A Study Group Presentation

Analysis and Design of Controllers for AQM Routers Supporting TCP Flows

C. V. Hollot, V. Misra, D. Towsley and W. Gong

Presented by: Tan Chee-Wei
Advisors: Professor Winston Chiu, Professor John Lui
The Chinese University of Hong Kong

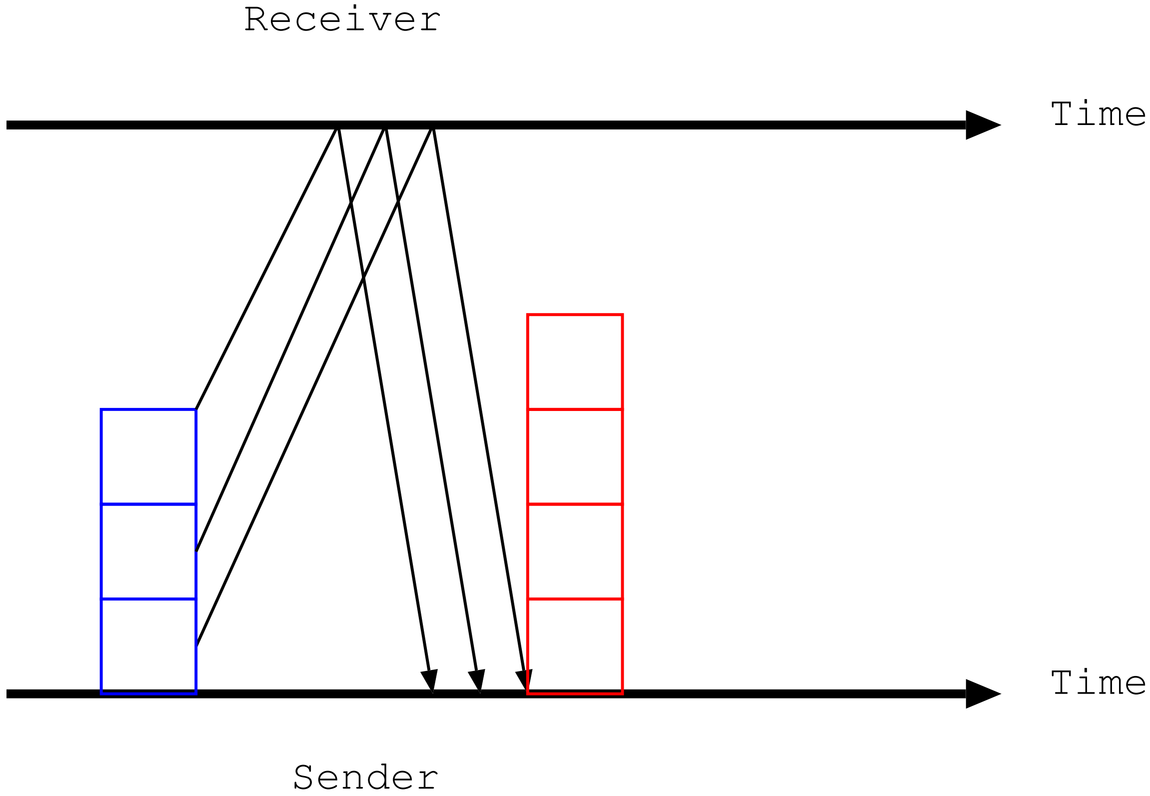
Contents

Introduction

- TCP is a protocol that sits between application and network
- The Internet is powered by TCP/IP
- TCP characteristics:
  - Acknowledgement
  - Retransmission
  - Flow control
  - Congestion avoidance

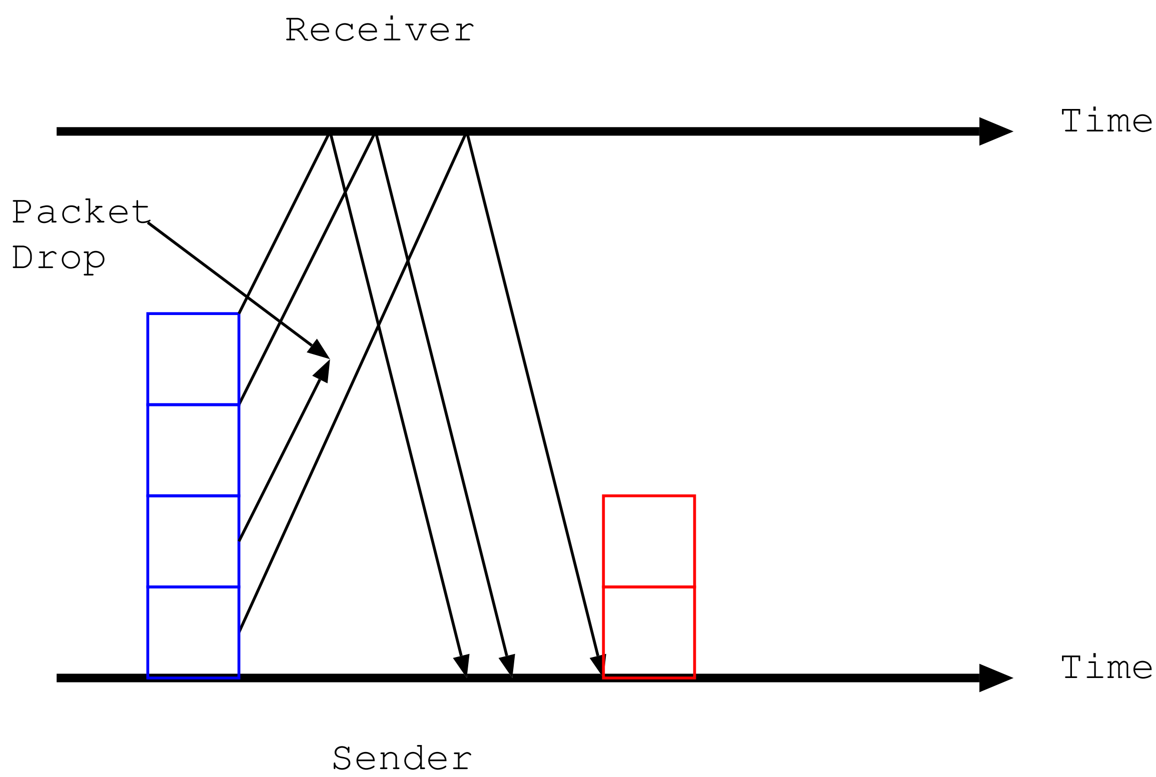
TCP Window Dynamics

- Additive Increase



TCP Window Dynamics

- Multiplicative Decrease



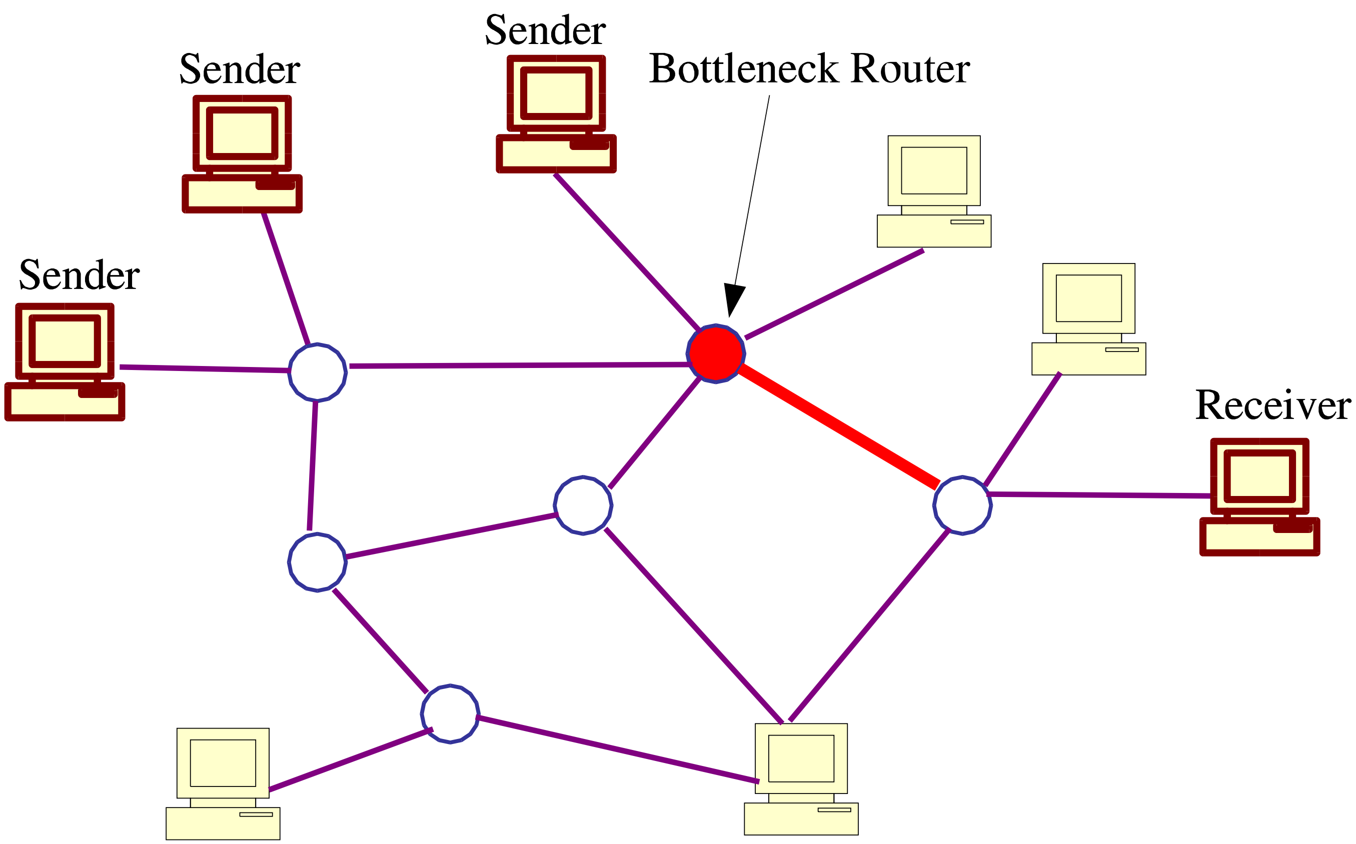
What Are The Factors Affecting TCP Throughput ?

- Number of TCP Sessions
- Round Trip Time
- Link Capacity of Network
- Buffer Capacity of Network

Why Active Queue Management ?

- Passive Queue Management has only two states
  - No packet drop $\Rightarrow$ No early congestion warning
  - 100% packet drop $\Rightarrow$ All senders back off
- Goal of AQM is to avoid the above problems
- Characteristics of AQM
  - Random packet drop is performed before buffer is full
  - The probability of packet drop increases with congestion level

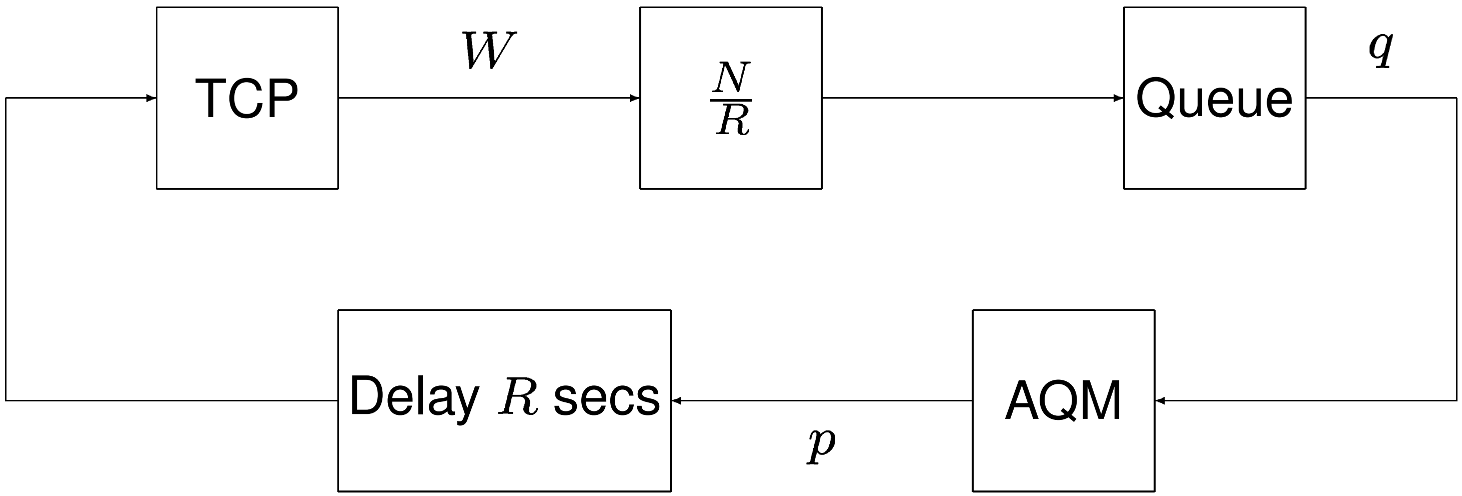
A Single Bottlenecked Queue



Contents

The Role Of Buffer-based AQM

- Introduction

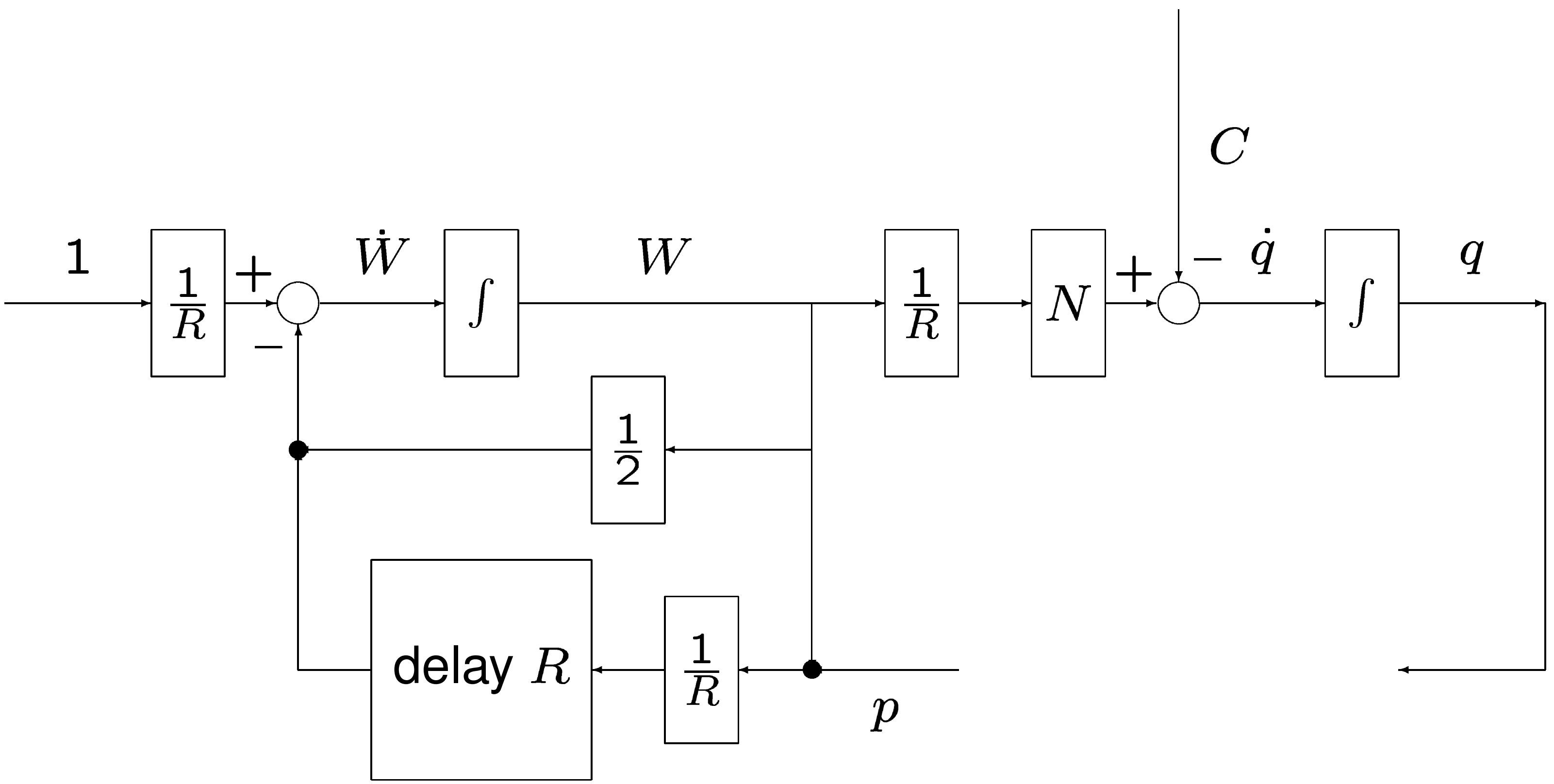


A Fluid-flow Model of TCP Behavior

- Windows dynamic
  - $\dot{W} = \frac{1}{R(t)} - \frac{W(t)}{2}\frac{W(t-R(t))}{R(t-R(t))}p(t - R(t))$
- Queue dynamic
  - $\dot{q} = -C + \frac{N(t)}{R(t)}W(t), \qquad q > 0$
  - $\dot{q} = max\{-C + \frac{N(t)}{R(t)}W(t), 0\}, \qquad q = 0$

$W \doteq$ average TCP time window size (packets);
$q \doteq$ average queue length (packets);
$R(t) \doteq$ round trip time $= \frac{q(t)}{C} + T_p$ (secs);
$C \doteq$ link capacity (packets/sec);
$T_p \doteq$ propagation delay (secs); $N \doteq$ number of TCP sessions;
$p \doteq$ probability of packet mark;

TCP Congestion Avoidance Mode

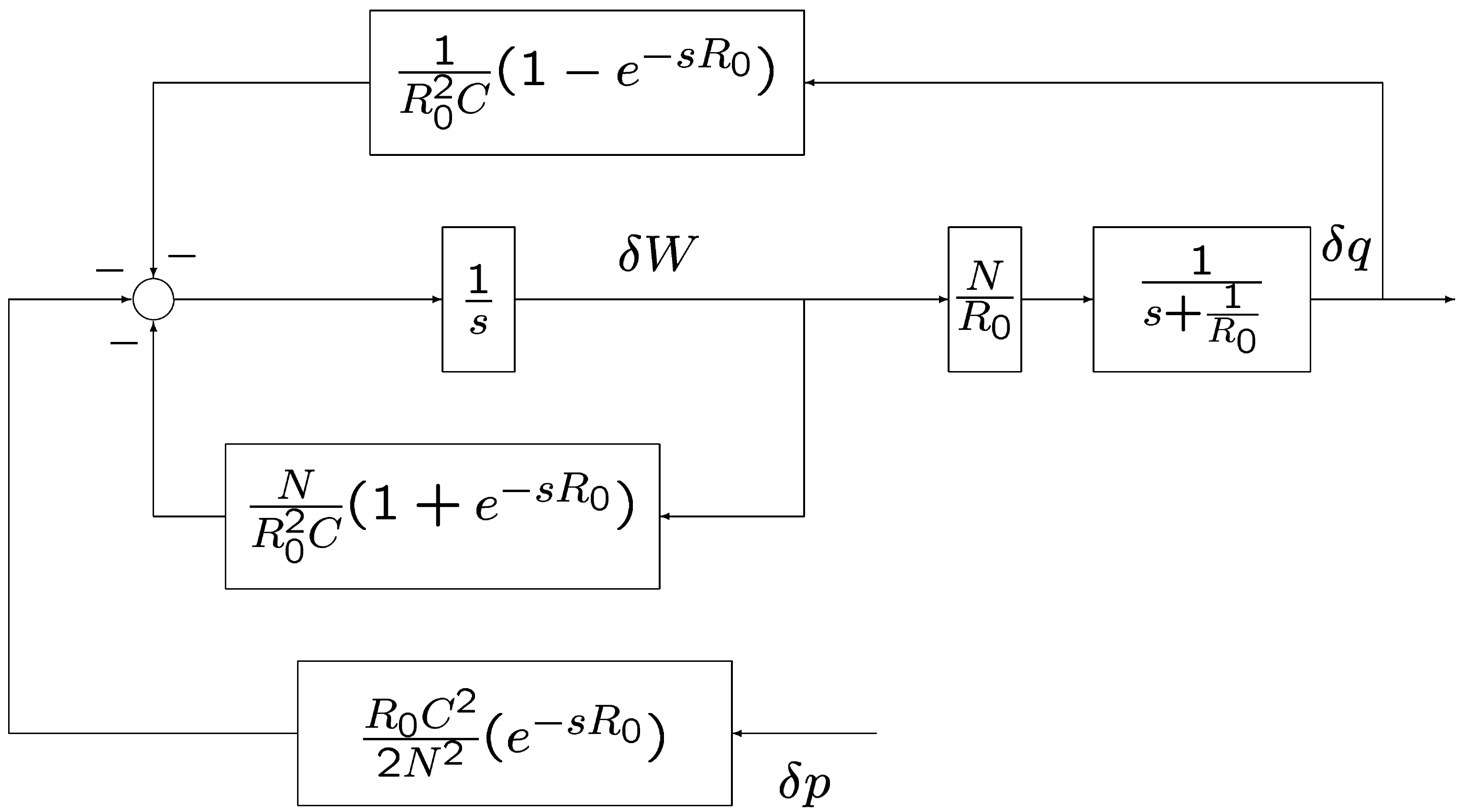


TCP Model Is Non Linear

- First step: Simplify TCP model
- Employ small signal linearization about steady state operating points (W_0, q_0, p_0)
- Perturbed variables about operating point
  - $\delta W \doteq W - W_0$
  - $\delta q \doteq q - q_0$
  - $\delta p \doteq p - p_0$
- Second step: Concentrate only on nominal dynamic behavior

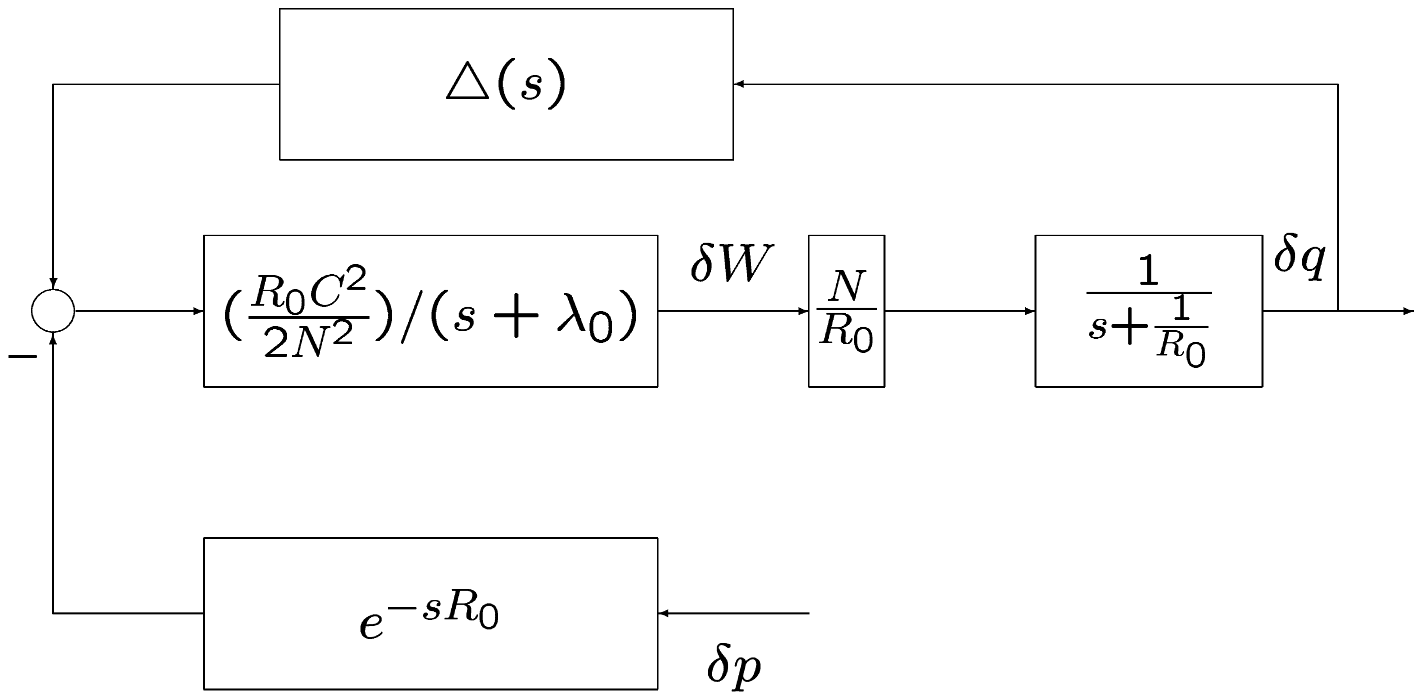
Linearalized TCP Connection

- First step: Simplify TCP model



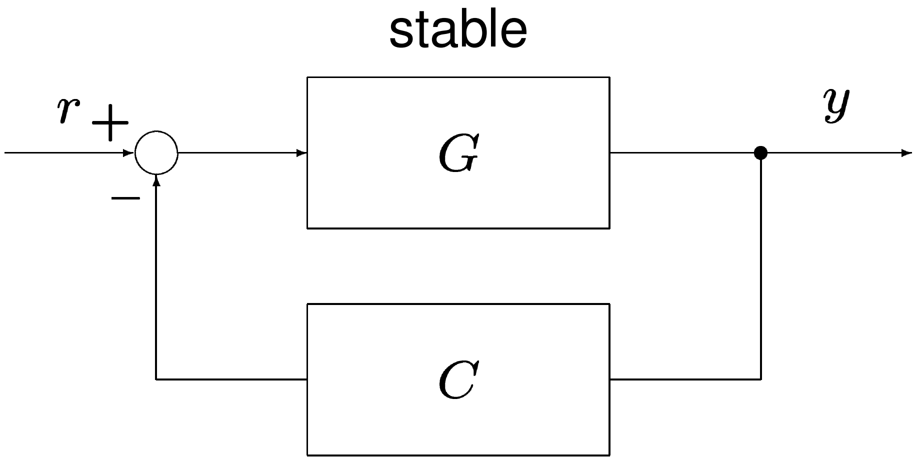
Linearized TCP Dynamic And High Frequency Parasitic

- Second step: Concentrate only on nominal dynamic behavior
- Isolate high frequency parasitic $\triangle(s)$



Negative Feedback Control Loop

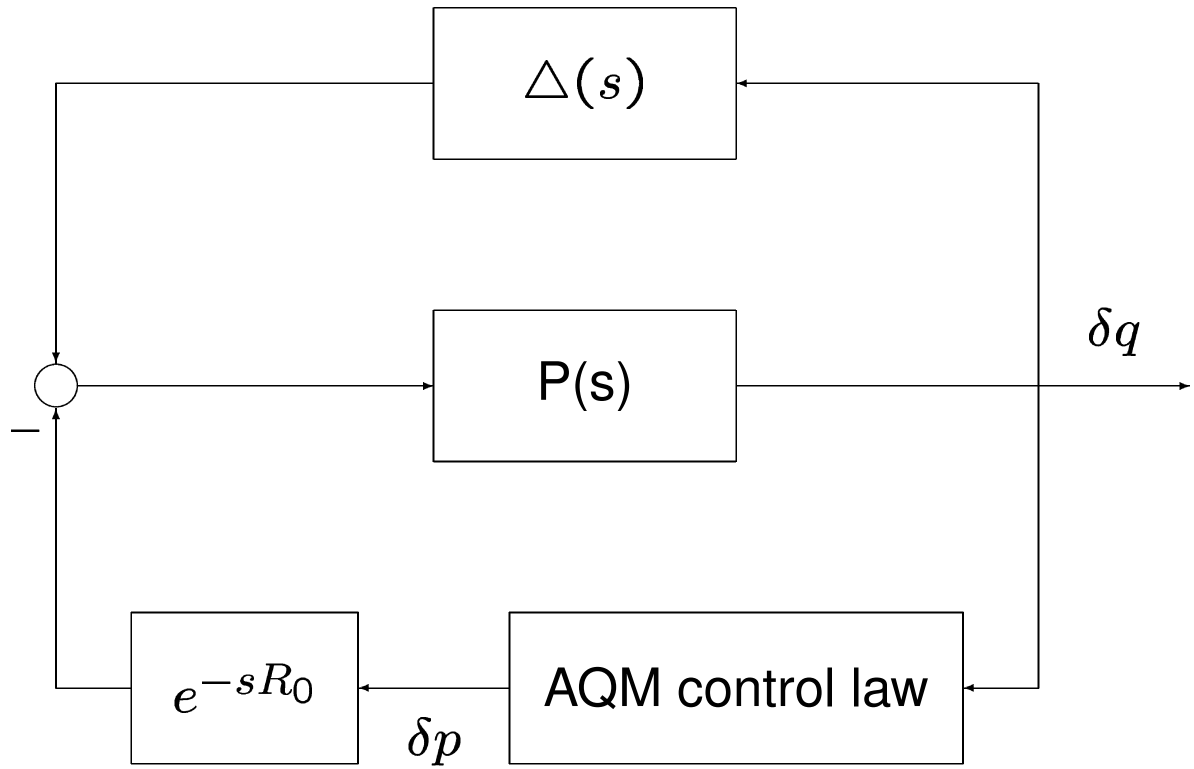
- Reference signal r
- Output signal y
- Stable plant G
- Controller C



- The case of a regulator with disturbance rejection, $r = 0$

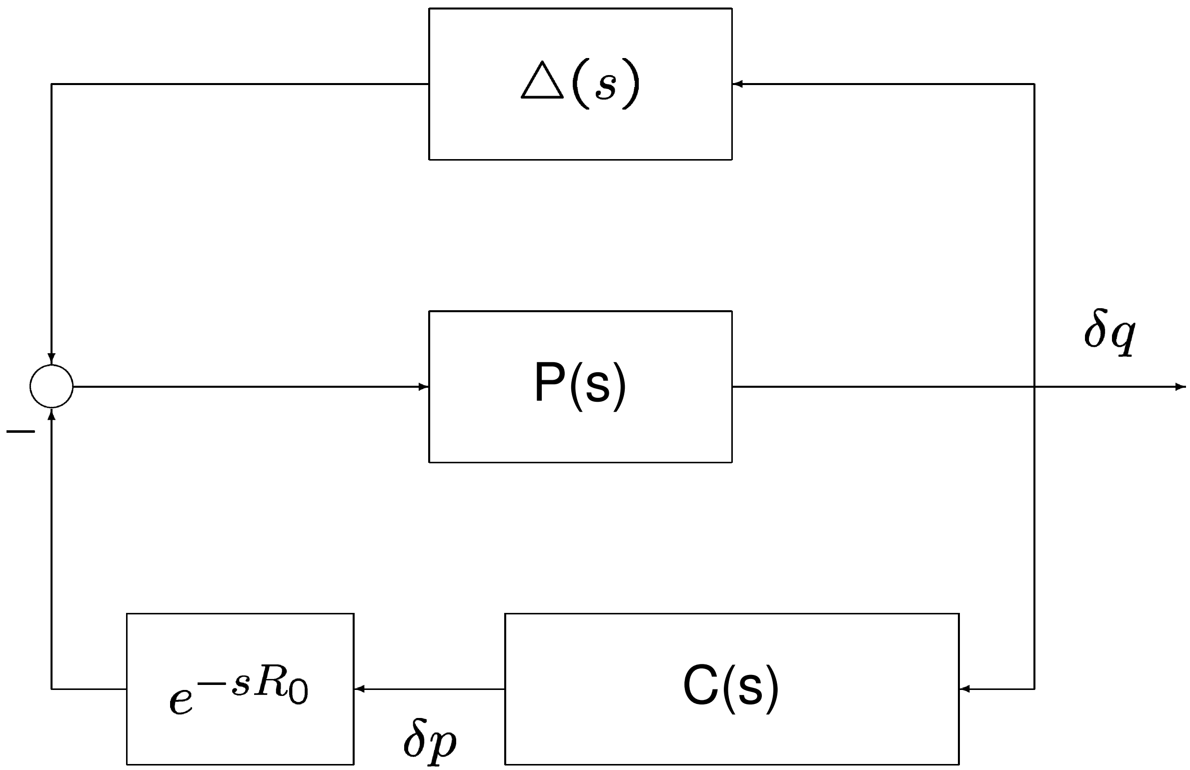
AQM As Feedback Control

- Plant model $P(s)$ consists of Window and Queue dynamics



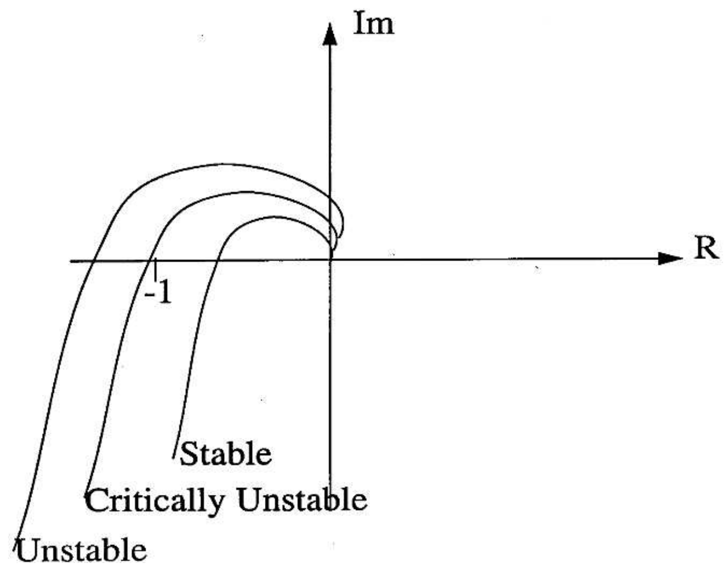
AQM As Feedback Control

- Design a $C(s)$ that stabilizes $P(s)e^{-sR_0}$ and gain-stabilizes $\triangle(s)$



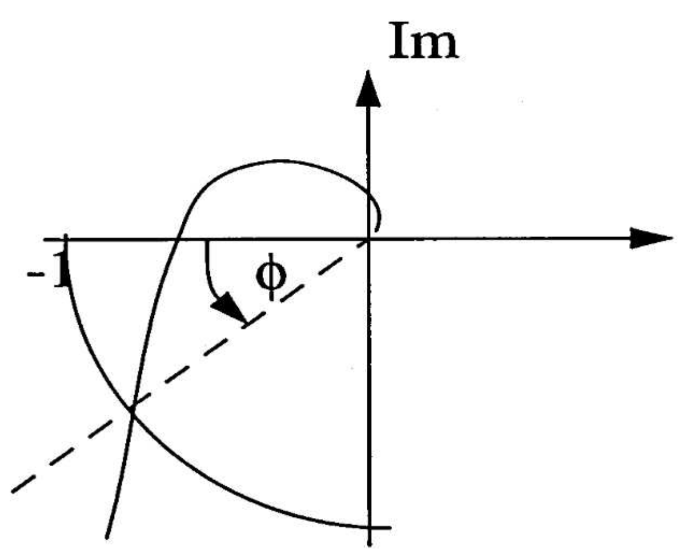
Nyquist Stability Criterion

- Assume $P(s)$ and $C(s)$ are stable. If the Nyquist plot of $L(s) = P(s)C(s)$ does not encircle the point $-1 + j0$ in the clockwise direction in the complex plane, then the system is asymptotically stable.



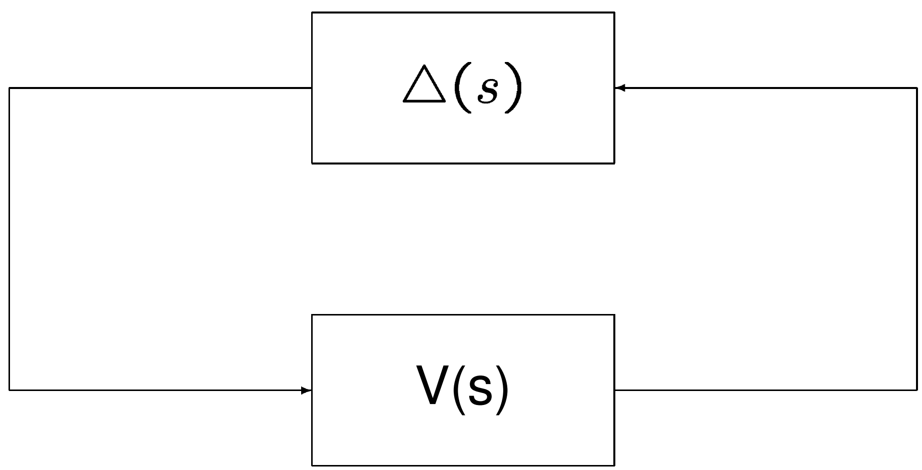
Relative Stability

- Frequency domain performance specifications
  - Phase Margin ϕ
  - Close Loop System Bandwidth w_b



Stabilizing AQM Control Laws

- $V(s) = \frac{P(s)}{1+P(s)C(s)e^{-sR_0}}$



- Using Small Gain Theorem, $| \triangle(jw)V(jw) | < 1 \Rightarrow \triangle(jw)V(jw)$ will not encircle -1 for all frequencies.

AQM Proposition 1

Given feasible network parameters $\eta = (N, C, T_p)$ and operating point $(W_0, q_0, p_0) \in \omega_\eta$, the linearized AQM control system is stable if

- $C(s)$ stabilizes the delayed nominal plant $P(s)e^{-sR_0}$;
- the high frequency parasitic $\triangle(s)$ is gain stabilized, i.e., $\mid \triangle(jw)V(jw) \mid < 1, \forall w > 0$.

AQM Proposition 2

Given feasible network parameters $\eta = (N, C, T_p)$ and operating point $(W_0, q_0, p_0) \in \omega_\eta$, assume $C(s)$ stabilizes the delayed nominal plant

$$P(s)e^{-sR_0} = \frac{\frac{C^2}{2N}e^{-sR_0}}{(s+\lambda_0)(s+\frac{1}{R_0})}.$$

Further, for feasible network parameters $\tilde{\eta} = (\tilde{N}, \tilde{C}, \tilde{T}_p)$ and operating point $(\tilde{W}_0, \tilde{q}_0, \tilde{p}_0) \in \omega_{\tilde{\eta}}$, suppose that

$$\tilde{N} \geq N,\ \tilde{C} \leq C,\ \frac{\tilde{q}_0}{\tilde{C}} + \tilde{T}_p \leq R_0$$

If $C(s)$ is stable, $\mid C(jw)P(jw) \mid$ is monotonically nonincreasing and $C(0) > 0$, then $C(s)$ stabilizes the perturbed plant

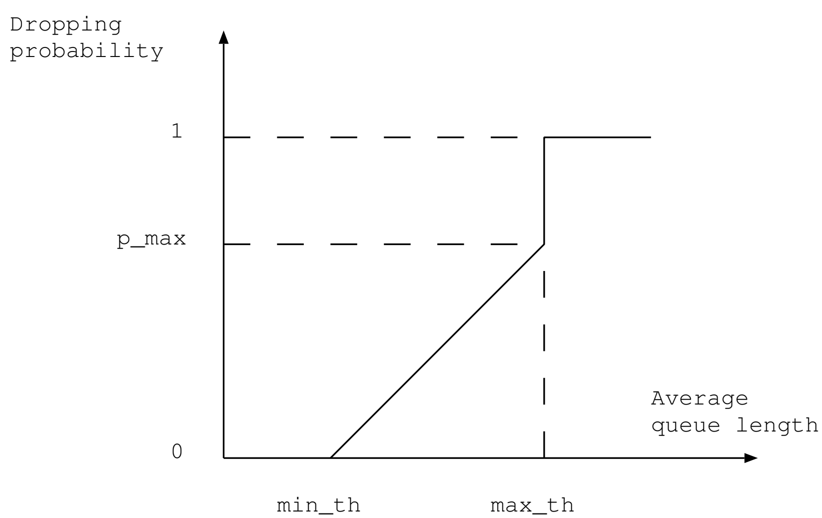
$$\tilde{P}(s)e^{-s\tilde{R}_0} = \frac{\frac{\tilde{C}^2}{2\tilde{N}}e^{-s\tilde{R}_0}}{(s+\tilde{\lambda}_0)(s+\frac{1}{\tilde{R}_0})}.$$

where $\tilde{R}_0 = (\tilde{q}_0/\tilde{C}) + \tilde{T}_p$.

Contents

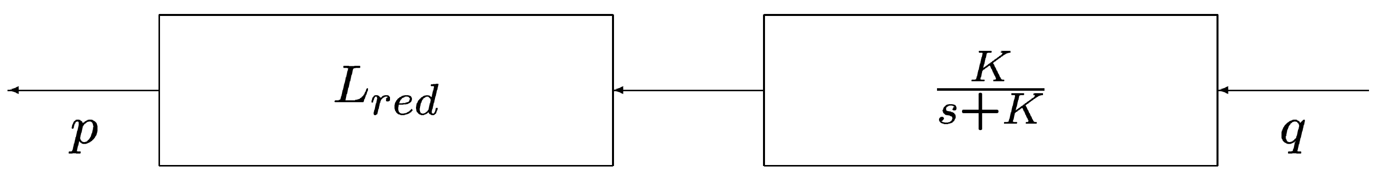
Tuning RED

- RED takes an average measure of the queue length and randomly drop packets that are within a threshold between $\min_{th}$ and $\max_{th}$
- Gradient $L_{red} = \frac{p_{max}}{max_{th} - min_{th}}$



AQM Using RED

- The low pass filter pole K is a function of the averaging weight and sampling frequency
- Transfer function of RED: $C(s) = \frac{K L_{red}}{s+K}$
- $L(s) = \frac{\frac{K L_{red} C^2}{2N} e^{-sR_0}}{(s+\lambda_0)(s+\frac{1}{R_0})(s+K)}$



- Select K such that it lies outside the loop's bandwidth or less than the corner frequencies of $P(s)$.
- Let pole K dominates the system's transient response.

AQM Using RED

- Evaluate system at unity gain cross-over frequency w_g
- Phase response of $L(jw)$ must satisfy $-w_g R_0 - \arctan\frac{w_g}{K} > -\pi$
- A tradeoff between w_g(Speed of response) and K(Queue averaging)
- Using the general relationship $w_g \leq w_b \leq 2w_g$
- For a desired phase margin, an increase in queue averaging leads to smaller system bandwidth w_b

An Example Of Using RED

- Consider the network parameters $C = 3750$pkts/s, $N = 60$ flows and $R_0 =$0.246s.
- Plant have poles at $\frac{2N}{R_0^2 C} = 0.53$ and $\frac{1}{R_0} = 4.1$
- Take $w_g = 0.1\min\{0.53, 4.1\} = 0.053\text{rads}^{-1}$
- To satisfy phase constraint, we get $K = 0.005$
- Dominant pole is too near to Imaginary Axis.

Contents

AQM Using Proportional Control

- RED results in small $w_b \Rightarrow$ Sluggish performance
- Remove the low pass filter in RED to improve transient response
  - Let $K \to \infty$
- $C(s) = K_P$
- $L(s) = \frac{\frac{K_P C^2}{2N} e^{-sR_0}}{(s+\lambda_0)(s+\frac{1}{R_0})}$
- Under the likely case $W_0 > 2$, phase response of $L(jw_g) > -147^o \Rightarrow$ Guarantees Close Loop Stability
- Using previous example, we have $w_g = 1.5\text{rads}^{-1}$ which is almost 30 times that of RED

Contents

AQM Using Proportional-Integral Control

- Both RED and Proportional Controller results in finite steady state error in queue length
- Use an Integrator with Proportional Controller to drive steady state error to zero
  - $\int_0^t u(\tau)d\tau$ =constant $\forall t > t_0$ if and only if $u(t) = 0 \forall t > t_0$
- $C(s) = K_{PI}\frac{s/z+1}{s}$
- $L(s) = \frac{\frac{K_{PI}C^2}{2N}(s/z+1)e^{-sR_0}}{s(s+\lambda_0)(s+\frac{1}{R_0})}$
- PI zero to coincide with TCP Window Pole $\Rightarrow$ First Order System
- Given a desired phase margin, select w_g that satisfies $-90^o - \frac{180}{\pi}w_g R_0 - \arctan(w_g R_0)$

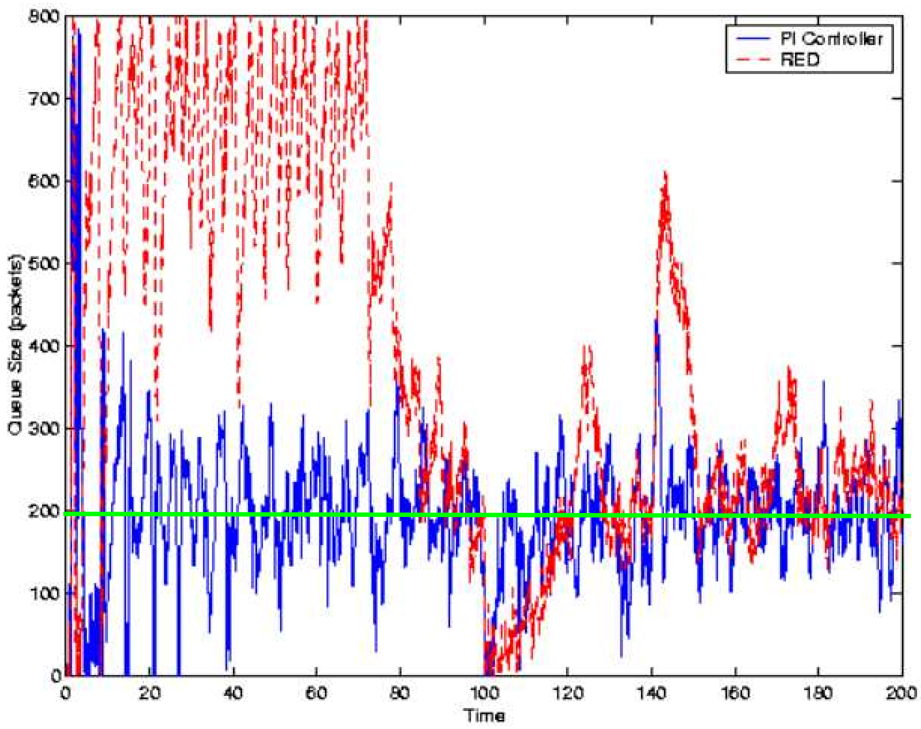
Designing PI Control

- Using the previous example, we select $w_g = 0.52\text{rads}^{-1}$, and we have a phase margin of 80^o and a system bandwidth of $0.52 \le w_b \le 1.04$
- Stability margin decreases with increased link capacity C, increased Round Trip Time or decreased number of TCP flows
- Need to avoid integrator windup due to control saturation since dropping probability p is $0 \le p \le 1$

Comparing PI Control With RED Using NS Simulations

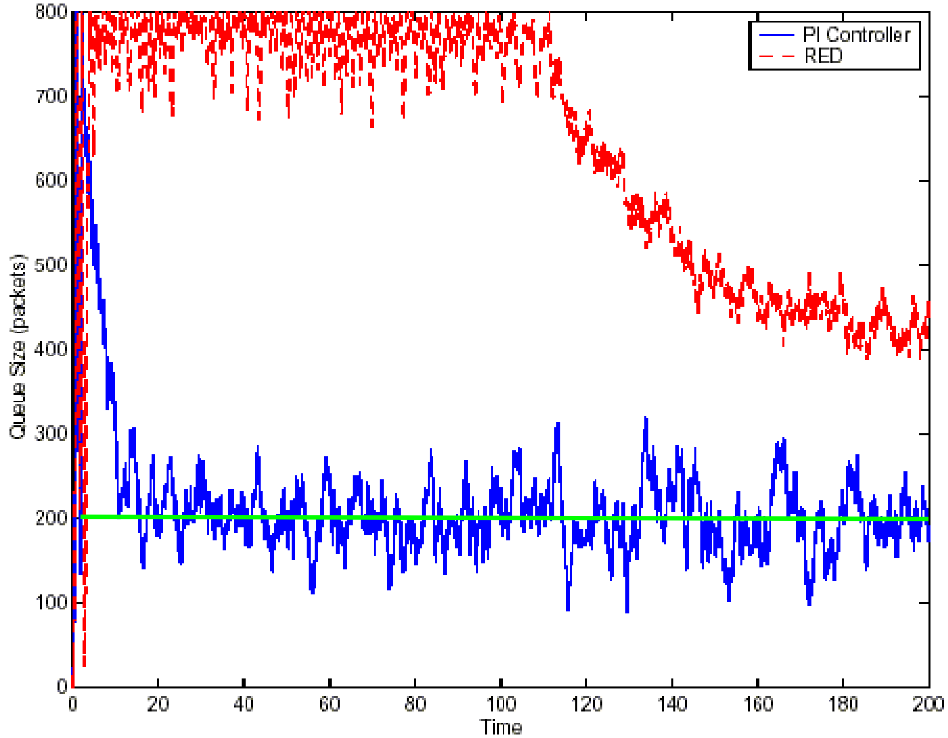
- Consider 60 TCP flows and 180 HTTP sessions
- A queue with buffer size 800 packets
- Desire a steady state queue length of 200 packets
- Load variation
  - At Time $t = 100$, 20 TCP flows drop out
  - At Time $t = 140$, 20 TCP flows return

PI Regulates Queue Length Independent Of TCP Flow Level



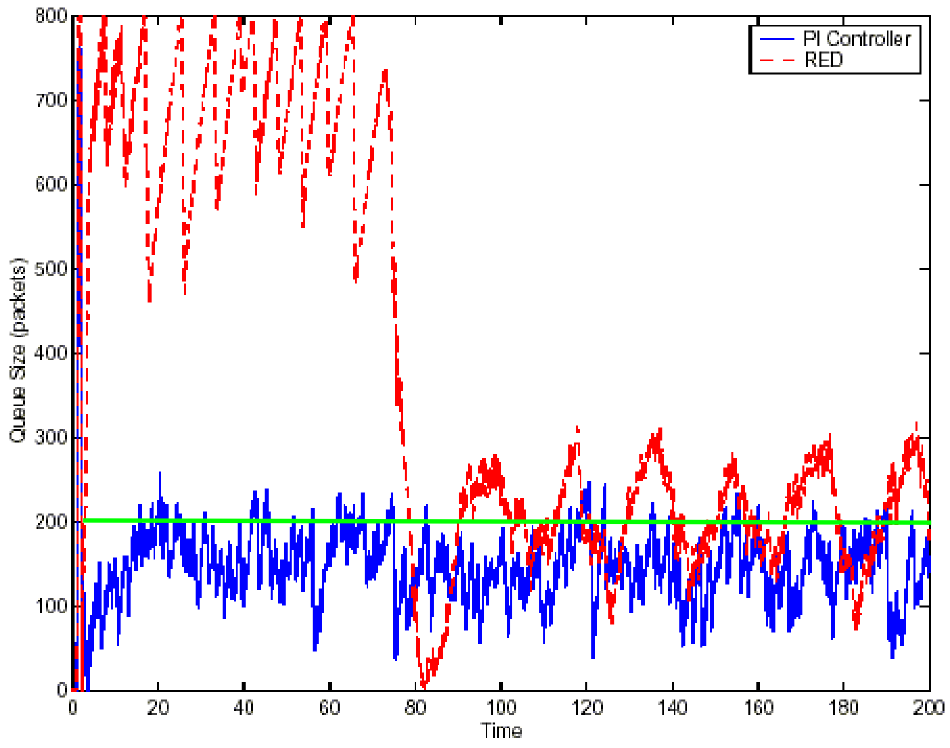
Increasing TCP Flow Decreases The System Bandwidth

- Smaller system bandwidth w_b dampens system transient response



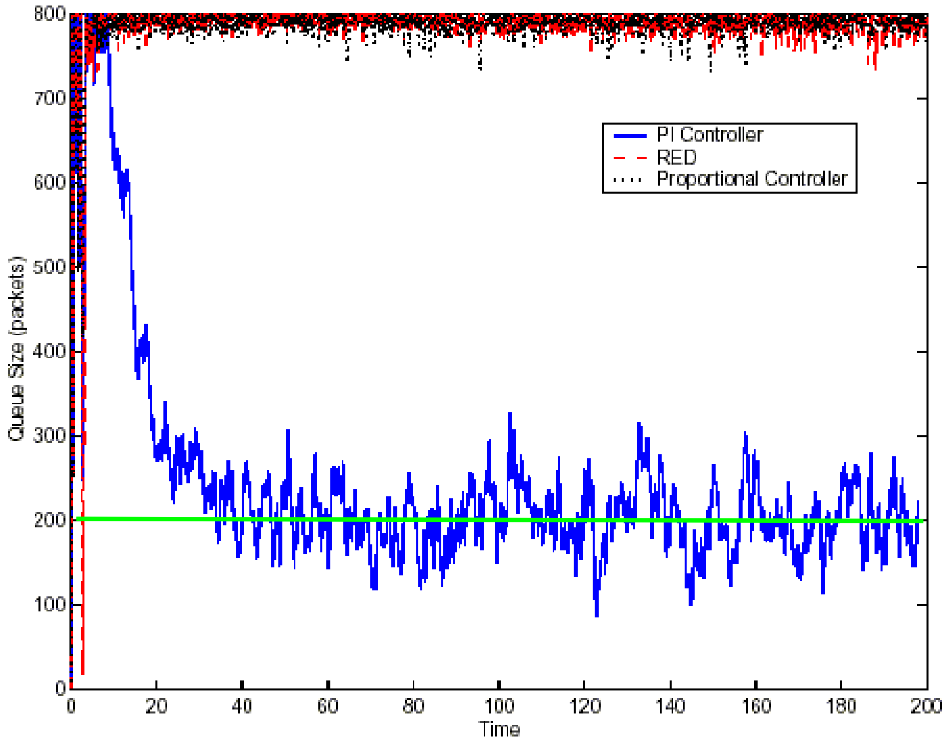
Phase Margin Decreases With Increasing TCP Flows

- Lower Phase Margin ⇒ More Oscillations



PI Continues To Regulate Queue Length At High TCP Load

- RED and Proportional controllers exhibit large steady state errors



Contents

Related Work

- L. Le, J. Aikat, K. Jeffay and F. D. Smith, "The Effects of Active Queue Management on Web Performance", ACM SIGCOMM 2003
- H. G. Zhang, C. V. Hollot, D. Towsley and V. Misra, "A Self-tuning Structure for Adaptation in TCP/AQM Networks", IEEE Globecom 2003
- Y. Gao and J. C. Hou, "A State Feedback Control Approach to Stabilizing Queues for ECN-Enabled TCP Connections", IEEE INFOCOM 2003
- P. F. Quet and H. $\ddot{O}$zbay, "On the Design of AQM Supporting TCP Flows Using Robust Control Theory", IEEE Transactions on Automatic Control, 2004
- C. G. Wang, B. Li, Y. T. Hou, K. Sohraby and Y. Lin, "LRED: A Robust Active Queue Management Scheme Based on Packet Loss Ratio", IEEE INFOCOM 2004

Contents

Conclusion

- A model for a single bottlenecked link
- Apply classical control theory to Internet congestion avoidance
- A set of design rules for tuning packet dropping probability
- Queue averaging in RED is not recommended
- Tradeoff between low queueing delay and high link utilization

Comments

- Other dynamic behavior
  - $\tilde{N} < N, \tilde{C} \leq C, \frac{\tilde{q}_0}{\tilde{C}} + \tilde{T}_p \leq R_0$
- Modeling issue
  - Congestion window W is not gradually decreased at a rate of $\frac{W_0^2 p}{2}$, but suddenly halved upon receipt of congestion
  - Model assumes only long-lived TCP connections (Elephants). Ignores short-lived HTTP connections (Mice) and UDP connections

What's Next ?

- More robust and adaptive AQM schemes for "Elephants" and "Mice"
- Multiple bottlenecks in network
- Exploit traffic characteristic, e.g., Long Range Dependence, for predictive control
- Feasibility of implementation of discretized PI controller in Today's high-speed routers