Real-Time Status Updating: Multiple Sources

Roy D. Yates
WINLAB, ECE Dept.
Rutgers University
email: *ryates@winlab.rutgers.edu*

Sanjit Kaul
IIIT-Delhi
email: *skkaul@iiitd.ac.in*

***Abstract*—We examine multiple independent sources providing status updates to a monitor through a first-come-first-served M/M/1 queue. We formulate a status-age timeliness metric and find the region of feasible average status ages for a pair of updating sources. In the presence of interfering traffic with a given offered load, we show the existence of an optimal rate at which a source should generate its updates.**

I. Introduction

Increasingly ubiquitous connectivity to communication networks and availability of portable devices have engendered a host of applications in which sources – people and environmental sensors – send updates of their status to interested recipients. These include news and weather reports and updates by individuals on Twitter about what is keeping them busy, updates by environmental sensors [1], and vehicular status (position, velocity, acceleration) updates that can assist drivers of nearby vehicles in an intelligent transportation system [2]. These applications need status updates at one or more monitors to be *as timely as possible*; however, this is typically constrained by limited network resources.

For example, consider status updates generated by sensors in a vehicle. The update packets are queued while they wait to be serviced by the car radio. The packet currently being serviced by the radio waits for medium access and transmission before it is received by other cars. Note that each sensor in the car may be a source or that the car may aggregate a collection of sensor measurements into a status update message that is transmitted as a single packet. The packet service time will depend on the wireless channel and may or may not incorporate retransmissions due to channel errors and backoff due to the activity of other wireless transmitters. While system models that incorporate these effects can be arbitrarily complex, we observe even in the simple setting of an $M/M/1$ queue that optimal updating policies are not well understood.

Maintaining the timeliness of data and state information in a network is a problem that has appeared in many forms, including, for example, data freshness in warehouses [3] and web caches [4], periodic updating of real time databases [5], and route caches in ad hoc networks [6]. However, no consistent analytic methodology has emerged. In this paper, we explore a new *status-age* timeliness metric as a basis for the evaluation and design of status update systems. When a monitor's most recently received update at time t is time-stamped $u(t)$, the *status update age*, which we will refer to as simply the *age*, is the random process $\Delta(t) = t - u(t)$. The monitor's requirement of timely updating corresponds to small average $\Delta(t)$. While status age is an application-independent metric, it can be useful in specific applications by designing the communication network to meet statistical requirements, such as expected value and variance, of the age process. For example, if a status updating system is forwarding sample values of a Wiener process $X(t)$ with variance αt [7], then the monitor's MMSE estimate of $X(t)$ given the status age $\Delta(t)$ is $\hat{X}(t) = X(t - \Delta(t))$. The variance of this estimate is $\alpha\Delta(t)$.

We will see that the goal of timely updating is neither the same as maximizing the utilization of the communication system, nor of ensuring that generated status updates are received with minimum delay. Utilization may be maximized by making the sensor send updates as fast as possible. However, this may lead to the monitor receiving delayed statuses because the status messages become backlogged in the communication system. In this case, delay suffered by the stream of status updates can be reduced by reducing the rate of updates. Alternatively, reducing the update rate can also lead to the monitor having unnecessarily outdated status information because of a lack of updates.

In this work, we start in Section II with a formulation of the time-averaged status-age $\Delta(t)$ that applies to a broad class of systems. This is used in Section III to analyze an FCFS $M/M/1$ queuing system that delivers the status updates of multiple independent sources to one or more monitors. For the FCFS multiuser system, we derive the region of feasible status-ages. In addition, in the presence of "interfering" traffic with a given offered load, we show the existence of an optimal rate at which a source should generate its updates to keep its status as timely as possible at the monitor. The key novelty found in this analysis is that it is in the self-interest of a source to limit its load on the service system. We conclude the paper with possible extensions and a summary of our contributions in Section IV.

II. FCFS Status Update Age

Figure 1 shows a sample variation of age $\Delta_1(t)$, for source 1 as a function of time t, at the monitor. Without loss of generality, assume that we begin observing at $t = 0$ when the queue is empty and the age is $\Delta_1(0) = \Delta_0$. The first status update of source 1 is generated at t_1, followed by updates at $t_2, t_3, \dots, t_n$. The status age of source 1 at the monitor increases linearly in time in the absence of any updates and

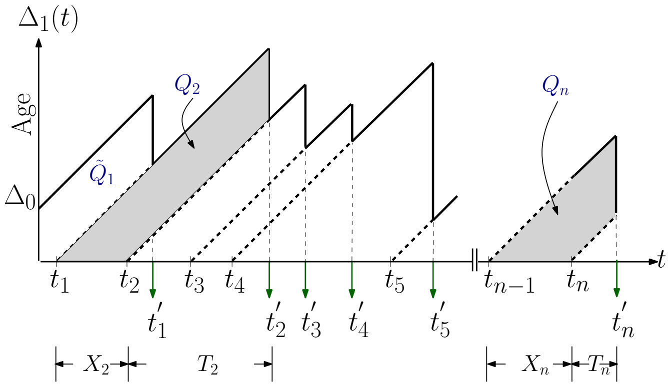

Fig. 1: Example change in status update age at a monitor for a system with a FCFS queue.

is reset to a smaller value when an update is received. Update i of source 1, generated at time t_i, finishes service and is received by the monitor at time t_i'. At t_i', the age $\Delta_1(t_i')$ at the monitor is reset to the age $T_i = t_i' - t_i$ of the received status update. The age T_i is also the system time of update packet i. Thus the age function $\Delta_1(t)$ exhibits the sawtooth pattern shown in Figure 1.

The time average age of the status updates is the area under the age graph in Figure 1 normalized by the time interval of observation. Over an interval $(0, \mathcal{T})$, the average age is

$$\langle \Delta_1 \rangle_{\mathcal{T}} = \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \Delta_1(t) dt. \tag{1}$$

For simplicity of exposition, the length of the observation interval is chosen to be $\mathcal{T} = t_n'$, as depicted in Figure 1. We decompose the area defined by the integral in (1) into a sum of disjoint geometric parts. Starting from $t = 0$, the area can be seen as the concatenation of the polygon area $\tilde{Q}_1$, the trapezoids Q_i for $i \geq 2$ (Q_2 and Q_n are highlighted in the figure), and the triangular area of width T_n over the time interval (t_n, t_n'). With $N_1(\mathcal{T}) = \max\{n | t_n \leq \mathcal{T}\}$ denoting the number of source 1 arrivals by time $\mathcal{T}$, this decomposition yields

$$\langle \Delta_1 \rangle_{\mathcal{T}} = \frac{\tilde{Q}_1 + T_n^2/2 + \sum_{i=2}^{N_1(\mathcal{T})} Q_i}{\mathcal{T}}. \tag{2}$$

From Figure 1, we see that the area Q_i can be calculated as the difference between the area of the isosceles triangle whose base connects the points t_{i-1} and t_i' and the area of the isosceles triangle with base connecting the points t_i and t_i'. Defining

$$X_i = t_i - t_{i-1} \tag{3}$$

to be the elapsed time between the generation of updates $i-1$ and i, it follows that

$$Q_i = \frac{1}{2}(T_i + X_i)^2 - \frac{1}{2}T_i^2 = X_i T_i + X_i^2/2. \tag{4}$$

When the generation of updates can be represented as the arrivals of a stochastic process, X_i is the interarrival time of update i. Substituting (4) in (2), some rearrangement yields the time-average age

$$\langle \Delta_1 \rangle_{\mathcal{T}} = \frac{\tilde{Q}}{\mathcal{T}} + \frac{(N_1(\mathcal{T}) - 1)}{\mathcal{T}} \frac{\sum_{i=2}^{N_1(\mathcal{T})} \left[X_i T_i + X_i^2/2\right]}{N_1(\mathcal{T}) - 1} \tag{5}$$

where, $\tilde{Q} = \tilde{Q}_1 + T_n^2/2$. We observe that the age contribution $\tilde{Q}$ represents a boundary effect that is finite with probability 1, so the first term in (5) will vanish as $\mathcal{T}$ grows. Let

$$\lambda_1 = \lim_{\mathcal{T} \to \infty} \frac{N_1(\mathcal{T})}{\mathcal{T}} \tag{6}$$

be the ergodic rate at which status update packets are generated. Furthermore, as $N_1(\mathcal{T}) \to \infty$, the remaining summation term in (5) is a sample average that will converge to its corresponding stochastic average. The average status update age can be obtained as

$$\Delta_1 = \lim_{\mathcal{T} \to \infty} \langle \Delta_1 \rangle_{\mathcal{T}} = \lambda_1 \left(\mathrm{E}[XT] + \mathrm{E}\left[X^2\right]/2 \right), \tag{7}$$

where $\mathrm{E}[\cdot]$ is the expectation operator, and X and T are the random variables that correspond to the interarrival time and system time of a source 1 update packet, respectively.

We note that the average update age in (7) holds under weak assumptions on the ergodicity of the service system. Furthermore, (7) is a general result for a broad class of service systems in which the update packets are processed FCFS. In particular, the analysis leading to (7) made no assumptions regarding other traffic that shares the queue with the update packets of source 1. That is, (7) holds for a queue in which the status update stream shares the service facility with other packet streams. However, evaluation of the age Δ_1 can be challenging. In particular, X is the random variable that describes the time between generation of an update packet and the one that preceded it while T is the system time of that same packet. The variables X and T are dependent. In a queue serving a single source, a large interarrival time X allows the queue to empty, yielding a small waiting time and typically a small system time T. In this case, X and T are negatively correlated. This correlation may be reduced when the queue serves packets from other sources; nevertheless, the evaluation of $\mathrm{E}[TX]$ tends to be nontrivial.

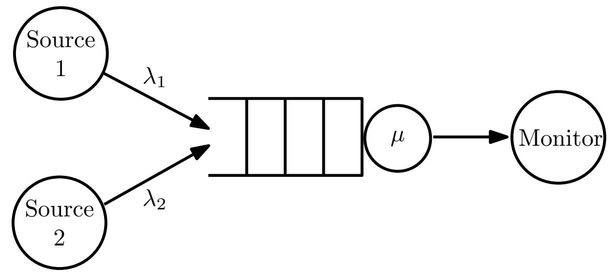

Fig. 2: Independent sources send through a queue to a monitor.

III. M/M/1 FCFS – Two Sources

In [8], we analyzed $\mathrm{E}[TX]$ for a variety of FCFS queues serving the status updates of only a single source. In that work, it was shown that the average status age for an M/M/1 queue with arrival rate λ_1, service rate μ and offered load $\rho_1 = \lambda_1/\mu$

is given by

$$\Delta_1 = \frac{1}{\mu}\left(1 + \frac{1}{\rho_1} + \frac{\rho_1^2}{1-\rho_1}\right). \tag{8}$$

The average age Δ_1 is minimized at $\rho_1^* \approx 0.53$.

Here we generalize this prior work to the FCFS $M/M/1$ system depicted in Figure 2 with independent sources $i = 1, 2$. Note that it is sufficient to consider the two source system as source 2 is equivalent to an aggregate of multiple Poisson streams. The service rate is μ for packets from either source. Source i has Poisson arrival rate λ_i, generating offered load $\rho_i = \lambda_i/\mu$. Update packets are generated as a rate $\lambda = \lambda_1 + \lambda_2$ Poisson process and the overall load is

$$\rho = \lambda/\mu = \rho_1 + \rho_2. \tag{9}$$

To analyze the average update age for each source, we refer to packet k, i for the ith packet from source k, $k = 1, 2$. We use X_{ki} to denote the interarrival time of packet k, i, *relative to the prior packet* $k, i-1$ *also from source* k.

In this notation, the average status age (7) for source k is

$$\Delta_k = \lambda_k \left(\mathrm{E}\left[X_{ki}T_{ki}\right] + \mathrm{E}\left[X_{ki}^2\right]/2\right). \tag{10}$$

In steady state, the system time T_{ki} of each packet k, i is stochastically identical to T, with the exponential PDF

$$f_T(t) = \mu(1-\rho)e^{-\mu(1-\rho)t}, \qquad (t \geq 0.) \tag{11}$$

For the average age Δ_1 in (10), we now derive an expression for $\mathrm{E}\left[T_{1i}X_{1i}\right]$ that accounts for the traffic offered by the other source. The system time of update $1, i$ is

$$T_{1i} = W_{1i} + S_{1i}, \tag{12}$$

where W_{1i} and S_{1i} are the respective waiting time and service time of source 1 packet i. Since S_{1i} is independent of X_{1i}, it follows from (12) that

$$\mathrm{E}\left[X_{1i}T_{1i}\right] = \mathrm{E}\left[X_{1i}W_{1i}\right] + \mathrm{E}\left[X_{1i}\right]\mathrm{E}\left[S_{1i}\right]. \tag{13}$$

We note that $\mathrm{E}\left[S_{1i}\right] = 1/\mu$ and that the rate λ_1 Poisson arrival process implies $\mathrm{E}\left[X_{1i}\right] = 1/\lambda_1$ and $\mathrm{E}\left[X_{1i}^2\right] = 2/\lambda_1^2$. It follows from (10) and (13) that

$$\Delta_1 = \lambda_1 \mathrm{E}\left[X_{1i}W_{1i}\right] + \frac{1}{\mu} + \frac{1}{\lambda_1}. \tag{14}$$

The key to the evaluation of $\mathrm{E}\left[X_{1i}W_{1i}\right]$ is to characterize the waiting time W_{1i} via the partition

$$B_i = \left\{X_{1i} < T_{1,i-1}\right\}, \quad L_i = \left\{T_{1,i-1} < X_{1i}\right\}. \tag{15}$$

That is, B_i denotes the event that the ith interarrival time for source 1 is brief, specifically, less than the system time of the preceding packet from source 1. By the same standard, L_i is the complementary event that X_{1i} is long. When the X_{1i} is brief, W_{1i} is the sum of the residual system time of packet $1, i-1$ plus the service times of those source 2 packets that arrived during the interarrival time. When X_{1i} is long, we verify that the number of source 2 packets in the system following the departure of packet $1, i-1$ is described by the geometric stationary distribution for source 2 packets. As shown in the Appendix, these facts yield the next lemma.

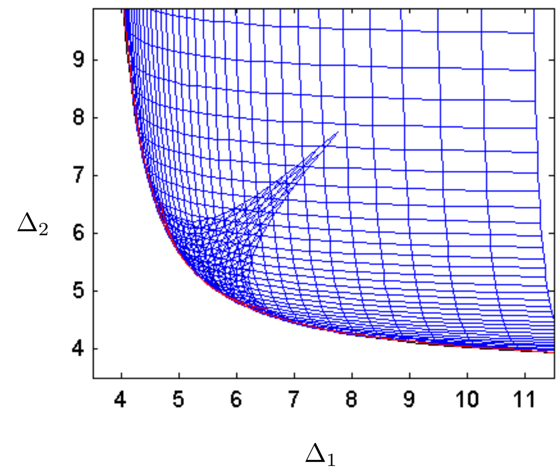

Fig. 3: The age region for two sources sharing a rate $\mu = 1$ M/M/1 queue. The the minimum sum age point in the lower left is achieved by $\rho_1 = \rho_2 = 0.306$.

Lemma 1:

$$\mathrm{E}\left[W_{1i}X_{1i}\right] = \frac{1}{\mu^2}\left[\frac{\rho_1(1-\rho\rho_2)}{(1-\rho)(1-\rho_2)^3} + \frac{\rho_2}{\rho_1(1-\rho_2)}\right].$$

Applying Lemma 1 to (14) yields

$$\Delta_1 = \frac{1}{\mu}\left[\frac{\rho_1^2(1-\rho\rho_2)}{(1-\rho)(1-\rho_2)^3} + \frac{1}{1-\rho_2} + \frac{1}{\rho_1}\right]. \tag{16}$$

Given a fixed alternate traffic load ρ_2, it is instructive to define the normalized load

$$\hat{\rho}_1 = \frac{\rho_1}{1-\rho_2}. \tag{17}$$

With some additional algebra, this permits us to write

$$\Delta_1 = \frac{\beta_1(\hat{\rho}_1, \rho_2)}{(1-\rho_2)\mu}, \tag{18a}$$

$$\beta_1(\hat{\rho}_1, \rho_2) = 1 + \frac{1}{\hat{\rho}_1} + \frac{\hat{\rho}_1^2}{1-\hat{\rho}_1} + \rho_2\hat{\rho}_1^2. \tag{18b}$$

From the perspective of source 1, $(1-\rho_2)\mu$ is the average service rate provided to its packets. Comparing (8) with load ρ_1 against (18) with normalized load $\hat{\rho}_1$, we see that the source 2 traffic effectively reduces the service rate of source 1 updates from μ to $\mu(1-\rho_2)$ as well as adding the penalty term $\rho_2\hat{\rho}_1^2$.

By reversing labels 1 and 2, (18) also enables us to characterize the average age for source 2 updates. With normalized service rate $\mu = 1$, the region of achievable average update ages is shown in Figure 3. The lower left "corner" point where the sum $\Delta_1 + \Delta_2$ is minimized is obtained at $\rho_1 = \rho_2 = 0.306$, yielding $\Delta_1 = \Delta_2 = 5.30$. By comparison, serving a single source with optimal load $\rho_1 = 0.531$, Equation (8) will yield $\Delta_1 = 3.48$. This implies that if we were to partition the resources to create two systems, each with service rate $\mu = 1/2$, to serve the sources independently, then each source would obtain age $\Delta_i = 6.96$. Thus we observe a trunking efficiency in having two status-updating sources share the service facility.

A second situation of interest occurs when source 1 is a status updater in the presence of "interfering" traffic from source 2. In this case, source 1 can choose its updating rate in order to minimize Δ_1. Given the interfering load ρ_2, we see from (18) that an optimal policy selects $\hat{\rho}_1$ to minimize $\beta_1(\hat{\rho}_1, \rho_2)$. While the exact minimum is the root of a fifth order polynomial, one can show that a second order approximation yields the approximately optimal linear solution

$$\hat{\rho}_1^*(\rho_2) = 1/2 + (1 - \rho_2)/32. \quad (19)$$

We further note that this linear approximation is exactly optimal at $\rho_2 = 1$ and very close to the numerically calculated optimum $\hat{\rho}_1 = 0.531$ at $\rho_2 = 0$, which corresponds to the single source case. In fact, performance differences between the linear approximation and the exact optimum are insignificant for all ρ_2. In particular, Figure 4 shows Δ_1 as function of ρ_2 for the optimal updating load $\hat{\rho}_1^*(\rho_2)$ (marked "opt" in the legend) as well as for heuristic choices of the form $\hat{\rho}_1 = \omega$. Note that $\hat{\rho}_1 = \omega$ implies $\rho_1 = \omega(1 - \rho_2)$. That is, source 1 uses a fixed fraction of the residual system capacity. In fact, we observe that the age with $\omega = 0.5$ is indistinguishable from that with the optimal updating rate and, further, that the age Δ_1 is relatively insensitive to the choice of ω in the vicinity of $\omega = 1/2$.

Now suppose that source 2 were also a status updater, and that both the sources selected an update rate that minimized their age given the other source's update rate. Note that the minimum age for source 1 can be obtained by selecting $\hat{\rho}_1$ that minimizes Δ_1 in (18), for a known ρ_2. The same can be achieved for source 2 by reversing labels 1 and 2 in (18). Using such an algorithm we confirm experimentally, for a range of initial server utilizations, that the ages of the sources converge to the pair $(5.4390, 5.4390)$. This age pair is a fixed point of unilateral optimization and is in the interior of the feasible age region in Figure 3 and leads to suboptimal sharing. The update rate pair at this fixed point is a Nash equilibrium and achieves $(\Delta_1, \Delta_2) = (5.4390, 5.4390)$ at $(\rho_1, \rho_2) = (0.342, 0.342)$.

IV. Conclusion

We have looked at the problem of multiple sources generating timely status updates at interested recipients. We have employed a simple approach in which the communication network is modeled by an FCFS M/M/1 queue. We derived the region of feasible status ages. We also resolved the optimal updating rate in the presence of interfering traffic. The preliminary insights lead us to believe that status updating is a potential way to address more complex problems in real-time process estimation through networks.

Appendix

The proof of Lemma 1 relies on the following basic properties of Poisson processes and exponential random variables.

Lemma 2: Let X_1 and X_2 be independent exponential random variables with $\mathrm{E}[X_i] = 1/\alpha_i$. Let $Y = X_2 - X_1$.

(a) $\mathrm{P}[X_1 < X_2] = \alpha_1/(\alpha_1 + \alpha_2)$.

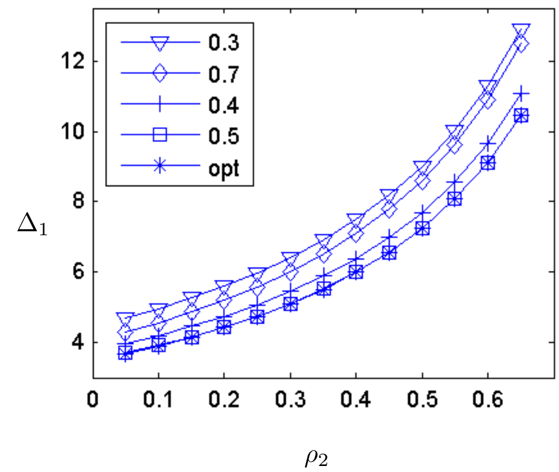

Fig. 4: Source 1 age Δ_1 as a function of interfering traffic with load ρ_2 for various normalized loads $\hat{\rho}_1$; "opt" marks the age-minimizing source 1 normalized load $\hat{\rho}_1^*$.

(b) Given $X_1 < X_2$, X_1 and Y are conditionally independent and have conditional exponential probability density functions (PDFs)

$$f_{X_1|X_1<X_2}(x) = (\alpha_1 + \alpha_2)e^{-(\alpha_1+\alpha_2)x}, \quad (x \geq 0,)$$
$$f_{Y|X_1<X_2}(y) = \alpha_2 e^{-\alpha_2 y}, \quad (y \geq 0.)$$

Lemma 3: Given a rate λ Poisson process $N(t)$ and an independent exponential (α) random variable X, the number of arrivals $N(X)$ in the interval $[0, X]$ has the geometric PDF

$$P_{N(X)}(n) = (1 - \gamma)\gamma^n, \quad (n \geq 0),$$

with $\gamma = \lambda/(\alpha + \lambda)$.

Proof: Lemma 1 The partition $\{B_i, L_i\}$ permits us to write

$$\mathrm{E}[X_{1i}W_{1i}] = \mathrm{E}[X_{1i}W_{1i}|L_i]\,\mathrm{P}[L_i] + \mathrm{E}[X_{1i}W_{1i}|B_i]\,\mathrm{P}[B_i]. \quad (20)$$

In steady state, the system time $T_{1,i-1}$ has the exponential PDF (11). Furthermore, $T_{1,i-1}$ depends on packets and their service times that arrived prior to packet $1, i-1$. Thus $T_{1,i-1}$ is independent of X_{1i}.

Given B_i, packet $1, i-1$ is still in the system when packet $1, i$ is generated. The waiting time W_{1i} depends on both the residual system time $T_{1,i-1} - X_{1i}$ and also on the workloads of source 2 packets that arrive during the source 1 interarrival period of length X_{1i}. Specifically, let $M_2 = N_2(X_{1i})$ denote the number of source 2 packets that arrive during the source 1 interarrival period and let $S_{21}, S_{22}, \ldots S_{2M_2}$ denote the service requirements of these source 2 packets. As these packets are all queued between packets $1, i-1$ and $1, i$, we observe that

$$W_{1i} = (T_{1,i-1} - X_{1i}) + \sum_{j=1}^{M_2} S_{2j}. \quad (21)$$

This implies

$$\mathrm{E}[X_{1i}W_{1i}|B_i] = E_1 + E_2 \quad (22)$$

where

$$E_1 = \mathrm{E}\left[X_{1i}(T_{1,i-1} - X_{1i})|B_i\right], \tag{23}$$

$$E_2 = \mathrm{E}\left[X_{1i}\sum_{j=1}^{M_2} S_{2j}|B_i\right]. \tag{24}$$

By Lemma 2(b),

$$\begin{aligned} E_1 &= \mathrm{E}\left[(T_{1,i-1} - X_{1i})|B_i\right]\mathrm{E}\left[X_{1i}|B_i\right] && (25)\\ &= \left(\frac{1}{\mu-\lambda_1-\lambda_2}\right)\left(\frac{1}{\lambda_1+(\mu-\lambda_1-\lambda_2)}\right) && (26)\\ &= \frac{1}{\mu^2(1-\rho)(1-\rho_2)}. && (27)\end{aligned}$$

For the second term, iterated expectation yields

$$\begin{aligned} E_2 &= \int_0^\infty \mathrm{E}\left[X_{1i}\sum_{j=1}^{M_2} S_{2j}|B_i, X_{1i}=x\right] f_{X_{1i}|B_i}(x)\,dx \\ &= \int_0^\infty \mathrm{E}\left[x\sum_{j=1}^{M_2} S_{2j}|X_{1i}=x\right] f_{X_{1i}|B_i}(x)\,dx. && (28)\end{aligned}$$

Given that $X_{1i} = x$, $M_2 = N_2(X_{1i}) = N_2(x)$ is the number of source 2 arrivals in a period of length x and is Poisson with conditional expectation $\mathrm{E}[M_2|X_{1i} = x] = \lambda_2 x$. In addition, each S_{2j} is an exponential (μ) random variable, independent of X_{1i}, implying $\mathrm{E}[S_{2j}|X_{1i} = x] = 1/\mu$. This implies

$$\begin{aligned} \mathrm{E}\left[x\sum_{j=1}^{M_2} S_{2j}|X_{1i}=x\right] &= x\,\mathrm{E}\left[M_2|X_{1i}=x\right]\mathrm{E}\left[S_{2j}|X_{1i}=x\right] \\ &= x(\lambda_2 x)(1/\mu) = \rho_2 x^2. && (29)\end{aligned}$$

By Lemma 2, X_{1i} given B_i is an exponential (α) random variable with $\alpha = \lambda_1 + (\mu - \lambda_1 - \lambda_2) = \mu - \lambda_2$. These facts imply

$$E_2 = \rho_2 \int_0^\infty x^2 \alpha e^{-\alpha x}\,dx = \frac{2\rho_2}{\alpha^2} = \frac{2\rho_2}{\mu^2(1-\rho_2)^2}. \tag{30}$$

It follows from (22), (27) and (30) that

$$\mathrm{E}\left[W_{1i}X_{1i}|B_i\right] = \frac{1}{\mu^2}\left[\frac{2\rho_2}{(1-\rho_2)^2} + \frac{1}{(1-\rho_2)(1-\rho)}\right]. \tag{31}$$

Given event L_i, packet $1, i-1$ has departed the system prior to the arrival of packet $1, i$. In this case, the waiting time for packet $1, i$ depends on the number of source 2 packets in the system when packet $1, i$ arrives. To characterize this, we now let M_2 denote the number of source 2 packets in the system at the departure instant of packet $1, i-1$. Since the queue is FCFS, M_2 is the number of source 2 packets that arrived and queued during the system time $T_{1,i-1}$ of packet $1, i-1$. Given $T_{1,i-1}$ is exponential and independent of X_{1i}, Lemma 2(b) tells us that $T_{1,i-1}$ given L_i is conditionally an exponential (α) random variable with $\alpha = (\mu - \lambda_1 - \lambda_2) + \lambda_1 = \mu - \lambda_2$. As $T_{1,i-1}$ is independent of the subsequent Poisson arrivals of source 2, Lemma 3 implies that M_2 has the geometric PMF

$$P_{M_2}(m) = (1-\gamma)\gamma^m, \qquad (m \ge 0), \tag{32}$$

where $\gamma = \lambda_2/(\alpha + \lambda_2) = \rho_2$.

Thus at the departure instant of packet $1, i-1$, the number of source 2 packets in the system, M_2, is described by the stationary distribution for an M/M/1 queue serving only source 2 packets at rate λ_2. Going forward from this instant, we wait an additional time of length $X_{1i} - T_{1,i-1}$ for the arrival $1, i$ from source i. In this time period, there may be either arrivals or departures of source 2 packets. Nevertheless, as the queue holds zero source 1 packets, the operation of the queue is identical to an M/M/1 queue for just source 2 packets. At all times up to the arrival of packet $1, i$, the number of source 2 packets in the queue remains stochastically identical to M_2. It follows that when packet $1, i$ does arrive, the number of queued packets is described by M_2, independent of the additional delay $X_{1i} - T_{1,i-1}$ until the arrival of packet $1, i$. If the jth queued source 2 packet has service requirement S_{2j}, then $W_{1i} = \sum_{j=1}^{M_2} S_{2j}$ and

$$\begin{aligned} \mathrm{E}\left[X_{1i}W_{1i}|L_i\right] &= \mathrm{E}\left[X_{1i}\sum_{j=1}^{M_2} S_{2j}|L_i\right] \\ &= \mathrm{E}\left[X_{1i}|L_i\right]\mathrm{E}\left[\sum_{j=1}^{M_2} S_{2j}|L_i\right] \\ &= \mathrm{E}\left[T_{1,i-1} + (X_{1i} - T_{1,i-1})|L_i\right]\frac{\mathrm{E}\left[M_2|L_i\right]}{\mu} \\ &= \left(\frac{1}{\mu-\lambda_2} + \frac{1}{\lambda_1}\right)\frac{\rho_2}{\mu(1-\rho_2)}. && (33)\end{aligned}$$

Next we recall from Lemma 2 that independence of $T_{1,i-1}$ and X_{1i} implies $\mathrm{P}[B_i] = \rho_1/(1-\rho_2)$. Combining this fact with (20), (31), and (33), some algebra yields the claim. □

REFERENCES

[1] A. Mainwaring, D. Culler, and et al., "Wireless sensor networks for habitat monitoring," in *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, ser. WSNA '02. Atlanta, Georgia, USA: ACM, 2002, pp. 88–97, ACM ID: 570751.

[2] P. Papadimitratos, A. La Fortelle, K. Evenssen, R. Brignolo, and S. Cosenza, "Vehicular communication systems: Enabling technologies, applications, and future outlook on intelligent transportation," *IEEE Communications Magazine*, vol. 47, no. 11, pp. 84–95, Nov. 2009.

[3] A. Karakasidis, P. Vassiliadis, and E. Pitoura, "ETL queues for active data warehousing," in *Proceedings of the 2nd international workshop on Information quality in information systems*, ser. IQIS '05. Baltimore, Maryland: ACM, 2005, pp. 28–39, ACM ID: 1077509.

[4] H. Yu, L. Breslau, and S. Shenker, "A scalable web cache consistency architecture," *SIGCOMM Comput. Commun. Rev.*, vol. 29, no. 4, pp. 163–174, Aug. 1999, ACM ID: 316219.

[5] M. Xiong and K. Ramamritham, "Deriving deadlines and periods for real-time update transactions," in *The 20th IEEE Real-Time Systems Symposium, 1999. Proceedings*. IEEE, 1999, pp. 32–43.

[6] Y. C. Hu and D. B. Johnson, "Ensuring cache freshness in on-demand ad hoc network routing protocols," in *Proceedings of the second ACM international workshop on Principles of mobile computing*, 2002, pp. 25–30.

[7] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.

[8] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE INFOCOM Mini Conference*, 2012.