A Restless Bandit Model for Energy-Efficient Job Assignments in Server Farms

Jing Fu, *Member, IEEE*, Xinyu Wang, Zengfu Wang, and Moshe Zukerman, *Life Fellow, IEEE*

***Abstract*— We aim to maximize the energy efficiency, gauged as average energy cost per job, in a large-scale server farm with various storage or/and computing components modeled as parallel abstracted servers. Each server operates in multiple power modes characterized by potentially different service and energy consumption rates. The heterogeneity of servers and multiple power modes complicate the maximization problem, where optimal solutions are generally intractable. Relying on the Whittle relaxation technique, we resort to a near-optimal, scalable job-assignment policy. Under a mild condition related to the service and energy consumption rates of the servers, we prove that our proposed policy approaches optimality as the size of the entire system tends to infinity; that is, it is asymptotically optimal. For the non-asymptotic regime, we show the effectiveness of the proposed policy through numerical simulations, where the policy outperforms all the tested baselines, and we numerically demonstrate its robustness against heavy-tailed job-size distributions.**

***Index Terms*— Restless bandit; job-assignment; asymptotic optimality.**

I. Introduction

THE ever-increasing demand for internet services in recent decades has led to explosive growth in data centers and the markets of computing and storage infrastructures to the so-called Zettabyte Era [1], [2]. In 2014, U.S. data centers were reported to consume around 70 billion kWh of annual electricity and this consumption is predicted to continue increasing [2]. Computing and storage components have been considered major contributors to power consumption in data centers [3], [4]. We study energy-efficient scheduling policies in a large server farm with widely deployed abstracted servers, each of which represents a physical component used to serve incoming customer requests.

Methodologies applicable to server farm scheduling or network resource allocation have been studied from several perspectives. Energy-efficient policies were considered in [5], [6] with only identical servers, and in [7], [8] through static scheduling mechanics without dynamic reuse of released physical resources. Nonetheless, to meet the various demands of internet customers, service providers have launched a large number of networked facilities with highly diverse physical features in backhaul networks, where frequent reuse of released resources is the preferred option because of its efficiency benefits. The heterogeneity of backhaul network facilities keeps increasing in terms of not only their characteristics of implementing specific functions but also their generations [9]. Advanced virtualization techniques abstract these physical components as network resources in modern Cloud computing platforms [10]. It is important to consider the heterogeneity of such abstracted components in the development of scheduling policy because it has substantial implications for the optimized profit of service providers and costs incurred by customers.

Research on the development of dynamic resource allocation methodologies for large-scale server farms (networks) with the reuse of released physical resources has been conducted under certain simplifying assumptions. The work in [11]–[13] considered heterogeneous servers but under the assumption of negligible power consumption of idle servers, while in [14], [15] it was assumed that servers either operate at their peak power consumption rates or stay idle. Then, in [16], [17], it was assumed that servers' power consumption linearly increases in their service rates. Such assumptions of specific power functions simplify the analysis of the relevant problems. However, the diversity of computing/storage components prevents a specific function of their power consumption from being always applicable and there is a need for a methodology applicable to any power consumption behavior. Publications focusing on the server, GPU, and storage component power consumption pointed out that real-life power behavior does not support the convexity (or linearity) assumption [18], [19]. In this paper, we do not assume convexity or linearity of power functions but consider more general and practical situations and in this way, our policy solutions can apply to a wider range of practical scenarios.

We focus on energy-efficient server farms consisting of a large number of abstracted components that are potentially different in terms of service rates, power consumption rates, and their abilities to serve different jobs. Optimizing resource allocation in server farms is achieved in the vein of the *restless multi-armed bandit problem* (RMABP) proposed in [20]. The RMABP is a special type of Markov decision process (MDP) consisting of parallel *bandit processes*, which are also MDPs evolving with binary actions, referred to as the *active* and *passive* modes. The RMABP includes a large

This work was supported in part by the Shenzhen Municipal Science and Technology Innovation Committee under Project JCYJ20180306171144091, a grant from the Research Grants Council of the Hong Kong SAR, P. R. China (CityU 11200721), and Dr Jing Fu's start-up funds from RMIT (Ref. 800132109).

Jing Fu is with the School of Engineering, STEM College, RMIT University, VIC 3000, Australia (e-mail: jing.fu@rmit.edu.au).

Zengfu Wang is with the School of Automation, Northwestern Polytechnical University, Xi'an 710072, P. R. China (e-mail: wangzengfu@nwpu.edu.cn).

Xinyu Wang and Moshe Zukerman are with the Department of Electrical Engineering, City University of Hong Kong, Hong Kong SAR, P. R. China (e-mail: xywang47-c@my.cityu.edu.hk; moshezu@cityu.edu.hk).

number of such bandit processes that are competing for limited opportunities of being evolved in the active mode. In [21], RMABP was proved to be PSPACE-hard in general. Whittle [20] proposed the classical *Whittle index policy* through the *Whittle relaxation technique* and conjectured its asymptotic optimality; that is, the Whittle index policy was conjectured to approach optimality as the scale of the entire system tends to infinity. The Whittle index policy is scalable for large problems and, if it is asymptotically optimal, its performance degradation is bounded and diminishes when the problem size becomes larger and larger. Nonetheless, in general, Whittle relaxation technique does not ensure either the existence of *Whittle indices*, the main parameters required to construct the Whittle index policy, or the bounded performance degradation. Whittle indices were originally defined in [20] under a condition, subsequently referred to as *Whittle indexability*. In [22], Weber and Weiss proved the asymptotic optimality of the Whittle index policy with an extra, non-trivial condition that requires the existence of a *global attractor* of a proposed process associated with the RMABP. In the research field of RMABP, the discussions on Whittle indexability and the global attractor remain open questions in the past several decades.

In [23], based on the results in [24], a scalable job-assignment policy was proposed and proved to be asymptotically optimal for a simplified server farm, where the servers/components were assumed to have only two power modes (corresponding to two power consumption rates). We refer to Section II for a detailed survey of RMABP and other related work.

The contributions of this paper are as follows.

- We provide a scalable policy that aims to maximize the energy efficiency of a large-scale system of deployed computing/storage clusters. This policy always prioritizes physical components (abstracted servers) according to the descending order of their associated *indices*, which are real numbers representing the marginal rewards gained by selecting these components. The indices are pre-computed and the complexity of implementing the index policy is only linear in the number of available physical components. We refer to the policy as *Multiple Power Modes with Priorities* (MPMP), reflecting its applicability to a parallel-server system with multiple power modes.
- When job sizes are exponentially distributed, under a mild condition related to the service and energy consumption rates of physical components, we prove that the index policy approaches optimality as the job arrival rates and the number of physical components in each cluster tend to be arbitrarily large proportionately; that is, it is asymptotically optimal. The asymptotic optimality is appropriate for computing/storage clusters with a rapidly increasing number of physical components. We prove that the performance deviation between our proposed MPMP policy and the optimal point in the asymptotic regime diminishes exponentially in the size of the problem. It implies that the MPMP policy is already near-optimal for a relatively small system.

  Recall that no previous results can be directly applied for scalable, asymptotically optimal policies in the large server farm, where the abstracted computing/storage components operate in multiple power/service modes. The complexity of the server farm problem requires a new analysis of the entire system, provided in this paper, including discussions on the indexability and the global attractor for proving asymptotic optimality in the continuous-time case.
- We numerically demonstrate the effectiveness of MPMP in the general case, where MPMP significantly outperforms baseline policies in all the tested cases. We further explore its performance with different job-size distributions and numerically show that the energy efficiency of MPMP is not very sensitive to different shapes of job-size distributions.

The remainder of the paper is organized as follows. In Section II, we discuss other related work for job assignments and RMABP. In Section III, a description of the server farm model is provided, and in Section IV, the underlying stochastic optimization problem is rigorously defined. In Section V, we discuss the indexability of the underlying stochastic process and propose the *indices* - the most important parameters for constructing our policy. In Section VI, we formally define the MPMP policy, and in Section VII we prove its asymptotic optimality. Section VIII presents extensive numerical results that demonstrate the effectiveness of MPMP in the general case. The conclusions of this paper are included in Section IX.

II. Other Related Work

Job-assignment policies with strict capacity constraints of physical resources have been studied in [23], [25], [26], where the release and reuse of physical resources were considered. Following the ideas of restless bandits [20], [22], the authors of [23], [25], [26] proposed scalable policies and proved that the policies, which do not necessarily perform well in small systems, approach the optimal solution in large-scale systems. Optimization problems for small systems can be solved by conventional algorithms, which cannot be directly applied in large cases because of high computational complexity. Nevertheless, these publications assumed either two power modes (that is, fixed power consumption values for busy and idle servers) or power consumption linearly increasing in servers' traffic loads. As mentioned in Section I, here, we overcome the weaknesses of past publications and provide general solutions that are applicable to realistic power functions.

Apart from the job-assignment problems, conventional RMABP has been widely studied and applied to scheduling problems. For instance, in [27], a set of identical servers/processors were scheduled to serve stochastically identical jobs that keep arriving. In [28], Borkar considered a special type of bandit processes, of which the state variables take binary values and are only partially observable. He proved the *Whittle indexability*, a key property for an RMABP, by generating and analyzing the indexability of an equivalent process of the original bandit process. In [29], Whittle indexability was proved for a channel-selecting problem where each bandit process was associated with a wireless channel and its state variable was defined as the number of successive transmission failures in that channel. In [30], a group maintenance problem

was modeled as a standard RMABP with a detailed analysis of its Whittle indexability.

In [31], Niño-Mora proposed the partial conservation law (PCL) and the PCL-indexability for RMABP. The latter was proved to imply (be stronger than) the Whittle indexability. Later in 2002, Niño-Mora [32] identified a set of optimization problems that satisfy PCL-indexability and thus are Whittle indexable. A detailed survey about the Whittle and PCL-indexability was provided in [33]. In [34], Niño-Mora defined the indexability and Whittle indices of a bandit process with continuous state space and proposed a method that verified the indexability and computed the corresponding Whittle indices. All in all, these studies have established computational methodologies for verifying Whittle indexability for the general RMABP that aims to maximize/minimize the expected cumulative rewards/costs. Our work, in this paper, aims to maximize a long-run average objective that prevents the existing off-the-shelf techniques from being applied directly. Although from [35], optimizing the long-run average rewards/costs of an MDP can usually be translated to a problem that optimizes the expected cumulative rewards/costs of the same process with an attached, real-valued criterion, this real-valued criterion is not known a priori and has a strong impact on the indexability of the underlying bandit processes. We refer to Section V-B for a detailed explanation for demonstrating the indexability with the long-run average objective for our server farm.

For a general RMABP, to further prove that the Whittle index policy approaches optimality as the number of bandit processes increases to infinity (that is, asymptotic optimality), Weber and Weiss [22] required another non-trivial condition; namely, there exists a fixed point such that the underlying stochastic process of the RMABP will almost surely enter a nearby neighborhood of the point. Such a point is referred to as a *global attractor* of the process. In [36], [37], similar assumptions related to the global attractor were required in the proofs of asymptotic optimality of Whittle index policy in channel selection problems, which are special cases of RMABP. In [24], for a system consisting of a special type of bandit processes, Fu *et al.* proved that such a global attractor exists and hence the resulting policy is asymptotically optimal. The idea of this technique was later applied to a server farm model in [23]. Nonetheless, due to the complexity of the server farm model considered in this paper, it does not fall in the scope of [24] and we cannot directly apply the same conclusions here. Recall that modeling the server farm as an RMABP or RMABP-like system cannot ensure the existence of a scalable near-optimal policy with theoretically bounded performance. The complexity of the problem requires a new and thorough analysis of the indexability and global attractor, which have been, in general, open questions in the past several decades. Here, we resort to scalable policies for such a challenging server farm model with a theoretical performance guarantee.

Our job-assignment problem can be modeled as a set of parallel, restless bandit processes coupled with action constraints. A detailed description of our model is provided in Section III. As discussed, in this paper, we generalize the power functions discussed in [23], which requires a new analysis. In particular, we prove that, for a mild condition related to the service and energy consumption rates of physical components (abstracted servers), the MPMP policy is asymptotically optimal with respect to energy efficiency. To the best of our knowledge, there is no published work applicable to a server farm model with strict capacity constraints and generalized power functions, where a scalable policy is proposed with theoretically guaranteed performance when the system is largely scaled.

III. Model

For any positive integer N, let $[N]$ represent the set $\{1, 2, \ldots, N\}$. Let $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{R}_0$ represent the sets of all real numbers, positive real numbers, and non-negative real numbers, respectively. Similarly, $\mathbb{N}$ and $\mathbb{N}_+$ are the sets of integers and positive integers, respectively.

There are I clusters of physical components. These components are identical within the same cluster in terms of their availability of accommodating jobs, hardware features and software profiles. Physical components in different clusters may be different.

Consider L classes of jobs, each of which is characterized by a tuple $(\lambda_\ell, \mathscr{I}_\ell)$ for $\ell \in [L]$, where the arrival process of jobs in class ℓ follows a Poisson process with rate $\lambda_\ell > 0$, and the set of clusters that are potentially able to serve jobs of class ℓ is given by $\mathscr{I}_\ell \in 2^{[I]} \setminus \{\emptyset\}$. The components in the clusters $i \in \mathscr{I}_\ell$ are referred to as the *available components* for jobs of class ℓ. The sizes of all jobs are considered as independently and identically distributed (i.i.d.) random variables with unit mean, and jobs are arriving sequentially with positive inter-arrival time. We refer to a job of class $\ell \in [L]$ as an ℓ-job.

Each physical component of cluster $i \in [I]$ has a finite capacity, $C_i \in \mathbb{N}_+$, as the maximal number of jobs it can serve simultaneously. Define its energy consumption and service rates as functions of carried load: $\varepsilon_i(n)$ and $\mu_i(n)$, $n \in \{0\} \cup [C_i]$. Note that the service units of each physical component are being reused and released dynamically along the timeline. For the purpose of this paper, assume that $\varepsilon_i(n)$ and $\mu_i(n)$ are finite, non-negative and increasing in n (that is, $0 \le \varepsilon_i(n) \le \varepsilon_i(n+1) < \infty$ and $0 \le \mu_i(n) \le \mu_i(n+1) < \infty$) with $\mu_i(0) \equiv 0$, $\varepsilon_i(0) \ge 0$ and, for $n > 0$, $\varepsilon_i(n) > \varepsilon_i(0)$ and $\mu_i(n) > 0$.

Consider $M_i^0 \cdot h$ components in cluster i where M_i^0 and h are positive integers. Let $\lambda_\ell = \lambda_\ell^0 \cdot h$ for some $\lambda_\ell^0 \in \mathbb{R}_+$. In this context, h is a constant that signifies the scale of the multi-cluster system, which is referred to as the *scaling parameter*. There are in total $J = h \sum_{i \in [I]} M_i^0$ components in our system, labeled by $j \in [J]$, with each computing cluster $i \in [I]$ consisting of physical components $j_1, j_2, \ldots, j_{hM_i^0} \in [J]$. For $j \in [J]$, let i_j represent the label of the cluster with $j \in \mathscr{J}_{i_j}$.

IV. Optimization Problem

At time $t \ge 0$, there are $N_j(t)$ jobs accommodated by physical component j, and define $\boldsymbol{N}(t) = (N_j(t) : j \in [J])$. The $N_j(t)$ is a random variable and is referred to as the *state variable* of component j. When a job is assigned to a component j at time t, the state of component j transitions from $N_j(t)$ to $N_j(t) + 1$; when a job on component j is completed, the state decreases to $N_j(t) - 1$. Preemption of jobs is not permitted in our system. The variable $N_j(t)$ is affected

by the underlying scheduling policy, denoted by ϕ. When an ℓ-job arrives, the scheduling policy selects a component in cluster $i \in \mathscr{I}_\ell$ with at least a vacant slot to accommodate this job or block it. For fairness, we do not allow the rejection of any job when there is a vacant slot in any of the available components for this job. In other words, rejection of a job occurs if and only if all the available components for the job are fully occupied.

To indicate the dependency of the counting process $\{\boldsymbol{N}(t), t \geq 0\}$ and the scheduling policy ϕ, we rewrite the state variables as $N_j^\phi(t)$ and the state vector $\boldsymbol{N}^\phi(t)$.

In this context, the set of all possible values of $N_j^\phi(t)$ of component j in cluster i is $\mathscr{N}_i \coloneqq \{0, 1, \ldots, C_i\}$; that is, $\mathscr{N}_i$ is the *state space* of the process $\{N_j^\phi(t), t \geq 0\}$ associated with a component j in cluster i, abbreviated as the state space of cluster i. The state vector $\boldsymbol{N}^\phi(t)$ of the entire system is then taking values in $\mathscr{N} \coloneqq \prod_{i\in[I]} (\mathscr{N}_i)^{h\cdot M_i^0}$, of which the size is increasing exponentially in h.

More precisely, define *action variable* $a_{\ell,j}^\phi(\boldsymbol{n}) \in \{0, 1\}$ as a function of a state vector $\boldsymbol{n} \in \mathscr{N}$ associated with policy ϕ: if $a_{\ell,j}^\phi(\boldsymbol{n}) = 1$, then a newly-arrived job of class ℓ will be accommodated by physical component j when $\boldsymbol{N}^\phi(t) = \boldsymbol{n}$ under policy ϕ; otherwise, the job will not be assigned to component j. To take account for the rejection of jobs, define a *virtual component* for each class ℓ, labeled by $j_\ell \coloneqq J + \ell$; that is, if $a_{\ell,j_\ell}^\phi(\boldsymbol{n}) = 1$, a new job of class ℓ will be blocked. For $\ell \in [L]$, assume without loss of generality that the virtual component j_ℓ belongs to a virtual cluster $i_\ell \coloneqq I + \ell$. Define a stochastic process $N_{j_\ell}^\phi(t) \equiv 0$, $t \geq 0$, associated with the virtual component j_ℓ, which has state space $\mathscr{N}_{i_\ell} \coloneqq \{0\}$.

To mathematically define the feasibility of policy ϕ, we introduce the following constraints for the action variables:

$$\sum\nolimits_{j\in\mathscr{I}_\ell\cup\{j_\ell\}} a_{\ell,j}^\phi(\boldsymbol{n}) = 1,\ \forall \ell \in [L], \boldsymbol{n} \in \mathscr{N}, \tag{1}$$

and

$$a_{\ell,j}^\phi(\boldsymbol{n}) = 0, \text{if } n_j = C_{i_j},\ \forall \ell \in [L], j \in \mathscr{J}_\ell, \boldsymbol{n} \in \mathscr{N}, \tag{2}$$

where $\mathscr{J}_\ell$ is the set of all available components of job class ℓ, and cluster i_j is the cluster in which component j is located. Define that $a_{\ell,j}^\phi(\cdot) \equiv 0$ for all $j \notin \mathscr{J}_\ell \cup \{j_\ell\}$. Constraints (1) ensure that only one component is selected for an arrived job, and Constraints (2) disable a fully-occupied component from accommodating more jobs.

Recall that a rejection of jobs is not permitted if there is a vacant slot on any available component. It can be addressed by introducing intermediate variables $\bar{a}_\ell^\phi(\boldsymbol{n})$ $(\boldsymbol{n} \in \mathscr{N})$ satisfying

$$\bar{a}_\ell^\phi(\boldsymbol{n}) + \sum\nolimits_{j\in\mathscr{J}_\ell} I(C_{i_j} - n_j) \leq 1,\ \forall \ell \in [L], \boldsymbol{n} \in \mathscr{N}, \tag{3}$$

where $I(x)$ $(x \in \mathbb{R})$ is a Heaviside function with $I(x) = 1$ for $x > 0$; and 0 for $x \leq 0$. For such $\bar{a}_\ell^\phi(\boldsymbol{n})$ that takes values in $(-\infty, 1]$, define

$$a_{\ell,j_\ell}^\phi(\boldsymbol{n}) \coloneqq I(\bar{a}_\ell^\phi(\boldsymbol{n})). \tag{4}$$

From (3), when there is at least one available component in $\mathscr{J}_\ell$, $\bar{a}_\ell^\phi(\boldsymbol{n}) \leq 0$ and so $a_{\ell,j_\ell}^\phi(\boldsymbol{n}) = 0$. If all these components are fully occupied, constraints in (1) and (2) force $a_{\ell,j_\ell}^\phi(\boldsymbol{n}) = 1$, which does not violate constraints in (3).

We aim to maximize the energy efficiency of the entire system; specifically, to maximize the ratio of the long-run average job throughput to the long-run average power consumption. Let

$$\mathfrak{L}^\phi = \lim_{T\to+\infty} \frac{1}{T}\mathbb{E}\int_0^T \sum\nolimits_{j\in[J]} \mu_j(N_j^\phi(t))dt \tag{5}$$

and

$$\mathfrak{E}^\phi = \lim_{T\to+\infty} \frac{1}{T}\mathbb{E}\int_0^T \sum\nolimits_{j\in[J]} \varepsilon_j(N_j^\phi(t))dt \tag{6}$$

represent the long-run average job throughput and the long-run average power consumption, respectively. Our optimization problem is

$$\max\nolimits_\phi\ \mathfrak{L}^\phi/\mathfrak{E}^\phi \tag{7}$$

subject to (1), (2) and (3). Let Φ represent the set of all policies constrained by (1), (2) and (3).

As in [23], [25], consider an optimal policy ϕ^* that maximizes the problem described in (7), (1), (2) and (3) and define a real number

$$e^* = \mathfrak{L}^{\phi^*}/\mathfrak{E}^{\phi^*}. \tag{8}$$

Following [11, Theorem 1], if $0 < \mathfrak{L}^\phi < +\infty$ and $0 < \mathfrak{E}^\phi < +\infty$, then a policy $\phi \in \Phi$ that maximizes

$$\max\nolimits_{\phi\in\Phi} \mathfrak{L}^\phi - e^*\mathfrak{E}^\phi, \tag{9}$$

subject to (1), (2) and (3) also maximizes the problem defined in (7), (1), (2) and (3).

The constraints (3) make our problem slightly different from a standard RMABP in the sense that in our problem, rejecting a job has the lowest priority among all the other actions. Unlike the *uncontrollable* variables constrained by (2), (3) guarantees the lowest priority for rejecting a job rather than disables this action. If $L = 1$ and constraints (3) do not exist, then our problem reduces to an RMABP, where the processes $\{N_j^\phi(t), t \geq 0\}$ $(j \in [J])$ are parallel restless bandit processes coupled by action constraint (1). Note that such an RMABP is no longer applicable to our server farm.

Similar to an RMABP, addressing our problem requires overcoming the challenge of its large state space, which is exponentially increasing in the number of components and optimal solutions are intractable. In [23], a scalable policy was proposed for a similar problem with simplified $\mu_i(n) \equiv \mu_i$ and $\varepsilon_i(n) \equiv \varepsilon_i$ for all $n = 1, 2, \ldots, C_i$. This policy was proved to be asymptotically optimal, for which the performance gap to optimality diminishes exponentially in the scale of the problem. However, in this paper, the generalized $\mu_i(n)$ and $\varepsilon_i(n)$ (that is, the multiple power states) prevent the same technique from being applied directly. For general RMABPs or some relevant problems, the Whittle relaxation technique does not ensure a good, scalable policy that asymptotically approaches optimality. From [20], [22], asymptotic optimality relies on two important but non-trivial properties related to the underlying stochastic process: *Whittle indexability* and the existence of a *global attractor* in the asymptotic regime. These two properties do not necessarily hold in general and remain open questions in the past several decades.

As mentioned in Section I, the existence of an appropriate global attractor has been discovered in a class of RMABPs [24], including the special case studied in [23]. In this paper, the server farm with more general service and power con-

sumption rates does not fall exactly in the scope of [24]. This requires a new analysis of the entire system. As mentioned in Section II, Whittle indexability for relatively general RMABPs has been widely studied with provided sufficient conditions for optimizing cumulative rewards/costs [31]–[34]. Nonetheless, the generalized transition and reward rates of the underlying stochastic process and the objective in this paper - maximization of an average reward - prevent these off-the-shelf techniques from being applied directly. Although the maximization of the average reward of an MDP can usually be translated to the maximization of the cumulative reward of the same process with an attached real-valued criterion, the exact value of this attached criterion cannot be known a priori and has a strong impact on the discussion of Whittle indexability. In [23], for the very special server farm with $\mu_i(n) \equiv \mu_i$ and $\varepsilon_i(n) \equiv \varepsilon_i$, the real-valued criterion was directly offset during the analysis of Whittle indexability, which significantly simplified the entire discussion. Whittle indexability ensures the existence of a scalable policy, derived from *Whittle indices*, for a standard RMABP [20]. Unfortunately, such an unknown criterion persists in this paper which requires a thorough analysis of the entire system. We will discuss in Section V the indexability of the server farm with the generic $\mu_i(n)$ and $\varepsilon_i(n)$ and in Section VII further results for asymptotically optimal policies.

V. Whittle Relaxation and Indexability

A. Whittle Relaxation

Following the idea of *Whittle relaxation* [20], we relax constraints (1) and (3) to

$$\lim_{t\to+\infty} \mathbb{E}\Big[\sum\nolimits_{j\in\mathscr{J}_\ell\cup\{j_\ell\}} a^\phi_{\ell,j}(\boldsymbol{N}^\phi(t))\Big] = 1,\ \forall \ell \in [L] \tag{10}$$

and

$$\lim_{t\to+\infty} \mathbb{E}\Big[\bar{a}^\phi_\ell(\boldsymbol{N}^\phi(t)) + \sum\nolimits_{j\in\mathscr{J}_\ell} I(C_{i_j} - N^\phi_j(t))\Big] \le 1,\ \forall \ell \in [L], \tag{11}$$

respectively. Similarly, (2) can be rewritten as

$$\lim_{t\to+\infty} \mathbb{E}\left[a^\phi_{\ell,j}(N^\phi_j(t)) \mid N^\phi_j(t) = C_{i_j}\right] = 0,\ \forall \ell \in [L], j \in [J]. \tag{12}$$

Define a special policy ϕ_0 with $a^{\phi_0}_{\ell,j}(\boldsymbol{n}) = 1$ for all $\ell \in [L]$, $j \in \mathscr{J}_\ell$, and $\boldsymbol{n} \in \mathscr{N}$. Note that $\phi_0 \notin \Phi$ because it violates the constraints on action variables. We apply the ϕ_0 to all stochastic processes $\{N^{\phi_0}_j(t), t \ge 0\}$ for $j \in [J]$. Define $A_\ell \coloneqq \sum_{j\in\mathscr{J}_\ell} \lim_{t\to+\infty} \mathbb{E}\left[I(C_{i_j} - N^{\phi_0}_j(t))\right]$, where $\lim_{t\to+\infty} \mathbb{E}\left[I(C_{i_j} - N^{\phi_0}_j(t))\right]$ is the proportion of time that $N^\phi_j(t) < C_{i_j}$. The value $1 - A_\ell$ represents the blocking probability of job class ℓ under the policy ϕ_0. Thus, we can further relax (11) as

$$I(\bar{\alpha}^\phi_\ell) \le I(1 - A_\ell),\ \forall \ell \in [L] \tag{13}$$

where $\bar{\alpha}^\phi_\ell = \lim_{t\to+\infty} \mathbb{E}\big[\bar{a}^\phi_\ell(\boldsymbol{N}^\phi(t))\big]$. Equations (9), (10), (12) and (13) comprise a *relaxed* version of our original problem described in (9), (1), (2) and (3). Define $\tilde{\Phi}$ as the set of all the policies ϕ satisfying (10), (12) and (13) so that $\Phi \subset \tilde{\Phi}$.

For clarity of presentation,

- define $\pi^\phi_j(n)$ as the steady state probability of state $n \in \mathscr{N}_{i_j}$ under $\phi \in \Phi$, and define $\boldsymbol{\pi}^\phi_j \coloneqq (\pi^\phi_j(n) :\ n \in \mathscr{N}_{i_j})$;
- for $n \in \mathscr{N}_{i_j}$, $j \in [J] \cup \{j_\ell : \ell \in [L]\}$, $\ell \in [L]$, define $\alpha^\phi_{\ell,j}(n) \coloneqq \lim_{t\to+\infty} \mathbb{E}\big[a^\phi_{\ell,j}(\boldsymbol{N}^\phi(t)) \mid N^\phi_j(t) = n\big]$;
- for $\boldsymbol{\nu} \in \mathbb{R}^L$ and $\boldsymbol{\omega} \in \mathbb{R}^{LJ}$, define, if $n < C_{i_j}$, $r^\phi_{j,n}(\boldsymbol{\nu},\boldsymbol{\omega}) \coloneqq \mu_j(n) - e^*\varepsilon_j(n) - \sum_{\ell:\ j\in\mathscr{J}_\ell} \nu_\ell \alpha^\phi_{\ell,j}(n)$; otherwise, $r^\phi_{j,n}(\boldsymbol{\nu},\boldsymbol{\omega}) \coloneqq \mu_j(n) - e^*\varepsilon_j(n) - \sum_{\ell:\ j\in\mathscr{J}_\ell} (\nu_\ell + \omega_{\ell,j})\alpha^\phi_{\ell,j}(n)$; and let $\boldsymbol{r}^\phi_j(\boldsymbol{\nu},\boldsymbol{\omega}) \coloneqq (r^\phi_{j,n}(\boldsymbol{\nu},\boldsymbol{\omega}) :\ n \in \mathscr{N}_{i_j})$;
- let $\boldsymbol{\imath}^\phi \coloneqq (I(\bar{\alpha}^\phi_\ell) : \ell \in [L])$, $\boldsymbol{\alpha}^\phi = (\alpha^\phi_{\ell,j_\ell} : \ell \in [L])$ and $\boldsymbol{I} = (I(1 - A_\ell) : \ell \in [L])$.

The Lagrangian dual function of the relaxed problem is

$$L(\boldsymbol{\nu},\boldsymbol{\omega},\boldsymbol{\gamma}) = \max\nolimits_{\phi\in\tilde{\Phi}} \sum\nolimits_{j\in[J]} \boldsymbol{\pi}^\phi_j \cdot \boldsymbol{r}^\phi_j(\boldsymbol{\nu},\boldsymbol{\omega}) - \boldsymbol{\nu}\cdot\boldsymbol{\alpha}^\phi - \boldsymbol{\gamma}\cdot\boldsymbol{\imath}^\phi + \boldsymbol{\nu}\cdot\mathbf{1} + \boldsymbol{\gamma}\cdot\boldsymbol{I}, \tag{14}$$

with Lagrangian multipliers $\boldsymbol{\nu}$, $\boldsymbol{\omega}$ and $\boldsymbol{\gamma}$ corresponding to (10), (12) and (13), respectively. Here, $\boldsymbol{x}\cdot\boldsymbol{y}$ is the inner product of vectors $\boldsymbol{x}$ and $\boldsymbol{y}$. In the same vein of [20], for given multipliers $\boldsymbol{\nu}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\omega}$, the optimal solution of the maximization problem in (14) is also optimal for

$$\begin{aligned}\max_{\phi\in\tilde{\Phi}} &\sum\nolimits_{j\in[J]} \boldsymbol{\pi}^\phi_j \cdot \boldsymbol{r}^\phi_j(\boldsymbol{\nu},\boldsymbol{\omega}) - \boldsymbol{\nu}\cdot\boldsymbol{\alpha}^\phi - \boldsymbol{\gamma}\cdot\boldsymbol{\imath}^\phi \\ &= \sum\nolimits_{j\in[J]} \max\nolimits_{\boldsymbol{\alpha}^\phi_j \in [0,1]^{L|\mathscr{N}_{i_j}|}} \boldsymbol{\pi}^\phi_j \cdot \boldsymbol{r}^\phi_j(\boldsymbol{\nu},\boldsymbol{\omega}) \\ &\quad + \sum\nolimits_{\ell\in[L]} \max\nolimits_{\alpha^\phi_{\ell,j_\ell}\in[0,1]} (-\nu_\ell\alpha^\phi_{\ell,j_\ell} - \gamma_\ell I(\bar{\alpha}^\phi_\ell)),\end{aligned} \tag{15}$$

where $\boldsymbol{\alpha}^\phi_j = (\alpha^\phi_{\ell,j}(n) :\ n \in \mathscr{N}_{i_j},\ \ell \in [L])$. Since the constraints of action variables are now interpreted by the multipliers, the problem described in (15) can be decomposed into the following $J + L$ independent sub-problems. For $j \in [J]$,

$$\max\nolimits_{\boldsymbol{\alpha}^\phi_j \in [0,1]^{L|\mathscr{N}_{i_j}|}} \boldsymbol{\pi}^\phi_j \cdot \boldsymbol{r}^\phi_j(\boldsymbol{\nu},\boldsymbol{\omega}), \tag{16}$$

where the steady state distribution $\boldsymbol{\pi}^\phi_j$ is determined by only the action vector $\boldsymbol{\alpha}^\phi_j$ associated with the underlying process $\{N^\phi_j(t), t \ge 0\}$. Similarly, for $\ell \in [L]$,

$$\max\nolimits_{\alpha^\phi_{\ell,j_\ell}\in[0,1]} (-\nu_\ell\alpha^\phi_{\ell,j_\ell} - \gamma_\ell I(\bar{\alpha}^\phi_\ell)). \tag{17}$$

Define $\tilde{\Phi}_1$ as the set of policies determined by action variables $\boldsymbol{\alpha}^\phi_j \in [0,1]^{L|\mathscr{N}_{i_j}|}$ ($j \in [J]$) and $\alpha^\phi_{\ell,j_\ell} \in [0,1]$ ($\ell \in [L]$).

Remark The sub-problems in (16) and (17) are independent problems, each of which has only one-dimensional state space and thus experiences remarkably lower computation time than the original problem. Nonetheless, non-trivial properties are generally required to establish theoretical connections between these one-dimensional sub-problems and the high-dimensional original problem. A detailed survey about RMABP has been provided in Section II.

B. Whittle Indexability

For a standard RMABP, Whittle [20] proposed the well-known *Whittle index policy* when a non-trivial property related to each bandit process was satisfied. This property was later referred to as the Whittle indexability. More precisely, for our problem defined in Section IV, when $L = 1$, the bandit process $\{N^\phi_j(t), t \ge 0\}$, for $j \in [J]$, reduces to a bandit process for a

standard RMABP. In this special case, based on [20], for each $j \in [J]$, if there exist an optimal solution ϕ^* for the problem described in (16) and a vector of real numbers $\boldsymbol{\upsilon}_j^* = (\upsilon_{\ell,j}^*(n) : n \in \mathscr{N}_{i_j}, \ell \in [L])$, satisfying, for all $n \in \mathscr{N}_{i_j} \backslash \{C_{i_j}\}$,

$$\alpha_{\ell,j}^{\phi^*}(n) = \begin{cases} 1, \text{ if } \nu_\ell < \upsilon_{\ell,j}^*(n), \\ a, \text{ if } \nu_\ell = \upsilon_{\ell,j}^*(n), \\ 0, \text{ otherwise}, \end{cases} \tag{18}$$

where a can be any number in $[0,1]$ and $\ell = L = 1$, then we say the bandit process $\{N_j^\phi(t), t \geq 0\}$ associated with $j \in [J]$ is Whittle indexable and the real number $\upsilon_{\ell,j}^*(n)$ is the *Whittle index* for state $n \in \mathscr{N}_{i_j}$ of the process. If all the bandit processes of an RMABP are Whittle indexable, then the RMABP is Whittle indexable. Note that, in (18), the real number $\upsilon_{\ell,j}^*(n)$ must be independent from ν_ℓ. Although the policy ϕ^* is optimal for the sub-problem described in (16), which is usually not applicable to the original problem, the Whittle index $\upsilon_{\ell,j}^*(n)$ intuitively represents the marginal reward of taking $a_{\ell,j}^\phi(\boldsymbol{N}^\phi(t)) = 1$ when $N_j^\phi(t) = n$ and hence brings a bird's-eye view of approximating optimality for the original problem. Whittle [20] proposed a scalable *index policy* by prioritizing states $n \in \mathscr{N}_{i_j}$ of bandit processes $j \in [J]$ according to the descending order of their Whittle indices, which was subsequently referred to as the Whittle index policy. In [20], Whittle conjectured asymptotic optimality of the Whittle index policy, and it was proved by Weber and Weiss [22] under another non-trivial condition - the existence of a global attractor related to the original stochastic process. As indicated in Section I, asymptotic optimality acts as an important performance guarantee for scalable policies in large-scale optimization problems.

Recall that, in general, bandit processes are not necessarily Whittle indexable, so does an RMABP. Please refer to Section II for a detailed survey of past studies on Whittle indexability. Here, we focus on the server farm problem described in Section IV.

Definition 1: We say a physical component $j \in [J]$ in cluster $i \in [I]$ is *energy-efficiently unimodal* if, for any $n_1, n_2 \in \mathscr{N}_i \backslash \{C_i\}$ with $n_1 < n_2$,

$$\begin{aligned} &(\mu_i(n_2+1) - \mu_i(n_2))(\varepsilon_i(n_1+1)\mu_i(n_1) - \varepsilon_i(n_1)\mu_i(n_1+1)) \\ &\leq \big(\mu_i(n_1+1) - \mu_i(n_1)\big)\big(\varepsilon_i(n_2+1)\mu_i(n_2) - \varepsilon_i(n_2)\mu_i(n_2+1)\big) \end{aligned} \tag{19}$$

Intuitively, the energy-efficient unimodality implies a mild relationship of the component energy efficiency, $\mu_i(n)/\varepsilon_i(n)$, in different states $n \in \mathscr{N}_i$: there is at most one bump on the curve of $\mu_i(n)/\varepsilon_i(n)$ as n tends from 0 to C_i. For example, if there exists $n_1 \in \mathscr{N}_i \backslash \{C_i, C_i - 1\}$ with $\mu_i(n_1) = \mu_i(n_1 + 1)$, then (19) indicates either $\varepsilon_i(n_1) = \varepsilon_i(n_1 + 1)$ or, for all $n_2 = n_1 + 1, n_1 + 2, \ldots, C_i - 1$, $\mu_i(n_2) = \mu_i(n_2 + 1)$; that is, the curve of $\mu_i(n)/\varepsilon_i(n)$ is flat as n tends from n_1 to C_i. For $n_1 < n_2$, if $\mu_i(n_1) < \mu_i(n_1 + 1)$ and $\mu_i(n_2) = \mu_i(n_2 + 1)$, then (19) holds all the time. For any $n_1, n_2 \in \mathscr{N}_i \backslash \{C_i\}$ with $n_1 < n_2$, $\mu_i(n_1) < \mu_i(n_1 + 1)$ and $\mu_i(n_2) < \mu_i(n_2 + 1)$, from (19), if $\mu_i(n_1 + 1)/\varepsilon_i(n_1 + 1) \leq \mu_i(n_1)/\varepsilon_i(n_1)$ then $\mu_i(n_2 + 1)/\varepsilon_i(n_2 + 1) \leq \mu_i(n_2)/\varepsilon_i(n_2)$. As a consequence, the curve of $\mu_i(n)/\varepsilon_i(n)$ has at most one bump as n tends from 0 to C_i. The energy-efficient unimodality is only a mild condition because for real-world computing components, such as CPUs and GPUs, energy efficiency increases with processing speed - $\mu_i(n)/\varepsilon_i(n)$ increases with n [19], [38].

Proposition 1: When the job sizes are exponentially distributed and all computing components are energy-efficiently unimodal, if, for all $\ell \in [L]$, if $\nu_\ell = \nu\lambda_\ell$ for some $\nu \in \mathbb{R}$, then there exist $H \in \mathbb{R}$ and a policy ϕ^* satisfying, for $n \in \mathscr{N}_{i_j} \backslash \{C_{i_j}\}$, $j \in \mathscr{J}_\ell$, $\ell \in [L]$,

$$\alpha_{\ell,j}^{\phi^*}(n) = \begin{cases} 1, \text{ if } \nu_\ell < \max\limits_{\substack{n'=n+1, \\ n+2,\ldots,C_{i_j}}} \frac{\lambda_\ell}{\mu_{i_j}(n')}\Big(R_{i_j}(n') - g_j^*(\boldsymbol{\nu})\Big), \\ a, \text{ if } \nu_\ell = \max\limits_{\substack{n'=n+1, \\ n+2,\ldots,C_{i_j}}} \frac{\lambda_\ell}{\mu_{i_j}(n')}\Big(R_{i_j}(n') - g_j^*(\boldsymbol{\nu})\Big), \\ 0, \text{ otherwise}, \end{cases} \tag{20}$$

where a is any value in $[0,1]$, $R_i(n) \coloneqq \mu_i(n) - e^*\varepsilon_i(n)$, and

$$\begin{aligned} g_j^*(\boldsymbol{\nu}) \coloneqq \max\nolimits_{\phi \in \tilde{\Phi}_1} \lim\nolimits_{T\to+\infty} \frac{1}{T}\mathbb{E}\int_0^T \Big(&R_{i_j}(N_j^\phi(t)) \\ &- \sum\nolimits_{\ell \in [L]} \nu_\ell \alpha_{\ell,j}^\phi(N_j^\phi(t))\Big)dt, \end{aligned} \tag{21}$$

such that, for all $h > H$, the policy ϕ^* is optimal for the maximization problem in (14).

The proof of Proposition 1 is provided in Appendix I. Here, the real number $g_j^*(\boldsymbol{\nu})$ is equal to the maximized average reward gained by the process $\{N_j^\phi(t), t \geq 0\}$ for $j \in [J]$, where the reward rate for state $N_j^\phi(t) = n$ is $\mu_j(n) - e^*\varepsilon_j^\phi(n) - \sum_{\ell \in [L]} \nu_\ell \alpha_{\ell,j}^\phi(n)$. For an MDP, the $g_j^*(\boldsymbol{\nu})$ in (20) is usually referred to as the *attached criterion* used to translate the maximization of the average reward to the maximization of the expected cumulative reward of the same process [35].

Equation (20) has the same form as (18) except that $g_j^*(\boldsymbol{\nu})$ is dependent on $\boldsymbol{\nu}$. Unlike the simplified case discussed in [23], the $g_j^*(\boldsymbol{\nu})$ cannot be offset or expressed in a closed form. This dependence between $g_j^*(\boldsymbol{\nu})$ and $\boldsymbol{\nu}$ significantly complicates the analysis of indexability and the computation of the indices and prevents the same technique in [23] from being applied directly. For the purpose of the server farm problem, instead of seeking perfect Whittle indexability, we consider a less stringent property referred to as the *asymptotic indexability*.

Definition 2: For $j \in [J]$, if there exist a real vector $\boldsymbol{\upsilon}_j^* \coloneqq (\upsilon_{\ell,j}^*(n) : \ell \in [L], n \in \mathscr{N}_{i_j})$, a policy ϕ^* with action variables $\alpha_{\ell,j}^{\phi^*}(n)$ ($n \in \mathscr{N}_{i_j} \backslash \{C_{i_j}\}$ and $\ell \in [L]$) satisfying (18), and $H > 0$ such that, for all $h > H$, ϕ^* is optimal for the maximization problem in (16), then we say that the process $\{N_j^\phi(t), t \geq 0\}$ is asymptotically indexable with indices $\boldsymbol{\upsilon}_j^*$.

Proposition 2: When the job sizes are exponentially distributed, if there exists $H > 0$ such that, for all $h > H$, an optimal solution ϕ^* for the maximization problem in (14) exists and satisfies (20), then, for any $j \in [J]$, the process $\{N_j^\phi(t), t \geq 0\}$ is asymptotically indexable.

The proof of Proposition 2 is provided in Appendix II. When the process $\{N_j^\phi(t), t \geq 0\}$ reduces to a standard bandit process with $L = 1$, asymptotic indexability indicates Whittle indexability with sufficiently large h. Similar to the Whittle indices, in a large system, for $\ell \in [L]$, $j \in [J]$,

and $n \in \mathscr{N}_{i_j}\backslash\{C_{i_j}\}$, the real number $\upsilon^*_{\ell,j}(n)$ represents the marginal reward of admitting a new job of class ℓ when there are n jobs being served by physical component j. For $\ell \in [L]$, $j \in [J]$ and $n \in \mathscr{N}_{i_j}\backslash\{C_{i_j}\}$, we refer to $\upsilon^*_{\ell,j}(n)$ as the *index* of state n of the process $\{N^{\phi}_j(t), t \geq 0\}$ for job class ℓ. Asymptotic indexability plays an important role in proposing an asymptotically optimal and scalable policy for the original problem defined in (9), (1), (2) and (3). From Propositions 1 and 2, when the job sizes are exponentially distributed and all computing components are energy-efficiently unimodal, if, for all $\ell \in [L]$, if $\nu_\ell = \nu\lambda_\ell$ for some $\nu \in \mathbb{R}$, then, for any $j \in [J]$, the process $\{N^{\phi}_j(t), t \geq 0\}$ is asymptotically indexable.

C. Existence of Indices

More importantly, we are interested in the exact values of the indices $\boldsymbol{\upsilon}^*$ that impose asymptotic indexability and further lead to a scalable, near-optimal policy for the original problem.

Proposition 3: When the job sizes are exponentially distributed and all computing components are energy-efficiently unimodal, if, for all $\ell \in [L]$, if $\nu_\ell = \nu\lambda_\ell$ for some $\nu \in \mathbb{R}$, then, for any $j \in [J]$, the process $\{N^{\phi}_j(t), t \geq 0\}$ is asymptotically indexable with indices satisfying

$$\upsilon^*_{\ell,j}(n) = \max_{\substack{n'=n+1,\\ n+2,\ldots,C_i}} \lambda_\ell\Big(1 - \frac{e^*\varepsilon_{i_j}(n') + \Gamma_j(\boldsymbol{\upsilon}^*_j(n))}{\mu_{i_j}(n')}\Big) \quad (22)$$

where $\boldsymbol{\upsilon}^*_j(n) = (\upsilon^*_{\ell,j}(n) : \ell \in [L])$, and

$$\Gamma_j(\boldsymbol{\upsilon}^*_j(n)) := \max_{n\in\mathscr{N}_{i_j}\backslash\{C_{i_j}\}} \boldsymbol{\pi}^{\psi_j(n)}_j \cdot \boldsymbol{r}^{\psi_j(n)}_j(\boldsymbol{\upsilon}^*_j(n), \cdot), \quad (23)$$

with policy $\psi_j(n)$ satisfying, for all $\ell \in \{\ell' \in [L] | j \in \mathscr{J}_{\ell'}\}$ and $n' \in \mathscr{N}_{i_j}\backslash\{C_{i_j}\}$, $\alpha^{\psi_j(n)}_{\ell,j}(n') = 1$ if $n' \leq n$; and $\alpha^{\psi_j(n)}_{\ell,j}(n') = 0$ otherwise.

The proposition is proved in Appendix III. The policy $\psi_j(n)$ satisfies constraints in (12) and thus is not affected by $\boldsymbol{\omega}$; we write the reward vector $\boldsymbol{r}^{\phi}_j(\boldsymbol{\nu}, \boldsymbol{\omega})$ as $\boldsymbol{r}^{\phi}_j(\boldsymbol{\nu}, \cdot)$ since it becomes independent of the second argument under $\psi_j(n)$.

The index values $\upsilon^*_{\ell,j}(n)$ mentioned in Proposition 3 can be obtained by solving Equations (22) and (23). Observing (23), function $\Gamma_j(\boldsymbol{\nu})$ is dependent on $j \in [J]$ and $\boldsymbol{\nu} \in \mathbb{R}^L$ only through i_j and the sum $\sum_{\ell:i_j\in\mathscr{I}_\ell} \nu_\ell$, respectively. We rewrite $\Gamma_j(\boldsymbol{\nu})$ as $\Gamma_j(\eta_j)$ with $\eta_j = \sum_{\ell:i_j\in\mathscr{I}_\ell} \nu_\ell$, and, for $\eta \in \mathbb{R}$, $i \in [I]$ and any $j \in \mathscr{J}_i$, define $\bar{\Gamma}^h_i(\eta^0) := \Gamma_j(h\eta^0)$. Let $\bar{\Gamma}^{h,\mathrm{EXP}}_i(\eta^0)$ represent the value of $\bar{\Gamma}^h_i(\eta^0)$ with assumed exponentially distributed job sizes. For $n \in \mathscr{N}_i\backslash\{C_i\}$ and $i \in [I]$, define

$$f^h_{i,n}(\eta^0) := \eta^0 + \hat{\lambda}^0_i \min_{\substack{n'=n+1,\\ n+2,\ldots,C_i}} \Big(\frac{\bar{\Gamma}^{h,\mathrm{EXP}}_i(\eta^0) + e^*\varepsilon_i(n')}{\mu_i(n')} - 1\Big) \quad (24)$$

where $\hat{\lambda}^0_i = \sum_{\ell:i\in\mathscr{I}_\ell} \lambda^0_\ell$. For $n \in \mathscr{N}_i\backslash\{C_i\}$, $j \in \mathscr{J}_i$, $i \in \mathscr{I}_\ell$ and $\ell \in [L]$, given $\eta^0_{i,n}$ satisfying $f^h_{i,n}(\eta^0_{i,n}) = 0$, we obtain that $\upsilon^*_{\ell,j}(n) = \bar{\upsilon}^*_{\ell,i}(n) := h\eta^0_{i,n}\lambda^0_\ell/\hat{\lambda}^0_i$. From (22) and (23), for any $i \in [I]$, $\upsilon^*_{\ell,j}(n)$ remains the same for all $j \in \mathscr{J}_i$. In this context, solving (22) and (23) is equivalent to finding a zero point of function $f^h_{i,n}(\eta^0)$. Proposition 3 ensures the existence of such a zero point when the sub-processes are asymptotically indexable. Without assuming asymptotic indexability, a solution also exists for $f^h_{i,n}(\eta^0) = 0$, so do the indices $\upsilon^*_{\ell,j}(n)$ satisfying (22) and (23).

Proposition 4: For $h \in \mathbb{N}_+ \cup \{+\infty\}$, $n \in \mathscr{N}_i\backslash\{C_i\}$, and $i \in [I]$, $f^h_{i,n}(\eta^0)$ is Lipschitz continuous in $\eta^0 \in \mathbb{R}$, and there exists $\eta^0_{i,n} \in \mathbb{R}$ such that $f^h_{i,n}(\eta^0_{i,n}) = 0$.

The proposition is proved in Appendix IV.

VI. Multiple Power Modes with Priorities

For the original problem described in (9), (1), (2) and (3), we propose a policy that prioritizes the components and jobs according to the descending order of the indices $\boldsymbol{\upsilon}^* := (\upsilon^*_{\ell,j}(n) : \ell \in [L], j \in [J], n \in \mathscr{N}_{i_j})$ satisfying (22) and (23) for all $j \in [J]$. These indices represent marginal rewards of serving the various jobs. Define a policy φ satisfying

$$a^{\varphi}_{j,\ell}(\boldsymbol{N}^{\varphi}(t)) = \begin{cases} 1, & \text{if } N^{\varphi}_j(t) < C_{i_j} \text{ and} \\ & j = \arg\max\limits_{j'\in\mathscr{J}_\ell}\Big[\frac{1}{h}\upsilon^*_{\ell,j'}(N^{\varphi}_j(t))\Big], \\ 0, & \text{otherwise,} \end{cases} \quad (25)$$

where $\upsilon^*_{\ell,j}(n)$ ($\ell \in [L]$,$j \in [J]$, $n \in \mathscr{N}_{i_j}$) is derived from (22) and (23), and

$$a^{\varphi}_{\ell,j_\ell}(\boldsymbol{N}^{\varphi}(t)) = 1 - \sum\nolimits_{j\in\mathscr{J}_\ell} a^{\varphi}_{j,\ell}(\boldsymbol{N}^{\varphi}(t)). \quad (26)$$

Here, the $\frac{1}{h}$ before $\upsilon^*_{\ell,j'}(N^{\varphi}_j(t))$ is used to keep the value finite for all $h \in \mathbb{N}_+ \cup \{+\infty\}$ and will not change the order of the indices $\boldsymbol{\upsilon}^*$. In (25), tie-breaking rules can be arbitrary if $\arg\max$ returns more than one argument. Note that all the theoretical results presented in this paper hold for arbitrary tie-breaking rules. Nonetheless, beyond the theoretical results, the policies φ with different tie-breaking rules can potentially have different performance. In Section VIII, for the numerical results, we will consider specific tie-breaking rules to complete the simulations and provide numerical setting details. Such a policy φ is a feasible policy in Φ; that is, the action variables determined by (25) and (26) satisfy constraints (1), (2) and (3). For implementation in a server farm system, φ is scalable with computational complexity at most linear to the number of physical components for each arriving job. Equation (25) indicates that, for an ℓ-job newly arrived at time t, policy φ always selects the component j with the highest index value $\upsilon^*_{\ell,j}(N^{\varphi}_j(t))$, among all components $j \in \mathscr{J}_\ell$ with at least a vacant slot ($N^{\varphi}_j(t) < C_{i_j}$). If all components $j \in \mathscr{J}_\ell$ are fully occupied, then, from (26), the virtual component j_ℓ is selected to reject this job. In Section VII, we prove in Proposition 7 that φ is asymptotically optimal under certain conditions.

A. Indices with Known Criterion

Recall that the server farm problem in (7), (1), (2) and (3) has been translated to the problem described in (9), (1), (2) and (3) by introducing a given real number $e^* \in \mathbb{R}$ that satisfies (8). As mentioned in Section V-C, from Proposition 4, with given $e^* \in \mathbb{R}$, the indices $\boldsymbol{\upsilon}^*$ satisfying (22) and (23) can always be obtained by solving $f^h_{i,n}(\eta^0) = 0$ ($n \in \mathscr{N}_i\backslash\{C_i\}$, $i \in [I]$). In particular, $\upsilon^*_{\ell,j}(n) = \bar{\upsilon}^*_{\ell,i_j}(n) = h\eta^0_{i_j,n}\lambda^0_\ell/\hat{\lambda}^0_{i_j}$ where $\eta^0_{i_j,n}$ satisfies $f^h_{i_j,n}(\eta^0_{i_j,n}) = 0$. Such zero points of $f^h_{i,n}$ ($n \in \mathscr{N}_i\backslash\{C_i\}$, $i \in [I]$) can be computed through a bisection method with a precision parameter $\epsilon > 0$. A pseudo-code

with detailed steps for the zero points of $f^h_{i,n}$ ($n \in \mathscr{N}_i \backslash \{C_i\}$, $i \in [I]$) is provided in [39, Algorithm 1].

Definition 3: For $e \in \mathbb{R}$, $h \in \mathbb{N}_+$, $\ell \in [L]$, $i \in [I]$, $j \in \mathscr{J}_i$, and $n \in \mathscr{N}_i \backslash \{C_i\}$, let $u_{\ell,i}(n,e)$ represent an estimate of $\frac{1}{h}\bar{\upsilon}^*_{\ell,i}(n)$, i.e., an estimate of the index $\frac{1}{h}\upsilon^*_{\ell,j}(n)$ of state n of the process $\{N^{\phi}_j(t), t \geq 0\}$ for job class ℓ with given $e^* = e$, through a bisection method with a precision parameter $\epsilon > 0$. Define $\boldsymbol{u}(e) \coloneqq (u_{\ell,i}(n,e) : n \in \mathscr{N}_i \backslash \{C_i\}, i \in \mathscr{I}_\ell, \ell \in [L])$.

If we simplify our problem by allowing e^* to be any given non-negative real number, the resulting objective defined in (9) becomes the maximization of the difference between the job throughput $\mathfrak{L}^{\phi}$ and the power consumption $\mathfrak{E}^{\phi}$ *weighted* by the known e^*. Such an objective is popular and has been widely considered and studied in the literature [24], [40], [41]. In this simplified problem, based on later results in Section VII, the policy φ with indices $\boldsymbol{\upsilon}^*$ approximated by $\boldsymbol{u}(e^*)$ is asymptotically optimal under certain conditions.

B. Approximating the Unknown Criterion

The value of e^* satisfying (8) is not known a priori. For the purpose of this paper, we need to obtain this specific e^*, for which the optimal solution maximizing objective (9) also maximizes (7). From the definition of e^* in (8), we obtain

$$\max\nolimits_{\phi \in \Phi} \left(\mathfrak{L}^{\phi} - e^* \mathfrak{E}^{\phi} \right) / h = 0, \tag{27}$$

where the maximization is subject to (1), (2) and (3), and the $\frac{1}{h}$ is used to keep the value at the left-hand side of (27) finite for all $h \in \mathbb{N}_+ \cup \{+\infty\}$. Following similar ideas of [11], e^* can be approximated by a bisection method with stopping condition (27). Nonetheless, as mentioned in Section IV, because of the complexity of solving the maximization in the left-hand side of (27) (or the right-hand side of (8)), the optimal solutions are intractable. We thus resort to scalable techniques that effectively approximate e^*.

Definition 4: For $e \in \mathbb{R}$, $h \in \mathbb{N}_+$, $\ell \in [L]$, $i \in [I]$, and $n \in \mathscr{N}_i \backslash \{C_i\}$, let $u^*_{\ell,i}(n,e) \in \mathbb{R}$ represent a solution of $f^h_{i,n}\big(\frac{\hat{\lambda}^0_i}{\lambda^0_\ell} u^*_{\ell,i}(n,e)\big) = 0$ with $e^* = e$. Let $\boldsymbol{u}^*(e) \coloneqq (u^*_{\ell,i}(n,e) : \ell \in [L], i \in [I], n \in \mathscr{N}_i \backslash \{C_i\})$.

From Proposition 4, $\boldsymbol{u}^*(e)$ also represents a vector of the estimates $\boldsymbol{u}(e)$ obtained through a bisection method with an ideal precision parameter $\epsilon \downarrow 0$. For the system with scaling parameter $h \in \mathbb{N}_+$, we consider a policy $\psi^h(\boldsymbol{\nu}, \boldsymbol{a}^h, e) \in \tilde{\Phi}_1$ with given $\boldsymbol{\nu} \in \mathbb{R}^L$, $e \in \mathbb{R}_0$ and $\boldsymbol{a}^h \in \{0,1\}^J$, satisfying, for $\ell \in [L]$, $j \in \mathscr{J}_\ell$ and $n \in \mathscr{N}_{i_j} \backslash \{C_{i_j}\}$,

$$\alpha^{\psi^h(\boldsymbol{\nu},\boldsymbol{a}^h,e)}_{\ell,j}(n) = \begin{cases} 1, & \text{if } \sum\limits_{\ell' : j \in \mathscr{J}_{\ell'}} \nu_{\ell'} < \sum\limits_{\ell' : j \in \mathscr{J}_{\ell'}} u^*_{\ell',i_j}(n,e), \\ a^h_j, & \text{if } \sum\limits_{\ell' : j \in \mathscr{J}_{\ell'}} \nu_{\ell'} = \sum\limits_{\ell' : j \in \mathscr{J}_{\ell'}} u^*_{\ell',i_j}(n,e), \\ 0, & \text{otherwise,} \end{cases} \tag{28}$$

$\alpha^{\psi^h(\boldsymbol{\nu},\boldsymbol{a}^h,e)}_{\ell,j}(C_{i_j}) = 0$ for $\ell \in [L]$ and $j \in \mathscr{J}_\ell$, and for $\ell \in [L]$,

$$\alpha^{\psi^h(\boldsymbol{\nu},\boldsymbol{a}^h,e)}_{\ell,j_\ell} = \max\Big\{0, 1 - \sum\nolimits_{j \in \mathscr{J}_\ell} \sum\nolimits_{n \in \mathscr{N}_{i_j}} \pi^{\psi^h(\boldsymbol{\nu},\boldsymbol{a}^h,e)}_j(n) \alpha^{\psi^h(\boldsymbol{\nu},\boldsymbol{a}^h,e)}_{\ell,j}(n) \Big\}, \tag{29}$$

where $\pi^{\phi}_j(n)$ is defined in Section V-A and represents the steady state probability of state $n \in \mathscr{N}_{i_j}$ under a policy ϕ. For such a policy $\psi^h(\boldsymbol{\nu}, \boldsymbol{a}^h, e)$, the process $\{N^{\psi^h(\boldsymbol{\nu},\boldsymbol{a}^h,e)}_j(t), t \geq 0\}$ is a birth-and-death process with state transition rates linear to h and finitely many states, leading to the existence of $\lim_{h\to+\infty} \boldsymbol{\pi}^{\psi^h(\boldsymbol{\nu},\boldsymbol{a}^h,e)}_j$. Define, for $\ell \in [L]$, $A^{h,\psi^h(\boldsymbol{\nu},\boldsymbol{a}^h,e)}_\ell \coloneqq \sum_{j \in \mathscr{J}_\ell \cup \{j_\ell\}} \sum_{n \in \mathscr{N}_{i_j}} \pi^{\psi^h(\boldsymbol{\nu},\boldsymbol{a}^h,e)}_j(n) \alpha^{\psi^h(\boldsymbol{\nu},\boldsymbol{a}^h,e)}_{\ell,j}(n)$, which is the expected average sum of the action variables, namely the left-hand side of (10), under the policy $\psi^h(\boldsymbol{\nu}, \boldsymbol{a}^h, e)$ in the asymptotic regime. Since $A^{\psi^h(\boldsymbol{\nu},\boldsymbol{a}^h,e)}_\ell$ is decreasing in $\boldsymbol{\nu}$ and increasing in $\boldsymbol{a}^h$, there exist $\boldsymbol{\nu} \in \mathbb{R}^L$ and $\boldsymbol{a}^h \in \{0,1\}^J$ such that $\lim_{h\to\infty} A^{h,\psi^h(\boldsymbol{\nu},\boldsymbol{a}^h,e)}_\ell = 1$ for all $\ell \in [L]$; that is, constraints (10) are satisfied by substituting the policy $\psi^h(\boldsymbol{\nu}, \boldsymbol{a}^h, e)$ for ϕ. More precisely, let $\mathscr{V}$ represent the set of $(\boldsymbol{\nu}, \boldsymbol{a}^h)$ such that $\lim_{h\to\infty} A^{h,\psi^h(\boldsymbol{\nu},\boldsymbol{a}^h,e)}_\ell = 1$. Define

$$(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{a}}^h) = \arg\max\nolimits_{(\boldsymbol{\nu},\boldsymbol{a}^h) \in \mathscr{V}} \lim\nolimits_{h\to\infty} \sum\nolimits_{j \in \mathscr{J}_\ell} \sum\nolimits_{n \in \mathscr{N}_{i_j}} \pi^{\psi^h(\boldsymbol{\nu},\boldsymbol{a}^h,e)}_j(n) \alpha^{\psi^h(\boldsymbol{\nu},\boldsymbol{a}^h,e)}_{\ell,j}(n). \tag{30}$$

For $\phi \in \tilde{\Phi}$, $h \in \mathbb{N}_+$ and $e \in \mathbb{R}$, define $\Gamma^{h,\phi}(e) \coloneqq \frac{1}{h} \sum_{j \in [J]} \boldsymbol{\pi}^{\phi}_j \cdot \boldsymbol{r}^{\phi,e}_j$, where $\boldsymbol{x} \cdot \boldsymbol{y}$ represents the inner product of vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, and $\boldsymbol{r}^{\phi,e}_j \coloneqq \big(\mu_{i_j}(n) - e\varepsilon_{i_j}(n) : n \in \mathscr{N}_{i_j}\big)$. The policy $\psi^h(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{a}}^h, e)$ is applicable to the relaxed problem but not necessarily to the original problem. We will discuss in Section VII that, when job sizes are exponentially distributed, for any given real number e^*, $\Gamma^{h,\varphi}(e^*)$ converges to $\Gamma^{h,\psi^h(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{a}}^h,e^*)}(e^*)$ as $h \to +\infty$. In the asymptotic regime, if the policy $\psi^h(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{a}}^h, e^*)$ is coincidentally optimal for the relaxed problem, it must also be optimal for the original problem because $\Gamma^{h,\varphi}(e^*) \leq \max_{\phi \in \Phi} \Gamma^{h,\phi} \leq \max_{\phi \in \tilde{\Phi}} \Gamma^{h,\phi}$. In this case, due to the simplicity of computing $\Gamma^{h,\psi^h(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{a}}^h,e)}(e)$, we can approximate the value of e^* satisfying (8) by utilizing the condition in (27). For $e \in \mathbb{R}$, let

$$\Gamma(e) \coloneqq \lim\nolimits_{h\to+\infty} \Gamma^{h,\psi^h(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{a}}^h,e)}(e). \tag{31}$$

Proposition 5: When job sizes are exponentially distributed, if, for a given $e \in \mathbb{R}$,

$$\lim\nolimits_{h\to\infty} \big| \Gamma^{h,\psi^h(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{a}}^h,e)}(e) - \max\nolimits_{\phi \in \tilde{\Phi}} \Gamma^{h,\phi}(e) \big| = 0, \tag{32}$$

then $\Gamma(e)$ is Lipschitz continuous and piece-wise linear in $e \in \mathbb{R}$, there exists a unique solution $e_0 \in \mathbb{R}$ for $\Gamma(e_0) = 0$, and, for the unique e_0 satisfying $\Gamma(e_0) = 0$,

$$e_0 = \lim\nolimits_{h\to+\infty} e^* \geq 0. \tag{33}$$

The proposition is proved in Appendix V. Note that e^* defined in (8) is dependent on the scaling parameter h. The proposition indicates that, if, for any given $e \in \mathbb{R}$, the policy $\psi^h(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{a}}^h, e)$ is optimal for the relaxed problem described in (9), (10), (12) and (13) in the asymptotic regime (that is, (32) is satisfied), then e^* can be approximated by the zero point e_0 asymptotically. When (32) does not hold, $\Gamma(e)$ may be discontinuous at some points and there may not exist e_0 satisfying $\Gamma(e_0) = 0$. For clarify, define

$$e_0 \coloneqq \inf\{e \in \mathbb{R} \mid \forall e' > e,\ \Gamma(e') < 0\}. \tag{34}$$

From Proposition 5, if job sizes are exponentially distributed and (32) holds, then e_0 defined in (34) coincides with the unique solution for $\Gamma(e) = 0$.

Condition 1: $A_\ell \leq 1$ in the asymptotic regime for all $\ell \in [L]$.

Condition 1 implies that the system is in *heavy traffic* with positive blocking probabilities of all job classes under policy ϕ_0 in the asymptotic regime.

Lemma 1: When job sizes are exponentially distributed, if the components are energy-efficiently unimodal and Condition 1 or $L = 1$, then (32) holds.

The lemma is based on Proposition 3 and proved in Appendix VI. From Lemma 1 and Proposition 5, when job sizes are exponentially distributed, if the components are energy-efficiently unimodal, and Condition 1 holds or $L = 1$, then (33) holds, e_0 defined in (34) is the unique solution for $\Gamma(e_0) = 0$, and $\Gamma(e)$ is Lipschitz continuous and piece-wise linear in $e \in \mathbb{R}$.

Given the simple form of $\psi^h(\boldsymbol{\nu}, \boldsymbol{a}^h, e)$ defined in (28) and (29), we can fit the values of $\bar{\boldsymbol{\nu}} \in \mathbb{R}^L$ and $\Gamma(e)$ to satisfy (10). See pseudo-code for fitting the values of $\bar{\boldsymbol{\nu}} \in \mathbb{R}^L$ and $\Gamma(e)$ in [39, Algorithm 2]. We can also estimate the value of e_0 through a bisection method with the precision parameter $\epsilon > 0$. See pseudo-code in [39, Algorithm 3]. Let $\bar{e}_0$ represent such an estimate of e_0. Lemma 1 and Proposition 5 guarantee that the estimate $\bar{e}_0$ is within $[e_0 - \epsilon, e_0 + \epsilon]$, where e_0 is the unique zero point of $\Gamma(e_0) = 0$, and e_0 is equal to e^* in the asymptotic regime, when the computing components are energy-efficiently unimodal and Condition 1 holds or $L = 1$.

Definition 5: For e_0 defined in (34), we construct a policy $\varphi(e_0)$ by substituting $\varphi(e_0)$ and $u^*_{\ell,i_j}(n, e_0)$ for φ and $\upsilon^*_{\ell,j}(n)$, respectively, in (25) and (26). We refer to it as the *Multiple Power Modes with Priorities (MPMP)* policy.

Similar to φ, defined in (25) and (26), the policy $\varphi(e_0)$ prioritizes physical components according to the descending order of $\boldsymbol{u}^*(e_0)$. This policy is applicable to the original problem, described in (9), (1), (2) and (3). From Lemma 1 and Proposition 5, if components are energy-efficiently unimodal, and Condition 1 holds or $L = 1$, MPMP asymptotically approaches φ, which will be proved to be optimal in the asymptotic regime in Proposition 7. For cases without assuming energy-efficient unimodality, Condition 1 or $L = 1$, numerical results are presented in Section VIII to demonstrate the effectiveness of MPMP.

In Algorithm 1, we provide a pseudo-code for implementing MPMP for each arrived job. In Algorithm 1, as an example, we select the component with the smallest label in the tie case for implementing MPMP. Recall that, based on the definition of $\varphi(e_0)$ (i.e., MPMP) above, the indices used for MPMP are expected to be $\boldsymbol{u}^*(e_0)$, considered as an ideal estimate with the precision parameter $\epsilon \to 0$. All the theoretical results presented in this paper apply to MPMP, i.e., $\varphi(e_0)$. For the numerical results in Section VIII, with slightly abused notation, we still refer to the policy output by Algorithm 1 as the MPMP policy although the indices are estimated with a small $\epsilon > 0$.

For any $e \in \mathbb{R}$ and its estimate $e + \epsilon$ with a small $\epsilon > 0$, the indices $\boldsymbol{u}^*(e)$ and $\boldsymbol{u}^*(e + \epsilon)$ may be completely different. Nonetheless, even if $\|\boldsymbol{u}^*(e) - \boldsymbol{u}^*(e + \epsilon)\|$ is large, the performance deviation of the resulting policies can still be negligible. For $e \in \mathbb{R}$, define a policy $\varphi(e)$ by substituting $\varphi(e)$ and $u^*_{\ell,i_j}(n, e)$ for φ and $\upsilon^*_{\ell,j}(n)$, respectively, in (25) and (26).

Proposition 6: When job sizes are exponentially distributed, for $e \in \mathbb{R}$, if (32) holds, then, for any $\epsilon > 0$, there exists a constant $C > 0$ such that

$$\lim_{h \to +\infty} \left| \Gamma^{h,\varphi(e)}(e) - \Gamma^{h,\varphi(e+\epsilon)}(e + \epsilon) \right| \leq C\epsilon. \tag{35}$$

The proposition is proved in Appendix VII. Together with Lemma 1, if the components are energy-efficiently unimodal, and Condition 1 holds or $L = 1$, then (35) holds. Although $\|\boldsymbol{u}^*(e) - \boldsymbol{u}^*(e + \epsilon)\|$ may be sensitive to e, the performance deviation of the resulting policies is negligible for sufficiently small ϵ. When the estimate $\bar{e}_0$ of e_0 is subject to a small $\epsilon > 0$, the performance deviation between policies $\varphi(\bar{e}_0)$ and $\varphi(e_0)$ (i.e., MPMP) is bounded by $C\epsilon$.

Input : The class label ℓ of an arrived job, the indices $u^*_{\ell,j}(n) \in \mathbb{R}$ for $j \in \mathscr{J}_\ell$ and $n \in \mathscr{N}_{i_j}$, and current system state $\boldsymbol{N}^{\text{MPMP}}(t)$ upon the arrival.
Output: The selected component j in $\mathscr{J}_\ell \cup \{j_\ell\}$ to accommodate this job.
Function `ImplementingMPMP`**:**
  $j \leftarrow j_\ell$ and $u \leftarrow -\infty$
  for $\forall j' \in \mathscr{J}_\ell$ **do**
    if $N^{MPMP}_{j'}(t) < C_{i_{j'}}$ *AND* $u^*_{\ell,j'}(N^{MPMP}_{j'}(t)) > u$ **then**
      $j \leftarrow j'$ and $u \leftarrow u^*_{\ell,j'}(N^{\text{MPMP}}_{j'}(t))$
    end
  end
return

Algorithm 1: Implementing MPMP with given indices.

The MPMP policy is scalable and applicable to the original server farm problem. Let $C = \sum_{i \in [I]} C_i$. The computational complexity of computing $\boldsymbol{u}(\bar{e}_0)$ is $O\big(P_2(P_1 C + C \ln C + IL)\big)$ where P_1 and P_2 are the depths of the convergence trees for the bisection processes implemented for estimating $\boldsymbol{u}^*(e_0)$ and e_0, respectively, and only dependent on the precision parameter ϵ. This complexity is linear in the number of clusters I and the number of job classes L, and log-linear in the capacity of each component C_i, resulting in a reasonably fast procedure of obtaining the estimated indices $\boldsymbol{u}(\bar{e}_0)$. Recall that the estimated indices $\boldsymbol{u}(\bar{e}_0)$ are pre-calculated, and the computational complexity for pre-computing $\boldsymbol{u}(\bar{e}_0)$ is different from that of implementing MPMP. As mentioned earlier in this section, for implementing MPMP, the computational complexity is at most linear to the number of physical components.

VII. Asymptotic Optimality

For given $e \in \mathbb{R}$, we say a policy $\phi \in \Phi$, applicable to the original problem described in (9), (1), (2) and (3) with substituted e for e^*, is *asymptotically optimal* if

$$\lim_{h \to \infty} \left| \Gamma^{h,\phi}(e) - \max_{\phi' \in \Phi} \Gamma^{h,\phi'}(e) \right| = 0. \tag{36}$$

If a policy is asymptotically optimal, it approaches optimality for the server farm problem as $h \to +\infty$. Recall that, when the parameter $e = e^*$ satisfies (8), a policy optimal for the problem described in (9), (1), (2) and (3) is also optimal for the server farm problem described in (7), (1), (2) and (3) aiming to maximize the ratio of the long-run average job throughput to the long-run average power consumption.

Definition 6: For any given $e \in \mathbb{R}$, let $\varphi(e)$ represent the policy described in (25) and (26) with substituted $\varphi(e)$ and $u^*_{\ell,i_j}(n, e)$ for φ and $\upsilon^*_{\ell,j}(n)$, respectively.

Proposition 7: For the problem described in (9), (1), (2) and (3) with given $e \in \mathbb{R}$ substituting for e^*, when job sizes are exponentially distributed, if (32) holds, then the policy $\varphi(e)$ is asymptotically optimal.

The proposition is proved in Appendix VIII. Together with Lemma 1, when job sizes are exponentially distributed, if the components are energy-efficiently unimodal, and Condition 1 holds or $L = 1$, then the policy $\varphi(e)$ is asymptotically optimal. The problem, described in (9), (1), (2) and (3) with given $e \in \mathbb{R}$ substituting for e^*, aims to maximize the difference between the long-run average job throughput and the long-run average energy consumption rate weighted by the given e. Since we assume quite general $e \in \mathbb{R}$ and $\mu_i(n)$ and $\varepsilon_i(n)$ for the component clusters, the average job throughput and energy consumption rate can be directly generalized as the average reward and cost of the process. This is a popular objective and has been widely used in the literature, such as [24], [40], [41]. For such an objective, the simple policy $\varphi(e)$ is asymptotically optimal under the provided conditions.

From Lemma 1 and Propositions 5 and 7, for the problem defined by (7), (1), (2) and (3), when job sizes are exponentially distributed, if the computing components are energy-efficiently unimodal, and Condition 1 holds or $L = 1$, then the policy $\varphi(e_0)$, i.e., the MPMP policy, is asymptotically optimal; that is, (36) holds with substituted $\varphi(e_0)$ for ϕ.

Asymptotic optimality implies that MPMP is approaching optimality as the system becomes large. Unlike previous work in [24], [25], Proposition 7 and the asymptotic optimality of MPMP apply to systems with a less stringent relationship between power consumption, service rate, and traffic load.

We now further discuss the relationship between the sub-optimality of the index policy $\varphi(e)$ and the scaling parameter h of the server farm system. We rank the state-component (SC) pairs (n, i), $n \in \mathscr{N}_i \backslash \{C_{i_j}\}$ and $i \in [I]$, according to the descending order of $\eta_{i,n}$ where $\eta_{i,n}$ satisfies $f^h_{i,n}(\eta_{i,n}) = 0$ with $f^h_{i,n}$ defined in (24). Recall that the indices $\upsilon^*_{\ell,j}(n) = \bar{\upsilon}^*_{\ell,i_j}(n) = \lambda^0_\ell \eta_{i_j,n}/\hat{\lambda}^0_{i_j}$ for $\ell \in [L]$, $j \in [J]$, and $n \in \mathscr{N}_{i_j} \backslash \{C_{i_j}\}$. Then, we place the remaining SC pairs (n, i) with $n = C_i$ for all $i \in [I]$ afterwards. To emphasize this ranking, we refer to the kth SC pair as SC pair k, where $k = 1, 2, \ldots, K$ for $K = \sum_{i\in[I]} |\mathscr{N}_i|$. Let $Z^{\phi,h}_k(t)$ represent the proportion of processes $\{N^\phi_j(t), t \geq 0\}$ ($j \in [J]$) that are in the kth SC pair at time t under policy ϕ; that is,

$$Z^{\phi,h}_k(t) = \frac{1}{J}\left|\{j \in [J] \mid i_j = i_k, N^\phi_j(t) = n_k\}\right|, \qquad (37)$$

where $i_k \in [I]$ and $n_k \in \mathscr{N}_{i_k}$ are the cluster and state labels of SC pair k. Recall that J is defined in Section III and is dependent on h. Define $\boldsymbol{Z}^{\phi,h}(t) \coloneqq (Z^{\phi,h}_k(t) : k \in [K])$. For any given $h \in \mathbb{N}_+$, the stochastic process $\{\boldsymbol{N}^\phi(t), t \geq 0\}$ can be translated to the process $\{\boldsymbol{Z}^{\phi,h}(t), t \geq 0\}$. Consider a server farm that starts with no job (that is, $\boldsymbol{N}^\phi(0) = \boldsymbol{0}$) and, correspondingly, we have $\boldsymbol{Z}^{\phi,h}(0) = \boldsymbol{z}^0$ for this empty system. Let $\mathscr{Z} \coloneqq [0, 1]^K$ represent a probability simplex.

Proposition 8: When job sizes are exponentially distributed, for given $e \in \mathbb{R}$, there exists $\boldsymbol{z}^{\varphi(e)} \in \mathscr{Z}$ such that, for any $\delta > 0$, there exist $s > 0$ and $H > 0$ satisfying, for all $h > H$,

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \mathbb{P}\Big\{\|\boldsymbol{Z}^{\varphi(e),h}(t) - \boldsymbol{z}^{\varphi(e)}\| > \delta\Big\}dt \leq e^{-sh}, \qquad (38)$$

where $\boldsymbol{Z}^{\varphi(e),h}(0) = \boldsymbol{z}^0$.

The proposition is proved in Appendix IX. Proposition 8 indicates that, under the policy $\varphi(e)$, the underlying stochastic process $\{\boldsymbol{Z}^{\varphi(e),h}(t), t \geq 0\}$ converges to a global attractor $\boldsymbol{z}^{\varphi(e)}$ almost surely as $h \to \infty$, and, more importantly, the deviation between $\boldsymbol{Z}^{\varphi(e),h}(t)$ and $\boldsymbol{z}^{\varphi(e)}$ is diminishing exponentially in h. Define $\boldsymbol{r}(e) \coloneqq (\mu_{i_k}(n_k) - e\varepsilon_{i_k}(n_k) : k \in [K])$. Then, for given $h \in \mathbb{N}_+$ and policy $\phi \in \Phi$, the normalized long-run average reward

$$\Gamma^{\phi,h}(e) = \sum\nolimits_{i\in[I]} M^0_i \boldsymbol{r}(e) \cdot \lim_{t\to\infty} \mathbb{E}[\boldsymbol{Z}^{\phi,h}(t)]. \qquad (39)$$

From Proposition 8, when job sizes are exponentially distributed, $\lim_{h\to\infty} \Gamma^{\varphi(e),h}(e) = \sum_{i\in[I]} M^0_i \boldsymbol{r}(e) \cdot \boldsymbol{z}^{\varphi(e)}$. If $\boldsymbol{z}^{\varphi(e)}$ coincides with an optimal point of the relaxed problem, described in (9), (10), (12) and (13) with e substituting for e^*, in the asymptotic regime, then the policy $\varphi(e)$ is asymptotically optimal. Together with asymptotic indexability discussed in Section V, Proposition 8 can lead to Proposition 7. Apart from asymptotic optimality, Proposition 8 also implies that, for a system with large h, if $\varphi(e)$ is asymptotically optimal, the performance deviation between $\varphi(e)$ and optimality in the asymptotic regime diminishes exponentially as $h \to \infty$. This conclusion extends [23, Proposition 2] to the more generalized server farm model discussed in this paper, and, as a straightforward result of Proposition 8, when job sizes are exponentially distributed, (38) also applies to the MPMP policy by setting $e = e_0$.

VIII. Numerical Results

Without assuming energy-efficient unimodality or Condition 1, we numerically demonstrate the effectiveness of MPMP by comparing it with two baseline policies. For all simulation results, the 95% confidence intervals based on the Student t-distribution are within 3% of the observed mean. We consider exponentially distributed job sizes in Section VIII-A and other job-size distributions in Section VIII-B.

A. Effectiveness of MPMP

We consider a scenario with Google cluster traces of job arrivals in 2011 [42], [43], where there are 12,500 physical components with arriving jobs classified into four groups ($L = 4$). We divide the system into ten clusters ($I = 10$), each of which includes 1,250 components; that is, setting $M^0_i = 1$ for all $i \in [I]$ with scaling parameter $h = 1250$. In this scenario, we no longer assume Poisson arrivals, and instead, consider the job arrivals of the Google cluster traces. The capacities of physical components are set to ten ($C_i = 10$) for all $i \in [I]$. For other detailed settings see Appendix X.

In Figure 1, we present the energy efficiencies of MPMP, PAS, and JSQ, averaged over an hour, where MPMP significantly outperforms PAS and JSQ. The total energy efficiency of MPMP is around 13% higher than that of PAS. For the same settings, we plot in Figure 2, the job throughput, normalized by the scaling parameter h, of MPMP, PAS, and JSQ. Given

the clear advantages of MPMP against PAS and JSQ with respect to energy efficiency, they still achieve almost the same job throughputs. It is because the blocking probabilities of the three policies are likely to be negligible, although the capacity for each component is relatively small.

As mentioned in Section VI, all the theoretical results in this paper apply to the index policy $\varphi(e)$ with arbitrary tie-breaking rules. Note that different tie-breaking rules may lead to different performance. Consider a tie-breaking rule that selects the component with the lowest label, and refer to it as the *Lowest-Label Tie-Breaking* (LLTB). In Figure 3, we explore the effects of different tie-breaking rules for the same setting as before. We consider another tie-breaking rule that always selects the component with the least number of holding jobs in the tie-breaking cases and refer to it as the *Shortest-Queue Tie-Breaking* (SQTB). In Figure 3, we demonstrate the energy efficiency of MPMP with LLTB and SQTB and observe that LLTB achieves slightly higher energy efficiency than that of SQTB. The total energy efficiency for LLTB is around 5% higher than that of SQTB. Based on Proposition 8, when job sizes are exponentially distributed, MPMP with different tie-breaking rules approaches the same performance as h increases and, for a large system, the performance deviation between different tie-breaking rules diminishes exponentially in h. The 5% difference in Figure 3 between SQTB and LLTB is marginal, considering our 95% confidence intervals. This paper focuses on scalable and asymptotically optimal policies in large-scale server farms. A thorough discussion on the effects of different tie-breaking rules in relatively small and practical systems is a fundamentally interesting topic on its own, which is beyond the scope of this paper.

B. Sensitivity

The theoretical results presented in this paper are based on exponential job-size distributions. From past studies [44], [45], real-world online applications exhibit job sizes that have heavy-tailed distributions. Here, we numerically demonstrate the performance of MPMP considering a heavy-tailed job-size (Pareto) distribution, as well as a mixed version with different job-size distributions for different job classes. In particular, we consider simulations involving three non-exponential job size distributions with unit mean: deterministic, Pareto with shape parameter 2.001 and Pareto with shape parameter 1.98. We refer to the Pareto distributions with shape parameters 2.001 and 1.98 as Pareto-F and Pareto-INF for short, as they have finite and infinite variances, respectively. Apart from the above-mentioned distributions, define a *mixed* case where different job classes have different job-size distributions.

Consider a server farm with ten clusters ($I = 10$) and four job classes ($L = 4$), where the peak service rate for each cluster (that is, $\mu_i(C_i)$ for $i \in [I]$) is uniformly randomly generated from $[10, 15]$ and the capacities C_i are set 5 for all $i \in [I]$. For cluster $i \in [I]$, we uniformly randomly generate the energy efficiency of its fully-occupied component, $\mu_i(C_i)/\varepsilon_i(C_i)$, from $[0.5, 1]$, and set the power consumption of its idle component $\varepsilon_i(0)$ to be $0.3\varepsilon_i(C_i)(0.9 - 0.1i)$. For other states $n \in \mathcal{N}_i \backslash \{C_i\}$ of cluster i, the service and energy consumption rates are obtained by setting $\mu_i(n) = \mu_i(n+1)\frac{n}{n+1}$ and

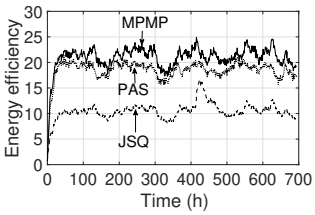

Fig. 1. Energy efficiency of MPMP, PAS and JSQ.

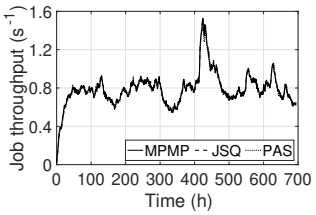

Fig. 2. Job throughput of MPMP, PAS and JSQ.

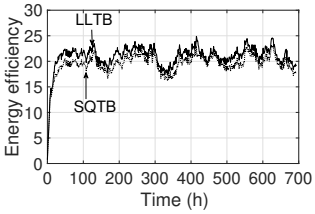

Fig. 3. Energy efficiency of MPMP with different tie-breaking rules.

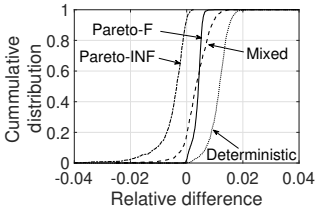

Fig. 4. Relative difference of energy efficiency with different job size distributions.

$\varepsilon_i(n) = (\varepsilon_i(n+1) - \varepsilon_i(0))\sqrt{\frac{n}{n+1}} + \varepsilon_i(0)$, respectively. In this case, a higher service rate indicates higher power consumption and higher energy efficiency, which follows realistic situations in [46], [47]. We take the scaling parameter $h = 10$ and the number of components in each cluster $M_i^0 = 1$ for all $i \in [I]$.

For job class $\ell \in [L]$, we uniformly randomly generate an integer $\kappa_\ell \in [I]$ as the number of clusters involving available components for ℓ-jobs, and then randomly select κ_ℓ clusters: all the components within the selected clusters are available components for ℓ-jobs and join the set $\mathcal{J}_\ell$. Define the *normalized offered traffic* of job class $\ell \in [L]$ as $\rho_\ell := \lambda_\ell / \sum_{j \in \mathcal{J}_\ell} \mu_{i_j}(C_{i_j})$ and the *relative difference* of policy ϕ_1 to ϕ_2 ($\phi_1, \phi_2 \in \Phi$) with respect to energy efficiency as $(\mathfrak{L}^{\phi_1}/\mathfrak{E}^{\phi_1} - \mathfrak{L}^{\phi_2}/\mathfrak{E}^{\phi_2})/(\mathfrak{L}^{\phi_2}/\mathfrak{E}^{\phi_2})$. In this subsection, we set $\rho_\ell = 0.35$ for all $\ell \in [L]$ and compute the arrival rates of different job classes.

Define $\mathfrak{T}^{\text{MPMP},\mathscr{D}}$ and $\mathfrak{E}^{\text{MPMP},\mathscr{D}}$ as the long-run average job throughput and power consumption under MPMP with job-size distribution $\mathscr{D}$, respectively. In Figure 4, we present the cumulative distribution of the relative difference of energy efficiency with job size distribution $\mathscr{D}$ from the one with exponentially distributed job sizes; that is, the cumulative distribution of

$$\frac{\mathfrak{T}^{\text{MPMP},\mathscr{D}}/\mathfrak{E}^{\text{MPMP},\mathscr{D}} - \mathfrak{T}^{\text{MPMP},exponential}/\mathfrak{E}^{\text{MPMP},exponential}}{\mathfrak{T}^{\text{MPMP},exponential}/\mathfrak{E}^{\text{MPMP},exponential}},$$

with $\mathscr{D}$ = deterministic, Pareto-F, Pareto-INF and mixed. For the case of mixed, we set the job-size distributions for the job classes 1-4 as deterministic, exponential, Pareto-F, and Pareto-INF, respectively. In Figure 4, we observe that the relative differences of energy efficiency for all the tested simulation runs are within $\pm 3\%$, indicating similar energy efficiencies for tested $\mathscr{D}$ and the exponential case. It follows that the energy efficiency of MPMP is not very sensitive to the tested distributions, including the heavy-tailed Pareto-INF.

IX. Conclusions

We have proposed the MPMP policy that always prioritizes physical components with the highest indices satisfying (22).

The indices are computable within linear time in the number of clusters I and the number of job classes L. It is log-linear in the capacity of each component, and the implementation of the MPMP policy is at most linear in the number of physical components. It follows that MPMP is scalable as we have demonstrated its applicability to a large-scale server farm with tens or hundreds of thousands of abstracted servers.

When job sizes are exponentially distributed and all the components are energy-efficiently unimodal, we have proved that, if $L = 1$ or Condition 1 holds, MPMP approaches optimality as the scaling parameter h tends to infinity; that is, MPMP approaches optimality as the numbers of components in clusters and the arrival rates of jobs increase proportionately to infinity. We have provided an analysis of the entire system, including discussions on the indexability and the global attractor for proving asymptotic optimality in the continuous-time case. For a large system, we have proved that the performance deviation between MPMP and an optimal solution in the asymptotic regime diminishes exponentially in the scaling parameter h. That is, MPMP becomes already close to optimality in a relatively small system.

For the non-asymptotic regime without assuming energy-efficient unimodality and the heavy traffic condition, we have numerically demonstrated the effectiveness of MPMP by comparing it to JSQ and PAS. When the job throughputs are compatible, MPMP has shown substantial advantages against the baseline policies with respect to energy efficiency. We have further investigated the performance of MPMP considering different job-size distributions and demonstrated that MPMP is robust in all the cases we tested.

APPENDIX I

For $\ell \in [L]$, $j \in \mathscr{J}_\ell$, let $\omega_{\ell,j} > -\nu_\ell$ all the time, such that an optimal solution of sub-problem (16) always has $\alpha^{\phi}_{\ell,j}(C_{i_j}) = 0$: constraints (12) are satisfied. In this context, we replace $\boldsymbol{r}^{\phi}_j(\boldsymbol{\nu},\boldsymbol{\omega})$ with $\boldsymbol{r}^{\phi}_j(\boldsymbol{\nu}) \coloneqq \boldsymbol{r}^{\phi,e^*}_j(\boldsymbol{\nu})$ (defined in Section VI-B) for the remainder of this appendix.

Consider the underlying stochastic process of the problem defined in (16) for component $j \in [J]$: it is a birth-and-death Markov process $\{N^{\phi}_j(t), t \geq 0\}$, for which $N^{\phi}_j(t)$ represents the number of jobs being served by this component at time t and $r^{\phi}_{j,n}(\boldsymbol{\nu})$ ($n \in \mathscr{N}_{i_j}$) is the reward rate in state n. We refer to this process $\{N^{\phi}_j(t), t \geq 0\}$ as the *sub-process* associated with component j or sub-process j.

Let $V^{\phi,\boldsymbol{\nu}}_j(n)$ and $T^{\phi,\boldsymbol{\nu}}_j(n)$ represent the expected accumulated reward and time, respectively, of sub-process j starting from state n and ending in state 0 under policy ϕ, where the reward rate in state n is $r^{\phi}_{j,n}(\boldsymbol{\nu})$ and $V^{\phi,\boldsymbol{\nu}}_j(0) = T^{\phi,\boldsymbol{\nu}}_j(n) \equiv 0$. For given $g \in \mathbb{R}$, define $V^{\phi,g,\boldsymbol{\nu}}_j(n) = V^{\phi,\boldsymbol{\nu}}_j(n) - gT^{\phi,\boldsymbol{\nu}}_j(n)$ and $V^{g,\boldsymbol{\nu}}_j(n) = \max_{\phi\in\tilde{\Phi}_1} V^{\phi,g,\boldsymbol{\nu}}_j(n)$.

Lemma 2: When job sizes are exponentially distributed, there exists a $g \in \mathbb{R}$ such that, for any policy $\phi^* \in \tilde{\Phi}_1$ that is optimal for the maximization problem in (14),

$$\alpha^{\phi^*}_{\ell,j}(n) = \begin{cases} 1, \text{ if } \nu_\ell < \lambda_\ell \left(V^{g,\boldsymbol{\nu}}_j(n+1) - V^{g,\boldsymbol{\nu}}_j(n)\right), \\ a, \text{ if } \nu_\ell = \lambda_\ell \left(V^{g,\boldsymbol{\nu}}_j(n+1) - V^{g,\boldsymbol{\nu}}_j(n)\right), \\ 0, \text{ if } \nu_\ell > \lambda_\ell \left(V^{g,\boldsymbol{\nu}}_j(n+1) - V^{g,\boldsymbol{\nu}}_j(n)\right), \end{cases} \quad (40)$$

for recurrent states $n \in \mathscr{N}_{i_j}\backslash\{C_{i_j}\}$, $j \in \mathscr{J}_\ell$, $\ell \in [L]$, where a can be any real number in $[0,1]$.

Proof. As described in (15), the maximization problem in (14) consists of $J+L$ independent sub-problems: the J sub-problems described in (16) and the L sub-problems described in (17) subject to $\alpha^{\phi^*}_{\ell,j}(C_{i_j}) \equiv 0$ ($j \in [J]$,$\ell \in [L]$) and $\alpha^{\phi^*}_{\ell,j}(n) \equiv 0$ ($n \in \mathscr{N}_{i_j}\backslash\{C_{i_j}\}$, $j \neq \mathscr{J}_\ell$, $\ell \in [L]$). Equation (40) is obtained by solving sub-problems in (16).

For a constant $g = g^*_j(\boldsymbol{\nu})$ (defined in (21)), $n \in \mathscr{N}_{i_j}\backslash\{C_{i_j}, 0\}$, $j \in \mathscr{J}_\ell$ and $\ell \in [L]$, equation (40) is obtained by solving the Bellman equations for $V^{g,\boldsymbol{\nu}}_j(n)$ ($n \in \mathscr{N}_{i_j}\backslash\{0\}$) in the cases with $\alpha^{\phi^*}_{\ell,j}(0) = 1$ for at least one $\ell \in \{\ell' \in [L] : j \in \mathscr{J}_{\ell'}\}$. We refer to the extended version of this paper [39, Equation (50)] for the description of the Bellman equations for $V^{g,\boldsymbol{\nu}}_j(n)$. If $\alpha^{\phi^*}_{\ell,j}(0) = 0$ for all $\ell \in \{\ell' \in [L] : j \in \mathscr{J}_{\ell'}\}$, then sub-process j will stay in state 0 all the time whatever the action variables of other states will be. Hence, the policy satisfying (40) is still optimal.

It remains to discuss (40) for $n = 0$. For $\ell \in \{\ell' \in [L] | j \in \mathscr{J}_{\ell'}\}$, $j \in [J]$, let $\bar{\lambda}^{\ell}_j = \sum_{\ell':j\in\mathscr{J}_{\ell'},\ell'\neq\ell} \lambda_{\ell'} \alpha^{\phi^*}_{\ell',j}(0)$ and $\bar{\nu}^{\ell}_j = \sum_{\ell':j\in\mathscr{J}_{\ell'},\ell'\neq\ell} \nu_{\ell'} \alpha^{\phi^*}_{\ell',j}(0)$. If $\bar{\lambda}^{\ell}_j > 0$, from the Bellman equation, we obtain that (40) holds for $n = 0$.

It remains to prove (40) for $n = 0$ when $\bar{\lambda}^{\ell}_j = 0$. Let $\bar{r}^{\phi}_j$ represent the average reward received when the process $\{N^{\phi}_j(t), t \geq 0\}$ is in the states $n \in \mathscr{N}_{i_j}\backslash\{0\}$ under ϕ. Since (40) holds for $n \in \mathscr{N}_{i_j}\backslash\{C_{i_j}\}$, $\bar{r}^{\phi^*}_j$ with any optimal solution $\phi^* \in \tilde{\Phi}_1$ is independent from the value of $\boldsymbol{\alpha}^{\phi}_j(0)$. When $\bar{\lambda}^{\ell}_j = 0$, if $R_j(0) \leq \frac{1}{1+\frac{\lambda_\ell}{\mu_j(1)}} \left(R_j(0) - \nu_\ell + \frac{\lambda_\ell}{\mu_j(1)} \bar{r}^{\phi^*}_j\right)$ with $R_j(n) = \mu_j(n) - e^* \varepsilon_j(n)$, then $\alpha^{\phi^*}_{\ell,j}(0) = 1$; that is,

$$\nu_\ell \leq \lambda_\ell (\bar{r}^{\phi^*}_j - R_j(0))/\mu_j(1). \quad (41)$$

Because $g^*_j(\boldsymbol{\nu}) \geq R_j(0)$ (based on its definition in (21)) and $V^{g^*_j(\boldsymbol{\nu}),\boldsymbol{\nu}}_j(1) = \lambda_\ell(\bar{r}^{\phi^*}_j - g^*_j(\boldsymbol{\nu}))/\mu_j(1)$ (Bellman equation), if $\nu_\ell \leq \lambda_\ell V^{g^*_j(\boldsymbol{\nu}),\boldsymbol{\nu}}_j(1)$ then (41) holds: $\alpha^{\phi^*}_{\ell,j}(0) = 1$.

We then show that $\nu_\ell \leq \lambda_\ell V^{g^*_j(\boldsymbol{\nu}),\boldsymbol{\nu}}_j(1)$ is necessary for $\alpha^{\phi^*}_{\ell,j}(0) = 1$. If $\alpha^{\phi^*}_{\ell,j}(0) = 1$, (41) holds, and $g^*_j(\boldsymbol{\nu}) = \big(R_j(0) - \nu_\ell + \lambda_\ell \bar{r}^{\phi^*}_j/\mu_j(1)\big)\big(1 + \lambda_\ell/\mu_j(1)\big)$. Together with (41), we obtain $\nu_\ell \leq \lambda_\ell V^{g^*_j(\boldsymbol{\nu}),\boldsymbol{\nu}}_j(1)$. This proves the lemma. ∎

Proof of Proposition 1. Consider a policy $\phi^* \in \tilde{\Phi}_1$ satisfying (40). There exists another policy ϕ_1 such that, for $n \in \mathscr{N}_{i_j}\backslash\{C_{i_j}\}$, $\alpha^{\phi_1}_j(n) = \alpha^{\phi^*}_j(n)$, if $n < m$ where $m = \min\{m' \in \mathscr{N}_{i_j} | \alpha^{\phi^*}_j(m') = 0\}$; otherwise, $\alpha^{\phi_1}_j(n) = 0$. Since all states $n > m$ under policy ϕ^* are transient, policy ϕ_1 leads to the same stationary distribution as ϕ^*, which is optimal for the maximization in (14).

For state $n \in \mathscr{N}_{i_j}\backslash\{C_{i_j}\}$ ($j \in [J]$), let $m^*_j(n) \in \{n, n+1, \ldots, C_{i_j}-1\}$ represent the state such that $\alpha^{\phi^*}_j(m^*_j(n)+1) = 0$, and, if $m^*_j(n) \geq 1$, $\alpha^{\phi^*}_j(n+1) = \ldots = \alpha^{\phi^*}_j(m^*_j(n)) = 1$. If $\alpha^{\phi^*}_j(n') = 0$ for all $n' \geq n+1$, then $m^*_j(n) = n$. From Lemma 2, for $n \in \mathscr{N}_{i_j}\backslash\{C_{i_j}\}$, we obtain that, if $m^*_j(n) = n$,

then

$$V_j^{g_j^*(\boldsymbol{\nu}),\boldsymbol{\nu}}(n+1) - V_j^{g_j^*(\boldsymbol{\nu}),\boldsymbol{\nu}}(n) = \frac{R_j(n+1) - g_j^*(\boldsymbol{\nu})}{\mu_{i_j}(n+1)}; \quad (42)$$

otherwise,

$$V_j^{g_j^*(\boldsymbol{\nu}),\boldsymbol{\nu}}(n+1) - V_j^{g_j^*(\boldsymbol{\nu}),\boldsymbol{\nu}}(n) = \frac{R_j(n+1) - g_j^*(\boldsymbol{\nu})}{\mu_{i_j}(n+1)} + \sum_{k=1}^{m_j^*(n)-n} \prod_{\iota=1}^{k} \frac{\hat{\lambda}_{i_j}}{\mu_{i_j}(n+\iota)} \Big(\frac{R_j(n+k+1) - g_j^*(\boldsymbol{\nu})}{\mu_{i_j}(n+k+1)} - \nu \Big), \quad (43)$$

where $\hat{\lambda}_{i_j} := \hat{\lambda}_{i_j}^0 h = \sum_{\ell: i_j \in \mathscr{I}_\ell} \lambda_\ell$, and $g_j^*(\boldsymbol{\nu})$ is a given real number satisfying (21). We refer to a detailed exposition for achieving (43) in the extended version of this paper [39, Equations (59)-(64)]. Recall that $V_j^{g_j^*(\boldsymbol{\nu}),\boldsymbol{\nu}}(0) \equiv 0$. If

$$\nu \geq \max_{\substack{n'=n+1,\\ n+2,\ldots,C_{i_j}}} \big(R_j(n') - g_j^*(\boldsymbol{\nu})\big)/\mu_{i_j}(n'), \quad (44)$$

where $R_j(n) = \mu_j(n) - e^* \varepsilon_j(n)$ for $n \in \mathscr{N}_{i_j}$, then, from (42)-(43), $V_j^{g_j^*(\boldsymbol{\nu}),\boldsymbol{\nu}}(n+1) - V_j^{g_j^*(\boldsymbol{\nu}),\boldsymbol{\nu}}(n) \leq \nu$; together with Lemma 2, $\alpha_j^{\phi^*}(n) = 0 = \alpha_j^{\phi_1}(n)$. It remains to prove that there exists a $H \in \mathbb{R}$ such that, for all $h > H$ and $n \in \mathscr{N}_{i_j} \backslash \{C_{i_j}\}$, if (44) does not hold, $\alpha_j^{\phi_1}(n) = 1$; equation (20) is then led by Lemma 2 and the continuity of $V_j^{g_j^*(\boldsymbol{\nu}),\boldsymbol{\nu}}(n+1) - V_j^{g_j^*(\boldsymbol{\nu}),\boldsymbol{\nu}}(n) - \nu$ in ν, where $\boldsymbol{\nu} = \boldsymbol{\lambda}\nu$.

For $n = C_{i_j} - 1$, if (44) does not hold, then, from Lemma 2, $V_j^{g_j^*(\boldsymbol{\nu}),\boldsymbol{\nu}}(C_{i_j}) - V_j^{g_j^*(\boldsymbol{\nu}),\boldsymbol{\nu}}(C_{i_j} - 1) = \frac{R_j(C_{i_j}) - g_j^*(\boldsymbol{\nu})}{\mu_{i_j}(C_{i_j})} > \nu$; that is, $\alpha_j^{\phi^*}(C_{i_j} - 1) = 1$ and (20) holds for $n = C_{i_j} - 1$.

We prove the remaining case with $n < C_{i_j} - 1$ through iterations. Assume that, for sufficiently large h, (20) holds for $n+1, n+2, \ldots, C_{i_j} - 1$. If $m_j^*(n) > n$, then, $\nu \leq \frac{R_j(m_j^*(n)+1) - g_j^*(\boldsymbol{\nu})}{\mu_{i_j}(m_j^*(n)+1)}$. By (42)-(43) and Lemma 2, if $m_j^*(n) > n$ and

$$\nu < \frac{\frac{R_j(n+1) - g_j^*(\boldsymbol{\nu})}{\mu_{i_j}(n+1)} + \sum_{k=1}^{m_j^*(n)-n} \prod_{\iota=1}^{k} \frac{\hat{\lambda}_{i_j}}{\mu_{i_j}(n+\iota)} \frac{R_j(n+k+1) - g_j^*(\boldsymbol{\nu})}{\mu_{i_j}(n+k+1)}}{1 + \sum_{k=1}^{m_j^*(n)-n} \prod_{\iota=1}^{k} \frac{\hat{\lambda}_{i_j}}{\mu_{i_j}(n+\iota)}}, \quad (45)$$

then $\alpha_j^{\phi^*}(n) = 1$. Let $\upsilon_j(n)$ represent the value of the right-hand side of (45). If (44) does not hold and component j is energy-efficiently unimodal, for any $\sigma > 0$, there exits $H \in \mathbb{R}$ such that, for all $h > H$, $\upsilon_j(n) > \nu$. Accordingly, $\alpha_j^{\phi_1}(n) = \alpha_j^{\phi^*}(n) = 1$. If $m_j^*(n) = n$ and (44) does not hold, then, together with the Bellman equation for $V_j^{g_j^*(\boldsymbol{\nu}),\boldsymbol{\nu}}(m_j^*(n))$ and Lemma 2, $\alpha_j^{\phi^*}(n) = 1$. This proves the proposition. ∎

APPENDIX II

For $j \in [J]$ and $n \in \mathscr{N}_{i_j} \backslash \{C_{i_j}\}$, define, for $\nu \in \mathbb{R}$,

$$\bar{f}_{j,n}(\nu) := \nu + \min_{\substack{n'=n+1,\\ n+2,\ldots,C_{i_j}}} \Big(\frac{g_j^*(\nu\boldsymbol{\lambda})}{\mu_j(n')} - \frac{R_j(n')}{\mu_j(n')} \Big). \quad (46)$$

Lemma 3: For any $j \in [J]$, $n \in \mathscr{N}_{i_j} \backslash \{C_{i_j}\}$, there exists $\nu \in \mathbb{R}$ such that $\bar{f}_{j,n}(\nu) = 0$. In particular, if job sizes are exponentially distributed and all computing components are energy-efficiently unimodal, there exists $H \in \mathbb{R}$ and $\upsilon_{j,n}^h \in \mathbb{R}$ such that, for all $h > H$,

$$\bar{f}_{j,n}(\nu) \begin{cases} > 0, \text{ if } \nu > \upsilon_{j,n}^h, \\ = 0, \text{ if } \nu = \upsilon_{j,n}^h, \\ < 0, \text{ otherwise.} \end{cases} \quad (47)$$

Proof. From the definition, $g_j^*(\nu\boldsymbol{\lambda})$ is piece-wise linearly decreasing and continuous in $\nu \in \mathbb{R}$. From the definition in (46), $\bar{f}_{j,C_{i_j}-1}(\nu)$ is piece-wise linear and continuous in $\nu \in \mathbb{R}$. For $\nu \in \mathbb{R}$, $\frac{d^-}{d\nu} \bar{f}_{j,C_{i_j}-1}(\nu) = 1 + \frac{d^-}{d\nu} \frac{g_j^*(\nu\boldsymbol{\lambda})}{\mu_j(C_{i_j})} > 0$, where $\frac{d^-}{d\nu}$ takes the left derivative. Together with the piece-wise linearity of $\bar{f}_{j,C_{i_j}-1}(\nu)$, $\bar{f}_{j,C_{i_j}-1}(\nu)$ is monotonically increasing in $\nu \in \mathbb{R}$. Since $\bar{f}_{j,C_{i_j}-1}(\nu)$ is also continuous in $\nu \in \mathbb{R}$ and $\bar{f}_{j,C_{i_j}-1}(\nu)$ tends to $\pm\infty$ as $\nu \to \pm\infty$, there exists ν such that $\bar{f}_{j,C_{i_j}-1}(\nu) = 0$. For any $n \in \mathscr{N}_{i_j} \backslash \{C_{i_j}, C_{i_j} - 1\}$, from the definition in (46), for any $\nu \in \mathbb{R}$, $\bar{f}_{j,n}(\nu) \leq \bar{f}_{j,C_{i_j}-1}(\nu)$. Since $\bar{f}_{j,n}(\nu)$ tends to $\pm\infty$ as $\nu \to \pm\infty$ and with the continuity in $\nu \in \mathbb{R}$, there exists a $\nu \in \mathbb{R}$ such that $\bar{f}_{j,n}(\nu) = 0$ for any $n \in \mathscr{N}_{i_j} \backslash \{C_{i_j}, C_{i_j} - 1\}$. Let $\upsilon_{j,n}$ represent the value of such a ν with $\bar{f}_{j,n}(\nu) = 0$ for $n \in \mathscr{N}_{i_j} \backslash \{C_{i_j}\}$.

We now discuss the uniqueness of the zero point $\upsilon_{j,n}$. When job sizes are exponentially distributed and the components are energy-efficiently unimodal, for $n \in \mathscr{N}_{i_j} \backslash \{C_{i_j}\}$ and $\nu \in \mathbb{R}$, if $\bar{f}_{j,n}(\nu) \geq 0$, then, by Proposition 1, there exists H such that, for all $h > H$, an optimal policy ϕ^* exists and satisfies that, for any $n' \geq n+1$, $\alpha_j^{\phi^*}(n') = 0$. For such ν with $\bar{f}_{j,n}(\nu) \geq 0$, $\frac{d^-}{d\nu} \bar{f}_{j,n}(\nu) > 0$, where $\pi_j^{\phi^*}(n')$ is the stationary distribution of sub-process j under ϕ^*, for which the sub-process achieves the maximal average reward $g^*(\nu\boldsymbol{\lambda})$. Accordingly, the zero point $\upsilon_{j,n}$ is unique and (47) is achieved by setting $\upsilon_{j,n}^h = \upsilon_{j,n}$. ∎

Proof of Proposition 2. It is proved invoking Proposition 1 and Lemma 3 by substituting $\upsilon_{\ell,j}^*(n) = \lambda_\ell \upsilon_{j,n}^h$. ∎

APPENDIX III

Proof of Proposition 3. From Proposition 2, for any $j \in [J]$, if $\boldsymbol{\nu} = \nu\boldsymbol{\lambda}$ for some $\nu \in \mathbb{R}$, then there exist a policy $\psi_j(n) \in \tilde{\Phi}_1$ satisfying, for all $\ell \in \{\ell' \in [L] | j \in \mathscr{J}_\ell\}$ and $n' \in \mathscr{N}_{i_j} \backslash \{C_{i_j}\}$, $\alpha_{\ell,j}^{\psi_j(n)}(n') = 1$; and $\alpha_{\ell,j}^{\psi_j(n)}(n') = 0$ otherwise, and $H > 0$ such that, for all $h > H$, $\psi_j(n)$ is optimal for the problem described in (16). In other words, for given $j \in [J]$, $n \in \mathscr{N}_{i_j}$ and multipliers $\boldsymbol{\nu} = \nu\boldsymbol{\lambda}$, there exists $H > 0$ such that, for all $h > H$, $\Gamma_j(\boldsymbol{\nu}) = g_j^*(\boldsymbol{\nu})$. Substituting $\Gamma_j(\boldsymbol{\upsilon}_j^*(n))$ for $g_j^*(\boldsymbol{\upsilon}_j^*(n))$ in (47), together with Proposition 1, we prove the proposition. In particular, the indices $\upsilon_{\ell,j}^*(n) = \lambda_\ell \upsilon_{j,n}^h$ where $\upsilon_{j,n}^h$ is the zero point for $\bar{f}_{j,n}(\upsilon_{j,n}^h) = 0$. ∎

APPENDIX IV

Proof of Proposition 4. From the definition, for given $h \in \mathbb{N}_+$, $i \in [I]$, and $n \in \mathscr{N}_i \backslash \{C_i\}$, $f_{i,n}^h(\eta^0)$ is piece-wise linear and continuous in $\eta^0 \in \mathbb{R}$, and

$$\Big| \frac{d^-}{d\eta^0} f_{i,n}^h(\eta^0) \Big| \leq 1 + \frac{h\hat{\lambda}^0}{\mu_i(n+1)} \sum_{n'=n^+}^{1} \prod_{n''=n^+}^{n'} \frac{\mu_i(n'')}{h\hat{\lambda}^0} \pi_j^{\psi_j(n^+-1)}(n^+), \quad (48)$$

where j is any element in $\mathscr{J}_i$, $n^+ \in \mathscr{N}_i\backslash\{0\}$ is the state such that the state n^+-1 maximizes the right-hand side of (23), $\pi_j^{\psi_j(n^+-1)}(n^+)$ is the steady state distribution of state n^+ under the policy $\psi_j(n^+-1)$. It follows that $|\frac{d^-}{d\eta^0} f_{i,n}^h(\eta^0)|$ is bounded with some finite constant $C \in \mathbb{R}_+$, where the j is any element in $\mathscr{J}_i$. Along similar lines, for the limit case with $h \to +\infty$, $|\frac{d^-}{d\eta^0} f_{i,n}^h(\eta^0)|$ is still bounded with some finite constant. Thus, $f_{i,n}^h(\eta^0)$ is Lipschitz continuous for given $h \in \mathbb{N}_+ \cup \{+\infty\}$, $i \in [I]$, and $n \in \mathscr{N}_i\backslash\{C_i\}$.

Observing that, for given $h \in \mathbb{N}_+$ and any $j \in \mathscr{J}_i$, $\bar{f}_{j,n}(\nu) = \frac{1}{\hat{\lambda}^0} f_{i,n}^h(\hat{\lambda}^0\nu)$. From Lemma 3, there exists $\eta^0 \in \mathbb{R}$ such that $\bar{f}_{i,n}^h(\eta^0) = 0$. Since, for any $h \in \mathbb{N}_+ \cup \{+\infty\}$, $f_{i,n}^h(\eta^0)$ is Lipchitz continuous in $\eta^0 \in \mathbb{R}$, is bounded with bounded $\eta^0 \in \mathbb{R}$ and approaches $\pm\infty$ as $\eta^0 \to \pm\infty$, there exists $\eta^0 \in \mathbb{R}$ such that $f_{i,n}^h(\eta^0) = 0$ when $h \to +\infty$. ∎

APPENDIX V

Define, for $e \in \mathbb{R}$,

$$\boldsymbol{z}^{\psi(e)} \coloneqq \lim_{h\to+\infty} \lim_{t\to+\infty} \mathbb{E}[\boldsymbol{Z}^{\psi^h(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{a}}^h,e),h}(t)],$$

where $\boldsymbol{Z}^{\phi}(t)$ has been defined in (37), and the existence of the limit is ensured by the existence of $\lim_{h\to+\infty} \boldsymbol{\pi}_j^{\psi^h(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{a}},e)}$ ($j \in [J]$).

Proof of Proposition 5. From [24, Propositions 4 and 5], we obtain that, when the job sizes are exponentially distributed, for any $\delta > 0$,

$$\lim_{h\to+\infty} \lim_{T\to+\infty} \frac{1}{T}\mathbb{P}\Big\{\|\boldsymbol{Z}^{\varphi(e),h}(t) - \boldsymbol{z}^{\psi(e)}\| > \delta\Big\} = 0, \quad (49)$$

where $\boldsymbol{Z}^{\varphi,h}(0) = \boldsymbol{z}^0$ and the policy $\varphi(e) \in \Phi$ is the index policy defined in (25) and (26) with substituted $\varphi(e)$ and $\boldsymbol{u}^*(e)$ for φ and $\boldsymbol{\upsilon}^*$, respectively. From (39) and (49), for any $\delta > 0$, there exists $H > 0$ such that, for all $h > H$,

$$\Gamma^{h,\psi^h(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{a}}^h,e)}(e) - \delta \le \Gamma^{h,\varphi(e)}(e) \le \max\nolimits_{\phi\in\Phi} \Gamma^{h,\phi}(e), \quad (50)$$

where the second inequality comes from $\varphi(e) \in \Phi$. Together with (32), for $\delta > 0$, there exists $H > 0$ such that, for $h > H$,

$$\max_{\phi\in\tilde{\Phi}} \Gamma^{h,\phi}(e) - \delta \le \max_{\phi\in\Phi} \Gamma^{h,\phi}(e) \le \max_{\phi\in\tilde{\Phi}} \Gamma^{h,\phi}(e), \quad (51)$$

where the second inequality is based on $\Phi \subset \tilde{\Phi}$.

Recall that, for any $j \in [J]$ and $e \in \mathbb{R}$, the existence of $\lim_{h\to+\infty} \boldsymbol{\pi}_j^{\psi^h(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{a}}^h,e)}$ leading to the existence of $\lim_{h\to+\infty} \Gamma^{h,\psi^h(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{a}},e)}(e)$. When (32) holds, $\lim_{h\to+\infty} \max_{\phi\in\tilde{\Phi}} \Gamma^{h,\phi}(e)$ also exists. Based on (51) and (32), for any $e \in \mathbb{R}$,

$$\lim_{h\to+\infty} \max_{\phi\in\Phi} \Gamma^{h,\phi}(e) = \Gamma(e) = \lim_{h\to+\infty} \max_{\phi\in\tilde{\Phi}} \Gamma^{h,\phi}(e). \quad (52)$$

Let $\Gamma^{h,*}(e) \coloneqq \max_{\phi\in\tilde{\Phi}} \Gamma^{h,\phi}(e)$, which is piece-wise linear, continuous and decreasing in $e \in \mathbb{R}$. Since $\lim_{h\to+\infty} \Gamma^{h,*}(e)$ exists for all given $e \in \mathbb{R}$ and $\lim_{h\to+\infty} \Gamma^{h,*}(e)$ tends to $\pm\infty$ as $e \to \mp\infty$, there exists a solution $e \in \mathbb{R}$ for $\lim_{h\to+\infty} \Gamma^{h,*}(e) = 0$. Together with (32), there exists a zero point $e \in \mathbb{R}$ such that $\lim_{h\to+\infty} \Gamma^{h,\psi^h(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{a}}^h,e)}(e) = 0$ and $\lim_{h\to+\infty} \Gamma^{h,\psi^h(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{a}}^h,e)}(e)$ is continuous in $e \in \mathbb{R}$.

Let e_0 represent a specific real number such that $\Gamma(e_0) = 0$, and, for $h \in \mathbb{N}_+$ and $e \in \mathbb{R}$, let $\phi^h(e)$ represent an optimal solution such that $\Gamma^{h,\phi^h(e)}(e) = \max_{\phi\in\Phi} \Gamma^{h,\phi}(e)$. From (52), for any $\delta > 0$, there exists $H > 0$ such that, for all $h > H$,

$$|\Gamma^{h,\phi^h(e_0)}(e_0)| = |\mathfrak{L}^{\phi^h(e_0)} - e_0\mathfrak{E}^{\phi^h(e_0)}| < \delta. \quad (53)$$

That is,

$$\lim_{h\to+\infty} \mathfrak{L}^{\phi^h(e_0)}/\mathfrak{E}^{\phi^h(e_0)} = e_0. \quad (54)$$

Based on (53), for any $\phi \in \Phi$ and $\delta > 0$, there exists $H > 0$ such that, for all $h > H$, $\mathfrak{L}^{\phi} - e_0\mathfrak{E}^{\phi} \le \mathfrak{L}^{\phi^h(e_0)} - e_0\mathfrak{E}^{\phi^h(e_0)} < \delta$. It follows that, for any $\phi \in \Phi$,

$$\lim_{h\to+\infty} \mathfrak{L}^{\phi}/\mathfrak{E}^{\phi} \le e_0. \quad (55)$$

From (54), (55) and (8), we obtain that $\lim_{h\to+\infty} e^* = \lim_{h\to+\infty} \max_{\phi\in\Phi} \frac{\mathfrak{L}^{\phi}}{\mathfrak{E}^{\phi}} = e_0$. This proves (33).

We then discuss the uniqueness of the zero point e_0 for $\Gamma(e_0) = 0$. If there exists $e_1 \neq e_0$ satisfying $\Gamma(e_1) = 0$, then, from the above discussion, $e_1 = \lim_{h\to+\infty} \max_{\phi\in\Phi} \frac{\mathfrak{L}^{\phi}}{\mathfrak{E}^{\phi}} = e_0$. Hence, e_0 is the unique solution for $\Gamma(e_0) = 0$.

Recall that, for given $h \in \mathbb{N}_+$, $\Gamma^{h,*}(e) \coloneqq \max_{\phi\in\tilde{\Phi}} \Gamma^{h,\phi}(e)$, which is piece-wise linear in $e \in \mathbb{R}$ with the left derivative $\left|\frac{d^-}{de}\Gamma^{h,*}(e)\right| < C$, where $C \in \mathbb{R}_+$ is a constant independent from h, ϕ^* is an optimal policy satisfying $\Gamma^{h,\phi^*}(e) = \Gamma^{h,*}(e)$, and $\boldsymbol{\varepsilon}_i \coloneqq (\varepsilon_i(n) : n \in \mathscr{N}_i)$ for $i \in [I]$. That is, for given $h \in \mathbb{N}_+$, $\Gamma^{h,*}(e)$ is Lipschitz continuous in $e \in \mathbb{R}$ with a bounded Lipschitz constant independent from h. For any given $e \in \mathbb{R}$, from (32), the function $\Gamma(e) = \lim_{h\to+\infty} \max_{\phi\in\tilde{\Phi}} \Gamma^{h,\phi}(e) = \lim_{h\to+\infty} \Gamma^{h,*}(e)$. It follows that, for any $e \in \mathbb{R}$, $\lim_{\Delta\to 0}|\Gamma(e+\Delta) - \Gamma(e)| \le 0$. That is, $\Gamma(e)$ is continuous in $e \in \mathbb{R}$. Similarly, $\lim_{\Delta\uparrow 0} \frac{1}{\Delta}|\Gamma(e+\Delta) - \Gamma(e)| = \lim_{\Delta\uparrow 0} \frac{1}{\Delta}|\lim_{h\to+\infty}\big(\Gamma^{h,*}(e+\Delta) - \Gamma^{h,*}(e)\big)| \le C$, leading to the Lipschitz continuity of $\Gamma(e)$. ∎

APPENDIX VI

Proof of Lemma 1. The lemma can be proved by showing that constructing $\boldsymbol{\nu}$, $\boldsymbol{\gamma}$, and ϕ^*, where $\boldsymbol{\alpha}_j^{\phi^*}$ and $\bar{\alpha}_\ell^{\phi^*}$ maximize the the objectives defined in (16) and (17); and showing that such $\boldsymbol{\nu}$, $\boldsymbol{\gamma}$, and ϕ^* achieve the complementary slackness for the relaxed problem defined by (9), (10), (12) and (13).

We start with the case with Condition 1. Let $\nu_\ell/\lambda_\ell = \nu$ where

$$\nu = \min\left\{0, \min\nolimits_{j\in\mathscr{J}_\ell} \upsilon_{\ell,j}^*(C_{i_j} - 1)/\lambda_\ell\right\} \quad (56)$$

with $\upsilon_{\ell,j}^*(n)$ ($n \in \mathscr{N}_{i_j}$, $j \in \mathscr{J}_\ell$, $\ell \in [L]$) given by (22), and let $\gamma_\ell = -\lambda_\ell\nu$, $\ell \in [L]$. In this context, from Proposition 2, for sufficiently large h, there is an optimal policy ϕ^* for the problem defined in (16), satisfying $\alpha_{\ell,j}^{\phi^*}(n) = 1$ for all $n \in \mathscr{N}_{i_j}\backslash\{C_{i_j}\}$, $j \in \mathscr{J}_\ell$. Let $\bar{\alpha}_\ell^{\phi^*} = 1 - A_\ell$, where $0 \le 1 - A_j \le 1$ under Condition 1. In other words, constraints (10) and (13) are satisfied with equality. Recall that Constraint (12) has been guaranteed by setting $\boldsymbol{\eta}$ to be sufficiently large: the complementary slackness of the relaxed problem is achieved. Hence, ϕ^* also maximizes the primal problem defined by (9), (10), (12) and (13).

We now consider the case with $L = 1$. If $A_\ell \le 1$, then this is a special case for $L \ge 1$ with Condition 1. It remains to discuss the case with $L = 1$ and $A_\ell > 1$. From Proposition 2, for sufficiently large h, there is an optimal policy ϕ^* that maximizes the problem defined in (16), for which $\boldsymbol{\alpha}_j^{\phi^*}$ is determined by (18) and (22). Since $A_\ell > 1$, there must exist

a ν_ℓ such that $\nu_\ell \geq \min_{j\in\mathscr{J}_\ell, n\in\mathscr{N}_{i_j}} v^*_{\ell,j}(n)$, where the ℓ is the only element in $[L]$, and $\sum_{j\in\mathscr{J}_\ell} \boldsymbol{\pi}_j^{\phi^*}\cdot\boldsymbol{\alpha}_j^{\phi^*}=1$.

If $A_\ell > 1$, then let $\gamma_\ell > \max\{0, -\nu_\ell\}$, which guarantees that policy ϕ^* with $\bar{\alpha}_\ell^{\phi^*} < 0$ maximizes the objective defined in (17). Then, $I(\bar{\alpha}_\ell^{\phi^*}) = 0 = I(1-A_\ell)$. Constraints (10) and (13) achieve equality. The complementary slackness conditions are also satisfied under this setting. The lemma is proved. ∎

Appendix VII

Proof of Proposition 6. Along similar lines as the proof of Proposition 5 in Appendix V, based on (32), we obtain (50) and (52). In other words, for any $e \in \mathbb{R}$, $\lim_{h\to+\infty}\Gamma^{h,\varphi(e)}(e) = \lim_{h\to+\infty}\Gamma^{h,\psi^h(\bar{\boldsymbol{\nu}},\bar{\boldsymbol{a}}^h,e)}(e) = \Gamma(e)$. From Proposition 5, if (32) holds, $\Gamma(e)$ is Lipschitz continuous in $e\in\mathbb{R}$. It follows with (35) and proves the proposition. ∎

Appendix VIII

Proof of Proposition 7. The proposition is proved by invoking [24, Propositions 4 and 5], which ensures the existence of a global attractor for the process $\{\boldsymbol{Z}^{\varphi(e),h}(t), t\geq 0\}$, where $\boldsymbol{Z}^{\varphi(e),h}(t)$ has been defined in (37) and the index policy $\varphi(e)$ has been defined in (25) and (26) with substituted $\varphi(e)$ and $u^*_{\ell,i_j}(n,e)$ for φ and $v^*_{\ell,j}(n)$, respectively. More precisely, from [24, Propositions 4 and 5], for any $\delta>0$, we obtain (49) with $\boldsymbol{Z}^{\varphi(e),h}(t) = \boldsymbol{z}^0$. The $\boldsymbol{z}^{\psi(e)}$ is the global attractor for the process $\{\boldsymbol{Z}^{\varphi(e),h}(t), t\geq 0\}$. That is, together with Lemma 2, we obtain $\lim_{h\to+\infty}\Gamma^{h,\varphi(e)}(e) = \lim_{h\to+\infty}\max_{\phi\in\tilde{\Phi}}\Gamma^{h,\phi}(e)$, which, since $\varphi(e)\in\Phi$ and $\Phi\subset\tilde{\Phi}$, indicates (36) with substituted $\varphi(e)$ for ϕ. ∎

Appendix IX

Proof of Proposition 8. Along similar lines as the proof of [23, Proposition 2], by invoking [48, Theorem 4.1 in Chapter 7 & Theorem 3.3 in Chapter 3], there exists a deterministic process $\boldsymbol{z}(t)$ taking values in $\mathscr{Z}$ such that, for any $\delta>0$, there exist $s>0$ and $H>0$ satisfying, for all $h>H$,

$$\lim_{T\to+\infty}\frac{1}{T}\int_0^T \mathbb{P}\Big\{\Big\|\boldsymbol{Z}^{h,\varphi(e)}(t)-\boldsymbol{z}(t)\Big\|>\delta\Big\}dt \leq e^{-sh}. \quad (57)$$

Together with (49), we obtain that, for $\delta>0$, there exist $s>0$ and $H>0$ such that, for all $h>H$, (38) is achieved. ∎

Appendix X

Consider a system with $I=10$ clusters, where the service and energy consumption rates of components are instances of pseudo-random variables:

- $\mu_1(C_1)=2.425$, $\varepsilon_1(0)=0.0655$, $\frac{\mu_1(C_1)}{\varepsilon_1(C_1)-\varepsilon_1(0)}=13.699$
- $\mu_2(C_2)=1.620$, $\varepsilon_2(0)=0.0333$, $\frac{\mu_2(C_2)}{\varepsilon_2(C_2)-\varepsilon_2(0)}=15.338$
- $\mu_3(C_3)=1.758$, $\varepsilon_3(0)=0.0315$, $\frac{\mu_3(C_3)}{\varepsilon_3(C_3)-\varepsilon_3(0)}=14.845$
- $\mu_4(C_4)=1.600$, $\varepsilon_4(0)=0.0189$, $\frac{\mu_4(C_4)}{\varepsilon_4(C_4)-\varepsilon_4(0)}=18.562$
- $\mu_5(C_5)=1.728$, $\varepsilon_5(0)=0.0116$, $\frac{\mu_5(C_5)}{\varepsilon_5(C_5)-\varepsilon_5(0)}=26.225$
- $\mu_6(C_6)=1.668$, $\varepsilon_6(0)=0.0069$, $\frac{\mu_6(C_6)}{\varepsilon_6(C_6)-\varepsilon_6(0)}=33.166$
- $\mu_7(C_7)=2.390$, $\varepsilon_7(0)=0.0055$, $\frac{\mu_7(C_7)}{\varepsilon_7(C_7)-\varepsilon_7(0)}=43.127$
- $\mu_8(C_8)=2.116$, $\varepsilon_8(0)=0.0026$, $\frac{\mu_8(C_8)}{\varepsilon_8(C_8)-\varepsilon_8(0)}=51.306$
- $\mu_9(C_9)=2.416$, $\varepsilon_9(0)=0.0011$, $\frac{\mu_9(C_9)}{\varepsilon_9(C_9)-\varepsilon_9(0)}=70.356$
- $\mu_{10}(C_{10})=2.224$, $\varepsilon_{10}(0)=0$, $\frac{\mu_{10}(C_{10})}{\varepsilon_{10}(C_{10})-\varepsilon_{10}(0)}=97.625$

and, for all clusters $i\in[I]$, the service and energy consumption rates for states $n\in\mathscr{N}_i\backslash\{0,C_i\}$ are given by $\mu_i(n) = \frac{n}{n+1}\mu_i(n+1)$ and $\varepsilon_i(n)=\frac{n}{n+1}(\varepsilon_i(n+1)-\varepsilon_i(0))+\varepsilon_i(0)$, respectively. In particular, the unit of service rates of all the clusters is 10^{-10} number of jobs per second: they are normalized to be sufficiently small that we can observe a positive number of blocked jobs and the heavy traffic condition can be achieved during peak hours. There are $L=4$ classes of jobs and, for all the classes $\ell\in[L]$, the sets of clusters able to serve an ℓ-job are $\mathscr{J}_1=\{1,5,6,10\}$, $\mathscr{J}_2=\{1,2,3,4,5,7,8,9\}$, $\mathscr{J}_3=\{1,6,7,10\}$, and $\mathscr{J}_4=\{2\}$.

References

[1] Cisco, "Cisco global cloud index: Forecast and methodology, 2016-2021," 2018, accessed: Jan. 10, 2024. [Online]. Available: https://virtualization.network/Resources/Whitepapers/0b75cf2e-0c53-4891-918e-b542a5d364c5_white-paper-c11-738085.pdf

[2] A. Shehabi, S. Smith, D. Sartor, R. Brown, M. Herrlin, J. Koomey, E. Masanet, N. Horner, I. Azevedo, and W. Lintner, "United states data center energy usage report," Jun. 2016.

[3] D. Kliazovich, P. Bouvry, F. Granelli, and N. L. S. da Fonseca, "Energy consumption optimization in cloud data centers," in *Cloud Services, Networking, and Management*, N. L. S. da Fonseca and R. Boutaba, Eds. John Wiley & Sons, Inc, Apr. 2015, pp. 191–215, accessed: Jan. 10, 2024. [Online]. Available: http://dx.doi.org/10.1002/9781119042655.ch8

[4] M. Dayarathna, Y. Wen, and R. Fan, "Data center energy consumption modeling: A survey," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 4, pp. 2180 – 2194, Aug. 2017.

[5] T. Lu, M. Chen, and L. L. H. Andrew, "Simple and effective dynamic provisioning for power-proportional data centers," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 6, pp. 1161–1171, Apr. 2013.

[6] M. E. Gebrehiwot, S. Aalto, and P. Lassila, "Near-optimal policies for energy-aware task assignment in server farms," in *Proc. CCGrid 2017*. Madrid, Spain: IEEE Press, May 2017, pp. 1017–1026.

[7] F. Esposito, D. Di Paola, and I. Matta, "On distributed virtual network embedding with guarantees," *IEEE/ACM Trans. Netw.*, vol. 24, no. 1, pp. 569–582, Feb. 2016.

[8] H. Feng, J. Llorca, A. M. Tulino, D. Raz, and A. F. Molisch, "Approximation algorithms for the NFV service distribution problem," in *Proc. IEEE INFOCOM 2017*, Atlanta, GA, USA, May 2017, pp. 1–9.

[9] W. Q. M. Guo, A. Wadhawan, L. Huang, and J. T. Dudziak, "Server farm management," Jan. 2014, US Patent 8,626,897. Accessed: Jan. 10, 2024. [Online]. Available: http://www.google.com/patents/US8626897

[10] A. Hameed, A. Khoshkbarforoushha, R. Ranjan, P. P. Jayaraman, J. Kolodziej, P. Balaji, S. Zeadally, Q. M. Malluhi, N. Tziritas, A. Vishnu, S. U. Khan, and A. Zomaya, "A survey and taxonomy on energy efficient resource allocation techniques for cloud computing systems," *Computing*, vol. 98, no. 7, pp. 751–774, Jun. 2016.

[11] Z. Rosberg, Y. Peng, J. Fu, J. Guo, E. W. M. Wong, and M. Zukerman, "Insensitive job assignment with throughput and energy criteria for processor-sharing server farms," *IEEE/ACM Trans. Netw.*, vol. 22, no. 4, pp. 1257–1270, Aug. 2014.

[12] E. Hyytiä, R. Righter, and S. Aalto, "Task assignment in a heterogeneous server farm with switching delays and general energy-aware cost structure," *Performce Evaluation*, vol. 75-76, pp. 17–35, 2014.

[13] T. Lin, T. Alpcan, and K. Hinton, "A game-theoretic analysis of energy efficiency and performance for cloud computing in communication networks," *IEEE Syst. J.*, vol. 11, no. 2, pp. 649–660, Jun. 2017.

[14] X. Wei and M. J. Neely, "Data center server provision: Distributed asynchronous control for coupled renewal systems," *IEEE/ACM Trans. Netw.*, vol. 18, no. 1, First Quarter 2016.

[15] S. K. Mishra, D. Puthal, J. J. Rodrigues, B. Sahoo, and E. Dutkiewicz, "Sustainable service allocation using a metaheuristic technique in a fog server for industrial applications," *IEEE Trans. Ind. Informat.*, vol. 14, no. 10, pp. 4497–4506, Oct. 2018.

[16] O. T. Akgun, D. G. Down, and R. Righter, "Energy-aware scheduling on heterogeneous processors," *IEEE Trans. Automat. Contr.*, vol. 59, no. 3, pp. 599–613, 2013.

[17] J. Li, Y. Zhu, J. Yu, C. Long, G. Xue, and S. Qian, "Online auction for IaaS clouds: Towards elastic user demands and weighted heterogeneous VMs," *IEEE Trans. Parallel Distrib. Syst.*, vol. 29, no. 9, pp. 2075–2089, Sep. 2018.

[18] N. Bansal, H.-L. Chan, and K. Pruhs, "Speed scaling with an arbitrary power function," *ACM Trans. Algorithms.*, vol. 9, no. 2, p. 18, Mar. 2013.

[19] X. Mei, Q. Wang, and X. Chu, "A survey and measurement study of GPU DVFS on energy conservation," *Digit. Commun. Netw.*, vol. 3, no. 2, pp. 89–100, May 2017.
[20] P. Whittle, "Restless bandits: Activity allocation in a changing world," *J. Appl. Probab.*, vol. 25, pp. 287–298, 1988.
[21] C. H. Papadimitriou and J. N. Tsitsiklis, "The complexity of optimal queuing network control," *Math. Oper. Res.*, vol. 24, no. 2, pp. 293–305, May 1999.
[22] R. R. Weber and G. Weiss, "On an index policy for restless bandits," *J. Appl. Probab.*, no. 3, pp. 637–648, Sep. 1990.
[23] J. Fu and B. Moran, "Energy-efficient job-assignment policy with asymptotically guaranteed performance deviation," *IEEE/ACM Trans. Netw.*, vol. 28, no. 3, pp. 1325–1338, 2020.
[24] J. Fu, B. Moran, and P. G. Taylor, "A restless bandit model for resource allocation, competition, and reservation," *Oper. Res.*, vol. 70, no. 1, pp. 416–431, 2022.
[25] J. Fu, B. Moran, J. Guo, E. W. M. Wong, and M. Zukerman, "Asymptotically optimal job assignment for energy-efficient processor-sharing server farms," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, Dec. 2016.
[26] Q. Wang, J. Fu, J. Wu, B. Moran, and M. Zukerman, "Energy-efficient priority-based scheduling for wireless network slicing," in *Proc. IEEE GLOBECOM 2018*, Abu Dhabi, UAE, Dec. 2018.
[27] Z. Yu, Y. Xu, and L. Tong, "Deadline scheduling as restless bandits," *IEEE Trans. Automat. Contr.*, vol. 63, no. 8, pp. 2343–2358, 2018.
[28] V. S. Borkar, "Whittle index for partially observed binary Markov decision processes," *IEEE Trans. Automat. Contr.*, vol. 62, no. 12, pp. 6614–6618, 2017.
[29] J. Wang, X. Ren, Y. Mo, and L. Shi, "Whittle index policy for dynamic multichannel allocation in remote state estimation," *IEEE Trans. Automat. Contr.*, vol. 65, no. 2, pp. 591–603, 2019.
[30] A. Abbou and V. Makis, "Group maintenance: A restless bandits approach," *INFORMS J. Comput.*, vol. 31, no. 4, pp. 719–731, 2019.
[31] J. Niño-Mora, "Restless bandits, partial conservation laws and indexability," *Adv. Appl. Probab.*, pp. 76–98, 2001.
[32] J. Niño-Mora, "Dynamic allocation indices for restless projects and queueing admission control: a polyhedral approach," *Math. Program.*, vol. 93, no. 3, pp. 361–413, 2002.
[33] J. Niño-Mora, "Dynamic priority allocation via restless bandit marginal productivity indices," *Trans. Oper. Res.*, vol. 15, no. 2, pp. 161–198, 2007.
[34] J. Niño-Mora, "A verification theorem for threshold-indexability of real-state discounted restless bandits," *Math. Oper. Res.*, vol. 45, no. 2, pp. 465–496, 2020.
[35] S. M. Ross, *Applied probability models with optimization applications*. Dover Publications (New York), 1992.
[36] W. Ouyang, A. Eryilmaz, and N. B. Shroff, "Downlink scheduling over Markovian fading channels," *IEEE/ACM Trans. Netw.*, vol. 24, no. 3, pp. 1801–1812, 2016.
[37] A. Maatouk, S. Kriouile, M. Assaad, and A. Ephremides, "Asymptotically optimal scheduling policy for minimizing the age of information," in *2020 IEEE International Symposium on Information Theory (ISIT)*. Los Angeles, CA, USA, USA: IEEE, Jun. 2020, pp. 1747–1752.
[38] I. Takouna, W. Dawoud, and C. Meinel, "Accurate mutlicore processor power models for power-aware resource management," in *Proc. 2011 IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing*. Sydney, NSW, Australia: IEEE, Dec. 2011, pp. 419–426.
[39] J. Fu, X. Wang, Z. Wang, and M. Zukerman, "A restless bandit model for energy-efficient job assignments in server farms," Oct. 2023, extended Version. [Online]. Available: https://arxiv.org/abs/2112.06275
[40] M. Pedram, "Energy-efficient datacenters," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 31, no. 10, pp. 1465–1484, 2012.
[41] H. Wu and K. Wolter, "Stochastic analysis of delayed mobile offloading in heterogeneous networks," *IEEE Trans. Mobile Comput.*, vol. 17, no. 2, pp. 461–474, 2017.
[42] J. Wilkes, "More Google cluster data," Google research blog, Nov. 2011, accessed: Jan. 10, 2024. [Online]. Available: http://googleresearch.blogspot.com/2011/11/more-google-cluster-data.html
[43] C. Reiss, J. Wilkes, and J. L. Hellerstein, "Google cluster-usage traces: format + schema," Google Inc., Mountain View, CA, USA, Technical Report, Nov. 2011, revised 2014-11-17 for version 2.1. Accessed: Jan. 10, 2024. [Online]. Available: https://github.com/google/cluster-data
[44] M. E. Crovella and A. Bestavros, "Self-similarity in World Wide Web traffic: evidence and possible causes," *IEEE/ACM Trans. Netw.*, vol. 5, no. 6, pp. 835–846, Dec. 1997.
[45] M. Harchol-Balter, *Performance Modeling and Design of Computer Systems: Queueing Theory in Action*. Cambridge University Press, 2013.
[46] R. A. Giri and A. Vanchi, "Increasing data center efficiency with server power measurements," *Document. Intel Information Technology. IT@ Intel White Paper*, 2010, accessed: Jan. 10, 2024. [Online]. Available: https://www.intel.co.za/content/dam/doc/white-paper/intel-it-data-center-efficiency-server-power-paper.pdf
[47] T. Kaur and I. Chana, "Energy efficiency techniques in cloud computing: A survey and taxonomy," *ACM Comput. Surv.*, vol. 48, no. 2, pp. 1–46, 2015.
[48] M. I. Freidlin and A. D. Wentzell, *Random perturbations of dynamical systems*. Springer Science & Business Media, 2012, translated by J. Szücs.

Jing Fu (S'15–M'16) received the B.Eng. degree in computer science from Shanghai Jiao Tong University, Shanghai, China, in 2011, and the Ph.D. degree in electronic engineering from the City University of Hong Kong in 2016. She has been with the School of Mathematics and Statistics, the University of Melbourne as a Post-Doctoral Research Associate from 2016 to 2019. She is now with RMIT University, Australia, as a lecturer. Her research interests now include energy-efficient networking/scheduling, resource allocation in large-scale networks, restless multiarmed bandit problems, stochastic optimization.

Xinyu Wang received the B.Eng. degree in Electronic Science and Technology from Zhejiang University, Hangzhou, and the M.Sc. degree in Electronic Information Engineering and Ph.D. degree in Electrical Engineering from City University of Hong Kong in 2016, 2018, and 2023, respectively. He is currently a research assistant with the Department of Electrical Engineering, City University of Hong Kong, China. His research interests include path planning and resource allocation in telecommunication networks.

Zengfu Wang received the B.Sc. degree in applied mathematics, the M.Sc. degree in control theory and control engineering, the Ph.D. degree in control science and engineering all from Northwestern Polytechnical University, China, in 2005, 2008, and 2013, respectively. Since 2014, he has been with the same university, where he is now an Associate Professor. During 2014-2015, he was a Postdoctoral Research Fellow with the Department of Electronic Engineering, City University of Hong Kong. Between 2019-2020, he visited Delft University of Technology, The Netherlands. His research interests include path planning, discrete optimization, and information fusion.

Moshe Zukerman (M'87–SM'91–F'07–LF'20) received a B.Sc. degree in industrial engineering and management, an M.Sc. degree in operations research from the Technion – Israel Institute of Technology, Haifa, Israel, and a Ph.D. degree in engineering from the University of California, Los Angeles, in 1985. He was an independent consultant with the IRI Corporation and a Postdoctoral Fellow with the University of California, Los Angeles, in 1985–1986. In 1986–1997, he was with Telstra Research Laboratories (TRL). During 1997-2008, he was with The University of Melbourne, Victoria, Australia. In 2008 he joined the City University of Hong Kong (CityU) as a Chair Professor of Information Engineering.