Optimal Control in Fluid Models of nxn Input-Queued Switches under Linear Fluid-Flow Costs

Yingdong Lu
IBM Research

Mark S. Squillante
IBM Research

Tonghoon Suk
IBM Research

ABSTRACT

We consider a fluid model of $n \times n$ input-queued switches with associated fluid-flow costs and derive an optimal scheduling control policy to an infinite horizon discounted control problem with a general linear objective function of fluid cost. Our optimal policy coincides with the $c\mu$-rule in certain parameter domains, but more generally, takes the form of the solution to a flow maximization problem. Computational experiments demonstrate the benefits of our optimal scheduling policy over variants of max-weight scheduling and the $c\mu$-rule.

1 INTRODUCTION

Input-queued switch architectures are widely used in modern computer and communication networks. The optimal scheduling control of these high-speed, low-latency switch networks is critical for our understanding of fundamental design and performance issues related to internet routers, cloud computing data centers, and high-performance computing. A large and rich literature exists around optimal scheduling in these computer and communication systems. This includes the extensive study of input-queued switches as an important mathematical model for a general class of optimal control problems of broad interest in both theory and practice.

Most of the previous research related to scheduling control in input-queued switches has focused on throughput optimality. In particular, the max-weight scheduling policy, first introduced in [24] for wireless networks and subsequently in [18] specifically for input-queued switches, is well-known to be throughput optimal. The question of delay-optimal scheduling control in such switch networks, however, is far less clear with much more limited results. This is due in large part because of the inherent difficulty of establishing delay (or equivalently, via Little's Law, queue length) optimality for these types of stochastic systems in general. Hence, previous research on optimal delay scheduling control in input-queued switches has focused on heavy-traffic and related asymptotic regimes; see, e.g., [1, 11, 20–22].

Such previous research includes showing that the max-weight scheduling policy is asymptotically optimal in heavy traffic for an objective function of the summation of the squares of the queue lengths with the assumption of complete resource pooling [23]. Max-weight scheduling was then shown to be optimal in heavy traffic for an objective function of the summation of the queue lengths under the assumption that all the ports are saturated [17]. This was subsequently extended to the case of incompletely saturated ports under the same objective function [16] and then to the case of general linear objective functions [13]. Nevertheless, beyond these and related recent results limited to the heavy-traffic regime, the question of delay-optimal scheduling control in input-queued switches remains open in general, as does the question of delay-optimal scheduling under more general objective functions.

In this paper, we seek to gain fundamental insights on optimal delay-cost scheduling in these stochastic systems by studying a fluid model of general $n \times n$ input-queued switches where each fluid flow has an associated cost. The objective of the corresponding optimal control problem is to determine the scheduling policy that minimizes the discounted summation over an infinite horizon of general linear cost functions of the fluid levels associated with each queue. Related research has been conducted in the queueing network literature; see, e.g., [2, 3, 7, 15]. In particular, similar problems have been studied within the context of fluid models of multiclass queueing networks [2, 3]. These previous studies take a classical optimal control approach based on exploiting Pontryagin's Maximum Principle [19], which itself only provides necessary conditions for optimality, to identify optimal policies. However, while this framework enables with relative ease the derivation of optimal policies for fluid models of basic queueing networks, the situation for input-queued switches is quite different and much more difficult. Specifically, the highly constrained structure of the input-queued switch networks requires us to pay special attention to the feasibility of the optimal control problem.

To address these issues, we implicitly move the capacity constraint into the objective and identify the appropriate Lagrangian multiplier through carefully designed search algorithms. Then, at any fluid level, we establish that the optimal scheduling policy is obtained through a solution to a flow maximization problem, which is also shown to be throughput optimal. Our optimal policy coincides with the $c\mu$-rule in certain parameter domains. These theoretical results reflect the high complexity nature of input-queued switches, and are expected to be of interest more broadly than input-queued switch networks and more broadly than related classes of fluid models of stochastic networks with constraints.

We observe important differences in the decisions made under our optimal scheduling control policy in comparison with those made under a cost-weighted variant of the max-weight scheduling policy and the $c\mu$-rule within the fluid model of general $n \times n$ input-queued switches. It is important to emphasize that our goal is to determine the optimal solution of the corresponding fluid control problem, which is at the core of the important scheduling-decision differences between our optimal policy and those of the other scheduling policies. Although we show that our flow maximization solution coincides with the $c\mu$-rule in some regions of the decision space, we also show that the $c\mu$-rule is not stable under

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
Performance 2020, November 02–06, 2020, Milan, Italy
© 2020 Association for Computing Machinery.
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

certain arrival rates and thus it cannot in general be the optimal scheduling policy. In contrast to the max-weight scheduling policy which does not use any arrival rate information, we show that the optimal policy from our flow maximization solution for the $n \times n$ input-queued switch fluid control problem can depend in general on the arrival rates, which is consistent with known results established for the original (non-fluid limit) 2×2 input-queued switch where the optimal policy takes into account the arrival processes in some regions of the decision space [12]. The cost-weighted max-weight scheduling policy has been shown to exhibit optimal queue-length scaling in the heavy traffic regime [13], suggesting that the importance of arrival-process information on the queue-length scaling of the optimal scheduling control policy tends to diminish asymptotically as the traffic intensity increases.

To further investigate these important differences, we conduct fluid-model computational experiments with our optimal scheduling policy, the max-weight scheduling policy, and the $c\mu$-rule to gain additional fundamental insights on various important theoretical issues with respect to optimal scheduling control in input-queued switch networks. In comparisons with the max-weight scheduling policy, we find that our optimal scheduling control policy provides improvements of at least 10% in most of the experiments, sometimes rendering improvements of more than 50%. Moreover, the improvements of our optimal policy over max-weight scheduling grow as the throughput increases. With respect to the $c\mu$-rule, we find that the comparisons with our optimal scheduling control policy fall into three different cases: (1) The $c\mu$-rule coincides with the optimal policy, and thus is fluid-cost optimal; (2) The $c\mu$-rule is unstable (not throughput optimal), and obviously not fluid-cost optimal; (3) The $c\mu$-rule is stable, but not fluid-cost optimal. Moreover, the greatest improvements observed for our optimal policy over stable $c\mu$-rule instances represent relative performance gaps of more than 70%.

The remainder of this paper is organized as follows. Section 2 presents our mathematical models, for both stochastic processes of input-queued switch networks and their mean-field limits, together with our formulation of the optimal scheduling control problems of interest. Section 3 then provides our analysis and results for optimal scheduling control and related theoretical properties. The results of computational experiments are presented in Section 4, followed by concluding remarks. We refer the reader to [14] for our proofs and additional technical details and theoretical results.

2 MATHEMATICAL MODELS

2.1 Technical Preliminaries

Let $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}^+$, $\mathbb{Z}$, $\mathbb{Z}_+$, and $\mathbb{Z}^+$ respectively denote the sets of real numbers, non-negative real numbers, positive real numbers, integers, non-negative integers, and positive integers. For positive integer $n \in \mathbb{Z}^+$, we define $[n] := \{1, 2, \ldots, n\}$ to be the set of all positive integers less than or equal to n. The blackboard bold typefaces is used for general sets, e.g., $\mathbb{I}$ and $\mathbb{J}$. When the set $\mathbb{I}$ is finite, we represent its cardinality by $|\mathbb{I}|$; e.g., we have $|[n]| = n$ for $n \in \mathbb{Z}^+$.

We use the bold font to represent vectors, matrices, and real-valued functions on a finite set. The function $\boldsymbol{\mu} : \mathbb{I} \to \mathbb{R}$, defined on the finite set $\mathbb{I}$, can be considered as an $|\mathbb{I}|$-dimensional vector $\boldsymbol{\mu} = [\mu(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{I}]$, where $\mu(\boldsymbol{s})$ is the value of $\boldsymbol{\mu}$ at $\boldsymbol{s}$. We denote by $\mathbb{R}^{\mathbb{I}}$ the set of all real-valued functions on $\mathbb{I}$. For finite sets $\mathbb{I}$ and $\mathbb{J}$, $\mathbb{R}^{\mathbb{I}\times\mathbb{J}}$ is the set of all real-valued functions from $\mathbb{I} \times \mathbb{J}$ in which an element $\boldsymbol{A}$ can also be represented by the matrix $\boldsymbol{A} = [A(\boldsymbol{s}, \boldsymbol{\rho}) : \boldsymbol{s} \in \mathbb{I}, \boldsymbol{\rho} \in \mathbb{J}]$, where $A(\boldsymbol{s}, \boldsymbol{\rho})$ is the value of the function $\boldsymbol{A}$ at $(\boldsymbol{s}, \boldsymbol{\rho}) \in \mathbb{I} \times \mathbb{J}$.

For $\boldsymbol{A} \in \mathbb{R}^{\mathbb{I}\times\mathbb{J}}$, $\boldsymbol{\eta} \in \mathbb{R}^{\mathbb{J}}$, and $\boldsymbol{\mu} \in \mathbb{R}^{\mathbb{I}}$, we respectively define $\boldsymbol{\mu A} \in \mathbb{R}^{\mathbb{J}}$, $\boldsymbol{Aq} \in \mathbb{R}^{\mathbb{I}}$, and $\mu A\eta \in \mathbb{R}$ by $(\mu A)(\rho) := \sum_{\boldsymbol{s}\in\mathbb{I}} \mu(\boldsymbol{s})A(\boldsymbol{s}, \boldsymbol{\rho})$, $(A\eta)(s) := \sum_{\boldsymbol{\rho}\in\mathbb{J}} A(\boldsymbol{s}, \boldsymbol{\rho})\eta(\boldsymbol{\rho})$, $\mu A\eta := \sum_{\boldsymbol{s}\in\mathbb{I}} \sum_{\boldsymbol{\rho}\in\mathbb{J}} \mu(\boldsymbol{s})A(\boldsymbol{s}, \boldsymbol{\rho})\eta(\boldsymbol{\rho})$, which is similar to matrix-vector multiplication. For $\boldsymbol{w}, \boldsymbol{\mu} \in \mathbb{R}^{\mathbb{I}}$, we also define $\boldsymbol{w} \cdot \boldsymbol{\mu} \in \mathbb{R}$ by $\boldsymbol{w} \cdot \boldsymbol{\mu} := \sum_{\boldsymbol{s}\in\mathbb{I}} w(\boldsymbol{s})\mu(\boldsymbol{s})$, which is the same as the inner-product of two vectors. We denote the 1-norm of a vector by $\|\cdot\|_1$, namely for $\boldsymbol{\mu} \in \mathbb{R}^{\mathbb{I}}$, $\|\boldsymbol{\mu}\|_1 := \sum_{\boldsymbol{s}\in\mathbb{I}} |\mu(\boldsymbol{s})|$. Finally, we use the sans serif font for random variables and use the bold sans serif font for random vectors, e.g., Q and $\boldsymbol{\mathsf{Q}}$, respectively.

2.2 Stochastic Models

The input-queued switch of interest consists of n input ports and n output ports. For each pair $(i, j) \in \mathbb{J} := [n] \times [n]$, packets that need to be transmitted from the i-th input port to the j-th output port are stored in a queue indexed by (i, j). We describe below how the number of packets in a queue (queue length) evolves over time. Time is slotted by nonnegative integers and the length of queue $\boldsymbol{\rho} \in \mathbb{J}$ at the beginning of the t-th time slot is denoted by $\mathsf{Q}_t(\boldsymbol{\rho})$.

External packets arrive at each queue according to an exogenous stochastic process. Let $\mathcal{A}_t(\boldsymbol{\rho}) \in \mathbb{Z}_+$ represent the number of arrivals to queue $\boldsymbol{\rho} \in \mathbb{J}$ until time t. Assume that $\{\mathcal{A}_{t+1}(\boldsymbol{\rho}) - \mathcal{A}_t(\boldsymbol{\rho}) : t \in \mathbb{Z}_+,\ \boldsymbol{\rho} \in \mathbb{J}\}$ are independent random variables and that, for fixed $\boldsymbol{\rho} \in \mathbb{J}$, $\{\mathcal{A}_{t+1}(\boldsymbol{\rho}) - \mathcal{A}_t(\boldsymbol{\rho}) : t \in \mathbb{Z}_+\}$ are identically distributed with $\mathbb{E}[\mathcal{A}_{t+1}(\boldsymbol{\rho}) - \mathcal{A}_t(\boldsymbol{\rho})] =: \lambda(\boldsymbol{\rho})$. We refer to the $|\mathbb{J}|$-dimensional vector $\boldsymbol{\lambda} \in [0, 1]^{\mathbb{J}}$ as the arrival rate vector and assume $\boldsymbol{\lambda}$ lies in the interior of the *capacity region*: $\{\boldsymbol{\lambda} \in [0, 1]^{\mathbb{J}}, \sum_i \lambda_{ij} < 1, \sum_j \lambda_{ij} < 1\}$.

During each time slot, packets in the queues can be simultaneously transmitted (or departed from the queues) subject to: (1) At most one packet can be transmitted from an input port; (2) At most one packet can be transmitted to an output port. Hence, we denote the departure of packets from the queues during a time slot by an n^2-dimensional binary vector $\boldsymbol{s} = [s(\boldsymbol{\rho}) : \boldsymbol{\rho} \in \mathbb{J}]$ such that $s(\boldsymbol{\rho}) = 1$ if a packet in queue $\boldsymbol{\rho}$ departs from the queue, and $s(\boldsymbol{\rho}) = 0$ otherwise. We refer to such $\boldsymbol{s}$ as a *basic schedule*, and let $\mathbb{I}$ denote the set of all basic schedules: $\mathbb{I} := \{\boldsymbol{s} \in \{0, 1\}^{\mathbb{J}} : \sum_{i\in[n]} s(i, j) \le 1, \sum_{j\in[n]} s(i, j) \le 1, \forall i, j \in [n]\}$. Note that $\mathbb{I}$ contains the empty basic schedule $\boldsymbol{s}$, such that $s(i, j) = 0$ for all $(i, j) \in \mathbb{J}$.

For $\boldsymbol{s} \in \mathbb{I}$, let $\mathcal{D}_t(\boldsymbol{s})$ denote the cumulative number of time slots devoted to basic schedule $\boldsymbol{s}$ until time t. We therefore have $\|\boldsymbol{\mathcal{D}}_t\|_1 = \sum_{\boldsymbol{s}\in\mathbb{I}} \mathcal{D}_t(\boldsymbol{s}) = t$ and $\|\boldsymbol{\mathcal{D}}_{t+1}\|_1 - \|\boldsymbol{\mathcal{D}}_t\|_1 = 1$ for every $t \in \mathbb{Z}_+$. From the description of arrivals and departures, we can see that $\boldsymbol{\mathsf{Q}}_t$ evolves according to the following dynamics $\boldsymbol{\mathsf{Q}}_t = \boldsymbol{\mathsf{Q}}_0 + \boldsymbol{\mathcal{A}}_t - \boldsymbol{\mathcal{D}}_t\boldsymbol{A}$, where $\boldsymbol{\mathsf{Q}}_0 = [\mathsf{Q}_0(\boldsymbol{\rho}) : \boldsymbol{\rho} \in \mathbb{J}]$ is the initial queue lengths and $\boldsymbol{A} \in \{0, 1\}^{\mathbb{I}\times\mathbb{J}}$ is the schedule-queue adjacency matrix such that $A(\boldsymbol{s}, \boldsymbol{\rho}) = s(\boldsymbol{\rho})$ for $\boldsymbol{s} \in \mathbb{I}$ and $\boldsymbol{\rho} \in \mathbb{J}$. We refer to a stochastic process $\{(\boldsymbol{\mathsf{Q}}_t, \boldsymbol{\mathcal{A}}_t, \boldsymbol{\mathcal{D}}_t) \in \mathbb{Z}_+^{\mathbb{J}} \times \mathbb{Z}_+^{\mathbb{J}} \times \mathbb{Z}_+^{\mathbb{I}} : t \in \mathbb{Z}_+\}$ that satisfies the above dynamics as a *discrete-time stochastic model for input-queued switches* with the (random) initial state $\boldsymbol{\mathsf{Q}}_0 \in \mathbb{Z}_+^{\mathbb{J}}$.

2.3 Fluid Models

This section introduces a deterministic process that represents our fluid models for input-queued switches, describes the scaled processes of the original stochastic process, and relates them to our fluid models. The basic set up and ideas can be found in the research literature on fluid limit models, especially the papers of Dai [10] and Dai and Prabhakar [9]. The key concepts concern the tightness and the measures of stochastic processes, which leads to the convergence of the subsequences of the scaled processes.

Definition 2.1. An absolutely continuous deterministic process $\{(\boldsymbol{q}_t, \boldsymbol{\delta}_t) \in \mathbb{R}^{\mathbb{J}} \times \mathbb{R}^{\mathbb{I}} : t \in \mathbb{R}_+\}$ is called a *(input-queued switch) fluid model* with initial state $\boldsymbol{q}_0 \in \mathbb{R}_+^{\mathbb{J}}$ and arrival rates $\boldsymbol{\lambda} \in [0,1]^{\mathbb{J}}$ if the following conditions hold: **(FM1)** $\boldsymbol{q}_t = \boldsymbol{q}_0 + \boldsymbol{\lambda} t - \boldsymbol{\delta}_t \boldsymbol{A}$ for $t \in \mathbb{R}_+$; **(FM2)** $\boldsymbol{q}_t \geq 0$ for $t \in \mathbb{R}_+$; **(FM3)** $\sum_{s \in \mathbb{I}} \delta_t(s) = t$ (i.e., $\|\boldsymbol{\delta}_t\|_1 = t$) and $\boldsymbol{\delta}_t \geq \boldsymbol{0}$ for $t \in \mathbb{R}_+$; **(FM4)** For any $\boldsymbol{s} \in \mathbb{I}$, $\delta_t(s)$ is non-decreasing with respect to t. Furthermore, a deterministic process $\{\boldsymbol{\mu}_t \in \mathbb{R}_+ : t \in \mathbb{R}_+\}$ is called an *(fluid-level) admissible policy* for the input-queued switch if and only if there exists a fluid model $(\boldsymbol{q}_t, \boldsymbol{\delta}_t)$ such that $\boldsymbol{\mu}_t = \dot{\boldsymbol{\delta}}_t$ for all $t \in \mathbb{R}_+$ at which $\dot{\boldsymbol{\delta}}_t$ exists.

Note that, since $(\boldsymbol{q}_t, \boldsymbol{\delta}_t)$ is absolutely continuous, $\dot{\boldsymbol{q}}_t$ and $\dot{\boldsymbol{\delta}}_t$ exist at almost every $t \in \mathbb{R}_+$. The following proposition introduces convenient alternative criteria for a fluid-level admissible policy.

Proposition 2.2. *Fix $\boldsymbol{q}_0 \in \mathbb{R}_+^{\mathbb{J}}$ and $\boldsymbol{\lambda} \in [0,1]^{\mathbb{J}}$. Let $\{\boldsymbol{\mu}_t \in \mathbb{R}_+^{\mathbb{I}} : t \in \mathbb{R}_+\}$ be an integrable deterministic process and $\{\boldsymbol{q}_t \in \mathbb{R}^{\mathbb{J}} : t \in \mathbb{R}_+\}$ a process satisfying $\dot{\boldsymbol{q}}_t = \boldsymbol{\lambda} - \boldsymbol{\mu}_t \boldsymbol{A}$ with initial state $\boldsymbol{q}_0$. Then, the following statements are equivalent: **(AP1)** $\boldsymbol{\mu}_t$ is a fluid-level admissible policy; **(AP2)** $\|\boldsymbol{\mu}_t\|_1 = 1$ and $\boldsymbol{q}_t \geq 0$ for all $t \in \mathbb{R}_+$; **(AP3)** $\|\boldsymbol{\mu}_t\|_1 = 1$ and $\boldsymbol{\mu}_t \in \mathbb{U}(\boldsymbol{q}_t)$ for all $t \in \mathbb{R}_+$, where $\mathbb{U}(\boldsymbol{q}) := \{\boldsymbol{\mu} \in [0,1]^{\mathbb{I}} : (\mu A)(\boldsymbol{\rho}) \leq \lambda(\boldsymbol{\rho}) \text{ if } q(\boldsymbol{\rho}) = 0\}$. In this case, $(\boldsymbol{q}_t, \boldsymbol{\delta}_t := \int_0^t \boldsymbol{\mu}_{t'} dt')$ is the fluid model associated with the fluid-level admissible policy $\boldsymbol{\mu}_t$.*

2.3.1 Scaled Queueing Processes. Fix index $r \in \mathbb{Z}^+$ and then let $\{(\boldsymbol{Q}_t^r, \boldsymbol{\mathcal{A}}_t^r, \boldsymbol{\mathcal{D}}_t^r) : t \in \mathbb{Z}_+\}$ be a discrete-time stochastic model with initial state $\boldsymbol{Q}^r$ as described in Section 2.2. We extend this discrete-time process to a continuous-time process by defining $\boldsymbol{\mathcal{A}}_t^r := (t - \lfloor t \rfloor)(\boldsymbol{\mathcal{A}}_{\lfloor t \rfloor+1}^r - \boldsymbol{\mathcal{A}}_{\lfloor t \rfloor}^r) + \boldsymbol{\mathcal{A}}_{\lfloor t \rfloor}^r$, $\boldsymbol{\mathcal{D}}_t^r := (t - \lfloor t \rfloor)(\boldsymbol{\mathcal{D}}_{\lfloor t \rfloor+1}^r - \boldsymbol{\mathcal{D}}_{\lfloor t \rfloor}^r) + \boldsymbol{\mathcal{D}}_{\lfloor t \rfloor}^r$, $\boldsymbol{Q}_t^r := (t - \lfloor t \rfloor)(\boldsymbol{Q}_{\lfloor t \rfloor+1}^r - \boldsymbol{Q}_{\lfloor t \rfloor}^r) + \boldsymbol{Q}_{\lfloor t \rfloor}^r = \boldsymbol{Q}^r + \boldsymbol{\mathcal{A}}_t^r - \boldsymbol{\mathcal{D}}_t^r \boldsymbol{A}$, where $\lfloor t \rfloor$ is the largest integer less than or equal to t.

Remark. *Processes $Q_t^r(\boldsymbol{\rho})$, $\mathcal{A}_t^r(\boldsymbol{\rho})$ and $\mathcal{D}_t^r(\boldsymbol{s})$ are random functions, and every sample path for $(\boldsymbol{Q}_t^r, \boldsymbol{\mathcal{A}}_t^r, \boldsymbol{\mathcal{D}}_t^r)$ is continuous. We use the notation ω^r to explicitly denote the dependency on the randomness in the r-th system and the notation $\boldsymbol{\omega} = [\omega^r : r \in \mathbb{Z}^+]$ to denote the overall randomness. For example, $Q_t^r(\rho; \boldsymbol{\omega}) = Q_t^r(\rho; \omega^r)$ and $\boldsymbol{Q}_t^r(\boldsymbol{\omega}) = \boldsymbol{Q}_t^r(\omega^r)$.*

For randomness $\boldsymbol{\omega}$, the scaled r-th system is defined by ($\hat{\boldsymbol{Q}}_t^r(\boldsymbol{\omega})$, $\hat{\boldsymbol{\mathcal{A}}}_t^r(\boldsymbol{\omega})$, $\hat{\boldsymbol{\mathcal{D}}}_t^r(\boldsymbol{\omega})$) := ($r^{-1}\boldsymbol{Q}_{rt}^r(\boldsymbol{\omega})$, $r^{-1}\boldsymbol{\mathcal{A}}_{rt}^r(\boldsymbol{\omega})$, $r^{-1}\boldsymbol{\mathcal{D}}_{rt}^r(\boldsymbol{\omega})$). We assume that the initial state of the r-th system satisfies $r^{-1}\boldsymbol{Q}_0^r \Rightarrow \boldsymbol{q}_0$, as $r \to \infty$, for a (deterministic) point $\boldsymbol{q}_0 \in \mathbb{R}_+^{\mathbb{J}}$, where the convergence is understood to be convergence in distribution.

2.3.2 Tightness and Convergence. For a fixed sample path $\boldsymbol{\omega}$, from the above definitions of the discrete-time departure process and the extension to its continuous-time process, we have $\hat{\mathcal{D}}_0(\boldsymbol{\rho}; \boldsymbol{\omega}) = 0$ and $\hat{\mathcal{D}}_t(\boldsymbol{\rho}; \boldsymbol{\omega}) \leq \|\hat{\boldsymbol{\mathcal{D}}}_t(\boldsymbol{\omega})\|_1 = t$ so that $\hat{\mathcal{D}}_t^r(\boldsymbol{\rho}; \boldsymbol{\omega}) - \hat{\mathcal{D}}_{t'}^r(\boldsymbol{\rho}; \boldsymbol{\omega}) \leq (t - t')$, for any $r > 0$ and $t \geq t' \geq 0$. This implies the tightness of the process $\hat{\mathcal{D}}_t^r$; see, e.g., [4]. Meanwhile, from the functional strong law of large numbers, we have $\lim_{r \to \infty} \sup_{0 \leq t \leq T} |\hat{\mathcal{A}}_t^r(\boldsymbol{\rho}; \boldsymbol{\omega}) - \lambda(\boldsymbol{\rho}) t| = 0$ almost surely (see, e.g., [6]). We therefore have that, almost surely, for each sample path $\boldsymbol{\omega}$ and any sequence $\{r_k\}$ such that $\lim_{k \to \infty} r_k = \infty$, there exists a subsequence $\{r_{k_l}\}$ and deterministic process $(\boldsymbol{q}_t, \boldsymbol{\delta}_t)$, which is a fluid model in Definition 2.1, such that $(\hat{\boldsymbol{Q}}_t^{r_{k_l}}(\boldsymbol{\omega}), \hat{\boldsymbol{\mathcal{D}}}_t^{r_{k_l}}(\boldsymbol{\omega}))$ converges to $(\boldsymbol{q}_t, \boldsymbol{\delta}_t)$ uniformly on all compact sets as $l \to \infty$.

Remark. *The conditions **FM1** to **FM4** are necessary conditions for all the fluid limits, and they do not uniquely determine a fluid limit, even under a fixed admissible scheduling policy. Such a lack of uniqueness for the fluid limits and its implications for queueing networks are discussed at length in [5]. For certain special cases, with extra conditions on the policies, fluid limits can be shown to be unique; see, e.g., [22] for input-queued switches. Our interest, however, is in solving optimal scheduling control problems within the context of the fluid models. With conditions such as **FM1** and **FM4**, fluid limit results are generally established for converging subsequences; similar results can be found in [10] for queueing networks.*

2.4 Fluid Model Optimal Control Problems

We now formulate the optimal scheduling control problem of interest within the context of the fluid models of input-queue switches. To this end, we define as follows the total discounted delay cost over the entire time horizon under a fluid-level admissible policy $\{\boldsymbol{\mu}_t : t \in \mathbb{R}_+\}$ with initial state $\boldsymbol{q}_0$: $c(\boldsymbol{\mu}_t; \boldsymbol{q}_0) := \int_0^\infty e^{-\beta t} \boldsymbol{c} \cdot \boldsymbol{q}_t dt$, where $\boldsymbol{q}_t$ is the deterministic function defined in **FM1** with $\boldsymbol{\delta}_t := \int_0^t \boldsymbol{\mu}_s ds$ and initial state $\boldsymbol{q}_0$, β is the discount factor, and $\boldsymbol{c} \in (\mathbb{R}^+)^{\mathbb{J}}$ is the vector of cost coefficients. Specifically, we seek to find a fluid-level admissible scheduling policy with the following objective:

Minimize $c(\boldsymbol{\mu}_t; \boldsymbol{q}_0)$ over all admissible policies $\{\boldsymbol{\mu}_t : t \in \mathbb{R}_+\}$.

From **(AP2)** in Proposition 2.2, this control problem can be formulated as

$$\begin{aligned} \text{minimize} \quad & \int_0^\infty e^{-\beta t} \boldsymbol{c} \cdot \boldsymbol{q}_t dt \\ \text{subject to} \quad & \dot{\boldsymbol{q}}_t = \boldsymbol{\lambda} - \boldsymbol{\mu}_t \boldsymbol{A}, \quad \forall t \in \mathbb{R}_+, \\ & \boldsymbol{q}_t \geq \boldsymbol{0}, \quad \forall t \in \mathbb{R}_+, \qquad \boldsymbol{\mu}_t \in \mathbb{U}, \quad \forall t \in \mathbb{R}_+, \end{aligned} \tag{1}$$

where $\mathbb{U} = \{\boldsymbol{\mu} \in [0,1]^{\mathbb{I}} : \|\boldsymbol{\mu}\|_1 = 1\}$ and the initial state of $\boldsymbol{q}_t$ is $\boldsymbol{q}_0$.

In the remainder of this section, we exploit results in optimal control theory and present necessary and sufficient conditions for the optimality of Problem (1). As previously noted, the Pontryagin Maximum Principle [19] typically only provides necessary conditions for optimality, but sufficient conditions can be shown to be the case for our optimal control problem.

Proposition 2.3. *Let $\boldsymbol{q}_0$ be the initial state of a fluid model. Let $\{\boldsymbol{\mu}_t^* \in \mathbb{R}_+^{\mathbb{I}} : t \in \mathbb{R}_+\}$ be a fluid-level admissible policy, and let $\boldsymbol{q}_t^* = \boldsymbol{q}_0 - \boldsymbol{\lambda} t + \int_0^t \boldsymbol{\mu}_{t'}^* \boldsymbol{A} dt'$ be the associated queue length process. Assume there exists a continuous process $\{\boldsymbol{p}_t \in \mathbb{R}^{\mathbb{J}} : t \in \mathbb{R}_+\}$ with piecewise continuous $\dot{\boldsymbol{p}}_t$ and a process $\{\boldsymbol{\eta}_t \in \mathbb{R}_+^{\mathbb{J}} : t \in \mathbb{R}_+\}$ such that the following conditions are satisfied: **(C1)** $\boldsymbol{\mu}_t^* \in \arg\max\{\mu A p_t : \boldsymbol{\mu} \in \mathbb{U}\}$; **(C2)** $\dot{\boldsymbol{p}}_t - \beta \boldsymbol{p}_t = \boldsymbol{c} - \boldsymbol{\eta}_t$; **(C3)** $\boldsymbol{q}_t^* \cdot \boldsymbol{\eta}_t = 0$, $\boldsymbol{q}_t^* \geq 0$, $\boldsymbol{\eta}_t \geq 0$; **(C4)** $\liminf_{t \to \infty} \boldsymbol{p}_t \cdot (\boldsymbol{q}_t^* - \boldsymbol{q}_t) \geq 0$ for any fluid model $(\boldsymbol{q}_t, \boldsymbol{\delta}_t)$ with initial condition $\boldsymbol{q}_0$. Then, $\{\boldsymbol{\mu}_t^* : t \in \mathbb{R}_+\}$ is a solution to the optimal control problem* (1).

3 OPTIMAL CONTROL

3.1 Technical Preliminaries

We refer to the stochastic model in Section 2.2 as the pre-limit model and refer to the fluid model in Section 2.3 as the limit system. For the pre-limit model, recall that a basic schedule is a collection of queues from each of which a packet can depart simultaneously and is represented by a $|\mathbb{J}|$-dimensional binary vector $\boldsymbol{s} = [s(\boldsymbol{\rho}) \in \{0,1\} : \boldsymbol{\rho} \in \mathbb{J}]$, where $s(\boldsymbol{\rho}) = 1$ if and only if $\boldsymbol{\rho}$ is in the collection composing the basic schedule and $\mathbb{J} := [n] \times [n]$ denotes the set of queues. For $\boldsymbol{\rho} \in \mathbb{J}$ and $\boldsymbol{s} \in \mathbb{I}$, with $\mathbb{I}$ the set of all basic schedules defined in Section 2, we use $\boldsymbol{\rho} \in \boldsymbol{s}$ if $s(\boldsymbol{\rho}) = 1$. For a basic schedule $\boldsymbol{s} \in \mathbb{I}$, we define the *weight* of $\boldsymbol{s}$ by $w(\boldsymbol{s}) := \sum_{\boldsymbol{\rho} \in \boldsymbol{s}} c(\boldsymbol{\rho})$, where $\boldsymbol{c} \in (\mathbb{R}^{+})^{\mathbb{J}}$ is the cost coefficient vector introduced in (1).

While time in the pre-limit system is discrete with queue-length vector $\boldsymbol{Q}_t \in \mathbb{Z}_{+}^{\mathbb{J}}$ at time t, time in the limit system is continuous with the state space of (fluid) queue-length vectors $\boldsymbol{q}_t \in \mathbb{R}_{+}^{\mathbb{J}}$. From Proposition 2.2, we define a *(fluid-level) schedule* by a convex combination of basic schedules and represent it as an $|\mathbb{I}|$-dimensional vector $\boldsymbol{\mu} = [\mu(\boldsymbol{s}) \in [0,1] : \boldsymbol{s} \in \mathbb{I}]$ with $\|\boldsymbol{\mu}\|_1 = 1$, where $\mu(\boldsymbol{s})$ is the coefficient of schedule $\boldsymbol{s}$. Furthermore, schedule $\boldsymbol{\mu}$ is *admissible* at state $\boldsymbol{q} \in \mathbb{R}_{+}^{\mathbb{J}}$ if and only if $\boldsymbol{\mu} \in \mathbb{U}(\boldsymbol{q})$, as defined in Proposition 2.2.

3.2 Critical Thresholds

We now introduce, for each state $\boldsymbol{q} \in \mathbb{R}_{+}^{\mathbb{J}}$, a family of linear programming problems, indexed by non-negative real numbers, from which we construct an associated (admissible) schedule. These schedules are instrumental to the development of the optimal control algorithms in Section 3.3. For a given state $\boldsymbol{q}$ and a real value $\tau \in \mathbb{R}_{+}$, define sets $\mathbb{I}_\tau \subset \mathbb{I}$ and $\mathbb{J}_{\boldsymbol{q}} \subset \mathbb{J}$ by $\mathbb{I}_\tau := \{\boldsymbol{s} \in \mathbb{I} : w(\boldsymbol{s}) \geq \tau\}$ and $\mathbb{J}_{\boldsymbol{q}} := \{\boldsymbol{\rho} \in \mathbb{J} : q(\boldsymbol{\rho}) = 0\}$, respectively, and define an $|\mathbb{I}_\tau|$-dimensional vector $\boldsymbol{w}_\tau := [w(\boldsymbol{s}) - \tau : \boldsymbol{s} \in \mathbb{I}_\tau] \in \mathbb{R}_{+}^{\mathbb{I}_\tau}$. Then, for τ with $\mathbb{I}_\tau \neq \emptyset$, we formulate the following problem:

$$\max \quad \boldsymbol{w}_\tau \cdot \boldsymbol{\nu}, \quad \text{s.t.} \quad \boldsymbol{\nu} \boldsymbol{A}_{\tau,\boldsymbol{q}} \leq \boldsymbol{\lambda}_{\boldsymbol{q}}, \quad \boldsymbol{\nu} \geq \boldsymbol{0}, \qquad (P_{\boldsymbol{q},\tau})$$

where $\boldsymbol{A}_{\tau,\boldsymbol{q}} := [A(\boldsymbol{s},\boldsymbol{\rho}) : \boldsymbol{s} \in \mathbb{I}_\tau, \boldsymbol{\rho} \in \mathbb{J}_{\boldsymbol{q}}] \in \{0,1\}^{\mathbb{I}_\tau \times \mathbb{J}_{\boldsymbol{q}}}$, $\boldsymbol{\lambda}_{\boldsymbol{q}} := [\lambda(\boldsymbol{\rho}) : \boldsymbol{\rho} \in \mathbb{J}_{\boldsymbol{q}}] \in [0,1]^{\mathbb{J}_{\boldsymbol{q}}}$, and $\boldsymbol{\nu} \in \mathbb{R}^{\mathbb{I}_\tau}$ is the vector of decision variables. Note that, if $\tau = 0$, then $\mathbb{I}_0 = \mathbb{I}$ and $\boldsymbol{w}_0 = \boldsymbol{A}\boldsymbol{c}$.

Remark. *The feasible region for Problem $(P_{\boldsymbol{q},\tau})$ is nonempty because $\boldsymbol{\nu} = \boldsymbol{0}$ obviously satisfies all constraints. From any feasible vector $\boldsymbol{\nu}$ for Problem $(P_{\boldsymbol{q},\tau})$, if we define $\boldsymbol{\mu} \in \mathbb{R}^{\mathbb{I}}$ by $\mu(\boldsymbol{s}) = \nu(\boldsymbol{s}) \mathbf{1}_{\mathbb{I}_\tau}(\boldsymbol{s})$, then we have $\boldsymbol{\mu} \in \mathbb{U}(\boldsymbol{q})$ due to the constraints in Problem $(P_{\boldsymbol{q},\tau})$. Thus, when $\|\boldsymbol{\mu}\|_1 = \|\boldsymbol{\nu}\|_1 = 1$, $\boldsymbol{\mu}$ is an admissible schedule at state $\boldsymbol{q}$.*

Theorem 3.1. *For any state $\boldsymbol{q}$, there exists a $\tau = \tau(\boldsymbol{q}) \in \mathbb{R}_{+}$ such that Problem $(P_{\boldsymbol{q},\tau})$ has an optimal solution $\boldsymbol{\nu}$ that can be extended to an admissible schedule at state $\boldsymbol{q}$; namely, $\|\boldsymbol{\nu}\|_1 = 1$. We call such τ a* critical threshold *of state $\boldsymbol{q}$.*

We next devise a search algorithm for critical thresholds that will terminate in a finite number of iterations. Due to space limitations, we refer the reader to [14] for a presentation of Algorithms 1 and 2, together with supporting theoretical results, that are intended to serve this purpose and that will be used to identify the desired critical thresholds. (Our Algorithm numbering herein is intended to maintain consistency with [14].)

3.3 Optimal Control Algorithm

By exploiting the critical threshold for any state $\boldsymbol{q}$ from the previous section, we now introduce an optimal control algorithm and show that it renders an optimal solution to the optimal control problem (1).

Algorithm 4 Optimal Control Algorithm for initial state $\boldsymbol{q}_{t=0}$

1: Set $k = 0$, $t_0 = 0$, and $\boldsymbol{q}_0^* = \boldsymbol{q}_{t=0}$
2: **while** $t_k < \infty$ **do**
3: Let τ_k be the critical threshold from a combination of Algorithms 1 and 2 with input $\boldsymbol{q} = \boldsymbol{q}_{t_k}^*$
4: Find an optimal point $\boldsymbol{\nu}_k$ to Problem $(P_{\boldsymbol{q},\tau})$ with $\boldsymbol{q} = \boldsymbol{q}_{t_k}^*$ and $\tau = \tau_k$ such that $\|\boldsymbol{\nu}_k\| = 1$.
5: Define $\boldsymbol{\mu}^* \in \mathbb{R}^{\mathbb{I}}$ by $\mu^*(\boldsymbol{s}) = \begin{cases} \nu_k(\boldsymbol{s}) & \text{if } \boldsymbol{s} \in \mathbb{I}_{\tau_k} \\ 0 & \text{otherwise} \end{cases}$
6: Set
$$t_{k+1} = t_k + \min\left\{ \frac{q_{t_k}(\boldsymbol{\rho})}{(\mu^* A)(\boldsymbol{\rho}) - \lambda(\boldsymbol{\rho})} \;:\; \boldsymbol{\rho} \in \mathbb{J} \backslash \mathbb{J}_{\boldsymbol{q}_{t_k}^*},\ (\mu^* A)(\boldsymbol{\rho}) - \lambda(\boldsymbol{\rho}) > 0 \right\}$$
7: Set $\boldsymbol{\mu}_t^* = \boldsymbol{\mu}^*$ for $t \in [t_k, t_{k+1})$ and $\boldsymbol{q}_t^* = \boldsymbol{q}_{t_k}^* + (t - t_k)\boldsymbol{\lambda} - (t - t_k)\boldsymbol{\mu}^* \boldsymbol{A}$ for $t \in [t_k, t_{k+1}]$
8: Set $k = k + 1$

Proposition 3.2. *In Algorithm 4, we have that $\boldsymbol{\mu}_t^*$ is a fluid-level admissible policy and $\boldsymbol{q}_t^*$ is the continuous process satisfying $\dot{\boldsymbol{q}}_t^* = \boldsymbol{\lambda} - \boldsymbol{\mu}_t^* \boldsymbol{A}$ with initial state $\boldsymbol{q}_{t=0}$.*

An importan mathematical property within the context of our results is weakly stable, the definition of which is as follows.

Definition 3.3 ([9, Definition 6]). A fluid-level admissible policy $\boldsymbol{\mu}_t$ is *weakly stable* if the corresponding fluid queue length process $\{\boldsymbol{q}_t : t \in \mathbb{R}_{+}\}$ with initial state $\boldsymbol{q}_0 = \boldsymbol{0}$ satisfies $\boldsymbol{q}_t = \boldsymbol{0}$ for all $t \geq 0$.

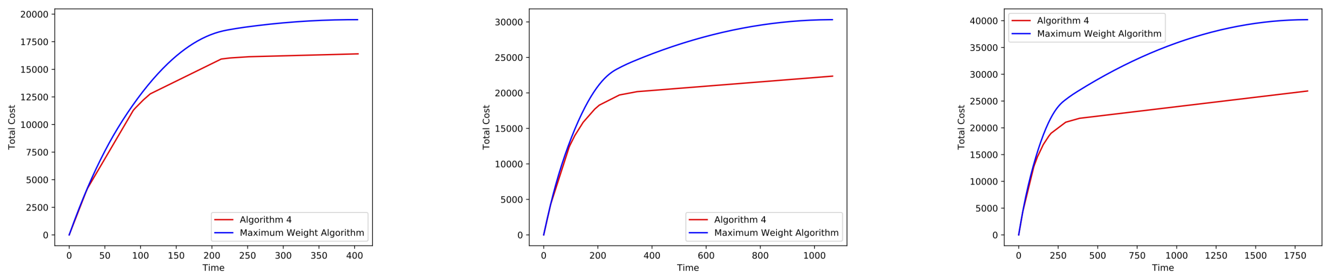
We then have the following main result on the optimality of our solution to the optimal control problem, for which we note that the corresponding optimal control policy is weakly stable.

Theorem 3.4. *Assume that the arrival rate vector $\boldsymbol{\lambda}$ is in the capacity region. Then, $(\boldsymbol{q}_t^*, \boldsymbol{\mu}_t^*)$ is an optimal solution to problem* (1).

3.4 Relationship with $c\mu$-Rule

Given an arrival rate vector $\boldsymbol{\lambda}$ and initial queue length $\boldsymbol{q}_0$ such that $\lambda(i,j) = q_0(i,j) = 0$ for all $i \in [n]$ and $j \in [n] \setminus \{1\}$, the $n \times n$ input-queued switch is equivalent to n parallel queues with one server. The $c\mu$-rule is well-known for this case to be an optimal policy that minimizes the discounted total cost over an infinite horizon in both the stochastic and fluid models (see [8] and [3]); and, in this case, Algorithm 4 follows the $c\mu$-rule in the fluid model.

However, the $c\mu$-rule is not optimal for the $n \times n$ input-queued switch in general. Consider a 3×3 input-queued switch fluid model such that $\lambda(i,j) = 0.45$ if $(i,j) = (1,1), (1,2), (2,1), (2,3)$, and zero otherwise; $c(i,j) = 1$ if $(i,j) = (1,2), (2,3)$, $c(i,j) = 0.5$ if $(i,j) = (2,1)$, $c(i,j) = 0.1$ if $(i,j) = (1,1), (2,3)$, and zero otherwise; $\boldsymbol{q}_0 = \boldsymbol{0}$. Then, according to the $c\mu$-rule, the admissible schedule at $\boldsymbol{q}$ with



(a) $\kappa = 0.70$, **Relative Gap is** 19% **(b)** $\kappa = 0.90$, **Relative Gap is** 35% **(c)** $\kappa = 0.95$, **Relative Gap is** 50%

Figure 1: Performance Comparisons of Total Costs under Optimal Policy (Algorithm 4) and Max-Weight Algorithm

$q(1,2) = q(2,3) = q(2,1) = 0$ becomes

$$\mu(\boldsymbol{s}) = \begin{cases} 0.45 & \text{for } \boldsymbol{s} \text{ such that } s(1,2) = s(2,3) = 1 \\ 0.45 & \text{for } \boldsymbol{s} \text{ such that } s(2,1) = 1 \\ 0.10 & \text{for } \boldsymbol{s} \text{ such that } s(1,1) = 1 \\ 0 & \text{otherwise} \end{cases}.$$

Hence, the queue lengths for $(1,2)$, $(2,3)$ and $(2,1)$ are maintained at zero but the queue length for $(1,1)$ increases with rate $0.45 - 0.10 = 0.35$, which shows that the $c\mu$-rule is not weakly stable.

On the other hand, according to Theorem 3.6 in [14], Algorithm 4 is weakly stable. In this example, the critical threshold at $\boldsymbol{q}_0 = \boldsymbol{0}$ is $\tau = 0$ and the admissible schedule is

$$\mu^*(\boldsymbol{s}) = \begin{cases} 0.45 & \text{for } \boldsymbol{s} \text{ such that } s(1,2) = s(2,1) = 1 \\ 0.45 & \text{for } \boldsymbol{s} \text{ such that } s(1,1) = s(2,3) = 1 \\ 0 & \text{otherwise} \end{cases},$$

which maintains the system to be empty.

4 COMPUTATIONAL EXPERIMENTS

In this section, we present computational experiments that compare the performance of our optimal control algorithm with that of the max-weight scheduling algorithm and the $c\mu$-rule in the fluid model context. We fix the number of input and output ports to be $n \in \mathbb{Z}_+$ and fix the throughput $\kappa \in (0,1)$. For $1 \le i,j \le n$, we randomly generate the costs $c(i,j) \in (0,1)$ and the arrival rates $\lambda(i,j) \in (0,1)$ such that $\max\{\sum_{k=1}^n \lambda(i,k), \sum_{k=1}^n \lambda(k,j) : i,j \in [n]\} = \kappa$. We also choose an initial queue length to be an integer between 1 and 100 uniformly at random for each $(i,j) \in [n] \times [n]$.

With these parameters, we apply Algorithm 4 until we reach the time T at which the queue length becomes 0 for all queues. During our experiments, we let $t_0, t_1, \ldots, t_K$ denote the epochs at which Algorithm 4 updates the admissible schedule, with $t_0 = 0$ and $t_K = T$. Then, the total cost $\int_0^\infty \boldsymbol{c} \cdot \boldsymbol{q}_t \, dt$ is given by $\sum_{k=0}^{K-1} \int_{t_k}^{t_{k+1}} \boldsymbol{c} \cdot \boldsymbol{q}_t dt = \sum_{k=0}^{K-1} \boldsymbol{c} \cdot (\frac{\boldsymbol{q}_{t_{k+1}} + \boldsymbol{q}_{t_k}}{2})(t_{k+1} - t_k)$ because on the interval $[t_k, t_{k+1}]$ the admissible schedule does not change and $\boldsymbol{q}_t$ is a linear function. Note that, even though the objective function in the optimal control problem (1) has a discount factor $\beta \in (0,1)$, we set $\beta = 1$ for the results of our computational experiments herein because Algorithm 4 does not depend on β.

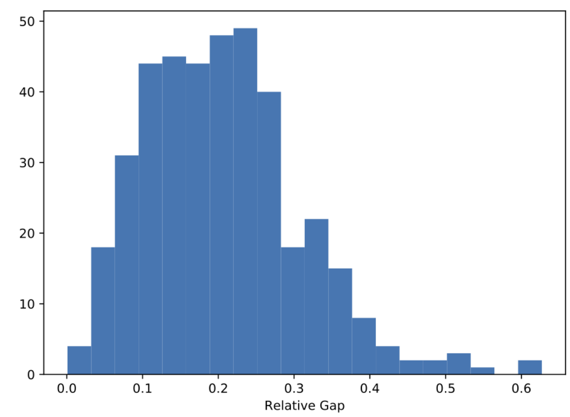
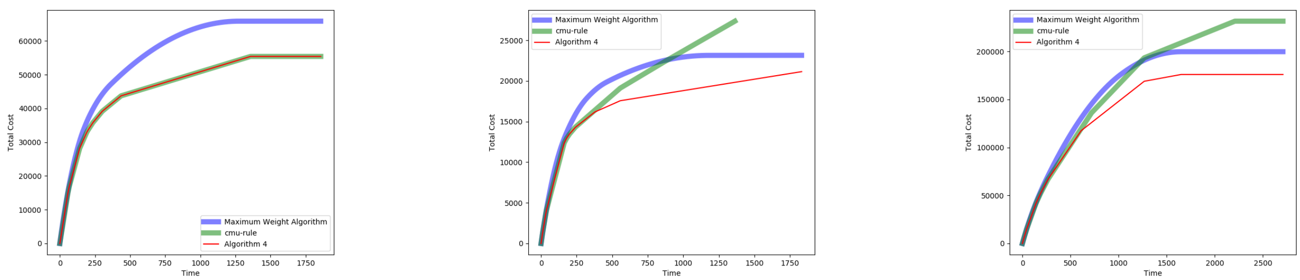


Figure 2: Histogram of Relative Gaps for $\kappa = 0.9$

While the existence and uniqueness of the fluid limit under the max-weight scheduling algorithm has been proven (see [9] and [22]), an explicit formula is not known. Hence, to numerically compute the max-weight scheduling algorithm in the fluid model, we partition the interval $[0,T]$ into slots of size Δt; then, for time slot $[t'_k, t'_k + \Delta t]$, we find a basic schedule of the max-weight algorithm with respect to $\boldsymbol{q}_{t_k}$, say $\boldsymbol{s} \in \mathbb{I}$, and use this schedule during that time slot. In other words, we set $q_{t_{k+1}}(i,j) = \max\{q_{t_k}(i,j) + (\lambda(i,j) - s(i,j))\Delta t, 0\}$ for $(i,j) \in [n] \times [n]$ and approximately measure the total cost on the interval $[0,T]$ by (assuming that $t'_{K'} = T$) $\int_0^T \boldsymbol{c} \cdot \boldsymbol{q}_t \, dt \approx \sum_{k=1}^{K'-1} \boldsymbol{c} \cdot \boldsymbol{q}_{t'_k}$, which is close to the actual total cost under the max-weight scheduling algorithm as $\Delta t \to 0$ and we selected Δt accordingly.

Figure 1 illustrates a representative sample of the total cost over time on $[0,T]$ for the 3×3 input-queued switch fluid model under our optimal control policy and the max-weight scheduling policy. The cost coefficients and the initial queue lengths are set to be the same in each of these three experiments. We vary the throughput κ, defined above, across the three experiments (i.e., $\kappa = 0.7, 0.9, 0.95$) while fixing the ratio among the arrival rates. As observed in the figure, the performance of our optimal policy (Algorithm 4) improves in comparison with that of the max-weight scheduling algorithm as



(a) Algorithm 4 coincides $c\mu$-rule **(b) Unstable $c\mu$-rule** **(c) Stable but not optimal $c\mu$-rule**

Figure 3: Performance Comparisons of Total Costs under Optimal Policy (Algorithm 4) and $c\mu$-rule

the throughput κ increases. To quantify this performance comparison, we calculate the *relative gap* defined by the difference between the total costs at time T under the two algorithms divided by the total cost at time T of the optimal algorithm. The growth in this relative performance gap as the throughput increases ranges from 19% for $\kappa = 0.7$, to 35% for $\kappa = 0.9$ and 50% for $\kappa = 0.95$.

Figure 2 illustrates a representative sample of the corresponding relative performance gap results for various combinations of costs, initial state, and arrival rates under a fixed throughput of $\kappa = 0.9$. We observe that the distribution of the relative gap demonstrates improved performance of at least 10%, in most cases, under Algorithm 4 in comparison with the max-weight scheduling. The sample average of the relative performance gap is around 20%.

We also compare the total cost under our optimal policy (Algorithm 4) and the $c\mu$-rule. Figure 3 illustrates a representative sample of the total cost over time on $[0, T]$ for the 3×3 input-queued switch fluid model, demonstrating three different types of behavior. In Figure 3a, the $c\mu$-rule and the optimal algorithm are identical and provide the same performance. We observe in Figure 3b, however, that the $c\mu$-rule is unstable and clearly not optimal. Moreover, even when the $c\mu$-rule is stable, it may not be optimal as shown in Figure 3c. The highest relative performance improvement of our optimal policy over instances of the stable $c\mu$-rule is more than 70%.

5 CONCLUSIONS

We studied a fluid model of general $n \times n$ input-queued switches where each fluid flow has an associated cost, and derived an optimal scheduling control policy under a general linear objective function based on minimizing discounted fluid cost over an infinite horizon. We demonstrated that, while in certain parameter domains the optimal policy coincides with the $c\mu$-rule, in general the optimal policy is determined algorithmically through a constrained flow maximization problem whose parameters, essentially Lagrangian multipliers of some key network constraints, were in turn identified by another set of carefully designed algorithms. Computational experiments within fluid models of input-queued switches demonstrated the significant benefits of our optimal scheduling policy over variants of max-weight scheduling and the $c\mu$-rule.

REFERENCES

[1] Matthew Andrews, Kyomin Jung, and Alexander Stolyar. 2007. Stability of the Max-weight Routing and Scheduling Protocol in Dynamic Networks and at Critical Loads. In *STOC '07*. 145–154.

[2] Florin Avram, Dimitris Bertsimas, and Michael Ricard. 1995. Fluid models of sequencing problems in open queueing networks; an optimal control approach. *Stochastic Networks* (1995), 199–234.

[3] Nicole Bäuerle and Ulrich Rieder. 2000. Optimal control of single-server fluid networks. *QUESTA* 35, 1 (01 Jul 2000), 185–200.

[4] P. Billingsley. 2013. *Convergence of Probability Measures*. Wiley.

[5] Maury Bramson. 1998. State space collapse with application to heavy traffic limits for multiclass queueing networks. *QUESTA* 30, 1 (01 Nov 1998), 89–140.

[6] H. Chen and D.D. Yao. 2001. *Fundamentals of Queueing Networks: Performance, Asymptotics, and Optimization*. Springer.

[7] Hong Chen and David D. Yao. 1993. Dynamic Scheduling of a Multiclass Fluid Network. *Op. Res.* 41, 6 (1993), 1104–1115.

[8] D. R. Cox and W. L. Smith. 1961. *Queues*. Methuen.

[9] J.G. Dai and B. Prabhakar. 2000. The throughput of data switches with and without speedup. In *INFOCOM '00*. 556–564.

[10] J. G. Dai. 1995. On Positive Harris Recurrence of Multiclass Queueing Networks: A Unified Approach Via Fluid Limit Models. *Ann. Appl. Probab.* 5, 1 (02 1995), 49–77. https://doi.org/10.1214/aoap/1177004828

[11] W. N. Kang and R. J. Williams. 2012. Diffusion approximation for an input-queued packet switch operating under a maximum weight algorithm. *Sto. Sys.* (2012).

[12] Y. Lu, S.T. Maguluri, M.S. Squillante, and T. Suk. 2016. Delay-optimal scheduling for some input-queued switches. *Perf. Eval. Rev.* 44, 2 (September 2016).

[13] Y. Lu, S.T. Maguluri, M.S. Squillante, and T. Suk. 2018. Optimal Dynamic Control for Input-Queued Switches in Heavy Traffic. In *ACC '18*.

[14] Yingdong Lu, Mark S. Squillante, and Tonghoon Suk. 2020. Optimal Scheduling Control in Fluid Models of General n×n Input-Queued Switches. (2020). arXiv:1910.13657v2.

[15] Constantinos Maglaras. 2000. Discrete-review policies for scheduling stochastic networks: trajectory tracking and fluid-scale asymptotic optimality. *Ann. Appl. Probab.* 10, 3 (08 2000), 897–929. https://doi.org/10.1214/aoap/1019487513

[16] Siva Theja Maguluri, Sai Kiran Burle, and R. Srikant. 2016. Optimal Heavy-Traffic Queue Length Scaling in an Incompletely Saturated Switch. *Preprint* (2016).

[17] Siva Theja Maguluri and R. Srikant. 2016. Heavy traffic queue length behavior in a switch under the MaxWeight algorithm. *Stoch. Syst.* 6, 1 (2016), 211–250.

[18] N. McKeown, V. Anantharam, and J. Walrand. 1996. Achieving 100switch. In *INFOCOM '96*, Vol. 1. 296–302 vol.1.

[19] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mishchenko. 1962. *The Mathematical Theory of Optimal Processes*. Interscience, New York.

[20] D. Shah, J. N. Tsitsiklis, and Y. Zhong. 2015. On queue-size scaling for input-queued switches. arxiv preprint http://arxiv.org/abs/1405.4764.

[21] D. Shah, N. S. Walton, and Y. Zhong. 2014. Optimal queue-size scaling in switched networks. *Ann. Appl. Probab.* 24, 6 (12 2014), 2207–2245.

[22] D. Shah and D. Wischik. 2012. Switched networks with maximum weight policies: Fluid approximation and multiplicative state space collapse. *Ann. Appl. Probab.* 22, 1 (2012), 70–127.

[23] Alexander L. Stolyar. 2004. MaxWeight scheduling in a generalized switch: State space collapse and workload minimization in heavy traffic. *Ann. Appl. Probab.* 14, 1 (02 2004), 1–53. https://doi.org/10.1214/aoap/1075828046

[24] L. Tassiulas and A. Ephremides. 1992. Stability Properties of Constrained Queueing Systems and Scheduling Policies for Maximum Throughput in Multihop Radio Networks. *IEEE Trans. Automat. Control* 37 (1992), 1936–1948.