Multi-Constrained Multi-Path Routing with Integrated Traffic Control Laws

X.Agnes Kala Rani
Lecturer
Veltech Engineering College
Chennai, India
agneskalarani@yahoo.co.in

T.Balaji Vishwanath
Veltech Engineering College
Chennai, India
tbalajivishwanath@gmail.com

***Abstract*— A large body of networking research has addressed the question of predicting steady state quality of service measures such as delay distributions and loss probabilities for various purposes such as network design (i.e., dimensioning links, buffers, switches, etc.) and admission control (making a decision whether or not to admit a connection or set of connections into the network).**

In a high-speed network environment the ratio of the propagation delay to the transmission delay can be very high, and by the time the control packet reaches the source, millions of packets may already be in transit and contributing to significant losses and further congestion in the network. To address this problem, the thesis currently investigating techniques of predicting the network traffic sufficiently far into the future (more than the round trip propagation delay) and controlling the non-real-time traffic based on the predicted value.

The proposed idea is that a predictive/control based solution must explicitly take into account the queuing behavior at the node(s) of interest. This approach is directly in contrast with traditional predictive flow control schemes which have tried to minimize the prediction error of the source by largely ignoring the underlying queuing structure because it was too difficult to handle. The scheme of the thesis is try to minimize the congestion in the queue rather than focus on minimizing an adhoc criterion such as the mean squared error of the source. The basic system is based on the predicted value of the aggregate traffic at the bottleneck node, a control vector is generated for the non-real-time traffic sources.

I. INTRODUCTION

Even if the design and admission control mechanism are good, the network may go into periods of congestion due to transient oscillations in the network traffic[2]. Fortunately, in most networks there is a significant fraction of the traffic which is not very delay sensitive, and hence its input rate to the network can be controlled. In real-systems this traffic corresponds to “data traffic” which is generated via email, file transfers, images, etc.

Such traffic is generally referred to as non-real-time traffic versus real-time traffic such as voice and video, which have more stringent delay requirements in which if packets arrive after a certain deadline they are assumed to be lost.

This thesis summarizes the technical basis for predictive flow control, including some fundamental advantages of credit which could be adopted by other mechanisms. It is thereby to speed the evolution of ATM flow control, and minimize the risk of standardizing inadequate solutions. This thesis avoids short-term pragmatic issues, such as migration paths and interoperability, noting that flow control mechanisms adopted now may be in use long after such issues are forgotten.

A. *Flow control problem*

Any data network has bottlenecks, points where more data can arrive than the network can carry. These points are often in switches with multiple ports, congestion arises when data, destined for a single output port, arrives at many inputs. The universal short-term solution involves buffer memory in which a switch can temporarily queue data directed at overloaded outputs. In the longer term, no amount of buffering is sufficient instead, each source of traffic flowing through a bottleneck must be persuaded to send no more than its fair share of the bottleneck’s capacity.

This is fundamentally a feedback control problem, and many control ideas and principles apply. Each network switch collects information about congestion, and informs, directly or indirectly, the sources of data. This feedback is usually based on the amount of buffer space available or in use in the switch. The sources act to control how much data they send. This control loop has a delay of at least twice the propagation delay between the switch and control point.

Control systems should seek to minimize this delay, since switches will need to buffer any data that arrives after they signal the congestion status but before the end of the delay. A variety of technical goals, some of them conflicting are desirable for any flow control mechanism. Data should rarely, if ever, be discarded due to exhaustion of switch buffer memory. Such data may have to be retransmitted after a possibly lengthy time-out period, further contributing to network congestion and the delay seen by the user.

978-1-4244-4711-4/09/$25.00 ©2009 IEEE

Links between switches should be used at full capacity whenever possible. For instance, if one connection sharing a link reduces the rate at which it sends, the others should increase as soon as possible. All the connections which are constrained by a bottleneck link should get fair shares of that link. The flow control mechanism should be robust, loss or delay of control messages, for instance, should not cause increased congestion. The network administrator should not have to adjust any complex parameters to achieve high performance. Finally, the flow control mechanism should have a cost commensurate with the benefits it provides.

B. Intermediary Links and Nodes of Interest

The basic system is described one where based on the predicted value of the aggregate traffic at the bottleneck node, a control vector is generated for the non-real-time traffic sources. Note that these sources could all be at different points in the network and hence the propagation delay to each of them from the bottleneck queue is different. The proposed system focused on the problem of controlling the traffic in such way as to minimize the asymptotic tail of the buffer occupancy distribution for a given constraint on the nodal utilization. The system has developed a heuristic scheme that satisfies the necessary condition for optimality. The interesting fact in this scheme is, since it explicitly takes into account the queuing behavior at the node, decreases the probability of overflow in the network by several orders of magnitude.

C. Predictive Flow Control Mechanism

Any prediction of how well a flow control scheme will work requires a model for the behavior of network traffic. A full-blown model might involve characteristics of applications and higher-level protocols. For our purposes it is enough to distinguish between smooth and bursty traffic.

A smooth traffic source offers a constant and predictable load, or only changes in time scales that are large compared to the amount of time the flow control mechanism takes to respond. Such traffic is easy to handle well, the sources can be assigned rates corresponding to fair shares of the bottleneck bandwidth with little risk that some of them will stop sending and lead to underutilized links. Switches can use a small amount of memory, since bursts in traffic intensity are rare.

Sources of smooth traffic include voice and video with fixed-rate compression.

The aggregate effect of a large number of bursty sources may also be smooth, particularly in a wide-area network where the individual sources are relatively low-bandwidth and uncorrelated. Rate-based flow control works well with smooth traffic.

Bursty traffic lacks any of the predictability of smooth traffic, as observed in some computer communications traffic. Some kinds of bursts stem from users and applications. These bursts are sporadic, and typically do not last long enough on a high-speed link to reach steady state over the link round-trip time.

D. Queuing Structure

The Predictive loss pattern (PLoP) algorithm aims at equally distributing necessary packets drops within a single queue between flows belonging to a certain group of flows with similar properties and QoS requirements (foreground traffic-FT). This is done to minimize violations of the given advance characterization of the flow's sensitivity to burst losses (drop profiles). The task of a drop profile is to translate the applications' end-to-end QoS requirements (i.e. the minimization of the conditional packet loss probability in our case) to a per-packet behavior of a queue management algorithm at nodes of interest.

E. Prediction Error Minimizer

When the queue length exceeds its threshold, a packet is selected to be dropped. After the first drop of a packet of a particular FT flow, the flow ID and the index referring to a corresponding drop probability of the profile for the next drop are recorded in the flow table. The flow ID is the [protocol ID, src addr/port, dst addr/port] tuple for IPv4. With IPv6 the flow label can be used.

When another FT packet should be dropped a drop experiment is performed. The table is checked, whether the ID of the selected packet has already been stored. If true, a random number is generated and the packet is dropped with a probability as found in the table record and the index into the profile within the flow table is up- dated. If this drop experiment does not result in an actual drop, the packet is marked as a “survivor" and the next packet matching the FT requirement is searched for in the queue. This procedure is repeated until an actual drop has taken place. If the end of the queue is reached (i.e. no adequate replacement packet for the original packet was found “force failure"), the original packet is dropped.

Given a bandwidth constraint B and a delay constraint D, the problem is to find a path that is most likely to satisfy both constraints. Specifically, the problem is to find a path r* such that for any other path p from s to t.

If the b(i,j)’s and d(i,j)’s are constants, the MP-BDCP problem reduces to the familiar bandwidth-delay constrained path problem, which can be easily solved in two steps.

(i) prune every link (i, j) for which $b(i, j) < B$, and

(ii) find the shortest path with respect to the delay parameter in the pruned graph.

Using the central limits theorem and Lagrange relaxation techniques, two complementary solutions for the delay case have been provided. These solutions are found to be highly efficient, requiring, on average, a few iterations of Dijkstra’s shortest path algorithm. The bandwidth case is rather simple, and is dealt with by transforming it into a variant of the shortest path problem. The solutions for the MP-BCP and MP-DCP problems are then combined to address the MP-BDCP problem.

MP-BDCP belongs to the class of multi-objective optimization problems, for which a solution may not even exist. To eliminate the potential conflict between the two optimization objectives, an approach is proposed in which a subset of nearly non dominated paths is computed for the given bandwidth and delay constraints.

II. EXPERIMENTAL EVALUATION

Extensive simulations are conducted to evaluate the performance and computational complexity of the aforementioned algorithmic solutions. The interest is not only to assess the good-ness of these solutions, but to also demonstrate the potential benefits of the probabilistic approach, as a means of reducing the protocol overhead at no loss in the routing performance.

In the probabilistic approach, routers are expected to maintain and advertise two parameters for each QoS measure (e.g., mean and variance for delay, minimum and maximum for available bandwidth). These parameters vary at a much slower pace than the instantaneous delay and bandwidth values. In our simulations, an assumption that these statistical parameters are computed and advertised once at the beginning of each simulation run is made. Source nodes then use the one-time advertised information to determine the most-probable path with respect to the delay constraint, the bandwidth constraint, or both. Once this path is computed, its feasibility according to the actual (instantaneous) link values (which are not available to the path selection algorithm) is checked. If the path is feasible according to the actual values, the attempt is called a ‘success’. The performance of a path selection algorithm is expressed in terms of the success rate (SR), which is the fraction of returned paths that are feasible.

To demonstrate the robustness of the probabilistic approach, contrast it with the standard threshold-based triggered approach. In the triggered approach, the instantaneous bandwidth and delay values are advertised once they exceed certain thresholds, indicated by THB and THD, respectively (for simplicity, express these thresholds in absolute terms). Consider, for example, the available bandwidth over a given link. If this bandwidth changes (e.g., following the addition of a new flow or the termination of an existing one) such that the absolute difference between the new value and the most recently advertised one exceeds THB , then a new link state advertisement (LSA) is generated and advertised throughout the domain.

The smaller the values of THB and THD, the higher are the SR of the triggered approach. But this performance gain comes at the expense of increased advertisement overhead. The algorithm treats the available state values as if they were exact. Compare the probabilistic and triggered approaches in terms of the normalized SR’s and the communications overhead. To measure the communications overhead of the triggered approach, compute the percentage of links whose bandwidth and delay values changed to the extent of triggering a state update within a given period of time. In the simulations the probabilistic approach uses the one-time advertised statistical information.

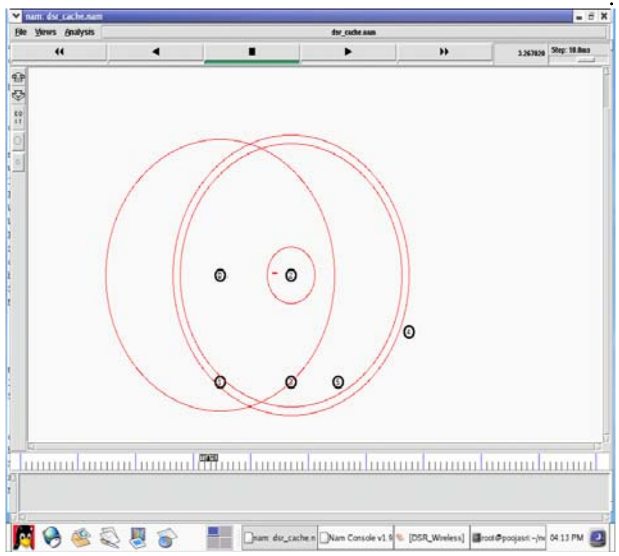

Figure 1: Network Node Establishment

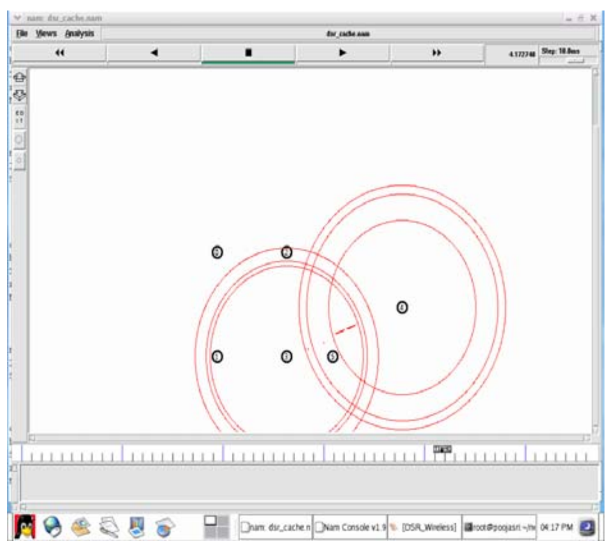

Figure 2: Data Communication Between Nodes

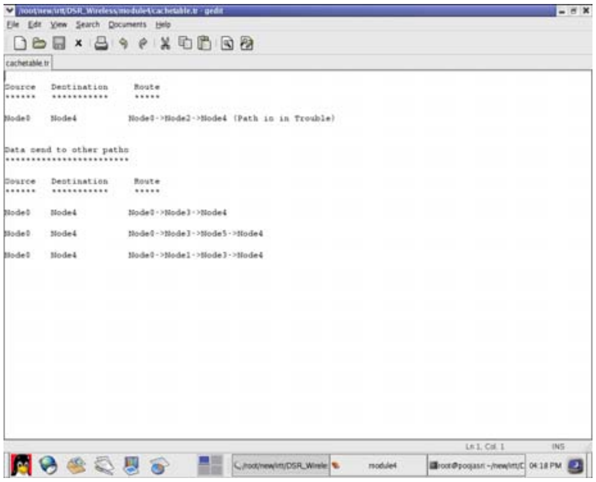

Figure 3: Path Through Which Data Sent

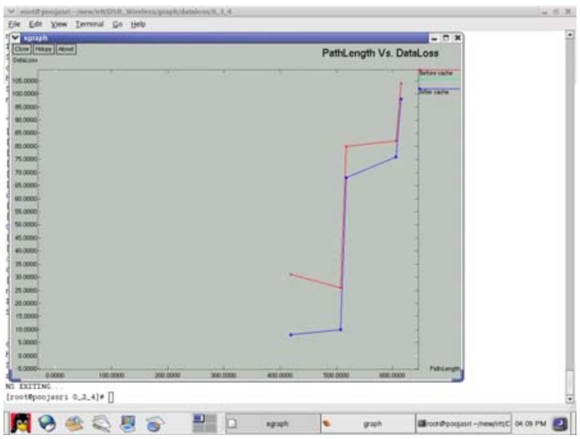

Figure 4: Graph 1: Path length Vs Data Loss

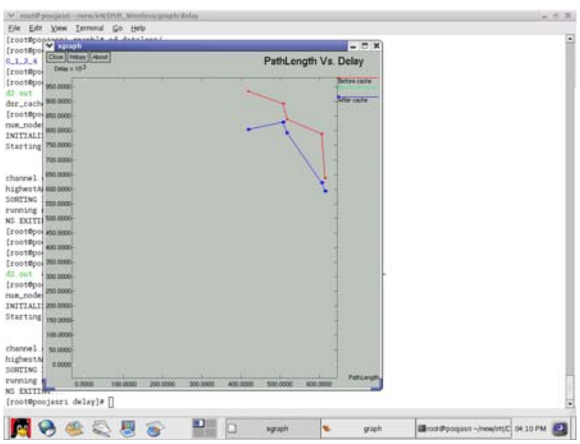

Figure 5: Graph 2: Path Length Vs Delay

III. CONCLUSION

The thesis have introduced simple metrics to describe the loss process of individual flows and presented simulation results in a multi-hop topology for a data service using queue management algorithms for burst loss control.

Two approaches of predictive flow algorithms have been analyzed, the first one uses packet marking to designate drop preference without keeping local per flow state. The second approach on the contrary operates purely local (the sender or other nodes are not involved in the scheme), but keeps state for the protected flows.

The system found that both approaches of algorithms have a significant impact on conventional traffic. It is possible to control the loss characteristics of individual flows while keeping their unconditional loss probability within a controlled bound around the value expected using conventional error feedback mechanisms.

For the given scenario algorithms using packet marking are found to be superior because a high probability for short bursts with potentially high perceptual impact can be traded against a higher probability for isolated losses as well as higher (but acceptable) probability for very long loss bursts.

REFERENCES

[1] G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. (references)

[2] S. Ramanathan, P.V. Rangan, and H. Vin, “ Frame-induced packet discarding: An efficient strategy for video networking," in Proceedings of NOSSDAV '93, 1993.

[3] L. Wolf, R. Herrtwich, and L. Delgrossi, “Filtering multimedia data in reservation based internetworks," Technical Report 43.9608, IBM European Networking Center, August 1994.

[4] [3] R. Wittmann and M. Zitterbart, “AMnet: Active multicasting network," in Proc. of Intern. Conf. on Communications (ICC'98), Atlanta, GA, June 1998.

[5] C. Perkins, O. Hodson, and V. Hardman, “A survey of packet loss recovery technique for streaming audio," IEEE Network Magazine, Sept./Oct. 1998.

[6] C. Perkins and O. Hodson, “ Options for the repair of streaming media," RFC 2354, IETF, June 1998.

[7] G. Siganos, M. Faloutsos, P. Faloutsos, and C. Faloutsos “Power-laws and the AS-level Internet topology” IEEE/ACM Transactions on Networking, 11(4):514–524, August 2003.

[8] B. Fortz and M. Thorup. “Optimizing OSPF/IS-IS weights in a changing world “ IEEE Journal on Selected Areas in Communications, pages 756–767, 2002.