Impact of Cross Traffic Burstiness on the Packet-scale Paradigm — An Extended Analysis

Rebecca Lovewell and **Jasleen Kaur**

Technical Report # TR11-007
Department of Computer Science
University of North Carolina at Chapel Hill

***Abstract*—The packet-scale paradigm is a novel framework for achieving ultra-high speed congestion control. Due to its reliance on finely-controlled inter-packet gaps, the paradigm is expected to be sensitive to transient burstiness in traffic encountered on bottleneck links. This paper uses a first-principles approach to study the impact of cross traffic burstiness on the efficiency of the packet-scale paradigm. It relies on a simple periodic on-off model for cross traffic and studies the interaction of the burstiness timescale, round-trip times, and the smoothing filters adopted by the paradigm. The analysis is validated against ns-2 simulations with a prototype. Our analysis helps gain fundamental insights on the impact of several factors.**

I. Introduction

The packet-scale paradigm is a recently proposed framework for congestion control that promises to scale up the state of the art in ultra-high speed congestion control by several orders of magnitude [1], [2].[1] A key feature of the paradigm is its use of short probing-streams to probe for available network bandwidth at fine timescales, using finely-controlled inter-packet spacings—this gives the paradigm its characteristic properties of scalability, adaptability to dynamic bandwidth, and queue-friendliness. Unfortunately, the excellent adaptability of the paradigm to dynamically varying bandwidth also raises significant concern about its sensitivity to transient short-scale burstiness in the cross traffic encountered on congested links. In particular, before this paradigm can be deployed in practice, it is important to ask: *to what extent does cross traffic burstiness impact the efficiency of the packet-scale paradigm?* In this paper, we use a first-principles analysis approach to partially study this issue.

Most congestion control protocols are analyzed for their steady-state throughput by conducting stochastic analysis of their window growth functions [3], [4], [5]. Most of these analyses incorporate the impact of packet losses and round trip-times (RTTs), considering these to be the major influences on the throughput of a transfer. In contrast, the packet-scale paradigm does not rely on window-based rate-control and relies on fundamentally different mechanisms for controlling the rate of a transfer. Furthermore, due to its agility at fine timescales, the paradigm is mostly limited by dynamics in the bottleneck queues (which, in turn, influence the available bandwidth, avail-bw). In particular, (i) by design, the steady-state throughput of long transfers is independent of path RTT; and (ii) due to its queue-friendly behavior, packet losses are expected to occur rarely (and are not expected to play a dominant role in limiting throughput). Consequently, existing analysis approaches are not suitable for studying the paradigm.

In this paper, we instead adopt a first-principles approach to analyze the impact of fine-scale traffic burstiness on the paradigm.[2] Our analysis considers a simple periodic on-off model for the cross traffic burstiness (parametrized by the burstiness timescale) and incrementally incorporates the impact of queue backlogs, RTTs, and the averaging timescales of the paradigm. We validate our analysis against experiments conducted with an ns-2 implementation of a prototype of the paradigm. Our analysis shows that:

- The degree to which traffic burstiness impacts the efficiency of the paradigm depends on both the average cross traffic load, as well as its characteristic ratio of burst-to-idle timescales.
- RTTs have little influence on the worst-case impact of traffic burstiness on the paradigm.
- The impact of cross-traffic burstiness on the paradigm can be alleviated by selecting the two smoothing timescales used by the paradigm such that: (i) these are reasonably larger than the typical burstiness timescale of the cross-traffic encountered, and (ii) the difference between the two smoothing timescales is small.

In the rest of this paper, we formulate the analysis framework in Section II. We develop and validate the analysis in Sections III-VI. We use the analysis to quantitatively study the performance of the paradigm under different conditions and settings in Section VII. We conclude in Section VIII.

II. Formulation

The available bandwidth of a link, observed during a time interval (t_1, t_2), is defined as:

$$B_{avail}(t_1, t_2) = C - \frac{B_{CT}(t_1, t_2)}{t_2 - t_1} \tag{1}$$

[1]This material is based upon work supported by the National Science Foundation under Awards CNS-1018596 and OCI-1127413.

[2]In this paper, we ignore dynamics that occur at timescales even smaller than a probing-stream and focus only on larger timescales.

where C is the transmission capacity of the link, and $B_{CT}(t_1, t_2)$ is the total amount of traffic that arrives on the link during the interval (t_1, t_2). The avail-bw of a link reflects its spare capacity. The end-to-end avail-bw of a path is defined as the minimum avail-bw among its constituent links [6].

The packet-scale paradigm attempts to measure the end-to-end avail-bw and adjust the sending rate of the transfer to match it. The Rapid protocol described in [1] is a prototype of the packet-scale paradigm. Below, we briefly describe the key mechanisms that control Rapid behavior in steady-state—for brevity, we provide only details that are relevant to the discussion in this paper. In particular, since our analysis focuses on timescales larger than probing-streams, our description is light on mechanisms that operate at sub-probing-stream timescales.

A. Rapid: Operation in Steady-state

The Rapid sender continuously sends back-to-back probing-streams (p-streams) of length N packets, where each packet is sent at a potentially different probing rate by controlling the inter-packet gaps. The Rapid receiver records the arrival time of packets at the receiver, and returns these to the sender. The sender compares the receiver-gaps to the original gaps and computes an estimate for the avail-bw according to the *highest probing rate for which the inter-packet gaps did not show an increasing trend*. Details of the bandwidth-estimation logic and the structure of the p-streams can be found in [1]. In this paper, we ignore packetization effects and assume fluid behavior at p-stream and lower timescales—this implies that a Rapid p-stream returns a perfect estimate of the avail-bw.

The avail-bw estimate, B_{est}, computed by the sender is then fed through a moving average filter that updates the average sending rate, $\widehat{R}$, used for the next p-stream as follows[3]:

$$\begin{aligned} \widehat{R} &= \widehat{R} + \frac{L}{\tau}(B_{est} - \widehat{R}), \quad \text{if } \widehat{R} > B_{est}, \\ \widehat{R} &= \widehat{R} - \frac{L}{\eta}(\widehat{R} - B_{est}), \quad \text{otherwise} \end{aligned}$$

where $L = \frac{N*P}{\widehat{R}}$ is the duration in time of the most recent p-stream, P is the packet size, and τ and η are the smoothing parameters employed by Rapid for increasing or decreasing the sending rate, respectively. Effectively, by employing the above moving average filter, when avail-bw increases (or decreases), Rapid linearly increases (or decreases) its average sending rate to the new avail-bw over τ (or η) time units.

B. Cross Traffic Model

According to Equation (1), the avail-bw is determined by the cross traffic arrival process. Note that the avail-bw influences a Rapid p-streams via dynamics in the (shared) packet queue at the bottleneck link. In particular, the queue-size grows when the collective bit-rate of the p-streams and the arriving cross traffic exceeds the link capacity—it is precisely this growth in the bottleneck queue that Rapid relies on for estimating the avail-bw.

[3]Note that this description is different from that of the Rapid protocol in [1], which does not use η (or equivalently, sets: $\eta = L$).

Fig. 1. Periodic On-off Cross Traffic Model

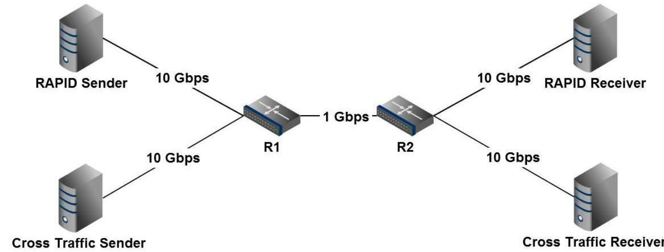


Fig. 2. Topology for Experimental Validation

Bursty cross traffic creates additional (transient) queue dynamics. Since Rapid is capable of estimating and adapting to avail-bw at fine timescales, it also reacts to the transient queue build-up caused by the cross traffic bursts. In order to study the impact of such bursts on Rapid p-streams and the steady-state throughput it achieves, we consider a simple model of the cross traffic in this paper—a periodic on-off source. Such a traffic source, illustrated in Fig 1, is characterized by the fixed duration of each burst, T_{on}, the inter-burst idle time, T_{off}, and the on-rate, R_{on}. The average cross traffic rate is then give by: $R_{avg} = \frac{R_{on} T_{on}}{T_{on} + T_{off}}$.

Note that for a given R_{avg}, traffic can be made more "bursty" by increasing R_{on}, and the timescale of burstiness can be controlled using T_{on}. Given this model for cross traffic, we analyze Rapid using the following metrics.

C. Metrics

In the ideal absence of burstiness in cross traffic, Rapid should be able to fully utilize the steady-state avail-bw [1]. That is, the *steady-state Rapid throughput*, $\mathcal{R}$, should be equal to $C - R_{avg}$. In order to see how cross traffic burstiness impacts Rapid throughput, in our analysis we study both $\mathcal{R}$ and the overall link utilization, given by:

$$\mathcal{U} = \frac{R_{avg} + \mathcal{R}}{C} \tag{2}$$

D. Experimental Setup

We use the ns-2 implementation of the Rapid protocol to validate our analysis against. We generate a simple dumbell topology illustrated in Fig 2. The Rapid transfer and the cross traffic share the 1 Gbps bottleneck link. All other links have a transmission capacity of 10 Gbps. The RTT of the Rapid transfer is controlled by varying the propagation delays on the links attached to the Rapid sender and the Rapid receiver. Sufficient queues are provisioned on all links to avoid packet losses. The cross traffic is driven by a periodic CBR on-off source—the source alternates between an *on* state, in which it send traffic at a fixed bit-rate of R_{on} for an interval of fixed length T_{on}, and an *off* state in which it does not send any traffic for an interval of fixed length T_{off}. The cross traffic parameters are varied across experiments.

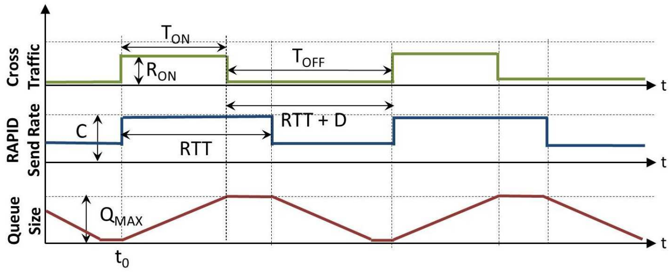
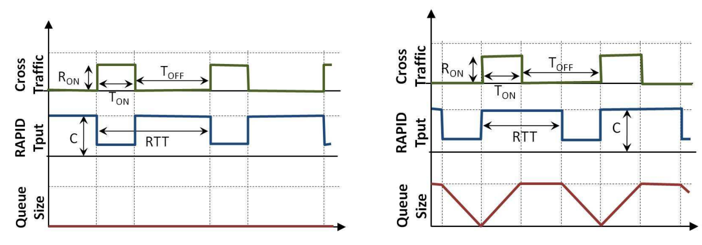


Fig. 3. Impact of RTT: Illustration



(a) No Queue Buildup

(b) Max Queue Buildup

Fig. 4. Best-case and Worst-case Queue Buildups

III. IMPACT OF QUEUE BACKLOG

We begin our analysis with a simple but important observation—which will be used in subsequent analysis—about the influence of a *residual* queue [4] at the bottleneck link on the avail-bw estimates yielded by a Rapid p-stream:

***Observation** 1:* The presence of a residual queue at the bottleneck link does not influence the avail-bw estimates yielded by an arriving Rapid p-stream—the avail-bw estimates in this case will simply depend on the "arrival" process of the cross traffic.

This is not an immediately obvious observation, since the packets of a p-stream that arrives behind a residual queue are likely to bunch up together while the residual queue drains. While this is certainly true, the inter-packet gaps at the time of *departure* will still be shaped according to the cross traffic that arrives *along with* the p-stream (and which gets inserted in between the p-stream packets)—the departure gaps are independent of the amount of residual queue encountered by the arriving p-stream. In Appendix IX, we formally establish this for the case when two successive probing packets belong to the same *busy period* of the router link.

IV. ROLE OF RTT

A Rapid sender learns of a change in the avail-bw only after a feedback delay of RTT time units. We first investigate the role that this feedback delay might play in influencing Rapid's efficiency in utilizing the avail-bw. In order to solely focus on this factor, we set: $\tau = \eta = L$ — this ensures that the Rapid sender sets its sending rate equal to a new avail-bw estimate immediately upon learning about it (no smoothing/averaging delay is involved).

Fig 3 illustrates a typical timeline, which includes the periodic on-off cross traffic arrival rate, the Rapid sending rate, and the queue buildup at the bottleneck link. The cross traffic alternates between bit-rates of R_{on} and 0. Observe that:

- When a cross traffic burst arrives at the bottleneck link at time t_0, it interacts with the arriving Rapid p-streams—the Rapid sender learns that the avail-bw has decreased to $C - R_{on}$ after a delay of 1 RTT at time: $t_0 + RTT$. The Rapid sender immediately reduces its sending rate from C to $C - R_{on}$ at this time.
- The cross traffic burst lasts till $t_0 + T_{on}$. Note that the queue-backlog at the bottleneck link keeps increasing during the time interval $[t_0, t_0 + T_{on}]$ (since the sum of Rapid and cross traffic arrival rates exceeds C). However, because of Observation 1, the Rapid p-streams arriving in this interval yield avail-bw estimates equal to $C - R_{on}$.
- The Rapid p-streams that arrive at the bottleneck link immediately after $t_0 + T_{on}$ will yield avail-bw estimates equal to C (despite the presence of a residual queue, from Observation 1). However, these p-streams will experience an inflated delay in reaching the receiver (due to the extra queuing delay waiting for the residual bottleneck queue to drain). In fact, the *first* notification that the avail-bw has now increased to C will reach the sender only after time: $t_0 + T_{on} + RTT + D$, where D is the queue drain time of the maximum residual queue, given by: $D = \frac{Q_{max}}{C}$. Therefore, the Rapid sender will increase its sending rate back to C only at time $t_0 + T_{on} + RTT + D$.
- This whole interaction repeats with the arrival of the next cross traffic burst.

The average Rapid throughput can be computed by averaging its sending rate over the interval $[t_0, t_0 + T_{on} + T_{off}]$ as:

$$\begin{aligned} \mathcal{R} &= \frac{C \cdot (T_{off} - D) + (C - R_{on}) \cdot (T_{on} + D)}{T_{on} + T_{off}} \\ &= C - R_{avg} \frac{T_{on} + D}{T_{on}} \end{aligned} \tag{3}$$

Note that while RTT does not appear directly in the above term, it may indirectly influence the maximum queuing delay, D, which adversely impacts the Rapid throughput.

D, in turn, depends on the phase lag between the cross traffic bursts and Rapid's response. This lag is governed by: $RTT_{rem} = remainder(\frac{RTT}{T_{on}+T_{off}})$. For instance, when $RTT_{rem} = 0$ (as illustrated in Fig 4(a)), there is *no* queue buildup at the bottleneck link since the cross traffic peaks overlap perfectly with the Rapid troughs. Thus, $Q_{max} = 0$ and $D = 0$. On the other hand, Fig 4(b) illustrates that the worst-case queue buildup occurs when the peaks and troughs do not overlap in time at all (for instance, when $RTT_{rem} > T_{on}$ and $T_{off} > T_{on}$). In this case, $Q_{max} = T_{on} R_{on}$, and $D = \frac{T_{on} R_{on}}{C}$.

Note that when the peaks and troughs overlap only partially (for instance, when $RTT_{rem} < T_{on}$), the queue-buildup is smaller than the worst-case and is proportional to RTT_{rem}. Since our objective is to quantify the extent to which burstiness can adversely impact Rapid performance, we focus only on the

[4]The residual queue refers to buildup which already exists on the queue before the first packet of a Rapid p-stream arrives.

worst-case queue buildup, for which:

$$\mathcal{R} = C - R_{avg} - \frac{R_{avg}^2}{C}\frac{T_{on}+T_{off}}{T_{on}}$$
$$\mathcal{U} = \frac{R_{avg}+\mathcal{R}}{C} = 1 - \frac{R_{avg}}{C}^2\frac{T_{on}+T_{off}}{T_{on}}$$

This implies that the role of RTT is limited to influencing the phase shift between the cross traffic peaks and Rapid troughs. *The worst-case loss in utilization is independent of the path RTT*. Henceforth, we focus only on analyzing the worst-case queue buildups.

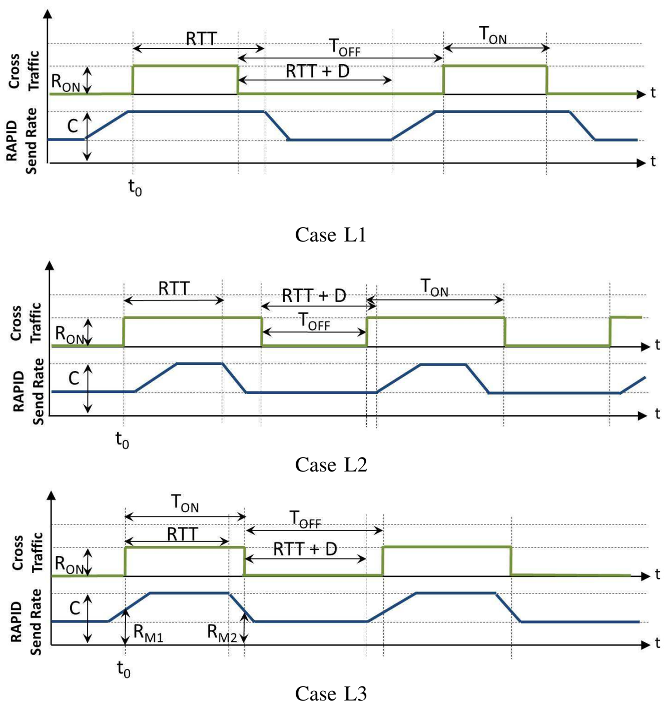


Fig. 5. Impact of Large-scale Bursts: Illustration

V. Large-scale Bursts

Note that the parameters τ and η introduce additional delay before Rapid completely adapts to an increase or decrease in avail-bw. For studying the influence of τ and η, we first consider the cases in which the cross traffic bursts are long enough to allow Rapid to converge to all avail-bw changes. We define large-scale cross traffic bursts to be such that $T_{on}+D \geq \eta$ and $T_{off} - D \geq \tau$ in the steady state, so the Rapid sender will always have enough time to completely adapt to a change in avail-bw before it changes again. Observe:

- At t_0, a cross traffic burst arrives at the bottleneck link and the router queue starts to build up. Rapid learns of the corresponding decrease in avail-bw after a delay of 1 RTT. It then takes η time units to converge its sending rate to the lower value of $C - R_{on}$.
- The cross traffic burst stops at $t_0 + T_{on}$. Since the queue builds up in the meantime, however, Rapid learns of the corresponding increase in avail-bw only after a delay of $RTT + D$, where D is the time it takes for the additional queue buildup to drain. It then takes τ time units to increase its sending rate from $C - R_{on}$ to C.
- All of this occurs before Rapid learns of another decrease in avail-bw.

In this case, the steady-state Rapid throughput $\mathcal{R}$ can be time-averaged as:

$$\begin{aligned}\mathcal{R} &= (C - \frac{R_{on}}{2})\frac{\tau+\eta}{T_{on}+T_{off}} + C\frac{T_{off}-\tau-D}{T_{on}+T_{off}} \\ &\quad +(C-R_{on})\frac{T_{on}-\eta+D}{T_{on}+T_{off}} \\ &= C - R_{avg}\frac{2T_{on}+2D+\tau-\eta}{2T_{on}}\end{aligned}$$

Loss in Rapid throughput is proportional to D, the delay in Rapid's response to an increase in avail-bw. This delay is directly proportional to the maximum queue buildup which occurs during a cross traffic burst and is described by $D = \frac{Q_{max}}{C}$. The worst-case maximum queue buildup occurs when Rapid traffic aligns with the cross traffic bursts in such a way that the number of bytes that Rapid sends during a cross traffic burst is maximized. We use RTT'_{rem} to denote an RTT that creates such a worst-case alignment. The three formulations below show how RTT'_{rem} and Q_{max} are influenced by the values of T_{on}, T_{off}, τ, and η in the presence of large bursts. In each of the proceeding cases, we solve for the given values using mathematical software. The closed-form solutions are in general too unwieldy to present here.

A. Case L1: $D \leq T_{off} - T_{on} - \tau$

When cross traffic bursts are large and $D \leq T_{off} - T_{on} - \tau$, the worst-case queue buildup will occur when $RTT'_{rem} \in [T_{on}, T_{off} - D - \tau]$. In this case, the Rapid sender sends at capacity during the entire cross traffic burst period, and the queue buildup is described by

$$\begin{aligned}Q_{max} &= R_{on}T_{on} \\ &= R_{avg}\,(T_{on}+T_{off})\end{aligned}$$

When the above Q_{max} is substituted for, the steady-state Rapid throughput and link utilization can be simplified as:

$$\begin{aligned}\mathcal{R} &= C - R_{avg} - \frac{R_{avg}^2}{C}\frac{T_{on}+T_{off}}{T_{on}} - R_{avg}\cdot\frac{\tau-\eta}{2T_{on}} \\ \mathcal{U} &= 1 - \left(\frac{R_{avg}}{C}\right)^2\frac{T_{on}+T_{off}}{T_{on}} - \frac{R_{avg}}{C}\cdot\frac{\tau-\eta}{2T_{on}} \qquad (4)\end{aligned}$$

B. Case L2: $D \geq T_{off} - T_{on} + \eta$

When cross traffic bursts are large and $D \geq T_{off} - T_{on} + \eta$, the worst-case queue buildup occurs when $RTT'_{rem} \in [T_{off} - D, T_{on} - \eta]$. In this case, the Rapid traffic burst lies completely within the cross traffic burst period, and the queue buildup is described by

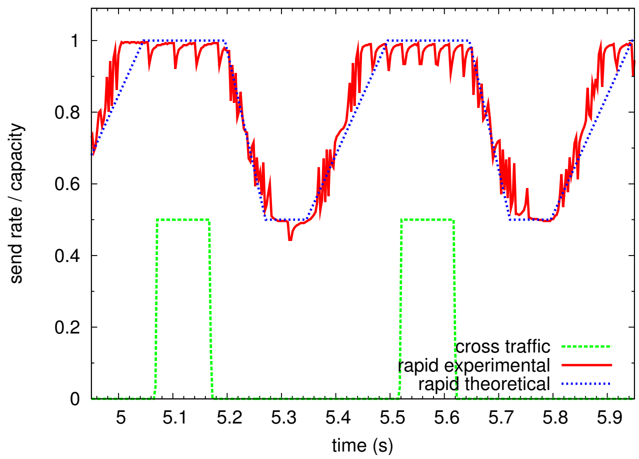


Fig. 6. Large-scale Bursts: Experimental Validation

$$
\begin{aligned}
Q_{max} &= C\left(T_{off} - D - \tau\right) + \left(C - \frac{R_{on}}{2}\right)(\tau + \eta) + \\
&\quad \left(C - R_{on}\right)\left(T_{on} - T_{off} + D - \eta\right) - \\
&\quad T_{on}\left(C - R_{on}\right) \\
&= R_{avg}\left(\frac{T_{on} + T_{off}}{T_{on}}\right)\left(\frac{2T_{off} - 2D - \tau + \eta}{2}\right)
\end{aligned}
$$

C. Case L3: $D \in (T_{off} - T_{on} - \tau, T_{off} - T_{on} + \eta)$

When cross traffic bursts are large and $D \in (T_{off} - T_{on} - \tau, T_{off} - T_{on} + \eta)$, only part of the Rapid traffic burst can fit within the cross traffic burst period. When the cross traffic burst begins, Rapid increasing its sending rate and has an instantaneous send rate of R_{M1}. When the cross traffic burst ends, Rapid is lowering its send rate and has an instantaneous send rate of R_{M2}. In Appendix X, we establish that the worst-case queue buildup occurs when RTT is such that $R_{M1} = R_{M2}$. This occurs when

$$
RTT'_{rem} = T_{on} - \left(\frac{\eta}{\tau + \eta}\right)\left(T_{on} - T_{off} + D + \tau\right)
$$

In this case, the queue buildup is described by

$$
\begin{aligned}
Q_{max} &= \left(\frac{C - R_M}{2}\right)\left(T_{off} - T_{on} - D - \tau\right) + \\
&\quad R_{on}T_{on} \\
R_M &= C - R_{on}\frac{T_{on} - T_{off} + D + \tau}{\tau + \eta}
\end{aligned}
$$

where $R_M = R_{M1} = R_{M2}$ is the rate at which the Rapid sender is sending upon both the start and the end of the cross traffic burst.

Validation We validate the above analysis against an ns-2 experiment in which: $T_{on} = 100ms$, $T_{off} = 350ms$, $R_{avg} = 111Mbps$, $\tau = 150ms$, $\eta = 75ms$, and $RTT = 125ms$. Fig 6 plots a sample of the theoretical and experimentally observed time series for the Rapid sending rate. We find that the experiment validates our fluid analysis quite well. We have also conducted several additional experiments (i) by varying the above parameters, and (ii) by using an exponential on-off cross traffic source —our measured values for link utilization match the predicted one very well in all cases.

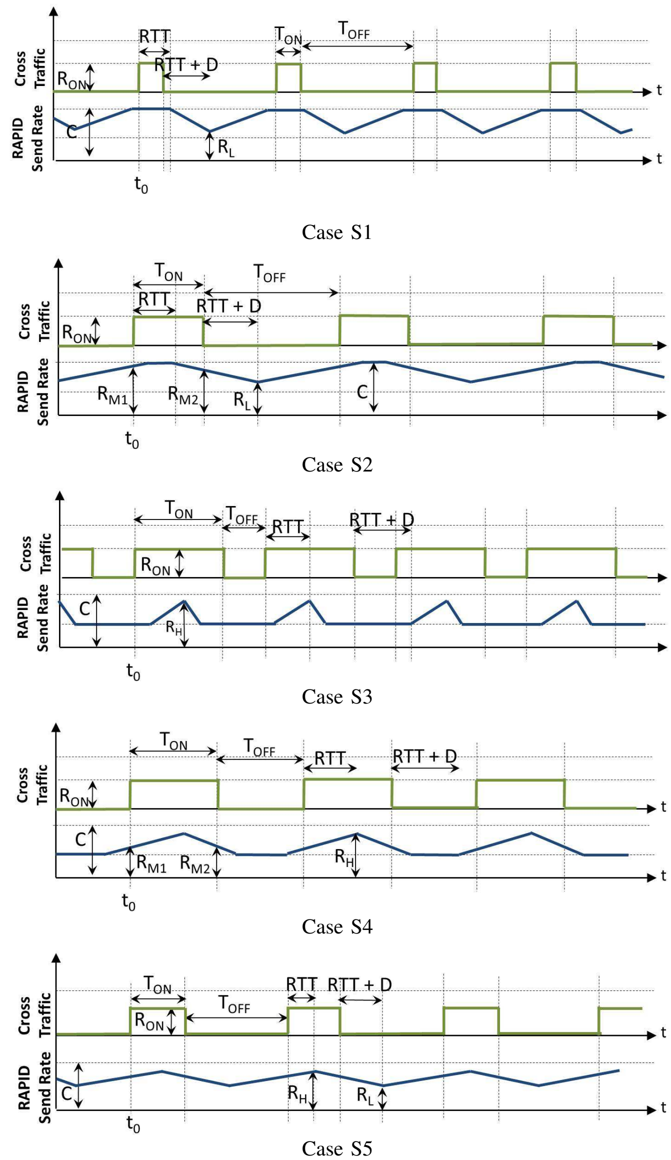


Fig. 7. Impact of Small-scale Bursts: Illustration

VI. SMALL-SCALE BURSTS

We next consider the cases in which the cross traffic bursts are fairly short and Rapid is unable to fully adapt to a change in avail-bw, before the latter changes again—this happens when $T_{on}+D < \eta$ and/or $T_{off}-D < \tau$. When $T_{off}-D < \tau$, Rapid will reach a maximum sending rate of only $R_H < C$ before learning of a decrease in avail-bw; similarly, when $T_{on}+D < \eta$, Rapid will be able to reach a mimimum sending rate of only $R_L > C - R_{on}$ before the avail-bw increases again.

As in the large-scale burst cases, Rapid will experience an additional delay when learning about an increase in available bandwidth due to queue buildup during the cross traffic burst.

This delay in Rapid's response is directly proportional to the maximum queue buildup during the cross traffic burst and is described by $D = \frac{Q_{max}}{C}$. In all small-burst cases the worst-case maximum queue buildup occurs when Rapid traffic aligns with the cross traffic bursts in such a way that the number of bytes that Rapid sends during a cross traffic burst is maximized. We use RTT'_{rem} to denote an RTT that creates such a worst-case alignment. The five formulations below show how RTT'_{rem}, Q_{max}, and Rapid's time-averaged throughput $\mathcal{R}$ are influenced by the values of T_{on}, T_{off}, τ, and η in the presence of small bursts. In each of the proceeding cases, we solve for the given quantities using mathematical software. The closed-form solutions are in general too unwieldy to present here.

A. *Case S1:* $T_{on}+D < \eta$, $T_{off}-D \geq \tau$, $D \leq T_{off}-T_{on}-\tau$

When $T_{off}-D \geq \tau$ but $T_{on}+D < \eta$, the Rapid sender will have enough time to converge on an higher sending rate when it receives a higher avail-bw estimate, but when it learns of a decrease in avail-bw it will not have sufficient time to converge on a lower sending rate before the avail-bw increases again. The lowest send rate it will achieve is $R_L > C - R_{on}$.

The lowest sending rate R_L and and the time-averaged steady-state Rapid throughput $\mathcal{R}$ are described by

$$\begin{aligned}
R_L &= C - R_{on}\left(\frac{T_{on}+D}{\eta}\right) \\
&= C - R_{avg}\left(\frac{T_{on}+T_{off}}{T_{on}}\right)\left(\frac{T_{on}+D}{\eta}\right) \\
\mathcal{R} &= \left(\frac{C+R_L}{2}\right)\left(\frac{T_{on}+D+\tau}{T_{on}+T_{off}}\right)+ \\
&\quad C\left(\frac{T_{off}-D-\tau}{T_{on}+T_{off}}\right) \\
&= C - R_{avg}\frac{(T_{on}+D)(T_{on}+D+\tau)}{2T_{on}\cdot\eta}
\end{aligned}$$

When $T_{on} \leq T_{off} - D - \tau$, the worst-case maximum queue buildup occurs when $RTT'_{rem} \in [T_{on}, T_{off}-D-\tau]$. Because the Rapid sender will be sending data at capacity during the entire cross traffic burst period, the maximum queue buildup is described by

$$\begin{aligned}
Q_{max} &= R_{on}T_{on} \\
&= R_{avg}(T_{on}+T_{off})
\end{aligned}$$

B. *Case S2:* $T_{on}+D < \eta$, $T_{off}-D \geq \tau$, $D > T_{off}-T_{on}-\tau$

As in Case S1, the Rapid sender will have enough time to converge on a higher sending rate when it receives a higher avail-bw estimate, but when it learns of a decrease in the avail-bw estimate it will not have sufficient time to converge on a lower sending rate before the avail-bw increases again. The lowest sending rate R_L and the time-averaged steady-state Rapid throughput $\mathcal{R}$ are calculated in the same manner as in Case S1:

$$\begin{aligned}
R_L &= C - R_{avg}\left(\frac{T_{on}+T_{off}}{T_{on}}\right)\left(\frac{T_{on}+D}{\eta}\right) \\
\mathcal{R} &= C - R_{avg}\frac{(T_{on}+D)(T_{on}+D+\tau)}{2T_{on}\tau}
\end{aligned}$$

However, because $T_{on} > T_{off} - D - \tau$, Rapid's traffic burst is too large to be contained entirely within the cross traffic burst period; only part will lie within the burst. In Appendix X, we establish that the worst-case maximum queue buildup occurs when RTT is such that Rapid's instantaneous send rate when the cross traffic burst begins, R_{M1}, is equal to the instantaneous send rate when the cross traffic burst ends, R_{M2}. This occurs when

$$RTT'_{rem} = T_{on} - \frac{T_{on}+D}{T_{on}+D+\tau}(T_{on}-T_{off}+D+\tau)$$

In this case, the maximum queue buildup is described by

$$\begin{aligned}
Q_{max} &= \left(\frac{C-R_M}{2}\right)(T_{off}-T_{on}-D-\tau)+ \\
&\quad R_{on}T_{on} \\
R_M &= C\left(\frac{T_{off}}{T_{on}+D+\tau}\right)+R_L\left(\frac{T_{on}-T_{off}+D+\tau}{T_{on}+D+\tau}\right)
\end{aligned}$$

where $R_M = R_{M1} = R_{M2}$ is the instantaneous rate at which the Rapid sender is sending upon both the start and the end of the cross traffic burst.

C. *Case S3:* $T_{on}+D \geq \eta$, $T_{off}-D < \tau$, $D \geq T_{off}-T_{on}+\eta$

When $T_{on}+D \geq \eta$ but $T_{off}-D < \tau$, the Rapid sender will have enough time to converge on a lower sending rate when it receives a lower avail-bw estimate, but when it learns of an increase in avail-bw it will not have sufficient time to converge on a higher sending rate before the avail-bw decreases again. The highest rate it will achieve is $R_H < C$.

The highest sending rate R_H and the time-averaged steady-state Rapid throughput $\mathcal{R}$ are described by

$$\begin{aligned}
R_H &= (C-R_{on})+R_{on}\left(\frac{T_{off}-D}{\tau}\right) \\
&= C+R_{avg}\left(\frac{T_{on}+T_{off}}{T_{on}}\right)\left(\frac{T_{off}-D-\tau}{\tau}\right) \\
\mathcal{R} &= \left(\frac{R_H+C-R_{on}}{2}\right)\left(\frac{T_{off}-D+\eta}{T_{on}+T_{off}}\right)+ \\
&\quad (C-R_{on})\frac{T_{on}+D-\eta}{T_{on}+T_{off}} \\
&= C - R_{avg}\frac{T_{on}+T_{off}}{T_{on}} - \\
&\quad R_{avg}\frac{(T_{off}-D)(T_{off}-D+\eta)}{2T_{on}\cdot\tau}
\end{aligned}$$

When $T_{on} < T_{off} - D + \eta$, the worst-case queue maximum buildup occurs when $RTT'_{rem} \in [T_{off}-D, T_{on}-\eta]$. In this case, the entire Rapid traffic burst will occur within the cross traffic burst period. The maximum queue buildup is described by

$$Q_{max} = \frac{R_H-C+R_{on}}{2}(T_{off}-D+\eta)$$

D. Case S4: $T_{on}+D \geq \eta$, $T_{off}-D < \tau$, $D < T_{off}-T_{on}+\eta$

As in Case S3, the Rapid sender will have enough time to converge on a lower sending rate when it receives a lower avail-bw estimate, but when it learns of an increase in the avail-bw estimate it will not have sufficient time to converge on a higher sending rate before the avail-bw decreases again. The highest sending rate R_H and the time-averaged steady-state Rapid throughput $\mathcal{R}$ are calculated in the same manner as in Case S3:

$$\begin{aligned} R_H &= C + R_{avg}\left(\frac{T_{on}+T_{off}}{T_{on}}\right)\left(\frac{T_{off}-D-\tau}{\tau}\right) \\ \mathcal{R} &= C - R_{avg}\frac{T_{on}+T_{off}}{T_{on}} - \\ &\quad R_{avg}\frac{(T_{off}-D)(T_{off}-D+\eta)}{2T_{on}\cdot\tau} \end{aligned}$$

When $T_{on} < T_{off} - D + \eta$, Rapid's traffic burst will fit partially within the cross traffic burst period. As we establish in Appendix X, the worst-case maximum queue buildup will occur when RTT is such that Rapid's instantaneous send rate when the cross traffic burst begins, R_{M1}, is equal to the instantaneous send rate when the cross traffic burst ends, R_{M2}. This occurs when

$$RTT'_{rem} = T_{on}\frac{T_{off}-D}{T_{off}-D+\eta}$$

In this case, the maximum queue buildup is described by

$$\begin{aligned} Q_{max} &= T_{on}\frac{R_H+R_M}{2} + (C-R_{on})T_{on} \\ R_M &= R_H\frac{T_{off}-T_{on}-D+\eta}{T_{off}-D+\eta} - \\ &\quad (C-R_{on})\left(\frac{T_{on}}{T_{off}-D+\eta}\right) \end{aligned}$$

where $R_M = R_{M1} = R_{M2}$ is the instantaneous rate at which the Rapid sender is sending upon both the start and the end of the cross traffic burst.

E. Case S5: $T_{on}+D < \eta$, $T_{off}-D < \tau$

The last small-scale burst scenario occurs when $T_{on}+D < \eta$ and $T_{off}-D < \tau$, and Rapid is neither able to adapt to an increase nor a decrease in avail-bw before the latter changes again. When avail-bw increases, the Rapid sender is able to reach a maximum sending rate of only $R_H < C$, and when avail-bw decreases, Rapid is able to reduce its sending rate to a minimum value of only $R_L > C - R_{on}$ before the avail-bw changes again.

In steady-state, the high and low sending rates are described by:

$$\begin{aligned} R_L &= R_H - (R_H-(C-R_{on}))\left(\frac{T_{on}+D}{\eta}\right) \\ R_H &= R_L - (C-R_L)\left(\frac{T_{off}-D}{\tau}\right) \end{aligned}$$

where the steady-state Rapid throughput can be averaged as $\mathcal{R} = \frac{R_L+R_H}{2}$.

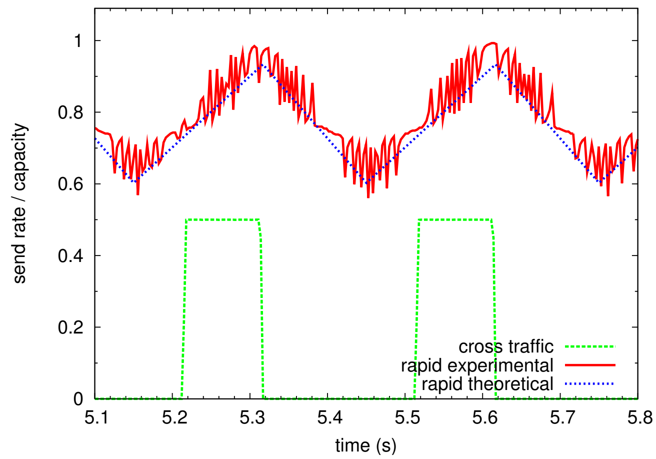


Fig. 8. Small-scale Bursts: Experimental Validation

As we establish in Appendix X, the worst-case maximum queue buildup occurs when RTT is such that Rapid's instantaneous send rate when the cross traffic burst begins, R_{M1}, is equal to the instantaneous send rate when the cross traffic burst ends, R_{M2}. This occurs when

$$RTT'_{rem} = \left(\frac{T_{on}}{T_{on}+T_{off}}\right)(T_{off}-D)$$

In this case, the maximum queue buildup is described by

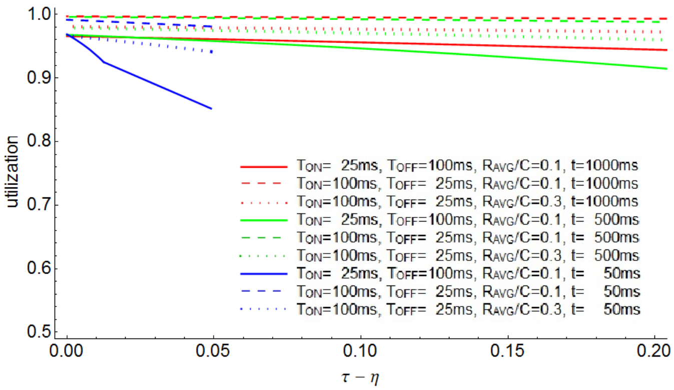
$$\begin{aligned} Q_{max} &= T_{on}\frac{R_M+R_H}{2} - (C-R_{on})T_{on} \\ &= T_{on}\left(\frac{R_H T_{off}-R_L T_{on}}{2(T_{on}+T_{off})}\frac{R_H}{2} - C + R_{on}\right) \\ R_M &= R_H\left(\frac{T_{off}}{T_{on}+T_{off}}\right) - R_L\left(\frac{T_{on}}{T_{on}+T_{off}}\right) \end{aligned}$$

where $R_M = R_{M1} = R_{M2}$ is the instantaneous rate at which the Rapid sender is sending upon both the start and end of the bike path.

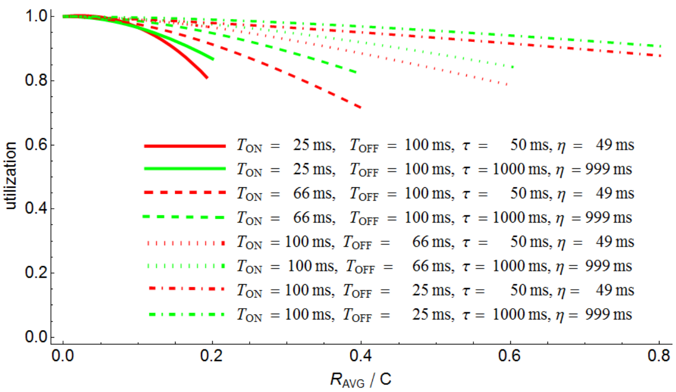
Validation We validate the above analysis against an ns-2 experiment in which: $T_{on} = 100ms$, $T_{off} = 200ms$, $R_{avg} = 111Mbps$, $\tau = 200ms$, $\eta = 175ms$, and $RTT = 100ms$. Fig 8 plots a sample of the theoretical and experimentally observed time series for the Rapid sending rate. We find that the experiment validates the analysis quite well. Our experiments with other parameter settings and with exponential on-off cross traffic also match the analysis equally well.

VII. Discussion

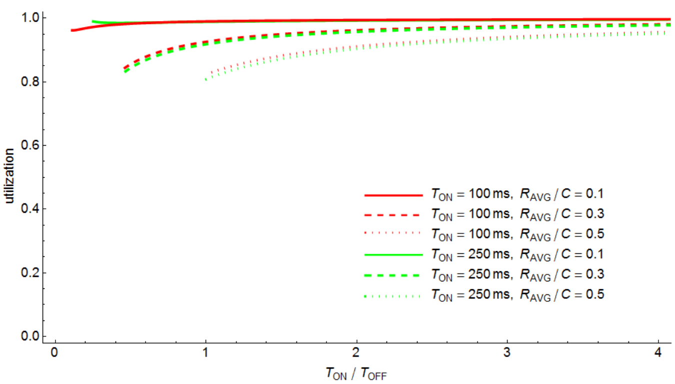
We next use the models we have developed to quantitatively study the impact that cross traffic burstiness can have on Rapid and the extent to which it can be alleviated. Sections V-VI establish that the magnitude of τ and η *relative* to T_{on} and T_{off} determines whether the cross-traffic interacts with Rapid as "large-scale" or "small-scale" bursts. Further, Equation (4) shows that the inability of Rapid to fully utilize the bottleneck link due to large-scale cross traffic bursts: (i) does not depend individually on τ or η, but rather on $\tau - \eta$, the difference between the two smoothing parameters; (ii) is



(a) Role of $\tau - \eta$



(b) Role of $\frac{R_{avg}}{C}$



(c) Role of $\frac{T_{on}}{T_{off}}$

Fig. 9. Impact of Cross Traffic Bursts: Influence of Factors

proportional to the relative load of the cross traffic, $\frac{R_{avg}}{C}$; (iii) is inversely proportional to T_{on}; and (iv) is proportional to the ratio: $\frac{T_{off}}{T_{on}}$. While the small-scale bursts analysis yields closed-form expressions that are not as simple, we use mathematical software to quantify the influence of the above factors and find that these are still the dominant ones.

Role of $\tau - \eta$: Fig 9(a) plots $\mathcal{U}$ as a function of $\tau - \eta$. C is set to 1 Gbps and the different curves correspond to different values of T_{on}, T_{off}, and R_{avg}.[5] $\tau \in \{50ms, 500ms, 1s\}$, while η is varied over $[1ms, \tau]$. We find that in all cases:

- $\mathcal{U}$ decreases with increase in $\tau - \eta$.
- The loss in utilization due to a large $\tau - \eta$ is influenced somewhat, but not significantly, by the values of R_{avg}, T_{on}, and T_{off}.
- For a given value of $\tau - \eta$, a larger value of τ seems to alleviate the loss in utilization due to cross-traffic burstiness. However, the $\tau = 500ms$ and $\tau = 1s$ curves are fairly close to each other—this suggests that the influence of τ seems to be limited to whether the bursts are "large-scale" or "small-scale" in comparison to τ. We further explore this issue below.

Observe that τ and η are the only model parameters that are under the control of Rapid. The above figures indicate that independent of the cross traffic load and burstiness, these two parameters should be set such that their difference is small. In the rest of this discussion, we set: $\tau - \eta = 1ms$.[6]

Role of $\frac{R_{avg}}{C}, \tau$: Fig 9(b) plots $\mathcal{U}$ as a function of $\frac{R_{avg}}{C}$.[7] The different curves correspond to 4 different values of (T_{on}, T_{off}) and 2 values for τ ($\tau - \eta = 1ms$).[8] We find that:

- $\mathcal{U}$ decreases with increase in $\frac{R_{avg}}{C}$. This trend is more pronounced for smaller values of $\frac{T_{on}}{T_{off}}$.
- For a given combination of T_{on} and T_{off}, Rapid performs better when bursts are "small-scale" (τ, η are sufficiently larger than T_{on}, T_{off}). Further, as long as the bursts are small-scale, the value of τ does not influence Rapid performance much — in fact, the curves for $\tau = 250ms, 500ms$ (not plotted to avoid clutter) are fairly close to the $\tau = 1s$ for all of the cases plotted in this figure.
  This suggests that even though Rapid performance is relatively independent of the individual values of τ and η (and depends mostly on $\tau - \eta$), it does get strongly influenced by whether or not these parameters are larger than the traffic bursts and idle durations. Using large values for τ and η can help ensure that bursts are relatively "small" in scale and Rapid performs well.

Role of $T_{on}, \frac{T_{on}}{T_{off}}$: Fig 9(c) plots $\mathcal{U}$ as a function of $\frac{T_{on}}{T_{off}}$ — here, $T_{on} = 25ms, 100ms$ and T_{off} is varied to get different values for this ratio. $C = 1Gbps$ and the different curves correspond to different values of $\frac{R_{avg}}{C}$. We find that:

- The link utilization achieved increases with the ratio $\frac{T_{on}}{T_{off}}$. In most cases, the influence flattens out beyond a ratio of

[5]Note that R_{on} can not be larger than C—this limits the maximum feasible choice of R_{avg} for a given $\frac{T_{off}}{T_{on}}$. Similarly, we have: $\tau - \eta \leq \tau$ —this limits the x-range in Fig 9(a).

[6]It can be shown that to ensure the stability of the paradigm, we need to select: $\eta < \tau$.

[7]$\mathcal{U}$ is independent of the individual values of R_{avg} and C, but is influenced by the ratio $\frac{R_{avg}}{C}$, for both large-scale and small-scale bursts. This is a significant property — it implies that the impact of burstiness dos not change as we consider upcoming networks with higher transmission capacities of 40-100 Gbps.

[8]Note again that $R_{on} < C$ limits the maximum feasible choice of R_{avg} for a given $\frac{T_{off}}{T_{on}}$.

$\frac{T_{on}}{T_{off}} = 4$.

- For a given $\frac{T_{on}}{T_{off}}$, the impact of the individual choice of T_{on} is minimal. This is a very interesting observation—the timescale of cross traffic burstiness does *not* significantly influence Rapid's performance (as long as the smoothing timescales τ and η are larger), but the *relative* durations of burst and idle periods do!

Our use of the model in this section reveals that once the Rapid protocol parameters are configured ideally (smaller $\tau - \eta$, large τ), the impact of cross traffic is dependent mostly on the relative load of cross traffic, $\frac{R_{avg}}{C}$, and the relative burstiness-to-idleness ratio: $\frac{T_{on}}{T_{off}}$. With ideal parameter configuration, the ability of Rapid to full utilize the bottleneck link can vary from 80% to 95%, depending on the combination of these two factors.

VIII. Concluding Remarks

The packet-scale paradigm is a promising framework for ultra-high speed networks. However, it can deliver on its promise only if network "noise" does not adversely impact its bandwidth estimation process. This paper conducts the first analysis of the interaction of the packet-scale paradigm with one source of noise: traffic burstiness. The analysis relies on a first-principles approach to gain several simple but fundamental insights on the influence of network, traffic, and protocol parameters on the efficiency of the paradigm.

While this analysis is an important first step, it is only a small one in understanding the interaction of the packet-scale paradigm with sources that can introduce "noise" in the inter-packet gaps. We are currently extending this analysis in the following directions: (i) An important source of "noise" is non-bottleneck buffering, which occurs in almost all store-and-forward systems that packets pass through. This buffering occurs at very fine timescales—the packet-scale paradigm is uniquely likely to get impacted by this due to the fine timescales at which it operates. (ii) On a similar note, bottleneck burstiness that occurs at timescale smaller than probing-streams are likely to impact the paradigm in ways that this paper does not study. We are currently relying on packet-level queuing theory to study these issues. (iii) Obviously, the periodic on-off cross traffic model is unrealistic. However, we believe it has helped us understand the relative role of several factors and it is likely this will hold as we study more complex and realistic models for traffic burstiness.

References

[1] V. Konda and J. Kaur, "RAPID: Shrinking the Congestion-control Timescale," in *Proceedings of the IEEE INFOCOM*, April 2009.

[2] "Packet-scale Congestion Control," http://rapid.web.unc.edu.

[3] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling TCP Throughput: A Simple Model and its Empirical Validation," in *Proceedings of ACM SIGCOMM*, September 1998.

[4] B. Sikdar, S. Kalyanaraman, and K. Vastola, "Analytic Models for the Latency and Steady-State Throughput of TCP Tahoe, Reno and SACK," in *Proceedings of IEEE Globecom*, November 2001.

[5] M. Weigle, L. Chen, J. Kaur, and V. Kulkarni, "Generalized Stochastic Performance Models for Loss-based Congestion Control," *Computer Communications Journal*, vol. 33, pp. 513–525, March 2010.

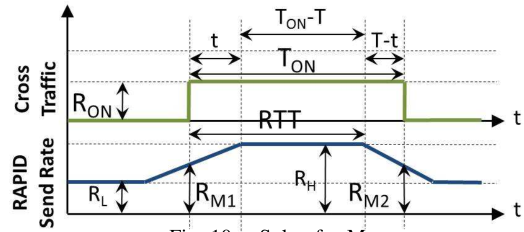


Fig. 10. Solve for M

[6] C. Dovrolis and M. Jain, "End-to-End Available Bandwidth: Measurement Methodology, Dynamics, and Relation with TCP Throughput," *IEEE/ACM Transactions in Networking*, August 2003.

IX. Influence of a Residual Queue

We formally derive Observation 1 for the case when all packets of a p-stream belong to the same *busy period* of the router link.

Let a_1 and a_2 be the arrival times at the bottleneck link of two successive packets, p_1 and p_2, of a p-stream. Let d_1 and d_2 be their departure times. Since p_1 and p_2 belong to the same busy period, we have:

$$d_2 - d_1 = \frac{P}{C} + \frac{B_{CT}(a_1, a_2)}{C} \quad (5)$$

where P is the packet size of p_2, and $B_{CT}(a_1, a_2)$ is the amount of cross-taffic that arrives in the interval (a_1, a_2) (and gets inserted in between the two p-stream packets).

If p_2 is sent at a probing rate of r_2, then we have: $a_2 - a_1 = \frac{P}{r_2}$. Let R_{CT} denote the average cross traffic arrival rate in the interval (a_1, a_2). Then we have: $B_{CT}(a_1, a_2) = R_{CT}(a_2 - a_1)$. Thus, we have:

$$d_2 - d_1 = \frac{P}{C}\left(1 + \frac{R_{CT}}{r_2}\right) \quad (6)$$

Rearranging and solving, we get: $d_2 - d_1 > a_2 - a_1$ iff $r_2 > C - R_{CT}$. Hence, the relation between the departure gaps and the arrival gaps (and consequently, the avail-bw estimates yielded by the p-stream) depend only on the cross traffic that arrives along with the p-stream packets (and is independent of the amount of residual queue).

X. Worst-case Alignment for Cases L3, S2, S4, and S5

In cases where the Rapid traffic burst can fit only partially inside the cross traffic burst period, Rapid is increasing its sending rate when the burst period begins and is decreasing its sending rate when the period ends. In such cases, the maximum queue buildup will depend on R_{M1} and R_{M2}, the instantaneous send rates of Rapid at the beginning and end of the burst period, respectively. The maximum queue size with respect to t, the length of time from the beginning of the burst period until Rapid completes ramping up to its highest sending rate R_H, is described by

$$Q = \frac{R_H}{2}(2T_{on} - T) + \left(\frac{R_{M1}}{2}\right)t + \left(\frac{R_{M2}}{2}\right)(T - t) - (C - R_{on})\,T_{on}$$

where T is the amount of time during the burst period during which Rapid is not sending at its highest rate R_H. By differentiating with respect to t, we determine that Q_{max} is maximized when $R_{M1} = R_{M2}$. Thus, worst-case queue buildup in the case where the Rapid burst fits partially within the cross traffic burst period is described by

$$Q = \frac{R_H}{2}(2T_{on} - T) + \left(\frac{R_M}{2}\right)T - (C - R_{on})T_{on}$$

where $R_M = R_{M1} = R_{M2}$; the relationship between R_{M1} and R_{M2} can be used to solve for R_M when analyzing Q_{max}.