Optimizing Cloud Utilization via Switching Decisions [*]

Eugene Feinberg
Applied Mathematics and Statistics
Stony Brook University
Stony Brook, NY 11794
efeinberg@notes.cc.sunysb.edu

Xiaoxuan Zhang
IBM T.J. Watson Research Center
Yorktown Heights, NY 10598
zhangxiaoxuan@live.com

ABSTRACT

This paper studies a control problem for optimal switching on and off a cloud computing services modeled by an $M/M/\infty$ queue with holding, running and switching costs. The main result is that an average-optimal policy either always runs the system or is an (M, N)-policy defined by two thresholds M and N, such that the system is switched on upon an arrival epoch when the system size accumulates to N and it is switched off upon a departure epoch when the system size decreases to M. We compare the optimal (M, N)-policy with the classical $(0, N)$-policy and show the non-optimality of it.

Keywords

Cloud Computing, M/M/∞, Queueing Control, Markov Decision Process

1. INTRODUCTION

As cloud computing starts to play an important role in the areas related to big data analytics, the performance evaluation and stochastic control of cloud computing centers are paid more and more attentions in recent years. Cloud providers like Amazon Elastic Compute Cloud [1], Google Compute Engine [5], HP Cloud Compute [**?**], IBM Smart Cloud [9] and Microsoft [15] offer computing and storage sharing facilities for lease at various prices depending on the type of services. A major business requirement in cloud computing is that it must be able to support extensive scalability at an acceptable and competitive cost. Most cloud computing is expected to be able to support thousands of servers with reasonable performances [10], which is realized by using peer-to-peer parallel and distributed techniques. Such facilities are highly scalable and capable of handling huge number of simultaneous user requests using queueing based models [15]. Large financial and environmental costs associated with high electricity consumption of running a data center is one of the major considerations for these cloud providers [6]. How to operate a *cloud* efficiently and environmental friendly is becoming more challenging nowadays. This paper studies the optimal decisions for a server provider of cloud computing services, who needs to make decisions on whether to pay for the cloud facility (switch on the service) or not.

Many research works on resource allocation and stochastic control of a cloud center consider modeling the cloud computing as a multi-server queueing system [11, 13]. However, a cloud center can have a huge number of server nodes, typically of the order

of hundreds or thousands [6, 1]. Given that n is large, traditional $M/M/n$ queueing system is not suitable in this case. Thus we look into $M/M/\infty$ queue in this paper.

In addition to cloud computing, another motivation comes from the application in IT software maintenance. [12] studied the software maintenance problem as a queue formed by software maintenance requests relevant to software bugs experienced by customers. Once a customer is served and the appropriate bug is fixed in the new software release or patch, it also provides service to some other customers in the queue. In [12] it was assumed that the number of customers leaving the queue at a service completion time has a binomial distribution, this problem was modeled in [12] as an optimal switching problem for an $M/G/1$ queue in which a binomially distributed number of customers is served each time, and the best policy was found among the policies that turn the system off when it is empty and turn it on when there are N or more customers in the system. Here we observe that after an appropriate scaling, this problem and the optimal switching problem for an $M/M/\infty$ queue have the same fluid approximations. So, the result on optimality of $(M.N)$-policies described below provide certain insights to the software maintenance problem in [12].

In this paper, we consider an $M/M/\infty$ queue with Poisson arrivals and independent exponential service times. The number of servers is unlimited. The system can be switched on and off any time. All occupied servers operate when the system is on, and all the servers are off when the system is off. The costs include the linear holding cost h for a unit time that a customer spends in the system, the start-up cost s_1, the shut-down cost s_0, and the running costs c per unit time. We assume that $h > 0$, and $s_0 + s_1 > 0$.

To simplify the initial analysis, we assume that the server can be turned on and off only at time 0, customer arrival times, and customer departure times. These times are jump epochs for the process $X(t)$ of the number of customers in the system at time t. Let $t_0, t_1, \ldots$ be the sequence of jump epochs. We initially consider the servers can be switched on and off only at jump epochs. Switching takes place only at these times is not restrictive, and the optimal policies described in the paper are also optimal when the system can be turned on and off any time.

The main result of this paper is that either the policy that always keeps the server on is average-optimal or for some integers M and N, where $N > M \geq 0$, the so-called (M, N)-policy, is average-optimal. The (M, N)-policy switches the running system off when the number of customers in the system is not greater than M and it turns the idling system on when the number of customers in the queue reaches or exceeds N. In particular, $(0, N)$-policies are known in the literature under the name of N-policies. It is well-known [8] that for an $M/G/1$ queue either the policy that always runs the sever is average-optimal or for some natural number N the N-policy is

average-optimal. We will show that N-policies may not be average-optimal for $M/M/\infty$ queues.

We model the problem as a Continuous-Time Markov Decision Process in Section 2. The transition rates in this CTMDP are not bounded. This is an additional complication. CTMDPs with unbounded transition rates were recently studied by [7]. However, it was assumed there that any stationary policy defines an ergodic continuous-time Markov chain. This condition does not hold for the problem we consider because the policy that always keeps the system off defines a transient Markov chain. First we study the expected total rewards in Section 3.1. Then, in Section 3.2 we investigate the discrete-time total-reward problem limited to the policies that never turn the running system off. Section 3.3 investigates the reduction of the original model to the finite state space model and proves the existence of average-optimal policies. In Section 4 we prove the existence of stationary average-optimal policies and describe their structure. Section 5 provides the conditions for the full-service policies and (M, N)-policies. We provide the linear programming method to compute the average-optimal polices in Sectin 6.

2. PROBLEM FORMULATION

We model the above described problem as a continuous time Markov decision process (CTMDP). The state space is $Z = \mathbb{N} \times \{0, 1\}$, where $\mathbb{N} = \{0, 1, \ldots\}$. If the state of the system at the decision epoch n is $z_n = (X_n, \delta_n) \in Z$, this means that the number of customers in the system is X_n and the state of the system is δ_n, with $\delta_n = 1$ if the system is on and $\delta_n = 0$ if the system is off. The action set is $A = \{0, 1\}$, where $a = 0$ indicating turning or keeping the system off and $a = 1$ meaning turning or keeping the system on.

The transition rate from state $z = (i, \delta)$ with action a to state $z' = (j, a)$ is $q(z'|z, a) = q(j|i, a)$ where

$$q(j|i, a) = \begin{cases} \lambda, & \text{if } j = i + 1, \\ i\mu & \text{if } i > 0, \ a = 1, \ j = i - 1, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

At state $z = (i, \delta)$, define $q(z, a) = q(i, a) = \sum_{z' \in Z} q(z'|z, a) = \sum_{j \in \mathbb{N}} q(j|i, a)$, for $z' = (j, a)$. Now we define the cost function. The cumulative cost $C(t)$ over the interval t is $C(t) = \int_0^t (hX(t) + c\delta(t))dt + \sum_{n=0}^{N(t)} s_{\delta(t_n)}|\delta(t_n) - \delta(t_n-)|$, where $X(t)$ is the system size, $\delta(t)$ is the system state at t, and $N(t)$ is the number of jump epochs up to t.

A *stationary* policy is defined by a mapping $\pi : Z \to A$ such that $\pi(z) \in A, z \in Z$. For each initial state of the system $x_0 = (i, \delta)$, and for any policy π, the policy π defines a stochastic sequence $\{x_n, a_n, t_n, n = 0, 1, \ldots\}$, where $t_0 = 0$ and $t_{n+1} \geq t_n$. We denote by $E_{x_0}^\pi$ the expectation of this process. For any initial state of the system $z_0 = (i, \delta)$, and for any policy π, the expected total discounted cost over the infinite horizon is

$$V_\alpha^\pi(i, \delta) = E_{(i,\delta)}^\pi \int_0^\infty e^{-\alpha t} dC(t)$$

$$= E_{(i,\delta)}^\pi \left[\int_0^\infty e^{-\alpha t}(hX(t) + \delta(t)c)dt + \sum_{n=0}^\infty e^{-\alpha t_n n}|a_n - \delta_n|s_{a_n} \right]. \quad (2)$$

Define the average cost per unit time as $v^\pi(i, \delta) = \limsup_{t \to \infty} t^{-1} E_{z_0}^\pi C(t)$.
Let $V_\alpha(i, \delta) = \inf_\pi V_\alpha^\pi(i, \delta)$, and $v(i, \delta) = \inf_\pi v^\pi(i, \delta)$. A policy φ is called discount-optimal if $V_\alpha^\varphi(i, \delta) = V_\alpha(i, \delta)$ for any policy π and for any initial state (i, δ). A policy φ is called average-optimal if $v^\varphi(i, \delta) = v(i, \delta)$ for any policy π and for any initial state (i, δ).

3. DISCOUNTED COST CRITERION

3.1 Reduction to Discrete Time and Existence of Stationary Discount-Optimal Policies

At state $(i, 1)$, the time until the next jump has an exponential distributions with intensity $\lambda + i\mu \to \infty$ as $i \to \infty$. Since the jump rates are unlimited, it is impossible to present the problem as a discounted MDP with the discount rate smaller than 1. Thus, we shall present our problem as a negative MDP. By Theorem 5.6 in [4], the expected transition time until the next jump epoch is $\tau_\alpha(z, a) = \tau_\alpha((i, \delta), a) = \dfrac{1}{\alpha + q(z, a)}$, and the one-step cost as $C_\alpha((i, \delta), a) = |a - \delta|s_a + (hi + ac)\tau_\alpha((i, \delta), a)$. We use $C(z, a)$ and $\tau(z, a)$ to denote $C_\alpha(z, a)$ and $\tau_\alpha(z, a)$ respectively if $\alpha = 0$. The value function $V(z)$ satisfies the optimality equation

$$V(z) = \min_{a \in A(z)} \{C_\alpha(z, a) + \sum_{z' \in Z} \frac{q(z'|z, a)}{\alpha + q(z, a)}(z'|z, a)V(z')\}, \ z \in Z. \quad (3)$$

By [4], for an MDP with a countable state set Z, action sets $A(Z)$, transition rate $q(z'|z, a)$, and nonnegative one-step rewards $C(z, a)$, a stationary a stationary policy ϕ is optimal if and only if for all $z \in Z$ it satisfies $V(z) = C_\alpha(z, \phi(z)) + \sum_{z' \in Z} \dfrac{q(z'|z, a)}{\alpha + q(z, \phi(z))} V(z')$, where $V(z)$ is the infimum of the expected total costs starting from state z. We follow the conventions that $q(-1|i, a) = 0$, $V_\alpha(-1, \delta) = 0$, $\sum_\emptyset = 0$, and $\prod_\emptyset = 1$.

THEOREM 1. *For any $\alpha > 0$ the following statements hold:*

(i) For $i = 0, 1, \ldots, V_\alpha(i, \delta) \leq (1 - \delta)s_1 + \dfrac{hi}{\mu + \alpha} + \dfrac{h\lambda}{\alpha(\mu + \alpha)} + \dfrac{c}{\alpha}.$

(ii) For all $i = 0, 1, \ldots$ and all $\delta = 0, 1$ the value function $V_\alpha(i, \delta)$ satisfies the optimality equation

$$V_\alpha(i, \delta) = \min_{a \in \{0, 1\}} \left\{ C((i, \delta), a) + \frac{q(i - 1|i, a)}{\alpha + q(i, a)} V_\alpha(i - 1, a) \right.$$
$$\left. + \frac{q(i + 1|i, a)}{\alpha + q(i, a)} V_\alpha(i + 1, a) \right\} \quad (4)$$

(iii) There exists a stationary discount-optimal policy, and a stationary policy ϕ is discount-optimal if and only if for all $i = 0, 1, \ldots$ and for all $\delta = 0, 1$

$$V_\alpha(i, \delta) = \min_{\phi(i, \delta) \in [0, 1]} \left\{ C_\alpha((i, \delta), \phi(i, \delta)) + \frac{q(i - 1|i, a)}{\alpha + q(i, \phi(i, \delta))} \times \right.$$
$$\left. V_\alpha(i - 1, \phi(i, \delta)) + \frac{q(i + 1|i, a)}{\alpha + q(i, \phi(i, \delta))} V_\alpha(i + 1, \phi(i, \delta)) \right\}.$$

3.2 Full Service Policies

The class of the policies that never turn the running server off is the class of all policies in the MDP with Z, A, $A(i, 0) = A$ and $A(i, 1) = \{1\}$. This is a sub-model of our original model. Defined by (2), let $U_\alpha(i, \delta)$ be the optimal total discount cost under the policy that always runs the system.

THEOREM 2. *For any $\alpha > 0$ the following statements hold:*

(i) For all $i = 0, 1, \ldots U_\alpha(i, 1) = \dfrac{hi}{\mu + \alpha} + \dfrac{h\lambda}{\alpha(\mu + \alpha)} + \dfrac{c}{\alpha}.$

(ii) For all $i = 0, 1, \ldots$, $U_\alpha(i, 0)$ satisfies the optimality equation

$$U_\alpha(i, 0) = \min \left\{ s_1 + \frac{hi + c}{\alpha + \lambda + i\mu} + \frac{\lambda}{\alpha + \lambda + i\mu} U_\alpha(i + 1, 1) + \right.$$

$$\left. \frac{i\mu}{\alpha + \lambda + i\mu} U_\alpha(i - 1, 1), \frac{hi}{\alpha + \lambda} + \frac{\lambda}{\alpha + \lambda} U_\alpha(i + 1, 0) \right\}. \quad (5)$$

DEFINITION 1. *For an integer $n \geq 0$, a policy is called n-full service if it never turns the running sever off and turns the inactive server on if and only if there are n or more customers in the system. In particular, the 0-full service policy turns on the server at time 0, if it is off, and always keeps it on. A policy is called full service if and only if it is n-full service for some $n \geq 0$.*

THEOREM 3. *A policy ϕ is discount optimal within the class of the policies that never turn off the server if and only if*

$$\phi(i, 0) = \begin{cases} 1, & \text{if } i > A(\alpha), \\ 0, & \text{if } i < A(\alpha), \end{cases}$$

for all $i = 0, 1, \ldots$, where $A(\alpha) = \dfrac{(\mu + \alpha)(c + \alpha s_1)}{h\mu}$.

3.3 Reduction to Finite State Space and Existence of Average-Optimal Policies

Let $Z_1 = \{i \geq 0 : V_\alpha^0(i, 1) \leq V_\alpha^1(i, 1)\}$, define

$$M_\alpha^* = \begin{cases} \max Z_1, & \text{if } Z_1 \neq \emptyset, \\ -1, & \text{otherwise.} \end{cases} \quad (6)$$

$$N_\alpha^* = \min\{i > M_\alpha^* : V_\alpha^1(i, 0) \leq V_\alpha^0(i, 0)\}. \quad (7)$$

THEOREM 4. *For each $\alpha > 0$ either the n_α-full service policy is optimal there exists a stationary discount-optimal policy ϕ_α with the following properties:*

$$\phi_\alpha(i, \delta) = \begin{cases} 1, & \text{if } i > M_\alpha^* \text{ and } \delta = 1; \\ 1, & \text{if } i = N_\alpha^* \text{ and } \delta = 0; \\ 1, & \text{if } i \geq n_\alpha \text{ and } \delta = 0; \\ 0, & \text{if } i = M_\alpha^* \text{ and } \delta = 1; \\ 0, & \text{if } M_\alpha^* \leq i < N_\alpha^* \text{ and } \delta = 0. \end{cases} \quad (8)$$

4. STRUCTURE OF AVERAGE-OPTIMAL POLICIES

DEFINITION 2. *For two nonnegative integers M and N with $N > M$, a stationary policy is called an (M, N)-policy if*

$$\phi(i, \delta) = \begin{cases} 1, & \text{if } i > M \text{ and } \delta = 1; \\ 1, & \text{if } i \geq N \text{ and } \delta = 0; \\ 0, & \text{if } i \leq M \text{ and } \delta = 1; \\ 0, & \text{if } i < N \text{ and } \delta = 0. \end{cases}$$

THEOREM 5. *There exists a stationary average-cost optimal policy and, depending on the model parameters, either the n-full service policy is optimal for $n = 0, 1, \ldots$, or an (M, N)-policy is optimal for some $N > M \geq 0$ and $N \leq n^*$, where*

$$n^* = \lfloor \frac{c}{h} + 1 \rfloor. \quad (9)$$

In addition, the optimal average-cost value $v(i, \delta)$ is the same for all initial states (i, δ); that is $v(i, \delta) = v$.

5. CONDITIONS FOR AVERAGE-OPTIMALITY OF N-FULL SERVICE POLICY AND (M, N)-POLICIES

THEOREM 6. *The following statements hold for $n^* = \lfloor c/h + 1 \rfloor$:*

(i) when $c < \dfrac{\lambda}{n^*}(s_0 + s_1) + \dfrac{h(n^* - 1)}{2}$, any n-full service policy is average-optimal for $n = 0, 1, \ldots$;

(ii) when $c = \dfrac{\lambda}{n^*}(s_0 + s_1) + \dfrac{h(n^* - 1)}{2}$, both the n-full service policy, $n = 0, 1, \ldots$, and the $(0, n^*)$-policy are average-optimal;

(iii) when $c > \dfrac{\lambda}{n^*}(s_0 + s_1) + \dfrac{h(n^* - 1)}{2}$, an (M, N)-policy is average-optimal for some $N > M \geq 0$ and $n^* \geq N$.

6. COMPUTATIONS OF AN AVERAGE-OPTIMAL POLICY

According to Theorem 6, there is an optimal policy ϕ with $\phi(i, \delta) = 1$ when $i \geq n^* = \lfloor \frac{c}{h} + 1 \rfloor$. Thus, the goal is to find the optimal values of $\phi(i, \delta)$ when $i = 0, 1, \ldots, n * -1$ and $\delta = 0, 1$. To do this, we truncate the state space Z to $Z' = \{0, 1, \ldots, n^* - 1\} \times \{0, 1\}$. If the action 1 is selected at state $(n^* - 1, 1)$, the system moves to the state $(n^* - 2, 1)$, if the next change of the number of the customer in the system is a departure and the system remains in $(n^* - 1, 1)$, if an arrival takes place. In the latter case, the number of customers increases by one at the arrival time and then it moves according to the random work until it hits the state $(n^* - 1, 1)$ again. Thus the system can jump from the state $(n^* - 1, 1)$ to itself and therefore it cannot be described as an CTMDP. However, it can be described as a Semi-Markov Decision Process (SMDP) [14, Chapter 5].

We describe our problem as an SMDP with the state set Z' and action set $A(z) = A = \{0, 1\}$. If an action a is selected at state $z \in Z'$, the system spent an average time τ' in this state until it moves to the next state $z' \in Z'$ with the probability $p(z'|z, a)$ and during this time the expected cost $C'(z, a)$ is incurred. In states $z' = (i, \delta)$ with $i = 0, 1, \ldots, n^* - 2$ and $\delta = 0, 1$, these characteristics are the same as for the original CTMDP, for $z = (i, \delta)$ with $i = 0, 1, \ldots, n^* - 2$ and $\delta = 0, 1$,

$$p(z'|z, a) = \begin{cases} 1, & \text{if } a = 0, \ z' = (i + 1, 0), \\ \dfrac{\lambda}{\lambda + i\mu}, & \text{if } a = 1, \ z' = (i + 1, 1), \\ \dfrac{i\mu}{\lambda + i\mu}, & \text{if } a = 1, \ z' = (i - 1, 1), \\ 0, & \text{otherwise,} \end{cases} \quad (10)$$

$$\tau'((i, \delta), a) = \begin{cases} \dfrac{1}{\lambda}, & \text{if } a = 0, \\ \dfrac{1}{\lambda + i\mu} & \text{if } a = 1, \end{cases} \quad (11)$$

and $C'((i, \delta), a) = |a - \delta| s_a + (hi + ac)\tau'((i, \delta), a)$. The transition probabilities in states $(n^* - 1, \delta)$ with $\delta = 0, 1$ are defined by $p((n^* - 2, 1)|(n^* - 1, \delta), 1) = (n^* - 1)\mu/(\lambda + (n^* - 1)\mu)$, $p((n^* - 1, 1)|(n^* - 1, \delta), 1) = \lambda/(\lambda + (n^* - 1)\mu)$, and $p((n^* - 1, 1)|(n^* - 1, \delta), 0) = 1$. In the latter case, the number of customers increases by 1 to n^*, the system turns on, and the number of customers becomes $n^* - 1$.

Let T_i be the expected time between an arrival sees i customers in an $M/M/\infty$ queue and the next time when a departure leaves i customers behind, $i = 0, 1, \ldots$. Applying the memoryless property of the exponential distribution, $T_i = B_{i+1} - B_i$, where B_i is the

expected busy period for $M/M/\infty$ starting with i customers in the system and $B_0 = 0$. $B_i = \frac{1}{\lambda}\left(e^\rho - 1 + \sum_{k=1}^{i-1}\frac{k!}{\rho^k}\left(e^\rho - \sum_{j=0}^{k}\frac{\rho^j}{j!}\right)\right)$, by formula (34b) in [2],where $\rho = \frac{\lambda}{\mu}$. Thus $T_{n^*-1} = B_{n^*} - B_{n^*-1} = \frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{\rho^{k+1}}{n^*(n^*+1)\dots(n^*+k)}$. The expected time $\tau'((n^*-1,\delta),1)$, where $\delta = 0, 1$, is the expected time until the next arrival plus T_{n^*-1}, if the next event is an arrival. Thus, $\tau'((n^*-1,\delta),1) = \frac{\lambda}{\lambda+(n^*-1)\mu}\left(\frac{1}{\lambda}+T_{n^*-1}\right)$, $\delta = 0, 1$. In addition $\tau'((n^*-1,\delta),0) = \frac{1}{\lambda} + T_{n^*-1}$, $\delta = 0, 1$.

To compute the one-step cost $C'((n^*-1,1),1)$, we define m_i as the average number of visits to state $(i,1)$ starting from state $(n^*-1,1)$ before returning to $(n^*-1,1)$, $i = n^*-1, n^*, \dots$. And define $m_{i,i+1}$ as the expected number of jumps from $(i,1)$ to $(i+1,1)$, $i = n^*-1, n^*, \dots$, and $m_{i,i-1}$ as the expected number of jumps from $(i,1)$ to $(i-1,1)$, $i = n^*, n^*+1, \dots$. Then $m_{i,i+1} = \frac{\lambda}{\lambda+i\mu}m_i$, $m_{i,i-1} = \frac{i\mu}{\lambda+i\mu}m_i$ and $m_{i,i+1} = m_{i+1,i}$. Since $m_{n^*-1} = 1$, $m_i = \prod_{j=0}^{i-n^*}\frac{\lambda}{\lambda+(n^*-1+j)\mu}\frac{\lambda+(n^*+j)\mu}{(n^*+j)\mu}$, $i = n^*, n^*+1, \dots$. Thus $C'((n^*-1,1),1) = \sum_{i=n^*-1}^{\infty}m_iC((i,1),1) = \sum_{i=n^*-1}^{\infty}m_i\frac{hi+c}{\lambda+i\mu}$, where $C((i,1),1) = \frac{hi+c}{\lambda+i\mu}$, $i = n^*-1, n^*, \dots$ is the cost incurred in state $(i,1)$ under action 1 for the original state space model; see Section 3.1. The one-step cost $C'((n^*-1,0),1) = s_1 + C'((n^*-1,1),1)$. Let C_{n^*} be the total cost incurred in $M/M/\infty$ until the number of customers becomes (n^*-1) if at time 0 there are n^* customers in the system and the system is running. Then $C'((n^*-1),1) = \frac{h(n^*-1)+c}{\lambda+(n^*-1)\mu} + \frac{\lambda}{\lambda+(n^*-1)\mu}C_{n^*}$, and this implies $C_{n^*} = \left(1 + \frac{(n^*-1)\mu}{\lambda}\right)C'((n^*-1),1) - \frac{h(n^*-1)+c}{\lambda}$. We also have $C'((n^*-1,0),0) = \frac{h(n^*-1)}{\lambda} + s_1 + C_{n^*}$ and $C'((n^*-1,1),0) = s_0 + C'((n^*-1,0),0)$.

We formulate the LP according to Section 5.5 in [14] as

$$\text{Minimize} \sum_{z\in Z'}\sum_{a\in A}C'(z,a)x_{z,a}$$

$$\text{s.t.} \quad \sum_{a\in A(z)}x_{z,a} - \sum_{z'\in Z'}\sum_{a\in A(z')}p(z|z',a)x_{z',a} = 0, \; z \in Z', \quad (12)$$

$$\sum_{z\in Z'}\sum_{a\in A(z)}\tau'(z,a)x_{z,a} = 1, \; x_{z,a} \geq 0, \; z\in Z', \; a\in A$$

Let x^* be the optimal basic solution of (12). According to general results on SMDPs [3, Section III], for each $z \in Z'$, there exists at most one $\delta = 0, 1$ such that $x^*_{z,a} > 0$. If $x^*_{z,a} > 0$, then for the optimal policy ϕ, $\phi(z) = a$, for $a = 0, 1$. If $x^*_{z,0} = x^*_{z,1} = 0$, then $\phi(z)$ can be either 0 or 1. Theorem 7 explains how to use x^* to construct an average-optimal policy ϕ with the properties in Theorem 6.

THEOREM 7. *For the optimal basic solution x^* of (12) and $z = (i,\delta) \in Z'$, $i = 0, 1, \dots, n^*-1$, $\delta = 0, 1$,*

(i) If $x^*_{(0,1),1} > 0$, then any n-full service policy is average-optimal, $n = 0, 1, \dots$.

(ii) Let $Z'_1 = \{i = 1, \dots, n^*-1 : x^*_{(i,0),1} > 0\}$. If $x^*_{(0,1),0} > 0$, then

the $(0, N)$-policy is average-optimal with N defined as

$$N = \begin{cases} n^*, & \text{if } \min Z'_1 = \emptyset; \\ \min Z'_1, & \text{if } \min Z'_1 \neq \emptyset. \end{cases} \quad (13)$$

(iii) If $x^*_{(0,1),0} = x^*_{(0,1),1} = 0$, then the (M, N)-policy is average-optimal with $M = \min\{i = 1, \dots, n^*-1 : x^*_{(i,1),0} > 0\} > 0$ and N being the same as in (13).

7. REFERENCES

[1] Amazon Elastic Compute Cloud, User Guide. http://aiweb.techfak.uni-bielefeld.de/content/bworld-robot-control-software/, December 2012. Online; accessed 03/01/2013.

[2] S. Browne and O. Kella. Parallel service with vacations. *Operations Research*, 43(5):870–878, 1995.

[3] E. Denardo. On linear programming in a Markov Decision Problem. *Management Science*, 16(5):281–288, 1970.

[4] E. Feinberg. Reduction of discounted continuous-time MDPs with unbounded jump and reward rates to discrete-time total-reward MDPs. In *In Optimization, Control, and Applications of Stochastic Systems, D. Hernaández-Hernaández and J.A. Minjárez-Sosa Birkhäuser (Eds.)*, pages 77–97. Springer, New York, 2012.

[5] Google Compute Engine. https://cloud.google.com/products/compute-engine, 2013.

[6] A. Greenberg, J. Hamilton, D. Maltz, and P. Patel. The cost of a cloud: research problems in data center networks. *ACM SIGCOMM Computer Communication Review*, 39(1):68–73, 2008.

[7] X. Guo and W. Zhu. Denumerable-state continuous-time Markov Decision Processes with unbounded transition and reward rates under the discounted criterion. *Journal of Applied Probability*, 39(2):233–250, 2002.

[8] D. Heyman. Optimal operating policies for M/G/1 queuing systems. *Operations Research*, 16(2):362–382, 1968.

[9] HP Cloud Compute Overview. https://docs.hpcloud.com/compute, 2013.

[10] Ibm SmartCloud. www.ibm.com/cloud, 2012. Online; accessed 03/01/2013.

[11] Ibm smartcloud provisioning, version 2.1. http://pic.dhe.ibm.com/infocenter/tivihelp/v48r1/topic/com.ibm.scp.doc_2.1.0/Raintoc.pdf, 2009.

[12] H. Khazaei, J. Misic, and V. Misic. Performance analysis of cloud computing centers using m/g/m/m+ r queuing systems. *Parallel and Distributed Systems, IEEE Transactions on*, 23(5):936–943, 2012.

[13] V. Kulkarni, S. Kumar, V. Mookerjee, and S. Sethi. Optimal allocation of effort to software maintenance: A queuing theory approach. *Production and Operations Management*, 18(5):506–515, 2009.

[14] M. Mazzucco, D. Dyachuk, and R. Deters. Maximizing cloud providers' revenues via energy aware allocation policies. In *Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on*, pages 131–138, 2010.

[15] H. Mine and S. Osaki. *Markovian Decision Processes*, volume 25. Elsevier Publishing Company, New York, 1970.

[16] Introducing the windows azure platform. David Chappell & Associates White Paper, 2010.