ABR Service in ATM Networks: Performance Comparison between Adaptive Stochastic, ERICA, BECN and BP Congestion Control Schemes with Guaranteed Minimum Cell Rate

Slobodan S. Petrovic (IEEE Student Member)
Department of Electrical Engineering J03, University of Sydney, Sydney, NSW 2006, Australia
Email: ssp@ee.usyd.edu.au Fax: +61.2. 9351.3847, Phone: +61.2.9351.2842

Abstract

Performances of four congestion *control schemes for ABR service in ATM networks: Explicit Rate Indication for Congestion Avoidance (ERICA), Backward Explicit Congestion Notification (BECN), Back Pressure (BP) and proposed Adaptive Stochastic (AS), are compared under heavy load conditions. Neither of the schemes dictates a particular switch architecture. The requirements tested include: utilization, queuing delay, queuing delay variance and queue size. A simple network topology involves an ATM switching node which is modeled as a single server queuing system fed by a number of highly bursty traffic sources. Neither packet discarding, nor packet retransmission mechanisms are included. The controllable sources' allowed cell rates (ACR) are dynamically shaped by the feedback messages from the switch.. The results of a simulation study are presented which suggest that AS scheme can provide for higher priority traffic considerably shorter queuing delay, queuing delay variance and average queue size, than the other three schemes, while retaining the same utilization.*

1 Introduction

Future high speed networks are expected to use the Asynchronous Transfer Mode (ATM) in which the information is transmitted using short fixed-size cells consisting of 48 bytes of payload and 5 bytes of header. The fixed size of the cells reduces the variance of delay making the networks suitable for integrated traffic consisting of voice, video and data. Proper traffic management helps ensure efficient and fair operation of networks in spite of constantly varying demand. The problem is especially difficult during periods of heavy load particularly if the traffic demands can not be predicted in advance as in the case of data traffic. Hence congestion control is critical in ATM networks and although only a part of the traffic management, is its most essential aspect. When two or more bursts arrive simultaneously at a node, the queue lengths may become large very fast resulting in buffer overflow. At the points where the total input rate is larger than the output link capacity, congestion becomes a problem. Several congestion control schemes for ATM networks have been developed [1] - [8].

The remainder of the paper is organized as follows: in Section 2 the assumed network structure and how it operates is described; in Section 3 the design of the compared control algorithms is outlined; in Section 4 the simulation model and obtained results are reported and in Section 5 conclusion remarks are made.

2 Network configuration

It is assumed that two types of traffic share the ATM network resources. The first type consists of Variable Bit Rate (VBR) traffic which is delay sensitive i.e. which has stringent delay requirements and is not subjected to any control and hence is uncontrollable. The second type of traffic requires good loss performance, but agrees to have its bit rate controlled when necessary, thus accepting the possibility of longer mean delays during periods of high demand, i.e. cell loss is avoided at the expense of delay and hence is controllable. All delay insensitive Available Bit Rate (ABR) traffic can join this category which is designed for normal data traffic such as file transfer and email. In these experiments delay sensitive traffic was given a higher priority. The ACRs of the ABR traffic sources take the values being conveyed by the feedback messages which are carried from the switch to the sources' terminals through backward Resource Management (RM) cells. Each source could also be a switch acting as a virtual source and generating RM cells periodically as if it was a source. The switch

acts as a virtual destination and returns all RM cells received from the source as if the switch was the destination. In this study RM cells are "in-rate" since they are counted toward the ACR, what means that the total rate of data and RM cells should not exceed the ACR of the source. During idle periods a source does not generate RM cells. A RM cell is always sent as the first cell after an idle period for the source. In the compared schemes each individual lower priority (controllable) traffic virtual connection (VC) is guaranteed by the network only a minimum bandwidth of a minimum cell rate (MCR). However, it is allowed to dynamically share the available resources left over from the higher priority (uncontrollable) traffic i.e. the lower priority traffic grabs whatever remains of the network resources (bandwidth and buffer space) after the higher priority traffic is served. Thus the ABR capacity is not fixed, but varies according to the VBR load. Switch complexity does not depend on the number of VCs since there is no per VC queuing. In the switch, output buffering is used, which means that the queuing happens at the output port. The cells of both traffic classes are treated here, in the switch output buffer, by equal priority and are served with FIFO discipline. The same service discipline applies also to all sources' queues. In order to support the feedback control schemes a simple connection admission control (CAC) policy ensures that the peak cell rate (PCR) of the total higher priority (VBR) traffic is within the ATM switch server capacity in every moment and that the long term mean rate of all the traffic (higher and lower priority together) is within the ATM switch server capacity. All compared schemes operate at the ATM switch output queuing point.

3 Description of Adaptive Stochastic (AS) Scheme

The traffic flowing from the sources to the ATM switch is assumed to be a process which takes place in random environment. The design of the control algorithm has been based upon input-output model of the process which allows analysis over the transfer function and realization of the control algorithm with minimal a priori knowledge about the dynamics and parameters of the process. The transfer function is defined as

$y(k)u^{-1}(k) = B(z^{-1})A^{-1}(z^{-1})$

where z^{-1} is delay operator meaning

$z^{-1}y(k) = y(k - 1), \quad z^{-1}u(k) = u(k - 1)$

and integer k is the control interval's time index. Since the process is modeled here as the first order linear dynamical system

$A(z^{-1}) = 1 + a(k)z^{-1}, \quad B(z^{-1}) = b(k)z^{-1},$

it is described by the stochastic difference equation

$y(k+1) = - a(k)y(k) + b(k)u(k).$

y(k) is the output equal to instantaneous queue length minus the reference level (threshold). The scalar u(k) is the control input generated by the algorithm. It is a function of the queue size measurement and is used for identification of parameters and for attaining the desired system response. u(k) determines the ACRs of the ABR sources with the goal to maintain the ATM switch output buffer occupancy as close as possible to a certain reference level. u(k) is partitioned among active ABR sources accordingly. ABR sources seen by the switch as inactive are allocated MCR only. An active ABR source's ACR is then formed as the sum of the source's long term mean cell rate and its share of u(k). In proportion to its absolute value, a negative u(k) reduces and a positive u(k) increases the ABR sources' ACRs. Hence, the ACRs fluctuations are dictated by u(k) but can not go below the MCR (agreed during the connection setup) or above the link rate. The unknown parameters are a(k) and b(k). Their actual values are random and time varying. The uncertainty in b(k) corresponds to uncertainty in the system gain while the uncertainty in a(k) corresponds to uncertainty in the system time constant. Not knowing a(k) exactly is equivalent to not knowing the pole location of the system function. Since the system is operating in a stochastic environment, both y(k) and u(k) are random sequences. The objective of the problem is to find in each control interval the optimal control input that minimizes the scalar real-valued cost functional i.e. the performance criterion

$J(u) = E\{qy^2(k+1) + ru^2(k) + qy^2(k)\} \qquad k = 0, 1, ...$

where E denotes expected value. q and r are positive scalars (weighting factors). As the only information about the process comes from the observation of the queue size, since the queue size is the only directly measurable quantity at the switch, the optimal control policy can be interpreted as follows: at the end of the k-th control interval an optimal control input $u^o(k|k)$ (superscript o stands for optimal) is generated, based on the queue length measurement. k on the left side of | denotes the control input's time index and k on the right side of | means that the measurement of the queue length was taken during the control interval k. The next control input $u^o(k+1|k+1)$ is again computed, but based on the information about the queue size available at the control interval k + 1. Thus, the optimal deterministic control shall be recomputed after new information becomes available at each control interval. The switch gives only one feedback value in a single control interval which is equal to the round-trip delay because there is little point in sending further control feedbacks to the sources until previous feedback has had time to take effect. The overall adaptive controller can be viewed as combined of

an identifier and a zero-memory controller in the feedback loop. The former is the learning device where the estimates of the parameters are generated in real time and can be designed independently of the controller objectives. The later computes the on-line optimal control for a deterministic system and is parameter adaptive. Accordingly, as this adaptive control problem involves nonlinear parameter estimation, the extended Kalman filter is used to generate the approximate conditional means and the associated error covariances of the unknown parameters. This approach led to the design of a simple proportional controller that has the desired adaptive properties. The closed loop feedback optimal control is given by

$u^o(k|k) = -g(k)y(k)$ $\quad k = 0, 1, ...$

The adaptive feedback gain g(k) depends not only on parameters a(k) and b(k) estimates ($a^{est}(k|k)$ and $b^{est}(k|k)$ resp.), but also on conditional error covariances

$e_a(k|k) = a^{est}(k|k) - a(k)$ and $e_b(k|k) = b^{est}(k|k) - b(k)$,

i.e. it is modulated by the current parameters uncertainty (vectors and matrices are denoted in bold letters, superscript T stands for 'transposed')

$$g(k) = -\{[R(k|k) + \mathbf{b}^T(k|k)\mathbf{Q}\mathbf{b}(k|k)]^{-1}\mathbf{b}^T(k|k)\mathbf{Q}\mathbf{\Phi}(k|k) + R^{-1}(k|k)\mathbf{d}^T(k+1|k)\}[1, 0, 0]^T$$

where $\mathbf{b}^T(k|k) = [b^{est}(k|k), S_{bb}(k|k), S_{ab}(k|k)]$,

$\mathbf{d}^T(k+1|k) = [qS_{ab}(k|k), qa^{est}(k|k), 0]$,

$\mathbf{\Phi}(k|k) = \mathbf{F}(k|k) - \mathbf{b}(k|k)R^{-1}(k|k)\mathbf{d}^T(k+1|k)$,

$$\mathbf{F}(k|k) = \begin{bmatrix} a^{est}(k|k) & 0 & 0 \\ S_{ab}(k|k) & a^{est}(k|k) & 0 \\ S_{aa}(k|k) & 0 & a^{est}(k|k) \end{bmatrix},$$

$S_{aa}(k|k) = e_a^2(k|k)$, $S_{bb}(k|k) = e_b^2(k|k)$,

$S_{ab}(k|k) = S_{ba}(k|k) = e_a(k|k)e_b(k|k)$,

$R(k|k) = r + qS_{bb}(k|k)$, $\quad \mathbf{Q} = \begin{bmatrix} q & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

The parameters estimates are generated via the equations

$\mathbf{p}^{est}(i+1|i+1) = \mathbf{p}^{est}(i|i) + \mathbf{G}(i+1)[y(i+1) - \mathbf{c}^T(i)\mathbf{p}^{est}(i|i)]$,

$\mathbf{p}^{est,T}(i|i) = [a^{est}(i|i), b^{est}(i|i)]$, $\quad \mathbf{c}^T(i) = [-y(i), u(i)]$.

The filter gain matrix $\mathbf{G}(i+1)$ satisfies the relation

$\mathbf{G}(i+1) = \mathbf{S}(i|i)\mathbf{c}(i)[\mathbf{c}^T(i)\mathbf{S}(i|i)\mathbf{c}(i)]^{-1}$ $\quad i = 0, 1, ... , k -1$

$\mathbf{S}(i+1|i+1) = [\mathbf{S}(i|i) - \mathbf{G}(i+1)\mathbf{c}^T(i)\mathbf{S}(i|i)] / ff$,

$$\mathbf{S}(i|i) = \begin{bmatrix} S_{aa}(i|i) & S_{ab}(i|i) \\ S_{ba}(i|i) & S_{bb}(i|i) \end{bmatrix},$$

with the initial estimates given by

$\mathbf{p}^{est,T}(0|0) = [a_0 , b_0]$,

$$\mathbf{S}(0|0) = \begin{bmatrix} S_{a0} & 0 \\ 0 & S_{b0} \end{bmatrix}.$$

The AS control algorithm expresses its dual nature by simultaneously performing two functions: perpetual parameters estimation i.e. "learning" and process control. The forgetting factor 'ff' allows the controller's capacity for adaptation on changes in the system parameters to be preserved, by preventing error covariance matrices to become approximately equal to zero. Estimating of parameters could continue only if the matrices are different from zero, i.e. if some level of parameters uncertainty remained. Owing to the forgetting factor, the identification of parameters even if they for some time remain constant, will never be completed. Fortunately, the exact identification of the unknown parameters is not necessary for good control. Through the parameters uncertainty the controller expresses its "cautiousness". The remaining schemes are only briefly outlined.

BECN scheme: Initially the sources transmit at their peak cell rates (PCR). When the queue length in an ATM switch output buffer exceeds a threshold, the feedback controller will send backward RM cells to the ABR sources currently submitting traffic to ask them to reduce their cell transmission rate to one half of the current rate. Successive RM cells will cause a transmitter to reduce its cell transmission rate to 1/2, 1/4, 1/8, 1/16, 1/32, 1/64 of its PCR after which a further RM cell will cause it to transmit at MCR. If no RM cells are received within a recovery time period, the current transmission rate for that source will be doubled, i.e. restored to the previous level once each recovery time period, until it reaches its original PCR.

BP scheme: Each source can emit its head-of-line cell until the queue length in a switch exceeds a predefined threshold A, when the switch will send a back pressure (choking) signal to the source terminals and ask the sources to stop transmission. When the queue length decreases to a threshold B (B<=A), a congestion release signal will be sent to the sources and ask the sources to resume transmission.

ERICA scheme: See [8].

4 Simulation model and numerical results

All four congestion control algorithms were tested and compared on the same network simulation model (Fig. 1), with the same MCR. AS scheme's, BECN scheme's and BP scheme's control intervals as well as ERICA scheme's averaging interval were equal to the round trip delay, i.e. the control rate was matched to the feedback rate. Each source S generates cells that are queued for transmission by the transmitter T. Each of the transmitters T independently removes cells from its source queue and transmits them to the receiver R. R is a point of congestion somewhere in the network, e.g. an

output buffer on an ATM switch. There is a transmission delay between each transmitter transmitting a cell and the cell being received at the receiver. The transmission delay represents the combination of the propagation delay and the switching delay from the source to the point of congestion. For simplicity all the sources are equidistant to the ATM switch and hence the propagation delay is the same between the ATM switch and all the sources. The performance of the algorithms was examined over source - switch distances (link lengths): 1 km, 10 km, 100 km and 1000 km, with corresponding control/averaging intervals 10 μsec, 100 μsec, 1 msec and 10 msec resp. (1 km implies 5 μsec propagation delay). Large buffers are assumed and the effect of cell losses is not modeled. The terminals are connected to the switch through links running at 155 Mbps (365566 cells/s). The switch receives traffic from terminals that are attached to its input ports. All input ports have traffic flowing to a common output port operating at 155 Mbps. The input to the ATM switch is a superposition of streams from 18 (12 ABR and 6 VBR) different ON/OFF sources, also referred to as Interrupted Bernoulli Process

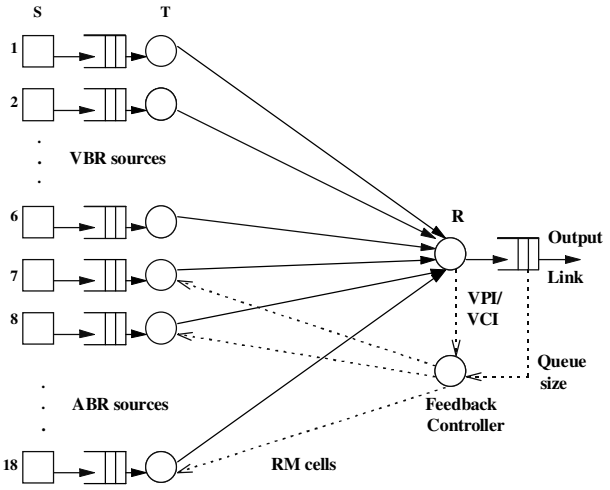

Figure 1: The simulation model

(IBP), which alternate active periods, transmitting at their PCR and idle periods, transmitting at zero bit rate. Both periods of burst and silence are exponentially distributed with constant average lengths during a simulation. Each burst in average contained 100 cells. The switch server rate was assumed to have a fixed ratio with respect to the long term mean arrival rate. The switch target utilization parameter was set at 0.95 where ABR traffic sources were given 60% and VBR traffic sources 35% bandwidth. In order to have more than a million cells generated, a simulation run time was 3 seconds. The higher priority (VBR) traffic was generated by the sources 1, 2, 3, 4, 5, 6 with PCR 50 Mbps, 40 Mbps, 30 Mbps, 20 Mbps, 10 Mbps, 5 Mbps resp., totally 155 Mbps, therefore equal to the switch server rate and with source utilizations 0.4, 0.5, 0.2, 0.2, 0.3, 0.25 resp.. The lower priority (ABR) traffic was generated by the sources 7, 8, ..., 18 with PCR 155 Mbps each, what means that each ABR traffic source was capable of filling the entire server capacity of the switch. The ABR traffic sources' utilization was 0.10 for sources 7 and 13, 0.08 for sources 8 and 14, 0.05 for sources 9 and 15, 0.04 for sources 10 and 16, 0.02 for sources 11 and 17, 0.01 for sources 12 and 18. The aggregate peak rate of the total traffic thus was 2015 Mbps. The purpose of the experiment was to examine how successfully the four compared schemes can cope with such stressful conditions when arrival rate to the switch output buffer can exceed 13 times the link rate/server capacity, while preserving the design goals. The output buffer threshold value used in AS and BECN schemes was 200 cells. A simpler form of BP scheme with one threshold only was used, i.e. B = A = 200 cells. A MCR value used was PCR/100 (1% link bandwidth) - in practice, it is desirable to allow some minimum capacity for processing aggregate ABR traffic when there is contention. Also inactive ABR traffic sources were allocated MCR , for two reasons mainly: in order to drain out their queues rapidly and because of the transmission delay the sources which may seem as inactive, could be in fact, active. As all ABR traffic sources had the same PCR = 155 Mbps, in AS scheme experiments each of the active ABR traffic sources received periodically the same amount of the feedback equal to u(k)/(number of currently active ABR traffic sources). In the case of different PCRs, the formula above could be easily adjusted . The forgetting factor equal to 0.96 was used in examining AS scheme; initial estimates were $S_{a0} = S_{b0} = 1$ and $a_0 = b_0 = 0.1$.

Performance evaluation showed that compared schemes achieved almost the same average utilization (Fig. 8) excepting ERICA scheme's utilization on 1000 km link which was about 1.4% higher, due to higher ABR utilization (Fig. 7). The number of VBR cells received (Fig. 6) for all four schemes was almost identical. As any MCR > 0 contradicts the initial stipulation that VBR traffic can take the entire server capacity whenever it needs it, the MCR employed in these experiments brought about conflicting demands by VBR and ABR traffic classes, as VBR traffic class, albeit able to fill the entire server capacity (if all VBR traffic sources are simultaneously transmitting), was guaranteed only 88% of it. Obviously, although the MCR was guaranteed to all 12 ABR traffic sources, due to their bursty nature they did not use it all the time and not all of them at the same time what allowed the VBR traffic to

have enough bandwidth to flow smoothly and unabatedly. However, the utilization metric gives only a partial picture of efficiency. Ideally, an efficient scheme should also control queuing delays within acceptable limits. With regard to the mean queuing delay (Fig. 2) and mean queue length (Fig. 5) in the switch output buffer, AS scheme performed visibly better than the other schemes, retaining the two variables practically unchanged for all four link lengths and many times shorter than ERICA scheme's (particularly on 1 km and 10 km links), BP scheme's and BECN scheme's (on 100 km and 1000 km links). In regard to the maximum queue length in the switch output buffer (Fig. 4) the situation was similar to that of the mean queue length, with main difference that AS scheme's maximum queue length with increasing link lengths was also increasing, so that on 100 km and 1000 km links was almost equal to ERICA scheme's values. Particularly interesting results were obtained for the queuing delay variance in the switch output buffer (Fig. 3) where AS scheme performed convincingly better than other three schemes. AS scheme's delay variance was smaller than ERICA scheme's on 1 km link, 10 km link, 100 km link and 1000 km link 878 times, 491 times, 2.3 times and 7 times resp.; smaller than BECN scheme's on 10 km link, 100 km link and 1000 km link 5 times, 26 times and 298 times resp.; smaller than BP scheme's on 10 km link, 100 km link and 1000 km link 4.4 times, 27 times and 307 times resp.. Only on 1 km link BECN scheme's and BP scheme's delay variance was 1.6 times and 1.5 times smaller, resp.. Another interesting result is that BECN and BP schemes in respect to all tested requirements performed almost identically. As the system was in a state of high variation and had no steady state, the aim was to keep the average utilization high, the average queue size small and prevent queuing delays from becoming unbounded. Under such stressful conditions the schemes performed differently. In respect to the maximum queue length, mean queue length, mean queuing delay and queuing delay variance in the switch output buffer AS scheme outperformed other schemes. BECN and BP schemes performed better than ERICA scheme on 1 km and 10 km links but worse on 100 km and 1000 km links. In respect to the average utilization all schemes attained very close results, excepting ERICA scheme's better performance on 1000 km link.

5 Conclusion

ERICA, BECN, BP and proposed AS congestion control schemes for ABR service in ATM networks are compared in the presence of two traffic classes - ABR and VBR, under heavy load conditions exceeding 0.95. The discussion relates to the traffic management at the segment of an ATM network between terminals generating traffic and the switch. By ensuring much shorter queues in the ATM switch output buffer, AS scheme provided for delay sensitive VBR traffic class sharply reduced queuing delay variance and mean queuing delay while attaining practically the same link utilization.

References

[1] Kalampoukas, A. Varma, K.K. Ramakrishnan, "An efficient rate allocation algorithm for ATM networks providing max-min fairness," In proceedings of the 6th IFIP International Conference on High Performance Networking, September 1995.

[2] Siu and T. Tzeng, " Intelligent congestion control for ABR service in ATM networks", Computer Communication Review, Volume 24, No. 5, pp. 81-106, October 1995.

[3] Tsang and W. Wong, "A new rate-based switch algorithm for ABR traffic to achieve max-min fairness with analytical approximation and delay adjustment", In Proceedings of the 15th IEEE INFOCOMM, pp. 1174-1181, March 1996.

[4] Charny, D. D. Clark, R Jain, " Congestion Control with Explicit rate Indication", In Proceedings of ICC'95, June 1995.

[5] Arulambalam, X. Chen and N. Ansari, "Allocating Fair Rates for Available Bit Rate Service in ATM Networks", IEEE Communications Magazine, 34(11):92-100, November 1996.

[6] Jain, S. Kalyanaraman and R. Viswanathan, "The OSU Scheme for Congestion Avoidance in ATM Networks: Lessons Learnt and Extensions", Performance Evaluation Journal, Vol. 31/1-2, December 1997.

[7] Kalyanaraman, "Traffic Management for the Available Bit Rate (ABR) Service in Asynchronous Transfer Mode (ATM) networks" Ph. D. Dissertation, Dept. of Computer and Information Sciences, The Ohio State University, August 1997.

[8] R Jain, S. Kalyanaraman, R. Goyal, R. Viswanathan and S. Fahmy, "ERICA: Explicit Rate Indication for Congestion Avoidance in ATM Networks", U.S. Patent Application (S/N 08/683,871), filed July 19, 1996.

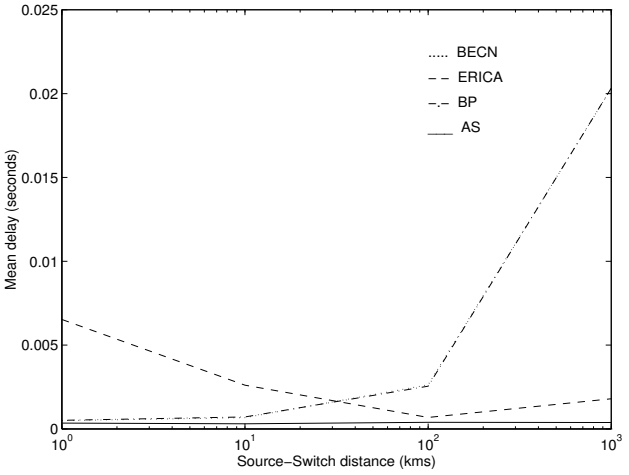

Figure 2: Mean queuing delay in the ATM switch output buffer

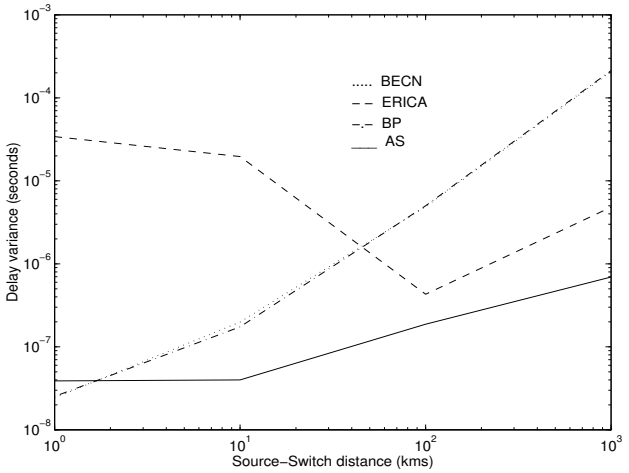

Figure 3: Queuing delay variance in the ATM switch output buffer

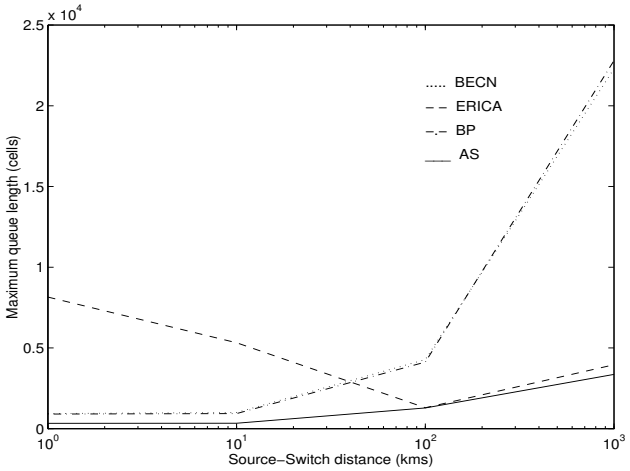

Figure 4: Maximum queue length in the ATM switch output buffer

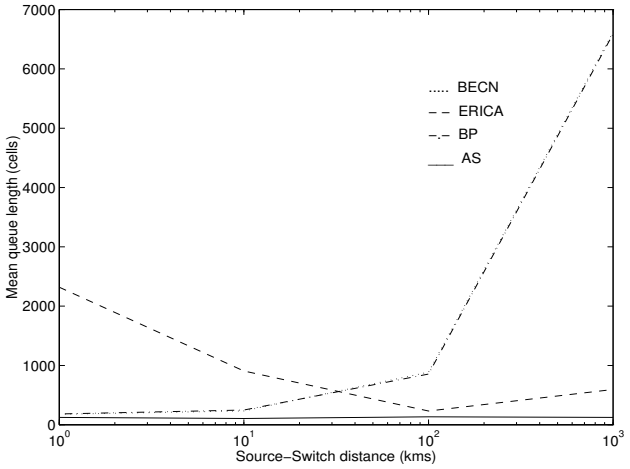

Figure 5: Mean queue length in the ATM switch output buffer

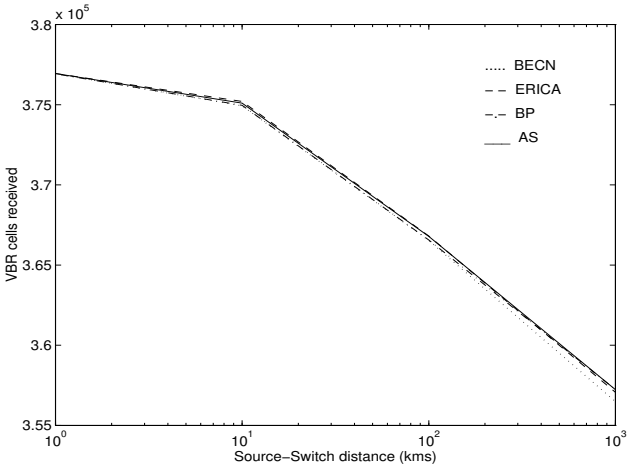

Figure 6: Number of VBR traffic cells received

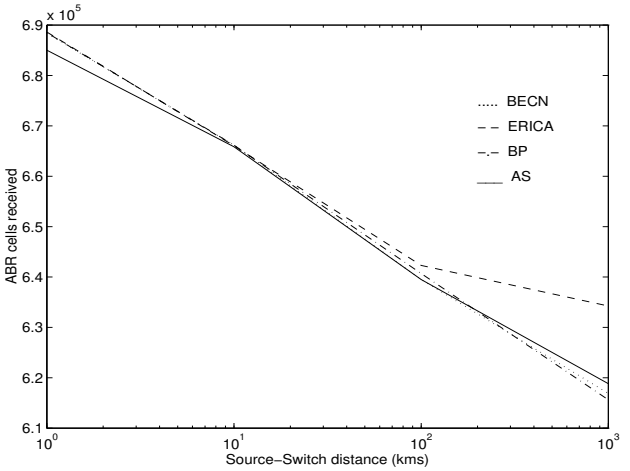

Figure 7: Number of ABR traffic cells received

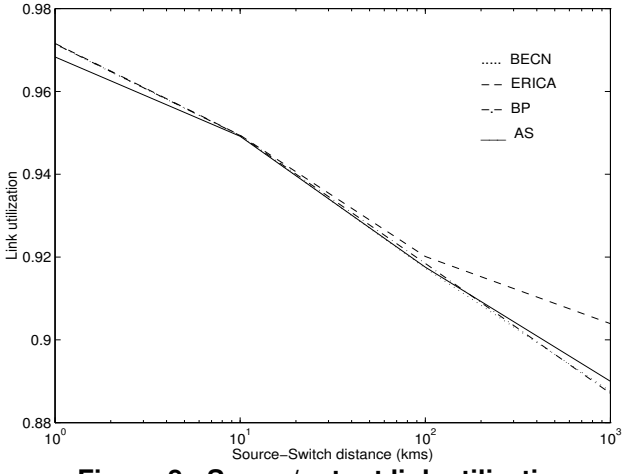

Figure 8: Server/output link utilization