Modeling Congestion Control: Part 2

Lecture 3

Subir Varma

Congestion Control as a Solution to an Optimization Problem

Congestion Control and Optimization

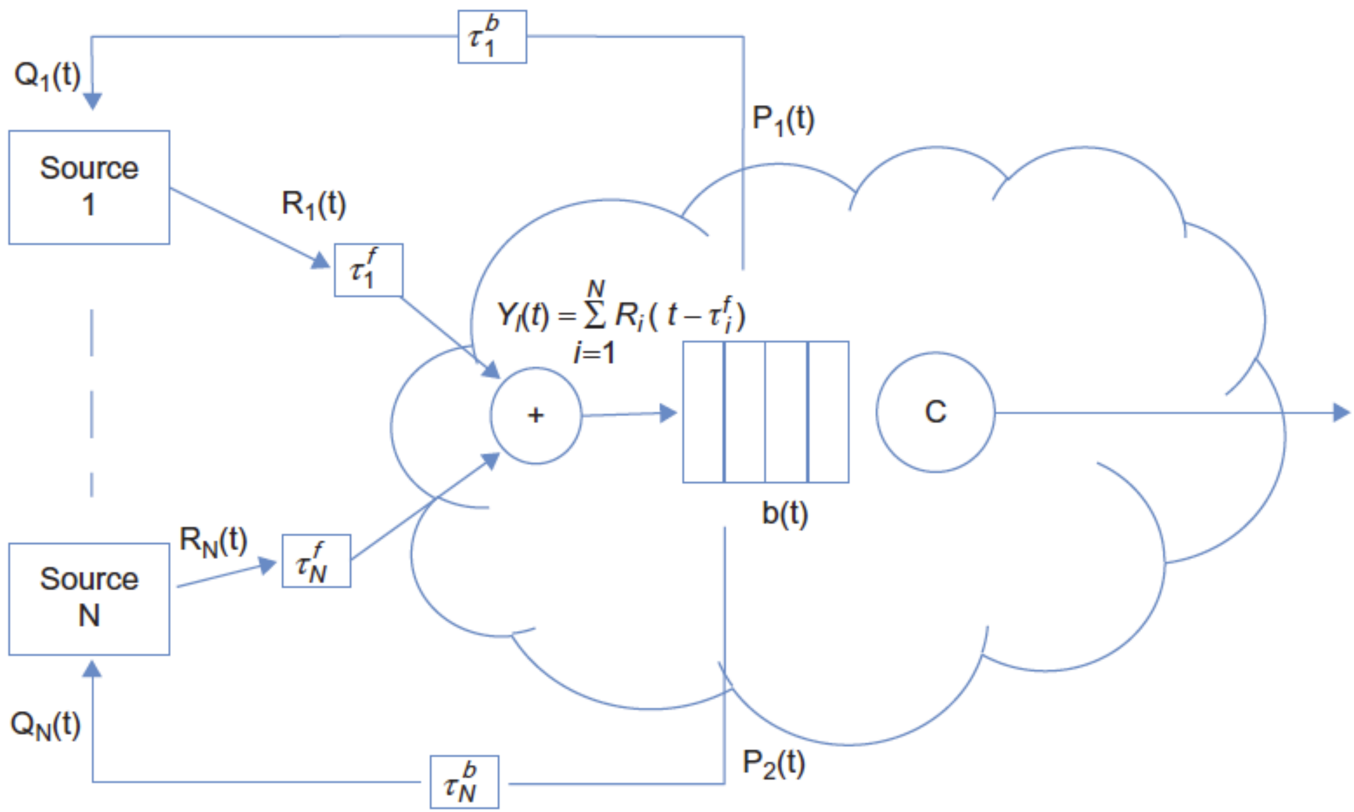
- So far we have introduced congestion control as a set of heuristic rules
- Can congestion control be looked upon as the solution to an optimization problem?
- If so, what is the Utility Function that algorithm is optimizing for?

Main Result

- We will use a network level model that represents entire networks with multiple connections
- By applying Lagrangian optimization theory to a fluid flow model of the network, it is possible to decompose the global optimization problem into independent local optimization problems at each source.
- The Lagrangian multiplier that appears in the solution can be interpreted as the congestion feedback coming from the network

- This provides an elegant theoretical result that provides a justification for the way congestion control protocols are designed.
- TCP can be put into this theoretical framework by modeling it in the fluid limit, and then the theory enables us to compute the global utility function that TCP optimizes.
- Alternately, we can derive new congestion control algorithms by starting from a utility function and then using the theory to compute the optimal rate function at the source nodes.

Network Level Model



Definitions

C_i: Capacity of the i[th] link, for $1 \le i \le L$, it is the i[th] element of the column vector C
L_i: Set of links that are used by source i
X_{li}: Element of a routing LXN matrix X, such that $X_{li} = 1$, if $l \in L_i$, and 0 otherwise
$R_i(t)$: Transmission rate of source i, for $1 \le i \le N$
r_i: Steady-state value of $R_i(t)$
$Y_l(t)$: Aggregate rate at link l from all the N sources, for $1 \le l \le L$
y_l: Steady-state value of $Y_i(t)$
$P_l(t)$: Congestion measure at link l, for $1 \le l \le L$. This is later identified as the buffer occupancy at the link.
p_l: Steady-state value of $P_l(t)$
$\tau^f_{li}(t)$: Propagation + Transmission + Queuing delay between the ith source and link l, in the forward direction
$\tau^b_{li}(t)$: Propagation + Transmission + Queuing delay between the ith source and link l, in the backward direction
$T_i(t) = \tau^f_{li}(t) + \tau^b_{li}(t)$ Total round trip delay
$Q_i(t)$: Aggregate of all congestion measures for source i, along its route for $1 \le i \le N$
q_i: Steady-state value of $Q_i(t)$
$b_l(t)$: Buffer occupancy at the link l

Model

Aggregate rate at link l

Note that

$$Y_l(t) = \sum_{i=1}^{N} X_{li} R_i(t - \tau_{li}^f(t)), \quad 1 \le l \le L, \quad \text{and}$$

Aggregate congestion feedback for source i

$$Q_i(t) = \sum_{l=1}^{L} X_{li} P_l(t - \tau_{li}^b(t)), \quad 1 \le i \le N$$

In steady state,

$$y = Xr \quad \text{and} \quad q = X^T p.$$

Assume that the equilibrium data rate is given by

$$r_i = f_i(q_i), \quad 1 \le i \le N$$

where f_i is a positive, strictly monotone decreasing function. This is a natural assumption to make because if the congestion along a source's path increases, then it should lead to a decrease in its data rate.

Local Optimization Problem for a Single Source

$$\max_{r_i}[U_i(r_i) - r_i q_i].$$

This equation has the following interpretation: If $U_i(r_i)$ is the utility that the source attains as a result of transmitting at rate r_i, and q_i is price per unit data that it is charged by the network, then the optimization leads to a maximization of a source's profit.

Optimum rate r_i^{max} occurs at

$$\frac{dU_i(r_i^{max})}{dr_i} = q_i = f^{-1}(r_i^{max})$$

Note that this equation is an optimization carried out by each source independently of the others, i.e., the solution r_i^{max} is individually optimal.

Global Optimization Problem

Find the rates r_i^{max} such that

$$\max_{r \geq 0} \sum_{i=1}^{N} U_i(r_i), \quad subject\ to$$

$$Xr \leq C.$$

Also called the Primal Problem

Are the solutions to the local and global optimization problems the same?

The Primal Problem

$$\max_{r \geq 0} \sum_{i=1}^{N} U_i(r_i), \quad \textit{subject to}$$

$$Xr \leq C.$$

A fully distributed implementation to solve the optimality problem is not possible because the sources are coupled to each other through the constraint equation.

There are two ways to approach this problem:

- By modifying the objective function for the primal problem by adding an extra term called the penalty or barrier function, or
- By solving the Dual Problem

Define the Lagrangian $L(r, \lambda)$

Lagrangian multipliers

$$L(r,\lambda) = \sum_{i=1}^{N} U_i(r_i) - \sum_{l=1}^{L} \lambda_l (y_l - C_l)$$

$$= \sum_{i=1}^{N} U_i(r_i) - \sum_{l=1}^{L} \lambda_l \sum_{i=1}^{N} X_{li} r_i + \sum_{l=1}^{L} \lambda_l C_l$$

$$= \sum_{i=1}^{N} U_i(r_i) - \sum_{i=1}^{N} r_i \sum_{l=1}^{L} X_{li} \lambda_l + \sum_{l=1}^{L} \lambda_l C_l$$

$$= \sum_{i=1}^{N} [U_i(r_i) - \overline{q}_i r_i] + \sum_{l=1}^{L} \lambda_l C_l \quad \text{where}$$

$$\overline{q}_i = \sum_{l=1}^{L} X_{li} \lambda_l \quad 1 \le i \le N$$

The Dual Problem

Define the Dual Function $D(\lambda)$

Convex Duality Theorem

$$D(\lambda) = \max_{r_i \geq 0} L(r, \lambda)$$

$$= \sum_{i=1}^{N} \max_{r_i \geq 0} [U_i(r_i) - \overline{q}_i r_i] + \sum_{l=1}^{L} \lambda_l C_l$$

This maximization can be done independently at each source
However since

$$r_i^{\max} = U_i'^{-1}(\overline{q}_i) = U_i'^{-1}\left(\sum_{l=1}^{L} X_{li}\lambda_l\right), \quad 1 \leq i \leq N$$

a source needs information from the network, in the form of $\overline{q}_i$, before it can compute its optimum rate.
This can be obtained by solving the Dual Problem

Find $\lambda_l, 1 \leq l \leq L,$ *such that* $\min_{\lambda \geq 0} D(\lambda)$

The Convex Duality Theorem

Find λ_l, $1 \leq l \leq L$, *such that* $\min_{\lambda \geq 0} D(\lambda)$

Substitute these back into to find the rates $r_i^{\max}$ at each source

$$r_i^{\max} = U_i'^{-1}(\overline{q}_i) = U_i'^{-1}\left(\sum_{l=1}^{L} X_{li}\lambda_l\right), \quad 1 \leq i \leq N$$

The Convex Duality Theorem states that these rates are also a solution to the original Primal Problem

Find the rates r_i^{max} such that

$$\max_{r \geq 0} \sum_{i=1}^{N} U_i(r_i), \quad \textit{subject to}$$

$$Xr \leq C.$$

Solving the Dual Problem

$$\min_{\lambda \geq 0} D(\lambda)$$

where

$$D(\lambda) = \sum_{i=1}^{N} \max_{r_i \geq 0} [U_i(r_i) - \overline{q}_i r_i] + \sum_{l=1}^{L} \lambda_l C_l$$

The Dual Problem can be solved by using the Gradient Projection Method

$$\lambda_l^{n+1} = \left[\lambda_l^n - \gamma \frac{\partial D(\lambda)}{\partial \lambda_l} \right]^+$$

$\gamma > 0$ is the step size and $[z]^+ = \max\{z, 0\}$.

Note that since $\overline{q}_i = \sum_{l=1}^{L} X_{li} \lambda_l$ it follows that

$$\frac{\partial D(\lambda)}{\partial \lambda_l} = C_l - \sum_{i=1}^{N} X_{li} r_i^{\max} = C_l - y_l(r^{\max}), \quad 1 \leq l \leq L$$

Solving the Dual Problem

So that

$$\lambda_l^{n+1} = [\lambda_l^n + \gamma(y_l(r_i^{\max}) - C_l)]^+$$

Note that the Lagrange multipliers λ_l behave as a congestion measure at the link because this quantity increases when the aggregate traffic rate at the link $y_l(r^{max})$ exceeds the capacity C_l of the link and conversely decreases when the aggregate traffic falls below the link capacity. Hence, it makes sense to identify the Lagrange multipliers λ_l with the link congestion measure p_l, so that $\lambda_l = p_l, 1 \leq l \leq L$, and

$$p_l^{n+1} = [p_l^n + \gamma(y_l(r_i^{\max}) - C_l)]^+$$

The Complete Solution

$$p_l^{n+1} = [p_l^n + \gamma(y_l(r_i^{\max}) - C_l)]^+$$

At link l:

1. Link l obtains an estimate of the total rate of the traffic from all sources that pass through it, y_l.
2. It periodically computes the congestion measure p_l using this equation, and this quantity is communicated to all the sources whose route passes through link l. This communication can either explicit as in ECN schemes or implicit as in random packet drops with RED.

At source i:

1. Source i periodically computes the aggregate congestion measure for all the links which lie along its route given by

$$q_i^n = \sum_{l=1}^{L} X_{li}\, p_l^n$$

2. Source i periodically chooses its new rate using the formula

$$r_i^n = U_i'^{-1}(q_i^n)$$

Applications

The solution can be used in two ways:

1. Given the Utility Function U, find the optimum Rate Control r that maximizes it, or
2. Given an exiting Rate Control r, find the Utility Function U that the Network is trying to maximize.

Utility Function for TCP Reno

Recall that $\frac{dU_i(r_i^{max})}{dr_i} = q_i = f^{-1}(r_i^{max})$, where $r = f(q)$

So that $U_i(r_i^{max}) = \int f^{-1}(r_i^{max})dr$

Hence we can find U_i by deriving the function f for TCP Reno

Utility Function for TCP Reno

Note that the total round trip delay is given by

$$T_i(t) = D_i + \sum_l X_{li} \frac{b_l(t)}{C_l}$$

where $D_i = T_{id} + T_{iu}$ is the total propagation delay and $b_l(t)$ is the i^{th} queue length at time t.

The source i rate $R_i(t)$ at time t is defined by $R_i(t) = \frac{W_i(t)}{T_i(t)}$

The aggregate congestion measure $Q_i(t)$ at a source can be written as

$$Q_i(t) = 1 - \prod_{l \in L_i} (1 - P_l(t - \tau_{li}^b(t))) \approx \sum_{l \in L_i} P_l(t - \tau_{li}^b(t))$$

Probability that none of the links are congested

Probability that at least one of the links are congested

Utility Function for TCP Reno

The rate of change in window size at source i for TCP Reno is given by the following equation in the fluid limit

Fraction of ACKs that are positive

$$\frac{dW_i(t)}{dt} = \frac{R_i(t - T_i(t))[1 - Q_i(t)]}{W_i(t)} - R_i(t - T_i(t))Q_i(t)\frac{1}{2}\frac{4W_i(t)}{3}$$

Window increase rate in the absence of congestion

Window decrease rate in the presence of congestion

We ignore the queuing delays so that the round trip delay is now fixed and approximated by

$$T_i = D_i + \sum_l \frac{X_{li}}{C_l} \quad \text{and} \quad R_i(t) \approx \frac{W_i(t)}{T_i}$$

It follows that

$$\frac{dR_i(t)}{dt} \approx \frac{R_i(t - T_i)[1 - Q_i(t)]}{R_i(t)T_i^2} - \frac{2R_i(t)R_i(t - T_i)Q_i(t)}{3}$$

$Q_i(t)$: Packet marking probability

Utility Function for TCP Reno

Assuming $R_i(t)=R_i(t-T_i)$

$$\frac{dR_i(t)}{dt} \approx \frac{[1-Q_i(t)]}{T_i^2} - \frac{2R_i^2(t)Q_i(t)}{3}$$

In Steady State
$$r_i = \frac{1}{T_i}\sqrt{\frac{3}{2}\frac{1-q_i}{q_i}} = f(q_i)$$

and
$$q_i = \frac{\frac{1}{T_i^2}}{\frac{1}{T_i^2}+\frac{2}{3}r_i^2} = \frac{1}{1+\frac{2}{3}r_i^2T_i^2} = f^{-1}(r_i)$$

So that
$$U_i(r_i) = \int \frac{\frac{1}{T_i^2}}{(\frac{1}{T_i^2}+\frac{2}{3}r_i^2)} dr_i = \frac{\sqrt{3/2}}{T_i}\tan^{-1}\left(\sqrt{\frac{2}{3}}r_iT_i\right)$$

Utility Function for TCP Reno

For small q_i $\quad r_i = \frac{1}{T_i}\sqrt{\frac{3}{2q_i}}$ and $q_i = \frac{3}{2r_i^2 T_i^2}$

So that $\quad U_i(r_i) = -\frac{1.5}{T_i^2 r_i}$

Utility functions of this form are known to lead to rates that minimize “potential delay fairness” in the network, that is, they minimize the overall potential delay of transfers in progress.

Other Utility Functions

Utility functions of the form

$$U_i(r_i) = w_i \log r_i$$

maximizes the "proportional fairness" in the network.
It can be shown that TCP Vegas's utility function is of this form, hence it achieves proportional fairness.

Utility functions of the form

$$U_i(r_i) = \lim_{\alpha \to \infty} \frac{r_i^{1-\alpha}}{1-\alpha}$$

lead to a max-min fair allocation of rates (cannot be achieved using AIMD type algorithms).

Generalized AIMD (GAIMD) Algorithms

$$W \leftarrow W + \frac{\alpha}{W^k} \quad \text{On positive ACK}$$

$$W \leftarrow W - \beta W^l \quad \text{On packet drop}$$

This results in the rate equation

$$\frac{dR_s(t)}{dt} = (1 - Q_s(t))R_s(t - T_s)\frac{\alpha}{R_s^{k+1}(t)T_s^{k+2}} - Q_s(t)R(t - T_s)\beta R_s^l(t)T_s^{l-1}$$

TCP Friendly GAIMD Algorithms

Define $\alpha_s(r_s) = \dfrac{\alpha}{r_s^{k+1} T_s^{k+2}}$ and $\beta_s(r_s) = \beta r_s^l T_s^{l-1}$

Which results in the rate equation

$$\frac{dR_s(t)}{dt} = (1 - Q_s(t))R_s(t - T_s)\alpha_s(R_s(t)) - Q_s(t)R_s(t - T_s)\beta_s(R_s(t))$$

In equilibrium $q_s = \dfrac{\alpha_s(r_s)}{\alpha_s(r_s) + \beta_s(r_s)} = f_s(r_s)$

For TCP Reno $q_s = \dfrac{1}{1 + r_s^2 T_s^2}$

GAIMD is TCP Friendly if the Utility Functions match, i.e., $\dfrac{\alpha_s(r_s)}{\beta_s(r_s)} = \dfrac{2}{r_s^2 T_s^2}$

TCP Friendly GAIMD Algorithms

Substituting $\alpha_s(r_s) = \dfrac{\alpha}{r_s^{k+1} T_s^{k+2}}$ and $\beta_s(r_s) = \beta r_s^l T_s^{l-1}$

We get $\dfrac{\alpha}{\beta} \dfrac{1}{(r_s T_s)^{k+l+1}} = \dfrac{2}{r_s^2 T_s^2}$

Hence the GAIMD is TCP Friendly if and only if

$$k + l = 1 \quad \text{and} \quad \frac{\alpha}{\beta} = 2$$

If equate $dR_s/dt = 0$ in equilibrium, then we obtain the following expression, which is the analog of the square-root formula for GAIMD algorithms:

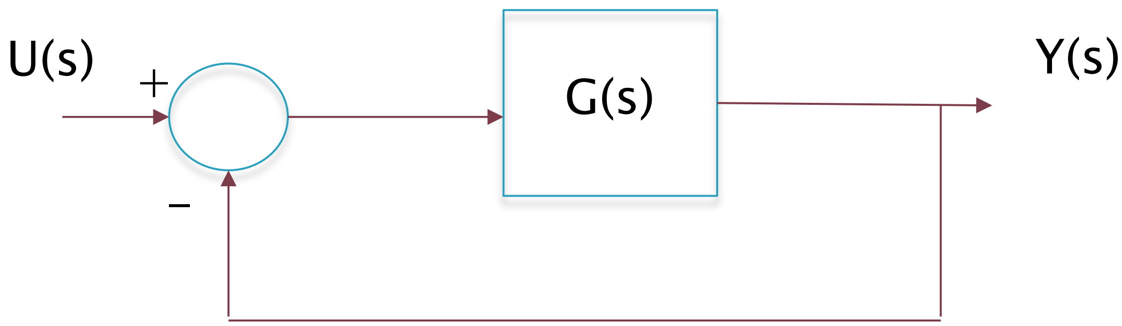
$$r_s = \left(\frac{\alpha}{\beta}\right)^{1/k+l+1} \frac{1}{T_s} \left(\frac{1}{q_s^{1/k+l+1}} - 1\right) \approx \left(\frac{\alpha}{\beta}\right)^{1/k+l+1} \frac{1}{T_s} \frac{1}{q_s^{1/k+l+1}} \tag{52}$$

Stability Analysis

What does Stability (or Instability) Look Like?

- The stability of a congestion control system is defined in terms of the behavior of the bottleneck queue size b(t).
- If the bottleneck queue size fluctuates excessively and very frequently touches zero, thus leading to link under utilization, then the system is considered to be unstable.
- Also, if the bottleneck queue size grows and spends all its time completely full, which leads to excessive packet drops, then again the system is unstable.
- Hence, ideally, we would like to control the system so that the bottleneck queue size stays in the neighborhood of a target length, showing only small fluctuations.

The Nyquist Stability Criterion



$$\frac{Y(s)}{U(s)} = \frac{G(s)}{1 + G(s)}$$

- To ensure stability, all the roots of the equation 1+G(s) = 0, which in general are complex numbers, have to lie in the left half of the complex plane.
- The Nyquist criterion is a technique for verifying this condition that is often easier to apply than finding the roots

Set $s = j\omega$ so that G(s) can be written as

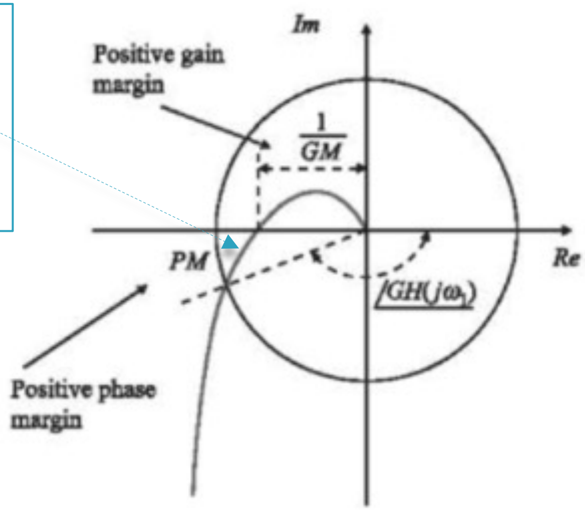
$$G(j\omega) = |G(j\omega)|e^{-j\, arg(G(j\omega))}$$

Vary ω from 0 to ∞ and plot the corresponding values of $G(j\omega)$ on the complex plane.

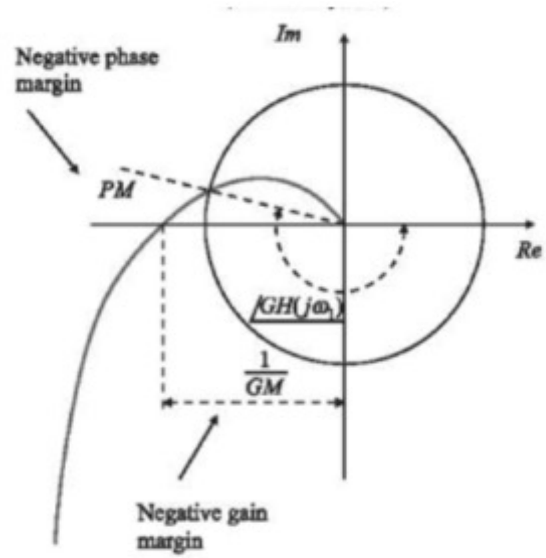
Nyquist Criterion for System Stability (with feedback): There should be no clockwise encirclements of the point (−1, 0) by the locus of $G(j\omega)$

The Nyquist Stability Criterion

Find a point in this part of the curve that satisfies $|G(j\omega_c)| \leq 1$ and $\arg\big(G(j\omega_c)\big) < 180^0$



Stable System

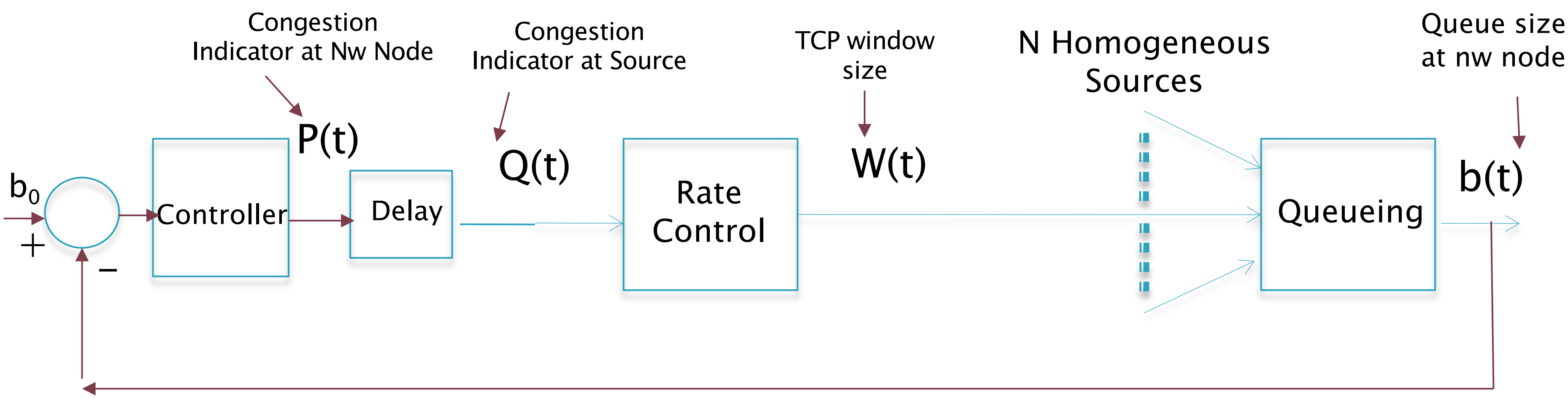


Un-Stable System

Alternative statement of the Nyquist Stability Criterion: Consider the point where the $G(j\omega)$ curve intersects the unit circle, this is called the gain crossover point. For a feedback system to be stable, the angle $\arg(G(j\omega))$ should be less that 180 degrees.

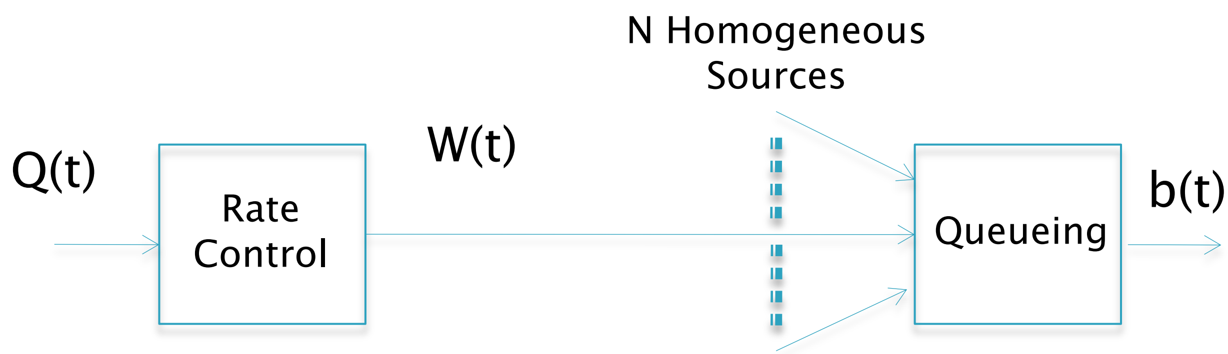
To show stability: It is sufficient to find a point ω_c such that $|G(j\omega_c)| \leq 1$ and $\arg\big(G(j\omega_c)\big) < 180^0$

Congestion Control System



Find expressions for $\frac{W(s)}{Q(s)}$ and $\frac{b(s)}{W(s)}$

TCP System Dynamics



Window control dynamics are given by

$$\frac{dW(t)}{dt} = \frac{R(t-T(t))[1-Q(t)]}{W(t)} - R(t-T(t))W(t)\frac{Q(t)}{2}$$

Substituting W(t) = R(t)T(t) in the first term on the RHS and making the approximations $R(t) \approx R(t-T(t)), 1-Q(t) \approx 1$ leading to the equation

$$\frac{dW(t)}{dt} = \frac{1}{T(t)} - \frac{W(t)W(t-T(t))}{2T(t-T(t))}Q(t) \quad \text{where} \quad T(t) = \frac{b(t)}{C} + D.$$

The fluid approximation for the queue length process at the bottleneck can be written as

$$\frac{db(t)}{dt} = \frac{W(t)}{T(t)}N(t) - C$$

These are Non-Linear Equations!

Linearization

Linearize around the operating point (W_0, b_0, Q_0) which is defined by dW/dt = 0 and db/dt = 0.

$$\frac{dW}{dt} = 0 \Rightarrow W_0^2 Q_0 = 2$$

$$\frac{db}{dt} = 0 \Rightarrow W_0 = \frac{CT_0}{N}$$

where $T_0 = \frac{b_0}{C} + D$

Define

$$W_\delta = W - W_0$$

$$b_\delta = b - b_0$$

$$Q_\delta = Q - Q_0$$

After Linearization

$$\frac{dW_\delta(t)}{dt} = -\frac{2N}{CT_0^2} W_\delta(t) - \frac{C^2 T_0}{2N^2} Q_\delta(t)$$

$$\frac{db_\delta(t)}{dt} = \frac{N}{T_0} W_\delta(t) - \frac{1}{T_0} b_\delta(t)$$

With Laplace Transforms

$$U_{tcp}(s) = \frac{W_\delta(s)}{Q_\delta(s)} = \frac{\frac{T_0 C^2}{2N^2}}{s + \frac{2N}{T_0^2 C}}$$

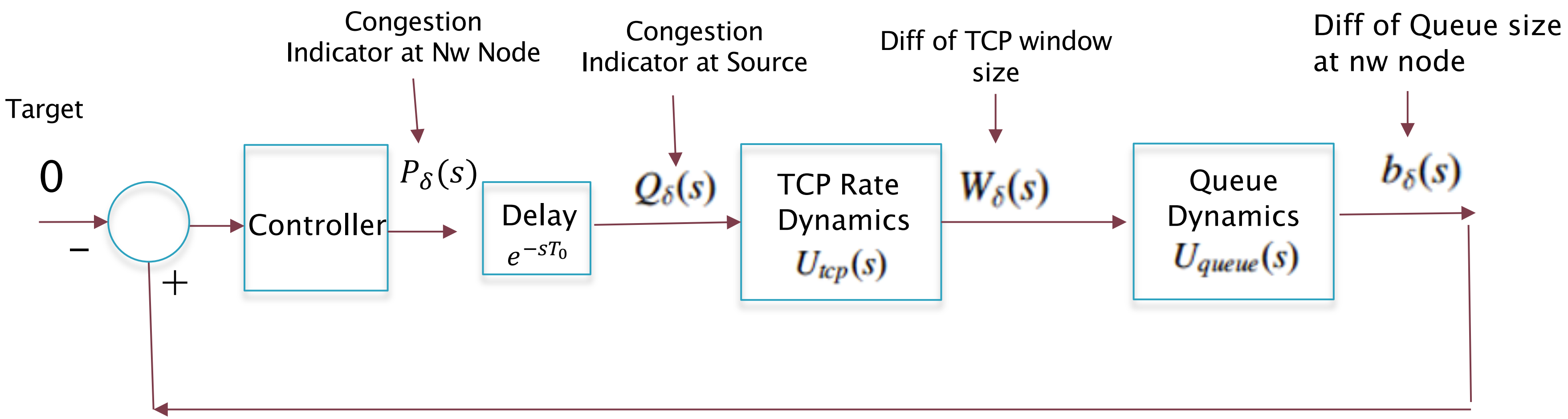
$$U_{queue}(s) = \frac{b_\delta(s)}{W_\delta(s)} = \frac{\frac{N}{T_0}}{s + \frac{1}{T_0}}$$

Linearized System

$W_\delta = W - W_0$

$b_\delta = b - b_0$

$Q_\delta = Q - Q_0$

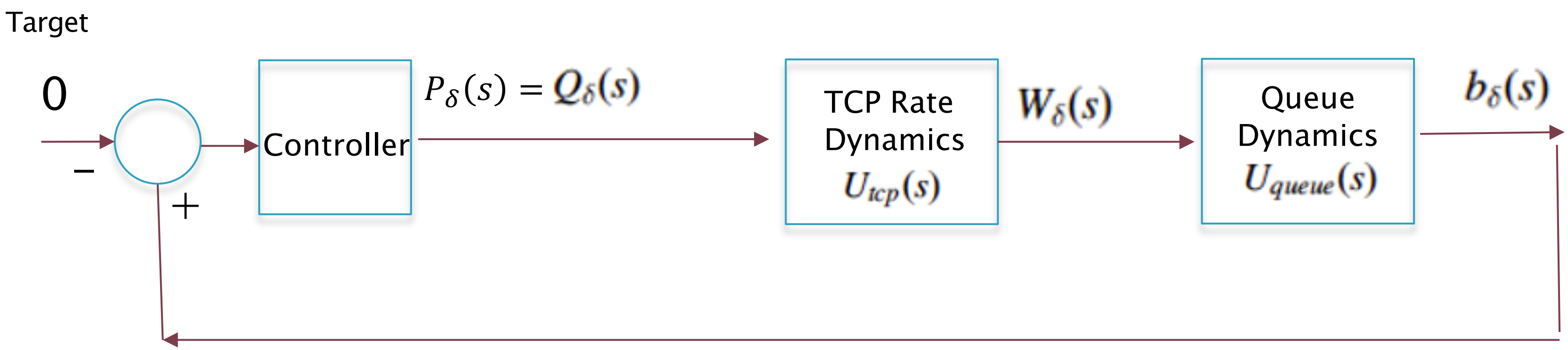


The objective of the control loop is to keep the target queue size at the network node close to b_0, i.e., drive the difference b_δ to zero.

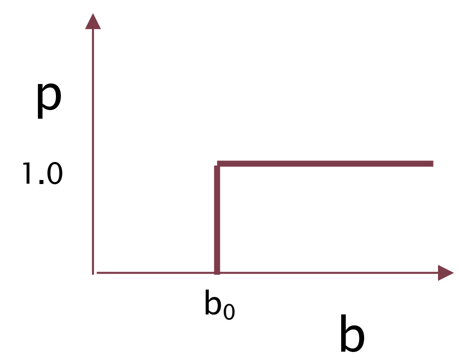
$$U_{tcp}(s) = \frac{W_\delta(s)}{Q_\delta(s)} = \frac{\frac{T_0 C^2}{2N^2}}{s + \frac{2N}{T_0^2 C}} \qquad U_{queue}(s) = \frac{b_\delta(s)}{W_\delta(s)} = \frac{\frac{N}{T_0}}{s + \frac{1}{T_0}}$$

On-Off Controllers

A Closed Loop System with Instant Feedback and On-Off Controller



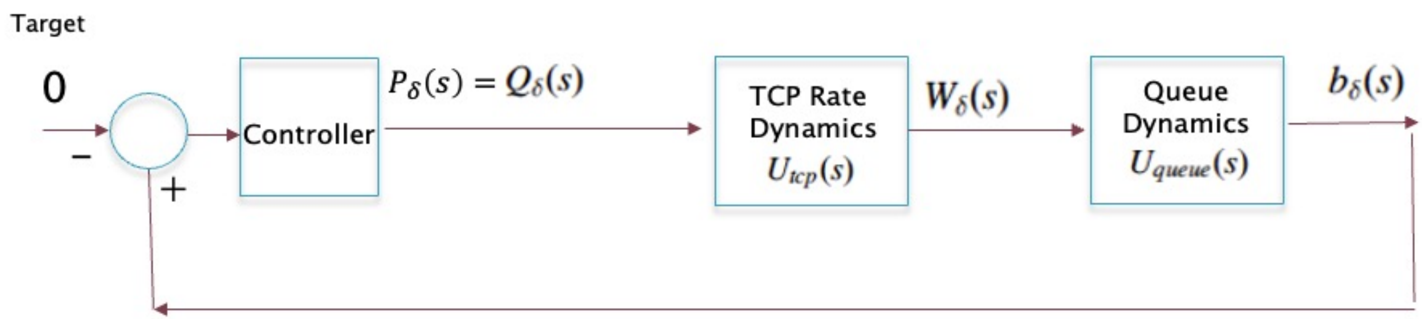
Simple On-Off Controller



The plant dynamics for this system is

$$U(s) = U_{tcp}(s)U_{queue}(s) = \frac{\frac{C^2}{2N}}{\left(s + \frac{2N}{T_0^2 C}\right)\left(s + \frac{1}{T_0}\right)} = \frac{\frac{(CT_0)^3}{(2N)^2}}{\left(\frac{s}{2N/CT_0^2} + 1\right)(sT_0 + 1)}$$

Instant Feedback: Transfer Function



The open loop transfer function of this system is of the form

$$U(s) = \frac{K}{(as+1)(bs+1)}$$

The corresponding close loop transfer function W(s) is given by

$$W(s) = \frac{U(s)}{1+U(s)} = \frac{K}{(as+1)(bs+1)+K}.$$

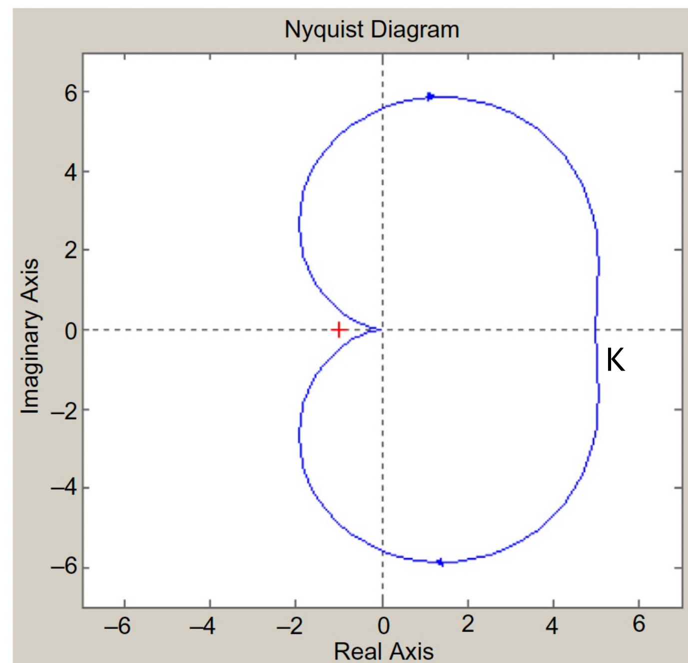
Instant Feedback: Stability Analysis

In order to apply Nyquist Critereon set $s = j\omega$ so that

$$U(j\omega) = \frac{K \exp(-j(\theta_1+\theta_2))}{(1+a^2\omega^2)(1+b^2\omega^2)} \text{ where } \theta_1 = tan^{-1}(\mathrm{a}\omega) \; and \; \theta_2 = tan^{-1}(\mathrm{b}\omega)$$

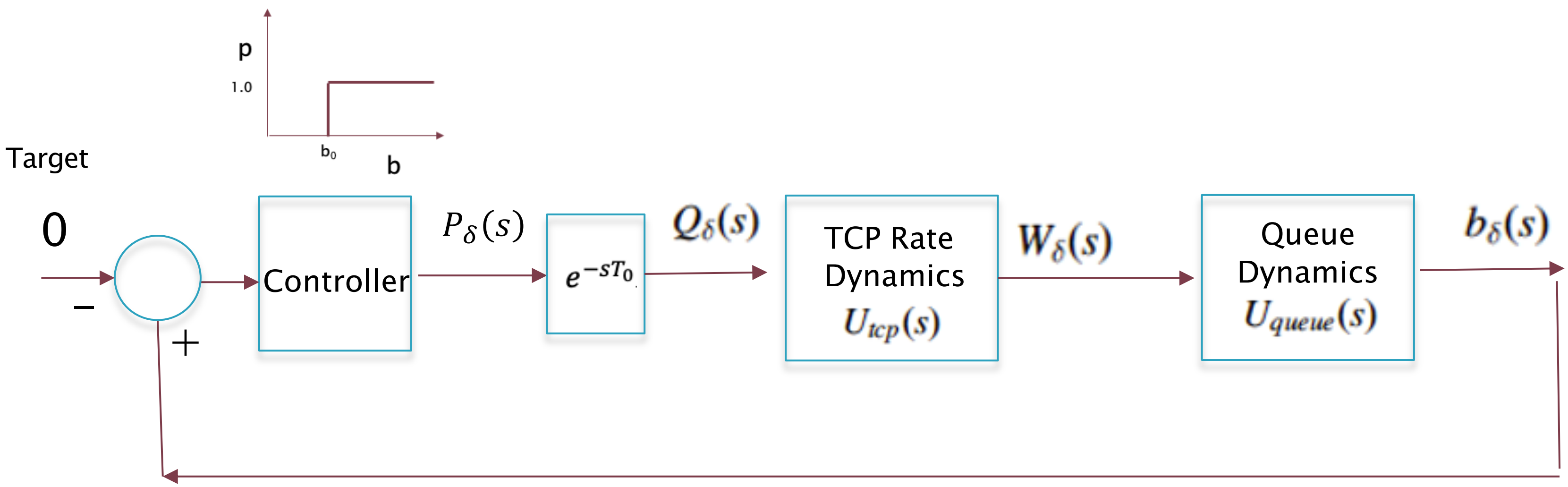
Thus U(0) = K and as ω increases $\theta_1 \to \frac{\pi}{2}$ and as does θ_2

Also as ω increases $|U(j\omega)| \to 0$



Using the Nyquist criterion, we can see that the locus of $U(j\omega)$ as ω varies from 0 to ∞, does not encircle the point (-1, 0), even for large values of K; hence, the system is unconditionally stable

A Closed Loop System with Delayed Feedback and On-Off Controller (TCP)



Next let's introduce the propagation delay into the system, so that

$Q_\delta(t) = P_\delta(t - T)$, so that $Q_\delta(s) = e^{-sT_0} P_\delta(s)$

The open loop and closed loop transfer functions for this system are given by

$$U'(s) = U(s)e^{-sT_0} = \frac{Ke^{-sT_0}}{(as+1)(bs+1)}$$

$$W'(s) = \frac{Ke^{-sT_0}}{(as+1)(bs+1) + Ke^{-sT_0}}$$

Given x = a +jb $= |x| e^{j\,\mathrm{Arg}(x)}$

$|x| = \sqrt{a^2 + b^2}$

Arg(x) $= tan^{-1} \frac{b}{a}$ radians

Delayed Feedback: Stability Analysis

In order to apply Nyquist Critereon set $s = j\omega$ so that

$$U(j\omega) = \frac{K\exp(-j(\omega T_0+\theta_1+\theta_2))}{(1+a^2\omega^2)(1+b^2\omega^2)}$$ where $\theta_1 = tan^{-1}(\mathrm{a}\omega)$ *and* $\theta_2 = tan^{-1}(\mathrm{b}\omega)$

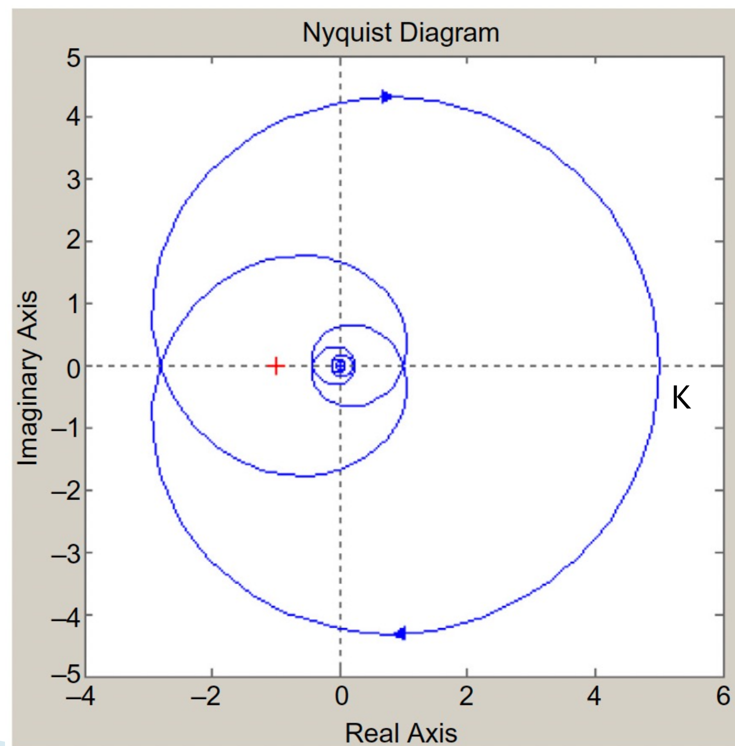
Thus U(0) = K and as ω increases ωT_o increase and $\theta_1 \rightarrow \frac{\pi}{2}$ as does θ_2

At the ω=w value at which

$$wT_o + tan^{-1}(\mathrm{a}w) + tan^{-1}(\mathrm{b}w) = \pi$$

$$|U(jw)| = \frac{K}{(1+a^2w^2)(1+b^2w^2)}$$

This value can be greater than 1,
Thus rendering the system unstable



Delayed Feedback: Stability Analysis

Recall that $$K = \frac{(CT_0)^3}{(2N)^2}$$

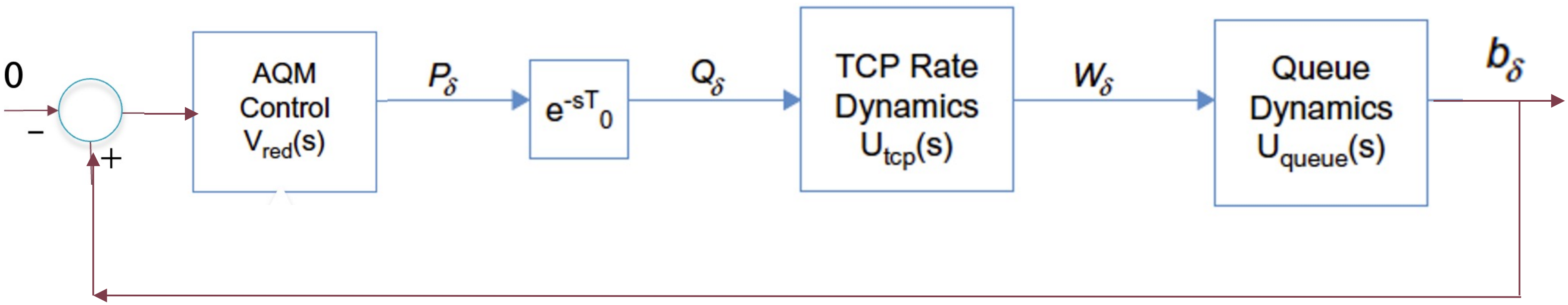
so that the system can become unstable in the presence of feedback if either
(1) C increases or
(2) T_0 increases or
(3) N decreases.

Thus, this shows that high link capacity or high round trip latency can cause system instability.

Instability can also be triggered for smaller values of N. Why?

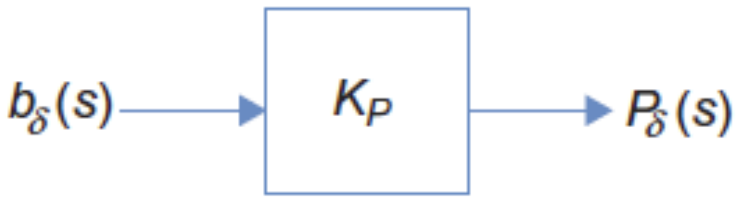
Proportional Controllers

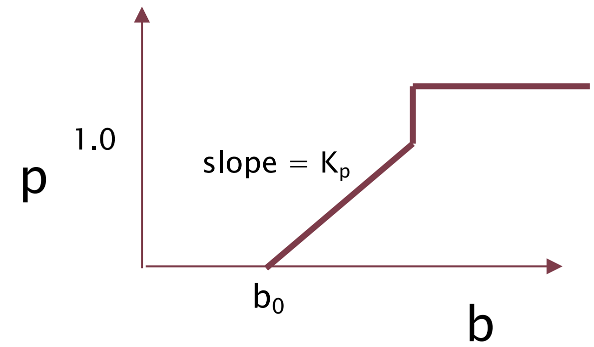
Controller with Probabilistic Drops



Because the On-Off Controller can become unstable, we now explore the option of dropping (or marking) packets in a probabilistic fashion if the queue size exceeds the threshold b_0.

Proportional Controllers

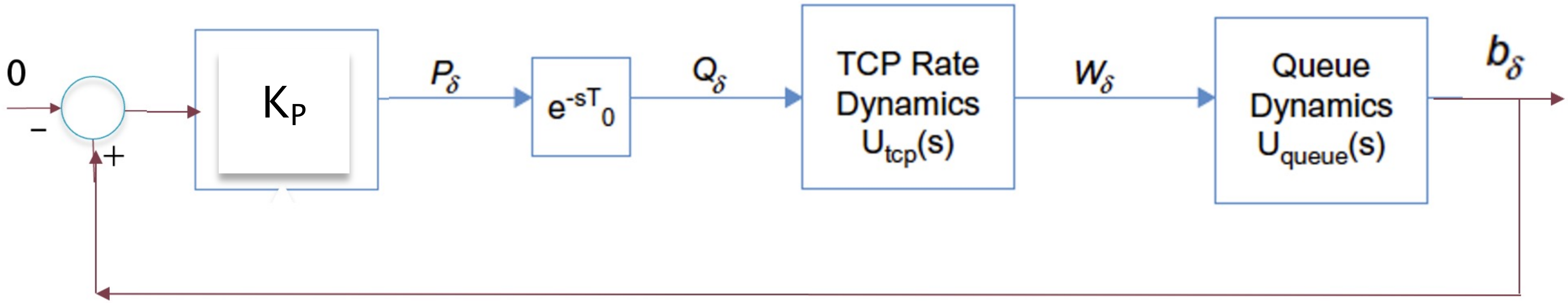




Once the queue length exceeds b_0, packets get marked with probability p

- The feedback signal is simply the regulated output (queue length) multiplied by a gain factor.
- In the RED context, it corresponds to obtaining the loss probability from the instantaneous queue length instead of the averaged queue length.

Proportional Controller



The Transfer Function is given by

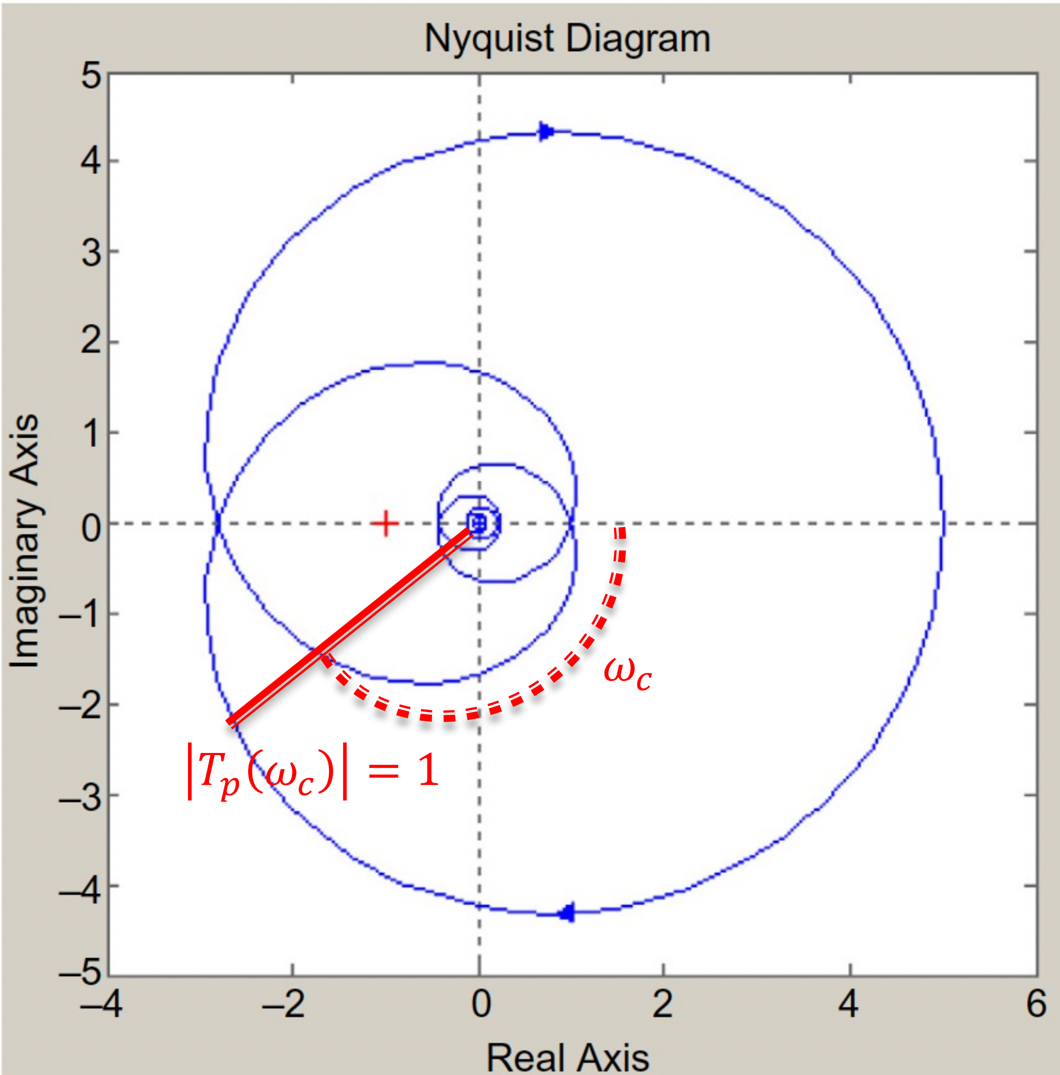
$$T_P(s) = U_{tcp}(s)U_{queue}(s)V_P(s)e^{-sT_0}$$

$$= \frac{\dfrac{K_P(CT_0)^3}{(2N)^2}e^{-sT_0}}{\left(\dfrac{s}{2N/CT_0^2}+1\right)\left(\dfrac{s}{1/T_0}+1\right)}$$

And in the Frequency Domain

$$T_P(j\omega) = |T_P(j\omega)|e^{-j\,\arg(T_P(j\omega))}$$

Proportional Controller: Stability Proof



Choose ω_c such that
$|T_p(\omega_c)| = 1$

Then show that
$\omega_c < 180^0$

This implies that the point where the locus touches the Real Axis, is also <1
Why?

Proportional Controller: Stability Proof

choose ω_c as the geometric mean of the roots $p_{TCP} = \frac{2N}{CT_0^2}$ and $p_{queue} = \frac{1}{T_0}$ so that

$$\omega_c = \sqrt{\frac{2N}{CT_0^3}}$$

Choose

$$K_P = \frac{\sqrt{\left(\frac{\omega_c}{2N/CT_0^2}\right)^2 + 1}\sqrt{\left(\frac{\omega_c}{1/T_0}\right)^2 + 1}}{\frac{(CT_0)^3}{(2N)^2}}$$

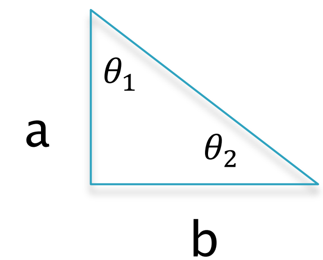
Then $|T_p(\omega_c)| = 1$

Note that this choice of K_p precisely cancels out the high loop gain caused by TCP's window dynamics.

Proportional Controller: Stability Proof

This equals 90 degrees (why?)

Also $$\arg(T_P(j\omega_c)) = \omega_c T_0 + \tan^{-1}\frac{\omega_c}{2N/C(T_0)^2} + \tan^{-1}\frac{\omega_c}{1/T_0}$$



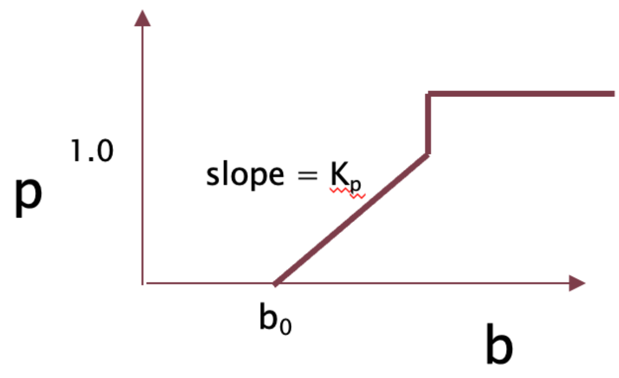
and $$\omega_c T_0 = \sqrt{\frac{2N}{CT_0}} = \sqrt{\frac{2}{W_0}} < 1 \; if \; W_0 > 2$$

So that $$\arg(T_P(j\omega_c)) < 1 + \tfrac{\pi}{2} \text{ radians} = 147^\circ$$

Hence, the Phase Margin is given by PM = 180^0 – 147^0 = 33^0 , which proves stability of the proportional controller.

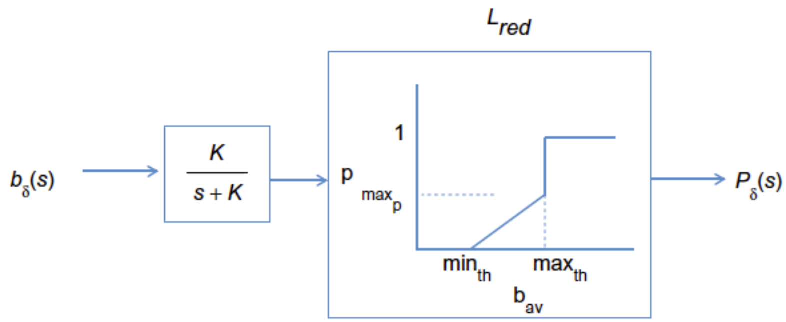
Performance

- The Proportional controller works better than the On-Off controller, however it has a few shortcomings
- The slope K_p is usually a small number to guarantee stability. However this means that if the buffer size is small, then the maximum drop rate may not be sufficient to keep the queue size near the operating point.
- This leads to a steady state error in the queue size. This can be driven down to zero using Proportional + Integral controllers.



RED Controller

RED Controller



$$P_\delta(s) = L_{red}(b - min_{th})$$

$$V_{red}(s) = \frac{L_{red}}{\frac{s}{K} + 1} \quad \text{where}$$

$$L_{red} = \frac{\max_p}{\max_{th} - \min_{th}} \quad \text{and} \quad K = -\frac{\log(1 - w_q)}{\delta}$$

The Transfer Function is given by

$$T(s) = U_{tcp}(s)U_{queue}(s)V_{red}(s)e^{-sT_0}$$

$$= \frac{\frac{L_{red}(CT_0)^3}{(2N)^2}e^{-sT_0}}{\left(\frac{s}{K} + 1\right)\left(\frac{s}{2N/CT_0^2} + 1\right)\left(\frac{s}{1/T_0} + 1\right)}$$

Stability Analysis

Note that

$$|T(j\omega)| = \frac{\frac{L_{red}(CT_0)^3}{(2N)^2}}{\sqrt{\left(\frac{\omega}{K}\right)^2+1}\sqrt{\left(\frac{\omega}{2N/CT_0^2}\right)^2+1}\sqrt{\left(\frac{\omega}{1/T_0}\right)^2+1}} \quad \text{and}$$

$$\arg(T(j\omega)) = \omega T_0 + \tan^{-1}\frac{\omega}{K} + \tan^{-1}\frac{\omega}{2N/CT_0^2} + \tan^{-1}\frac{\omega}{1/T_0}$$

Let L_{red} and K satisfy:

$$\frac{L_{red}(T^+C)^3}{(2N^-)^2} \le \sqrt{\frac{\omega_c^2}{K^2}+1} \quad \text{where}$$

$$\omega_c = 0.1\min\left\{\frac{2N^-}{(T^+)^2C}, \frac{1}{T^+}\right\}$$

then the linear feedback control system in Figure 3.7 using a RED controller is stable for all $N \ge N^-$ and all $T_0 \le T^+$.

Stability Analysis

Note that

$$|T(j\omega)| \leq \frac{\frac{L_{red}(CT_0)^3}{(2N)^2}}{\sqrt{\left(\frac{\omega}{K}\right)^2 + 1}}$$

if $\omega = \omega_c, N \geq N^-, T_0 \leq T^+$, then $|T(j\omega_c)| \leq 1$

From this, we conclude that the critical frequency ω_* at which $|T(j\omega_*)| = 1$, satisfies the relation $\omega_* \leq \omega_c$ (this is because $|T(j\omega)|$ is monotonically decreasing in ω). Hence, it follows from the Nyquist criterion that if we can show that the angle arg $T(j\omega_c)$ satisfies the condition

$$\arg T(j\omega_c) < \pi \text{ radians} \tag{72}$$

then the system has a positive phase margin (PM) given by

$$PM \geq \pi - \arg T(j\omega_c) > 0$$

Stability Analysis

From Equation 67 it follows that

$$\arg(T(j\omega_c)) = \omega_c T^+ + \tan^{-1}\frac{\omega_c}{K} + \tan^{-1}\frac{\omega_c}{2N^-/C(T^+)^2} + \tan^{-1}\frac{\omega_c}{1/T^+}$$

Since

$$\omega_c = 0.1 \min\left\{\frac{2N^-}{(T^+)^2 C}, \frac{1}{T^+}\right\}$$

it follows that

$$\omega_c T^+ = 0.1 \min\left\{\frac{2N^-}{T^+ C}, 1\right\} \quad \text{so that } \omega_c T \leq 0.1 \text{ radians.}$$

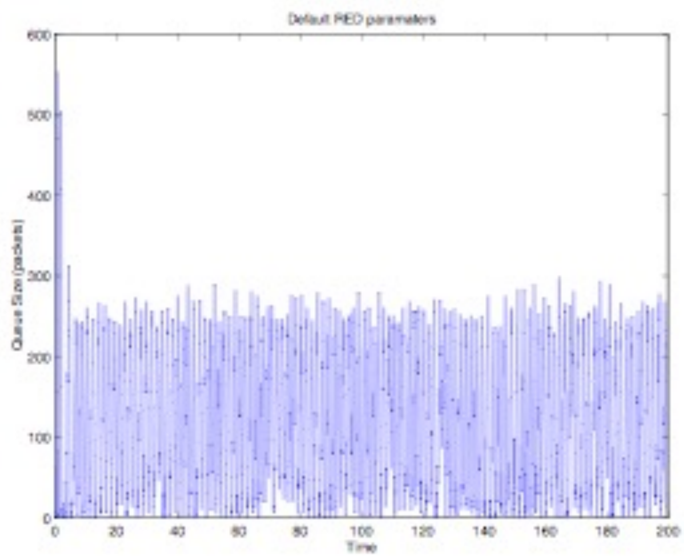
Also $\frac{\omega_c}{2N^-/C(T^+)^2} \leq 0.1$ and $\frac{\omega_c}{1/T^+} \leq 0.1$ so that

$\tan^{-1}\frac{\omega_c}{2N^-/C(T^+)^2} \leq 0.1$ and $\tan^{-1}\frac{\omega_c}{1/T^+} \leq 0.1$. Lastly, because

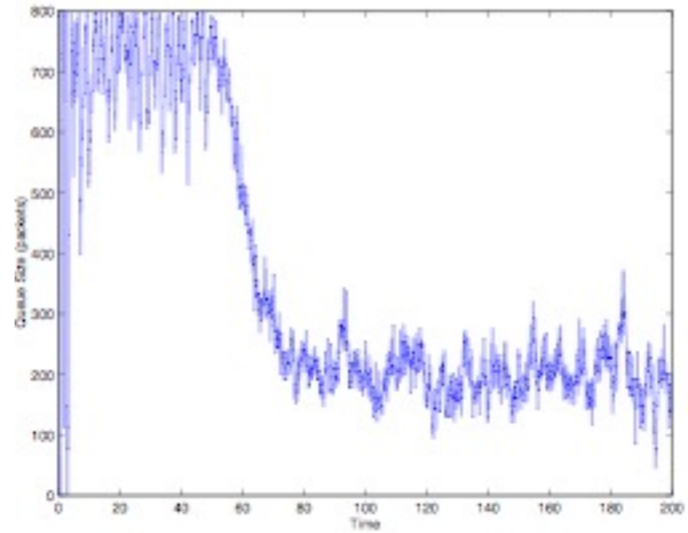
$\tan^{-1}\frac{\omega_c}{K} \leq \frac{\pi}{2}$, it follows that

$$\arg(T(j\omega_c)) \leq 0.1 + \frac{\pi}{2} + 0.1 + 0.1 = 1.87 \text{ radians or } 107^\circ,$$

Variation of the Queue Size under RED

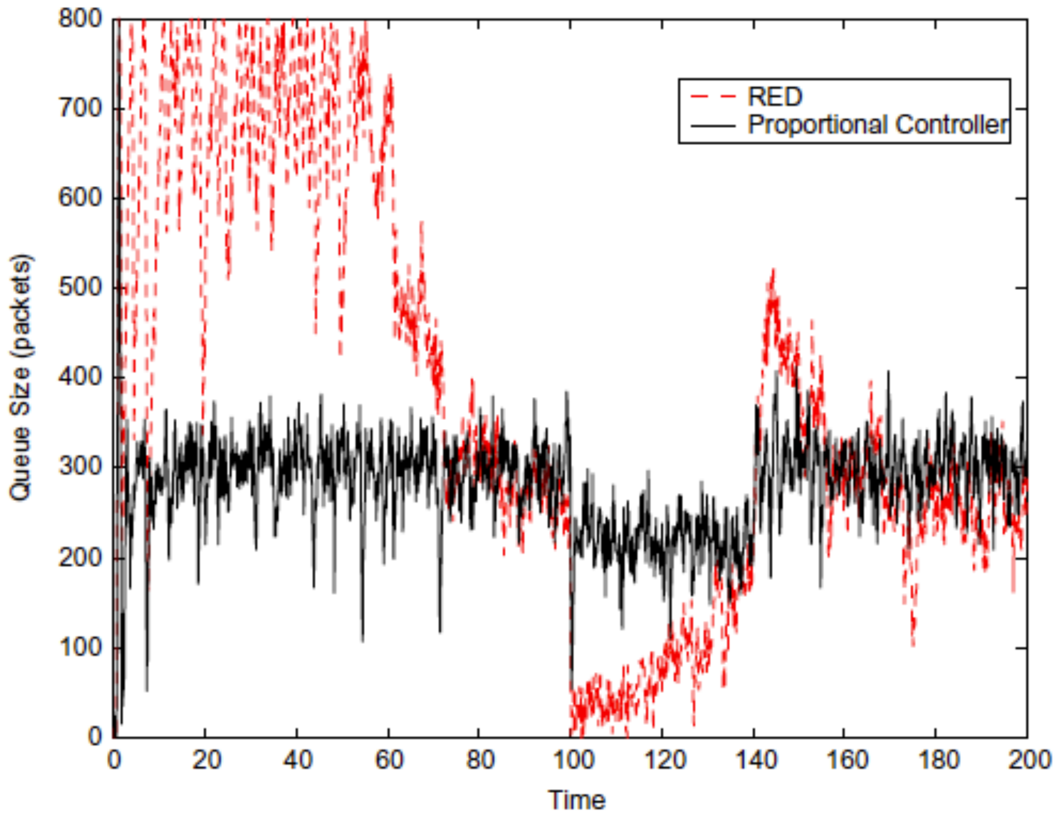


(a)



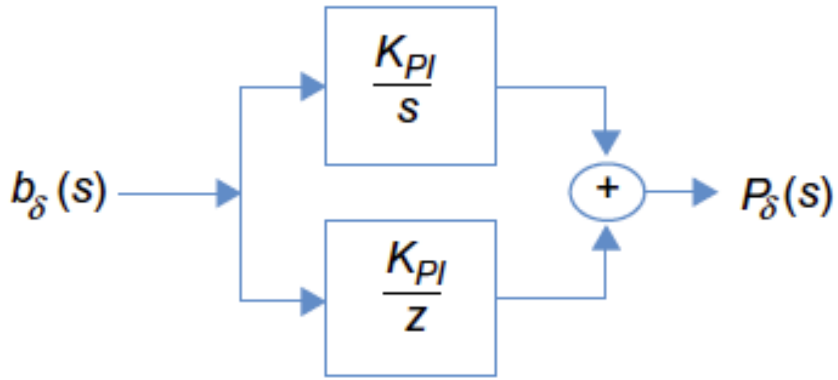
(b)

Comparison of the RED and Proportional Controllers



Proportional + Integral Controller

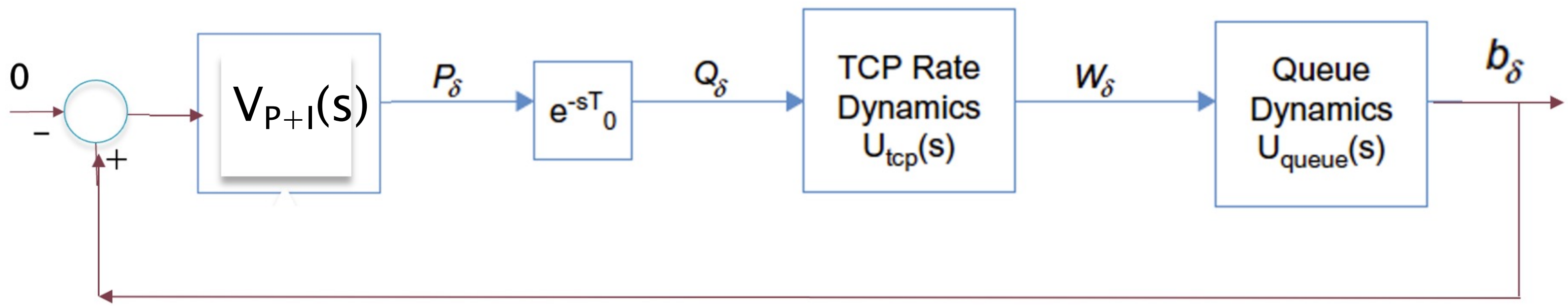
Proportional + Integral Controllers



Transfer Function of Controller

$$V_{PI}(s) = K_{PI}\frac{\left(\frac{s}{z}+1\right)}{s}$$

Proportional + Integral Controller



The Transfer Function of the system with PI control is given by

$$T_{PI}(s) = U_{tcp}(s)U_{queue}(s)V_{PI}(s)e^{-sT_0}$$

$$= \frac{\frac{K_{PI}(CT_0)^3}{(2N)^2}\left(\frac{s}{z}+1\right)e^{-sT_0}}{s\left(\frac{s}{2N/CT_0^2}+1\right)\left(\frac{s}{1/T_0}+1\right)}$$

And in the Frequency Domain

$$T_P(j\omega) = |T_P(j\omega)|e^{-j\ \arg(T_P(j\omega))}$$

P+I Controller: Stability

Choose the zero as $z = p_{tcp} = \frac{2N}{CT_0^2}$ and the critical frequency ω_c as $\omega_c = \frac{\beta}{T_0}$

If $K_{PI} = \frac{\omega_c\sqrt{(\omega_c T_0)^2 + 1}}{\frac{(CT_0)^3}{(2N)^2}}$ then $|T_{PI}(j\omega_c)| = 1$

To show stability, it is sufficient to show that $\arg(T_{PI}(j\omega_c)) < \pi$ radians

Note that

$$\arg(T_{PI}(j\omega_c)) = \frac{\pi}{2} + \omega_c T_0 + \tan^{-1}\omega_c T_0$$
$$= \frac{\pi}{2} + \beta + \tan^{-1}\beta$$

Given x = a +jb =|x|$e^{j\,\mathrm{Arg}(x)}$

|x| =$\sqrt{a^2 + b^2}$

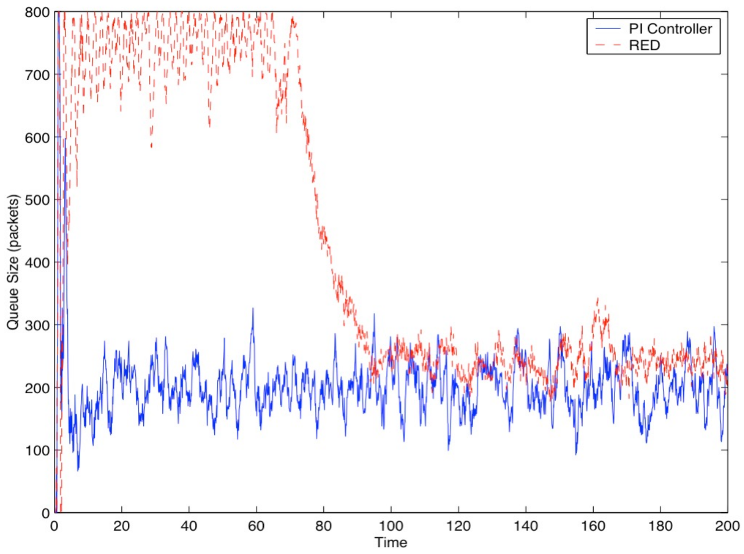
Arg(x) = $tan^{-1}\frac{b}{a}$ radians

If $0 < \beta < 0.85\ then\ \arg\big(T_{PI}(j\omega_c)\big) < \pi$

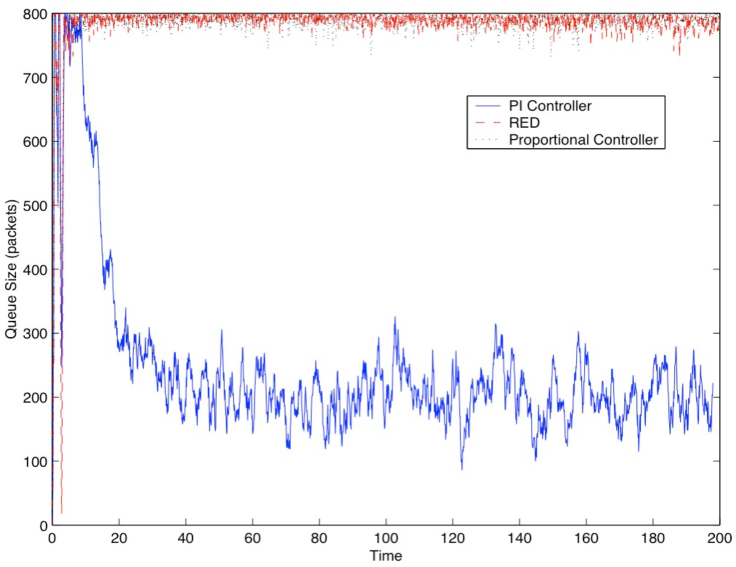
$= (3.142/2) + 0.85 + tan^{-1}(0.85) < \pi$

$$T_{PI}(s) = \frac{\frac{K_{PI}(CT_0)^3}{(2N)^2}\left(\frac{s}{z} + 1\right)e^{-sT_0}}{s\left(\frac{s}{2N/CT_0^2} + 1\right)\left(\frac{s}{1/T_0} + 1\right)}$$

Comparison between P+I and RED



(a)



(b)

- Simulation with a bottleneck queue size of 800 packets, in which the reference b_0 of the PI controller was set to 200 packets and the traffic consisted of a mixture of http and ftp flows.
- (a) clearly shows the faster response time PI compared with the RED controller as well as the regulation of the buffer occupancy to the target size of 200 packets.
- (b) shows the case when the number of flows has been increased to the point that the system is close to its capacity. Neither RED nor the proportional controllers are able to stabilize the queue because high packet drop rates have pushed the operating point beyond the size of the queue.

P+I Digital Implementation

$$V_{PI}(s) = K_{PI}\frac{\left(\frac{s}{z}+1\right)}{s}$$ becomes $$V_{PI}(z) = \frac{az-c}{z-1}$$

So that $$\frac{P(z)}{b_\delta(z)} = \frac{az-c}{z-1} = \frac{a-cz^{-1}}{1-z^{-1}}$$ where $$b_\delta = b - b_0$$

$P_\delta = P$ with $P_0 = 0$

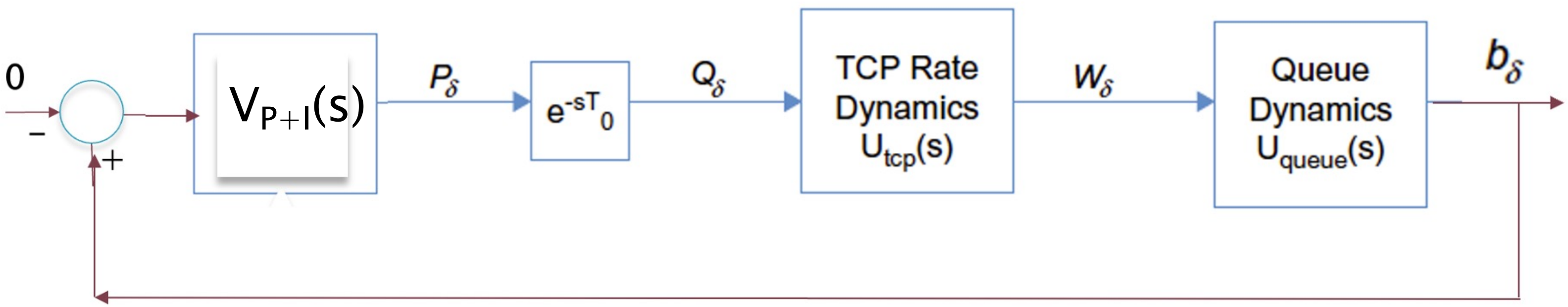
So that

$$P(kS) = ab_\delta(kS) - cb_\delta((k-1)S) + P((k-1)S)$$ $t = kS$, where $S = 1/f_s$

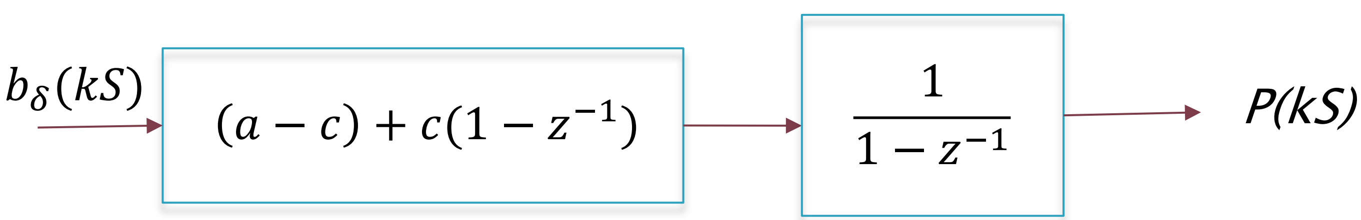
This can also be written as

$$P(kS) = (a-c)b_\delta(kS) + c(b_\delta(kS) - b_\delta((k-1)S) + P((k-1)S)$$

P+I Digital Representation



$$P(kS) = (a-c)b_\delta(kS) + c(b_\delta(kS) - b_\delta((k-1)S) + P((k-1)S)$$



The Importance of $\frac{db_\delta}{dt}$

$$P(kS) = (a - c)b_\delta(kS) + c(b_\delta(kS) - b_\delta((k-1)S) + P((k-1)S)$$

- The system converges when both $b_\delta(kS)$ and $(b_\delta(kS) - b_\delta((k-1)S)$ go to zero
- This implies that the queue length has converged to the reference value, and also the derivative of the queue length (since $(b_\delta(kS) - b_\delta((k-1)S)$ is an approximation to the derivative) has converged to zero.
- The derivative of the queue length converging to zero implies that the input rate to the flows to the router exactly matches the link capacity and there is no growth or drain in the router queue.
- If the input rate is lower than the link capacity, then the queue starts to drain, making the derivative negative and the marking probability gets correspondingly reduced.

This is the critical discovery: The best control signal should not only incorporate information about b_δ but also $\frac{db_\delta}{dt}$

$$\frac{db(t)}{dt} = \frac{W(t)}{T(t)}N(t) - C$$

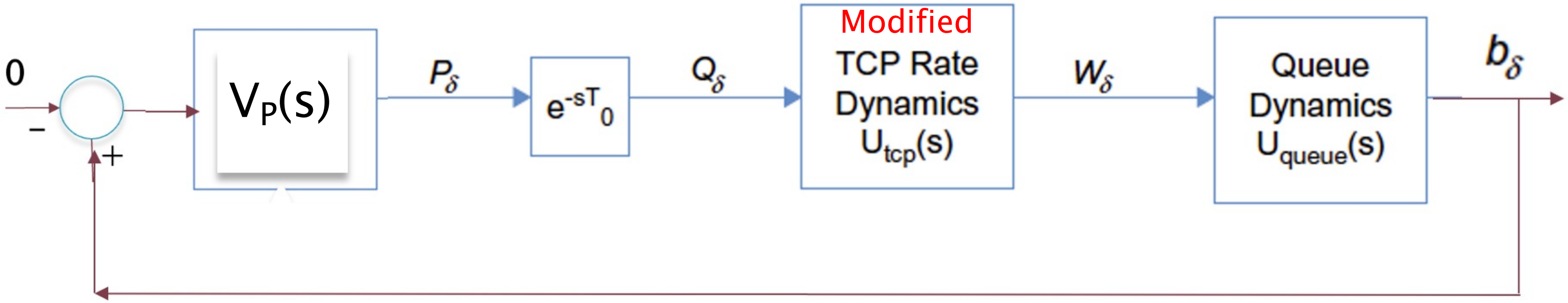
Influence of P+I Controller

- The TIMELY protocol for use in Data Center Networks
  - Uses derivative of delay rather than queue length.
- The QCN protocol that is part of the IEEE 802.1Qau Standard for use in Ethernet networks
  - The queue length + derivative value is fed back by means of a 6 bit field in the Ethernet Header
- Protocols such as XCP and RCP
  - P+I control is carried out at the nodes and the resulting rate fed back to the source.

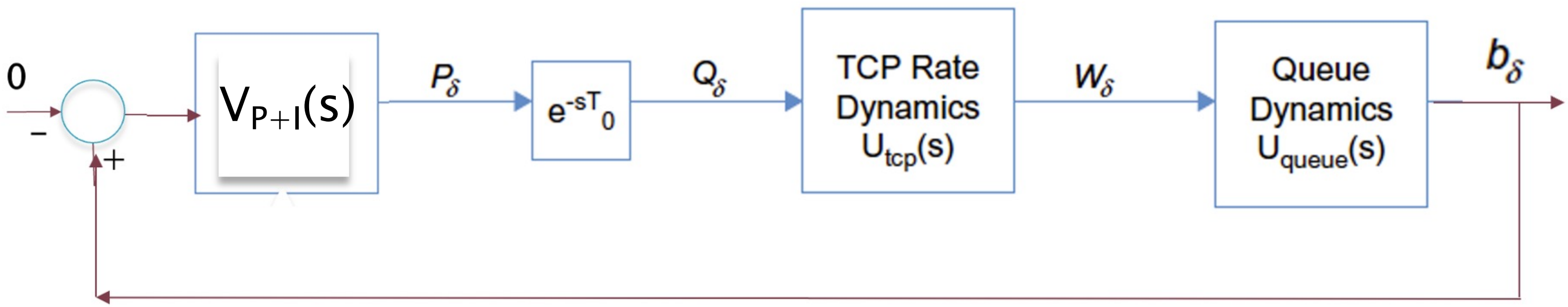
The Averaging Principle

"Stability Analysis of QCN: The Averaging Principle" Alizadeh et.al.

P Controller vs P+I Controller



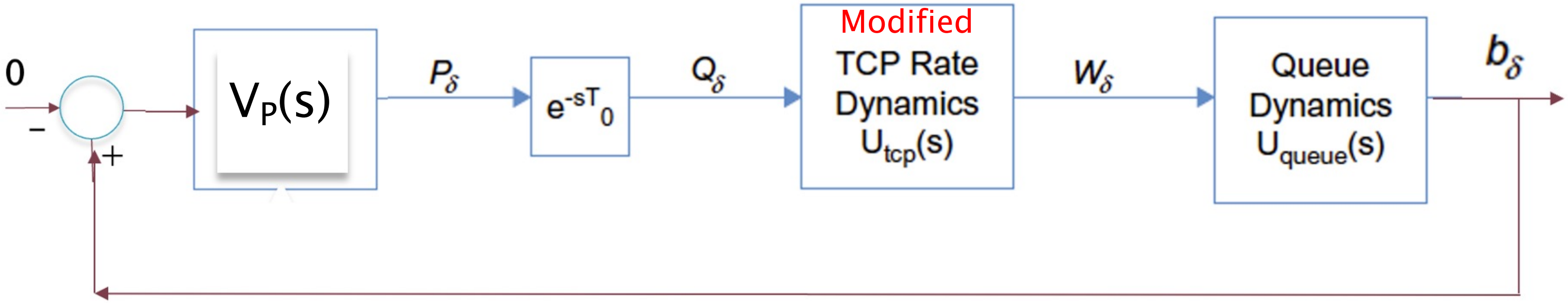
$$P(kS) = Kb_\delta(kS) + P((k-1)S)$$



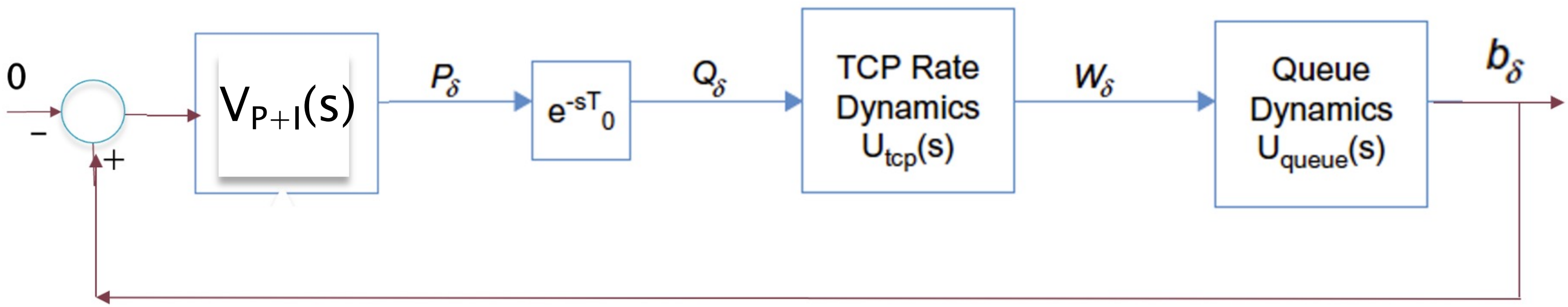
$$P(kS) = (a-c)b_\delta(kS) + c(b_\delta(kS) - b_\delta((k-1)S) + P((k-1)S)$$

Can we modify the Source Rate Control Algorithm so as to mimic the effect of the derivative information?

P Controller vs P+I Controller

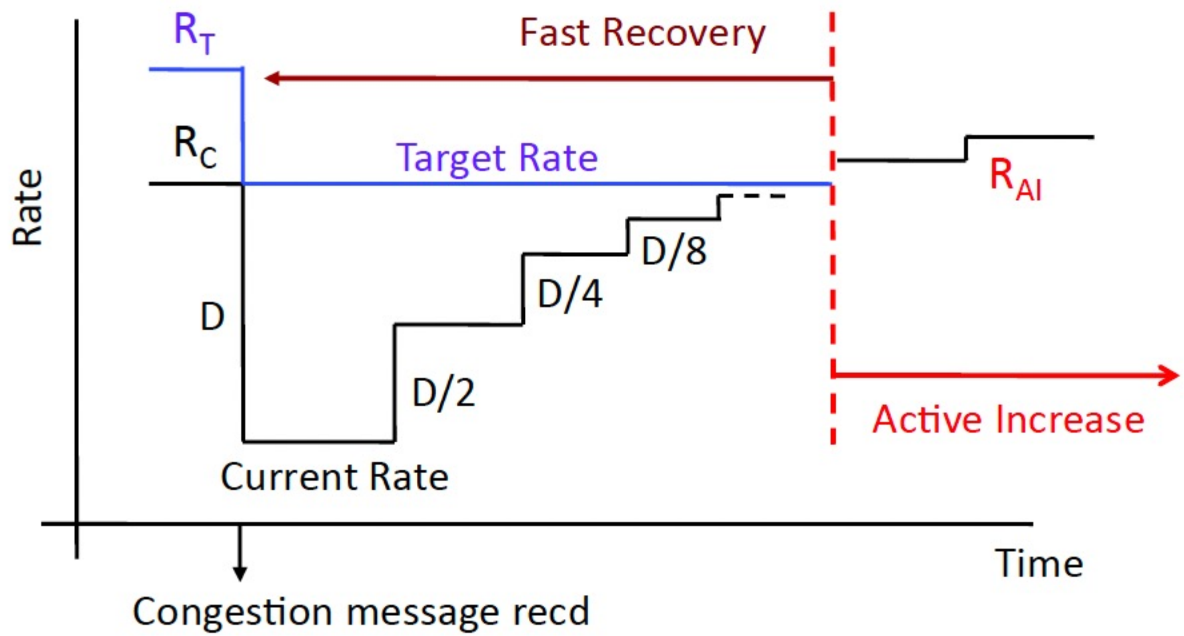


vs



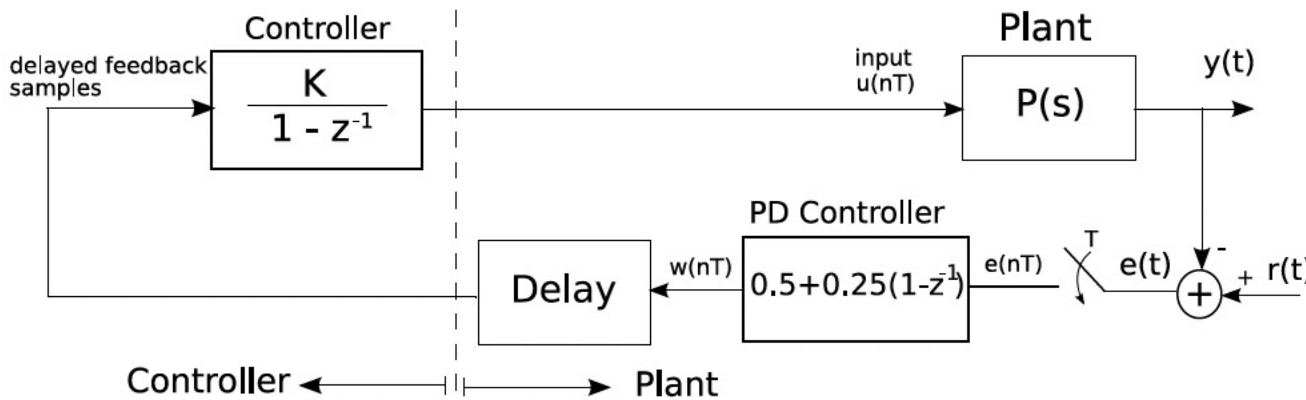
The two controllers can be shown to be equivalent, provided the TCP window control for the Proportional Controller obeys the Averaging Principle

Averaging Principle



In the absence of congestion, increase the TCP window using binary search, rather that increasing it by 1 every RTT.

Proof of the Averaging Principle



The Averaging Principle was shown to be true under the following assumptions:

- Rate control is being used at the source, as opposed to window based control.
- The error information is fed back directly to the source. As opposed to the error information being used to modify the packet marking probability, which in turn influences the TCP window dynamics.

Implications for Congestion Control Algorithms

- AQM schemes that incorporate the first or higher derivatives of the buffer occupancy process lead to more stable and responsive systems. This was shown to be the case in the analysis of the PI controller. Also, because the first derivative of the buffer occupancy process can be written as

$$\frac{db(t)}{dt} = C - \sum_i R_i$$

- It also follows that knowing the derivative is equivalent to knowing how close the queue is to its saturation point C.
- Some protocols such as QCN feed back the value of db/dt directly, and others such as XCP and RCP feed back the rate difference in the RHS of the equation.

Implications (cont)

- If the network cannot accommodate sophisticated AQM algorithms (which is the case for TCP/IP networks), then an AP-based algorithm can have an equivalent effect on system stability and performance as the derivative-based feedback. Examples of algorithms that have taken this route include the BIC and CUBIC algorithms

Further Reading

- Chapter 3 of Internet Congestion Control