Variable Frame Based Max-Weight Algorithms for Networks with Switchover Delay

G. Çelik, S. Borst, P. Whiting, E. Modiano

Abstract—**This paper considers the scheduling problem for networks with interference constraints and switchover delays, where it takes a nonzero time to reconfigure each service schedule. Switchover delay occurs in many telecommunication applications such as satellite, optical or delay tolerant networks (DTNs). Under zero switchover delay it is well known that the Max-Weight algorithm is throughput-optimal without requiring knowledge of the arrival rates. However, we show that this property of Max-Weight no longer holds when there is a nonzero switchover delay. We propose a class of variable frame based Max-Weight (VFMW) algorithms which employ the Max-Weight schedule corresponding to the beginning of the frame during an interval of duration dependent on the queue sizes. The VFMW algorithms dynamically adapt the frame sizes to the stochastic arrivals and provide throughput-optimality *without requiring knowledge of the arrival rates*. Numerical results regarding the application of the VFMW algorithms to DTN and optical networks demonstrate a good delay performance.**

I. Introduction

Dynamic scheduling of servers in stochastic networks with interference constraints has been a very active field [3], [4],[5],[9],[10]. However, the significant effects of server switchover delays or the time to reconfigure schedules have been largely ignored. Switchover delay occurs in many practical communications systems. In satellite systems it can take the satellite antenna about 10ms to switch from one ground station to another [2]. Electronic beamforming in wireless radios and laser tuning for optical transceivers can take μs-ms [3], [2]. Furthermore, in DTNs significant switchover delays occur when mobile servers (e.g., unmanned aerial vehicles (UAV)) are used as data gatherers from sensors in a field. We consider a general queueing model for stochastic networks which include wireless uplinks/downlinks, optical networks, or DTNs as special cases and study the impact of switchover delays on optimal policies.

Scheduling algorithms for wireless networks with interference constraints have received considerable attention in the past two decades. The seminal papers by Tassiulas and Ephremides [9], [10] characterized the stability region of such systems and proved throughput-optimality for the Max-Weight scheduling policy that works without requiring arrival rate information. These results were generalized to many different settings such as power allocation and routing or optimal scheduling in switches (e.g., [5], [6], [8]). These models do not consider the server switchover delays and we show in Section II-A that they fail to provide stability when there are switchover delays.

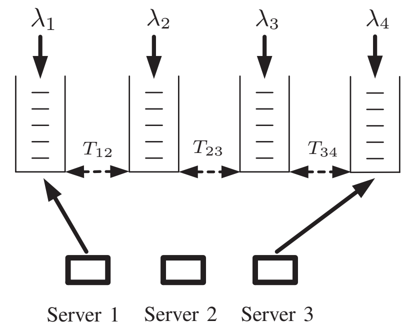

Fig. 1: System model with $N = 4$ queues and $M = 3$ servers. Server 2 is forced to be idle due to interference constraints.

For a simple two-queue system, [4] shows that simultaneous presence of randomly varying connectivity for users and switchover delays for servers reduces the stability region of the system significantly. It is known that in the absence of randomly varying connectivity, switchover delay does not reduce the stability region [3]. However, the celebrated Max-Weight policy is not throughput-optimal as it decides to reconfigure the schedule too often, incurring large throughput losses during reconfiguration. In order to overcome the negative effects of reconfiguration delays, [3] considered a frame based scheme which persists with a Max-Weight schedule during a frame of fixed duration. This scheme was shown to be throughput-optimal *when the arrival rate vector was known in advance*. The fluid limit of the system was considered in [3] and throughput-optimality was established under rate stability[1].

In this paper we propose a new class of policies called *Variable Frame Based Max-Weight* (VFMW) policies that operate over frames of dynamically changing duration based on queue states and show that they are throughput-optimal *without requiring knowledge of the arrival rates* for systems with nonzero switchover delay. Setting the frame size as a suitably increasing sublinear function of the queue lengths dynamically adapts the frame duration to the stochastic arrivals. This scheme gives infrequent reconfiguration decisions for large queue lengths, while enabling frequent reconfiguration

Güner D. Çelik and Eytan Modiano are with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, gcelik@mit.edu, modiano@mit.edu.

Sem Borst and Phil Whiting are with Bell Laboratories, Alcatel-Lucent, P.O. Box 94079, Murray Hill, NJ 07974, sem@research.bell-labs.com, pwhiting@research.bell-labs.com.

Sem Borst is also affiliated with the Department of Mathematics and Computer Science, Eindhoven University of Technology, The Netherlands.

This work was supported by NSF grants CNS-0626781 and CNS-0915988, and by ARO Muri grant number W911NF-08-1-0238.

[1] A queue of length $Q_i(t)$ at time t is rate stable if $\lim_{t\to\infty} Q_i(t)/t = 0$. This is a weaker notion of stability as compared to strong stability in Definition 1 which implies bounded first moments of a stationary measure.

for small queue lengths resulting in good delay performance.

This paper is organized as follows. In Section II we introduce the system model and provide an example showing that the Max-Weight policy is not throughput-optimal. We introduce the class of VFMW algorithms and prove its throughput-optimality in Section III. We present numerical examples regarding the application of the VFMW algorithm to DTN and Optical Networks in Section IV.

II. Queueing Model and Preliminaries

Consider a discrete time (slotted) system of N parallel queues served by M identical servers as in Fig. 1. Let $A_\ell(t)$, denote the number of arrivals to queue ℓ at time slot t. We assume that the processes $A_\ell(t)$ can take nonnegative real values, are independent of each other and are i.i.d. over time with $\mathbb{E}\{A_\ell(t)^2\} \leq A_{\max}^2$, $\mathbb{E}\{A_\ell(t)\} = \lambda_\ell \leq A_{\max}, \forall \ell = 1, \cdots, N$. We assume that it takes any server T_{ij} slots to switch from queue i to queue j. We define the system reconfiguration time as $T_r \triangleq \max\{T_{ij} | i, j \in \{1, ..., N\}\}$. The interference constraints in the system are given by the set of all possible activation vectors, $\mathcal{I} = \{\mathbf{I}^1, ..., \mathbf{I}^{|\mathcal{I}|}\}$, where $\mathcal{I}$ consists of vectors of at most M non zero entries. We include the zero vector $\mathbf{I} = \mathbf{0}$ in $\mathcal{I}$ for convenience. If at time slot t the activation vector $\mathbf{I}(t) = [D_1(t), D_2(t), ..., D_N(t)]' \in \mathcal{I}$ is used, then $\min\{D_\ell(t), Q_\ell(t)\}$ packets depart from queue ℓ. We assume there is a uniform departure rate bound, $\mu_{\max}$ over all schedules and queues: $D_\ell(t) \leq \mu_{\max}, \forall \ell, t$. Also let $\mu_{\min}$ be the minimum level of departure rate given to an *active* queue over all schedules. Finally, we assume that the queues are initially empty and that the arrivals take place after the departures in any given time slot.

Definition 1 (Strong Stability [5], [6]): The system is strongly stable under a given control policy π if

$$\limsup_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{\ell=1}^{N} \mathbb{E}_\pi [Q_\ell(t)] < \infty.$$

Note that for the case of integer arrival and service variables, this stability criterion implies the existence of a long-run stationary measure with bounded first moments [5].

Definition 2 (Stability Region [5], [6]): The stability region $\mathbf{\Lambda}$ is the closure of the set of all arrival rate vectors $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, ..., \lambda_N]'$ such that there exists a control policy that stabilizes the system under $\boldsymbol{\lambda}$.

A policy is said to be throughput-optimal if it stabilizes the system for all input rates strictly inside $\mathbf{\Lambda}$.

When $T_r = 0$, the stability region of this system, $\mathbf{\Lambda}^0$, consists of all arrival rate vectors $\boldsymbol{\lambda}$ in the convex hull of the vectors in $\mathcal{I}$ [5], i.e., $\mathbf{\Lambda}^0 = \{\boldsymbol{\lambda} | \boldsymbol{\lambda} \in Conv\{\mathcal{I}\}\}$. The Max-Weight algorithm applies the activation vector $\mathbf{I}^*(t) = \arg\max_{\mathbf{I} \in \mathcal{I}} \mathbf{Q}(t).\mathbf{I}$ in each time slot, where $.$ denotes the dot product. When $T_r > 0$, we lose service opportunities during the reconfiguration times. Therefore, the stability region of our system satisfies $\mathbf{\Lambda} \subseteq \mathbf{\Lambda}^0$. We will establish that $\mathbf{\Lambda} = \mathbf{\Lambda}^0$.

A. Instability of Ordinary Max-Weight Algorithm

In this section we show the instability of the Max-Weight policy for a system of 2 queues and a single server with i.i.d. Bernoulli arrivals with arrival probability $p < 1/2$. The set of available activation vectors is $\mathcal{I} = \{(0,0), (1,0), (0,1)\}$, and the switching delay is $T_r = 1$ slot. The stability region of this simple system is $\{p | p \leq 1/2\}$ and it is achieved by the policy that serves the queues until exhaustion. The Max-Weight policy decides to switch whenever the boundary $Q_1 = Q_2$ is crossed. By construction, there are an infinite number of service switches almost surely (a.s.) and we have

$$|Q_1(t) - Q_2(t)| \leq 3, \ \forall t. \tag{1}$$

***Lemma** 1: Max-Weight policy is not throughput-optimal. Furthermore, there exists an arrival rate $\hat{p} < 0.5$ such that both queues grow to infinity a.s. for all $p > \hat{p}$.*

We omit the proof for brevity but highlight the basic ideas. The probability of having no arrivals at the queue in service and 2 arrivals at the other queue during two consecutive time slots is $p^2(1-p)^2$ and such an arrival sequence leads to reconfiguration. Hence the fraction of time slots that the server spends switching is bounded from below by $1/3p^2(1-p)^2$, which converges to $1/48$ as p tends to $1/2$, while it must be less than $1 - 2p$ in order for the system to be stable.

For the second part of the lemma, let $t_k, k = 0, 1, 2, ...$ be the k^{th} reconfiguration epoch. It is easy to show by a drift analysis over the expected time to reconfiguration that

$$\mathbb{E}\left[Q_1(t_{k+1}) + Q_2(t_{k+1}) | \mathbf{Q}(t_k)\right] \geq Q_1(t_k) + Q_2(t_k) + \eta, \tag{2}$$

as long as $p > \hat{p} = 0.421$ where η is a fixed constant. From (2), it follows that the process $R_k \doteq \frac{1}{Q_1(t_k) + Q_2(t_k) + 1} \leq 1$ is a nonnegative supermartingale for $p > \hat{p}$:

$$\mathbb{E}\left[R_{k+1} | R_k\right] \leq R_k - \eta' R_k^2, \tag{3}$$

where η' is an appropriate constant. Therefore, $\lim_k R_k$ exists a.s. [1] and so does $\lim_k Q_1(t_k) + Q_2(t_k)$. Using a telescoping series argument over $k = 0, 1, 2, ...$, and Fatou's lemma gives $\liminf_k R_k = 0$ a.s. and since the limit exists we have $\lim_k Q_1(t_k) + Q_2(t_k) = \infty$ a.s. Therefore, using (1) and $|Q_\ell(t) - Q_\ell(t_k)| \leq 4$ for $t_k \leq t < t_{k+1}$, we have that the number of packets in both queues diverges to infinity a.s.

III. Variable Frame Based Max-Weight Policy

We propose a class of policies termed the Variable Frame Based Max-Weight (VFMW) policies that are throughput-optimal without requiring the arrival rate information. The VFMW policies operate over frames whose duration is dynamically changing based on queue states. Specifically, let t_k be the first slot of the kth frame, let $\mathbf{Q}(t_k)$ be the queue lengths at t_k and let $S(\mathbf{Q}(t_k)) \doteq \sum_i Q_i(t_k)$. The VFMW policy calculates the Max-Weight schedule with respect to $\mathbf{Q}(t_k)$ and applies this schedule during the frame, where the frame length is set as a sublinear and increasing function of $S(\mathbf{Q}(t_k))$. The VFMW policy is defined in detail in Algorithm 1.

Algorithm 1 VFMW ALGORITHM WITH FRAME LENGTH $\chi_k = T_r + F(S(\mathbf{Q}(t_k)))$:

1: Find the Max-Weight activation vector at time t_k, $\mathbf{I}^*(t_k)$:

$$\mathbf{I}^*(t_k) = \arg\max_{\mathbf{I}\in\mathcal{I}} \mathbf{Q}(t_k).\mathbf{I}$$

2: Invoke reconfiguration for the next T_r slots.

3: Apply $\mathbf{I}^*(t_k)$ for an interval of duration $F(S(\mathbf{Q}(t_k)))$ slots where $\chi_k \doteq T_r + F(S(\mathbf{Q}(t_k)))$, $F(.) > 0$ is a monotonically increasing function that satisfies

$$\lim_{y\to\infty} \frac{F(y)}{y} = 0.$$

4: Repeat above for the next frame starting at $t_{k+1} = t_k + \chi_k$.

The VFMW algorithm sets the frame length as a suitably increasing sublinear function of the queue sizes, which dynamically adapts the frame duration to the stochastic arrivals. For instance, $\chi_k = T_r + (\sum_i Q_i(t_k))^\alpha$ with $\alpha \in (0,1)$ satisfies the criteria for the frame duration. Under the VFMW policy the frequency of service reconfiguration is small when the queue sizes are large, limiting the fraction of time spent to switching. Note that this frequency should not be too small otherwise the system becomes unstable as it is subjected to a bad schedule for an extended period of time. Indeed, frame sizes linear in queue lengths do not guarantee stability in our framework. When the queue sizes are small, the VFMW policy gives frequent reconfiguration decisions, becoming more adaptive and providing a good delay performance.

***Theorem** 1: The VFMW policy stabilizes the system for all arrival rates $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}^0$ without requiring knowledge of $\boldsymbol{\lambda}$.*

An immediate corollary to this theorem is as follows:

***Corollary** 1:* $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}^0$.

The proof of Theorem 1 is given in the Appendix and is presented using the frame length function $\chi_k = T_r + (\sum_i Q_i(t_k))^\alpha$ for a fixed $\alpha \in (0,1)$ for ease of exposition. It is based on establishing a negative drift over the switching epochs t_k using a quadratic Lyapunov function, and then utilizing this result to establish stability of the overall system. It is novel in that the proof performs a drift analysis over a variable length interval whose duration is set as a function of the queue sizes. The basic intuition behind the proof is that if the queue sizes are large, the VFMW policy accumulates sufficient negative drift during the frame, which overcomes the cost accumulated during the reconfiguration time. Note that choosing the frame length as a *sublinear* function of the queue sizes is critical. This is because the VFMW algorithm uses the Max-Weight schedule corresponding to the beginning of the frame, which "loses weight" as the frame goes on. Therefore, one needs to make sure that the system is not subjected to this "light-weight" schedule for too long. In particular, frame lengths sublinear in queue sizes work, however, frame lengths that are linear in queue sizes do not guarantee stability.

We establish additional results when the arrival and departure processes have extra structure:

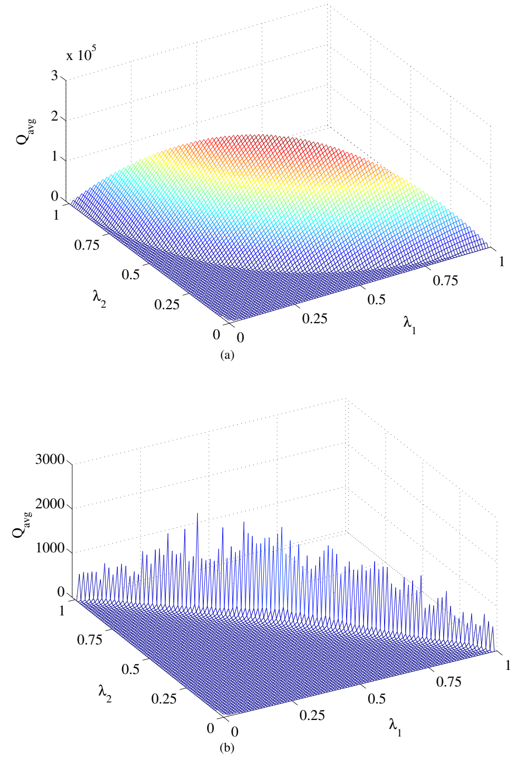

Fig. 2: The total average queue size for (a) the Max-Weight policy and (b) the VFMW policy for 2 queues and $T_r = 2$ slots.

***Corollary** 2: When the arrivals and service rates take integer values, and there is a positive probability of no arrivals to each queue, we have the following:*

- *the queues are all empty infinitely often with finite mean recurrence times,*
- *the queue sizes follow an irreducible and positive-recurrent Markov chain,*
- *the stationary measure of the Markov chain has bounded first moments.*

The proof of these results is omitted for brevity.

The VFMW algorithm has much lower computational complexity than the ordinary Max-Weight algorithm since it performs scheduling computation only once per frame. As first suggested in [7], by using out-of-date queue length information, the VFMW algorithm can be implemented to perform the computation of the next schedule during the current frame, without reducing the stability region. Therefore, letting C_{MW} denote the computational complexity of the ordinary Max-Weight algorithm per time slot, the VFMW algorithm has C_{MW}/χ computational complexity per time slot, where χ is the steady-state expected frame length.

IV. NUMERICAL RESULTS

We performed simulation experiments that determine average queue occupancy values for the VFMW policy, the ordinary Max-Weight policy and the Max-Weight policy with fixed frame sizes (MWFF). The average queue occupancy of queue ℓ over T_s slots is given by $\frac{1}{T_s}\sum_{t=1}^{T_s} Q_\ell(t)$ and the frame length for the VFMW policy is chosen as $\chi_k = T_r + (\sum_i Q_i(t_k))^{0.9}$. Through Little's law, the long-run packet-average delay in the system is equal to the time-average number of packets divided by the total arrival rate into the system. We used the 2-queue network described in Section II-A, except that we have different Bernoulli arrivals to each queue and that the switchover delay T_r is taken to be 2 or 20 in the experiments. In all the reported results, we have $(\lambda_1, \lambda_2) \in \mathbf{\Lambda}$ with 0.01 increments, where $\mathbf{\Lambda} = \{(\lambda_1, \lambda_2) | \lambda_1 + \lambda_2 \leq 1\}$. Furthermore, for each data point the simulation length was $T_s = 1,000,000$ slots.

Fig. 2 compares the stability regions of the VFMW and the Max-Weight policies when $T_r = 2$. This experiment can model single-hop optical networks where the reconfiguration time T_r is usually small. Fig. 2 (a) confirms that the plain Max-Weight policy is not throughput-optimal, and the points corresponding to the sudden jump in the plot represent the boundary of the region stabilized by the Max-Weight policy. Fig. 2 (b) shows that the VFMW policy has bounded queue sizes for all arrival rates inside the stability region $\mathbf{\Lambda}$.

Fig. 3 presents the delay as a function of throughput for the VFMW, Max-Weight and the MWFF (with frame sizes $T = 30$ and $T = 80$) policies along the main diagonal line. The switchover delay in this experiment, $T_r = 20$ slots, is relatively large, which could represent a DTN application such as mobile elements gathering data from sensors in a field. Fig. 3 confirms that the VFMW policy is throughput-optimal for this system and that the system quickly becomes unstable under the Max-Weight policy as the arrival rate is increased.

In Fig. 3, the MWFF policy with frame length $T = 30$ has similar delay performance to the Max-Weight policy for small arrival rates, however, under this policy the system becomes unstable around $\lambda_1 = \lambda_2 = 0.2$. Increasing the frame length improves the stability region of the MWFF policy at the expense of delay performance for small arrival rates. As opposed to fixed frame lengths, the VFMW policy dynamically adapts the frame length as a function of the queue states and stabilizes the system whenever possible, while providing a delay performance that is similar to that of the Max-Weight policy for small arrival rates.

V. Conclusions

We investigated the scheduling problem of multiple servers over parallel queues under arbitrary interference constraints and server switchover times. We showed that the Max-Weight scheduling algorithm is not throughput-optimal for such systems and we developed the class of Variable Frame Based Max-Weight (VFMW) algorithms that provide throughput-optimality without requiring the knowledge of the arrival rates. The VFMW algorithms persist with the Max-Weight schedule during an interval of duration dependent on the queue sizes, which dynamically adapts the frame sizes to stochastic arrivals and provides a good delay performance in addition to stability.

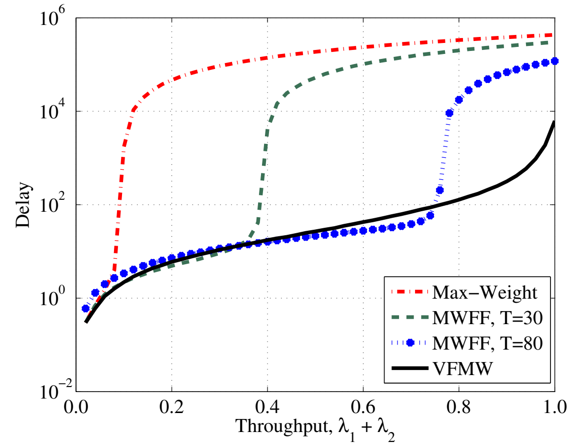

Fig. 3: Delay (total average queue size) vs the throughput under the Max-Weight, MWFF and the VFMW policies for 2 queues and $T_r = 20$ slots.

In the future, we intend to develop low-complexity distributed algorithms for systems with nonzero reconfiguration times that achieve a good performance. Finally, joint scheduling and routing in multihop networks with interference constraints and switchover times is another future direction.

Appendix-Proof of Theorem 1

Consider the following χ_k-step queue evolution expression:

$$Q_i(t_k+\chi_k) \leq \max\left\{Q_i(t_k) - \sum_{\tau=0}^{\chi_k-1} D_i(t_k+\tau), 0\right\} + \sum_{\tau=0}^{\chi_k-1} A_i(t_k+\tau).$$

To see this, note that if $\sum_{\tau=0}^{\chi_k-1} D_i(t_k+\tau)$, the total service opportunity given to queue i during the k^{th} frame, is smaller than $Q_i(t_k)$, then we have an equality. Otherwise, the first term is 0 and we have an inequality. This is because some of the arrivals during the frame might depart before the end of the frame. We first prove stability at the frame boundaries. Squaring both sides, using $\max(0,x)^2 \leq x^2, \forall x \in \Re$, and $D_i(t) \leq \mu_{\max}, \forall t$ we have

$$\begin{aligned} Q_i(t_k+\chi_k)^2 - Q_i(t_k)^2 &\leq \chi_k^2 \mu_{\max}^2 + \Big(\sum_{\tau=0}^{\chi_k-1} A_i(t_k+\tau)\Big)^2 \\ &- 2Q_i(t_k)\Big(\sum_{\tau=0}^{\chi_k-1} D_i(t_k+\tau) - \sum_{\tau=0}^{\chi_k-1} A_i(t_k+\tau)\Big). \quad (4) \end{aligned}$$

Define the quadratic Lyapunov function $L(\mathbf{Q}(t)) = \sum_{i=1}^{N} Q_i^2(t)$, and the χ_k-step conditional Lyapunov drift

$$\Delta_{\chi_k}(t_k) \triangleq \mathbb{E}\left[L(\mathbf{Q}(t_k+\chi_k)) - L(\mathbf{Q}(t_k)) \middle| \mathbf{Q}(t_k)\right].$$

Summing (4) over the queues, taking conditional expectation, using the assumption $\mathbb{E}[A_i(t)^2] \leq A_{\max}^2, \forall t$ (which also implies $\mathbb{E}[A_i(t_1)A_i(t_2)] \leq \sqrt{\mathbb{E}[A_i(t_1)]^2 \mathbb{E}[A_i(t_2)]^2} \leq A_{\max}^2$ for all t_1 and t_2), we have

$$\begin{aligned} \Delta_{\chi_k}(t) &\leq NB\chi_k^2 + 2\chi_k \sum_i Q_i(t_k)\lambda_i \\ &- 2\sum_i Q_i(t_k)\mathbb{E}\left[\sum_{\tau=0}^{\chi_k-1} D_i(t_k+\tau) \middle| \mathbf{Q}(t_k)\right] \quad (5) \end{aligned}$$

where $B = A_{\max}^2 + \mu_{\max}^2$ is a constant and we used the fact that the arrival processes are i.i.d. over time, independent of the queue lengths. Recalling that the system is idle for the first T_r slots of the frame, we have $\sum_i Q_i(t_k)\mathbb{E}\left[\sum_{\tau=0}^{\chi_k-1} D_i(t_k+\tau)\big|\mathbf{Q}(t_k)\right] = (\chi_k - T_r)\mathbf{Q}(t_k)\mathbf{.I}^*(t_k)$. Using this in (5) we have

$$\Delta_{\chi_k}(t_k) \leq NB\chi_k^2 + 2\chi_k \sum_i Q_i(t_k)\lambda_i - 2(\chi_k - T_r)\mathbf{Q}(t_k)\mathbf{.I}^*(t_k).$$

Now using the fact that for any arrival rate vector $\boldsymbol{\lambda}$ that is strictly inside $\boldsymbol{\Lambda}^0$, there exist real numbers $\beta^1, ..., \beta^{|\mathcal{I}|}$ such that $\beta^j \geq 0, \forall j \in 1, ..., |\mathcal{I}|$, $\sum_{j=1}^{|\mathcal{I}|}\beta^j = 1-\epsilon$ for some $\epsilon > 0$ and $\boldsymbol{\lambda} = \sum_{j=1}^{|\mathcal{I}|}\beta^j\mathbf{I}^j$ [3], we have

$$\begin{aligned}\Delta_{\chi_k}(t_k) \quad &\leq \quad NB\chi_k^2 + 2\chi_k\mathbf{Q}(t_k)\mathbf{.}\left(\sum_{j=1}^{|\mathcal{I}|}\beta^j\mathbf{I}^j\right) \\ &\quad - \quad 2(\chi_k - T_r)\mathbf{Q}(t_k)\mathbf{.I}^*(t_k) \\ &= \quad \chi_k\left(NB\chi_k - 2\left(\epsilon - \frac{T_r}{\chi_k}\right)\mathbf{Q}(t_k)\mathbf{.I}^*(t_k)\right),\end{aligned}$$

where we used the fact that $\mathbf{Q}(t_k)\mathbf{.I}^*(t_k) \geq \mathbf{Q}(t_k)\mathbf{.I}, \forall \mathbf{I} \in \mathcal{I}$. If $\chi_k = T_r + (\sum_i Q_i(t_k))^\alpha \leq \frac{T_r}{\epsilon}$ then $\Delta_{\chi_k}(t_k) \leq C_0$ where C_0 is a constant. Otherwise, there exists a small $\delta_1 > 0$ such that $\epsilon - \frac{T_r}{\chi_k} > \delta_1$. Hence, for $\sum_i Q_i(t_k) > \frac{T_r^{1/\alpha}(1-\epsilon)^{1/\alpha}}{\epsilon^{1/\alpha}}$, we use $\mathbf{Q}(t_k)\mathbf{.I}^*(t_k) \geq \frac{\mu_{\min}}{N}\sum_i Q_i(t_k)$ to have

$$\Delta_{\chi_k}(t_k) \leq \chi_k^2 NB - 2\chi_k\delta_1\frac{\mu_{\min}}{N}\sum_i Q_i(t_k).$$

Therefore, there exists a constant C such that

$$\Delta_{\chi_k}(t_k) \leq C - \delta\left(\sum_i Q_i(t_k)\right)^{1+\alpha}, \tag{6}$$

where $\delta \triangleq \delta_1\mu_{\min}/N$. Taking expectations with respect to $\mathbf{Q}(t_k)$, writing a similar expression over the frame boundaries $t_k, k \in \{0, 1, 2, ..., K\}$, summing them and telescoping these expressions leads to

$$L(\mathbf{Q}(t_K)) - L(\mathbf{Q}(0)) \leq KC - \delta\sum_{k=0}^{K-1}\mathbb{E}\left[\left(\sum_i Q_i(t_k)\right)^{1+\alpha}\right].$$

Using $L(\mathbf{Q}(t_K)) \geq 0$ and $L(\mathbf{Q}(\mathbf{0})) = 0$ we have

$$\frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}\left[\left(\sum_i Q_i(t_k)\right)^{1+\alpha}\right] \leq \frac{C}{\delta} < \infty.$$

This implies that

$$\limsup_{K\to\infty}\frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}\left[\left(\sum_i Q_i(t_k)\right)^{1+\alpha}\right] \leq \frac{C}{\delta} < \infty. \tag{7}$$

This establishes stability (as defined in Definition 1) at the frame boundaries $t_k, k \in \{0, 1, 2, ...\}$.

Now for any given $t \in (t_k, t_{k+1})$ we have

$$Q_i(t) \leq Q_i(t_k) + \sum_{\tau=0}^{\chi_k-1} A_i(t_k+\tau). \tag{8}$$

This gives

$$\sum_{\tau=0}^{\chi_k-1} Q_i(t_k+\tau) \leq \chi_k Q_i(t_k) + \sum_{\tau=0}^{\chi_k-1}\sum_{\tau_1=0}^{\chi_k-1} A_i(t_k+\tau_1). \tag{9}$$

Summing over the queues and taking conditional expectation we have for any $t \in (t_k, t_{k+1})$

$$\sum_{\tau=0}^{\chi_k-1}\sum_i \mathbb{E}\left[Q_i(t_k+\tau)|\mathbf{Q}(t_k)\right] \leq \chi_k\sum_i Q_i(t_k) + \chi_k^2\sum_i\lambda_i,$$

where we used the fact that arrival processes are i.i.d. and independent of the queue lengths. Recalling $\chi_k = T_r + (\sum_i Q_i(t_k))^\alpha$ with $0 < \alpha < 1$ we have

$$\begin{aligned}\sum_{\tau=0}^{\chi_k-1}\sum_i \mathbb{E}\left[Q_i(t_k+\tau)|\mathbf{Q}(t_k)\right] \leq &\left(\sum_i Q_i(t_k)\right)^{1+\alpha} + T_r\sum_i Q_i(t_k) \\ &+ \left(T_r + \left(\sum_i Q_i(t_k)\right)^\alpha\right)^2\sum_i\lambda_i.\end{aligned}$$

Now, for any given large T, let K_T be the number of frames up to and including T. We have

$$\begin{aligned}\sum_{t=0}^{T-1}\sum_i \mathbb{E}[Q_i(t)] \leq &\sum_{k=0}^{K_T-1}\mathbb{E}\left[\left(\sum_i Q_i(t_k)\right)^{1+\alpha} + T_r\sum_i Q_i(t_k)\right] \\ &+ \sum_{k=0}^{K_T-1}\mathbb{E}\left[\left(T_r + \left(\sum_i Q_i(t_k)\right)^\alpha\right)^2\right]\sum_i\lambda_i.\end{aligned}$$

Dividing both sides by T, using the fact that $T > K_T$ for any T, taking the $\limsup_T$ of both sides, using (7) and $0 < \alpha < 1$,

$$\limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{i=1}^{N}\mathbb{E}\left[Q_i(t)\right] < \infty. \tag{10}$$

Therefore, the system is stable.

References

[1] P. Billingsley, "Convergence of Probability Measures", Wiley, 1968.

[2] L. Blake and M. Long, "Antennas: Fundamentals, Design, Measurement," *SciTech*, 2009.

[3] A. Brzezinski and E. Modiano, "Dynamic Reconfiguration and Routing Algorithms for IP-over-WDM networks with Stochastic Traffic," *IEEE Journal of Lightwave Tech.*, vol. 23, no. 10, pp. 3188-3205, Oct. 2005

[4] G. D. Celik, L. B. Le, and E. Modiano, "Scheduling in Parallel Queues with Randomly Varying Connectivity and Switchover Delay," In *Proc. IEEE INFOCOM'11, Mini Conference*, Apr. 2011.

[5] M. J. Neely, "Stochastic Network Optimization with Application to Communication and Queueing Systems," Morgan and Claypool, 2010.

[6] M. J. Neely, E. Modiano, and C. E. Rohrs, "Dynamic Power Allocation and Routing for Time Varying Wireless Networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 1, pp. 89–103, Jan. 2005.

[7] M. J. Neely, E. Modiano, and C. E. Rohrs, "Tradeoffs in delay guarantees and computation complexity in nxn packet switches," I n *Proc. CISS'02*, Mar. 2002.

[8] D. Shah and D. J. Wischik, "Optimal Scheduling Algorithms for Input-Queued Switches," In *Proc. IEEE Infocom'06*, Mar. 2006.

[9] L. Tassiulas and A. Ephremides, "Stability Properties of Constrained Queueing Systems and Scheduling Policies for Maximum Throughput in Multihop Radio Networks," *IEEE Trans. Auto. Control*, vol. 37, no. 12, pp. 1936-1948, Dec. 1992.

[10] L. Tassiulas and A. Ephremides, "Dynamic Server Allocation to Parallel Queues with Randomly Varying Connectivity," *IEEE Trans. Infor. Theory*, vol. 39, no. 2, pp. 466-478, Mar. 1993.