# JOINTLY COORDINATING ECN AND TCP FOR RAPID ADAPTATION TO VARYING BANDWIDTH

Archan Misra
Teunis Ott

Telcordia Technologies
445 South Street,
Morristown, NJ 07960.

*Abstract*—**The introduction of service differentiation in the Internet implies that the residual bandwidth available to best-effort traffic becomes highly variable. We explore the design of a rapidly-reactive congestion control framework, where the ECN-aware best-effort flows aggressively go after any unused capacity. By making routers mark packets in a much more aggressive manner, we are able to achieve fast backoff in the network without resorting to TCP's current drastic step of halving the congestion window. Simulations indicate that our ECN-mod protocol is better than ECN-NewReno in exploiting rapid variations in the available bandwidth. Moreover, the milder backoff policy of ECN-mod also makes the link utilization less dependent on the exact values of the parameters in the router marking function.**

## I. Introduction

With the adoption of explicit service differentiation and work-conserving scheduling policies in Internet routers, the bandwidth available to best-effort traffic becomes a highly variable quantity. To effectively adapt to such variable 'residual' capacity, transport protocols for best-effort traffic face two conflicting requirements:

- During periods of congestion, the adaptive flows must backoff rapidly to prevent *congestion collapse.*
- Whenever additional bandwidth becomes available, such flows must rapidly and aggressively increase their transmission rate, thereby preventing unnecessary under-utilization of capacity.

In other words, the best-effort flows should be *rapidly reactive*, exploiting changes in capacities that occur over relatively shorter time scales.

In this paper, we show how the design of such a rapidly reactive window-based congestion control protocol can benefit from an intelligent use of the Explicit Congestion Notification (ECN) [1], [2] feature becoming available in the Internet. In particular, we study a generalized class of window-based protocols that conform to the 'TCP paradigm', whereby the congestion window is modified on the receipt of every acknowledgement packet. We leverage our work in [3], which analyzed the properties of such generalized congestion control in detail and provided recommendations on the choice of various adaptation parameters.

While ECN provides a much more explicit form of congestion feedback than packet dropping algorithms, we believe that the full power of ECN has not been effectively harnessed. Current implementations of ECN use the same algorithm (such as RED [4]) to determine both the packet *marking* and dropping probabilities. Since ECN feedback does not involve loss of transmitted packets, the marking probability can, however, be as high as 1 without causing deleterious side effects. Such aggressive marking behavior by the routers allows us to design an ECN-aware transport protocol that reacts to individual marked packets in a much more gentle way than TCP's current *congestion avoidance* [5] algorithm, yet can reduce its transmission rate very rapidly during a congestion episode. Such rapid response to congestion feedback also permits the flow to increase its window more aggressively than TCP in the absence of congestion; if the increased transmission rate proves to be too large, the flow can take rapid corrective action. We thus believe that the potential of ECN-based feedback can be exploited only if the window adaptation mechanism of an ECN-capable protocol is designed in tandem with the marking function employed in Internet routers.

We first present an overview of our analysis of the generic window-based adaptation mechanism. We also present the rationale behind our preferred choice for the values of the adaptation parameters and explain why our ECN-aware 'modified TCP' algorithm is expected to provide more rapid adaptation than the current TCP congestion avoidance scheme, especially in environments characterized by rapid variation in available bandwidth. We then report on a set of simulation results that show how our ECN-modified protocol is able to achieve higher utilization than conventional ECN-aware TCP. To simulate an environment where the bandwidth for best-effort traffic varied realistically, we used Voice-over-IP (VoIP) to simulate the variable load offered by higher priority traffic classes. We also provide results that illustrate the effect of stronger (more aggressive) marking functions in ECN-capable buffers on our ECN-modified and conventional ECN-TCP flows. In particular, we show that using such modified marking algorithms in tandem with our ECN-aware protocol results in higher capacity utilization over a wide variation in the maximum marking probability.

It is well-known that using packet losses as the sole indicator of congestion can lead to very low effective TCP throughput, especially over wireless links characterized by larger error rates. While introducing ECN support does improve the link utilization, the adoption of a more rapidly reactive ECN-aware protocol can lead to even better resource utilization, especially in wireless networks characterized by variable available capacity. Accordingly, while our current studies simulate variability in the available bandwidth through changes in the higher priority traffic load, our conclusions should be specially relevant for environments involving satellite links, characterized by large bandwidth-delay products and rapid changes in the available link capacities.

## II. Generalized Congestion Control and Current ECN-TCP

Assuming that routers in the Internet can use ECN-based *packet marking* to indicate network congestion, a generalized congestion window-based protocol can be described by the following behavior:

*Whenever an acknowledgement arrives for an unmarked data packet, the congestion window increases from its current value* $W$ *by* $incr(W)$*. If, however, the acknowledgement indicates that the data packet had been marked in the forward path, the congestion window is decreased from* $W$ *by* $decr(W)$*.*

If we assume that the marking probability of an individual packet remains constant at $p$, the drift associated with the congestion window process for a value $W$ is given by

$$\begin{aligned} drift(W,p) &= E[W_{n+1} - W_n | W_n = W] \\ &= (1-p).incr(W) - p.decr(W) \qquad (1) \\ &= p.incr(W).\left(\frac{1-p}{p} - \frac{decr(W)}{incr(W)}\right). \end{aligned}$$

Let $q(W)$ denote the function

$$q(W) = \frac{decr(W)}{incr(W)}. \qquad (2)$$

Any sensible choice for $q(W)$ must ensure that the window always increases if the marking probability $p$ is $0$, and that the window does not grow without any bound for any marking probability. Accordingly, we require

$$q(1) = 0, \quad \lim_{W \to \infty} q(W) = \infty. \qquad (3)$$

Equation (2) shows that, under constant $p$, the 'zero-drift' value or 'center' of the congestion window is obtained when

$$q(W) = \frac{1-p}{p} \qquad (4)$$

From a conceptual viewpoint, manipulating $q(W)$ provides one-way of devising Internet congestion control algorithms: $q(W)$ essentially determines a flow's "average" window size (and hence, the maximum number of unacknowledged packets) for a given level, $p$, of congestion indication. Thus, $q(.)$ defines a "response surface" of $W$ to $p$. Designers of congestion control algorithms, who first decide on the function $q(W)$, still have considerable leeway in devising the precise adaptation functions $incr(W)$ and $decr(W)$. Conversely, designers can explicitly choose individual $incr(W)$ and $decr(W)$, thereby specifying the response surface $W(p)$ implicitly.

From a practical perspective, we follow the procedure of [3], [6] and focus on the 'polynomial' class of adaptation algorithms, characterized by

$$incr(w) = c_1 w^{\alpha}, \ decr(w) = c_2 w^{\beta}. \qquad (5)$$

For such polynomial window adaptation algorithms, the response surface $q(W)$ is given by:

$$q(w) = \frac{decr(w)}{incr(w)} = \frac{c_2}{c_1} w^{\beta - \alpha}. \qquad (6)$$

Thus, a flow transporting a very large file and subject to a constant marking probability $p$, will see its congestion window fluctuate around $w(p)$, defined as

$$w(p) = \left(\frac{c_1}{c_2}\,\frac{1-p}{p}\right)^{\frac{1}{\beta-\alpha}}. \qquad (7)$$

Clearly, to ensure conformance with condition (3), we need $\alpha < \beta$.

### A. Current TCP Practice and Limitations

TCP's famous *congestion avoidance* algorithm [5] is based on an additive-increase, multiplicative-decrease principle: in the absence of congestion, the congestion window $cwnd$ increases by 1 once every round trip time (RTT); on detection of a congestion episode, $cwnd$ decreases by $\frac{1}{2}$. Neglecting transients such as fast recovery and slow-start, TCP's congestion control mechanism is thus a member of the polynomial class, with the parameters $c_1 = 1, \ \alpha = -1, \ c_2 = \frac{1}{2}, \ \beta = 1$. Of course, most modern TCP versions, such as NewReno or Vegas, halve their window only once for multiple packet losses occuring within a single window.

For TCP, the desire to have rapid multiplicative backoff governed the choice of $decr(W) = \frac{W}{2}$. Such rapid backoff was clearly the correct choice for contemporaneous networks, where routers implemented congestion notification only implicitly via tail-drop buffer management policies. It is well-known that using packet losses as an indirect congestion indicator can lead to very poor performance over wireless links characterized by larger link error rates, even if the intermediate buffers implement more advanced randomized packet dropping algorithms, such as RED. Such poor utilization occurs principally because such ECN-unaware TCP implementations cannot distinguish between losses due to link errors and those due to congestion at intermediate buffers.

Clearly, the introduction of an explicit marking-based congestion notification mechanism, such as ECN, should improve matters significantly. What is perhaps not as clearly appreciated is the fact that even ECN-aware TCP applications can fail to use resources efficiently in networks characterized by large and rapidly varying bandwidth-delay products, such as satellite networks. The current ECN recommendations [1], [2] state that a TCP source must react to the notification of an ECN-marked packet in exactly the same way as it reacts to the discovery of congestion via packet loss. *We do not share this opinion- we believe that proper exploitation of ECN-based feedback allows us to develop a much more responsive window-based transport* *protocol.* ECN-aware TCP suffers from performance degradation primarily due to the sharp drop in the transmission rate of a TCP flow after the detection of a congestion event. Clearly, a TCP flow that experiences congestion when $cwnd = 100$ would drop its congestion window to $cwnd = 50$ and subsequently take 49 RTTs to regain the optimal window value.

A trivial fix, that rectifies this drawback to some degree, is to make $decr(W)$ and $incr(W)$ both much smaller, without disturbing their ratio. While this would leave $q(W)$ (and hence the basic bandwidth sharing paradigm) unchanged, the flow would clearly be less responsive to changes in $p$. It is thus necessary to study the design and performance of congestion control protocols in an ECN-enabled environment more carefully. To motivate this statement, we also note that it is fallacious to believe that non-multiplicative decrease algorithms necessarily result in milder backoff. For example, if $decr(W) = 1$ (MSS) in a scheme with marking instead of dropping, a router that marks incoming packets with a probability $p = 1$ during an entire RTT reduces the windows of all constituent flows to 1 MSS or less – more draconic than halving all windows! In later sections, we shall illustrate how such modifications to the marking function in the router buffers can be combined with alternative choices for $\alpha, \beta, c_1$ and $c_2$ to devise a more responsive congestion control mechanism, especially for environments characterized by rapid variation in the available bandwidth.

## III. Choosing the Generalized Window Adaptation Parameters

Analysis and results presented in [7] show that alternative settings of the polynomial parameters, $\alpha,\ \beta,\ c_1$ and $c_2$, may be preferable in an ECN-enabled environment, for both theoretical and practical reasons. It must be noted that, since marking does not cause any loss of packets, the marking probability for even conventional TCP can be made as high as $100\%$, without giving rise to loss-related transients such as timeouts and slow-start. It is precisely this flexibility that allows us to design a more rapidly reactive congestion control environment, using ECN-aware flows that are less drastic than TCP in reducing their windows in response to a single marked packet, yet can be rapidly throttled using aggressive marking strategies at intermediate buffers.

### A. Choosing $\beta$

While a potentially infinite number of values for $\beta$ may be possible, analyses in [3], [7] shows that choosing $\beta = 1$ has certain theoretical appeal.

*Theorem 1:* If $\beta < 1$ (and $\alpha < \beta$), and $c_1 > 0,\ c_2 > 0$, then the congestion window process has the coefficient of variation

$$Coeff.Var(W) = \frac{st.dev(W)}{E[W]} \sim \frac{c_1^{-\frac{1-\beta}{2(\beta-\alpha)}} c_2^{\frac{1-\alpha}{2(\beta-\alpha)}}}{\sqrt{2(\beta-\alpha)}}\ p^{\frac{1-\beta}{2(\beta-\alpha)}}. \tag{8}$$

Clearly, if $\beta = 1$, then the coefficient of variation of $W$ becomes independent of $p$ for $p \downarrow 0$ (i.e. when the congestion window is allowed to be very large). When $\beta = 1$, the distribution of $\left(\frac{p}{c_1}\right)^{\frac{1}{1-\alpha}} W$ becomes independent of $p$ and $c_1$ for $p$ small. The non–dependence on $p$ implies non-dependence on the average value of $W$: *scale invariance!* Accordingly, we choose $\beta = 1$, thereby retaining the multiplicative decrease behavior of current TCP. Of course, $c_2$ should be much smaller than TCP's choice of $\frac{1}{2}$ to prevent the transmission rate from fluctuating wildly in response to a single marked packet.

### B. Choosing $\alpha$

If the marking probability remains constant at $p$, then the average number of marked packets within a single congestion window $W$ worth of packets is clearly $p * W$. By definition of the congestion window, these $W$ packets are transmitted over one RTT. Furthermore, since the average congestion window for a flow subject to a constant marking probability $p$ and performing 'polynomial' window adaptation is given by (7), it is easy to see that the number of marked packets per RTT is roughly

$$p^{\frac{\beta-\alpha-1}{\beta-\alpha}} (1-p)^{\frac{1}{\beta-\alpha}} \left(\frac{c_1}{c_2}\right)^{\frac{1}{\beta-\alpha}}. \tag{9}$$

Having $\beta - \alpha \leq 1$ guarantees that, as the marking probability $p \downarrow 0$, the average number of marked packets per RTT does not go to zero. This makes it possible for routers to give frequent feedback to the sources on the state of congestion even under lightly loaded conditions, and makes it possible for sources to reactly gently to marked packets and aggressively to non–marked packets. From an implementation standpoint, if $\beta = 1$, $\alpha = 0$ provides the simplest choice that satisfies the constraint $\beta - \alpha \leq 1$. Thus, unlike conventional TCP, our ECN-aware congestion control algorithm increases the window by a fixed amount $c_1$ on receiving an acknowledgement for an unmarked packet. For large values of $W$, our ECN-modified protocol goes after available bandwidth much more aggressively than TCP. It is this aggressive behavior that enables our protocol to rapidly escalate its transmission rate, even over smaller time-scales.

### C. Choosing $c_1$ and $c_2$

While we have theoretically motivated our preference for $\beta = 1$ and $\alpha = 0$, we do not have corresponding theoretical preferences for $c_1$ and $c_2$. Indeed, to a first approximation, $c_1$ and $c_2$ are scaling constants that merely scale the response surface function $W(\cdot)$, without fundamentally changing its shape. We would simply like our choices for $c_1$ and $c_2$ to ensure two things-

1. the average window size for moderate marking rates does not become extremely high.
2. the average number of packets marked per RTT should not become unusually low.

The precise choice of $c_1$ and $c_2$ also depends on the appropriate choice of a specific marking function in the router buffers. For the simulation studies presented here, as well as in [3], we have experimented with a number of $c_1$ and $c_2$ values, which effectively ensure that the expected number of marked packets per RTT (for small $p$) lies in the range $(\frac{1}{5}, 5)$.

## IV. Simulation Studies and Observations

We have performed extensive simulations using the ns-2 [8] simulator to verify the performance of our proposed changes.

Our "gently-greedy" adaptive protocol was obtained by modifying the TCP simulator code. For the results presented here, we choose $\alpha = 0$ and $\beta = 1$, as per the recommendations of the previous section. The objectives of our simulation-based studies are two-fold:

- To demonstrate that our "modified-ECN" flows adapt to variable link capacity and achieve better utilization than corresponding "ECN-TCP" flows.
- To show how changing the marking function in a router buffer, in tandem with the "modified-ECN" window adaptation algorithm, provides better link utilization than that achieved by a conventional implementation of ECN marking with conventional ECN-aware TCP flows. In particular, ECN-mod flows are less sensitive than corresponding ECN-aware TCP NewReno flows to changes in the exact marking probabilities.

In [3], we have also investigated the fairness issues involved in having our "modified-ECN" flows coexist with conventional ECN-capable, as well as ECN-unaware, TCP flows. Our results, which we omit due to space limitations, also show that our modified-ECN algorithm, while more aggressive than conventional TCP, does not unduly starve such competing TCP flows.

### A. Simulation Environment

To simulate an environment where the link capacity available to best-effort flows was variable, we used the simulation topology of figure 1, where VoIP and best-effort traffic are buffered in two separate queues and Class Based Weighted Fair Queuing (more precisely, SCFQ) is used to isolate the two classes. To provide voice higher priority, the VoIP class had a weight of 0.8, compared to 0.2 for TCP traffic, even though the offered load of VoIP traffic was often much lower than that of TCP; such a large relative weight effectively shields VoIP flows from best-effort traffic. The TCP flows then effectively utilize the portion of the link bandwidth that is unused by higher-priority VoIP flows.

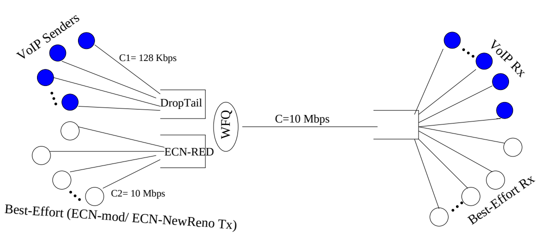


Figure 1: Simulation Topology

Each VoIP flow was modeled as an exponentially modulated on-off process, with the mean on and off times, as per the ITU P.59 [9] recommendations, being 1.008 and 1.587 secs respectively. During the on-period, the voice source generates an 80 byte packet every 10 msecs (similar to that in typical G.711 codecs). Since each packet has an additional header of 40 bytes, the peak rate of a single source is 96 Kbps. To simulate variations in the traffic load offered by high priority VoIP traffic, we modeled the total number of instantaneous calls as a birth-death process. New voice calls arrived according to a Poisson process with rate $\lambda$; the duration of each call was exponentially distributed with a mean $\frac{1}{\mu}$.

The best-effort traffic consisted of two alternative types:

- "ECN–NewReno": To model conventional ECN-aware TCP, we used ECN-aware TCP NewReno flows, where a source reacts to marked and dropped packets in an (almost) identical manner.
- "ECN–mod", or "Modified ECN": This is our 'responsive' ECN-capable protocol, where the source reacts to marked packets as described in Section III, and to dropped packets in the same way as TCP NewReno. We always have $\beta = 1$, $\alpha = 0$, $c_2 = 2^{-3}$. We report on three choices for $c_1$: $2^{-2}$, $2^{-3}$, and $2^{-4}$. These choices lead to $pW(p) \in \{2, 1, 1/2\}$. We found that $c_1 = 2^{-4}$, $pW(p) = 1/2$ gave the best results. Unlike current TCP versions, our ECN-mod protocol decreases its window by $c_2 W^{\beta}$ on receiving an acknowledgement for *every* marked packet, even if such congestion marking occurs multiple times within a single window worth of packets.

The router dropping function for ECN-NewReno flows, $p(Q)$, was based on the linear-drop model:

$$p(Q) = \begin{cases} 0 & \text{if } 0 \le Q < min_{th}, \\ p_{max} * \frac{Q - min_{th}}{max_{th} - min_{th}} & \text{if } min_{th} \le Q \le max_{th}, \\ p_{max} & \text{if } max_{th} < Q. \end{cases} \tag{10}$$

To ensure a fair comparison for different values of $\alpha$ and $\beta$, we modified the marking function, $p_{mod}(Q)$, for ECN-mod flows, such that given a specific buffer occupancy, the congestion windows are the same in all cases. This is achieved by defining $p_{mod}(Q)$, the marking function[1] for ECN-mod packets as:

$$p_{mod}(Q) = \left(1 + \frac{c_2}{c_1} * \sqrt{\frac{2 * (1 - p(Q))}{p(Q)}}\right)^{-1} \tag{11}$$

### B. Bandwidth Utilization of ECN-mod and ECN-NewReno

For the plots provided here, the bottleneck link capacity $C$ is 10 Mbps; the VoIP queue was sized to have a maximum drain time of 20 msecs. The number of best effort flows (either all ECN-NewReno or all ECN-mod) was 20. The best-effort queue had RED parameters (in packets) of $min_{th} = 25$, $max_{th} = 75$, $p_{max} = 0.2$ and buffer size $B = 150$ (following the recommendations in [11]). The RTT of the TCP connections are uniformly spaced out over the interval $(25, \ldots, 250)$ msecs.

Figure 2 plots the total goodput (VoIP+ TCP), as well as the TCP goodput alone, as $\lambda$ is varied to change the average number of simultaneous voice calls. We see that ECN-mod sources (especially with $c_1 = 0.0625$) are better than ECN NewReno in responding faster to instantaneous fluctuations in the available capacity. In all cases, the VoIP throughput was unaffected by the choice of the best effort protocol, demonstrating the effectiveness of WFQ in effectively giving VoIP traffic higher priority. The plots also show that choosing $c_1 = 0.25$ causes ECN-mod

[1] For experiments that simply compare the effectiveness of ECN-mod and ECN-NewReno in exploiting unused capacity, we left unchanged RED's drop-biasing mechanism that results in a uniform distribution between two consecutive packet drops. In such a case, equation (11) was suitably modified to reflect the fact that the average unconditional dropping probability is not $p(Q)$ but $2 * p(Q)$. For experiments that studied the capacity utilization as a function of either the marking probability or the number of adaptive flows, we modified RED to implement the *Geometric* drop-biasing policy [10], where equation (11) applies unchanged.

to perform worse than ECN-NewReno. Our modified window increase procedure is too aggressive in this case; rapid increases in the congestion window are followed by rapid decreases (as the router marking always lags behind the instantaneous rate). Such high fluctuations in the window size increase the likelihood of buffer underflow and lead to poorer link utilization.

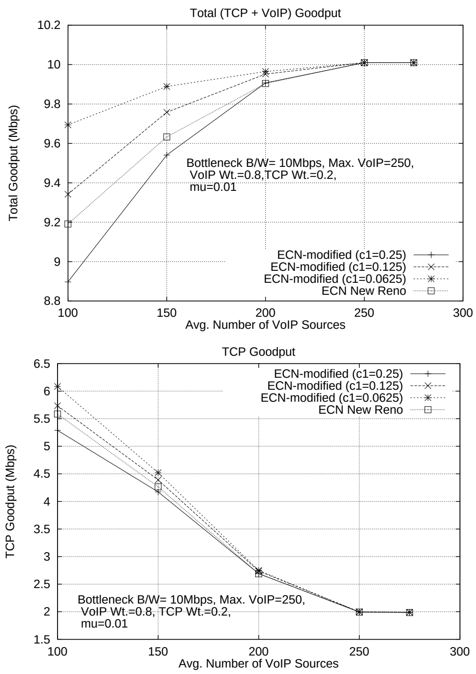


Figure 2: Comparative Capacity Utilization

Figure 3 shows the packet marking rates for the best-effort traffic, as well as the coefficient of variation of the occupancy of the best-effort queue. With the router performing ECN-based feedback, the packet loss rates for the adaptive traffic were practically zero in all cases.

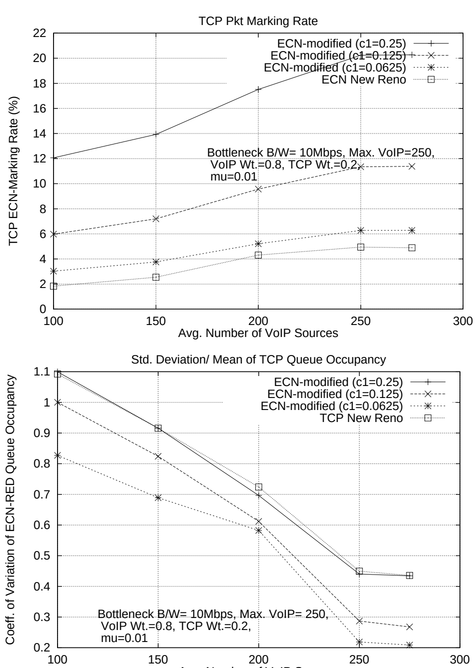


Figure 3: Flow Marking Rates and Queue Variability

The plot of the marking probabilities also shows that our *parameter adjustment* procedure causes the ECN-mod flows to experience much higher marking rates than ECN-NewReno flows. The superiority of ECN-based congestion avoidance is clearly evident from these plots, since the router can mark packets with rates as high as 20% with no degradation in the achieved goodput. The plot of the coefficient of variation (defined as $\frac{Std.Deviation}{Mean}$), also shows that ECN-mod flows result in in a lower coefficient of variation of the queue occupancy, as opposed to ECN-NewReno flows. Unlike ECN-NewReno, where the drastic reduction in the window size leads to larger variability in the queue occupancy, ECN-mod sources react more gently to individual packet markings; the variation in the queue occupancy is consequently much smoother. This was, of course, not true for ECN-mod with $c_1 = 0.25$, where the overly aggressive increase in the window size actually increased the variability in the queue occupancy.

### C. *Sensitivity to Buffer Marking Parameters*

One of the big advantages of using ECN for congestion feedback is that the marking probability can be made as high as 1, without giving rise to undesirable loss-related transients. It is precisely this feature of aggressive marking that allows ECN-mod flows to respond much less drastically to an individual acknowledgement for a marked packet. However, it is well-known that the optimal setting of the various thresholds and parameters in randomized feedback algorithms such as RED varies with the number of flows. While various algorithms, such as SRED [12] and BLUE [13], have been suggested to adaptively alter the feedback parameters (such as $max_{th}$ or $p_{max}$) based on estimates of the number of active flows or link loads, such adaptive algorithms have not yet been practically implemented. It is naturally interesting to study the relative degree to which a sub-optimal parameter choice in non-adaptive implementations of random marking affects the performance of ECN-mod and ECN-NewReno flows. Accordingly, for the simulations in this sub-section, we used the same settings as the previous sub-section for the best-effort queue, except that the maximum marking probability $p_{max}$ (in equation (10)) was varied between $0.05$ and $1$. Moreover, the average number of VoIP call was maintained at $200$ by setting $\lambda = 2$ in all these simulations.

Figure 4 plots the total goodput (VoIP+ TCP), as well as the TCP goodput alone, as $p_{max}$ is varied to change the maximum marking probability. Analytical techniques that predict the queue occupancy as a function of $p_{max}$ (see, for example, [14]) show that increasing $p_{max}$ decreases the average queue occupancy. We see that the utilization obtained by ECN-mod sources (especially with $c_1 = 0.0625$) is much less sensitive to an increase in $p_{max}$ compared to that attained by ECN-NewReno flows. Indeed, as $p_{max}$ is increased from $0.05$ to $1$, the TCP goodput for 20 ECN-mod sources ($c_1 = 0.0625$) drops from $\sim 2.77$ Mbps to $\sim 2.7$ Mbps only; in contrast, for ECN-NewReno flows, the TCP goodput drops from $\sim 2.765$ Mbps to $\sim 2.64$ Mbps for ECN-NewReno flows. The reason for the relatively worse performance of ECN-NewReno is not too difficult to find. As $p_{max}$ is increased, the average queue occupancy (and indeed the individual average $cwnd$ of each flow) decreases. Since each ECN-NewReno flow halves its window on receipt of an acknowledgement for a marked packet, the potential for buffer underflow (and bandwidth wastage) becomes stronger. In contrast, the response of each ECN-mod flow to a single marked packet is much milder; since, the flow decreases its $cwnd$ by a much smaller value, the potential for buffer underflow is greatly diminished.

Additional plots (omitted due to space constraints) confirm, as expected, that the average marking rates are higher for ECN-mod flows than ECN-NewReno flows. Moreover, the smoother variation in the window evolution of an individual ECN-mod flow also translates into a smoother evolution (lower coefficient of variation) of the occupancy of the best-effort buffer. As before, these studies demonstrate how more-responsive congestion control can be achieved by combining aggressive packet marking with a less drastic decrease in window size by individual flows. More importantly, the studies demonstrate that such a congestion control strategy can make the link utilization less sensitive to the choice of individual marking parameters. Such insensitivity is practically important, since the variation in the best-effort traffic load can make the determination of optimal parametric values for the marking algorithm quite difficult.

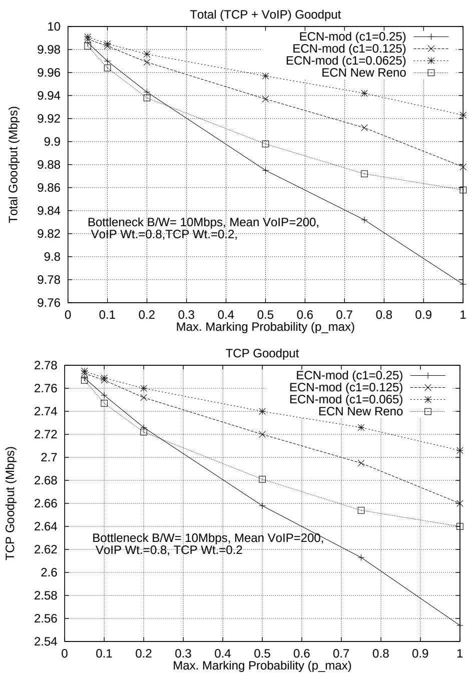


Figure 4: Capacity Utilization with Changing $p_{max}$

## V. Conclusion

In this paper, we have studied a framework for "rapidly reactive congestion control" in the Internet. This framework is motivated by the need for future best-effort adaptive traffic to not only exhibit fast backoff during congestion, but also to aggressively consume the rapidly-varying bandwidth left unused by higher priority traffic. The framework is based on using ECN-based marking in routers and developing a window-based protocol whereby sources react to individual marked packets much more gently than conventional TCP. We first leveraged our work in [3] to develop a framework for such ECN-aware generalized adaptation protocols and motivated our choice for having $\alpha = 0$, $\beta = 1$ and small $c_1$ and $c_2$ in the polynomial generalized congestion avoidance algorithm. The key idea was to note that, unlike packet dropping rates, the marking probability in a router could be as high as $100\%$.

Although by no means conclusive, our simulations have shown that, by such cooperation between router and rapidly adaptive sources, it becomes possible for "best effort" flows to better utilize the bandwidth left unused by higher priority traffic, especially if the unused bandwidth is quickly varying over time. Additional simulation-based results in [3] show that, while a well-designed ECN-mod protocol does obtain a larger share of the available capacity than either ECN-aware TCP NewReno or ECN non-capable TCP NewReno, it does not starve either of them. Moreover, our ECN-mod protocol requires very minor changes in current TCP implementations.

Our simulations also establish an additional benefits of using our ECN-mod adaptation algorithm. In contrast to ECN-NewReno flows, the utilization attained by ECN-mod flows was relatively unaffected by changes to the $p_{max}$ parameter in the router marking function. These simulations also attest to a big benefit with ECN: *the maximum marking rate can be as high as* $1$, *without adversely impacting utilization in any significant way.*

Although our simulations used a variable amount of higher-priority VoIP traffic to change the best-effort bandwidth, our conclusions should apply to any environment where the available capacity is rapidly changing. In particular, we believe that such ECN-based rapid adaptation is crucial over satellite links, which are characterized by variable IP-layer capacities, large bandwidth-delay products and high link error rates.

## References


[1] Floyd, S., "TCP and Explicit Congestion Notification", *ACM Computer Communications Review* 21 no 5, 1994.

[2] Ramakrishnan, K.K. and Floyd, S., "A Proposal to add Explicit Congestion Notification(ECN) to IP", *RFC 2481*, January 1999.

[3] Ott, T.J. and Misra, A., "ECN and 'Meekly Aggressive' Flows", *under submission*, 2001.

[4] Floyd, S. and Jacobson V., "Random Early Detection Gateways for Congestion Avoidance", *IEEE/ACM Transactions on Networking*, August 1993.

[5] Jacobson, V. and Karels. K., "Congestion Avoidance and Control', *Proceedings of ACM SIGCOMM*, 1988.

[6] Bansal, D. and Balakrishnan, H., "Binomial Congestion Control Algorithms", *Proceedings of IEEE INFOCOM*, April 2001.

[7] Ott, T.J., "ECN Protocols and the TCP Paradigm", *ftp://ftp.research. telcordia.com/pub/tjo/ECN.ps*, 1999.

[8] "The ns-2 network simulator", *http://www-mash.CS.Berkeley.EDU/ ns.*

[9] International Telecommunications Union, "Recommendation P.59, Artificial Conversational Speech", 1993.

[10] Misra, A., Ott, T.J. and Baras, J., "Using 'Drop-Biasing' to Stabilize the Occupancy of Random-Drop Queues with TCP Traffic", *Proceedings of IEEE International Conference on Communications Systems (ICCS)*, November 2000.

[11] Floyd, S., "Recommendations on using the 'gentle' variant of RED", *http://www.aciri.org /floyd/red.html*, March 2000.

[12] Ott, T., Lakshman, S. and Wong, L., "SRED: Stablized RED", *Proceedings of IEEE INFOCOM*, March 1999.

[13] Feng, W., Kandlur, W., Saha, D. and Shin, K., "Blue: A New Class of Active Queue Management Algorithms", *UM CSE-TR-387-99*, April 1999.

[14] Misra, A., Ott, T.J. and Baras, J., "The Window Distribution of Multiple TCPs with Random Loss Queues", *Proceedings of IEEE GLOBECOM*, December 1999.