OPTIMAL WORKING POINT ANALYSIS FOR THE HETEROGENEOUS PACKET NETWORKS

JULIJE OZEGOVIC

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. The network is considered a black box, and modeled with G/G/1 queuing system. Numeric results are obtained for a range of arrival and service variances. The optimal network working point strategy is proposed.

KEYWORDS: flow control, congestion control, G/G/1, ATM, Internet

1. INTRODUCTION

The evolution of telecommunications has advanced in two directions, through the telephone and data networks growth. Nowadays, telephone (voice transfer) network is implemented as a fixed capacity (64 kb/s) channel switching system, spreading worldwide as a public switched telephone network (PSTN). Data networks use the packet switching techniques to achieve statistical multiplexing of the user data, where Internet is the most grown up public network carrying the same amount of bits per month as worldwide PSTN does.

Two types of traffic, voice and data, have in some way been integrated from the very beginning of the telecommunications era. In the period of analogue technology, frequency multiplexing has been used to embed multiple telegraph channels in the telephone one, or to use the residual capacity of the physical channel. PSTN is used for modem data transfer, acting as the access network or providing extra channels for the data network nodes interconnections. Digital technology offers more choices, however data networks at present dominantly use leased channels from the telephone network.

We can define four levels of integration. On the **primary channel level** some physical channels are used for the data transfer. On the **secondary channel level** secondary channels realized using the frequency (FDM) or time (TDM) domain multiplexing are used for data transfer, so data and telephone networks use the same physical channel simultaneously. On the **ISDN specification level** the user access network integration is achieved through the 2B+D specification, while the further information flows take separate secondary channels. On the **B-ISDN specification level** using the small fixed length packet (cell) switching, the full integration of various information types transfer is achieved. The current trend is to achieve higher level of integration. This process is slowed down because of the existing differences between the packet and telephone networks.

The telephone network is characterized with channel switching, where primary channels are manifold used through the fixed channel multiplexing (PCM). Once the switched channel is established, the propagation delay is equal to the sum of delays on the individual segments, and is mainly dependent on the geographical distance between nodes. The quality of service (QoS) is described in terms of signal to noise ratio, and user is provided with the constant 64kb/s capacity of the switched channel.

The data networks use packet switching as the basic mechanism of statistical multiplexing of the user data flows. The delay is mainly dependent on queuing delays in the network nodes, with the addition of propagation delays. The QoS is expressed in terms of delay and bit error rate (BER). User is provided with variable capacity dependent on his own as well as the needs of other users, meaning dependent on the network load.

The architecture of the modern packet switching data networks is realized as the hierarchical layered structure based on the concepts of protocols, interfaces and headers. The ISO-OSI specification is often used as the reference hierarchical network architecture consisting of seven layers.

The process(n), running on the layer(n) provides the service to the higher layer(n+1) using the services of the lower layer(n-1). Real data flow is carried through the interfaces of the adjacent layers, and finally through the communications channel to the remote computer. Seemingly, process(n) communicates directly to the corresponding process(n) in the remote computer using the protocol(n) rules.

The basic protocol mechanisms are addressing, synchronization, flow control and error control. The supervisory information for these mechanisms is contained in headers. Process(n) generates protocol(n) data unit (PDU) adding its own header(n) to the service(n) data unit (SDU). PDU(n) is sent through the interface(n-1) to the lower layer as the SDU(n-1). In the reverse direction, process(n) discards the header(n) from the PDUs received, thus forming the SDU(n), and forwards it to the higher layer through the interface(n). For the process(n+1), SDU(n) serves as PDU(n+1).

The primary role of addressing mechanism is the SDU user identification, which can be data terminal connected to the network, or the user process in that terminal. In more sophisticated structures, the protocol identification is needed. Addressing can be realized strictly, when higher layer address implies real addresses of all of the lower layers. This kind of addressing contributes to the layer functions separation. Addressing can be distributed, when total address is combination of separate layer addresses. This approach requires more connections between layers, but provides for easier address management and usage of well known addresses for some network services.

The synchronization mechanism is related to the PDU separation from the data bits stream as well as to the synchronous operation of the communicating processes of the same layer. Process(n) achieves its function through the assumed knowledge of the corresponding process state. The only information about process running on the remote computer is gained from the PDUs received. The process(n) algorithms should provide data transfer in the case of lost PDUs and incorrect assumptions about remote process state.

The error control mechanism makes the data flow restitution possible after the PDU is lost, mainly using the retransmission of the lost PDU. The flow control mechanism controls the PDUs sending rate not to exceed the capabilities of the corresponding receiver and network. The two mechanisms are connected tightly, because the main reason for PDU losses in modern networks is the network congestion. PDU losses due to communications errors are very rare. The PDU loss is often taken into account as the indication of the network congestion.

The retransmission procedure requires the lost PDU identification, which is achieved by PDU modulo m (cyclical) numbering. Concerning uniqueness of the numbering scheme, only m PDUs can be sent to the network simultaneously. This limits the number of PDUs, that transmitter is allowed to send to the network before any acknowledge is noticed from the receiver, to m. Constant m is called a window. Window can be determined explicitly, to express the ability of the receiver to accept some quantity of PDUs. Modern protocols use window mechanism for flow control as the congestion avoidance measure.

PDU numeration requires initial numbering synchronization and its consistent usage. Numbering synchronization is performed at the beginning of the data transfer, when a connection is established. This is called a logical channel establishment. Concerning error and flow control mechanisms, connection oriented and connectionless protocols are distinguished. Connectionless protocols lack error and (often) flow control mechanisms, so PDU numbering is not implemented. Logical channel is not needed. Processes using such protocols provide to the higher layers the PDU transfer service without QoS guarantees. On the contrary, connection oriented protocols use error control and (often) flow control mechanisms, establishing thus logical channels and providing QoS guarantees. After data transfer is finished, the logical channel is cleared. The important issue of the particular network architecture is the choice of the connection oriented and connectionless protocols disposal.

The example of the working network architecture is the Internet model, Figure 1.1. Network layer of the Internet model is fully comparable to the same layer of the ISO-OSI model. Internet transport layer combines the ISO-OSI transport and session layers. The application layer of the Internet performs functions of the ISO-OSI application and presentation layers (and some functions of the session layer). The most important difference is in the lower layers. While ISO-OSI model assumes strict standardization of the data and physical layer, Internet specification allows the usage of any network resource available. Under the network access layer, the complex networks with network and higher layers of their own can be found (X.25

or frame-relay). The most commonly used Internet protocols are listed in Figure 1.1. The Internet terminals (called hosts) obtain the functions of all layers, while Internet nodes (called routers or gateways) realize only functions of the network and lower layers.

APPLICATION LAYER	FTP, TELNET, WWW, SMTP, SNMP
TRANSPORT LAYER	TCP, UDP, GATEWAY PROTOCOLS
NETWORK LAYER	IP, ICMP ARP, RARP
NETWORK ACCESS LAYER	ETHERNET II, 802.3, 802.2, 802 SNAP

Figure 1.1 - Internet network architecture and related protocols

Another working architecture is the ATM network one, Figure 1.2. The homogeneous ATM network is based on the concept of virtual circuits (VC), which are used to transfer user information regarding the service type and QoS negotiated.

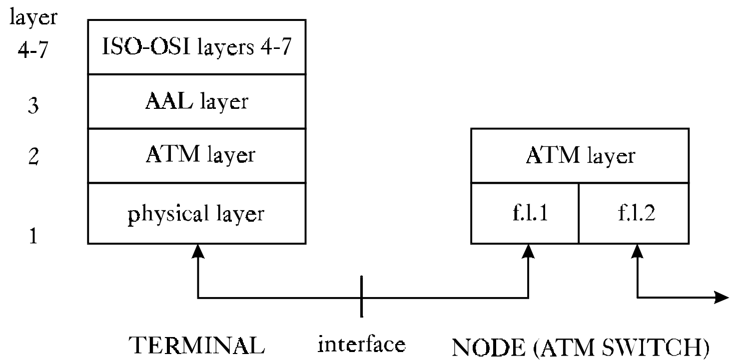

Figure 1.2 - ATM network architecture

Network nodes (ATM switches) fulfill only tasks of the ATM and physical layer. The ATM network provides the service of cell (48+5 octets) transfer and switching through the network. That is equivalent to the data layer of the ISO-OSI architecture. Terminals perform functions of all layers, where the ATM adaptation layer (AAL) provides interface between user data (PDU) and ATM network (cells).

Traffic management is performed in all types of telecommunications networks. Its goal is to achieve the optimal network utilization and QoS. Traffic management is accomplished through the procedures of access control, traffic routing and congestion control.

Customer access control (CAC) procedures verify the ability of the network to provide the QoS requested. If not, connection request is canceled. It is assumed that the user is going to inquire the same service later, when needed resources would eventually be free.

Traffic routing is performed in all networks as the main mechanism to obtain reachability between users. In packet networks, it can be done through individual packet routing (Internet), or using virtual channels (X.25, ATM) where all packets belonging to the same data flow (connection) pass across the same, on the beginning established path. Using alternate routing is the basic method of congestion avoidance in circuit switching networks. In packet networks, alternate routing is only an auxiliary method of congestion avoidance. Alternative routes are longer, and congestion is spread to previously unaffected parts of the network. It is recommended to use other methods for congestion control in packet networks.

Congestion control in packet networks includes congestion avoidance and congestion elimination procedures. Avoidance measures are performed when congestion have not been achieved yet, while elimination measures take place after the congestion has occurred because of any possible reason.

2. CONGESTION CONTROL AND FLOW CONTROL

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.

2.1. 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.

2.1.1. Congestion Types by Time Scale

The congestion definition above lacks the precise declaration of the observation period. Regarding its length we can observe permanent, periodic, temporary and momentary congestion.

Permanent congestion is the result of the increased user demands along with the late investments in resource development. It can be eliminated by resource leasing and avoided by on-time resource planning and building.

Periodic congestion is induced by the life and working dynamics of users, who tend to demand the service simultaneously. This type of congestion can be eliminated through the usage of other networks resources, or resources from other time zones. It can be avoided establishing CAC, stimulating users to change their behavior by tariff policy, and by traffic routing.

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 the flow control. Momentary congestion can be avoided by source traffic shaping, and eliminated by providing sufficient buffer capacity at routers.

2.1.2. 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.

On the **physical layer**, the only possibility is to switch the transmitter on and off for all data flows simultaneously. This is done under the control of 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 the 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 and admission control are performed here, including eventually flow control. 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 the nodes have all information about the 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 prepared to be transmitted are already stored in the source buffers. The natural solution is to implement the 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 the congestion or allowed transmission rate from the network layer.

The congestion control of a particular network architecture is closely related with the disposal of connectionless and connection oriented protocols. Although it is not imperative, flow control is mainly realized in connection oriented protocols. Some network architectures favor connection oriented protocols (X.25), specifying flow control mechanisms from the data layer up. Other prefer connectionless protocols (Internet), where flow control is obtained in the lowest connection oriented protocol (mainly on the transport layer within transport control protocol, TCP).

2.1.3. QoS and congestion control

In the channel switching networks, user utilizes the whole channel capacity, and the QoS is guaranteed in terms of channel bandwidth and signal to noise ratio. For digital channels, BER and bandwidth (bit rate) are declared.

In connectionless packet networks (Internet), the service is provided at the best effort policy, without any guarantees on errors, data rate and delay. Packets are routed according to the channel weights, excess packets are discarded, and users are obliged to supervise their own data integrity (TCP on transport layer). Internet is efficient for data transfer, but not appropriate for integration of the real time traffic (voice and video).

In connection oriented packet networks with virtual channel routing (X.25, ATM) error control and flow control are easily achieved. Network guarantees QoS, but regarding the finite possibility of data losses, users still need to supervise the data integrity.

ATM networks will be realized with integration of various traffic types as the main goal. Each traffic type requires specific QoS guarantees, which should be kept in spite of possible misbehaving users sending more than negotiated amount of data. ATM networks will provide services for non compressed voice and video (constant bit rate, CBR), compressed voice and video (real time variable bit rate, rt-VBR), data (nrt-VBR), residual capacity for data transfers with QoS guarantees (available bit rate, ABR) and residual capacity without QoS guarantees (unspecified bit rate, UBR). Table 2.1. lists the negotiated parameters for services provided:

<table>
<tr><th>negotiated parameter</th><th>CBR</th><th>rt-VBR</th><th>nrt-VBR</th><th>UBR</th><th>ABR</th></tr>
<tr><td>PCR, CDVT</td><td colspan="3">YES</td><td>YES</td><td>YES</td></tr>
<tr><td>SCR, MBS, CDVT</td><td>NO</td><td colspan="2">YES</td><td colspan="2">NO</td></tr>
<tr><td>MCR</td><td colspan="3">NO</td><td>NO</td><td>YES</td></tr>
<tr><td>CDV_{pp}</td><td>YES</td><td>YES</td><td>NO</td><td>NO</td><td>NO</td></tr>
<tr><td>CTD_{mean}</td><td>NO</td><td>NO</td><td>YES</td><td colspan="2">NO</td></tr>
<tr><td>CTD_{max}</td><td>YES</td><td>YES</td><td>NO</td><td colspan="2">NO</td></tr>
<tr><td>CLR</td><td colspan="3">YES</td><td>NO</td><td>YES</td></tr>
<tr><td>feedback</td><td colspan="3">NO</td><td>NO</td><td>YES</td></tr>
</table>

Table 2.1. - Negotiated QoS parameters for ATM services

where we have:

- PCR (Peak Cell Rate)
- SCR (Sustained Cell Rate)
- MCR (Minimal Cell Rate)
- CDVT (Cell Delay Variation Tolerance)
- MBS (Maximum Burst Size)
- CDV (Cell Delay Variation)
- CTD (Cell Transfer Delay)
- CLR (Cell Loss Ratio)

CBR users negotiate PCR, which serves as SCR, CDV, CDVT, CTD CLR and CLR. The delay variance is essential because of the playback process in the receiver. Late cells are useless and can be discarded in the network. Some cell losses are allowable without significant QoS degradation.

VBR users negotiate PCR and SCR in combination with MBS, CDVT and CLR. Real time users declare CDV and CTD_{max}, while non real time users negotiate CTD_{mean}. With real time users, QoS policy is somewhat similar to CBR, while non real time users negotiate less CLR and more CDV.

ABR virtual channels provide data transfer service that uses residual network capacity with some QoS guarantees. PCR, MCR and CDVT are negotiated along with CLR. Closed flow control feedback loop is used to signal the users about the actual cell rate available.

UBR service provides residual network capacity without any QoS guarantee. PCR and CDVT are negotiated. Network simply discards excess cells to eliminate congestion and inform the users to slow down transmission rates.

All services negotiations include CDVT, the parameter that is used to judge the conformance of every individual cell approaching the network. Non conforming cells are marked low priority or discarded at the entry node. After that, network guarantees QoS to CBR and VBR users by resource reservation during CAC. Flow control is fulfilled with open loop (no feedback).

With ABR service, mean cell rate is dynamically changed, and consequently is the conformance testing. SCR is signaled to the users. After that, network guarantees QoS. Flow control is fulfilled with closed loop (with feedback). UBR service is provided without QoS, with control loop opened and with cell loss indication of the congestion.

2.2. 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.

2.2.1. Optimal Network Working Point

The optimal network working point can be defined as the state vector of all network elements, needed to achieve the 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, while 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.

2.2.2. Fairness

The optimal working point of the network should be achieved in such a manner that all of the 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]:

$$F = \frac{\left(\sum_{i=1}^{n} A_i / A_i^*\right)^2}{n \sum_{i=1}^{n} \left(A_i / A_i^*\right)^2} \qquad \text{2-1}$$

where A_i is the achieved and A_i^* is the ideal portion of the network capacity belonging to the data flow observed. If the ideal fairness is accomplished, the value of function 2-1 is 1. When all the capacity belongs to only one customer, the value of 2-1 is 1/n (0 for the infinite number of customers).

2.2.3. Network Elements Functions

The network elements, nodes and terminals, perform different roles to accomplish the flow control functions of the network.

Nodes forward packets from input channels to output ones, storing them in the meantime in output queues. Packets from the queues are sent to channels according to the service algorithm, which provides accepted QoS. Data flows should be separated to protect responsive users from misbehaving ones. We distinguish services without specific algorithm (FIFO, first in first out), with stochastic and with deterministic algorithms.

- With FIFO algorithm, a node forwards customer packets in the order of arrivals, discarding the most recent packet when buffer overflows (drop-tail).
- Stochastic algorithms are used to discard packets at random, based on the assumption that more aggressive user should send more packets to the queue and thus have higher probability of random packet loss. One such algorithm is RED (Random Early Detection) algorithm [3], characterized with good performance in the heavy congested network when the numbers of packets in the queues are high.
- Deterministic algorithms take care about all data flows and make decisions about the individual packet services. One of the deterministic algorithms is FQ (Fair Queuing) algorithm [7,12], which tries to simulate the fixed multiplexing at the octet level. It requests significant amount of the processor power to process individual packets.

The data flow is any flow of PDUs recognized by node as a single unity, and the flow control functions are applied to it. Here we can observe the procedure of multiplexing, when multiple flows of the higher layer are sent as the single data flow of the lower layer. The node resolution is the ability to distinguish individual flows from the total traffic transpassing the channel, with more or less granularity. The resolution can be low, when source and destination subnetworks are recognized, through medium one when terminal pairs are identified, to high resolution when user process pairs are taken into account. The highest granularity opens the possibility that the same pair of customers uses multiple logical channels to acquire more than a fair portion of the network capacity.

Sources receive data (requests) from the higher layer, form packets (segmentation) and decide when to send them to the network. The packet sent is stored in memory until the acknowledgment is received. If a packet loss is detected, i.e. by acknowledgment timeout, packet is retransmitted. Sources perform flow control algorithms adjusting their windows and sending rates.

Destinations receive packets and send acknowledgments in the form of separate short packets or in the headers of data packets from the opposite direction. They make decisions about exact time to send acknowledgment and about the movement of the receive window upper bound. Excess number of small packets should be avoided, as well as segmentation of user data to small packets. For the flow control purposes, acknowledgments would be sent with minimum delay for each individual packet received.

2.2.4. Congestion Detection

Flow control mechanism is in its nature a feedback system with the purpose to keep the optimal network working point. Feedback information about the network state is fed to the sources which adjust their sending rates in accordance with the network capacity available. We distinguish between implicit and explicit congestion detection, while in both cases transmitters and network nodes perform the 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 the fairness criteria (FQ, [7]). With implicit congestion detection, the network is considered a black box.

Examples of the 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 the network and the 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 the 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 the 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].

2.2.5. Source Algorithms

After congestion is detected or communicated to the source, it should adjust its sending rate in accordance with new situation. This is achieved by performing the source algorithms. New rate is to be determined to reestablish optimal network working point. One can not rely on centralized, but only on distributed algorithms.

Measured or communicated parameters should be filtered to eliminate noise from the useful information. Exponential averaging filters are most common because of their simplicity and efficiency.

With explicit notification of the available capacity, sources can simply carry on with new rate. In practice, this is often filtered out to avoid sharp transitions. One such specification is for ATM ABR service [13], where the available (operating) rate (ACR) is increased to the explicit rate (ER) in steps determined by the rate increase factor (RIF).

When congestion is explicitly or implicitly detected, the available capacity is not known and the rate is increased or decreased using the source algorithms implemented. New rate is often calculated using linear functions with additive and multiplicative factors:

$$r(n+1) = \begin{cases} \alpha r(n) + \beta \ , \ r(n+1) > r(n) \\ \gamma r(n) - \delta \ , \ r(n+1) < r(n) \end{cases} \qquad \text{2-2}$$

It is shown that optimal source algorithm should use additive increase (α=0) and multiplicative decrease (δ=0) [6]. Users with higher rates slow down quickly, while free resources are divided with equal chances to all users.

Two groups of flow control algorithms are known, the window control and the rate control algorithms.

The window control is based on the number of packets limitation, which source can send to the network before acknowledgment has been received. A window was originally set to the receiver buffer capacity, and used to prevent sending packets which could not be stored in the receiver buffer. The resulting rate can be estimated from the receiver window (RWIN) and round trip time (T). The same mechanism is used as a congestion avoidance algorithm correcting the congestion window (CWIN). Source sends to the network the minimum of two windows of packets:

$$L = \frac{min(RWIN, CWIN)}{T} \qquad \text{2-3}$$

If the resources are limited with buffer capacity, window control is the algorithm of choice. Window control limits the total number of packets in the network efficiently. Its flaw is that the inbound traffic is not supervised and sources can generate bursts of packets.

The rate control is accomplished by transmission periods correction. Throughput is here defined as the number of the packets (octets, cells) N sent in the observation period t:

$$L = \frac{N}{t} \qquad \text{2-4}$$

The advantage of rate control is that the packet burst generation is avoided and the inbound traffic is controlled efficiently. The source characteristics approach the deterministic model. The flaw of rate control is that the number of packets in the network is not bounded and node buffers are not protected from overflow. Rate control is the algorithm of choice when resources are limited with the channel or processor capacity.

To achieve stability of feedback systems, the sources should take care of sending rate and window [8]. Regarding the advantages and flaws of the window and rate control, the combined algorithms should be developed and used.

3. OPTIMAL NETWORK WORKING POINT ANALYSIS

The packet network can be observed as a black box, Figure 3.1.

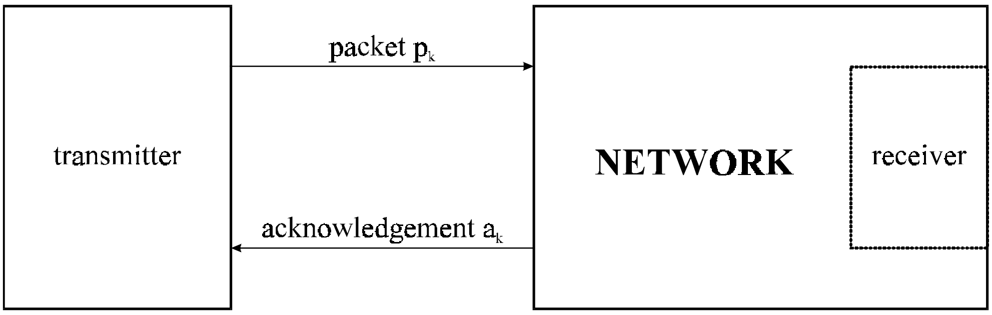

Figure 3.1. - Black box model of the network

The transmitter sends data packets and receives acknowledgments for individual packets, which makes it possible to keep track of round trip times (RTT):

$$T_k = t(a_k) - t(p_k) \qquad 3\text{-}1$$

and window (W):

$$W_k = p_k - a_j \;\; ; \;\; t(a_j) = t(p_k) \qquad 3\text{-}2$$

The receiver can be considered part of the network as its role is to issue an acknowledgment for each packet received. It is often assumed that a packet has exited the network at the moment of acknowledgment arrival, and the difference in packet and acknowledgment lengths is neglected.

With the black box approach, the network is optimally modeled with a single server queuing system, Figure 3.2.

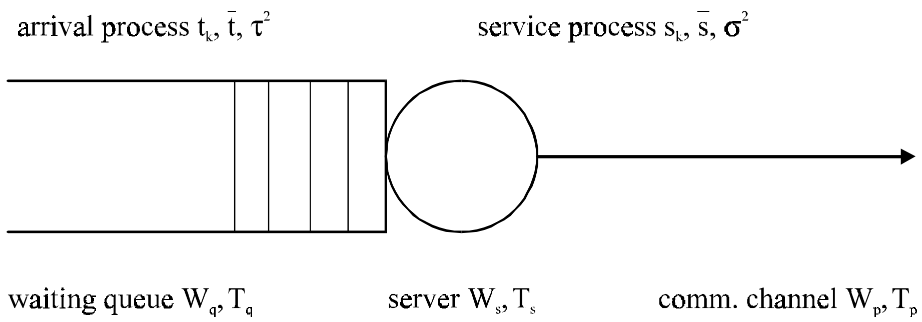

Figure 3.2. - Single queue model of the network

The statistical properties of packet arrivals and service are modeled each by a single stochastic process. Obtained results are acceptable compared to the queuing networks analysis, because arrival and service process simplifications are the same in both cases. Using the single queue model, analysis is much simpler and faster, and results are interpreted easily.

The time period spent by a packet in the network consists of the queuing time T_q and overall service time T_0, for communications system further divided in transmission (or service) and propagation times:

$$\begin{gathered} T = T_q + T_0 = T_q + T_s + T_p \\ T_0 = T_s + T_p \end{gathered} \qquad \text{3-3}$$

The service (transmission) time T_s is the time period from the moment of the first packet bit to the moment of the last packet bit have been sent to the channel, while propagation time T_p is the time period from the moment of the last packet bit has been sent to the channel to the moment that bit has exited at the receiver side of the channel. In our model, this is the moment the last bit of the acknowledgment has been received at the transmitter.

During the data transmission, W packets can be found in the system: W_q packets in the queue, $W_s = 1$ in the server (one server) and W_p on the communications channel:

$$\begin{gathered} W = W_q + W_0 = W_q + W_s + W_p \\ W_0 = W_s + W_p = 1 + W_p \end{gathered} \qquad \text{3-4}$$

If the average service rate is μ packets per second and average service time 1/μ seconds per packet, channel capacity C b/s and average packet length $\overline{M}$ bits per packet, the following relations stand:

$$\begin{gathered} T_s = W_s \frac{\overline{M}}{C} \underset{W_s=1}{=} \frac{\overline{M}}{C} = \frac{1}{\mu} \\ T_p = W_p \frac{\overline{M}}{C} = W_p \frac{1}{\mu} \Rightarrow W_p = T_p \mu = T_p \frac{C}{\overline{M}} = \frac{T_p}{T_s} \end{gathered} \qquad \text{3-5}$$

meaning that the average number of packets on the channel is proportional to the service rate (inverse of the service time), because both values depend on the average packet length and channel capacity.

For the sake of simplicity, the arrival and service processes are often considered Poisson. More general distributions are found, however, by practical measurements. Finally, we can consider determined processes, which can be 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. The analysis of the G/G/1 model is presented, and the results of M/M/1 and D/D/1 models are easily obtained.

G/G/1 model, often referred to as GI/GI/1 to emphasize the independence of subsequent events, is characterized by three basic properties:

1. Little law states, $E[W] = \lambda E[T]$.
2. The stability condition is $\lambda < \mu$ *or* $\rho = \lambda/\mu < 1$.
3. In stationary state, the system is statistically equal in the moments just before the packet arrival and in the moment just after the packet service.

We can define G/G/1 model such as:

1. $s_k > 0$ is service time of packet k, independent random variable with mean $E[s_k] = \bar{s} = 1/\mu$ and variance $E[s_k^2] - E[s_k]^2 = \sigma^2$, with probability density function $f_1(s)$ and cumulative density function (distribution) $G_1[\bar{s}, \sigma^2]$, where μ is the service rate.
2. $t_k > 0$ is interarrival time between packets k and k+1, independent random variable with mean $E[t_k] = \bar{t} = 1/\lambda$ and variance $E[t_k^2] - E[t_k]^2 = \tau^2$, with probability density function $f_2(t)$ and cumulative density function (distribution) $G_2[\bar{t}, \tau^2]$, where λ is the arrival rate.
3. one server sends packet from the queue to the channel of capacity C and propagation delay $T_p = 0$.

This system is exponentially bounded [14]:

$$E[T_q] \leq \frac{1}{\Theta^*} \tag{3-6}$$

where Θ^* is the biggest root of the Laplace-Stieltjes transform (LST) $\Phi(\Theta)$:

$$\Theta^* = \max_i(\Theta_i) \; : \; \begin{cases} \Theta_i > 0 \\ \Phi(\Theta_i) \leq 1 \end{cases} \tag{3-7}$$

defined as the function of the random variable z:

$$\Phi(\Theta) = E[e^{\Theta z_k}] \; ; \; z_k = s_k - t_k \tag{3-8}$$

To obtain the delay vs. utilization curves, one has to compute the worst case of Θ^* for a range of values $\rho \in (0,1)$, repeatedly for each $G_1[\bar{s},\sigma^2]/G_2[\bar{t},\tau^2]/1$ system. For simplicity, the constant propagation time is set to zero, $T_p = 0$, which can be added to the calculated waiting times easily.

In this analysis, the general distributions are substituted with the uniform ones, with their means and variances preserved:

$$\begin{aligned} G_1[\bar{s},\sigma^2] &\to U[\bar{s},\sigma^2] \\ G_2[\bar{t},\tau^2] &\to U[\bar{t},\tau^2] \end{aligned} \qquad 3\text{-}9$$

Substituting general distributions with normal ones leads to generally unsolvable transcendental equations.

Uniform distribution $U[\bar{x},\chi^2]$, defined with its mean and variance, is characterized with the probability $P(x) > 0$ existing in the region $a \le x \le b$, and with the probability density function plotted in Figure 3.3:

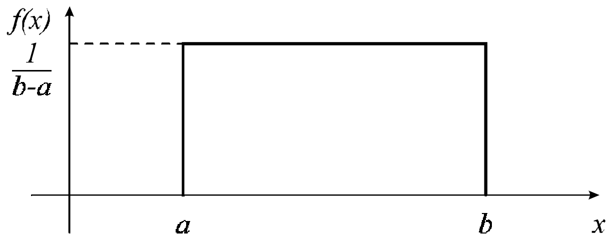

Figure 3.3. - Probability density function for uniform distribution

It stands:

$$P(x) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & x < a \;,\; x > b \end{cases} \qquad 3\text{-}10$$

$$D(x) = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a \le x < b \\ 1 & x \ge b \end{cases} \qquad 3\text{-}11$$

where:

$$\bar{x} = \frac{1}{2}(a+b) \qquad 3\text{-}12$$

$$\chi^2 = \frac{1}{12}(b-a)^2 \qquad 3\text{-}13$$

Borders are calculated from $\bar{x},\chi^2$:

$$a = \bar{x} - \sqrt{3\chi^2}$$
$$b = \bar{x} + \sqrt{3\chi^2} \qquad 3\text{-}14$$

with result:

$$P(x) = \frac{1}{2\sqrt{3\chi^2}} \qquad 3\text{-}15$$

In our case we have $U[\bar{s},\sigma^2]$ given with the mean:

$$\bar{s} = \frac{1}{\mu} \qquad 3\text{-}16$$

where μ is service rate, and with the variance σ^2. Probability is $P(s) > 0$ in range (a_1,b_1), so we have:

$$a_1 = \bar{s} - \sqrt{3\sigma^2}$$
$$b_1 = \bar{s} + \sqrt{3\sigma^2} \qquad 3\text{-}17$$

$$P(s) = \frac{1}{b_1 - a_1} = \frac{1}{2\sqrt{3\sigma^2}} \;\; ; \;\; a_1 \le s \le b_1 \qquad 3\text{-}18$$

$$D(s) = \frac{s - \bar{s} + \sqrt{3\sigma^2}}{2\sqrt{3\sigma^2}} \;\; ; \;\; a_1 \le s < b_1 \qquad 3\text{-}19$$

Similarly, $U[\bar{t},\tau^2]$ is defined with the mean:

$$\bar{t} = \frac{1}{\lambda} \qquad 3\text{-}20$$

where λ is arrival rate, and with the variance τ^2. Probability is $P(t) > 0$ in range (a_2,b_2) so we have:

$$a_2 = \bar{t} - \sqrt{3\tau^2}$$
$$b_2 = \bar{t} + \sqrt{3\tau^2} \qquad 3\text{-}21$$

$$P(t) = \frac{1}{b_2 - a_2} = \frac{1}{2\sqrt{3\tau^2}} \;\; ; \;\; a_2 \le t \le b_2 \qquad 3\text{-}22$$

$$D(t) = \frac{t - \bar{t} + \sqrt{3\tau^2}}{2\sqrt{3\tau^2}} \;\; ; \;\; a_2 \le t < b_2 \qquad 3\text{-}23$$

$U[\bar{s},\sigma^2]$ and $U[\bar{t},\tau^2]$ make two dimensional random vector defined above the rectangle region, Figure 3.4:

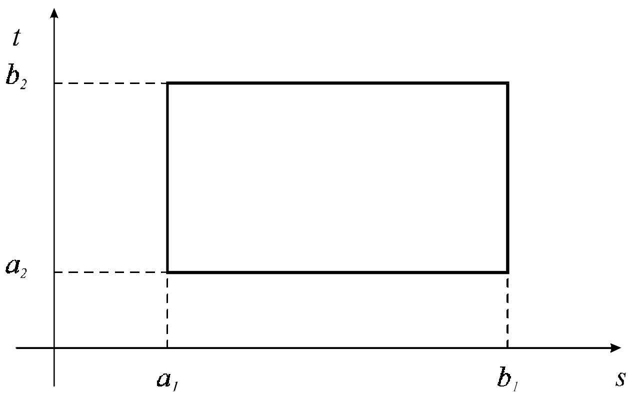

Figure 3.4. - Two dimensional random space (s,t)

Regarding s and t are independent random variables, each dot of the quadratic region is assigned the probability:

$$P(s,t) = P(s)P(t) = \frac{1}{4\sqrt{3\sigma^2}\sqrt{3\tau^2}} \;;\; \begin{matrix} a_1 \leq s \leq b_1 \\ a_2 \leq t \leq b_2 \end{matrix} \qquad 3\text{-}24$$

while for all other dots of the plane we have:

$$P(s,t) = 0 \;;\; \begin{matrix} s < a_1 \;,\; s > b_1 \\ t < a_2 \;,\; t > b_2 \end{matrix} \qquad 3\text{-}25$$

The stability condition of the system is:

$$\lambda < \mu \;\; or \;\; \rho = \frac{\lambda}{\mu} = \frac{\bar{s}}{\bar{t}} < 1 \qquad 3\text{-}26$$

with result:

$$\bar{s} < \bar{t} \qquad 3\text{-}27$$

This was expected, because system is going to serve all requests if, in the long enough period of observation, the possible number of services exceeds the total number of arrivals.

Regarding 3-8, we must find the expectation of the random variable $z_k = s_k - t_k$. Generally, for some function $Z(s,t)$ it stands [15,16]:

$$G(z) = P(Z \leq z) = \iint_{A_z} f(s,t)\,dt\,ds \qquad 3\text{-}28$$

where $f(s,t)$ is probability density function for the two dimensional random space (s,t). The set A_z contains all points of the plane that satisfy the condition $Z \leq z$. In our case it is:

$$Z(s,t) = s - t \qquad 3\text{-}29$$

and for the edge condition $P(Z = z)$ we have:

$$z = s - t \;\;,\;\; t = s - z \qquad 3\text{-}30$$

which is the formula of the line with parameter z in the (s,t) plane, Figure 3.5.

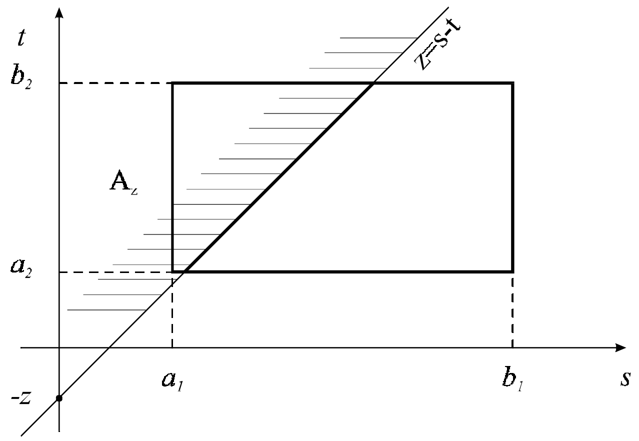

Figure 3.5. - Edge line in the (s,t) plane

The set A_z contains the points of the half-plane $z \le s-t$ left from the line (dashed). In the case of uniform distribution, this is reduced to the points of the rectangle left from the line. The part of the line inside the rectangle consists of points with same z, with probability greater than zero. We can write:

$$G(z) = P(s-t \le z) = \int_{-\infty}^{\infty} \int_{s-z}^{\infty} f(s,t)\,dt\,ds \qquad 3\text{-}31$$

and after derivation on z and elimination of integral on z (because $t = s - z$):

$$\frac{dG(z)}{dz} = g(z) = P(s-t = z) = \int_{-\infty}^{\infty} f(s, s-z)\,ds \qquad 3\text{-}32$$

The integral 3-32 sums all the probabilities of points on the line $t = s - z$, which for the uniform distributions means the probabilities of points inside the rectangle (thick).

Regarding s and t are independent random variables, it states:

$$f(s,t) = f_1(s)\ f_2(t) \qquad 3\text{-}33$$

and in this case:

$$f(s,s-z) = f_1(s)\ f_2(s-z) \qquad 3\text{-}34$$

which put into 3-32 gives:

$$g(z) = \int_{-\infty}^{\infty} f_1(s)\ f_2(s-z)\,ds \qquad 3\text{-}35$$

For uniform distributions $U[\bar{s},\sigma^2]$ and $U[\bar{t},\tau^2]$ where the probabilities are constant we have:

$$g(z) = P(s)P(t)\int_a^b ds = \frac{1}{4\sqrt{3\sigma^2}\sqrt{3\tau^2}}\int_a^b ds \qquad 3\text{-}36$$

The margins of integration are to be determined, Figure 3.6.

One can observe three areas of integration. In the area 1, the probability rises proportionally with z, in the area 2 probability is constant, and in the area 3 the probability decays proportionally with -z. Regarding:

$$z = s - t \,\Big|\, s = a_1, t = b_2 \qquad \text{3-37}$$

point z_1 can be obtained, and similarly all the others:

$$\begin{aligned} z_1 &= a_1 - b_2 = \bar{s} - \bar{t} - \sqrt{3\sigma^2} - \sqrt{3\tau^2} \\ z_2 &= a_1 - a_2 = \bar{s} - \bar{t} - \sqrt{3\sigma^2} + \sqrt{3\tau^2} \\ z_3 &= b_1 - b_2 = \bar{s} - \bar{t} + \sqrt{3\sigma^2} - \sqrt{3\tau^2} \\ z_4 &= b_1 - a_2 = \bar{s} - \bar{t} + \sqrt{3\sigma^2} + \sqrt{3\tau^2} \end{aligned} \qquad \text{3-38}$$

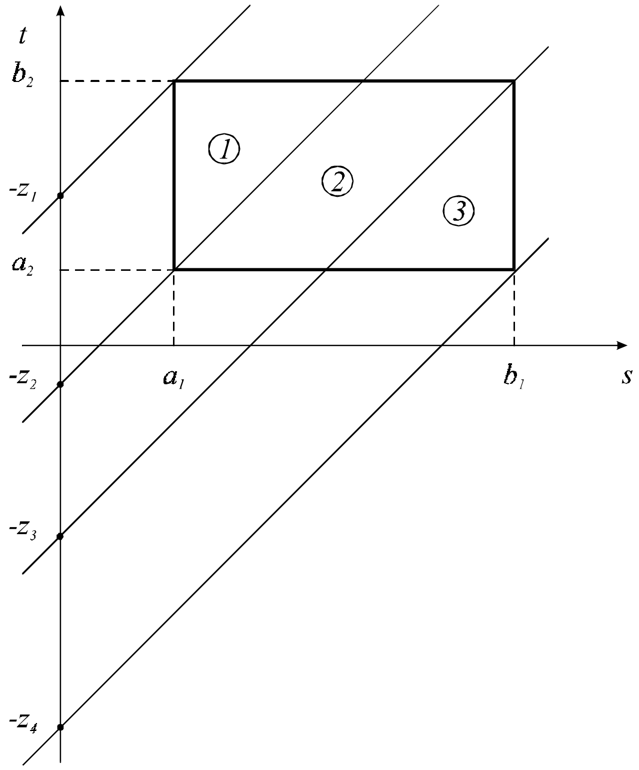

Figure 3.6. - Margins of integration for g(z)

Depending on the shape of the rectangle, points z_2 and z_3 can change their places. Now one can sketch the graph of g(z), Figure 3.7:

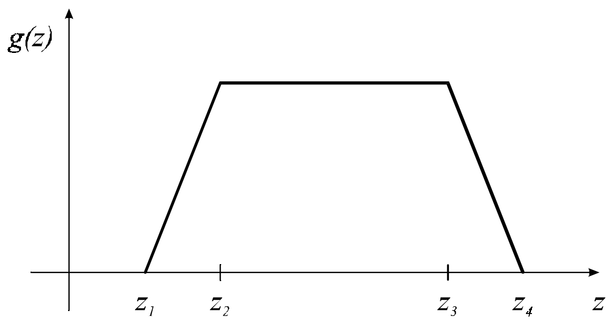

Figure 3.7. - Approximate shape of g(z)

or analytically:

$$g(z)=\frac{1}{4\sqrt{3\sigma^2}\sqrt{3\tau^2}}\begin{cases} z-z_1 & (z_1,z_2) \\ z_2-z_1=z_4-z_3 & (z_2,z_3) \\ z_4-z & (z_3,z_4) \end{cases} \qquad 3\text{-}39$$

The expectation from 3-8 can be obtained using the method of probability density function transformation [15,16]. If we have:

$$y=\varphi(x)\ ;\ x: P(x)=f(x) \qquad 3\text{-}40$$

then:

$$E[\varphi(x)]=\int_{-\infty}^{\infty}\varphi(x)f(x)dx \qquad 3\text{-}41$$

In our case:

$$E[e^{\Theta z}]=\int_{-\infty}^{\infty}e^{\Theta z}g(z)dz \qquad 3\text{-}42$$

in three areas of integration defined before. We have:

$$E[e^{\Theta z}]=K\left[\int_{z_1}^{z_2}e^{\Theta z}(z-z_1)dz+\int_{z_2}^{z_3}e^{\Theta z}(z_2-z_1)dz+\int_{z_3}^{z_4}e^{\Theta z}(z_4-z)dz\right] \qquad 3\text{-}43$$

where $K=P(s,t)$ regarding 3-24. After arrangement we get:

$$\Phi(\Theta)=\frac{K}{\Theta^2}\left(e^{\Theta z_1}-e^{\Theta z_2}-e^{\Theta z_3}+e^{\Theta z_4}\right) \qquad 3\text{-}44$$

It is convenient that 3-44 does not depend on relation between points z_2 and z_3.

The same result can be obtained regarding independence of s and t when calculating expectancy:

$$\begin{aligned}\Phi(\Theta)&=E[e^{\Theta z}]=E[e^{\Theta(s-t)}]=E[e^{\Theta s}e^{-\Theta t}]=E[e^{\Theta s}]E[e^{-\Theta t}]\\ &=\int_{-\infty}^{\infty}e^{\Theta s}f_1(s)ds\int_{-\infty}^{\infty}e^{-\Theta t}f_2(t)dt=P(s)\frac{1}{\Theta}e^{\Theta s}\Big|_{a_1}^{b_1}\ P(t)\frac{1}{\Theta}e^{-\Theta t}\Big|_{a_2}^{b_2}\end{aligned} \qquad 3\text{-}45$$

After the integration margins are inserted, the 3-44 is readily obtained.

To simplify the analysis and graph plotting, relation 3-44 is normalized to the unity value of the service rate μ, so varying the arrival rate from 0 to 1 corresponds with the network load ρ in range from 0 to 1:

$$\mu=1\ ,\ \lambda\in(0,1)\Rightarrow\hat{s}=\frac{1}{\mu}=1\ ,\ \hat{t}=\frac{1}{\lambda}\in(1\cdots\infty) \qquad 3\text{-}46$$

The service and interarrival times can not be negative, so variance for uniform distribution is limited when lower border is defined with $a=0$:

$$a_1 = \bar{s} - \sqrt{3\sigma^2} \geq 0 \Rightarrow \sigma^2 \leq \frac{\bar{s}^2}{3} \tag{3-47}$$

In the calculations, variances can be changed relatively from the maximum value of 3-47 towards 0, modifying parameter c_1 in the range $(0,1)$:

$$\sigma^2 = c_1 \frac{\bar{s}^2}{3} \ ; \ c_1 \in (0,1) \tag{3-48}$$

Similarly we use c_2 for the arrival distribution:

$$\tau^2 = c_2 \frac{\bar{t}^2}{3} \ ; \ c_2 \in (0,1) \tag{3-49}$$

For the analysis, it is enough to calculate curves for the case where relative variances for arrival and service processes are the same, $c_1 = c_2 = c$. Normalized to $\mu = 1$ we finally obtain:

$$\begin{aligned} \hat{\sigma}^2 &= \frac{c}{3} & c \in (0,1) \\ \hat{\tau}^2 &= \frac{c}{3\lambda^2} & c \in (0,1)\ ,\ \lambda \in (0,1) \end{aligned} \tag{3-50}$$

We introduce:

$$\bar{z} = \bar{s} - \bar{t} = \frac{1}{\mu} - \frac{1}{\lambda} = \frac{\lambda - \mu}{\lambda\mu} \tag{3-51}$$

and normalize to $\mu = 1$:

$$\hat{z} = \frac{\lambda - 1}{\lambda} \tag{3-52}$$

Including 3-50 and 3-52 in 3-38 we obtain:

$$\begin{aligned} \hat{z}_1 &= \frac{\lambda(1-c) - (1+c)}{\lambda} \\ \hat{z}_2 &= \frac{\lambda(1-c) - (1-c)}{\lambda} \\ \hat{z}_3 &= \frac{\lambda(1+c) - (1+c)}{\lambda} \\ \hat{z}_4 &= \frac{\lambda(1+c) - (1-c)}{\lambda} \end{aligned} \tag{3-53}$$

The constant K is defined as:

$$\hat{K} = \hat{P}(s,t) = \hat{P}(s)\hat{P}(t) = \frac{1}{4\sqrt{c}\,\frac{\sqrt{c}}{\lambda}} = \frac{\lambda}{4c} \tag{3-54}$$

We want to find the greatest root Θ of the transcendent equation 3-44 under the condition:

$$\Phi(\Theta) = E[e^{\Theta z}] \leq 1 \qquad 3\text{-}55$$

The problem with 3-55 is the inequality, meaning that the numerical result should be searched for in the predefined range. Observing the convergence of the process, adequate value for Θ^* should be chosen. On can write:

$$\beta\Theta^2 = K\left(e^{\Theta z_1} - e^{\Theta z_2} - e^{\Theta z_3} + e^{\Theta z_4}\right) \qquad 3\text{-}56$$

and the numeric results are obtained varying parameters in the following ranges:

$$c \in (0,1)\ ,\ \beta \in (-1,1)\ ,\ \lambda \in (0,1) \qquad 3\text{-}57$$

The set of values for Θ^* that uniquely defines the family of curves for $c \in (0,1)$, depending on $\rho \in (0,1)$, is obtained, Table 3.4:

ρ: c:	1,00	0,75	0,50	0,25	0,125	0,0625
0,05	25,000	50(0)	50(0)	50(0)	50(0)	50(0)
0,10	2,916	17,290	50(0)	50(0)	50(0)	50(0)
0,15	2,590	7,112	50(0)	50(0)	50(0)	50(0)
0,20	2,362	4,932	38,705	50(0)	50(0)	50(0)
0,25	2,160	3,925	13,642	50(0)	50(0)	50(0)
0,30	1,995	3,312	8,430	50(0)	50(0)	50(0)
0,35	1,841	2,879	6,194	50(0)	50(0)	50(0)
0,40	1,698	2,543	4,921	31,18	50(0)	50(0)
0,45	1,560	2,265	4,076	16,457	50(0)	50(0)
0,50	1,426	2,022	3,454	10,965	50(0)	50(0)
0,55	1,292	1,801	2,961	8,087	41,199	50(0)
0,60	1,157	1,592	2,545	6,282	21,591	50(0)
0,65	1,021	1,390	2,178	5,009	13,952	50(0)
0,70	0,881	1,192	1,839	4,026	9,865	33,595
0,75	0,739	0,994	1,518	3,211	7,247	19,277
0,80	0,594	0,797	1,208	2,495	5,338	12,398
0,85	0,448	0,599	0,903	1,835	3,798	8,160
0,90	0,299	0,399	0,601	1,210	2,453	5,054
0,95	0,149	0,200	0,300	0,601	1,206	2,428
0,96	0,120	0,160	0,240	0,481	0,963	1,934
0,97	0,090	0,120	0,180	0,360	0,721	1,446
0.98	0,060	0,080	0,120	0,240	0,480	0,962
0,99	0,039	0,047	0,062	0,120	0,240	0,480

Table 3.4. - Numeric results for Θ^*

The values written as 50(0) could not be obtained because of the convergence instability for low loads and variances, resulting with slight deviations in the left parts of the graphs. This is not of the importance for this analysis, because our main concern is about the system behavior under the heavy load.

We finally obtain the total delay using 3-3 and 3-6, $T_p = 0$, taking the worst case results:

$$T = T_q + T_s = E\left[T_q\right] + 1 = 1 + \frac{1}{\Theta^*} \qquad \text{3-58}$$

which are presented in the Figure 3.8:

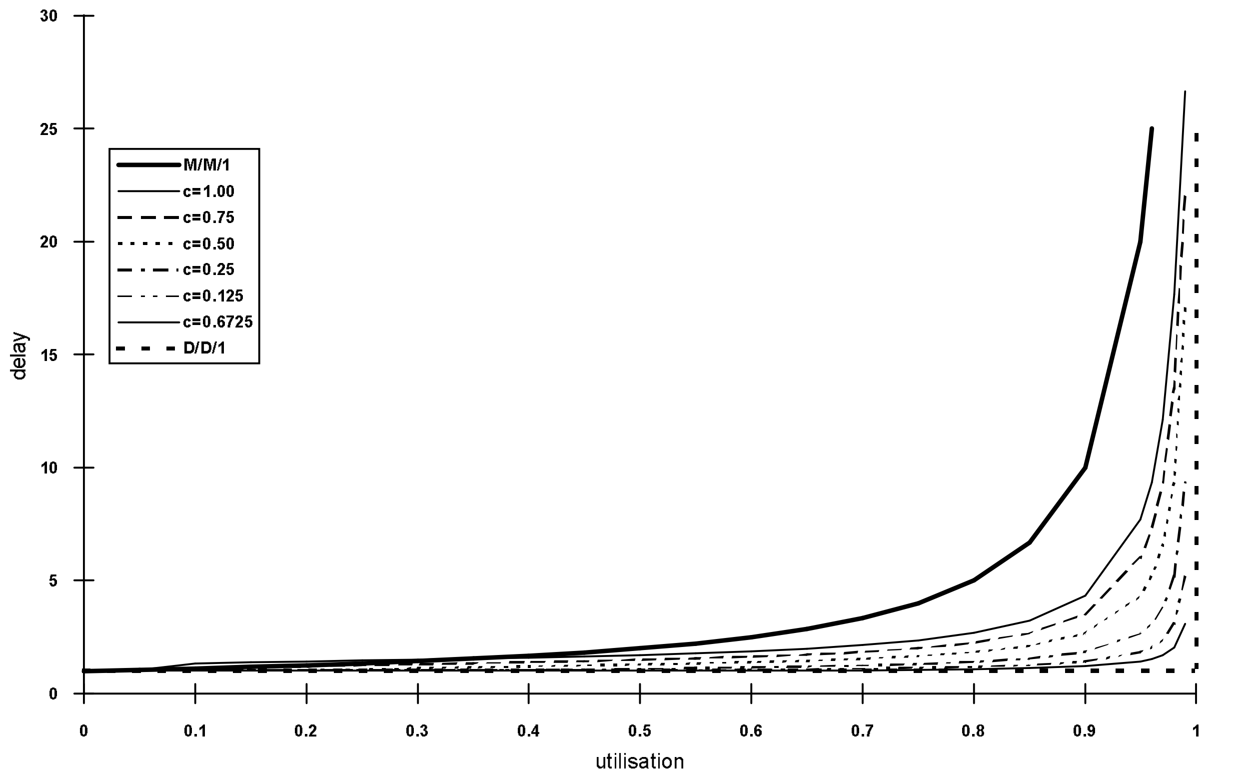

Figure 3.8. - Delay for range of queuing systems

It is clear that lowering the parameter c, thus lowering the variances for arrival and service distributions, delay curves approach the one for D/D/1 model (dashed thick). The M/M/1 curve is presented for reference (thick). The significant difference between the M/M/1 and the first G/G/1 curve is from the fact that maximum variance for uniform distribution is limited by the positive bound of random variables s and t. In the same time, the concentration of values for exponential distribution is around zero, leading to large variance. Considering that for $c = 0,25$ and $\rho = 0,9$ we have $T = 1,83$, one can conclude that it is possible to obtain the delay comparable to that for M/M/1 system for $\rho = 0,5$, in spite of the high relative variance and load.

The optimal working point is often determined for the maximum of "network power", defined as the ratio between throughput and delay:

$$P = \frac{L}{T} \; [b / s^2] \qquad \text{3-59}$$

where L is throughput in b/s and T total delay. Throughput is defined as the number of bits (octets) transferred in the unit of time. If user follows the packet number conservation law, i.e. keeps the number of its packets in the network constant, next packet should be sent to the network after the acknowledgment for the oldest one has been received. When the flow control is window controlled, the number of packets on the network from one transmitter is equal to the window W. The acknowledgment for a packet will arrive in the RTT, after the full window of packets had passed through the network. Thus, throughput can be expressed as:

$$L = \frac{W}{RTT} = \frac{W}{T} \qquad 3\text{-}60$$

where W is finally the number of packets in the single server queue model, and T total waiting time regarding 3-4 and 3-3. 3-60 can be rewritten as:

$$L = \frac{E(N)}{E(T)} \qquad 3\text{-}61$$

Using Little formula $E(N) = \lambda \cdot E(T)$ for all queuing models, we have:

$$L = \frac{\lambda \cdot E(T)}{E(T)} = \lambda \qquad 3\text{-}62$$

or the useful traffic is equal to the offered one in the case of the network utilization lower than maximal, when network operates without losses. We can write:

$$P = \frac{L}{T} = \frac{W/T}{T} = \frac{W}{T^2} = \frac{\lambda T}{T^2} = \frac{\lambda}{T} \qquad 3\text{-}63$$

and calculate it directly from the results in Table 3.4. The power curves for a range of queuing systems are plotted on the Figure 3.9:

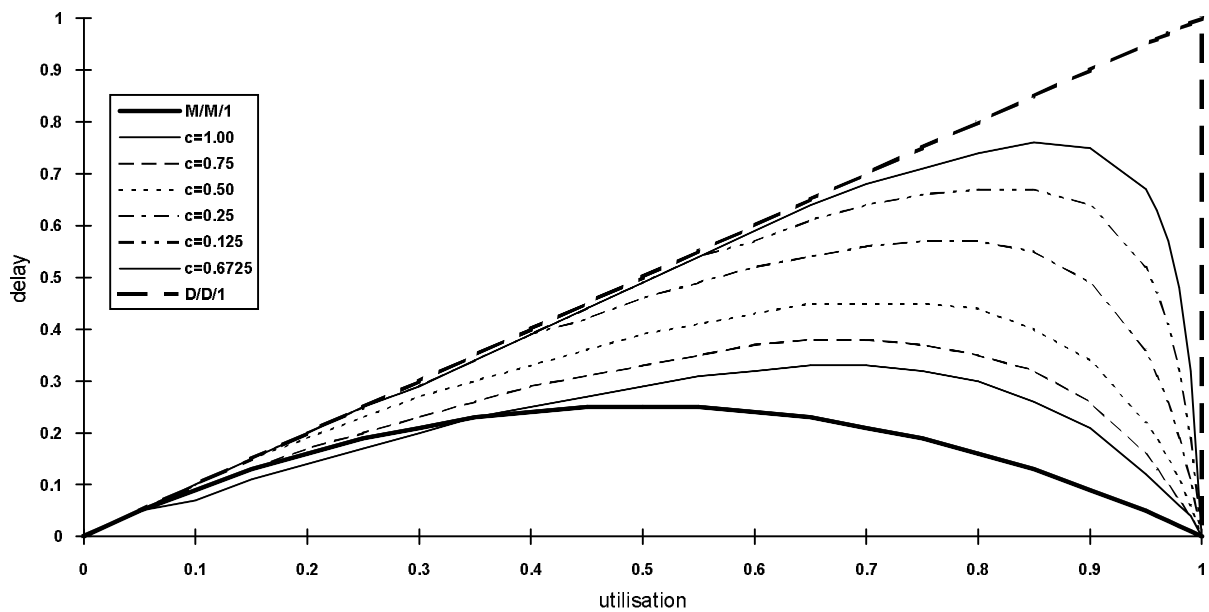

Figure 3.9. - Power for range of queuing systems

Obviously, lowering the variance the power increases, reaching its maximum for higher load of the network. Finally, the curve for D/D/1 system is approached (dashed thick). The power of M/M/1 system is shown for reference (thick). The inconsistency on the left side of graph results from the low precision of numeric data mentioned above.

One can notice that the maximum of power for $c = 0{,}25$ is obtained when $\rho = 0{,}8$, confirming earlier conclusion that for a relatively big variance the network utilization can be high. The accuracy of this result can be discussed estimating how appropriate was to use the uniform distributions in place of general ones.

Suppose we have data network with users sending their packets at (optimal) constant rates, changed only when their adaptation to the available capacity of the network is needed. The distribution of interarrival times in such a network should have low variance equal to the sum of individual flow variances (the property of normal distribution), and mean equal to the sum of means of the individual flow means. It is possible to achieve a low variance of arrival process, and one can conclude that the use of uniform distribution for arrival process modeling is justified.

The distribution of service times depends on the distribution of packet lengths, see 3-5. The packet lengths depend on network services and customer structure, and it is not easy to control them. It was observed (Internet) that packet lengths tend to concentrate around 50 octets (telnet), and in the range between 500 and 1500 (FTP, MAIL, WWW). The probability of other packet lengths is low. From the network node point of view, short packets load the processor in the same way as long ones do. Short packets place less burden on the node memory and the channel itself. Because of the big difference in length (50:1000), even when the number of short packets is the same as the number of long ones, the majority of traffic is carried on using long packets. This can verify the usage of uniform distribution as the packet length model, with mean around 1000 octets, and with variance that define borders between 500 and 1500 octets.

The more precise results could be achieved by more sophisticated models of arrival and service processes, i.e. modeling individual groups of customers separately. This can be useful when measurement data from the real network are used to make projections of future demands. In the case of flow control analysis, concerning the simplicity of single queue model, the gained results are acceptable.

The maximum network power vs. variance characteristics is shown on Figure 3.10. 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. This curve can be approximated with a straight line through the points (0,5;2) and (1;1).

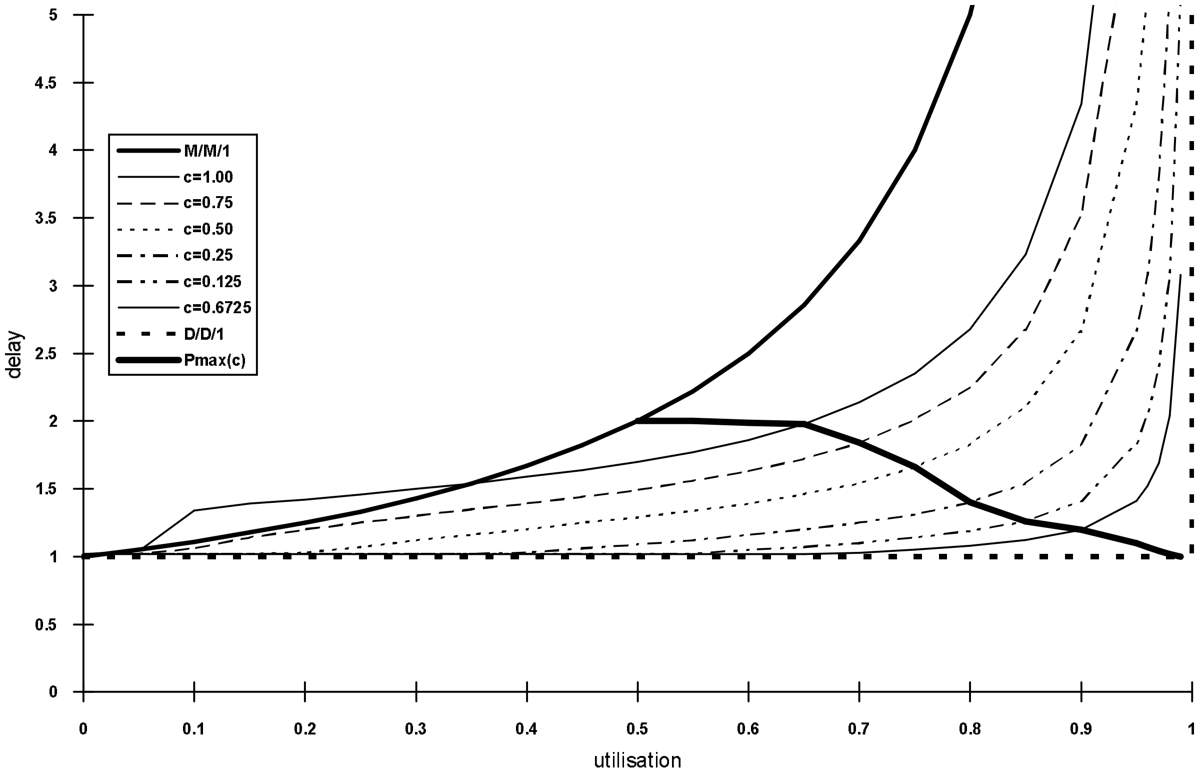

Figure 3.10. - Power vs. variance for range of queuing systems

Two strategies for the achievement of the optimal working point can be defined, Figure 3.11:

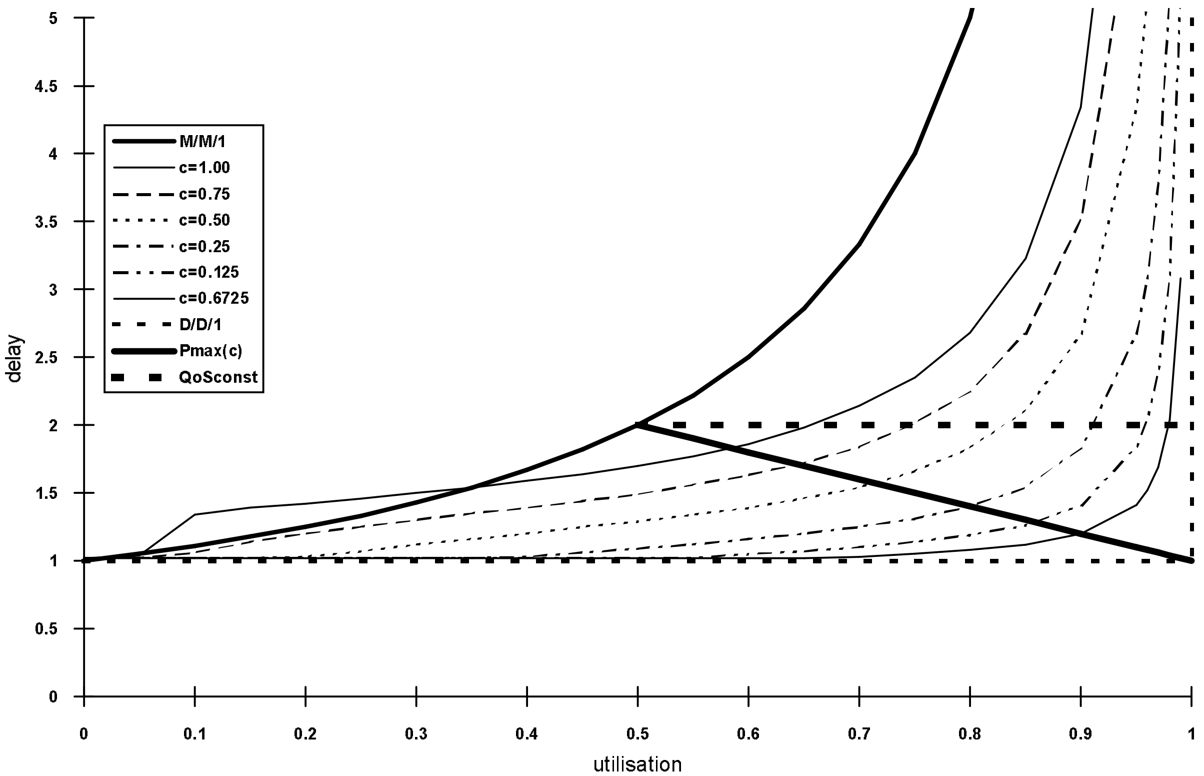

Figure 3.11. - Optimal working point strategies

In the first case (thick), the maximum power should be achieved. Lowering the variance of both input and output processes, the network utilization rises slowly, together with the improvement of QoS. In the second case (dashed thick) the declared QoS is maintained and the network utilization rises fast with the variance decrease. The approximate numeric data are listed in Table 3.5.

	Power maximized		QoS constant	
c:	utilization	delay:	utilization	delay:
M/M/1	0,50	2,00	0,50	2,00
1,00	0,59	1,80	0,66	2,00
0,75	0,65	1,65	0,75	2,00
0,50	0,71	1,50	0,84	2,00
0,25	0,80	1,35	0,92	2,00
0,125	0,86	1,25	0,96	2,00
0,0625	0,90	1,20	0,98	2,00
D/D/1	1,00	1,00	1,00	2,00

Table 3.5. - Approximate numeric data for working point strategies

The clear advantage can be found in the constant QoS strategy. In both cases, the variance of interarrival and service times should be kept as low as possible.

4. CONCLUSION

The mainstream of telecommunications research concentrates on the problems of traffic integration. The emerging technology is the cell switching ATM with important advantages in the WAN field. 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 SVCs are to be used.

The congestion control, consisting of the avoidance and elimination measures, is used to obtain declared QoS along with the optimal network utilization, taking the fairness criteria into account. Implementing the node and source algorithms, the flow control is performed as the main congestion avoidance measure in the packet networks. Its goal is to keep the network in the optimal working point ruling the source packet sending rates.

The optimal working point is achieved when the number of packets in the network allows high network utilization and low packet delay, recognized as the main QoS parameter. The optimal working point depends on the stochastic characteristics of the arrival and service processes.

The network is considered a black box, and modeled with the single server queuing system. The M/M/1 system is often used because of its simplicity, actually taking role of the worst case situation because of the low utilization and high delay results.

The D/D/1 system is achieved when the fixed multiplexing is used, leading to high utilization and low delay, however with lost flexibility of statistical multiplexing. In between these two extremes, a range of various G/G/1 systems can be interpolated.

The analysis of the G/G/1 systems is difficult, because results can be obtained using numerical methods only when actual distributions are known. In this work, the substitution of general with uniform distribution is used, to analyze the dependence of utilization-delay curves on the variance changes. Uniform distributions are acceptable because the packet interarrival times can be controlled with rate flow control mechanisms, while packet lengths tend to concentrate in the range between 500 and 1500 octets (Internet), or can be fixed (ATM).

The range of the utilization-delay curves is obtained, which conform with the expected result that lowering both the arrival and service variances, a system approaches the ideal D/D/1 performance. However, the main result is the quantitative measure of that approach. It is shown that with relatively large variances, high network utilization with low delay can be obtained.

The maximal network power is often considered the optimal network working point indicator. Using this criteria, variance should be reduced significantly to achieve high network utilization. However, the delay is lowered (QoS improved) at the same time.

The second possibility is to maintain some optimal delay (QoS kept). The high network utilization is achieved without significant reduction of the variances here.

We conclude that it is possible to obtain high network utilization with high QoS if the variances of interarrival and service times distributions are kept low. While the service times depend on packet lengths, interarrival times are function of the transmitter and node flow control algorithms involved. The future research on network models and flow control mechanisms should provide low variance arrivals. This can be achieved using combined, window and rate controlled algorithms at the sources. Traffic shaping becomes significant criteria for flow control algorithms development.

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.

[12] Zhang H. and Ferrari D, "Rate-Controlled Service Disciplines", Journal of High Speed Networks: Special Issue on Quality of Service, Vol. 3 No. 4, 1994.

[13] "ATM Traffic Management Specification Version 4.0", ATM-Forum 1996.

[14] Nain P. and Towsley D. (ed.) "Performance Evaluation of Computer Systems, Lecture Notes", rev. 1995, University of Massachusetts, Department of Computer Science, Amherst 1995.

[15] Pavlic I. "Statistical theory and application", III edition, Tehnicka knjiga, Zagreb 1985.

[16] Pause Z. "Probability: Information, Stochastic Processes", Skolska knjiga, Zagreb 1978.