Performance analysis of Compound TCP with AQM

Haseen Rahman, Krishnamurthy Giridhar and Gaurav Raina
India-UK Advanced Technology Centre of Excellence in Next Generation Networks
Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai 600 036, India
Email: {ee10m011, giri, gaurav}@ee.iitm.ac.in

***Abstract*—We study Compound TCP (C-TCP), the transport protocol in the Windows operating system, in different buffer sizing regimes along with Drop-Tail and Random Exponential Marking (REM). The buffer sizing regimes we focus on are the widely deployed bandwidth-delay rule and a small buffer regime. The performance metrics we consider are stability of the queue size, queuing delay, link utilisation and packet loss.**

We analyse the following models: (i) a non-linear model for C-TCP with Drop-Tail and small buffers, (ii) a stochastic variant of REM along with C-TCP, and (iii) the original REM proposal as a continuous time non-linear model with delayed feedback. We develop conditions to ensure local stability and show that variations in system parameters can induce a Hopf bifurcation, leading to the emergence of limit cycles.

With Drop-Tail, and small buffers, the Compound parameters and the buffer size play a key role in maintaining stability. With the stochastic variant of REM, larger thresholds for marking/dropping packets can destabilise the system. With REM, an increase in the feedback delay, or variations in the queue management parameters, can induce a Hopf bifurcation. Design guidelines for Compound parameters, based on analysis, to ensure stability are provided. Packet-level simulations corroborate some of the analysis.

***Index Terms*—Compound TCP, active queue management, buffer sizing, performance.**

I. Introduction

Compound TCP (C-TCP) [15] is widely deployed as it is the default transport layer protocol in the Windows operating system. So its evaluation can help understand the performance of today's networks. Network performance is also influenced by the choice of Active Queue Management (AQM) policies and the size of router buffers [3][10]. Thus far there is no consensus on the optimal transport protocol, queue management strategy, or buffer sizing rule.

There are numerous proposals for AQM schemes; for example, Random Early Detection (RED) [2], the Proportional and the PI Controller [7], and Random Exponential Marking (REM) [1]. In this paper, we study Drop-Tail as it is commonly used. We also study REM in some detail as it has been proposed for both wired and wireless networks, and targets both a *desired queue size* and *high link utilisation*. There are two extreme asymptotic regimes for sizing router buffers: the currently deployed *bandwidth-delay* rule and a *small buffer* regime; see [13], [16] for more details. We focus our evaluation of Compound, along with Drop-Tail and the REM queue policy, in the aforementioned buffer regimes.

We consider the following models: (i) a recently proposed non-linear model for C-TCP [14] with Drop-Tail and small buffers, (ii) a stochastic variant of REM along with C-TCP, and (iii) the original REM proposal as a continuous time non-linear dynamical system with time delayed feedback. All the models are non-linear, so we study both the local *stability* and the local *bifurcation* properties. To that end, we develop conditions to ensure local stability and show that these conditions would be violated via a Hopf bifurcation [6], [9] which would lead to periodic oscillations in the form of limit cycles. The key performance metrics considered are stability of the queue size, link utilisation and packet loss. Queuing delay is also a very important metric for performance [10]; it could be significant in routers that use the bandwidth-delay rule [3], and would be negligible with small buffers [12].

With small Drop-Tail buffers, and with the stochastic variant of REM, we show that the threshold for dropping packets should be chosen in conjunction with Compound parameters to ensure stability. The analysis offers design guidelines for system parameters to ensure sufficient, as well as necessary and sufficient, conditions for stability. Our analysis of the original REM proposal shows that an increase in the feedback delay, or variations in the REM parameters, can cause the system to lose stability via a Hopf bifurcation. The stability analysis can guide REM parameters to ensure locally stable operation. Some of our analysis is complemented with packet-level simulations, using the Network Simulator (NS2) [17], under different traffic mixes.

The rest of the paper is organised as follows. In Section II, we briefly discuss Compound TCP, router buffer sizing and Active Queue Management. In Section III, we analyse the stability and bifurcation properties of the various fluid models. In Section IV we conduct some packet-level simulations, and summarise our contributions in Section V.

II. Compound TCP, Buffer Sizing and AQM

We outline some pertinent aspects of Compound, buffer sizing and queue management as related to this paper.

Compound TCP: The Compound protocol aims to use both *queuing delay* and *packet loss* as feedback to regulate its flow and congestion control algorithms. Compound maintains both $cwnd$ (the *loss window*) and $dwnd$ (the *delay window*). The loss window is the same as in the standard TCP Reno algorithm [5], which aims to control the loss based component. The delay window was introduced to cater for delay as an additional feedback mechanism from the network. The sending window w is calculated as

$$w = \min(cwnd + dwnd, awnd),$$

where $awnd$ is the minimum window size. To estimate the transmission delay of a packet, a variable called $baseRTT$ is maintained which is the the minimum round-trip time (RTT) observed till that time. Then the number of backlogged packets of the connection $diff$ can be estimated as

$$diff = \left(\frac{w}{baseRTT} - \frac{w}{RTT}\right) baseRTT.$$

The algorithm to find the value of $dwnd$ is as follows:

$$dwnd(t{+}1) = \begin{cases} dwnd(t) + \left(\alpha\, w(t)^k - 1\right)^+ & \text{if } diff < \gamma_{th} \\ \left(dwnd(t) - \zeta\, diff\right)^+ & \text{if } diff \geq \gamma_{th} \\ \left(w(t)(1-\beta) - cwnd/2\right)^+ & \text{if loss,} \end{cases}$$

where $(z)^+$ is defined as $\max(z, 0)$, γ_{th} is the threshold for $diff$ and $\zeta > 0$. The parameters α, β and k influence the scalability, smoothness and responsiveness of the window function, and their default values are $\alpha = 0.125$, $\beta = 0.5$ and $k = 0.75$ [15]. If $diff < \gamma_{th}$, the network path is assumed to be under utilised and C-TCP acts aggressively by increasing the sending rate. Otherwise, the network path is considered to be congested and the delay based component reduces its window. This algorithm also tries to ensure fairness among competing TCP flows.

Buffer Sizing: In the literature, three buffer sizing regimes have been identified [13] [16]. First, the currently deployed large buffer regime where the buffer size B is dimensioned to be $C * \overline{RTT}$, where C is the line rate and $\overline{RTT}$ is the average round-trip time of all flows using that link. In practice, $\overline{RTT}$ is taken to be 250 ms [16]. Second, an intermediate buffer regime, where B is $C * \overline{RTT}/\sqrt{N}$ and N is the number of active TCP connections. Lastly, a small buffer regime where buffers are sized to the order of tens of packets and the dimensioning rule does not depend on C, $\overline{RTT}$ or N. We focus on the large buffer regime, as it is currently widely deployed, and the small buffer regime as it is appealing not to have network parameters influence the choice of buffer size.

Active Queue Management: In this paper, we focus on a modern AQM like REM [1] and the commonly used Drop-Tail policy. Drop-Tail simply drops packets only when the buffer is full. In contrast, REM aims to control the queue size to a predefined value while simultaneously targeting high link utilisation. The two main design aspects of REM are (i) Match Rate and Clear Buffer and (ii) Sum Price, which are now briefly outlined.

Match Rate and Clear Buffer: The idea is to stabilise the input rate around the link capacity and the queue size around a specified target, regardless of the number of users. REM maintains a variable called *price* as a congestion measure. The price is updated on the basis of *rate mismatch* (difference between input rate and link capacity) and *queue mismatch* (difference between the current and the target queue size). An increase in the number of users would cause the mismatch in the rate and queue to increase, which would increase the price and hence the marking probability. A high marking probability would send a strong signal to the end-systems to reduce their rates. For a queue indexed l, the price $p_l(t)$ is updated as

$$p_l(t{+}1) = \Big[p_l(t) + \gamma\Big(\alpha_l\big(b_l(t) - b_l^*\big) + x_l(t) - c_l(t)\Big)\Big]^+, \quad (1)$$

where $(z)^+ = \max(z, 0)$, γ and α_l are positive constants, $b_l(t)$ is the aggregate *buffer occupancy*, b_l^* is the *targeted queue size*, $x_l(t)$ is the aggregate *arrival rate*, $c_l(t)$ is the *link capacity*, $x_l(t) - c_l(t)$ measures *rate mismatch*, and $b_l(t) - b_l^*$ measures *queue mismatch*. The price may increase, but the queue size should stabilise around the predefined value.

Sum Price: The central idea is to sum all the individual link prices, and then embed them within an end-to-end marking probability function. Suppose a packet travels through links $l = 1, 2, ..., L$ which have prices $p_l(t)$, then the *marking probability* $m_l(t)$ at queue l at time t is given as

$$m_l(t) = 1 - \phi^{-p_l(t)}.$$

Therefore the end-to-end marking probability for the packet is given by [1]

$$1 - \prod_{l=1}^{L}\Big(1 - m_l(t)\Big) = 1 - \phi^{-\sum_l p_l(t)}, \quad (2)$$

where $\phi > 1$ is a tunable parameter, with the default value $\phi = 1.001$. From (2), the end-to-end marking probability is high when the congestion measure along the path $\sum p_l(t)$ is large. For small $p_l(t)$, an approximation for the end-to-end marking probability is $\log_e \phi \sum_l p_l(t)$ [1].

III. Models and Analysis

In this section, we consider the following non-linear models for our stability analysis and performance evaluation: (i) a non-linear model for C-TCP [14] with Drop-Tail and small buffers, (ii) a stochastic variant of REM along with C-TCP, and (iii) the original REM queue management proposal as a continuous time non-linear dynamical system with delayed feedback.

A. C-TCP in a small buffer regime

A *many flows*, non-linear, fluid model for C-TCP has recently been proposed [14]. For notation, let w denote the current window size, $i(w)$ represent the increase of the window per positive acknowledgement, and $d(w)$ represent the decrease of the current window size per negative acknowledgement. In a small buffer regime, the congestion avoidance algorithm of C-TCP is [14]

$$w(t+1) = \begin{cases} w(t)\Big(1 + \alpha w(t)^{k-1}\Big) & \text{if no loss} \\ w(t)\big(1 - \beta\big) & \text{if loss,} \end{cases}$$

where α, k, β are protocol parameters. Thus functional forms for $i(w(t))$ and $d(w(t))$ may be represented as

$$i(w(t)) = \alpha w(t)^{k-1} \quad \text{and} \quad d(w(t)) = \beta w(t). \quad (3)$$

A generalised non-linear model for the congestion avoidance phase of transport protocols is [12]

$$\frac{dw(t)}{dt} = \Big(i\big(w(t)\big) - d\big(w(t)\big)p(t-\tau)\Big)\frac{w(t-\tau)}{\tau}, \quad (4)$$

with equilibrium

$$i(w^*) = d(w^*)p(.), \quad (5)$$

were $p(.)$ is the packet drop probability, τ is the round-trip time and w^* is the equilibrium window size. If the average window size of all flows is $w(t)$, then the average rate at which packets are sent is $x(t) = w(t)/\tau$. Different forms of $p(.)$ would be appropriate for different queue management policies.

Substituting the functional forms (3) into equation (4) we get the following non-linear model for C-TCP

$$\frac{dw(t)}{dt} = \Big(\alpha w(t)^{k-1} - \beta w(t)p(t-\tau)\Big)\frac{w(t-\tau)}{\tau}. \quad (6)$$

Assume that $p(.)$ is a function of $w(.)$. Let $w(t) = u(t)+w^*$ and linearising (4), about equilibrium, we get

$$\frac{du}{dt} = -au(t) - bu(t-\tau), \quad (7)$$

where

$$a = -\frac{i(w^*)}{\tau}(k-2)$$
$$b = \frac{i(w^*)}{\tau}\frac{w^*p'(w^*)}{p(w^*)},$$

were $p'(w^*) = dp/dw|_{w=w^*}$. We now recapitulate some results pertaining to a linear delay equation [11]. Consider the following equation

$$\frac{du}{dt} = -au(t) - bu(t-\tau), \quad (8)$$

where $a \geq 0$, $b > 0$, $b > a$ and $\tau > 0$. A sufficient condition for stability of (8) is

$$b\tau < \frac{\pi}{2},$$

and the equation will undergo a Hopf bifurcation at

$$\tau\sqrt{b^2 - a^2} = \cos^{-1}(-a/b)$$

with period $2\pi\tau/\cos^{-1}(-a/b)$.

Thus the sufficient condition for local stability for (6) is

$$i(w^*)\frac{w^*p'(w^*)}{p(w^*)} < \frac{\pi}{2}, \quad (9)$$

and the Hopf condition is

$$i(w^*)\sqrt{\left(\frac{w^*p'(w^*)}{p(w^*)}\right)^2 - (k-2)^2} = \cos^{-1}\left(\frac{(k-2)p(w^*)}{w^*p'(w^*)}\right), \quad (10)$$

with period $2\pi\tau/\cos^{-1}\left(\frac{(k-2)p(w^*)}{w^*p'(w^*)}\right)$.

As $i(w^*) = \alpha w^{*k-1}$, both the sufficient condition and the Hopf condition depend on the protocol parameters and the choice of queue management policy which is represented by $p(w^*)$. Of course, if the queue policy also had parameters in it then the stability conditions could also depend on them. The queue management policy has to chosen carefully as the stability conditions also depends on $p'(w^*)$.

The Hopf bifurcation: In non-linear systems, it is common to find the emergence of limit cycles, which may themselves be stable or unstable, with variations in model parameters. The analysis associated with the Hopf bifurcation [6] [9] is a way to analyse the emergence and stability of limit cycles bifurcating from a stable equilibrium. We now discuss the Hopf bifurcation theory informally to convey the basic intuition. Consider a system of differential equations $dx/dt = f_\eta(x)$ on $\mathbb{R}^n$, with a locally unique equilibrium x^* that is stable for $\eta < \eta_c$ and unstable for $\eta > \eta_c$. Further assume that $Df(x^*)$ and the characteristic exponents at x^* are continuous in η and the stability changes when one pair of complex conjugate characteristic exponents crosses the imaginary axis. Now let $\xi, \bar{\xi}$ be the corresponding eigenvectors of $Df(x^*)$, then at η_c the linearised system has periodic solutions lying in the plane of Re (ξ) and Im (ξ). A geometric approach (based on the central manifold theorem) shows that for η near η_c, there is a 2-manifold invariant under the flow tangent to Re (ξ) and Im (ξ). This is the place where we get to observe interesting dynamical behaviour. It is feasible to analyse the motion on this central manifold. One way to do it is by parametrising the central manifold by a single complex variable, and then essentially using the method of averaging [6] [9].

B. Small Drop-Tail buffers

When there are many flows in the system, the following fluid level approximation has been suggested as a model for small Drop-Tail buffers [12]

$$p(x) = \big(x/C\big)^B, \quad (11)$$

where B is the router buffer size, C is the service rate and recall that $w(t) = x(t)\tau$.

Using this model as a representation for the queue management policy, a sufficient condition for local stability with C-TCP and small Drop-Tail buffers is [14]

$$\alpha B w^{*k-1} < \frac{\pi}{2}.$$

Note that both *protocol parameters* and the *queue policy parameters* feature in this condition. The associated Hopf condition, with this model, is at

$$\alpha w^{*k-1}\sqrt{B^2 - (k-2)^2} = \cos^{-1}\big((k-2)/B\big),$$

where the period of the bifurcating solutions is $2\pi\tau/\cos^{-1}\big((k-2)/B\big)$. This condition also suggests that protocol parameters should not be chosen independently of the resource design parameters, at least for this choice of queue management policy.

Design Guidelines: First note that the conditions for stability do not explicitly depend on the parameter β. However, to ensure stability the protocol parameters (k and α) and the router buffer size have to be chosen carefully. We note that the parameter B could also represent the dropping/marking threshold for a queue management policy. We consider *three* cases for the choice of parameters.

Case (i): $k = 0.75, \alpha = 0.125$ (default parameter values). The sufficient condition for local stability is

$$\alpha B w^{*-0.25} < \pi/2,$$

and the necessary and sufficient condition is

$$\alpha w^{*-0.25}\sqrt{B^2 - 1.25^2} < \cos^{-1}(-1.25/B).$$

The analysis tells us that if $\alpha = 0.125$, then with $B = 8$ local stability is easy to satisfy for large window sizes. This is an attractive design consideration, and highlights the importance of designing C-TCP parameters together with network parameters.

Case (ii): $k = 0$. The sufficient condition is

$$\alpha B/w^* < \pi/2,$$

and the necessary and sufficient condition is

$$\alpha/w^*\sqrt{B^2 - 4} < \cos^{-1}(-2/B).$$

If we set $\alpha = 1/B$, the sufficient condition for stability will be satisfied if the window size is greater than $2/\pi$. Again note that the system will be stable as the equilibrium window sizes increase. Observe that with $k = 0$, we get an Additive Increase and Multiplicative Decrease (AIMD) response function which is similar to TCP Reno.

Case (iii): $k = 1$. In this case, the sufficient condition for local stability is

$$\alpha B < \pi/2,$$

and the necessary and sufficient condition is

$$\alpha\sqrt{B^2 - 1} < \cos^{-1}(-1/B).$$

With this choice of k, the dependence of the equilibrium window on the stability conditions are removed. Then again, as above, choosing $\alpha = 1/B$ ensures that stability will not be violated as the window size increases. Figure 1 shows the relationship between α and B to ensure stability. Thus, if we choose $B = 15$, then $\alpha = 1/15$ ensures that the necessary and sufficient condition for stability will also be satisfied.

A key conclusion is that, for stability, C-TCP parameters will possibly have to be jointly designed with network parameters like router buffer sizes (or thresholds for queue management schemes). We now consider some refined models which could also act as queue management policies.

C. A stochastic REM variant

We first consider a model for a *stochastic marking function* and then a *stochastic variant of REM* with C-TCP. Suppose that the work load arriving at the resource over a time-period δ is Gaussian, with mean $x\delta$ and variance $x\delta\sigma^2$, and that a packet is marked on arrival if the workload present in the queue is larger than a threshold B. Then from the stationary distribution for a reflected Brownian motion [4], we get [8]

$$p(x) = \exp\left(\frac{-2B(C-x)}{x\sigma^2}\right). \qquad (12)$$

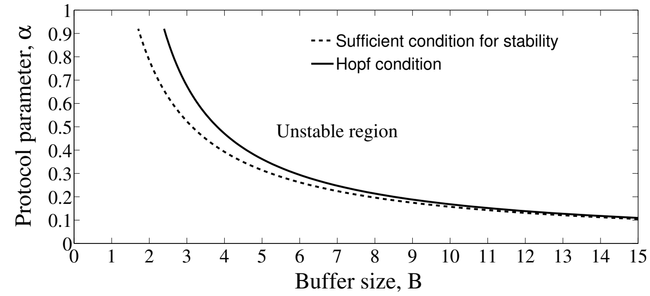


Fig. 1: Stability chart for α versus B in Drop-Tail with C-TCP for the protocol parameter $k = 1$.

Thus $p'(x)$ can be obtained and simplified as

$$p'(x) = p(x)\frac{2B}{\sigma^2}\frac{C}{x^2}.$$

With the aforementioned queue management policy, a *sufficient* condition for local stability of C-TCP is

$$\alpha w^{*k-1}\frac{2B}{\sigma^2} < \frac{\pi}{2}.$$

There are a few interesting design considerations that arise. First is that the larger the threshold for marking/dropping packets, the greater the chances for violating the stability conditions. Interestingly, variability of the traffic at the packet-level also influences stability. Once again, we observe the interplay between protocol and resource management parameters to ensure stability. This clearly highlights the need for the joint design of transport protocols embedded in end-systems and algorithms that are usually implemented in network routers.

Choosing the C-TCP parameter α to be the inverse of the threshold marking parameter B does help to ensure local stability. Such a relationship was also readily deduced from the previous small buffer Drop-Tail approximation.

We now consider some stability properties of a stochastic variant of REM with C-TCP. Suppose we mark a packet with a probability of $1 - \exp(-qW)$, where W is the workload already present, then the probability that a packet is marked can be shown to be [8]

$$p(x) = \frac{qx\sigma^2}{qx\sigma^2 + 2(C-x)}.$$

We can deduce that a *sufficient* condition for local stability of C-TCP, with the stochastic variant of REM, is

$$\alpha w^{*k-1}\frac{2}{q\sigma^2} < \frac{\pi}{2}.$$

Note the relation between $1/q$ and the threshold B in the above stochastic marking function. The relation between these two is readily seen as randomly choosing the threshold level upon each arrival, according to an exponential distribution with mean q^{-1}. Thus, even with refined models for the queue, the basic insight that queue policy thresholds can readily impact stability still holds.

Following on from the analysis, we highlight two points. First, the precise choice of buffer size, or the thresholds used in queue management schemes, can readily influence stability. Using packet-level simulations, later in the paper, this point will be corroborated. Then, we wish to reiterate the importance of the joint design of transport protocols, and their parameters, along with resource design functions, and their parameters, in providing stability and performance.

In the models considered thus far, we did not model the queue dynamics explicitly. Instead, the focus was on an operating regime where the transport protocol was acting to control the *distribution* of the queue size. The resource design functions were considering a regime where stochastic effects, relevant on a queuing time-scale, form an essential component of the system dynamics. Such stochastic effects were assumed to be averaged out over the time-scale of round-trip times.

We now analyse the original proposal of REM [1], where we model the instantaneous queue size along with a dynamical representation of the price which represents the congestion measure at the resource.

D. Non-linear model for REM

We now consider the original REM proposal; however, as a continuous time non-linear model with time delayed feedback. Consider the following non-linear model for REM:

$$\begin{aligned} \frac{dp_l(t)}{dt} &= \gamma\Big(\alpha_l\big(b_l(t)-b_l^*\big)+x_l(t-\tau)-CI_{\{p_l(t)>0\}}\Big) \\ \frac{db_l(t)}{dt} &= x_l(t-\tau)-C, \qquad \text{if } b_l>0 \end{aligned} \tag{13}$$

where $x_l(t)=\mathcal{D}(p_l(t))$ with $\mathcal{D}(p_l)$, $p_l \geq 0$, a non negative, continuous, decreasing demand function. Here $b_l(t)$ is the aggregate *buffer occupancy* at the queue, b_l^* is the *target queue size*, p_l represents the *price* as the congestion measure which takes the notion of *clear buffer* and *match rate* into account. C is the total link service rate, γ and α_l are positive REM constants, and τ is the round-trip time of the users. Note that the queue dynamics are represented by an integrator, and not by the distribution of the queue size.

Let p_l^* and b_l^* denote the equilibrium, and let $p_l(t)=u_1(t)+p_l^*$ and $b_l(t)=u_2(t)+b_l^*$. We now linearise the non-linear system (13) to obtain

$$\begin{aligned} \frac{du_1(t)}{dt} &= \gamma\alpha_l u_2(t)+\gamma\mathcal{D}' u_1(t-\tau) \\ \frac{du_2(t)}{dt} &= \mathcal{D}' u_1(t-\tau), \end{aligned} \tag{14}$$

where $\mathcal{D}'=\mathcal{D}'(p_l^*)<0$. The characteristic equation, obtained by looking for exponential solutions, associated with the linearised system is

$$\lambda^2+\zeta_1\lambda e^{-\lambda\tau}+\zeta_2 e^{-\lambda\tau}=0, \tag{15}$$

where $\zeta_1=-\gamma\mathcal{D}'>0$, $\zeta_2=-\gamma\alpha_l\mathcal{D}'>0$.

Consider the associated characteristic equation

$$\lambda^2+\zeta_1\lambda e^{-\lambda\tau}+\zeta_2 e^{-\lambda\tau}=0,$$

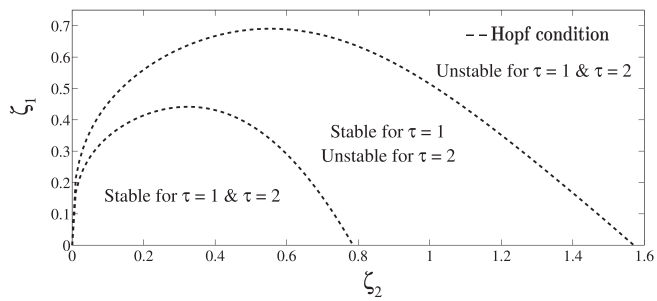


Fig. 2: Local stability chart for a non-linear model of REM, equation (13), where $\zeta_1=-\gamma\mathcal{D}'>0$ and $\zeta_2=-\gamma\alpha_l\mathcal{D}'>0$.

where $\zeta_1>0$, $\zeta_2>0$ and $\tau\geq 0$. We start by assuming that $\lambda=i\omega$, $\omega>0$ is a root of the above characteristic equation for some τ. We get

$$\begin{aligned} -\omega^2+\zeta_1\omega\sin(\omega\tau)+\zeta_2\cos(\omega\tau) &= 0 \\ \zeta_1\omega\cos(\omega\tau)-\zeta_2\sin(\omega\tau) &= 0, \end{aligned} \tag{16}$$

which yields

$$\omega^4-\zeta_1^2\omega^2-\zeta_2^2=0,$$

giving

$$\omega=\sqrt{\frac{\zeta_1^2+\sqrt{\zeta_1^4+4\zeta_2^2}}{2}}.$$

We need to satisfy the transversality condition of the Hopf bifurcation, i.e. we need to determine the sign of the derivative of Re$(\lambda(\tau))$ at the points where $\lambda(\tau)$ is purely imaginary.

From the characteristic equation we have

$$\big(2\lambda+[\zeta_1-\tau(\zeta_1\lambda+\zeta_2)]\,e^{-\lambda\tau}\big)\frac{d\lambda(\tau)}{d\tau}=\lambda(\zeta_1\lambda+\zeta_2)e^{-\lambda\tau}.$$

Instead of $d\lambda(\tau)/d\tau$, we can analyse $\big(d\lambda(\tau)/d\tau\big)^{-1}$. Thus

$$\left(\frac{d\lambda(\tau)}{d\tau}\right)^{-1}=\frac{2\lambda e^{\lambda\tau}+\zeta_1}{\lambda(\zeta_1\lambda+\zeta_2)}-\frac{\tau}{\lambda},$$

and

$$\lambda^2 e^{\lambda\tau}=-(\zeta_1\lambda+\zeta_2).$$

Hence

$$\begin{aligned} \text{sign}\left(\frac{\text{Re}\big(d\lambda(\tau)\big)}{d\tau}\right)\Big|_{\lambda=i\omega} &= \text{sign}\left(\text{Re}\left(\left(\frac{d\lambda(\tau)}{d\tau}\right)^{-1}\right)\right)\Big|_{\lambda=i\omega} \\ &= \text{sign}\left(\text{Re}\left(\frac{-2\lambda}{\lambda^3}+\frac{\zeta_1}{\lambda(\zeta_1\lambda+\zeta_2)}\right)\right)\Big|_{\lambda=i\omega} \\ &= \text{sign}\left(\frac{2}{\omega^2}-\frac{\zeta_1^2}{\zeta_2^2+\zeta_1^2\omega^2}\right) \\ &= \text{sign}\left(2\zeta_2^2+\zeta_1^2\omega^2\right) \\ &> 0. \end{aligned}$$

In our equation, with $\zeta_2>0$, only one imaginary root exists. Hence the only crossing of imaginary axis is from the left to the right as τ increases. From (16) we get

$$\tau_c = \frac{\theta}{\omega} \text{ where } 0 \leq \theta \leq 2\pi \text{ and}$$
$$\cos(\theta) = \frac{\zeta_2 \omega^2}{\zeta_2^2 + \zeta_1^2 \omega^2}$$
$$\sin(\theta) = \frac{\zeta_1 \omega^3}{\zeta_2^2 + \zeta_1^2 \omega^2}.$$

We now collect the above analysis in the form of a Theorem.

Theorem: The non-linear model for REM, equation (13), is locally asymptotically stable for $\tau = 0$, is stable until $\tau < \tau_c$, undergoes a Hopf bifurcation at $\tau = \tau_c$ and is unstable for $\tau > \tau_c$.

This shows that it would be hard to ensure local stability of REM in networks where the delays may be large. See Figure 2 for the associated stability chart. The condition for stability is also influenced by REM parameters, and so variations in REM parameters may violate the stability conditions.

The models analysed so far have represented fluid approximations to the underlying packet-based system. In the next section, we conduct packet-level simulations to complement some of the analytical results.

IV. Packet-level Simulations

We now conduct some packet-level simulations using the Network Simulator (NS2) [17]. The topology we use is a single bottleneck, with a capacity of 100 Mbps. We use the default parameter values of REM: $\gamma = 0.001$, $\alpha_l = 0.1$, and $\phi = 1.001$, and the default parameter values of Compound. We use the REM scripts available in NS2 version 2.35.

In subsection A, we conduct simulations of C-TCP, with Drop-Tail and REM, in large and small buffers. The *large buffer* regime (bandwidth-delay rule) translates into 2084 packets. This is because the link capacity is 100 Mbps, and in practice a value of 250 ms is used irrespective of the actual delays of the TCP flows. In the *small buffer* regime, parameters like the link capacity, feedback delays and the number of users do not influence the choice of buffer size. In this regime, we vary buffers in the range of 15 to 100 packets, as guided by our stability analysis. The traffic consists of 60 long-lived C-TCP flows, each with a 2 Mbps link feeding into the bottleneck. The packets have a size of 1460 bytes.

In subsection B, we explore the ability of REM to control queue sizes and maintain high link utilisation with C-TCP. We use the large buffer regime as this is widely deployed. In this simulation, the traffic mix is a combination of FTP, UDP and HTTP flows. The traffic consists of 55 FTP flows, each with a 2 Mbps link, 8 UDP flows contributing a total of 8 Mbps, and 50 short-lived (HTTP) flows which arrive and depart randomly and contribute an average of 2 Mbps. The packet sizes vary between $1000 - 1500$ bytes. The REM parameters are chosen to target a desired queue size of 15 and 100 packets. The average round-trip time of the FTP flows present is varied between 10 and 200 ms.

A. C-TCP, REM & Drop-Tail, Small & Large buffers

With *small buffers*, as router buffer sizes vary from 15 to 100 packets, with both Drop-Tail and REM we see stable limit cycles in the queue size. See Figure 3(a) (buffer size = 15 packets) and Figure 3(b) (buffer size = 100 packets). Our analysis with Drop-Tail predicted the loss of local stability, via a Hopf, as buffers got larger; see Figure 1 in Section III A. In such a small buffer regime, the transport layer protocol acts to control the *distribution of the queue size* and thus it is plausible that REM effectively degenerates into Drop-Tail. With small buffers, the simplicity of Drop-Tail appears to be preferable to the more involved REM.

With *large* buffers, REM finds it difficult to control the queue size to a desired value, especially with larger round-trip times; see Figure 3(c). With large delays we witness the formation of limit cycles in the queue size, which also hurt link utilisation. Thus the twin objectives of REM of controlling the queue size to a desired value and maintaining a high link utilisation need not be satisfied. Drop-Tail, with large buffers, is expected to have full link utilisation but this comes at the cost of enormous queuing delays.

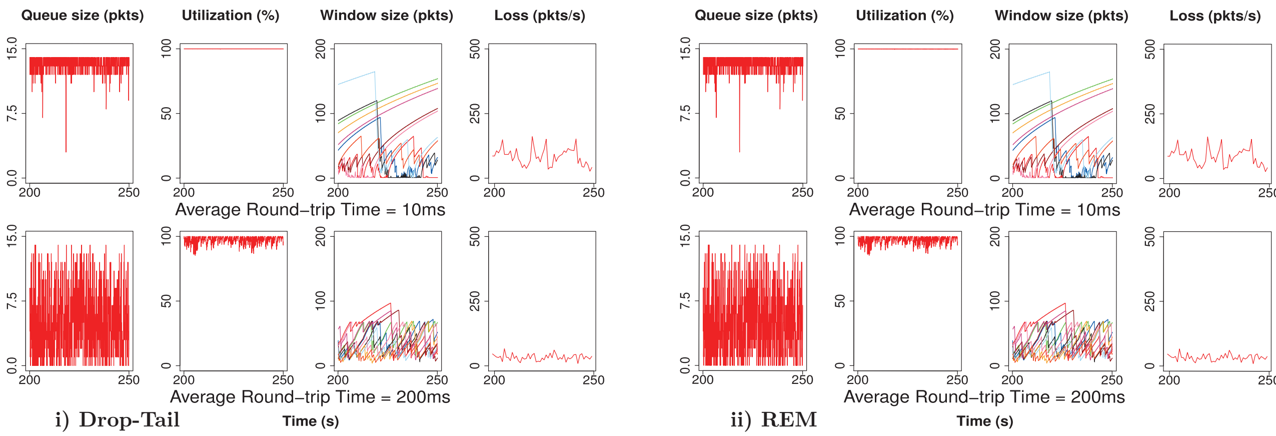
B. C-TCP, REM, Large buffers: varying queue thresholds

Large buffers are widely deployed, and thus a queue management policy should be able to ensure low latency, stability of queue sizes, and decent link utilisation. We conduct two simulations, where the desired queue size in REM is specified to be 15 and 100 packets respectively; see Figure 4(a) (15 packets) and Figure 4(b) (100 packets). The traffic consists of a combination of 55 FTP flows, each with a 2 Mbps link, 8 UDP flows (contributing 8 Mbps) and 50 HTTP flows (contributing a total of 2 Mbps). With smaller round-trip times for the FTP flows, we observed full link utilisation and no evidence of instabilities in the queue size. With larger round-trip times, the queue sizes developed limit cycles which in turn hurt link utilisation. Even with the presence of UDP and HTTP flows, the emergence of limit cycles was rather evident. So while latency could be reduced, link utilisation would be adversely impacted. We observe that the simple model of REM, analysed in Section III D, alerted us to the fact that large delays could be destabilising to the system.

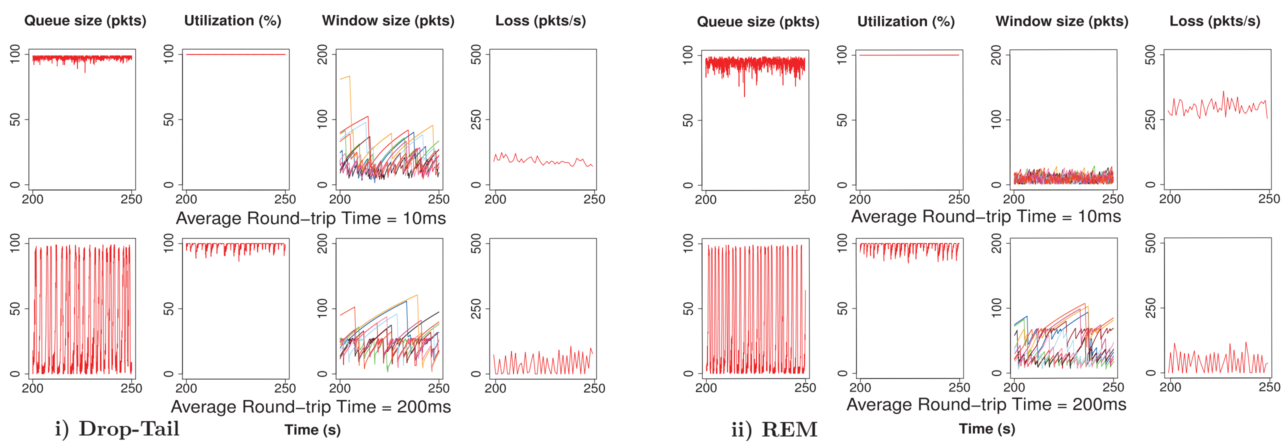
Clearly, the combination of Compound, queue management schemes and buffer size impacts network performance. Buffer size appears to play a central role, as variations in it influence both latency and queue stability. With the analysis and simulations thus far, there appears to be no clear advantage of using REM in either small or large buffers. However, a study of the stochastic variant of REM merits further investigation.

V. Conclusions

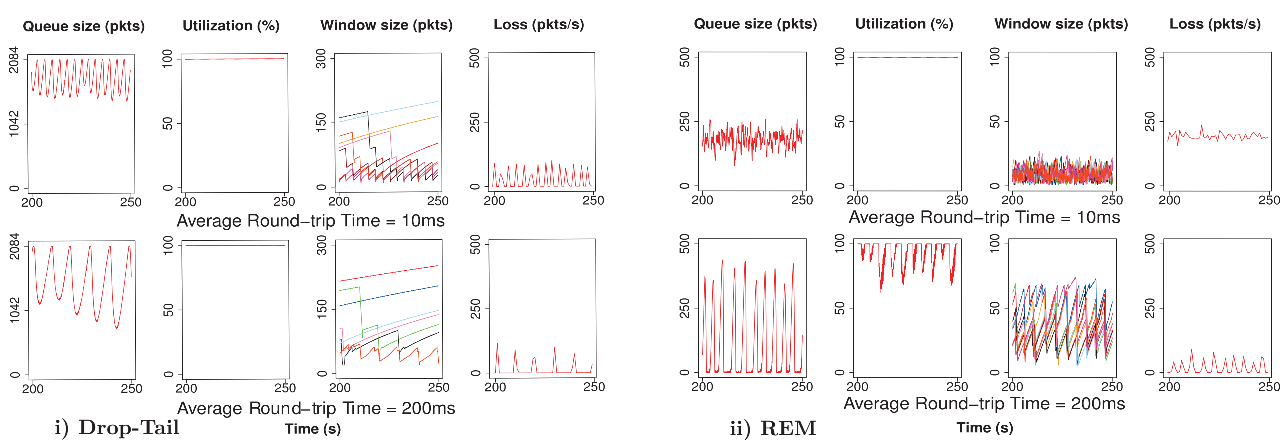
The presence of queuing delay in the Internet is detrimental to the development of low latency applications. Recent papers have the following strongly worded one line abstracts: "Networks without effective AQM may again be vulnerable to congestion collapse" [3], and "A modern AQM is just one piece of the solution to bufferbloat" [10]. These papers



(a) With 15 packet buffers, the queues with C-TCP and Drop-Tail do not exhibit limit cycles. This is in agreement with the stability analysis presented in Section III A. Figure 1 shows the relationship between α (a C-TCP parameter) and the buffer size to maintain local stability. With small buffers, REM effectively degenerates into Drop-Tail and thus we observe similar performance. The target queue size for REM was set to 10 packets.

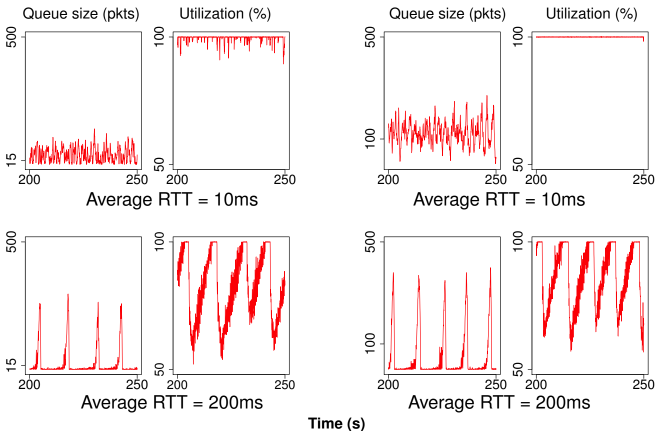


(b) With 100 packet buffers, the stability analysis in Section III A shows that larger buffer sizes could destabilise the system. With both queue policies, note the emergence of limit cycles in the queue size with larger feedback delays. These non-linear oscillations can start to hurt link utilisation. The target queue size for REM was set to 15 packets.



(c) With large buffers, both queue policies show degraded performance. Drop-Tail has full utilisation, but no control over latency. With larger feedback delays, with REM, the limit cycles also hurt link utilisation. The target queue size for REM was set to 100 packets.

Fig. 3: Performance of C-TCP, with Drop-Tail and REM, in a small and a large buffer regime. The traffic consists of 60 long-lived C-TCP flows, each with a 2 Mbps link, over a single bottleneck with capacity 100 Mbps. The average round-trip time was varied between 10 ms and 200 ms. The REM and Compound parameters were maintained at their default values.



(a) Target queue size = 15 (b) Target queue size = 100

Fig. 4: REM, in large buffers, with a mix of FTP, UDP and HTTP flows. The left panel has a target queue size of 15 packets and the right panel has a target of 100 packets. In both cases, larger feedback delays give rise to stable limit cycles in the queue size which clearly also hurt link utilisation. Note that larger threshold values tend to have a destabilising effect.

highlight the importance of the latency problem, that the problem is expected to get worse, and that as yet there has been no acceptable resolution.

We conducted a performance evaluation of Compound TCP coupled with Drop-Tail and REM operating over a large (bandwidth-delay product) and a small buffer regime. Our study was based on a combination of analysis, of numerous fluid models, and packet-level simulations using NS2. The analysis used queuing models coupled with control and bifurcation theory. The performance metrics considered were stability of queue sizes, latency, link utilisation and packet loss.

We first analysed a non-linear model for C-TCP along with small Drop-Tail buffers. Stability was found to be very sensitive to the precise choice of buffer size. Variations in the buffer size could readily induce a Hopf bifurcation. The analysis highlighted the relationships between C-TCP parameters and buffer sizes to maintain local stability. With small buffers we can get low latency and stable queue sizes, but have no explicit control over link utilisation. The analytical insights, for C-TCP over small Drop-Tail buffers, were corroborated using packet-level simulations. A key performance metric turned out to be stability of the queue size. Formation of limit cycles in queues causes packet losses to occur in bursts, which synchronises TCP flows, and this in turn leads to a loss in link utilisation.

The analysis of Compound with a stochastic variant of REM provided two key observations. First, that smaller REM thresholds for marking/dropping packets would help to maintain stability. Second, that for system stability both C-TCP and REM parameters should be jointly designed.

We finally modelled the original REM proposal as a set of continuous time, non-linear, equations with delayed feedback. For this model, we derived the necessary and sufficient conditions for local stability. The analysis showed that variations in delays, or in REM parameters, can destabilise the system. With large feedback delays, we explicitly characterised the loss of local stability to occur via a Hopf bifurcation.

With *large buffers*, with C-TCP, Drop-Tail is unable to ensure low latency whereas REM was unable to stabilise the queue sizes. In *small buffers*, REM does not appear to offer any clear advantage over Drop-Tail. In this regime we also suggested guidelines for the C-TCP parameters to be designed in conjunction with Drop-Tail router buffers to ensure stability.

Avenues for further research

It is important to enhance the mathematical understanding of the models. One certainly needs to analytically characterise the stability of the limit cycles that were predicted in theory, and observed in packet-level simulations. The performance analysis of Compound with other queue management proposals, in different buffer sizing regimes, will help in building a comprehensive understanding of network performance.

ACKNOWLEDGEMENTS

The authors acknowledge the IU-ATC project funded by the Department of Science and Technology (DST) Government of India and the UK EPSRC Digital Economy Programme.

REFERENCES

[1] S. Athuraliya, S.H. Low, V.H. Li and Q. Yin, "REM: Active Queue Management", *IEEE Network*, vol. 15, pp. 48–53, 2001.
[2] S. Floyd and V. Jacobson, "Random Early Detection gateways for congestion avoidance", *IEEE/ACM Transactions on Networking*, vol. 1, pp. 397–413, 1993.
[3] J. Gettys and K. Nichols, "Bufferbloat: dark buffers in the Internet", *Communications of the ACM*, vol. 55, pp. 57–65, 2012.
[4] J.M. Harrison, *Brownian Motion and Stochastic Flow Systems*. New York: Krieger, 1985.
[5] M. Hassan and R. Jain, *High Performance TCP/IP Networking*. Pearson Prentice Hall, 2004.
[6] B.D. Hassard, N.D. Kazarinoff and Y. Wan, *Theory and Applications of Hopf Bifurcation*. Cambridge, U.K.: Cambridge University Press, 1981.
[7] C.V. Hollot, V. Misra, D. Towsley and W.B. Gong, "On designing improved controllers for AQM routers supporting TCP flows", in *Proceedings of IEEE INFOCOM*, 2001.
[8] F.P. Kelly, "Models for a self-managed Internet", *Philosophical Transactions of the Royal Society A*, vol. 358, pp. 2335–2348, 2000.
[9] Y.A. Kuznetsov, *Elements of Applied Bifurcation Theory*. Applied Mathematical Sciences: Springer–Verlag, 2004.
[10] K. Nichols and V. Jacobson, "Controlling queue delay", *Communications of the ACM*, vol. 55, pp. 42–50, 2012.
[11] G. Raina, "Local bifurcation analysis of some dual congestion control algorithms", *IEEE Transactions on Automatic Control*, vol. 50, pp. 1135–1146, 2005.
[12] G. Raina, D. Towsley and D. Wischik, "Part II: Control theory for buffer sizing", *ACM SIGCOMM Computer Communication Review*, vol. 35, pp. 79–82, 2005.
[13] G. Raina and D. Wischik, "Buffer sizes for large multiplexers: TCP queueing theory and instability analysis", in *Proceedings of Next Generation Internet Networks*, 2005.
[14] P. Raja and G. Raina, "Delay and loss-based transport protocols: buffer-sizing and stability", in *Proceedings of International Conference on Communication Systems and Networks*, 2012.
[15] K. Tan, J. Song, Q. Zhang and M. Sridharan, "A Compound TCP approach for high-speed and long distance networks", in *Proceedings of IEEE INFOCOM*, 2006.
[16] D. Wischik and N. McKeown, "Part I: Buffer sizes for core routers", *ACM SIGCOMM Computer Communication Review*, vol. 35, pp. 75–78, 2005.
[17] The Network Simulator (NS2). [Online]. Available: http://nsnam.isi.edu/nsnam/index.php/User_Information