A Study of the Applicability of a Scaling Hypothesis

Rong Pan*, **Balaji Prabhakar**†, **Konstantinos Psounis*** and **Mayank Sharma***

* Department of Electrical Engineering, Stanford University
email: {rong,kpsounis,msharma}@stanford.edu

† Departments of Computer Science and Electrical Engineering, Stanford University
e-mail: balaji@stanford.edu

Abstract

Differential equations and fluid models have been increasingly used with significant success to model network behavior and the interaction of TCP-flows with Active Queue Management (AQM) schemes. A crucial assumption for these limiting approximations to work is that packet drops be distributed as a Poisson process. It has also been demonstrated [2] that in a wide variety of settings networks satisfy a scaling behavior, in that many important performance metrics remain invariant when the network is suitably scaled. In this paper we present the results of our study on how the choice of parameters of AQM schemes influence the validity of the scaling hypothesis and the ability of fluid models to predict network dynamics. We find that the fidelity of an approximation method is directly related to how *clumped* the packet drops it induces are, which in turn is related to the deviation of the packet-drop process from that of a Poisson process.

1 Introduction

The complexity and size of the Internet presents a problem to those attempting to model traffic behavior and design schemes to ensure a fair and efficient use of the network. A lot of recent work has focused on using differential equations to model network dynamics, in particular the interaction between TCP-flows and AQM schemes [1, 3, 4, 5]. This approach has met with remarkable success and has provided a handle by which the behavior of the Internet can be better understood. A by-product of this framework is that it allows one to view the network from a control systems perspective and analyze the stability/performance of various congestion control mechanisms.

Recently, a scaling hypothesis has been proposed in [2] which states that if a network controlled by a variety of AQM schemes is suitably scaled, then performance measures such as queueing delay and drop probability are left virtually unchanged. In that context too, the differential equation model for a network with TCP-like flows and standard AQM schemes is used to explain why such a scaling behavior is exhibited.

Differential equation models developed in [1] rely on the assumption that packet losses for each flow are described by a Poisson process of some rate. This assumption is crucial to establishing the validity of these deterministic models. For example, the *DropTail* AQM scheme, which simply drops packets that arrive at a full buffer, causes a number of consecutive packets to be lost. This makes the packet-drop process bursty and correlated. Separately, we observe in this paper that if *DropTail* is used as the AQM scheme, then indeed both the differential equation model and the scaling behavior fail. It is therefore a worthwhile exercise to study whether or not conventional AQM schemes drop packets according to a Poisson-like distribution.

While obtaining the exact distribution for the packet-loss process is an intractable problem, we can characterize it as being close or far from Poisson in nature by utilizing properties of the Poisson process. It is well known that a Poisson process tends to avoid clumps; i.e the probability that the number of points in a fixed interval is large is exponentially small. Thus one would expect that if an AQM scheme were to drop packets in clumps, thus inducing correlations between successive packet drops, then the packet-drop process will be far from Poisson. For such an AQM scheme we will expect both the differential equation model and the scaling hypothesis to be inaccurate. A way to achieve this objective is to gradually increase the rate at which the AQM scheme drops packet as a function of some parameter until a point are lost faster than the rate at which they leave the router. when successive packets get dropped with high probability leading to clumps in the packet-drop process.

Our simulations have shown that this is indeed the case and, in fact, there is a cut-off point beyond which if we change the parameters of the AQM scheme, the fidelity of these methods in approximating the behavior of the real network is lost. We will define two parameterized variants of the Random Early Detection (RED) AQM scheme in Section 2 and present our findings in Section 3. In Section 4 we will perform some analysis to explain why after the variation of parameters have pushed an

AQM scheme into the regime where it starts dropping packets clumpily, the previously mentioned models fail to predict network behavior. Finally we will conclude in Section 5.

2 AQM schemes

A significant number of routers in the Internet today use RED for the purpose of congestion control. The following two equations define how RED works, and together they specify the drop (or marking) probability. RED maintains a moving average q_a of the instantaneous queue size q; and q_a is updated whenever a packet arrives, according to the rule

$$q_a := (1-w)q_a + wq,$$

where the parameter w determines the size of the averaging window. The average queue size determines the drop probability $p(q_a)$, according to the equation

$$p(q_a) = \begin{cases} 0 & : \quad q_a < min_{th} \\ pmax\left(\frac{q_a - min_{th}}{max_{th} - min_{th}}\right) & : \quad min_{th} \le q_a < max_{th} \\ 1 & : \quad q_a > max_{th} \end{cases} \tag{1}$$

where p_{max}, min_{th} and max_{th} are design parameters chosen to meet some performance guidelines.

As the current design of RED stands, packet drops are relatively well spaced out. For our experiments, we wish to modify the drop probability function used in RED in a way so as to cause it to drop packets more and more burstily, with the burstiness of drops increasing as a function of some parameter. We will achieve this end by altering RED so as to increase the rate of packet drops.

Specifically, we use the following two variants of RED:

1. *SlopeRED*: We fix max_{th} and vary min_{th} from some nominal value up to max_{th} in steps of size δ. Changing min_{th} in this manner has the effect of increasing the slope of the drop probability function thus inducing a larger number of drops. The parameter in SlopeRED is the slope of the drop function. When $min_{th} = max_{th}$, SlopeRED becomes *DropTail* (if we identify max_{th} with the physical buffer size).
2. *PowerRED*: We define the drop function as:

$$p(q_a) = \begin{cases} 0 & : \quad q_a < min_{th} \\ pmax\left(\frac{q_a - min_{th}}{max_{th} - min_{th}}\right)^n & : \quad min_{th} \le q_a < max_{th} \\ 1 & : \quad q_a \ge max_{th} \end{cases} \tag{2}$$

   where n is an exponential parameter. Note that here again, as n grows larger, PowerRED tends to *DropTail*.

For the scaling experiments, the parameters p_{max}, min_{th}, max_{th} and w will be scaled in a way as specified in [2]. We will multiply min_{th} and max_{th} by α. Recall that we are multiplying the buffer size by α: thus min_{th} and max_{th} are fixed to be a constant fraction of the buffer size. We will keep p_{max} fixed at 10%, so that the drop probability is kept under 10% as long as the buffer is slightly congested. The averaging parameter w will be multiplied by α^{-1}.

3 Simulations

3.1 The Basic Setup

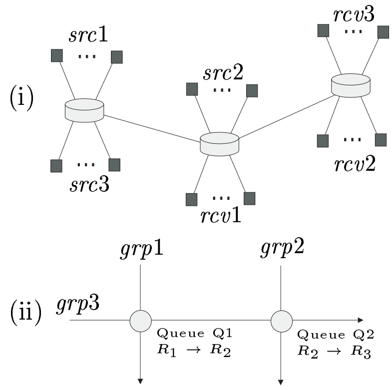

Figure 1: Basic network topology (i) physical view and (ii) logical view

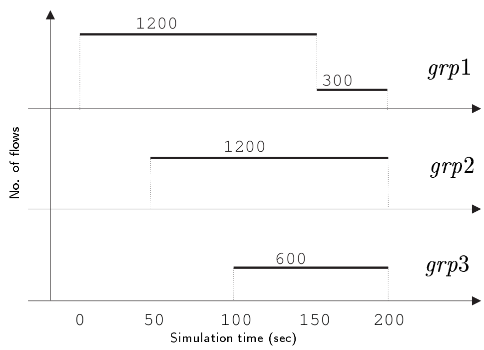

Figure 2: Traffic Patterns

We consider two congested links in tandem, as shown in Figure 1. There are three routers, $R1$, $R2$ and $R3$; and three groups of flows, *grp*1, *grp*2, and *grp*3, with group i connecting sources in *srci* to receivers in *rcvi*. The link speeds are 100Mbps and the buffers can hold 8000 packets. The RED parameters are $min_{th} = 1000$, $max_{th} = 3000$ and $w = 0.000005$. For the flows: *grp*0 consists of 1200 TCP flows each having a propagation delay of 150ms, *grp*1 consists of 1200 TCP flows each having a propagation delay of 200ms, and *grp*2 consists of 600 TCP flows each having a propagation delay of 250ms. The flows switch on and off as shown in the timing diagram of Figure 2. Note that 75% of *grp*0 flows switch off at time 150s.

This network is scaled-down by a factor $\alpha = 0.1$ and the parameters are modified as described above. Figures 3 and 4 depict the excellent match between the average queueing delays experienced at queues 1 and 2 respectively, before and after scaling.

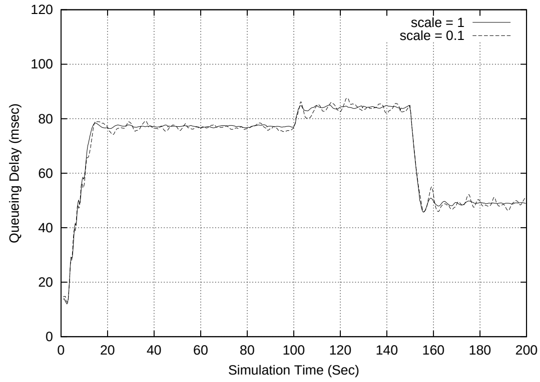

Figure 3: Basic Setup: Average Queuing Delay at Q1

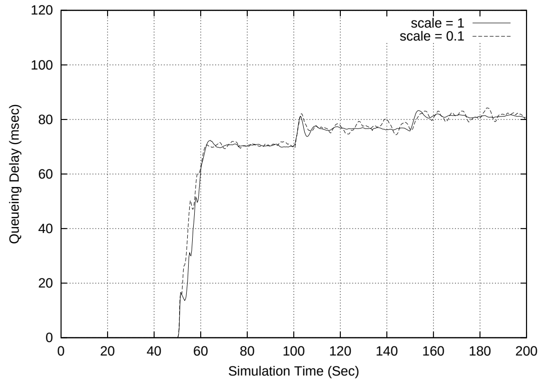

Figure 4: Basic Setup: Average Queuing Delay at Q2

3.2 Experiment 1: SlopeRED

For this experiment we kept $max_{th} = 3000$ and increased min_{th} from 1000 in steps of $\delta = 250$ up to max_{th}. Figures 5 and 6 show us that the behavior of the scaled system matches well with that of the original network when $min_{th} = 1500$. Thus the scaling hypothesis appears to hold for that value of min_{th}.

But we notice in Figures 7 and 8 that for $min_{th} = 2750$, the scaled system has queueing delays that differ substantially from those experienced in the original network. This we claim is evidence of the fact that as the slope of SlopeRED increases and the packet drops become more clumpy, the scaling behavior no longer holds.

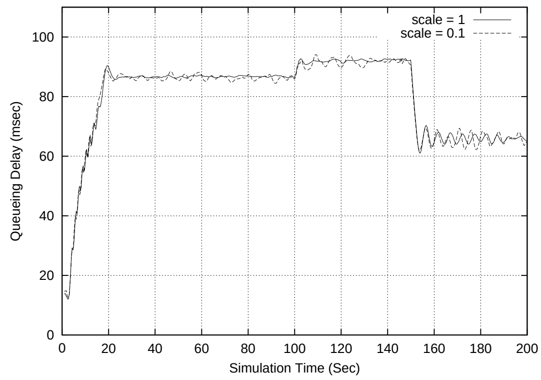

Figure 5: $min_{th} = 1500$: Average Queuing Delay at Q1

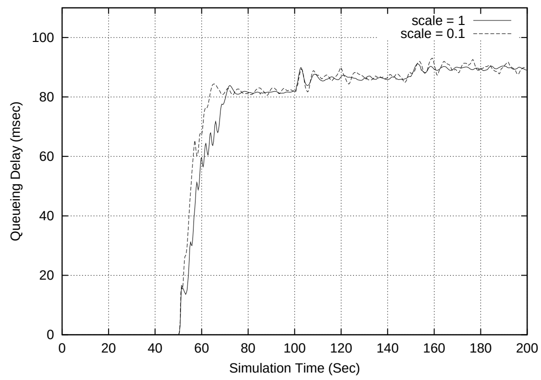

Figure 6: $min_{th} = 1500$: Average Queuing Delay at Q2

3.3 Experiment 2: PowerRED

Here we fix the values of max_{th} and min_{th} at 3000 and 1000 respectively. As in SlopeRED, we notice here too (Figures 9-12) that the scaled system's behavior deviates significantly from that of the original one as the exponent is increased, in fact, for as low an n as 2. Thus we do not even need to take $n \to \infty$ to observe the degradation in fidelity that *DropTail* displayed.

4 Differential Equation Models

4.1 Scaling of solutions

In [2], the recently proposed theoretical fluid model of TCP/RED [1] and the associated differential equations have been used to explain the observed scaling behavior. We reproduce the key concepts here.

Consider N flows sharing a link of capacity C. Let $W_i(t)$ and $R_i(t)$ be the window size and round-trip time of flow i at time t. Here $R_i(t) = T_i + q(t)/C$, where T_i is the propagation delay and $q(t)$ is the queue size at time t. Let $p(t)$ be the drop probability at time t, and $q_a(t)$ the average queue size used by RED.

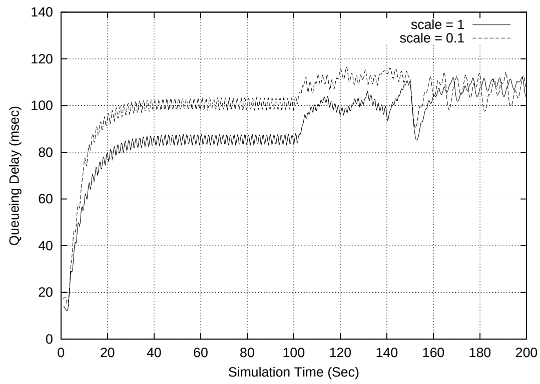

Figure 7: min_{th} = 2750: Average Queuing Delay at Q1

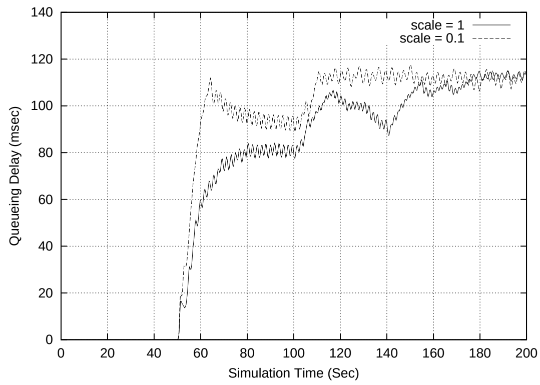

Figure 8: min_{th} = 2750: Average Queuing Delay at Q2

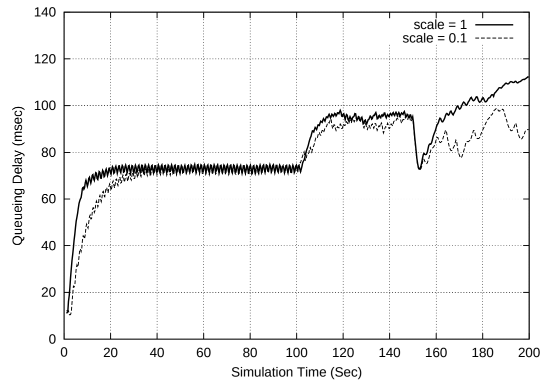

Figure 9: n = 2: Average Queuing Delay at Q1

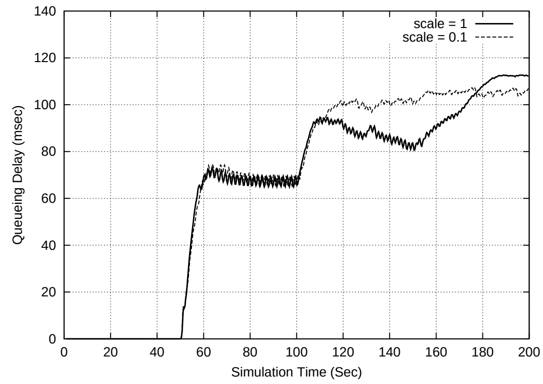

Figure 10: n = 2: Average Queuing Delay at Q2

The fluid model describes how these quantities evolve; or rather, since these quantities are random, the fluid model describes how their expected values evolve. Let $\bar{X}$ be the expected value of random variable X. Then the fluid model equations are these:

$$\begin{aligned}
\frac{d\bar{W}_i(t)}{dt} &= \frac{1}{R_i(\bar{q}(t))} - \frac{\bar{W}_i(t)\bar{W}_i(t-\tau_i)}{1.5R_i(\bar{q}(t-\tau_i))}\bar{p}(t-\tau_i) \\
\frac{d\bar{q}(t)}{dt} &= \sum_{i=1}^{N} \frac{\bar{W}_i(t)}{R_i(\bar{q}(t-\tau_i))} - C \\
\frac{d\bar{q}_a(t)}{dt} &= \frac{\log(1-w)}{\delta}\bar{q}_a(t) - \frac{\log(1-w)}{\delta}\bar{q}(t) \\
\bar{p}(t) &= p_{RED}(\bar{q}_a(t))
\end{aligned}$$

where $\tau_i = \tau_i(t)$ solves $\tau_i(t) = R_i(\bar{q}(t-\tau_i(t)))$, δ is the average packet inter-arrival time, and p_{RED} is as in Equation (1). The accuracy of the fluid model is shown in Figure 13 were we compare the numerical solution of the fluid model equations with the simulation results[1].

Remarks. While the applicability of these equations is not yet fully understood [6], [1] indicates that empirically they are reasonably accurate. Also, note that we have the constant 1.5 in Equation (3), not 2 as in [1]. This change improves the accuracy of the fluid model, for reasons elaborated on in [2]. Finally, note that while these equations describe a single link, the extension to networks is straightforward, and is given in [1].

Returning to the differential equations, suppose we have a solution to these equations

$$\big(\bar{W}_i(\cdot), \bar{q}(\cdot), \bar{q}_a(\cdot), \bar{p}(\cdot)\big).$$

Now, suppose the network is scaled and denote by C', N', etc the parameters of the scaled system. When the network is scaled, the fluid model equations change, and so the solution changes. Let $\big(\bar{W}'_i(\cdot), \bar{q}'(\cdot), \bar{q}'_a(\cdot), \bar{p}'(\cdot)\big)$ be the solution of the scaled system. It is shown in [2] that, in fact,

$$\big(\bar{W}'_i(\cdot), \bar{q}'(\cdot), \bar{q}'_a(\cdot), \bar{p}'(\cdot)\big) = \big(\bar{W}_i(\cdot), \alpha\bar{q}(\cdot), \alpha\bar{q}_a(\cdot), \bar{p}(\cdot)\big).$$

Thus the queueing delay $\bar{q}'/C' = \alpha\bar{q}/\alpha C$ is identical to that in the unscaled system. Also the drop probability is the same in each case ($\bar{p}(t) = \bar{p}'(t)$). We therefore have theoretical support for the scaling hypothesis.

Now consider PowerRED: we will show that $\bar{p}'(t)$ scales in the desired fashion (a similar argument applies

[1]The small discrepancies are due to the fact the fluid model assumes the physical buffer size to be infinite. If we increase the physical buffer size in the simulation, without altering the value of the max_{th}, the two lines coincide exactly.

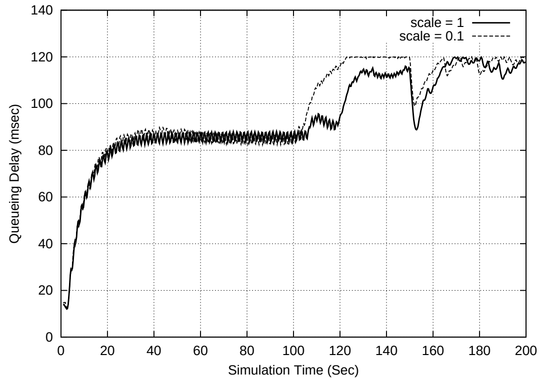

Figure 11: n = 3: Average Queuing Delay at Q1

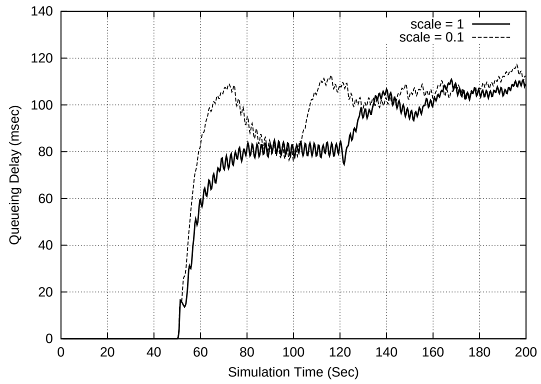

Figure 12: n = 3: Average Queuing Delay at Q2

to SlopeRED). Recall that $p'_{max} = p_{max}$, and that $min'_{th} = \alpha min_{th}$ and $max'_{th} = \alpha max_{th}$. It is then clear that

$$\bar{p}'(t) = p'_{max}\Big(\frac{\bar{q}'_a(t) - min'_{th}}{max'_{th} - min'_{th}}\Big)^n.$$

Since the other differential equations remain unchanged, all other quantities scale in exactly the same way as in [2], and thus the solutions to the original and scaled differential equations are identical for the modified AQM schemes we use.

4.2 Simulation results

From Figure 13 it is clear that the fluid model does a pretty good job of capturing the dynamics of the original network. But in Figures 14 and 15 we observe that the solution of the fluid model doesn't match the behavior of the network when the parameters for SlopeRED ($min_{th} = 2750$) and PowerRED ($n = 2$) are chosen so that drops occur in clumps. This demonstrates the fact that because of the clumpy nature of the packet-drop process, the differential equations no longer model the network dynamics faithfully. Also, since the solutions to the original and the scaled differential equations are identical, the inaccuracy of the fluid model will carry over to a scaled network. This in

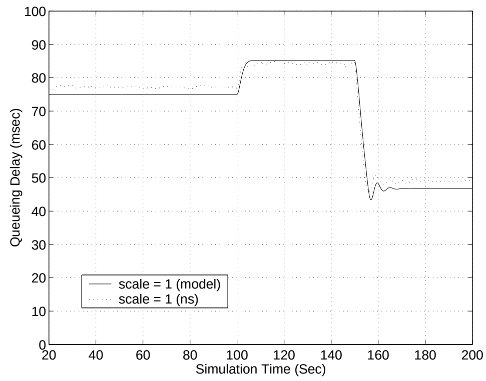

Figure 13: Basic Setup: Fluid Model vs. Ns Simulation

itself should give us reason to suspect that the scaling hypothesis will breakdown when the packet-drop process is clumpy; a fact already demonstrated in Section 3.

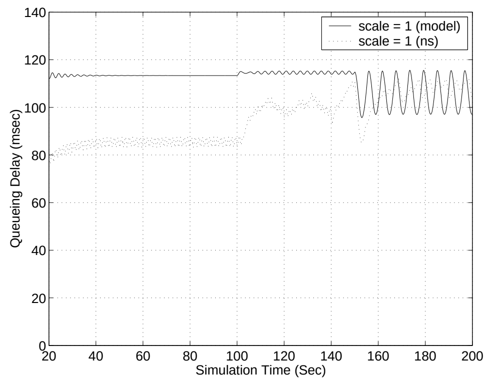

Figure 14: $min_{th} = 2750$: Fluid Model vs. Ns Simulation

5 Conclusion

In the regime where the differential equation model for an IP network with TCP flows is accurate, it has been argued in [2] that this model provides a possible explanation for why the scaling hypothesis holds. Specifically, it is shown there that for a large class of AQM schemes, the solutions of the differential equations that govern the dynamics of both the original and the scaled network are identical. But, we have observed in this paper that the ability of the fluid model and the scaled system to predict the behavior of the original network is seriously degraded when the packet drops occur in clumps. This is because an underlying assumption while applying these models is that the packet-drop process should be Poisson. If an AQM scheme drops consecutive packets then drops will occur in bursts, making the packet-drop process correlated and far from Poisson. Under such a scenario, these models will fail to accurately emulate the original network's dynamics. It is clear that the differential

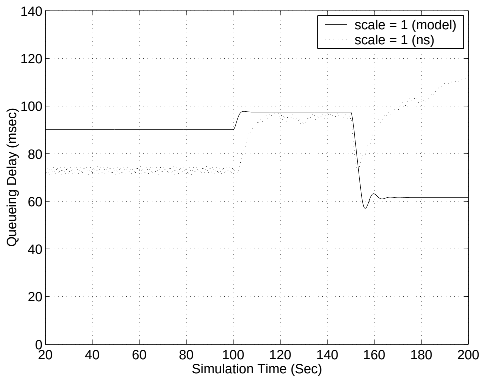

Figure 15: n = 2: Fluid Model vs. Ns Simulation

equation model and the scaling hypothesis have significant implications for network performance prediction and design. Hence, further work is required in identifying the regime of applicability of these models; and rigorously establishing their validity in that regime.

Acknowledgments

This work was supported in part by the National Science Foundation under Grant ANI-9985446; and by the Air Force Office of Scientific Research under Grant AF F49620-01-1-0365.

References

[1] V. Misra, W. Gong, D. Towsley. "Fluid-based Analysis of a Network of AQM Routers Supporting TCP Flows with an Application to RED", In *Proceedings of SIGCOMM*, 2000.

[2] R. Pan, K. Psounis, B.Prabhakar, D. Wischik. "The Scaling Hypothesis: Simplifying the Prediction and Measurement of Network Performance", Unpublished manuscript

[3] C.V. Hollot, V. Misra, D. Towsley and W. Gong, "A Control Theoretic Analysis of RED", In proceedings of *IEEE INFOCOM*, 2001.

[4] S. H. Low, F. Paganini, J. Wang, S. Adlakha and J. C. Doyle, "Dynamics of TCP/RED and a Scalable Control", To appear in proceedings of *IEEE INFOCOM*, 2002.

[5] F. Paganini, J. C. Doyle and S. H. Low, "Scalable Laws for Stable Network Congestion Control", In proceedings of *IEEE CDC*, Orlando, FL, December 2001.

[6] , S. Shakkottai and R. Srikant, "How Good are Deterministic Fluid Models of Internet Congestion Control", To appear in proceedings of *IEEE INFOCOM*, 2002.