Improving Bandwidth Utilization based on Deterministic Delay Bound in Connection-Oriented Networks

Peerapon Siripongwutikorn and Sujata Banerjee

University of Pittsburgh

Abstract—**Packet scheduling disciplines play an important role in providing Quality of Service (QoS) guarantees to applications traffic in high speed networks. Several scheduling disciplines have emerged in the literature that are capable of providing guarantees on various QoS parameters, such as packet delay, jitter, loss and throughput, while maintaining fairness among various connections. However, there is a trade-off between the performance and the simplicity of operation and implementation performance of the scheduling schemes. Thus the First-Come-First-Serve (FCFS) scheduling discipline has remained popular due to its overriding simplicity in spite of its poor capabilities in providing QoS guarantees. In previous work it has been shown that a practical worst case bound can be obtained for a deterministic delay guarantee in networks using the FCFS discipline, by imposing a source rate condition on each connection, that restricts the overall network utilization. In this paper, a modified source rate condition is proposed that can significantly increase the bandwidth utilization, while still maintaining the deterministic delay guarantee.**

I. Introduction and Motivation

Packet scheduling disciplines constitute a crucial component in providing Quality of Services (QoS) guarantees in high-speed integrated services networks. Typically, the more complex the scheduling discipline, the better is its performance, flexibility and the ability to maintain fairness among multiple connections. However, in practice, it is those scheduling disciplines that are simpler to implement and operate without tunable parameters that have found their way into commercial products, despite their poorer capabilities. The First-Come-First-Serve (FCFS) scheduling discipline is a prime example of such a scheme. Also, in recent years it has become apparent that most applications can make do with relatively simple guarantees, which has further bolstered the use of simpler schemes such as FCFS. With the above motivating factors, research into providing guarantees with simpler schemes such as FCFS, while maintaining high network utilization, has resurfaced with renewed interest.

The FCFS worst-case delay bound can be determined given that traffic generated by sources are bounded. The original work on FCFS worst-case delay bound computations are done by Cruz [1], [2]. In that seminal work, a traffic source is described by its (σ, ρ) leaky bucket parameters. Traffic flows in the network can be either feed-forward or non-feed-forward. However, per-hop traffic characterization inside the network is needed, and the bound is not easy to compute for non-feed-forward networks. Many attempts have been made to tighten the FCFS bound [3], [4], [5]. Nevertheless, these bounds are still not tight enough for practical use.In very recent work [6], this problem has been approached from a different viewpoint. By using a new route system parameter instead of the traditional network parameters such as the number of nodes or links, end-to-end queueing delay bounds can be easily computed without the need for per-hop traffic characterization or any limitation on traffic flows patterns.

The rest of the paper is organized as follows. In section II, the deterministic delay bounds proposed by [6], [7] and the necessary source rate condition (SRC) that must be imposed on traffic sources are reviewed. Section III discusses and illustrates how the SRC causes an inefficient bandwidth utilization. In section IV, a modification on the SRC based on a heuristic method is proposed. In addition, the maximum achievable bandwidth utilization is empirically determined in order to appreciate the extent to which the SRC restricts the bandwidth usage. Section V concludes the paper and discusses some limitations, issues, and other potential problems in this approach.

II. Background

In this section, the deterministic delay bound proposed in [6], and its improved version in [7] are reviewed. The network model under study can have an arbitrary complex topology. Both papers impose two important assumptions on the network model – homogeneous network, (all the links have the same capacity), fixed-sized packets and output buffering scheme at each node. In addition, only constant rate sources are considered. With the first two assumptions, time is divided into equal slots, each contains either zero or one packet. Both bounds are based on the concept of a route system parameter which is the number of interfering flows on a given connection path. Each flow, at an outgoing link of a node it traverses, interferes with flows that arrive at that node on a different incoming link but shares the same outgoing link. Thus the number of interfering flows depends only on a route that the flow traverses, which is why it is called a route system parameter. For a given flow, the total number of interfering flows, summed at each node it traverses along its path, is referred in [6], [7] as its *route interference number* (RIN). The concept of RIN is best explained by an illustrative example. Consider the network topology with six flows (A to F) shown in Fig. 1. As an example, flow A originates at node 0, terminates at node 9, and traverses links 0-3, 3-6, 6-8, 8-9. The RIN of flow j is the sum of the number of interfering flows of all links along flow j's path. For a given outgoing link l at node i, the number of flows interfering with flow j is denoted by $\alpha_j^{(l)}$. Let us determine the RIN of flow A. At node 0, $\alpha_A^{(0-3)}$ = 1, since flow A has only one interfering flow, B. At node 3, $\alpha_A^{(3-6)}$ = 1, since flow A also has only one interfering flow, D. At node 6, $\alpha_A^{(6-8)}$ = 3, because there are three flows - B, E, and F - joining flow A from other links. Note that flow D is not counted because it has the same incoming link as flow A. At node 8, $\alpha_A^{(8-9)}$ = 0. Thus the RIN of flow A, denoted by RIN_A, is given by

$$\mathrm{RIN}_A = \alpha_A^{(0-3)} + \alpha_A^{(3-6)} + \alpha_A^{(6-8)} = 1 + 1 + 3 + 0 = 5.$$

This research is supported in part by a National Science Foundation CAREER award under grant no. ANI-9702389

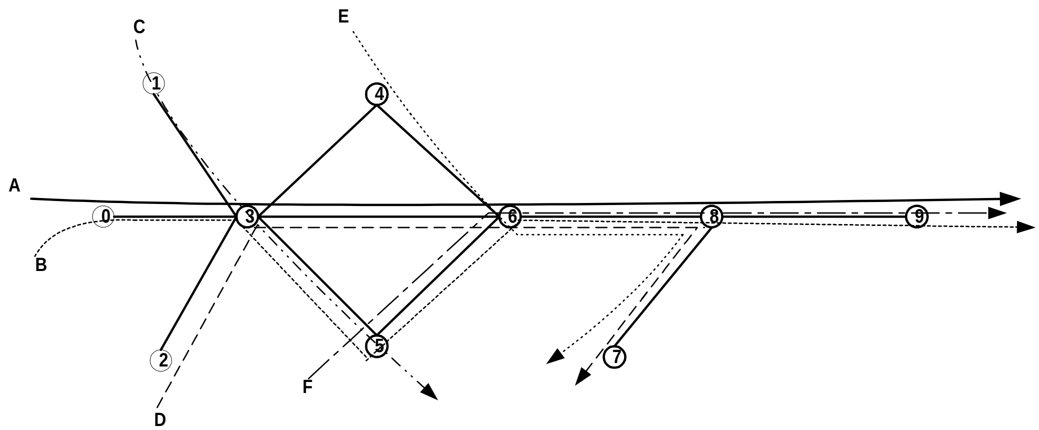


Fig. 1. Example network topology

Next we describe the source rate condition (SRC) [6], [7] necessary for traffic sources if the end-to-end delay bound is to be determined. Let τ_j denote the interpacket gap (in the unit of slots) between two consecutive packets generated at a traffic source of flow j. The SRC imposed on a traffic source is

$$\tau_j \quad \geq \mathrm{RIN}_j, \tag{1}$$

which is simply the minimum allowable interpacket gap in the unit of slot times. The source peak rate is thus inversely proportional to the inter-packet time[1] and is equal to $\frac{1}{\tau_j+1}$. From the previous example, the minimum allowable interpacket time at the source for flow A is τ_A = 5, and its source peak rate is $\frac{1}{6}$. Consider a given node i and define $\mathcal{M}_i$ as the set of all of its incoming links, link $e \in \mathcal{M}_i$, and f as one of its outgoing links. Also define N_e as the number of flows that has e and f on their path at node i, and $N_{\min}$ is the smallest among all the incoming links, i.e., $N_{\min} = \min_{\forall e \in \mathcal{M}_i} N_e$. If each flow satisfies its SRC, and the buffer size of each outgoing link f, denoted by $b(f)$, at each node is at least,

$$b(f) = \sum_{e \in \mathcal{M}_i} N_e - N_{\min}, \tag{2}$$

then the maximum end-to-end queueing delay of each flow is bounded by its RIN, and no packet is lost due to buffer overflow. We will refer to this end-to-end delay bound as the RIN bound. Thus all the traffic sources in the network must limit their peak rate to be at most $\frac{1}{\mathrm{RIN}+1}$ for the RIN bound and loss-free delivery of all flows to be satisfied.

Building on the work in [6], a tighter bound with the same SRC (as in (1)) has been developed in [7]. We will refer to this modified bound as the RIN′ bound, which is defined below. Using the same notation as defined above, let us first define

$$d(f) = \sum_{e \in \mathcal{M}_i} N_e - N_{\max}, \tag{3}$$

where $N_{\max}$ is the largest among N_e, i.e., $N_{\max} = \max_{\forall e \in \mathcal{M}_i} N_e$. For flow j traversing a set of links $\mathcal{L}_j$, if the SRC is satisfied by all the traffic sources, and a buffer size of each outgoing link f at each node is at least $d(f)$, the maximum end-to-end queueing delay of flow j is bounded by its RIN′ which is given by

$$\mathrm{RIN}'_j = \sum_{f \in \mathcal{L}_j} d(f), \tag{4}$$

and no packet is lost due to buffer overflow.

III. The Network Utilization Problem

In this section, the quality of the RIN and RIN′ bounds are evaluated by comparing them with the actual maximum delay obtained using simulations with traffic sources that satisfy the SRC. Intuitively, the performance measure commonly used in evaluating a deterministic delay bound is the ratio of the actual maximum delay over that bound, which we will refer to as the *bound tightness*. As mentioned earlier, theoretical bounds tend to be loose (much larger than the actual maximum delay) as verified by simulation in [8] because they are obtained from worst-case analysis which hardly occurs in practice. Furthermore, in case of RIN and RIN′ bounds where the source peak rate of a connection is approximately the inverse of its RIN, the bandwidth utilization becomes a problem due to the following causes:

- Since the actual maximum delay can be much lower than its corresponding theoretical bound, especially when the theoretical bound is large (as will be shown later), it follows that the source peak rate will be much lower than it should be, and so will the bandwidth utilization.
- Both RIN and RIN′ bounds are tightly coupled with the number of interfering flows, which strongly depends on the total number of flows in the network. As the number of flows increases, the bound itself will increase while the source peak rate decreases. This prevents the bandwidth utilization from increasing as the number of flows increases.

In this section, we illustrate these points by simulation. With the same network configuration and simulation setup as in [6], we repeat the simulation experiments and calculate the bandwidth utilization. The network model used in [6] is a 16-node, 4x4 grid network, with 48 unidirectional links. In each run, a total of 16 flows are generated, and the path of each flow is chosen using random routing, and then the RIN of each flow is computed. For each flow, the interpacket time at a source is either

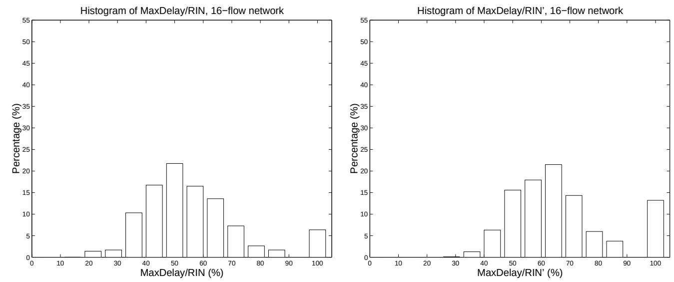
[1] Note that the packet interarrival time includes the time to transmit a packet in addition to the inter-packet gap and is equal to τ_j + 1.

RIN or RIN+1 idle slots, randomly selected with equal probability. The actual maximum end-to-end queueing delay on each flow is measured over the time period of 100,000 slots. The new set of 16 routes are then regenerated and the actual maximum end-to-end queueing delay are again measured. This procedure is repeated for 1,000 runs and the results from all runs are combined together.

A. Bound Tightness and Bandwidth Utilization

For each flow, we can calculate the ratio of the actual maximum delay experienced by the flow to the worst case theoretical bound, expressed as a percentage. This is defined as θ, which is a measure of the bound tightness. A flow having θ equal to 100% means that its actual maximum queueing delay equals its worst-case bound.

Fig. 2(a) plots the histograms of θ obtained from the simulation runs using RIN and RIN' bound (Recall that both RIN and RIN' bounds use the same SRC: $\tau_j \geq \mathrm{RIN}_j$). The RIN' bound apparently is tighter since its histogram is more skewed to the right. An average θ in case of RIN bound is 56.1% while that of RIN' bound is 67.1%. Because, for a given flow, the RIN bound is never better than the RIN' bound, we will consider only RIN' bound for the rest of the paper.



(a) Using RIN Bound (b) Using RIN' Bound

Fig. 2. The percentage of the actual maximum end-to-end queueing delay over the worst-case bound in a 4x4 grid, 16-flow network

We also examine the correlation between the number of hops (H) that a flow traverses, the bound (B), and the bound tightness (θ) by calculating correlation coefficients[2] $\rho_{H\theta}$ and $\rho_{B\theta}$ from the data points. For the RIN' bound, $\rho_{H\theta}$ is -0.24, and $\rho_{B\theta}$ is -0.366. The negative values of $\rho_{H\theta}$ and $\rho_{B\theta}$ imply that as the number of hops or the delay bound of a flow increases, an actual maximum delay is likely to be further lower than its corresponding RIN' bound, i.e., the bound becomes looser.

Consider the bandwidth utilization (ρ). Since all traffic sources generate packets separated by either RIN or RIN+1 with equal probability, i.e., the packet interarrival time can be either $\mathrm{RIN}+1$ or $\mathrm{RIN}+2$, the average rate is thus $\frac{1}{\mathrm{RIN}+1.5}$. In each run, ρ can readily be obtained by summing the average rate of all connections on all links that have traffic flows and then computing the mean over all the links in the network. From a total of 1,000 runs, ρ is 0.28, which could have been higher if the SRC could be relaxed somewhat.

B. The Effect of the Number of Flows

Now we show the relationship between the number of flows in the network and bandwidth utilization. With the same 4x4 grid network, the number of flows in the network is increased from 16 to 64. The results are shown in Fig. 3. The RIN' bound has become looser as the number of flows in the network increases. As indicated in the Fig., over 50% of the actual maximum queueing delay is only one-tenth of the RIN' bound, and an average θ has reduced from 67.1% to 23.6%. Note that this result is consistent with the conclusion from the correlation coefficient $\rho_{H\theta}$ we have calculated earlier. That is, with a longer-hop path, the bound tightness tends to decrease. This is because having each flow traverse a larger number of hops in the same network is, to a large extent, equivalent to having more interfering flows. Therefore, increasing the number of flows in the network would decrease the bound tightness.

The bandwidth utilization ρ in the 64-flow network reduces from 0.28 to 0.24. When the number of flows increases, it is obvious that each flow will have larger number of interfering flows on its path. It turns out that sources peak rate decreases approximately at the same rate of the increase in the number of flows. This fact is indicated by an increase in an average interpacket time τ from 8.0 to 31.8, which is almost exactly four times greater than that of the 16-flow case. Consequently, ρ will not increase with the number of flows. Table I shows the effect of the number of flows on ρ in two different network topologies (a 4x4 grid network and an actual voice network spanning the continental US. obtained from [9] in Fig. 4). Note that the values of ρ in the 8- and 16-flow cases are higher than those in the others since not all links carry traffic and ρ is computed over only those links having traffic flows. But in all the other cases where all links in the network contain traffic flows, with sources decreasing their peak rate proportional to the total number of flows, ρ is kept constant.

TABLE I

OVERALL NETWORK UTILIZATION (ρ) WITH DIFFERENT NETWORK TOPOLOGIES AND NUMBER OF FLOWS

	Number of flows in the network				
	8	16	64	256	1024
4x4 Grid	0.37	0.28	0.24	0.24	0.24
Real Network	0.34	0.25	0.21	0.21	0.21

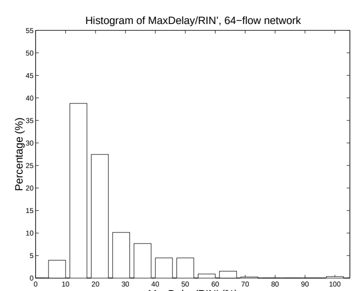


Fig. 3. The percentage of the actual maximum end-to-end queueing delay over the worst-case bound in a 4x4 grid, 64-flow network

[2]The correlation coefficient ρ_{xy} measures the correlation between two variables x and y, which is defined as $\frac{\sigma_{xy}^2}{\sigma_x \sigma_y}$ where σ_{xy}^2 is a covariance between x and y and σ_x (σ_y) is a standard deviation of x (y).

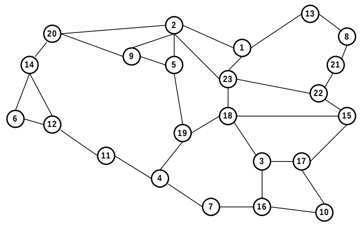


Fig. 4. A Realistic Example Network Topology[9]

IV. Modified Source Rate Condition

From the results in the previous section, it can be seen that the network utilization problem has two roots (1) the loose delay bound and (2) the dependence of the source peak rate on the number of flows in the network. The loose bound problem is analytically unavoidable because it is based on the worst-case analysis. On the other hand, knowing that the bound tends to be loose, we could increase a source peak rate in such a way that an actual maximum delay of all flows still satisfies the theoretical bound. In doing so, the bandwidth utilization will automatically increase. Furthermore, provided such a method exists, the source rate condition will be less dependent on the number of flows, alleviating the second problem.

Observe that at each node, packets from a given flow are delayed by those from interfering flows. However, interfering flows that join at the previous node may split to other links at the current node. If the interfering flows at the current node are to replace those splitting flows, the sum of total interfering flows along the path should be reduced. Based on the idea described above, we have come up with a heuristic method to improve the network utilization as follows. Suppose flow j traverses a set of links $\mathcal{L}_j$. At a given link $l \in \mathcal{L}_j$ at node i, let $\beta_j^{(l)}$ be the number of flows that shares the same incoming link as flow j but not the outgoing link l. Let $\alpha_j^{(l)}$ be the number of flow j's interfering flows on link l, as defined in section II. Note that for any flow j, $\beta_j^{(l)}$ = 0 if l is the first link on the path of flow j. Then if an interpacket time at a source of flow j, τ_j satisfies

$$\tau_j \;\geq \min\,\{\mathrm{RIN}_j,\ 1+\sum_{l\in\mathcal{L}_j}\max(\alpha_j^{(l)}-\beta_j^{(l)},0)\}, \tag{5}$$

the *actual* maximum end-to-end queueing delay of flow j is still bounded by its RIN' bound. We will refer to this modified source rate condition as SRC'.

Consider the example for calculating SRC' of flow A in Fig. 1. At node 0, $\alpha_A^{(0-3)}$ = 1 since flow A has only one interfering flow, which is B, and $\beta_A^{(0-3)}$ = 0. At node 3, $\alpha_A^{(3-6)}$ = 1 since flow A also has only one interfering flow, D, and $\beta_A^{(3-6)}$ = 1 since flow B splits to another link. At node 6, $\alpha_A^{(6-8)}$ = 3, because there are three flows - B, E, and F - joining flow A from other links, and $\beta_A^{(6-8)}$ = 0. Note that flow D is not counted because it has the same incoming link as flow A. At node 8, $\alpha_A^{(8-9)}$ = 0, and $\beta_A^{(8-9)}$ = 2 since flow D and E split to link 8-7. Thus with SRC', τ_A is given by

$$\begin{aligned}\tau_A \quad \geq \quad & \min\{\mathrm{RIN}_A, 1+\max(\alpha_A^{(0-3)}-\beta_A^{(0-3)},0)\\ & +\max(\alpha_A^{(3-6)}-\beta_A^{(3-6)},0)\\ & +\max(\alpha_A^{(6-8)}-\beta_A^{(6-8)},0)\\ & +\max(\alpha_A^{(8-9)}-\beta_A^{(8-9)},0)\}\\ \geq \quad & \min\{5,6\}=5.\end{aligned}$$

In the example provided, SRC' gives the same value as SRC because the number of splitting flows is small. If we have larger number of flows in the network, the term $1+\sum_{l\in\mathcal{L}_j}\max(\alpha_j^{(l)}-\beta_j^{(l)},0)$ will be smaller than RIN_j, increasing the source peak rate and hence the bandwidth utilization. This will be demonstrated by simulation results later.

A. Performance Evaluation

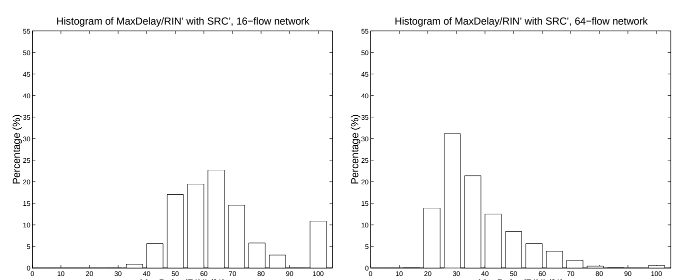
In this section, the effects of the modified source rate condition are evaluated using simulation experiments as were done in Section III. The only difference is that now the modified source rate condition SRC' is imposed instead of SRC. By changing traffic sources peak rate using SRC', the maximum end-to-end queueing delay becomes closer to RIN' bound, especially in the 64-flow network, as shown by the histograms of θ in Fig. 5. The average θ obtained for the 16- and 64-flow cases is now 66.2% and 37.4% respectively.

Fig. 6 plots the overall network utilization with the number of flows for SRC and SRC'. With SRC', sources peak rate decrease with a rate slower than the rate of increase in the total number of flows in the network – a significant advantage of SRC' over SRC. This is attributed to the fact that flows joining and splitting at each node causes the term $1+\sum_{l\in\mathcal{L}_j}\max(\alpha_j^{(l)}-\beta_j^{(l)},0)$ to be much smaller than RIN, particularly in a network with a large number of flows.

B. Empirical Peak Rate

Even with the modified source rate condition SRC', the histograms of θ in Fig. 5 show that a large portion of the actual maximum end-to-end queueing delay is still well below their bound. This suggests that the source peak rate of each flow could have been increased further, which, of course, would improve the bandwidth utilization. We are interested in finding out the maximum possible ρ in order to establish a guide line to improve a source rate condition and see how much SRC' has restricted the bandwidth usage.

To achieve this, we adopt a greedy algorithm, in which the basic idea is that at every step, a locally optimal choice is made,



(a) 16-flow Network (b) 64-flow Network

Fig. 5. The percentage of the maximum end-to-end queueing delay over the RIN' bound using SRC'

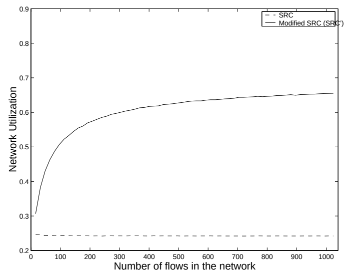


Fig. 6. Bandwidth Utilization Improvement with SRC'

with the hope that it will lead to a globally optimal solution [10]. The algorithm used is as follows. First, the modified source rate condition SRC' is computed for each flow j. Then, for a flow j, its SRC' is iteratively decreased by one unit, until the actual maximum queueing delay of *any* flow in the network exceeds its RIN' bound (checked by simulation). At this point, the algorithms stops decreasing the SRC' of flow j, and moves on decreasing the SRC' of other flows. The algorithm terminates when all the flows has been visited. The algorithm is extremely time-consuming because, for every decrease of SRC', checking for any bound violation requires a single simulation run. Therefore, only the 16-flow case has been tested.

Fig. 7 plots a histogram of θ from the experiment with a 4x4 grid, 16-flow network with the greedy algorithm applied. The utilization (ρ) obtained is 0.57 (compared to 0.34 with SRC'), and an average θ of 0.81. We can see that ρ increases about 67%, indicating that SRC' could further be relaxed. It is important to note that because a greedy algorithm does not always lead to a globally optimum solution, the bandwidth utilization obtained by the above algorithm does not imply the maximum achievable utilization, and the actual improvement might be greater than 67%. For the correlation coefficients, $\rho_{H\theta}$ is -0.474, and $\rho_{B\theta}$ is -0.541. Once again, this indicates that the bound tightness tends to decrease as the number of hops or the RIN' bound itself increases.

V. Concluding Remarks

As shown in [6], [7], when the peak rate of traffic sources are restricted, a simple worst-case queueing delay bound in a FCFS network with arbitrary traffic flow patterns can be obtained. However, this results in low bandwidth utilization. With the same worst-case delay bound, we have presented a new source rate condition using a heuristic method to increase traffic sources peak rate. Based on the simulation results, the new source rate condition considerably improves the bandwidth utilization. Further, unlike in previous work, the utilization continues to increase with the number of flows in the network. Although our heuristic source rate condition has not been mathematically proved, or it may actually theoretically result in the bound violation, the point being made here is that practically, a source rate condition such as the one proposed might work very well without violating the bound and improving the utilization. Moreover, we have shown by means of a greedy algorithm that the bandwidth utilization could have further been increased.

The following important issues are subject of our future work.

- The situation gets much worse for VBR traffic sources, where their peak rate and average rate can be considerably different. This source rate condition will result in very low network utilization.
- Because of a random routing, the RIN of each flow is never smaller than when a shortest path routing is used. Thus using the random routing results in a longer-hop path, and hence a larger number of interfering flows. This may effect the bound tightness and bandwidth utilization, for these performance measures has been shown to depend on the RIN. The effect of routing strategy is yet to be determined.
- What if the assumptions of a fixed packet size and homogeneous network are to be relaxed? A varying packet size could easily be dealt with by using the maximum packet size for the analysis so that we still have a fixed packet size assumption. This, of course, would likely to result in an even more loose bound. Relaxing the homogeneous network assumption, on the other hand, is much more difficult to deal with.

Acknowledgments

The authors would like to thank Prof. David Tipper for his helpful comments and discussion on using the RIN approach.

References

[1] R.L. Cruz, "A Calculus for Network Delay, part I: Network Elements in Isolation," *IEEE Transactions on Information Theory*, vol. 37, no. 1, January 1991.
[2] R.L. Cruz, "A Calculus for Network Delay, part II: Network Analysis," *IEEE Transactions on Information Theory*, vol. 37, no. 1, January 1991.
[3] J. Kurose, "On Computing per-session Performance Bounds in High-Speed Multi-hop Computer Networks," in *Proceedings of ACM SIGMETRICS'92*, June 1992.
[4] O. Yaron and M. Sidi, "Calculating Performance Bounds in Communication Networks," in *Proceedings of INFOCOM'93*, April 1993.
[5] C.-S. Chang, "Stability, Queue Length and Delay, part ii: Stochastic Queueing Networks," Technical Report 17709, IBM, May 1992.
[6] I. Chlamtac, A. Farago, and H. Zhang, "A Deterministic Approach to the End-to-End Analysis of Packet Flows in Connection-Oriented Networks," *IEEE/ACM Transactions on Networking*, vol. 6, no. 4, August 1998.
[7] H. Zhang, "A Note on the Deterministic End-to-End Delay Analysis," in *Proceedings of ICC'99*, June 1999.
[8] D.J. Yates, J.F. Kurose, D. Towsley, and M.G. Hluchyj, "On Per-session End-to-end Delay Distributions and the Call Admission Problem for Real-Time Applications with QOS Requirements," in *Proceedings of ACM SigComm Conference, San Francisco, CA*, September 1993.
[9] D. Medhi, "Multi-Hour, Multi-Traffic Class Network Design for Virtual Path-Based Dynamically Reconfigurable Wide-Area ATM Networks," *IEEE/ACM Transactions on Networking*, vol. 3, no. 6, December 1995.
[10] C.E. Leiserson T.H. Cormen and R.L. Rivest, *Introduction to Algorithms*, McGraw-Hill, 1994.

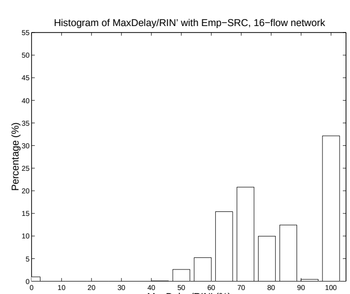


Fig. 7. The percentage of the maximum end-to-end queueing delay over the RIN' bound using the empirical SRC