Scheduling and Control in Wide-Area All-Photonic Networks

Nahid Saberi and Mark Coates
Department of Electrical and Computer Engineering
McGill University
Montreal, QC, Canada
E-mail: nahid.saberi@mail.mcgill.ca, coates@ece.mcgill.ca

Abstract—This paper introduces a feedback control system based on Smith’s principle for bandwidth reservation and scheduling of all-photonic networks with large propagation delay. The approach is also applicable to any single-hop communication network with significant signalling delay (such as satellite-TDMA systems). Scheduling in wide-area networks must be based on predictions of traffic demand and the resultant errors can lead to instability and unfairness. The feedback control system we propose reduces the effect of prediction errors, increasing the speed of the response to the sudden changes in traffic arrival rates and improving the fairness in the network through equalization of queue-lengths.

I. INTRODUCTION

The development of all-photonic switches that can switch in a fraction of a microsecond makes the deployment of all-photonic network cores attractive, because it eliminates the need for opto-electronic conversion inside the core. Unfortunately, the all-photonic switches are currently incapable of buffering packets, so scheduling of packet transmissions at edge switches must be carefully controlled. The scheduling task is much easier if the network architecture is of a simple nature that permits global network synchronization, thereby enabling the application of optical time-division multiplexing (OTDM) approaches for sharing link and switch capacity.

The authors of [1], [2] advocate an overlaid star topology in their proposal of an agile all-photonic network (AAPN) architecture. With an overlaid star topology, the scheduling problem consists of two tasks: (i) dividing the load amongst the stars, and (ii) scheduling the time-slots on each wavelength for each individual star. In this paper, we address the second task, assuming that the first is accomplished using a flow-based load balancing approach. The problem is the familiar one of designing a schedule for a one-hop communication system, as faced in an array of networking scenarios, particularly satellite TDMA systems (see [3], for example). The schedule determines when a source edge node has ownership of a given time-slot and is allowed to transmit to a specific destination edge node.

We focus on wide-area all-photonic networks that are designed to carry large volumes of traffic. In such networks it is more efficient to schedule frames, or blocks of slots, rather than allocate capacity on a slot-by-slot basis [4]. We address the problem of scheduling frames of fixed length, because this greatly simplifies protocol design, synchronization and signalling. The use of variable-length frames can reduce drop-rates or delay in certain network conditions [5], [6], but the evidence is less conclusive for high-volume, wide-area networks. In networks with substantial signaling delay, scheduling must be performed based on predictions of traffic demand in advance of traffic arrivals. If variable-length frames are used, the frame length must be calculated based on predicted load, and prediction errors eliminate most of the performance improvement that a variable-length frame makes possible. Even for fixed-length frame scheduling, the prediction becomes a source of error in resource allocation, potentially leading to instability and unfairness; this is the problem we attempt to rectify in this paper.

Contribution: In [7], [8] we described a scheduling algorithm for all-photonic networks that explicitly addresses scenarios of inadmissible demands and small offered loads, achieving zero rejection for admissible demands and providing fair allocation of free time-slots. Here we introduce a closed loop control architecture that is designed to interact with such an open loop scheduling mechanism. We employ Smith’s principle to design a linear feedback control architecture that compensates for the sources of error (prediction, rounding and rejection), resulting in a stable and fair system. We use a modified version of the Smith predictor to control the depletion rate of the virtual output queues (VOQs) at the edge nodes of the all-photonic network in order to achieve a desired queue level for each VOQ. We report the results of simulations conducted using OPNET Modeler [9] which indicate that the feedback control algorithm we propose successfully compensates for prediction errors and the effects of past rejection. It allocates spare capacity in a fair manner and responds to traffic variations.
faster than an open loop scheduling algorithm. Although the design we present focuses on concreteness on the case of an agile all-photonic network, it is applicable to any single-hop communication network with significant signalling delay (such as satellite-TDMA systems).

**Related Work:** Feedback congestion control has been examined from a control theoretic perspective by many authors, with the primary focus being controlling the rates at which sources inject best-effort traffic into a network in order to reduce the congestion at bottleneck queues whilst maintaining high utilization. In [10], Zhao et al. formulate the available bit-rate (ABR) resource allocation problem as a variant of the classical disturbance rejection problem. In a slightly different approach, Altman et al. pose the same task as a stochastic control problem, modelling the disturbance as an autoregressive process that is estimated by the controller using recursive least squares [11]. In [12], Hollot et al. analyze the combination of TCP and Active Queue Management (AQM) model from a control theoretic standpoint. There are many other examples of the application of linear control theory; see [13], [14] for surveys.

In the work most closely related to the controller design presented in this paper, Mascolo combines classical control theory and Smith’s principle to design a simple congestion control law that guarantees no packet loss and efficient use of bandwidth [15]. Bauer et al. propose a new class of time-variant Smith predictors using time-variant network delay models for forward and backward paths [16]. The proposed model features better tracking and faster rise and settling time. In both of these designs, the dynamic behavior of each network queue in response to data input is modelled as an integrator followed by a time delay. The use of Smith’s principle, which alleviates the stability difficulties of control systems with large delays, makes Mascolo’s design applicable to network paths with a wide range of propagation delays. Although the theoretical techniques we adopt in our design are similar to those used by Mascolo, the problem we address differs significantly. We assume that we have no control over arrival rates; instead we can adjust, through scheduling, the resources allocated within the network. This results in an inverted version of the standard congestion control problem: switch resources are controlled rather than source rates.

**Structure of the paper:** Section II provides an overview of the scheduling problem in agile all-photonic networks and discusses frame-based approaches. Section III illustrates how the frame-based scheduling algorithms act as feed-forward control systems. Section IV introduces Smith’s principle and describes the design of a modified Smith controller that interacts with the open loop scheduling algorithm to produce a stable resource allocation mechanism for AAPNs. Section V describes the simulation experiments we have executed to assess the performance of the control system and discusses the results. Finally, Section VI draws conclusions and indicates future research directions.

**II. AAPN OPEN LOOP SCHEDULING**

The agile all-photonic network (AAPN) architecture proposed in [1], [2] consists of edge nodes, where the opto-electronic conversion takes place, connected via selector/multiplexor devices to photonic core crossbar switches, which act independently of one another. Figure 1 shows an AAPN network with only one photonic switch. Each edge node maintains a separate VOQ for the traffic destined to each of the other edge nodes. These VOQs perform traffic aggregation, collecting together packets and transmitting them as a bundle in a single optical slot. Controllers, which receive requests from the edge nodes, derive the schedules and govern the operation of the photonic switches, are collocated with the photonic core nodes. Further details on the AAPN architecture and related research can be found in [17].

Before we can generate a schedule to allocate the available switch resources to the various edge nodes, the central controller must construct a predicted demand matrix $\mathbf{D}$, where $D_{ij}$ is the anticipated number of slots required by source node $i$ for destination $j$ during the fixed-length frame occurring $T$ seconds into the future. The signalling delay, $T$, exceeds the largest propagation delay between any edge node and the core. Many approaches can be adopted for performing this prediction, ranging from a naive predictor, where the estimate equals the number of slots required to accommodate the traffic that arrived in the current frame, to more elaborate techniques based on sophisticated traffic models such as those presented in [18]. The edge nodes can generate
their own estimates and send these to a central controller collocated with the photonic core switch, or they can send the raw measurements of traffic volumes and the central controller can form the estimates.

We define the following line sums of the demand matrix. The row sum, \( r_i = \sum_{j=1}^{N} D_{ij} \), is the total demand at source \( i \), and the column sum, \( c_j = \sum_{i=1}^{N} D_{ij} \), is the total demand for destination \( j \). It is important to achieve zero rejection if the demand is admissible. A demand matrix \( D \) is admissible for a frame of length \( L \) if

\[
\max\{\max_i \{r_i\}, \max_j \{c_j\}\} \leq L, \quad (1)
\]

Scheduling in cross-connect switches and star-topology networks has been investigated in depth for the past forty years, so there are naturally many approaches to generating the schedule once the demand matrix is available. The majority of frame-scheduling algorithms have focused on variable-length frames (for example, see [3], [19], [20] and the references therein). Algorithms designed for the variable-length frame scheduling problem can be applied to the case of fixed-length frames, but there must be adjustment when demand is low or inadmissible. When the predicted demand is insufficient to fill the schedule completely, we need a policy to divide the extra time slots amongst active connections (these extra slots alleviate the effect of potential underestimates in the demand matrix, which is merely a prediction of future traffic arrivals). On the other hand, when the demand is inadmissible, some of the predicted demand must be rejected. The choice of which requests to reject depends on whether the goal is to minimize the total amount of rejection or to achieve some form of fair rejection.

The authors of [21]–[24] have considered the problem of scheduling a frame of fixed length for star-coupled networks with tunable transmitters/receivers, but do not address the allocation of unused time slots or rejection of inadmissible demand. In more recent work, scheduling algorithms have been proposed for AAPNs with various objectives such as achieving max-min fairness of free slot allocation or rejection [7], minimizing total rejection [8], or maximizing the similarity of the modified traffic matrix to the original traffic matrix [25]. These techniques involve multiplicative modification of the original demand matrix \( D \) to form a modified demand matrix \( D' \) with the property that all of its line sums equal \( L \). After this modification, maximum cardinality matching algorithms such as EXACT [19], [20] are applied to \( D' \) to produce a full schedule (every time slot in the frame is allocated to a source-destination pair).

For concreteness, we focus on developing a feedback system for an architecture that employs the Fair Matching Algorithm (FMA) described in [8], but note that the algorithms proposed in [7], [25] could be employed. FMA represents a clamping procedure: resource demand (and allocation) is clamped to the full capacity available in a frame whether or not the full capacity has actually been requested (predicted) by the edge-nodes. FMA achieves zero rejection when the demand matrix is admissible. Its allocation of spare capacity (or rejection) is performed in a weighted max-min fair manner, where the weight associated with the connection between source \( i \) and destination \( j \) is \( \omega_{ij} = \frac{D_{ij}}{D'_{ij}} \), implying that subject to capacity constraints, extra slots are allocated (or rejection is assigned) in proportion to original demand. This is desirable if it is expected that the prediction error scales with the magnitude of the demand.

For Poisson arrivals and exponential distribution of the packet lengths, each \( VQ_{ij} \) can be approximated as a \( M/M/1 \) queueing system, with input rate \( \lambda_{ij} = D_{ij} \) (slots per frame) and output rate \( \mu_{ij} = D'_{ij} \) (slots per frame). Note that this model does not take into account prediction errors. Then the average number of packets in each \( VQ_{ij} \) is

\[
N_{ij} = \frac{\lambda_{ij}}{\mu_{ij} - \lambda_{ij}} = \frac{\rho_{ij}}{1 - \rho_{ij}}, \quad (2)
\]

where \( \rho_{ij} = \frac{\lambda_{ij}}{\mu_{ij}} \). When the FMA algorithm is applied, \( \rho_{ij} \) is the same for all \( VQs \) belonging to a bottleneck link as defined in [7], so the average numbers of packets in each of these queues are equal. FMA thus acts to equalize the waiting-time experienced by packets in the network, irrespective of arrival node or destination. See [8] for more details about FMA.

In common with the other open loop scheduling approaches discussed in this section, FMA does not record how many rejections occur for each connection in a frame if the demand is too great, acting as a memoryless system. In addition, there is a rounding procedure embedded in the scheduler, since time slots must be allocated in an integral fashion. The memoryless clamping and rounding act as further sources of imperfection in the resource allocation system. The remainder of this paper describes an architecture that addresses the error sources that exist when the resource allocation is a feed-forward system based on a clamping scheduler that receives predicted demands. We describe a congestion control feedback mechanism that interacts with an FMA scheduler to produce a stable bandwidth allocation system.

III. QUEUE CONTROL AND STABILITY

Scheduling algorithms that are solely based on the predicted demands arriving as signals from the edge nodes can be described as feed-forward or open loop systems. Figure 2 shows the schematic of the open loop system and its relationship with the FMA scheduler. In this section, we develop a control system model for resource
allocation in an agile all-photonic network. Initially we adopt a continuous-time model for the control system. Since scheduling is performed once per frame, we later sample the data with sampling period \( T_s \) to obtain a discrete-time system (see Section IV-C). Although the depletion rate varies within a frame due to the specific allocation of time slots, we model it as constant throughout the duration of a frame. Under this model, the depletion rate \( \text{dep}_{ij} \) has the following relationship with the number of time slots given to source-destination \((i, j)\) by the FMA scheduler:

\[
\text{dep}_{ij}(t + T) = \frac{D_{ij}(k)C}{L} k T_s \leq t \leq (k+1)T_s, \quad (3)
\]

The demand signal \( d_{ij}(t) \) is equal to the predicted arrivals:

\[
\hat{a}_{ij}(t) = d_{ij}(t) = \frac{D_{ij}(k)C}{L} k T_s \leq t \leq (k+1)T_s. \quad (4)
\]

Here \( D_{ij}(k) \) is the number of time slots demanded for a source-destination pair \((i, j)\) during one frame duration, \( T_s \), \( D_{ij} \) is the adjusted number of allocations based on the FMA algorithm, \( C \) is the line rate in bits-per-second, and \( L \) is the frame-length in time slots.

Demand matrix adjustment is performed by a clamping algorithm (e.g., FMA) which clamps the line sums of the demand matrix up or down to \( L \). FMA multiplies the predicted arrival rate \( \hat{a}_{ij} \) by a factor, \( x_{ij} = \frac{D_{ij}}{D_{ij}} \). Since this factor changes with the overall arrival rates the gain of the controller is tuned each frame.

If the system relies on only the feed-forward control then the effect of errors is not taken into account and this can lead to instability and unfairness. A closed loop control system is needed to achieve stability (i.e., steady state queue size variation), fairness, and faster response to traffic variations. Figure 3 shows a feedback control model for an agile all-photonic network with a central controller. Following the approach in [15], we use a simple integrator as the dynamic model for a VOQ. This control system has a similar structure to the system presented by Mascolo [15] with the major difference that we want to control the depletion rates from the queues in contrast to controlling the arrival rates to the queues. Let \( q_{ij}(t) \) be the length of the virtual queue of packets at edge node \( i \) destined to edge node \( j \). Let \( a_{ij} \) be the input rate to \( \text{VOQ}_{ij} \), and let \( \text{dep}_{ij} \) be the depletion rate of this queue. In the control model the length of each VOQ is compared with a reference signal, \( r_{ij}(t) \), and the difference is the input to the controller, which calculates the necessary adjustment rate, \( a_{ij}(t) \), for the predicted traffic arrival rate \( \hat{a}_{ij}(t) \).

Provided that the queue does not empty \( (q_{ij} > 0) \), the depletion rate can be expressed as the sum of the predicted arrival rate \( \hat{a}_{ij} \) and the adjustment due to the feedback \( ac_{ij} \), suitably delayed in time, i.e., \( \text{dep}_{ij}(t) = \hat{a}_{ij}(t - T) - ac_{ij}(t - T) \). Note that this adjustment is subtracted from the predicted arrival rate because it is proportional to the discrepancy between the desired state \( r_{ij} \) and the current state \( q_{ij} \). When this discrepancy is negative, the depletion rate should be increased to address the additional packets in the queue.

The queue length based on the flow conservation equation [15] with initial condition \( q_{ij}(0) = 0 \), is:

\[
q_{ij}(t) = \int_0^t [a_{ij}(\tau) - \text{dep}_{ij}(\tau)]d\tau, \quad (5)
\]

This model does not consider the case when the queue is empty: according to (5), departures from an empty queue result in a negative queue length. A more precise model for the network requires the inclusion of the operation "\( q_{ij} = \max(0, q_{ij}) \)" immediately prior to transmission of the feedback signal. Since this function can only be realized with a non-linear component we do not incorporate it, choosing instead to model the queues as always-occupied.

For this control system we aim to minimize the error between the queue length and the reference signal. The reference signal is interpreted as the desired queue length and may be calculated based on the state of the network. For example, if the desired state is equal queue lengths for all of the VOQs, then the reference signal should be the average of the VOQ lengths. As discussed in the previous section, this is also the effect of FMA, so this choice of reference signal aligns the feedback controller with the feed-forward controller. As an alternative, different reference signals could be provided to each source-destination pair, thereby inducing different average delays and providing the ability to provide different levels of service.

It is important to stress that because FMA is a clamping algorithm, the combination of the controller and
The Smith predictor, introduced by Smith in [26], makes modifications to the structure of the predictor, whereas others develop methods for tuning the system parameters. However, the original design requires a perfect representation of the “actual” plant, which results in insufficient robustness to modelling errors.

Many improvements and extensions have been proposed, particularly addressing the case of integrative or unstable plants [27]–[31]. Some authors introduce modifications to the structure of the predictor, whereas others develop methods for tuning the system parameters. Aström et al. described a modified Smith predictor in [27] that decouples the disturbance response from the setpoint response and improves the response time, and Liu et al. developed a technique for controlling unstable plants with long deadtime [31]. Our controller design is based on the modified Smith predictor by Matausek et al. [28].

### IV. AAPN Controller Design Based on the Smith Predictor

In this section we introduce a controller for an agile all-photonic network based on Smith’s principle. We first review the Smith predictor [26] and proposed modifications that improve robustness and address integral processes and unstable plants [27], [28].

#### A. Smith’s Principle

Instability is a common problem in delayed systems, since the addition of delays introduces extra phase lag, resulting in a less stable system (delay decreases the phase margin). Essentially, there is a substantial time lag before changes made by the controller have any effect. If the controller is not properly tuned to allow for this deadtime, it can overcompensate substantially. The Smith predictor, introduced by Smith in [26], makes the controller aware of the deadtime and adjusts its behavior so that it waits for the deadtime to elapse before expecting results. The effect of delay is eliminated from the characteristic equation of the closed loop control system. However, the original design requires a perfect representation of the “actual” plant, which results in insufficient robustness to modelling errors.

Many improvements and extensions have been proposed. Figure 4 shows the modified version of this controller for the AAPN network. The inputs to the system are $r_{ij}(t)$, $a(t)$ and $\dot{a}(t)$, and the output is $q_{ij}(t) = q_{ij}(t - T)$. We consider the arrival rate $a(t)$ and its prediction $\dot{a}(t)$ as disturbances. The reference signal, $r_{ij}$, represents the desired VOQ length. The response of the system is:

$$H_r(s) \triangleq \frac{Q_{p_{ij}}(s)}{R(s)} \bigg|_{a=0,\dot{a}=0} = \frac{x_{ij} K_r}{s + x_{ij} K_r} e^{-sT}, \quad (6)$$

$$H_d(s) \triangleq \frac{Q_{p_{ij}}(s)}{R(s)} \bigg|_{r=0,\dot{r}=0} = \frac{e^{-sT} [s - x_{ij} K_r (1 - e^{-2sT})]}{(s + x_{ij} K_r) (s + K_0 x_{ij} e^{-2sT})}, \quad (7)$$

$$\dot{H}_d(s) \triangleq \frac{Q_{p_{ij}}(s)}{R(s)} \bigg|_{r=0,\dot{r}=0} = \frac{x_{ij} e^{-2sT} [s - x_{ij} K_r (1 - e^{-2sT})]}{(s + x_{ij} K_r) (s + K_0 x_{ij} e^{-2sT})} = x_{ij} e^{-sT} H_d(s). \quad (8)$$

We strive to eliminate the steady-state effect of variations in the traffic arrival rate on the VOQ lengths. This corresponds to eliminating the load disturbance steady-state response and requires that $\lim_{s \to 0} H_d(s) = 0$, which is possible if $K_0 \neq 0$. Based on the final value theorem we have:

$$\lim_{t \to \infty} q_{ij}(t) = \lim_{s \to 0} R_{ij}(s) H_r(s) = r_{ij}. \quad (9)$$
From (7) it follows that the stability of the system depends on the roots of the characteristic equation

\[(s + x_{ij}K_r)(s + K_0x_{ij}e^{-sT}) = 0.\]  \hspace{1cm} (10)

The first term implies that \(x_{ij}K_r > 0\) must be satisfied. In order to find the roots of the second term and the ultimate gains for which the roots are located in the left-half plane we follow the approach of [28] and consider a hypothetical closed loop system with controllable parameter \(K_0\) and characteristic equation

\[(s + K_0x_{ij}e^{-2sT}).\] \hspace{1cm} (11)

For the analysis of this hypothetical system, (11) is rewritten in the form \(1 + W(s)\) where

\[W(s) = \frac{K_0x_{ij}}{s}e^{-2sT}.\]

The Nyquist criterion can then be applied to find the ultimate gain \(K_{0u}\), which indicates the supremum of \(K_0\) for which the hypothetical system is stable. \(K_{0u}\) is obtained by setting the phase margin \(\phi M\) and gain margin \(GM\) of the hypothetical system to zero. Note that these margins are not those of the actual control system. The construction of a control system described by (11) permits the application of standard techniques for finding the locations of the poles of a closed loop system. We have:

\[W(j\omega) = \frac{K_0x_{ij}}{j\omega}e^{-j2\omega} = \frac{K_0x_{ij}}{\omega}e^{-j(2\omega+\pi/2)}\]

\[\phi M = \pi + \angle W(j\omega) = \pi/2 - 2T\omega = 0\]

\[GM = -20\log |W(j\omega)| = -20\log \frac{K_0x_{ij}}{\omega} = 0.\]

Thus the ultimate gain \(K_{0u}\) is

\[K_{0u} = \frac{\pi}{4x_{ij}T},\] \hspace{1cm} (12)

and for all \(K_0 < K_{0u}\) the system is stable \((\phi M > 0)\).

C. Discrete-Time System Equations

Scheduling and signalling are only performed once per frame. In order to obtain the equivalent discrete-time system equations a simple approach is to design a digital control system using the Delta transform. The input to the plant is then converted to continuous form with zero-order-hold [32]. In the Delta transform approach we approximate the differential equation \(\frac{dy}{dt}\) with \(\frac{y(t+\Delta)-y(t)}{\Delta}\) [32]. For the control system presented in figure 5 the discrete time equations are approximated from the continuous form as:

\[d_{ij}(k) = \hat{a}_{ij}(k) - u_{rij}(k) + K_0q_{pij}(k) - K_0y_2(k),\]

\[y_1(k) = y_1(k-1) + x_{ij}(k-1)u_{rij}(k-1)T_s,\]

\[y_2(k) = y_1(k) - \frac{2T}{T_s},\]

\[u_{rij}(k) = K_r(y_1(k) - y_2(k) - q_{pij}(k) + r_{ij}(k)).\] \hspace{1cm} (13)

Defining \(\lambda \triangleq \frac{T_s}{T_s}\), we have:

\[u_{rij}(k) = K_r(\sum_{p=1}^{\lambda} x_{ij}(k-p)u_{rij}(k-p)T_s - q_{pij}(k) + r_{ij}(k)).\] \hspace{1cm} (14)

The rate adjustment thus depends, through the controller parameters \(K_0\) and \(K_r\), on the divergence of each queue length from the average queue length, \(r_{ij}\), as well as the
amount of the queue backlog $q_{pij}(k)$. The role of the Smith controller is to take into account the effect of rate adjustment on the queues during the $\lambda$ previous frames for which there is no feedback available.

D. Control Parameters and Adaptive Gain

The Smith controller can introduce undesirable fluctuations in the queue lengths by responding to small variations in the traffic arrivals. To avoid these fluctuations we set the controller such that it affects the system only when there is a queue length increase greater than $k$ time slots during one frame, where $k$ is some small threshold (in the simulations we use $k = 5$). This filters small, high-frequency fluctuations in the queue lengths so that they do not interfere with the controller.

The gain of the controller $K_r$ should be chosen such that the equivalent discrete-time system is stable. Standard digital control theory suggests that the sampling period should be at most half the time constant of the continuous system ($1/x_{ij}K_r$). Since our sampling time is the frame duration ($T_s$), we set:

$$K_r < \frac{1}{2x_{ij}T_s}. \quad (15)$$

Using a fixed controller gain can result in undesirable behavior. A small gain does not provide sufficiently fast response to traffic changes, but a large gain results in overreaction to minor fluctuations. An adaptive gain can provide a good compromise. We design the controller such that the gain $K_r$ adapts to the size of the queue variations:

$$K_r(k) = \min\{A \exp(C\Delta q_p), \frac{1}{2x_{ij}T_s}\}, \quad (16)$$

where $\Delta q_p = q_p(k) - q_p(k - 1)$. The choice of the constants $A$ and $C$ determines how fast the system reacts to traffic changes and whether there are residual oscillations. Simulations are used to determine a suitable range of values. Also to avoid overcompensation due to large control gains we use a fast-start slow-finish compensation procedure in which we reduce the gains of the controller by a factor of 0.05 two frames after activation of the Smith controller.

V. SIMULATION PERFORMANCE

In this section we report the results of simulations of the scheduling approaches performed using OPNET.
The Smith controller decreases the response time substantially, reducing the queue length of the heavy connection too fast, causing the other connections to starve and develop (relatively) large queue lengths, as shown in the middle panel. The bottom panel of Figure 6 compares the average divergences. During the initial period of heavy traffic, the fast draining of the long queue improves fairness, but later (around frame 100) the overcompensation results in a slight increase in average divergence. As outlined in Section IV-D, the use of fast-start slow-finish compensation can improve the performance of the controller. Figure 7 shows the performance of the Smith controller with this gain adjustment. The initial response is very fast but the queue of the heavy connection then drains slower (top panel), which smooths the effect on the queues of the non-heavy connections (middle panel). The average divergence has been improved compared to Figure 6 (bottom panel).

We performed simulations on a 16 edge-node star topology network. The links in the network have capacity 10 Gbps and the distance between each edge node and the photonic switch is 5 msec. A time slot is of length 10 μsec, and a frame has a fixed length of 1 msec (or 100 slots). Every experiment was run for a duration of 0.5 sec (equal to 500 frame durations). We investigate two traffic scenarios. In both scenarios, the average arrival rates to the VOQs are equal except for two periods (frames 20-32 and frames 130-132) during which the arrival rate of traffic from one source to one destination increases by a factor of 10. The two traffic scenarios are as follows. **Scenario A:** The arrival distribution of the data packets is Poisson with average arrival rate of 9 Gbps during the baseline periods; the packet size distribution is exponential with mean size of 1000 bits. **Scenario B:** Six Pareto ($\alpha = 1.9$) on-off sources are connected to each edge node. The mean on-period is 0.33 msec and mean off-period is 1.6 msec. The average rates are 9 Gbps during the on-period.

The top panel of Figure 6 compares the queue lengths of the VOQ carrying the heavy connection when using FMA with and without the Smith controller for the case of adaptive gains with $A = 63/x_{ij}$ and $C = 0.08$ in (16). The Smith controller decreases the response time substantially, reducing the queue length of the heavy connection much faster than the pure feed-forward controller derived by applying only FMA. This rapid response in the following frames empties the queue of the heavy connection too fast, causing the other connections to starve and develop (relatively) large queue lengths, as shown in the middle panel. The bottom panel of Figure 6 compares the average divergences. During the initial period of heavy traffic, the fast draining of the long queue improves fairness, but later (around frame 100) the overcompensation results in a slight increase in average divergence. As outlined in Section IV-D, the use of fast-start slow-finish compensation can improve the performance of the controller. Figure 7 shows the performance of the Smith controller with this gain adjustment. The
response to large changes in the queue length remains fast, but there is no starvation of the other queues, so the divergence remains low throughout the simulation.

Figure 8 examines the performance in response to bursty traffic as described in Scenario B, which is more unpredictable and thus poses a greater challenge for the Smith controller. The simulations indicate that the Smith controller still provides better drainage of the queues experiencing severe congestion. There is minimal negative effect on other queues or long-term fairness.

VI. CONCLUSION AND FUTURE WORK

This paper describes a feedback control system that compensates for scheduling errors due to mispredictions and rejection in single-hop communication networks with large propagation delays. The controller design is based on the Smith principle, which removes the destabilizing delays from the feedback loop by using a “loop cancellation” technique. We have shown through simulations that our controller reduces the response time to a sudden change in a queue length and imparts fairness by controlling the divergence from the average queue length. We are currently developing simple procedures for tuning the control system parameters that are insensitive to changes in traffic. In future work, we plan to explore the incorporation of methods from robust control to address the scenario where there are delay variations over time.

REFERENCES


