques for efficient regulation of ABR queue length in ATM switches, Proceedings. Second IEEE Symposium on Computers and Communication (Cat. No. 97TB100137), p 80-4
3. C. Douligeris, G. Develekos. A fuzzy logic approach to congestion control in ATM networks. 1995 IEEE International Conference on Communication, Seattle, pp. 1969-73, Vol. 3.
4. A. Pitsillides, Y. A. Sekercioglu, and G. Ramamurthy. Effective Control of Traffic Flow in ATM Networks Using Fuzzy Explicit Rate Marking (FERM) IEEE Journal on Selected Areas of Communications. Vol. 15, No. 2, February 1997, pp. 209 – 225.
5. G. Ascia, G. Ficili, D. Panno. Design of a VLSI Fuzzy Processor for ATM Traffic Sources Management. Proc. of the 20Th Conference on Local Computer Networks, Minneapolis, Oct. 1995, pp. 62-71
6. Aboelela, E.; Douligeris, C. Fuzzy Metric Approach for Routing in B-ISDN. ICC '99. 1999 IEEE International Conference on communications. 6-10 June, 1999 p. 484 – 488
7. J.-S. R. Jang. ANFIS: Adaptive-Network-Based Fuzzy Inference System. IEEE Transactions on Systems, Man, and Cybernetics. Vol. 23 no. 3. May/June 1993.
8. J.-S.R. Jang, C.-T. Sun, Predicting chaotic time series with fuzzy if-then rules. Second IEEE International Conference on Fuzzy Systems 1993. Pages 1079-1084. vol 2.
9. D. Bertsekas, R. Gallager. Data Networks. Second edition. Prentice-Hall, Englewood Cliffs, New Jersey, 1992. Section 6.5.2.
10. Mischa Schwartz. Broadband Integrated Networks. Prentice-Hall, A Simon & Schuster Company, Upper Saddle River, New Jersey.