Pathwise Performance of Debt Based Policies for Wireless Networks with Hard Delay Constraints

Rahul Singh, I-Hong Hou, and P. R. Kumar

Department of Electrical and Computer Engineering, Texas A&M University

Email: {rsingh1,ihou,prk}@tamu.edu

Abstract— **Hou et al have introduced a framework to serve clients over wireless channels when there are hard deadline constraints along with a minimum delivery ratio for each client's flow. Policies based on "debt," called maximum debt first policies (MDF) were introduced, and shown to be throughput optimal. By "throughput optimality" it is meant that if there exists a policy that fulfils a set of clients with a given vector of delivery ratios and a vector of channel reliabilities, then the MDF policy will also fulfill them. The debt of a user is the difference between the number of packets that should have been delivered so as to meet the delivery ratio and the number of packets that have been delivered for that client. The maximum debt first (MDF) prioritizes the clients in decreasing order of debts at the beginning of every period. Note that a throughput optimal policy only guarantees that $\liminf_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{1}\{\text{client } n\text{'s packet is delivered in frame } t\} \geq q_i$, where the right hand side is the required delivery ratio for client i. Thus, it only guarantees that the debts of each user are $o(T)$, and can be otherwise arbitrarily large. This raises the interesting question about what is the growth rate of the debts under the MDF policy. We show the optimality of MDF policy in the case when the channel reliabilities of all users are same, and obtain performance bounds for the general case. For the performance bound we obtain the almost sure bounds on $\limsup_{t\to\infty} \frac{d_i(t)}{\phi(t)}$ for all i, where $\phi(t) = \sqrt{2t\log\log t}$.**

I. INTRODUCTION

Consider a wireless network consisting of an access point serving N clients as shown in Figure 1. We will assume that time is divided into slots. Let us call the set of slots $k\tau, \ldots, (k+1)\tau - 1$ as the k-th frame, and τ as the period. At the beginning of each frame, the access point generates one packet for each of the N clients. If such a packet is to be useful, it should be delivered in the same frame. That is, if a packet generated at the beginning of slot $k\tau$ is not delivered to the client by slot $(k+1)\tau - 1$, then it is considered to have expired and is dropped. We shall call the throughput of packets per frame that are delivered for client i as its timely throughput. We shall suppose that each client i requires a timely throughput q_i. In each slot, the access point can transmit one packet. So it has to choose which client's

packet to transmit, from among those clients that still have undelivered packets in that frame. The channels between the access point and the clients are however not reliable. A packet transmitted to client i is successfully delivered in that slot with probability p_i, and fails with probability $(1 - p_i)$. The access point can retransmit a failed packet at a later slot in the frame, before it expires. This model of a wireless network serving real-time flows with per packet hard deadlines was introduced in Hou et al [1]. It is useful in applications such as video streaming, voice over IP, and real-time applications such as networked control and cyberphysical systems where delay is critical.

The problem above is characterized by a period τ, and a channel reliability p_i and a timely throughput requirement q_i for each flow $i = 1, \ldots, N$. The first question that arises is: Can the access point meet the timely throughput requirements of each client? This question is answered in [1], where the set of timely throughput vectors $(q_1, \ldots, q_N)$ that are feasible is characterized.

To describe the characterization, let γ_i denote the geometrically distributed random variable with parameter p_i that denotes the number of attempts needed to deliver a packet of client i. If $\sum_{i=1}^{N} \gamma_i < \tau$, then that implies that all the packets that arrived at the beginning of a frame have been delivered, and so the remaining time in the frame $\left[\tau - \sum_{i=1}^{N} \gamma_i\right]$ is idle time where the access point
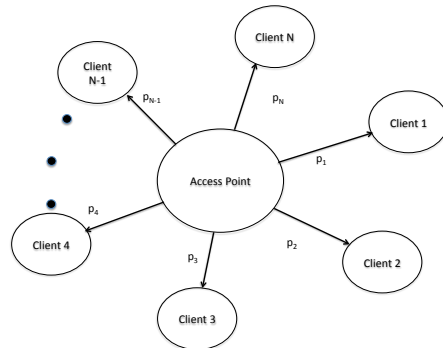

Fig. 1: An access point serving N real-time flows.

has no more packets to transmit. The quantity

$$I_{\{1,\ldots,N\}} := \frac{1}{\tau} E\left(\left[\tau - \sum_{i=1}^{N} \gamma_i\right]^+\right), \qquad (1)$$

where $x^+ := \max(x, 0)$, denotes the expected proportion of time in a frame that the access point is idle when the set of clients is $\{1, \ldots, N\}$.

Hou et al [1] have shown that a vector $(q_1, \ldots, q_N)$ of timely throughputs is feasible if and only if

$$\sum_{i \in S} \frac{q_i}{\tau p_i} \leq 1 - I_S, \text{ for every } S \subseteq \{1, \ldots, N\}, \qquad (2)$$

where $I_S = \frac{1}{\tau} E\left(\left[\tau - \sum_{i \in S} \gamma_i\right]^+\right)$ is the proportion of idle time when the set of clients is S. We may call this the "rate region".

It is also shown in [1] that the following class of weighted-debt policies will satisfy any set of clients for which the timely throughput vector is feasible. Let

$$d_i(t) := \frac{1}{\alpha_i}\Big[tq_i - \text{number of packets delivered for client}$$
$$i \text{ in the frames } \{0, \ldots, t-1\}\Big] \qquad (3)$$

denote the weighted packet-debt owed to client i at the beginning of frame t, where $\alpha_i > 0$ is a weighting parameter. They showed that if the clients are ordered in terms of decreasing weighted debt at the beginning of a frame, and then served in that order during the frame, then the resulting weighted debt policy is feasibility optimal. By this is meant that if a set of clients can be satisfied, i.e., satisfies (2), then this policy will meet their timely throughput requirements.

The timely throughput of client i, defined as

$$\liminf_{T \to \infty} \frac{\text{Number of packets of user delivered till frame } T}{T},$$

however only captures the long term average. It could be that there are long runs of frames where packets are not delivered. However such fluctuations in short term behaviour are ignored by the long term time average used in the definition of timely throughput. That is, the timely throughput does not capture a finite grained sense of performance. This is similar to queueing networks where there may be many policies that are throughput optimal, but whose delays may be exceedingly large.

Motivated by this, we address the issue of performance of real-time scheduling policies in this paper. In this study, we suppose that the system is in heavy traffic. In our context that corresponds to the vector of timely throughput requirements lying on the boundary of the rate region.

We assume the same framework as in [1], and study a scaled version of the unweighted debt process where $\alpha_i \equiv 1$, i.e.,

$$d_i(t) := tq_i - \text{number of packets delivered for client } i$$

in the frames $\{0, \ldots, t-1\}$. (4)

We scale the debt process by $\phi(t) := \sqrt{2t \log \log t}$ and study the almost sure limit set of the scaled process $\frac{d_i(t)}{\sqrt{2t \log \log t}}$. This is akin to the quantity studied in the law of the iterated logarithm [2]. We determine bounds on the limit set.

This is analogous to studying the "workload process" in queueing systems in the heavy traffic limit. The difference is that in our case the system is taken to be exactly in heavy traffic, and no limit need to be taken. Also, our characterization by the law of the iterated logarithm is a precise sample path characterization of the performance in heavy traffic.

The optimality of throughput-optimal policies has been considered in the "heavy trafic regime" in [3], [4]. Our problem differs from the general switch considered there in that the debt processes can become negative. Moreover the approach differs in that a diffusive scaled version of the "workload process" is studied and the results are given for a compact time interval. Another approach to show optimality of the MaxWeight scheduler under heavy traffic is taken in [4], where the expected value of steady state queue length, multiplied by the distance of the vector of arrival rates from the hyperplanes is studied and MaxWeight is shown to be optimal. Our approach differs from the previous works in that we consider the throughput requirements to lie on the boundary of the rate region and give results for a scaled version of the debt process as limit $t \to \infty$.

It has been shown in [1], that each of the 2^N subsets of clients imposes a "workload, idle-time" constraint, i.e. the sum of workloads over users in a subset must be less than the total time available to that subset of clients, which is the total time minus time spent idling. Our results proceed by showing that the debt vector evolves along the normal to the hyperplane on which the $\mathbf{q}$ vector lies. Then we derive an upper bound on the sums of the scaled debt processes, which must be satisfied by any non-idling throughput-optimal policy. These two together give us the bound on individual debt processes. We formulate the problem in Section II. Section III contains some preliminary results which will be used in later sections. Remaining sections contain results for various cases. In Section VI we introduce a notion of optimality and associate a cost with a policy. We show that the MWDF policy is optimal with respect to this cost. Finally Section VII summarizes the results.

II. Problem Formulation

For the remainder of this section, we will assume that $\tau = 1$, i.e. frame-length is one time slot. We consider the case where there are two users with $q_i, p_i, i = 1, 2$ as the parameters. We will analyze the scheduling policy that compares $\frac{d_1(t)}{p_1}$ and $\frac{d_2(t)}{p_2}$ and serves the client with higher weighted debt, and refer to it as the Maximum

Weighted Debt Policy (MWDF). First we will show that,

$$\lim_{t\to\infty} \frac{1}{\phi(t)} \left| \frac{d_1(t)}{p_1} - \frac{d_2(t)}{p_2} \right| = 0 \text{ a.s.,} \qquad (5)$$

where $\phi(t) = \sqrt{2t\log\log t}$. To this end we introduce the following lemma taken from [4].

Lemma 1: For an irreducible and aperiodic Markov Chain $\{X[t]\}_{t\geq 0}$ over a countable state space $\mathcal{X}$, suppose $Z : \mathcal{X} \to \mathbb{R}_+$ is a non-negative valued function. Define drift of Z at X as

$$\Delta Z(X) \triangleq [Z(X[t+1]) - Z(X[t])]\mathbb{1}\{X[t] = X\}. \quad (6)$$

Let the drift satisfy the following conditions:

- Condition $\mathcal{C}_1$: There exists an $\eta > 0$, and a $\kappa < \infty$ such that

$$E[\Delta Z(X)|X[t] = X] \leq -\eta, \forall X \in \mathcal{X} \text{ with } Z(X) \geq \kappa.$$

- Condition $\mathcal{C}_2$: There exists a $D < \infty$ such that

$$\Pr(|\Delta Z(X)| \leq D) = 1 \forall X \in \mathcal{X}.$$

Then there exists a $\theta^\star > 0$ and a $C^\star < \infty$ such that

$$\limsup_{t\to\infty} E[\exp^{\theta^\star Z(X[t])}] \leq C^\star.$$

Furthermore, if the Markov Chain is assumed to be positive recurrent, then $Z(X[t])$ converges in distribution to a random variable $\bar{Z}$ for which

$$E[\exp^{\theta^\star \bar{Z}}] \leq C^\star,$$

which implies that all moments of $\bar{Z}$ exist and are finite. Applying Lemma 1 we obtain,

Lemma 2: Under the MWDF policy (5) holds true.

Proof: Let $Z(\boldsymbol{d}) = |\frac{d_2}{p_2} - \frac{d_1}{p_1}|$. Let $\kappa = 1 + \frac{1}{\min(p_1, p_2)}$. Clearly when $Z(\boldsymbol{d}(t)) > \kappa$, the client with higher debt will be served in the time slot. Let $\frac{d_2(t)}{p_2} - \frac{d_1(t)}{p_1} > \kappa$. Then, if $\boldsymbol{d}(t)$ is such that $\frac{d_2(t)}{p_2} - \frac{d_1(t)}{p_1} > \kappa$,

$$E(\Delta Z(\boldsymbol{d})|\boldsymbol{d}) = p_2(\frac{d_2 + q_2 - 1}{p_2} - \frac{d_1 + q_1}{p_1})$$
$$+ (1-p_2)(\frac{d_2 + q_2}{p_2} - \frac{d_1 + q_1}{p_1}) - (\frac{d_2}{p_2} - \frac{d_1}{p_1})$$
$$= \frac{q_2}{p_2} - \frac{q_1}{p_1} - 1$$
$$= -2\frac{q_1}{p_1}.$$

Similarly, if $\frac{d_1(t)}{p_1} - \frac{d_2(t)}{p_2} > \kappa$, then $E(\Delta Z(\boldsymbol{d})|\boldsymbol{d}) < -2\frac{q_2}{p_2}$. Note also that if $\frac{d_2(t)}{p_2} - \frac{d_1(t)}{p_1} > \kappa$, then

$$\frac{d_2(t+1)}{p_2} - \frac{d_1(t+1)}{p_1} = \frac{d_2(t)}{p_2} - \frac{d_1(t)}{p_1} + \frac{q_2}{p_2} - \frac{q_1}{p_1}$$
$$- \frac{\mathbb{1}\{\text{packet for 2 is delivered}\}}{p_2}$$
$$\geq \frac{d_2(t)}{p_2} - \frac{d_1(t)}{p_1} - 1 - \frac{1}{p_2}$$
$$\geq 0,$$

where we have used the fact that $0 < \frac{q_i}{p_i} < 1$ and $\frac{d_2(t)}{p_2} - \frac{d_1(t)}{p_1} > \kappa$. Hence the clients maintain the order of debts if the difference is greater than κ. So we can take $\eta = -2\min\frac{q_1}{p_1}, \frac{q_2}{p_2}$. It remains to show that the drifts are bounded for all $\boldsymbol{d}$. If $\frac{d_2}{p_2} - \frac{d_1}{p_1} > \kappa$ then since the debts don't change the order, the drift is bounded simply by

$$\frac{q_1}{p_1} + \frac{q_2}{p_2} + \frac{1}{p_1} + \frac{1}{p_2}.$$

On the other hand, if $\frac{d_2}{p_2} - \frac{d_1}{p_1} < \kappa$, then

$$\left| \frac{d_2(t+1)}{p_2} - \frac{d_1(t+1)}{p_1} \right| - \left| \frac{d_2(t)}{p_2} - \frac{d_1(t)}{p_1} \right|$$
$$\leq \left| \frac{d_2(t+1)}{p_2} - \frac{d_2(t)}{p_2} + \frac{d_2(t)}{p_2} \right.$$
$$\left. + \frac{d_1(t)}{p_1} - \frac{d_1(t+1)}{p_1} - \frac{d_1(t)}{p_1} \right| + \kappa$$
$$\leq \left| \frac{d_2(t+1)}{p_2} - \frac{d_2(t)}{p_2} \right| + \left| \frac{d_1(t)}{p_1} - \frac{d_1(t+1)}{p_1} \right|$$
$$+ \left| \frac{d_2(t)}{p_2} - \frac{d_1(t)}{p_1} \right| + \kappa$$
$$\leq \frac{1+q_2}{p_2} + \frac{1+q_1}{p_1} + 2\kappa.$$

Hence for some finite $N(\epsilon)$, and $\theta^\star > 0$,

$$E[\exp^{\theta^\star |\frac{d_2(t)}{p_2} - \frac{d_1(t)}{p_1}|}] < C^\star, \forall t > N(\epsilon). \qquad (7)$$

Now applying the Borel Cantelli lemma, we have for any $\delta > 0$,

$$\sum_{t=1}^{\infty} \Pr\left(\frac{1}{\phi(t)} \left| \frac{d_1(t)}{p_1} - \frac{d_2(t)}{p_2} \right| > \delta \right)$$
$$\leq \frac{BN(\epsilon)}{\delta} + \frac{C^\star + \epsilon}{\exp(\theta^\star \delta \phi(t))}$$
$$\leq \frac{BN(\epsilon)}{\delta} + \frac{C^\star + \epsilon}{\exp(\delta \theta^\star \sqrt{t})}$$
$$< \infty, \qquad (8)$$

where the first inequality follows from inequality (7) applied with Markov inequality, and the second inequality follows by the fact that $\phi(t) > \sqrt{t}$. Now (8) completes the proof. ∎

III. PRELIMINARY RESULTS

Now we will introduce some auxiliary random variables and results, which will be useful in the later proofs. Let

$$m_j(t) = \mathbb{1}\{\text{user } j \text{ is attempted at time } t\}$$
$$- \frac{1}{p_j}\mathbb{1}\{\text{user } j \text{ is delivered at time } t\}$$
$$= u_j(t) - \frac{1}{p_j}g_j(t). \qquad (9)$$

Clearly

$$M_j(t) = \sum_{l=1}^{t} m_j(l), \qquad (10)$$

is a martingale since

$$E[M_j(t+1)|\mathcal{F}_t] = E[M_j(t) + m_j(t+1)|\mathcal{F}_t]$$

$$= M_j(t) + E\left[u_j(t+1) - \frac{1}{p_j}g_j(t+1)|\mathcal{F}_t\right]$$

$$= M_j(t) + E\left[E\left[u_j(t+1)\right.\right.$$
$$\left.\left. - \frac{1}{p_j}g_j(t+1)\Big|\mathcal{F}_t, u_j(t+1)\right]\Big|\mathcal{F}_t\right]$$

$$= M_j(t)$$

We will now provide some limit results for any policy for which q_1, q_2 is feasible. First we state the law of large numbers taken from [5]

Theorem 1: Let $Y_n = \sum_{i=1}^{n} X_i$ be a Martingale such that $\sum_{n=1}^{\infty} \frac{E(X_n^2)}{n^2} < \infty$. Then $\frac{Y_n}{n} \to 0$ a.s.
This immediately gives us the following for any throughput optimal policy.

Lemma 3: The following is true for any non-idling policy fulfilling q_1, q_2:

$$\lim_{T \to \infty} \frac{\sum_{t=1}^{T} u_j(t)}{T} = \frac{q_j}{p_j}.$$

Proof: Clearly $E(m_j^2(t)) \le (1 + \frac{1}{p_j})^2$ and so the conditions of Theorem (1) are satisfied. Hence,

$$\frac{M_j(t)}{t} \to 0, \qquad (11)$$

and so

$$\frac{\sum_{l=1}^{t} u_j(l)}{t} \to \frac{1}{p_j} \frac{\sum_{l=1}^{t} s_j(l)}{t}. \qquad (12)$$

But since $\liminf_{t\to\infty} \frac{\sum_{l=1}^{t} s_j(l)}{t} \ge q_j$, we have

$$\liminf_{t\to\infty} \frac{\sum_{l=1}^{t} u_j(l)}{t} \ge \frac{q_j}{p_j}. \qquad (13)$$

Now,

$$\sum_{j=1}^{N} \frac{\sum_{l=1}^{t} u_j(l)}{t} = 1,$$

holds for all t. If $\limsup_{t\to\infty} \frac{\sum_{l=1}^{t} u_j(l)}{t} > \frac{q_j}{p_j}$, then this would imply $\liminf_{t\to\infty} \frac{\sum_{l=1}^{t} u_k(l)}{t} < \frac{q_k}{p_k}$ for some k since $\sum_{j=1}^{N} \frac{q_j}{p_j} = 1$, which contradicts (13). This completes the proof. ∎
Let $(X_n, \mathcal{F}_n)_{n\ge 0}$ be a Martingale with $Y_n = X_n - X_{n-1}$ and $s_n^2 = \sum_{i=1}^{n} E[Y_i^2|\mathcal{F}_{i-1}]$. Also let $u_n = \sqrt{2n \log\log n}$
The following is the law of iterated logarithm for martingales [6].

Theorem 2: If $s_n^2 \to \infty$ and $|Y_n| \le \frac{K_n s_n}{u_n}$, where K_n are $\mathcal{F}_{n-1}$ measurable with $K_n \to 0$, then

$$\limsup_{n\to\infty} \frac{X_n}{s_n u_n} = 1. \qquad (14)$$

The following lemma uses Theorem (2) and will be crucial in later results. Equalities are to be understood in almost sure sense.
Lemma 4: Consider the Martingale $M_j(t)$. The following are true for any non-idling policy that fulfils q_1, q_2:

(i). $E(m_j^2(t)|\mathcal{F}_{t-1}) = u_j(t)v_j$, with $v_j = \frac{(1-p_j)}{p_j}$.

(ii). $\mathcal{S}_j^2(T) \triangleq \sum_{j=1}^{T} E(m_j^2(t)|\mathcal{F}_{t-1}) = \sum_{t=1}^{T} u_j(t)v_j$.

(iii). $\lim_{T\to\infty} \frac{\mathcal{S}_j^2(T)}{T} = \frac{q_j}{p_j}v_j = \frac{q_j(1-p_j)}{p_j^2} \triangleq c_j$.

(iv). $\lim_{T\to\infty} \frac{\mathcal{S}_j(T)}{\sqrt{T}} = \sqrt{c_j}$

Using the above three, also noting that $m_j(t)$ is bounded for all t, and applying Theorem 2 with, $K_n = \frac{\sqrt{\log\log n}}{\sum_{t=1}^{n} E(m_j^2(t)|\mathcal{F}_{t-1})} \to 0$ gives us,

$$\limsup_{T\to\infty} \frac{M_j(t)}{\phi(t)} = \sqrt{c_j}. \qquad (15)$$

Theorem 3: Under the MWDF policy,

$$\limsup_{t\to\infty} \frac{d_j(t)}{\phi(t)} \le p_j \frac{\sqrt{c_1} + \sqrt{c_2}}{2}. \qquad (16)$$

Proof: By (9) and (10), we have,

$$M_1(t) + M_2(t) = t - \frac{s_1(t)}{p_1} - \frac{s_2(t)}{p_2}$$
$$= t(\frac{q_1}{p_1} + \frac{q_2}{p_2}) - \frac{s_1(t)}{p_1} - \frac{s_2(t)}{p_2}$$
$$= \frac{d_1(t)}{p_1} + \frac{d_2(t)}{p_2}. \qquad (17)$$

Now,

$$2\frac{d_1(t)}{p_1} + \frac{d_2(t)}{p_2} - \frac{d_1(t)}{p_1} = M_1(t) + M_2(t)$$

$$\implies \frac{1}{\phi(t)}\left[2\frac{d_1(t)}{p_1} + \frac{d_2(t)}{p_2} - \frac{d_1(t)}{p_1}\right] = \frac{M_1(t)}{\phi(t)} + \frac{M_2(t)}{\phi(t)}$$

$$\implies \frac{2}{\phi(t)}\frac{d_1(t)}{p_1} = \frac{M_1(t)}{\phi(t)} + \frac{M_2(t)}{\phi(t)}$$
$$- \left[\frac{d_2(t)}{p_2} - \frac{d_1(t)}{p_1}\right]$$

$$\implies \limsup_{t\to\infty} \frac{2}{\phi(t)}\frac{d_1(t)}{p_1} \le \limsup_{t\to\infty} \frac{M_1(t)}{\phi(t)}$$
$$+ \limsup_{t\to\infty} \frac{M_2(t)}{\phi(t)}$$
$$+ \limsup_{t\to\infty}\left[\frac{d_2(t)}{p_2} - \frac{d_1(t)}{p_1}\right].$$

The third of the above quantities on right side is 0 by Lemma 2. So applying Lemma 4 we get the required result. ∎

IV. TWO USERS WITH GENERAL FRAME LENGTH

We will look at the case when the timely throughputs lie on exactly one hyperplane, given by

$$\frac{q_1}{p_1} + \frac{q_2}{p_2} = \tau(1 - I_{1,2}). \qquad (18)$$

where $I_{1,2}$ is the expected idle time in a frame after delivering both packets. The MWDF policy compares $\frac{d_1(t)}{p_1}$ and $\frac{d_1(t)}{p_1}$ and serves them in the decreasing order. As before we will consider a Lyapunov function to show (5) .

Lemma 5: For the system (18) under the MWDF policy (5) holds true.

Proof: Let $Z(\boldsymbol{d}) = |\frac{d_2}{p_2} - \frac{d_1}{p_1}|$. Let $\kappa = 1 + \frac{2}{\min(p_1,p_2)}$. Let $\frac{d_2(t)}{p_2} - \frac{d_1(t)}{p_1} > \kappa$. Then, the order $\{2,1\}$ is used in frame $t+1$. Let π_2 and π_1 denote the probabilities that packets for respective users are delivered under the ordering $\{2,1\}$. It has been shown in [1] that,

$$\frac{\pi_2}{p_2} = \tau(1 - I_{\{2\}}),$$
$$\frac{\pi_2}{p_2} + \frac{\pi_1}{p_1} = \tau(1 - I_{\{1,2\}}).$$

For $\boldsymbol{d}$ such that $\frac{d_2(t)}{p_2} - \frac{d_1(t)}{p_1} > \kappa$, we have

$E\left(\Delta Z\left(\boldsymbol{d}\right)|\boldsymbol{d}\right)$

$=E\left[\frac{d_2(t+1)}{p_2} - \frac{d_1(t+1)}{p_1} - \left(\frac{d_2(t)}{p_2} - \frac{d_1(t)}{p_1}\right)\bigg|\text{order }\{2,1\}\right]$

$=E\bigg[\left(\frac{d_2(t+1)}{p_2} - \frac{d_2(t)}{p_2}\right)$

$\qquad - \left(\frac{d_1(t+1)}{p_1} - \frac{d_1(t)}{p_1}\right)\bigg|\text{order }\{2,1\}\bigg]$

$=\frac{q_2}{p_2} - \frac{\pi_2}{p_2} - \left(\frac{q_1}{p_1} - \frac{\pi_1}{p_1}\right)$

$=\frac{q_2}{p_2} - 2\frac{\pi_2}{p_2} - \frac{q_1}{p_1} + \frac{\pi_1}{p_1} + \frac{\pi_2}{p_2}$

$=\frac{q_2}{p_2} - 2\tau\left(1 - I_{\{2\}}\right) - \frac{q_1}{p_1} + \tau\left(1 - I_{\{1,2\}}\right)$

$=2\frac{q_2}{p_2} - 2\tau\left(1 - I_{\{2\}}\right) - \left(\frac{q_1}{p_1} + \frac{q_2}{p_2}\right) + \tau\left(1 - I_{\{1,2\}}\right)$

$=2\frac{q_2}{p_2} - 2\tau\left(1 - I_{\{2\}}\right) - \tau\left(1 - I_{\{1,2\}}\right) + \tau\left(1 - I_{\{1,2\}}\right)$

$=2\left(\frac{q_2}{p_2} - \tau\left(1 - I_{\{2\}}\right)\right)$

$<0.$

Similarly

$$E(\Delta Z(\boldsymbol{d})|\boldsymbol{d}) < 0,$$

when $\frac{d_1(t)}{p_1} - \frac{d_2(t)}{p_2} > \kappa$. So we can take $\eta = \max 2\left(\frac{q_2}{p_2} - \left(1 - I_{\{2\}}\right)\right), 2\left(\frac{q_1}{p_1} - \left(1 - I_{\{1\}}\right)\right)$.

We need to show that the drift is uniformly bounded. As in previous section, once again if $\frac{d_2}{p_2} - \frac{d_1}{p_1} > \kappa$, drift is bounded simply by

$$\frac{1 + q_1}{p_1} + \frac{1 + q_2}{p_2}.$$

If $\frac{d_2}{p_2} - \frac{d_1}{p_1} < \kappa$, then once again,

$$\left|\frac{d_2(t+1)}{p_2} - \frac{d_1(t+1)}{p_1}\right| - \left|\frac{d_2(t)}{p_2} - \frac{d_1(t)}{p_1}\right|$$

$$\leq \frac{1 + q_2}{p_2} + \frac{1 + q_1}{p_1} + 2\kappa$$

For the remaining part, we can once again apply Borel Cantelli to get the claim. ∎

Let $i(t)$ denote the time spent idling in frame t. Note that $i(t)$ are independent and identically distributed with mean $\tau I_{\{1,2\}}$ and finite variance, which will be denoted as σ_I^2. Next we state without proof the results similar to previous section. The derivation of the first result is exactly the same as in the previous section, while the second follows using Kolmogorov's Law of the Iterated Logarithm.

Lemma 6: 1)

$$\limsup_{T\to\infty} \frac{M_j(t)}{\phi(t)} = \sqrt{c_j}$$

2)

$$\limsup_{t\to\infty} \frac{\sum_{l=1}^{t}\left(i(l) - \tau I_{1,2}\right)}{\phi(t)} = \sigma_I.$$

Theorem 4: Under the MDF policy,

$$\limsup_{t\to\infty} \frac{d_j(t)}{\phi(t)} \leq p_j \frac{\sqrt{c_1} + \sqrt{c_2} + \sigma_I}{2}. \qquad (19)$$

V. Multiple Users With $\tau = 1$

In this section we consider the case of N users with frame size of one time-slot. Since there is no idling in this case, the rate region is given by,

$$\sum_{j=1}^{N} \frac{q_j}{p_j} \leq 1. \qquad (20)$$

For the remainder of this section $\boldsymbol{q}$ will denote the vector $(q_1, \ldots, q_N)$, and it will be assumed to lie on the hyperplane defined by equation in (20). First we will show that,

$$\lim_{t\to\infty} \frac{1}{\phi(t)}\left|\frac{d_j(t)}{p_j} - \frac{d_k(t)}{p_k}\right| = 0 \text{ a.s.}, \forall j, k = 1, \ldots, N. \qquad (21)$$

We introduce the following Lyapunov function,

$$Z(\boldsymbol{x}) = \sum_{j=1}^{N}(x_j - x_{\min}), \qquad (22)$$

where $x_{\min}$ is the minimum entry of $\boldsymbol{x}$. Let $c := 1 + \max_{j=1,\ldots,N} \frac{1 + 2q_j}{p_j}$.

Lemma 7: Under the MWDF policy (21) holds true.

Proof: Consider $Z(\boldsymbol{x})$ as in (22). We will use the notation $\frac{\boldsymbol{d(t)}}{\boldsymbol{p}}$ to refer to the vector with entries $\frac{d_j(t)}{p_j}$. Note that

$$Z(\boldsymbol{x}) > Nc \implies \sum_{j=1}^{N}(x_j - x_{\min}) > Nc$$
$$\implies N(x_{\max} - x_{\min}) > Nc$$
$$\implies (x_{\max} - x_{\min}) > c.$$

Suppose now that

$$Z\left(\frac{\boldsymbol{d(t)}}{\boldsymbol{p}}\right) > Nc.$$

Then,

$$\left(\frac{\boldsymbol{d(t)}}{\boldsymbol{p}}\right)_{\max} > c.$$

Let $j \in \arg\max \frac{d_j(t)}{p_j}$. Then,

$$\frac{d_j(t+1)}{p_j} - \left[\frac{\boldsymbol{d(t+1)}}{\boldsymbol{p}}\right]_{\min} > \frac{d_j(t)}{p_j} - \frac{1+q_j}{p_j}$$
$$- \left[\left(\frac{\boldsymbol{d(t)}}{\boldsymbol{p}}\right)_{\min} + \max_{j=1,\ldots,N}\frac{q_j}{p_j}\right]$$
$$> 0,$$

where the inequalities are the result of the facts that the client being served cannot decrease in debt by more than $-1 + q_j$, and the maximum increment in the minimum value can be $\max_{j=1,\ldots,N}\frac{q_j}{p_j}$. Set $\kappa = Nc$. Also note that since the frame size is one slot, only one user $j \in \arg\max \frac{d_j(t)}{p_j}$ will be receiving service in the present frame and hence the debts of users at time $t+1$, other than this maximum weighted debt user, are known at time t itself. Also if $Z\left(\frac{\boldsymbol{d(t)}}{\boldsymbol{p}}\right) > Nc$ we know that the max weighted debt user will not be the minimum weighted debt user in the next frame. We will use Z_{t+1} to mean $Z(\frac{\boldsymbol{d(t+1)}}{\boldsymbol{p}})$ and j_t, j_{t+1} to refer to the indices of the clients with the least weighted debt at times t and $t+1$ respectively, and by $j^{\max}$ to the index of the client with highest weighted debt at the beginning of frame $t+1$. This gives us, for $Z_t > Nc$,

$$E\left[Z_{t+1}|Z_t\right] = \sum_{j \neq j_{t+1}, j^{\max}} \left(\frac{d_j(t)}{p_j} + \frac{q_j}{p_j}\right)$$
$$- \left(\frac{d_{j_{t+1}}(t)}{p_{j_{t+1}}} + \frac{q_{j_{t+1}}}{p_{j_{t+1}}}\right) + \left(\frac{d_{j^{\max}}(t)}{p_{j^{\max}}} + \frac{q_{j^{\max}}}{p_{j^{\max}}} - 1\right)$$
$$- \left(\frac{d_{j_{t+1}}(t)}{p_{j_{t+1}}} + \frac{q_{j_{t+1}}}{p_{j_{t+1}}}\right) \quad (23)$$

Also,

$$Z_t = \sum_{j \neq j_t, j^{\max}} \left(\frac{d_j(t)}{p_j}\right) - \left(\frac{d_{j_t}(t)}{p_{j_t}}\right)$$
$$+ \left(\frac{d_{j^{\max}}(t)}{p_{j^{\max}}}\right) - \left(\frac{d_{j_t}(t)}{p_{j_t}}\right). \quad (24)$$

Subtracting the above two, and using (20) we get for $d(t) = d$ such that $Z\left(\frac{\boldsymbol{d(t)}}{\boldsymbol{p}}\right) > Nc$,

$$E\left[\Delta Z(\boldsymbol{d})|\boldsymbol{d}\right]$$
$$= N\left[\frac{d_{j_t}(t)}{p_{j_t}} - \frac{d_{j_{t+1}}(t)}{p_{j_{t+1}}} - \frac{q_{j_{t+1}}}{p_{j_{t+1}}}\right]$$
$$< 0. \quad (25)$$

In the last inequality we have used the fact,

$$\frac{d_{j_{t+1}}(t)}{p_{j_{t+1}}} \geq \frac{d_{j_t}(t)}{p_{j_t}} \quad (26)$$

and also that,

$$\frac{d_{j_t}(t)}{p_{j_t}} + \frac{q_{j_t}}{p_{j_t}} \geq \frac{d_{j_{t+1}}(t)}{p_{j_t+1}} + \frac{q_{j_{t+1}}}{p_{j_{t+1}}}, \quad (27)$$

and hence

$$\frac{d_{j_t}(t)}{p_{j_t}} - \frac{d_{j_{t+1}}(t)}{p_{j_t+1}} \geq \frac{q_{j_{t+1}}}{p_{j_{t+1}}} - \frac{q_{j_t}}{p_{j_t}}. \quad (28)$$

Combining (26) and (28), when $Z\left(\frac{\boldsymbol{d(t)}}{\boldsymbol{p}}\right) > Nc$

$$-N\left(\frac{q}{p}\right)_{\max} \leq E\left[\Delta Z(\boldsymbol{d})|\boldsymbol{d}\right] \leq -N\left(\frac{q}{p}\right)_{\min}. \quad (29)$$

The boundedness of drift when $Z(\boldsymbol{d}) \leq Nc$ follows trivially by boundedness of the increments of individual debts by $\frac{q_j+1}{p_j}$. Hence we can use Lemma 1 in conjunction with the Borel Cantelli lemma to infer that

$$\lim_{t \to \infty} \frac{1}{\phi(t)}\left[\sum_{j=1}^{N}\frac{d_j(t)}{p_j} - \left(\frac{\boldsymbol{d(t)}}{\boldsymbol{p}}\right)_{\min}\right] = 0 \text{ a.s.}$$

Therefore,

$$\lim_{t \to \infty} \frac{1}{\phi(t)}\left[\frac{d_j(t)}{p_j} - \frac{d_k(t)}{p_k}\right] = 0 \text{ a.s.}, \forall j, k.$$

∎

The main part of the rest of the analysis lies in showing that the debt vector evolves along the perpendicular to the hyperplane. The rest of the analysis is exactly along the lines of section II, and is omitted. We directly state the main result.

Theorem 5: Under the MWDF policy,

$$\limsup_{t \to \infty} \frac{d_j(t)}{\phi(t)} \leq p_j \frac{\sum_{j=1}^{N}\sqrt{c_j}}{N}. \quad (30)$$

VI. GENERAL CASE WITH SYMMETRIC USERS

We will consider the case where there are N users with the vector of timely throughput requirements denoted as $\boldsymbol{q}$ with all entries as q, i.e. $q_i \equiv q, \forall i$. The channel reliabilities of each user will be the same and denoted as p, i.e. $p_i \equiv p, \quad \forall i$. The frame size is τ time slots. The rate-region is characterized as the intersection of the following 2^N halfspaces:

$$\sum_{j \in S}\frac{q_j}{p_j} \leq 1 - I_S, S \subseteq \{1, \ldots, N\}. \quad (31)$$

We wil assume that $\boldsymbol{q}$ lies on the hyperplane

$$N\frac{q}{p} = \tau\left(1 - I_{\{1,\ldots,N\}}\right), \quad (32)$$

and specifically in its relative interior. The following holds for any non-idling policy,

$$\sum_{j=1}^{N}\frac{d_j(t+1) - d_j(t)}{p}$$

$$= \sum_{j=1}^{N} \frac{q - \mathbb{1}\{\text{packet of user } j \text{ is delivered in frame } t+1 \}}{p}$$

$$= \sum_{j=1}^{N} \frac{q}{p}$$

$$- \sum_{j=1}^{N} \frac{\mathbb{1}\{\text{packet of user } j \text{ is delivered in frame } t+1 \}}{p}$$

$$\tag{33}$$

Denoting the probability that a packet of user j is delivered in frame $t+1$ by π_j,

$$E\left[\sum_{j=1}^{N} \frac{d_j(t+1) - d_j(t)}{p}\right] = \sum_{j=1}^{N} \frac{q}{p} - \sum_{j=1}^{N} \frac{\pi_j}{p}$$

$$= \sum_{j=1}^{N} \frac{q}{p} - \tau\left(1 - I_{\{1,\ldots,N\}}\right)$$

$$= 0, \tag{34}$$

where we have used $\frac{\pi_j}{p} = \tau\left(1 - I_{\{1,\ldots,N\}}\right)$ from [1] page 5 and (32). Also note that the distribution of $\sum_{j=1}^{N} \frac{d_j(t+1) - d_j(t)}{p}$ is the same for any non-iding policy, and the distribution of

$$\sum_{j=1}^{N} \mathbb{1}\{\text{packet of user } j \text{ is delivered in frame } t+1 \}$$

$$\triangleq y(t)$$

is given by

$$\Pr\left(y(t) = x\right) = \begin{cases} \binom{\tau}{x} p^x (1-p)^{\tau-x} & \text{if } x < N, \\ \sum_{y=N}^{\tau} \binom{y}{N} p^N (1-p)^{y-N} & \text{if } x = N, \\ 0 & \text{if } x > N. \end{cases}$$

Denote by $\sigma_{p,\tau}$ the variance of $\sum_{j=1}^{N} \frac{d_j(t+1) - d_j(t)}{p}$. Summing (33) over t, using the fact that $\sum_{j=1}^{N} \frac{d_j(t+1) - d_j(t)}{p}$ are i.i.d. with mean zero (34) and finite variance, we can use Kolmogorov's law of the iterated logarithm to infer that,

$$\limsup_{t\to\infty} \frac{1}{\phi(t)} \left[\sum_{j=1}^{N} \frac{d_j(t)}{p}\right] = \sigma_{p,\tau} \text{ a.s., and}$$

$$\liminf_{t\to\infty} \frac{1}{\phi(t)} \left[\sum_{j=1}^{N} \frac{d_j(t)}{p}\right] = -\sigma_{p,\tau} \text{ a.s.} \tag{35}$$

Lemma 8: For the system (32),

$$\lim_{t\to\infty} \frac{1}{\phi(t)} [d_j(t) - d_k(t)] = 0 \text{ a.s. .} \tag{36}$$

Proof: Note that

$$d_j(t) - d_k(t) \le 1, \forall j, k,$$

for all t since the MWDF policy gives preference to the client with higher debt. ∎

Theorem 6: Under the MWDF policy for the system (32), we have

$$\limsup_{t\to\infty} \frac{1}{\phi(t)} \frac{d_j(t)}{p} = \frac{\sigma_{p,\tau}}{N} \text{ a.s., } \forall j. \tag{37}$$

Proof: This follows by using Lemma 8 in conjunction with (35). ∎

A. A Notion of Optimality.

We now introduce an appropriate notion of optimality for the heavy traffic regime and show that the MWDF policy is optimal for the heavy traffic regime for the symmetric scenario of this section. We will restrict ourselves to policies satisfying $\limsup_{t\to\infty} \frac{1}{\phi(t)} \left|\frac{d_j(t)}{p}\right|$ is finite a.s.. Also we restrict ourselves to policies such that,

$$\limsup_{t\to\infty} \frac{1}{\phi(t)} \frac{d_j(t)}{p} = v_j^\pi \text{ a.s. } j = 1,\ldots,N \text{ and}$$

$$\liminf_{t\to\infty} \frac{1}{\phi(t)} \frac{d_j(t)}{p} = w_j^\pi \text{ a.s. } j = 1,\ldots,N$$

$$\text{such that } \sum_{j=1}^{N} v_j^\pi = \sigma_{p,\tau},$$

$$\sum_{j=1}^{N} w_j^\pi = -\sigma_{p,\tau}, \tag{38}$$

where the superscript denotes the variables corresponding to policy π. Note that v_j^π and w_j^π can be negative or positive. The set of policies satisfying (38) will be denoted as Π. It is obvious that we do not lose anything by restricting to policies in Π since (35) is true for any non-idling policy. The cost associated with a policy is

$$\max_{j=1,\ldots,N} v_j^\pi.$$

More precisely, we are interested in the following optimization problem,

$$\min_{\pi\in\Pi} \max_{j=1,\ldots,N} v_j^\pi$$

$$\text{such that } \sum_{j=1}^{N} v_j^\pi = \sigma_{p,\tau}, \tag{39}$$

where v_j are as in (38), and π denotes a policy. It is clear that the optimal value of this problem is bounded below by $\frac{\sigma_{p,\tau}}{N}$ and occurs when all v_j are equal. Since the MWDF policy attains this bound (Theorem 6), it is optimal.

Optimality of MWDF policy can also be shown in the cases considered in the previous sections. The performance cost defined as above can be seen as guaranteeing a sort of "fairness" amongst different flows.

VII. CONCLUSIONS AND FURTHER WORK

We have analyzed the performance of real-time wireless networks in heavy traffic. We have performed an analysis of the scaled version of debt process and provided bounds on the process as $t \to \infty$. We have also introduced a "notion of optimality" and have shown that

the MWDF policy is optimal in this sense. We believe an analysis of the general case is possible, the primary step being to show that the debt process evolves along the perpendicular to the hyperplane on which the timely throughput vector lies. A careful construction of the Lyapunov function is to be done to achieve this. Once this is done, it should be possible to show the optimality of the MWDF policy with respect to the performance measure defined in the previous section.

REFERENCES

[1] I-Hong Hou, Vivek Borkar and P. R. Kumar, "A theory of QoS for wireless," in *In Proc. of IEEE INFOCOM*, 2009.

[2] A. Kolmogoroff, "Uber das gesetz des iterierten logarithmus," *Mathematische Annalen*, pp. 126–135, 1929.

[3] A. Stolyar, "Max Weight Scheduling in a Generalized Switch: State Space Collapse and Workload Minimization in Heavy Traffic," in *Annals of Applied Probability*, vol. 14, no. 1, 2004, pp. 1–53.

[4] A. Eryilmaz and R. Srikant, "Asymptotically tight steady-state queue length bounds implied by drift conditions," *Queueing Systems*, vol. 72.

[5] Y. Chow, "On a strong law of large numbers for martinagles," *The Annals of Mathematical Statistics*, vol. 38, no. 2, pp. 610–610, 1967.

[6] W. Stout, "A martingale analogue of kolmogorov's law of the iterated logarithm," *Zeitschrift fr Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 15, no. 4, pp. 279–290, 1970. [Online]. Available: http://dx.doi.org/10.1007/BF00533299