# FLOW CONTROL ARCHITECTURE IN HETEROGENEOUS PACKET NETWORKS

JULIJE OŽEGOVIĆ

FESB Split, tel. 021 305725, fax. 563877
Julije.Ozegovic@fesb.hr

*Abstract: Congestion control, consisting of congestion avoidance and elimination, is of special interest in high speed - high delay packet networks. The flow control mechanisms are used as main congestion control technique. A model of heterogeneous future networks is introduced, and flow control problems are foreseen. Three flow control architectures are proposed. Conclusions are developed.*

KEYWORDS: flow control, congestion control, ATM, Internet

## INTRODUCTION

Congestion control is a key element of the network architecture. It provides optimal network utilization, while assuring declared quality of service (QoS). Congestion is defined as a situation when in the time interval of observation the offered traffic is greater then the transfer capacity of the network. When congestion occurs, some data are lost and QoS declines. Congestion control is accomplished through congestion avoidance and congestion elimination [6].

Congestion avoidance is a set of actions designed to keep the network at optimal working point, one that satisfies network utilization and QoS criteria. Congestion elimination is another, possibly overlapping set of actions, intended to take the network out of congestion, once it had happened because of any possible reason. In packet networks, flow control is one of the most important congestion avoidance mechanisms.

Chapter 1. contains an introduction to the congestion control problems. In Chapter 2. the flow control is elaborated. In Chapter 3. the architecture of heterogeneous networks is introduced and the flow control architecture of heterogeneous networks is analyzed. Conclusions and future work proposals are developed in Chapter 4.

## 1. CONGESTION CONTROL

Congestion avoidance and elimination measures are accomplished in multiple time scales and on all levels of hierarchical network architectures. They are applicable to networks with various switching techniques. In this article we are concerned with congestion control in packet switching networks, including ATM and heterogeneous networks.

### 1.1. Congestion Types by Time Scale

The definition of congestion (see Introduction) lacks precise declaration of the observation period. Regarding its length we can observe permanent, periodic, temporary and momentary congestion, Table 1.

| Congestion: | Avoidance measures: | Elimination measures: |
|---|---|---|
| PERMANENT | on-time resource planning and building | resource leasing |
| PERIODIC | pricing policy, admission control, dynamic routing | using resources from other time zones and user profiles |
| TEMPORARY | flow control | packet discarding |
| MOMENTARY | traffic shaping | providing enough buffer capacity |

*Table 1. - Congestion types and measures*

**Permanent congestion** is the result of increased user demands along with late investments in resource development.

**Periodic congestion** is induced by life and working dynamics of users, who tend to demand the service simultaneously.

**Temporary congestion** is in the time scale of seconds and minutes, and its duration is often declared as several (more than one) round trip times (RTT). It happens as the result of stochastic character of the connection setup and clearing process. It can be avoided with flow control mechanisms, and eliminated by packet discarding.

**Momentary congestion**, lasting less then few hundred milliseconds, is shorter than RTT. It is the result of uneven source traffic intensity and possible delay of flow control. Momentary congestion can be avoided by source traffic shaping, and eliminated by providing sufficient buffer capacity at routers.

### 1.2. Congestion Control by Hierarchy Layers

Congestion control algorithms can be systematized by network hierarchy layers, since they are implemented as parts of protocol mechanisms. Protocols can be connection oriented or connectionless. Connection oriented protocols implement error control mechanisms with Protocol Data Unit (PDU) enumeration and retransmission. Those mechanisms need initialization and synchronization, which are often referred to as logical channel establishment and maintenance. Connectionless protocols lack error control (limited to rejection of damaged PDUs), so there is no need for logical channel establishment. A survey of the congestion control methods systematized by architecture layers is shown in Table 2.

| Layer: | Connectionless: | Connection oriented: |
|---|---|---|
| PHYSICAL | transmission turning on and off, under data link layer control | |
| DATA LINK | transmission on and off for all data flows (back-off) | transmission on and off for single logical channel |
| NETWORK | packet routing, congestion indication, admission control, packet discarding | virtual channel, window flow control, admission control, packet discarding |
| TRANSPORT | flow control with explicit or implicit congestion indication, without QoS guarantees | window flow control, with explicit or implicit congestion indication, QoS guarantees |

*Table 2. - Congestion control by architecture level*

On the **physical layer**, the only possibility is to switch the transmitter on and off for all data flows simultaneously. This is done under control of the data link layer.

On the **data link layer**, congestion control is performed on physical or logical channel between two nodes. Delay on this straight connection is low (except for satellite channels), and direct flow control can be implemented. Transmission is switched on or off. The receiving node simply signals that memory buffers are filled above the limit, asking for temporary cessation of transmission. Packet buffering now occurs at the transmitting nodes and congestion spreads fast all over the network. Finally, the whole network is congested, even those parts that were not affected by initial congestion. Data link layer congestion should be avoided.

The **network layer** is used to establish data flows through the network, with connection oriented or connectionless protocols. Routing, admission control and flow control are performed here. Nodes are equipped with servicing algorithms which can ensure fair resource allocation and efficient operation. With virtual channel routing (X.25, ATM), flow control can be implemented easily because nodes have all information about data flows. With individual packet routing, data flows should be estimated first, counting with the possibility that a significant number of packets take different trajectories. In both cases, it is essential to inform the sources to decrease transmission rate for indicated data flows. That can be done with traffic shaping on the network layer, or indicating congestion to transport processes.

On the **transport layer**, connections between sources and destinations are realized. User data ready to be transmitted are already stored in source buffers. The natural solution is to implement flow control in this layer. The basic rule is to send as many packets as optimal for congestion avoidance, but no more than the advertised capacity of the receiver's buffer. The sender should receive explicit or implicit information about congestion or allowed transmission rate from the network layer.

## 2. FLOW CONTROL

The flow control, as basic measure for the temporary congestion avoidance, is the most important congestion control mechanism in packet switching networks. Flow control algorithms are accomplished in real time to keep the network at the optimal working point by means of ruling the transmitters' packet sending rate.

### 2.1. Optimal Network Working Point

The optimal network working point can be defined as the state vector of all network elements such as to allow the achievement of optimal relationship between network utilization and QoS. The network element state can be interpreted as number of packets in the output channel queues. Some common interest between network and user can be established considering minimal delay as main QoS parameter, and fast service makes network resources free for future demands.

Packets are stored in the output queues in the case of their arrivals being faster than the service rate. When node buffer is overflowed, packets are dropped. This initializes retransmission with double waste: the network capacity used for the dropped packet is lost, and QoS is degraded for the retransmission time. Packet losses by congestion should be avoided by all means. QoS needs empty queues, while network utilization requests some optimal number of packets in queues. The queuing theory gives answers about direction in which the solution is to be found.

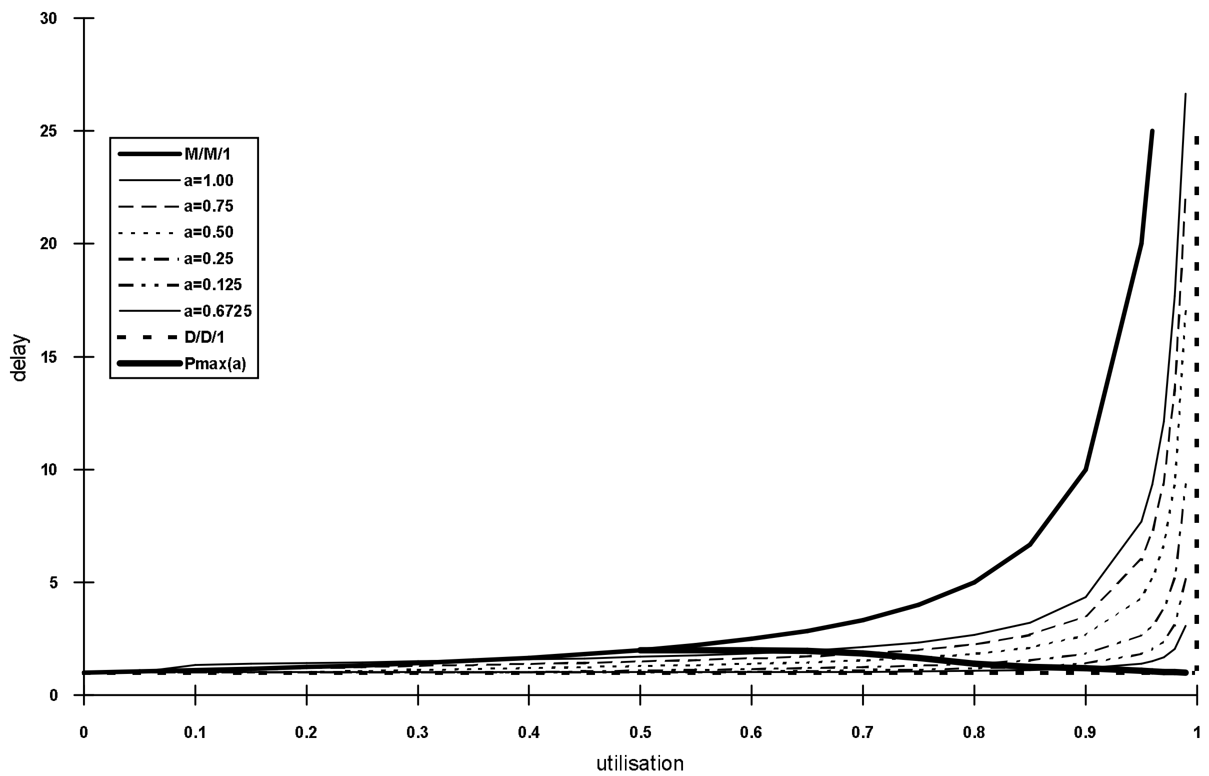


*Figure 1. - Delay for range of queuing systems*

The packet network can be observed as a black box, and modeled with a single server queuing system. The transmitter sends data packets and receives acknowledgments for individual packets, which makes it possible to keep track of round trip times (RTT) and window (W). The receiver can be considered part of the network as its role is to issue acknowledgment for each packet received.

For the sake of simplicity arrival and servicing processes are often considered Poisson. More general distributions are found, however, by practical measurements. Finally, we can consider determined processes, such as found in synchronous TDM systems like PCM hierarchy. Corresponding M/M/1, G/G/1 and D/D/1 queuing systems can be taken into account. Waiting times for a range of these systems are plotted in Figure 1.

The bold line plots the maximum network power points for subsequent distributions. It can be observed that, as variance decreases, network power is maximized for shorter delay and better network utilization. Two strategies for the achievement of the optimal working point can be defined, Figure 2.

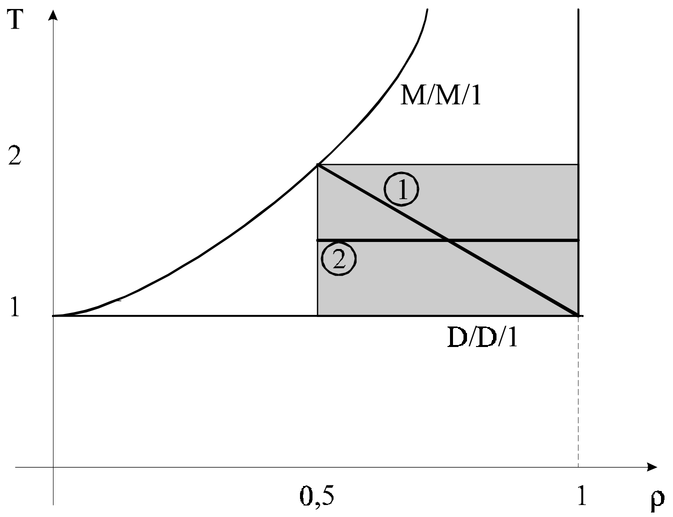


*Figure 2. - Optimal working point strategies*

In the first case, maximum power should be achieved. Lowering the variance of both input and output processes, network utilization rises slowly, together with improvement of QoS. In the second case declared QoS is maintained and network utilization rises fast with variance decrease. In both cases, variance of interarrival and service times should be kept as low as possible.

The optimal working point of the network should be achieved in such a manner that all users acquire equal part of the network capacity, weighted by the fraction of the capacity requested. This is not a trivial issue, since data flows take different paths through the network. The most common approach is the fair partitioning of the bottleneck link capacity, which leads to the min-max algorithm [9]. Fairness is judged by various criteria functions, such as the one in [5].

## 2.2. Congestion Detection and Algorithms

Flow control mechanism is in its nature a feedback system with the purpose to keep the optimal working point. Feedback information about network state is fed to the sources which adjust their sending rate in accordance with the network capacity available. We distinguish between implicit and explicit congestion detection, while in both cases transmitters and network nodes perform flow control algorithms.

In the case of **implicit** congestion notification, the sources are required to estimate the network state measuring the available values: RTT, W and packet loss rate. The switching nodes can enforce some of those parameters, i.e. induce packet losses in a random manner (RED), [3] or delay individual data flows in accordance with fairness criteria, [7]. With implicit congestion detection, the network is considered a black box.

Examples of implicit congestion indication protocols are TCP [4], TCP-Vegas [1] and fair queuing with packet-pair probes [7].

In the **explicit** congestion notification systems, network nodes inform the sources about their state with more or less resolution, ranging from single bit indications to multiple parameters delivery. Since the network is distributed, the nodes should estimate their local measure of congestion, observing queues' lengths, packet waiting times or packet loss rates for each output channel and class of service.

After the local congestion measure has been evaluated, that information should be communicated to the sources. Here we confront the gap between network and transport layer of the hierarchy structure. While network operations take place on the network level, distinguishing only terminal equipment, actual data flows are maintained on the transport level, where user processes are identified. In the case of strict layer independence, the network has no knowledge about transport protocol parameters. The nodes can recognize packets belonging to the same source-destination terminal pair as a single data flow. However, this approach is not fair in case of the firewall protected large subnetworks, or other multiuser hosts originating multiple connections. On the other hand, with the increased resolution of data flows to a single connection, the user can intentionally establish multiple connections to acquire more than a fair portion of the network capacity.

Examples of explicit congestion indication protocols are source quench messages of ICMP [10], EFCI systems as DEC-BIT [6], and explicit rate indication systems as ERICA [2].

To achieve stability of feedback systems, the sources should take care of sending rate and window [8]. Measured or communicated parameters should be filtered to eliminate noise from useful information. Exponential averaging filters are most common because of their simplicity and efficiency.

## 3. FLOW CONTROL IN HETEROGENEOUS NETWORKS

In the flow control analysis it is common practice to consider the network to be homogenous, meaning based on the same technology. Where Internet is concerned this assumption is true in the network and higher layers, while other networks are used just as channel providers. Those networks can maintain a flow control of their own, but the IP protocol is unaware of that possibility.

### 3.1. Heterogeneous Networks Architecture

The mainstream of network development is concentrated in the field of different traffic category integration. Starting with low delay jitter voice communications and ending with low loss data communications, the compromise is found in asynchronous transfer mode (ATM), a constant length packet (cell) switching technology. The CBR, VBR, ABR and UBR services are provided.

The ATM is highly perspective at the moment because of its compatibility with the SONET/SDH optical transmission system and its interchangeability with the STM fixed switching technology, emulating classical PCM hierarchy. We can suppose that the ATM technology will be predominantly used as wide area networks (WAN) infrastructure.

In the field of local area networks (LAN), the fast adoption of the ATM is doubtful. The 1Gb/s Ethernet technology is highly competitive, considering compatibility with millions of existing installations. Therefore, we think of future networks as Ethernet LAN installations connected through public ATM WAN. Such architecture is called heterogeneous, in Figure 3 based on the Internet framework:

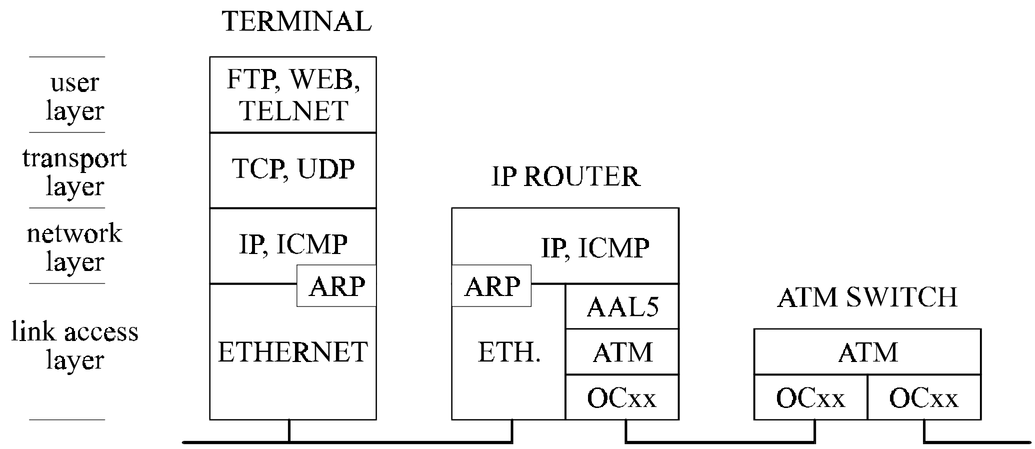


*Figure 3. - Architecture of heterogeneous networks*

The introduction of the ATM technology will take several steps. At the first moment, the IP layer should use only permanent virtual channels (PVC) of the ATM network, Figure 4. This solution does not require any signaling between the IP and the underlying ATM network. The logical IP channels (LCs) are just multiplexed over PVCs as over any other communications channel.

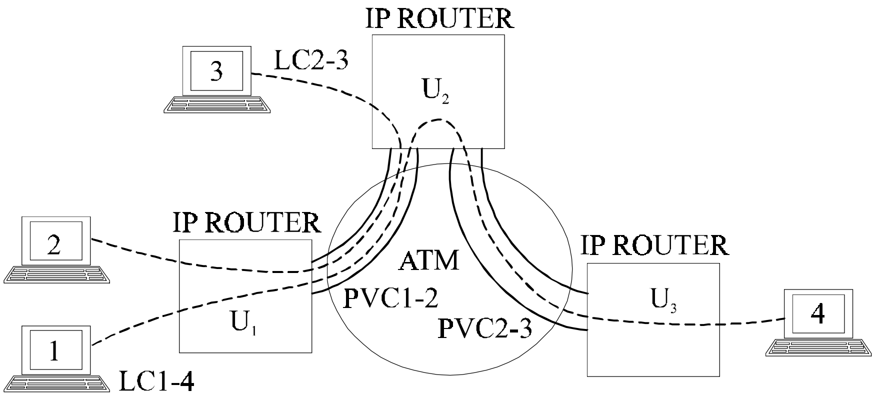


*Figure 4. - First step: IP using PVCs*

The main drawback of this solution is the lack of flexibility that can be provided by the switched virtual channels (SVC). The next step will introduce PVCs established on user's demand and with the user's specified service parameters, Figure 5.

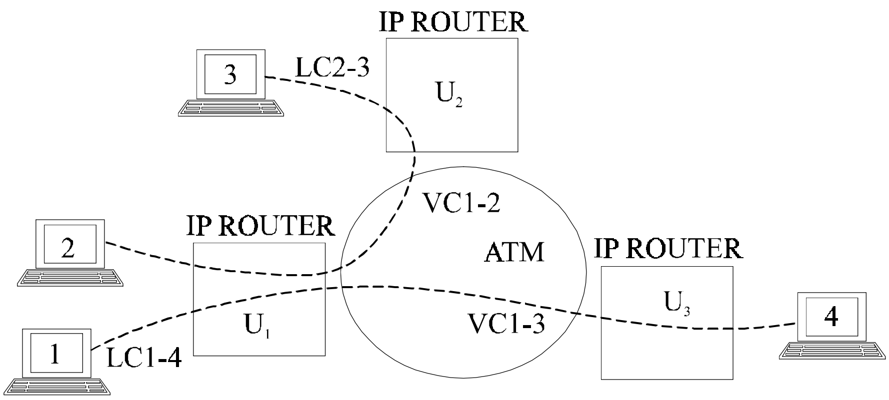


*Figure 5. - Second step: IP using SVCs*

In this solution, signaling is to be established between the IP and the ATM network to control the SVC connecting and clearing process. The IP users can now access the ATM services and use optimal routing of the global public ATM network, but the IP to ATM address conversion should be provided.

### 3.2. Flow Control in Heterogeneous Networks

Recently, researchers have been concerned with the performance of TCP protocol over ATM network [11]. The results have shown that a network utilization between 10% and 95% can be expected, being sensitive to parameter variations as the consequence of discrepancy between the IP and the ATM flow control.

On the heterogeneous networks, the situation is even worse. Workstations may not be directly connected to the ATM network, the flow control signaling is difficult, and the flow control integration is almost impossible.

We propose three perspective flow control architectures, without parameter signaling, with parameter signaling and with local flow control loops.

Networks without flow control parameter signaling make possible only data shaping on the edge of the ATM network, Figure 6.

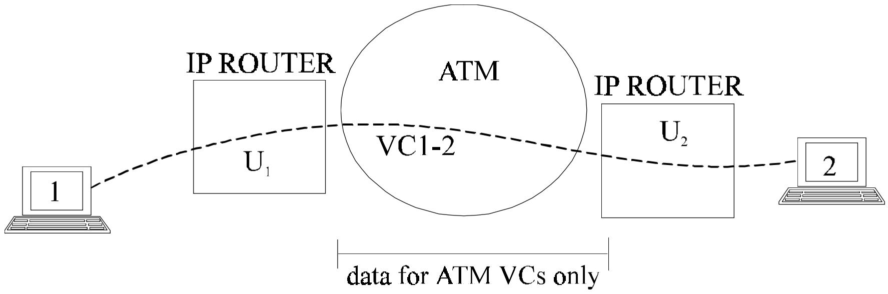


*Figure 6. - Network without flow control signaling*

This solution ensures minimal loss of inconsistent cells in the ATM network, but induces packet queues growth on the edges of the ATM network. The ATM part is protected from undisciplined users. The scenario is the only possible for UBR channels. It does not demand external protocol specification changes.

Networks with flow control parameter signaling make it possible to coordinate the transmitter rate with the SVC capacity, Figure 7.

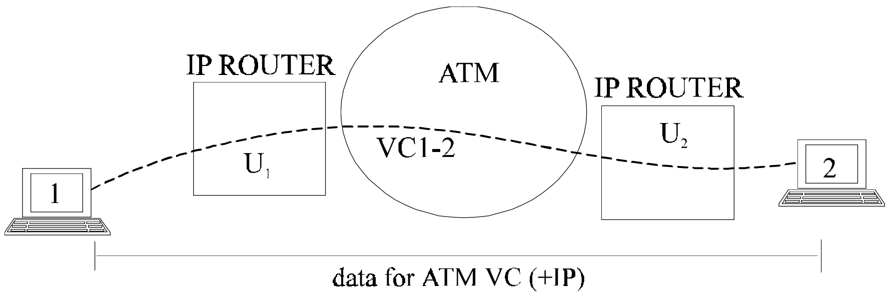


*Figure 7. - Network with flow control signaling*

If the problem of congestion on the IP part of the network is adequately solved, the source should transmit at the lower of the two signaled rates. This approach can be applied to ABR, CBR and VBR channels. The compatibility with the existing protocol standards is lost, and the external protocol specification is to be changed.

Closing of the global flow control loop and induced overall delay can be avoided if the local flow control loops are closed, Figure 8.

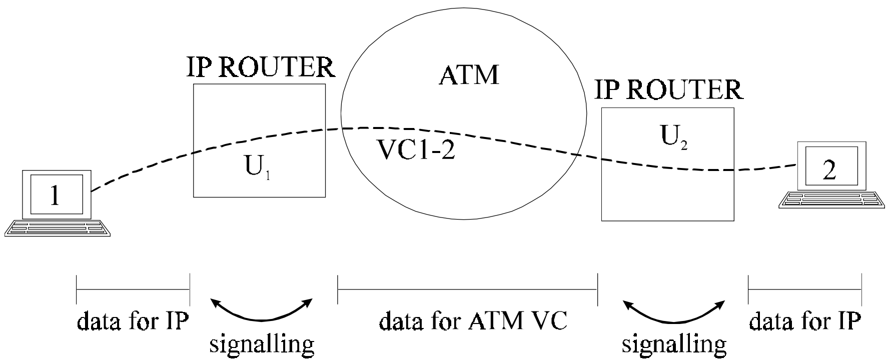


*Figure 8. - Network with local flow control loops*

The logical (TCP) channel should be divided in segments on the homogenous parts of the network. On the edges, this concept uses virtual receivers - transmitters (agents) which have knowledge about the state of both parts of the network. The current IP addressing architecture makes it difficult to segment a logical channel, while the users expect homogenous end to end connection. Segmentation should be made transparent to the users.

## 4. CONCLUSION

In the near future, the process of the ATM technology introduction will result in heterogeneous networks of LANs interconnected through the public ATM network, under the Internet services framework. The first step shall involve PVCs, while in the second the advantages of PVCs are to be used. The flow control in heterogeneous networks is difficult because of the incompatibility between TCP and ATM algorithms.

Three possible flow control architectures are discussed: without parameter signaling, with parameter signaling, and with local flow control loop closing. Future research should concentrate on developing enhanced flow control algorithms based on proposed architectures. The problems of congestion indication in heterogeneous networks should be solved.

**References:**


[1] Brakmo, L. S. O'Malley, S. W. and Peterson. L. L. "TCP Vegas: New techniques for congestion detection and avoidance”, In Proc. of ACM SIGCOMM '94, pp 24-35, May 94.

[2] Charny, A. Clark, D. D. and Jain, R. "Congestion Control with Explicit Rate Indication", Proceedings ICC'95, June 1995.

[3] Floyd, S. and Jacobson, V. "Random Early Detection Gateways for Congestion Avoidance", IEEE/ACM Trans. on Networking, Vol. l, No. 4, August 1993, pp. 397-413.

[4] Jacobson, V. "Congestion Avoidance and Control", Proc. ACM SIGCOMM '88, Stanford, Calif., August 1988, pp. 314-329.

[5] Jain, R. Chiu, D. M. and Hawe, W. "A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Computer Systems", Digital Equipment Corporation, Technical Report DEC-TR-301 Sept. 1984, 37 pp.

[6] Jain, R. Ramakrishnan, K. K. and Chiu, D. M. "Congestion Avoidance in Computer Networks with a Connectionless Network Layer", In C. Partridge, Ed., Innovations in Internetworking, Artech House, Norwood, MA, 1988, pp. 140-156.

[7] Keshav, S. Agrawala, A. and Singh, S. "Design and Analysis of A Flow Control Algorithm for a Network of Rate Allocating Servers", In Protocols for High Speed Networks II, Elsevier Science Publishers/North-Holland, April 1991

[8] Keshav, S. "A control-theoretic approach to flow control”, In Proceedings of ACM SIGCOMM'91, pp 3-15, Zurich, Switzerland, September 1991. ACM.

[9] Ma, Q. Steenkiste, P. and Zhang, H. "Routing High-bandwidth Traffic in Max-min Fair Share Networks" Proceedings of SIGCOMM'96, Stanford, California, August 1996.

[10] Postel J. “Internet Control Message Protocol”, RFC 792, IETF, September 1981,

[11] Romanow A. and Floyd, S. "The dynamics of TCP traffic over ATM networks”, IEEE Journal on Selected Areas in Communications, Vol. 13, No. 4, pp. 633-641, May 1995.