The HKU Scholars Hub   The University of Hong Kong   香港大學學術庫

Title	**A Study of Competitive Cloud Resource Pricing under a Smart Grid Environment**
Author(s)	**Jin, X; Kwok, YK; Yan, Y**
Citation	**The 5th IEEE International Conference on Cloud Computing Technology and Science (CloudCom 2013), Bristo, United Kingdom, 2-5 December 2013. In the Proceedings of the IEEE International Conference on Cloud Computing Technology and Science, 2013, v. 1, p. 655-662, article no. 6753858**
Issued Date	**2013**
URL	**http://hdl.handle.net/10722/204073**
Rights	**Proceedings of the IEEE International Conference on Cloud Computing Technology and Science. Copyright © I E E E, Computer Society.**

Competitive Cloud Resource Pricing under A Smart Grid Environment

Xin Jin, Yu-Kwong Kwok, and Yong Yan
Department of Electrical and Electronic Engineering
The University of Hong Kong
Pokfulam, Hong Kong
Email: {tojinxin,ykwok}@eee.hku.hk, fuxin04@gmail.com

Abstract—In the current IaaS cloud market, to achieve profit maximization, multiple cloud providers compete non-cooperatively by offering diverse price rates. At the same time, tenant consumers judiciously adjust demands accordingly, which in turn affects cloud resource prices. In this paper, we tackle this fundamental but daunting cloud price competition problem with Bertrand game modeling, and propose a dynamic game to achieve Nash equilibrium in a distributed manner. Specifically, we realistically consider spot electricity prices under a smart grid environment, and systematically investigate the impact of different system parameters such as network delay, renewable availability, and cloud resource substitutability. We also perform stability analysis to investigate the convergence of the proposed dynamic game to Nash equilibrium. Cooperation among cloud providers can achieve aggregate cloud profit maximization, but is subject to strategic manipulations. We then propose our *Striker* strategy to stimulate cooperation, the efficiency of which is validated by repeated game analysis. Our evaluation is augmented with realistic electricity prices in the spot energy market, and reveals insightful observations for both theoretic analysis and practical pricing scheme design.

Keywords—Cloud computing, competitive resource pricing, resource allocation, game theory, dynamic game.

I. INTRODUCTION

Cloud computing has fundamentally transformed the way of business operations in many industries. In particular, the Infrastructure-as-a-Service (IaaS) view is adopted by large companies such as Amazon, Google, and Microsoft [1]–[3] to deploy Internet-scale data centers where Internet users, small startups, and even large service providers (e.g., Hulu) can host their applications by dynamically renting computing and storage resources as tenant consumers [4]–[7]. Specifically, cloud providers offer cloud resources such as CPU, memory, and bandwidth as sellers. Tenants as buyers dynamically access cloud resources in bundles of virtual instances. Currently, serveral large cloud facilities such as Amazon EC2 [1], Google App Engine [2], and Microsoft Azure [3] dominate the entire market. Therefore, in such an oligopoly cloud market, the few cloud providers as oligopolists compete strategically in terms of offered price rates, to achieve their own profit maximization.

Instance prices intrinsically dictate resource allocation in IaaS clouds. Cloud price competition, largely unexplored, fundamentally determines instance price dynamics, which in turn affects tenant demand variations. Such tenant demand variations further influence pricing strategies of all cloud providers. In cloud pricing scheme design, cloud providers should consider not only profit maximization, but also optimal tenant demand responses. Therefore, a study of cloud price competition is challenging but helps to better understand the sustainable profitability of the cloud business. To further aggravate the problem, in a smart grid environment, temporal and spatial variations of electricity prices expose cloud providers to the risk of operational cost fluctuations.

In this paper, we tackle the problem of competitive cloud resource pricing by proposing a non-cooperative game to tractably investigate the price competition among cloud providers and its impact on cloud profit, tenant surplus, and instance prices. Under the practical scenario in which prices of different cloud providers are observable, the Bertrand game model for price competition is used to analyze and derive the equilibrium prices for a cloud market consisting of multiple cloud providers. To the best of our knowledge, we are among the first to study competitive cloud resource pricing. Specifically, our contributions are three-fold.

Firstly, we build a general model to realistically capture cloud resource pricing scheme design (Section II). Tenants make optimal demand response decisions to maximize their surplus (i.e., tenant utility minus dollar cost), given experienced service qualities and instance prices. Bearing tenant optimal demands and instance prices of the other cloud providers, one cloud provider optimizes its pricing decisions for profit maximization. Rigorous equilibrium analysis is provided, together with a dynamic game based on bounded rationality for cloud providers to achieve equilibrium prices in a distributed manner using local information only. Stability analysis is performed to investigate the convergence of the dynamic game.

Secondly, cloud providers can achieve higher profits than their equilibrium profits via cooperation, and attain aggregate profit maximization at the same time (Section III). This is credible in that only few cloud providers exist and compete in an oligopoly cloud market. However, in a one-shot static game, all the cloud providers adopt Nash equilibrium prices due to strategic interactions. Therefore, we model strategic cloud pricing as a repeated game, based on which we propose our Striker strategy for coercing cloud providers to cooperate

IEEE
computer
society

in instance pricing. The key idea of the Striker strategy is to provide enough threats to selfish cloud providers and thwart them from deviating from cooperation.

Thirdly, we conduct extensive performance evaluation to validate our analytical model and obtain insightful observations (Section IV). Our evaluation is augmented with realistic electricity prices in spot energy markets. We systematically investigate the impact of network delay, resource substitutability, and electricity prices on equilibrium prices. For instance, a negative correlation exists between instance prices and resource substitutability. It is validated that the dynamic game converges to Nash equilibrium, and that there is a tradeoff between cloud profit and tenant surplus. We have also obtained conditions of effective threatening for our Striker strategy.

II. CLOUD RESOURCE PRICING COMPETITION AND EQUILIBRIUM ANALYSIS

In this section, we build a game theoretic model for competitive cloud resource pricing. We first discuss tenant surplus and optimal demand response, followed profit analysis of cloud providers. Bertrand game is proposed to model cloud price competition, and dynamic game is used to achieve Nash equilibrium in a distributed manner.

A. Tenant Surplus and Optimal Demand Responses

To quantify payoffs obtained from resource consumption, tenant utility explicitly considers demand responses and the experienced service qualities. The service quality is dependent on both network delay (i.e., transmission delay due to request routing) and queueing delay (i.e., delay incurred by waiting for cloud service). Due to the illusion of infinite capacity in cloud computing, we assume no queueing delay in our analysis.

Denote by N the number of cloud providers in the cloud market. Cloud provider i sells cloud resources at price rate p_i per virtual instance. Denote by d_i the demand from cloud provider i. Then, $\mathbf{d} = \{d_1, \cdots, d_i, \cdots, d_N\}$ is the vector of demands from all cloud providers. Denote by γ_i the incurred network delay due to resource consumption from cloud provider i. Γ represents the maximum experienced network delay. Then, the payoff of unit virtual instance can be modeled as:

$$b_i = K \cdot \ln\left(1 + (\Gamma - \gamma_i)\right), \tag{1}$$

where K is a constant. Recall that p_i is the price offered by cloud provider i. Then, we define the tenant i's surplus as the following commonly adopted quadratic function:

$$U(\mathbf{d}) = \sum_{i=1}^{N} d_i \cdot b_i - \frac{1}{2} \cdot \left(\sum_{i=1}^{N} d_i^2 + 2\mu \cdot \sum_{i \neq j} d_i \cdot d_j\right) - \sum_{i=1}^{N} d_i \cdot p_i \tag{2}$$

Our model considers resource substitutability through parameter $\mu \in [-1.0, 1.0]$. When $\mu = 1.0$, the tenant user can freely switch among multiple cloud resource providers for virtual instance reservation. The function is concave, and thus reflects the law of diminishing return, with the saturation of user satisfaction as the amount of resource reservation

increases. We also consider the service quality of different cloud providers via parameter b_i.

THEOREM 1. The optimal demand response of tenant i is given by:

$$D_i(\mathbf{p}) = \frac{(b_i - p_i)\left(\mu\left(N - 1\right) + 1\right) - \mu \sum_{j=1}^{N}\left(b_j - p_j\right)}{(1 - \mu)\left(\mu\left(N - 1\right) + 1\right)}, \tag{3}$$

where $\mathbf{p} = \{p_1, \cdots, p_i, \cdots, p_N\}$ is the vector of prices offered by all the cloud providers.

Proof. To derive the optimal demand of the tenant user at cloud provider i, we differentiate $U(\mathbf{b})$ with respect to the demand level d_i:

$$\frac{\partial U(\mathbf{b})}{\partial d_i} = b_i - d_i - \mu \sum_{j \neq i} d_j - p_i = 0. \tag{4}$$

By solving the above equations, we can obtain the optimal demand response. $\square$

B. Profit of Cloud Providers

The profit of cloud providers is the revenue collected from tenants minus the operating cost. We can easily get the revenue from the sale of cloud resources:

$$R_i = p_i \cdot d_i. \tag{5}$$

To obtain operating cost, we consider energy cost, which constitutes the majority of the operating cost. We consider both electricity draw from the utility grid, and available renewables. The electricity price Pr_i has both temporal and spatial variations under a smart grid environment. The renewable availability for cloud provider is r, which is in the units of the number of virtual instances that can be powered. Denote by E_i renewable prices at cloud provider i. Then, we obtain the energy cost as expressed below:

$$C_i = Pr_i \cdot (d_i - r)^+ + E_i \cdot \min(r, d_i), \tag{6}$$

where $(x)^+ = \max(x, 0)$.

Therefore, cloud provider i's profit is given by

$$\begin{aligned} P_i(\mathbf{p}) &= R_i - C_i \\ &= p_i \cdot d_i + Pr_i \cdot (d_i - r)^+ + E_i \cdot \min(r, d_i). \end{aligned} \tag{7}$$

In practice, the aggregate tenant consumption is no smaller than renewable provision. That is, cloud providers have to procure energy from the electricity market. Then, we have

$$\begin{aligned} P_i(\mathbf{p}) &= R_i - C_i \\ &= p_i \cdot d_i - Pr_i \cdot (d_i - r) - E_i \cdot r. \end{aligned} \tag{8}$$

C. Bertrand Game Model

Based on the above analysis, we use a Bertrand game to model competitive pricing. The players in this game are cloud providers. The strategy of each player (i.e., cloud provider i) is the price per virtual instance. The payoff of each cloud provider i is the profit earned from the sale of cloud resources. The solution concept is Nash equilibrium.

The Nash equilibrium of a game is a solution that no player can increase his own payoff by unilaterally choosing a different strategy. The Nash equilibrium can be obtained by using the best response function, which is the optimal strategy of one player given the others' strategy choices. That is, the best response function of cloud provider i can be formulated as:

$$BR_i(\mathbf{p}_{-i}) = \arg\max_{p_i} P_i(\mathbf{p}), \quad (9)$$

where $\mathbf{p} = \mathbf{p}_{-i} \cup \{p_i\}$ and $\mathbf{p}_{-i}$ is the set of prices offered by cloud providers other than i.

Denote by $\mathbf{p}^* = \{p_1^*, \cdots, p_i^*, \cdots, p_N^*\}$ the Nash equilibrium of the pricing game. Then, we have:

$$p_i^* = BR_i(\mathbf{p}_{-i}^*), \quad \forall i \quad (10)$$

where $\mathbf{p}_{-i}^*$ denotes the set of best responses for all the cloud providers other than i. To this end, we can obtain the Nash equilibrium by solving the following equation array:

$$\frac{\partial P_i(\mathbf{p})}{\partial p_i} = 0. \quad (11)$$

By definition of cloud profit, the marginal profit function is:

$$\frac{\partial P_i(\mathbf{p})}{\partial p_i} = D_i(\mathbf{p}) + (p_i - Pr_i) \cdot \frac{\partial D_i(\mathbf{p})}{\partial p_i}. \quad (12)$$

From optimal demand response given by Equation 3, we get:

$$\frac{\partial D_i(\mathbf{p})}{\partial p_i} = -\frac{\mu(N-2)+1}{(1-\mu)(\mu(N-1)+1)}. \quad (13)$$

The solution $\mathbf{p}^*$ of equations given by $\frac{\partial P_i(\mathbf{p})}{\partial p_i} = 0$ is a Nash equilibrium. In practice, the cloud providers set prices using the Nash equilibrium, and the optimal demand response of the tenant user can be obtained from the demand function $D_i(\mathbf{p}^*)$.

THEOREM 2. The Nash equilibrium solution $\mathbf{p}^*$ is given by

$$p_i^* = \frac{Y \cdot Q - X \cdot b_i + Z \cdot Pr_i}{Z - X}, \quad (14)$$

where $X = \frac{1}{1-\mu}$, $Y = \frac{\mu}{(1-\mu)(\mu(N-1)+1)}$, $Z = -\frac{\mu(N-2)+1}{(1-\mu)(\mu(N-1)+1)}$, and $Q = \frac{Z\sum_{j=1}^{N}(b_j - Pr_j)}{Z+YN-X}$.

Proof. From Equation 3, we get:

$$
\begin{aligned}
D_i(\mathbf{p}) &= \frac{b_i - p_i}{1-\mu} - \frac{\mu}{(1-\mu)(\mu(N-1)+1)} \cdot \sum_{j=1}^{N}(b_j - p_j) \\
&= X(b_i - p_i) - Y\sum_{j=1}^{N}(b_j - p_j), \quad (15)
\end{aligned}
$$

where $X = \frac{1}{1-\mu}$ and $Y = \frac{\mu}{(1-\mu)(\mu(N-1)+1)}$.

Substitute the above into Equation 11 and Equation 12. We have

$$\frac{\partial P_i(\mathbf{p})}{\partial p_i} = X(b_i - p_i) - Y\sum_{j=1}^{N}(b_j - p_j) + Z(p_i - Pr_i), \quad (16)$$

where $Z = \frac{\partial D_i(\mathbf{p})}{\partial p_i} = -\frac{\mu(N-2)+1}{(1-\mu)(\mu(N-1)+1)}$. Then, we get:

$$
\begin{aligned}
\sum_{i=1}^{N}\frac{\partial P_i(\mathbf{p})}{\partial p_i} &= (X - YN - Z)\sum_{j=1}^{N}(b_j - p_j) \\
&\quad + Z\sum_{j=1}^{N}(b_j - Pr_j) \\
&= 0. \quad (17)
\end{aligned}
$$

Then, we have

$$\sum_{j=1}^{N}(b_j - p_j) = \frac{Z\sum_{j=1}^{N}(b_j - Pr_j)}{Z + YN - X}. \quad (18)$$

Substituting the above equation into Equation 16, we have

$$p_i = \frac{Y \cdot Q - X \cdot b_i + Z \cdot Pr_i}{Z - X}, \quad (19)$$

where $Q = \frac{Z\sum_{j=1}^{N}(b_j - Pr_j)}{Z+YN-X}$. $\qquad\square$

D. Dynamic Bertrand Game and Best Response Algorithms

To obtain the Nash equilibrium in Theorem 2, we need complete information of all cloud providers. However, in practice, one cloud provider may not know others' profit information. To this end, we propose distributed learning algorithms for dynamic price adjustments so as to gradually achieve Nash equilibrium for competitive pricing.

Denote by $p_i(t)$ the price offered by cloud provider i at time slot t. The price vectors $\mathbf{p}_{-i}(t)$ and $\mathbf{p}(t)$ are defined similarly. The pricing strategies of cloud providers approach to Nash equilibrium via iterative strategy updating. The ideal case is that the pricing strategies of all the other cloud providers at time t is observable by cloud provider i. Then, we have the best response algorithm:

$$p_i(t+1) = BR_i(\mathbf{p}_{-i}(t)), \quad \forall i. \quad (20)$$

However, as discussed above, in a realistic cloud system, the assumption of perfect information may not be substantiated. Therefore, each cloud provider can employ only local information and tenant demands to adapt offered prices. Intuitively, each cloud provider should adjust its strategy in the direction of profit maximization. Then, we can derive the strategy in the next iteration based on the price in the current iteration:

$$p_i(t+1) = p_i(t) + \delta_i \cdot \left(\frac{\partial P_i(\mathbf{p}(t))}{\partial p_i(t)}\right), \quad (21)$$

where δ_i is the updating step and determines the learning rate of the iterative algorithm. This dynamic game is based on the concept of bounded rationality, which does not update the strategy to the optimal one immediately but approaches to the optimal one gradually. This is reasonable when the collected information about the environment is not reliable or accurate enough.

E. Stability Analysis

Stability analysis is critical for both the dynamic learning algorithms to ensure that the Nash equilibrium can be achieved at the steady state. We analyze the stability of both algorithms by considering the eigenvalues of the Jacobian matrix of the self-mapping functions in Equation 20 and Equation 21. In particular, the dynamic algorithm is stable if and only if the eigenvalues λ_i are all inside the unit circle in the complex plane (i.e., $|\lambda_i| < 1$). In our analysis, the Jacobian matrix is calculated as follows:

$$
\mathbf{J} = \begin{bmatrix} \frac{\partial p_1(t+1)}{\partial p_1(t)} & \frac{\partial p_1(t+1)}{\partial p_2(t)} & \cdots & \frac{\partial p_1(t+1)}{\partial p_N(t)} \\ \frac{\partial p_2(t+1)}{\partial p_1(t)} & \frac{\partial p_2(t+1)}{\partial p_2(t)} & \cdots & \frac{\partial p_2(t+1)}{\partial p_N(t)} \\ \vdots & \vdots & & \vdots \\ \frac{\partial p_N(t+1)}{\partial p_1(t)} & \frac{\partial p_N(t+1)}{\partial p_2(t)} & \cdots & \frac{\partial p_N(t+1)}{\partial p_N(t)} \end{bmatrix}. \quad (22)
$$

Denote by j_{mn} the element at row m and column n in the Jacobian matrix. We employ the case of two cloud providers to perform stability analysis for clarity. The same approach can be applied to the case of arbitrary number of cloud providers for stability analysis.

Denote the matrix form of the dynamic game in Equation 20 by:

$$
\mathbf{p}(t+1) = F_1(\mathbf{p}(t)), \quad (23)
$$

and the matrix form of the dynamic game in Equation 21 by:

$$
\mathbf{p}(t+1) = F_2(\mathbf{p}(t)), \quad (24)
$$

where $F_1(\cdot)$ and $F_2(\cdot)$ are the corresponding self-mapping functions.

For both dynamic algorithms, at the equilibrium, we have $\mathbf{p}(t+1) = \mathbf{p}(t) = \mathbf{p}$, where $\mathbf{p}$ is the fixed point. Then, we have:

$$
\mathbf{p} = F_k(\mathbf{p}), \quad \forall k \in \{1, 2\}. \quad (25)
$$

In the following, we consider the two dynamic games respectively. For clarity, we consider the special case of two cloud providers (i.e., $N = 2$). From Theorem 2, we have the Nash equilibrium:

$$
\mathbf{p}^* = \left(\frac{Y \cdot Q - X \cdot b_1 + Z \cdot Pr_1}{Z - X}, \frac{Y \cdot Q - X \cdot b_2 + Z \cdot Pr_2}{Z - X} \right). \quad (26)
$$

Algorithm 1. First, we consider the bet response algorithm, (i.e., that defined by Equation 20). From Equation 25, we can obtain the fixed point, if any, by solving the following equations:

$$
p_i = F_1(p_i), \quad \forall i, \quad (27)
$$

which obviously gives the Nash equilibrium as the solution for the fixed point (i.e., $\mathbf{p} = \mathbf{p}^*$). From equations given by 16, we also have

$$
p_i(t+1) = BR_i(\mathbf{p}_{-i}(t)) = \frac{Y \sum_{j \neq i}^{N}(b_j - p_j) + ZPr_i - (X - Y)b_i}{Y + Z - X}, \quad (28)
$$

from which, we obtain the Jacobian matrix of the algorithm:

$$
\mathbf{J}_1 = \begin{bmatrix} 0 & -\frac{Y}{Y+Z-X} \\ -\frac{Y}{Y+Z-X} & 0 \end{bmatrix}, \quad (29)
$$

which is a diagonal matrix and the diagonal elements are the two eigenvalues $\lambda_i = -\frac{w_i}{x_i}$, $\forall i \in \{1, 2\}$. The fixed point is stable if and only if $|\lambda_i| < 1$. Therefore, the stability condition is $|-\frac{Y}{Y+Z-X}| < 1$. That is, $|\frac{\mu}{2\mu(N-2)+2}| < 1$, implies that the fixed point of Nash equilibrium is always stable.

Algorithm 2. Second, we consider the dynamic game requiring no complete information (i.e., that defined by Equation 21). From Equation 25, we can obtain the fixed point, if any, by solving the following equations:

$$
p_i = F_2(p_i), \quad \forall i, \quad (30)
$$

which gives

$$
\delta_i \cdot \left(\frac{\partial P_i(\mathbf{p})}{\partial p_i} \right) = 0. \quad (31)
$$

Therefore, the unique fixed point is also the Nash equilibrium:

$$
\mathbf{p} = \mathbf{p}^*. \quad (32)
$$

Again, from equations given by Equation 16 and

$$
p_i(t+1) = p_i(t) + \delta_i \cdot \frac{\partial P_i(\mathbf{p}(t))}{\partial p_i(t)}, \quad (33)
$$

$$
(34)
$$

we obtain the Jacobian matrix of the algorithm:

$$
\mathbf{J}_2 = \begin{bmatrix} 1 + \delta_1(Z + Y - X) & \delta_1 Y \\ \delta_2 Y & 1 + \delta_2(Z + Y - X) \end{bmatrix}, \quad (35)
$$

which is neither diagonal nor triangular. Therefore, we derive the eigenvalues from the characteristic equation $\lambda^2 - \lambda(j_{11} + j_{22}) + (j_{11}j_{22} - j_{12}j_{21}) = 0$, the roots of which are the eigenvalues of the Jacobian matrix:

$$
\begin{aligned} (\lambda_1, \lambda_2) &= \frac{(j_{11} + j_{22}) \pm \sqrt{4j_{12}j_{21} + (j_{11} - j_{22})^2}}{2} \\ &= 1 + \frac{(\delta_1 + \delta_2)(Z + Y - X)}{2} \\ &\quad \pm \frac{\sqrt{4\delta_1\delta_2 Y^2 + (\delta_1 - \delta_2)^2(Z + Y - X)^2}}{2} \end{aligned} \quad (36)
$$

The stability condition is again $|\lambda_i| < 1$. The condition is determined by X, Y, and Z, which is independent of electricity prices and the benefit obtained by the tenant users from the reserved resources. However, the stability condition is highly dependent on the resource substitutability (i.e., μ) and the updating step sizes (i.e., δ_i).

III. COERCION FOR AGGREGATE PROFIT MAXIMIZATION

A. Optimal Pricing for Aggregate Profit Maximization

The total profit for all the cloud providers is given by $\sum_{i=1}^{N} P_i(\mathbf{p})$. The optimal price for all the cloud providers can therefore be obtained by solving the following problem:

$$
\max_{\mathbf{p} \succeq 0} \sum_{j=1}^{N} P_j(\mathbf{p}). \quad (37)
$$

The problem can be solved by solving the following linear equations:

$$\frac{\partial \sum_{j=1}^{N} P_j(\mathbf{p})}{\partial p_i} = 0, \quad \forall i. \tag{38}$$

This gives optimal prices, different from equilibrium prices achieved in the Bertrand game (refer to Theorem 2).

THEOREM 3. The optimal prices for cloud profit maximization is given by:

$$\overline{p}_i^* = \frac{b_i + Pr_i}{2}. \tag{39}$$

Proof. From Equation 3 and Equation 8, we have

$$\sum_{j=1}^{N} P_j(\mathbf{p}) = \sum_{j=1}^{N} p_j D_j(\mathbf{p}) - \sum_{j=1}^{N} Pr_j(D_j(\mathbf{p}) - r) - \sum_{j=1}^{N} E_j r.$$

From Equation 15, we have

$$\frac{\partial D_j(\mathbf{p})}{\partial p_i} = Y, \quad \forall i \neq j, \tag{40}$$

$$\frac{\partial D_i(\mathbf{p})}{\partial p_i} = Y - X, \quad \forall i. \tag{41}$$

Then, we get the marginal function of the total profit:

$$\begin{aligned} \frac{\partial \sum_{j=1}^{N} P_j(\mathbf{p})}{\partial p_i} &= D_i(\mathbf{p}) + \sum_{j=1}^{N} (p_j - Pr_j) \cdot Y \\ &\quad - (p_i - Pr_i) \cdot X \\ &= 0. \end{aligned} \tag{42}$$

Substituting Equation 15 into the above equation, we obtain

$$\begin{aligned} 0 &= 2X(b_i - p_i) - X(b_i - Pr_i) - 2Y \sum_{j=1}^{N} (b_j - p_j) \\ &\quad + Y \sum_{j=1}^{N} (b_j - Pr_j), \quad \forall i. \end{aligned} \tag{43}$$

Summing up the left side and right side of the above equations, we get

$$\sum_{j=1}^{N} (b_j - p_j) = \frac{\sum_{j=1}^{N} (b_j - Pr_j)}{2}. \tag{44}$$

From the above two equations, we obtain

$$\overline{p}_i^* = \frac{b_i + Pr_i}{2}. \tag{45}$$

$\square$

1) Non-Cooperation in a Static Game: As demonstrated above, optimal prices for aggregate profit maximization are different from resource prices in the Nash equilibrium. Therefore, in a static game (i.e., the game is played only once), one or more of cloud providers can boost the experienced utility by unilaterally deviating from the optimal prices incurred by cooperation. Denote by C the cooperative strategy (i.e., adopting optimal prices for aggregate profit maximization), and D the non-cooperative strategy (i.e., adopting the deviation

prices using best response function). Denote by N the strategy to adopt the equilibrium prices. By definition, we get the following theorem.

THEOREM 4. There exists a unique N.E. $\langle N, N \rangle$ for the static cloud resource pricing game.

B. Coercing Non-Cooperative Cloud Providers

In practice, the cloud resource pricing game is played repeatedly among cloud providers. In this section, we demonstrate the feasibility of coercion leading to cooperation in this repeated cloud pricing game. Then, we propose a novel *Striker strategy* to coerce non-cooperative cloud providers to cooperate. The crux is to provide enough threat to non-cooperative behaviors so as to prevent strategic deviation from cooperation.

1) Repeated Game Modeling: We now model the repeated aspects in the cloud pricing game. Denote by $P_i(t)$ the profit of cloud provider i for strategy $s_i(t)$ taken in time round t. We utilize the discounted average of cloud profits obtained in different time periods to model and evaluate cloud utilities. Thus, we can evaluate an arbitrary strategy sequence $\{s_i(t), 0 \leq t \leq T\}$ by

$$\overline{P}_i = \sum_{t=0}^{T} \left[P_i(t) \cdot \lambda_i^{\ t} \right]. \tag{46}$$

In a realistic cloud system, T denotes how long they care about the future. The discount factor, denoted as λ_i ($0 \leq \lambda_i \leq 1$), models the shadow of the future (i.e., the importance of profits obtained from subsequent moves relative the previous move). In this paper, we adopt the infinitely repeated game (i.e., $T = +\infty$) because all cloud providers have no idea about exactly when the game will stop (i.e., the ending date of the cloud business).

2) Coercion Feasibility: Before further discussions on incentive scheme design for coercion leading to cooperation, we need to first answer whether it is feasible to enforce an N.E. in which both players are cooperative with our infinite game modeling. This is important because according to the backward induction principle, $\langle N, N \rangle$ is also N.E. in a finitely repeated game with the game termination time explicitly known to all cloud providers. Fortunately, we can always enforce a strategy path yielding players payoffs larger than the minmax payoff in N.E. The minmax payoff is defined below [8].

DEFINITION 1. The *minmax payoff* of player x is defined as

$$\min_{s_y \in S_y} \left(\max_{s_x \in S_x} P_x(s_x, s_y) \right), \tag{47}$$

where $x \in \{\tau, i\}$ and $\{y\} = \{\tau, i\} \setminus \{x\}$. Here, S_x is the strategy space of player x.

This proves the feasibility to enforce an N.E. $\langle C, C \rangle$, which yields the profit, larger than the minmax payoff. Indeed, the minmax payoff is the profit obtained in the N.E. in the static game (i.e., $\langle N, N \rangle$).

3) Striker Strategy: In a repeated game, the cloud providers play the pricing game for multiple rounds, and the outcome of the previous play is observable by other players. Consequently, the cloud providers may adjust their actions by coordinating with each other so as to achieve the desirable outcome. Therefore, we propose our *Striker strategy*, taken by cloud providers for coercing others to cooperation. In particular, cloud providers apply limited punishments as threatening to non-cooperative ones: $\forall t \geq 0$, cloud providers select

1. strategy C in time round $t+1$, if all others adopt strategy C in time round t;
2. strategy N in time rounds from $t+1$ to $t+T_i$, if some cloud providers adopt strategy N in time round t.

C. Effective Threatening

Responding to the threat, non-cooperative cloud providers can continue the non-cooperation or switch to the choice of cooperation. This depends on the experienced discounted profits. Intuitively, when the punishment intensity is large enough, non-cooperative ones do not have incentives to deviate from cooperation.

DEFINITION 2. We define a threat to be *effective* if non-cooperative cloud providers switch to cooperation after the threat.

We now derive conditions on which threatening in our Striker strategy is effective. For cloud provider i, denote by P_i^c, P_i^d, P_i^p the profits obtained by adopting cooperation strategy C, deviation strategy D, and punishment strategy N (i.e., Nash equilibrium profits), respectively. Denote by $\widetilde{P}_i^d$ the profits obtained by a cooperative cloud provider when some others adopt strategy D.

THEOREM 5. If and only if

$$\frac{P_i^d - P_i^c}{P_i^c - P_i^p} \leq \frac{\lambda_i \cdot \left(1 - \lambda_i^{T_i}\right)}{1 - \lambda_i}, \quad (48)$$

the striker strategy taken by cloud providers yields a unique and strict N.E. $\langle C, C \rangle$.

Proof. We have three cases.

Case 1. If all the cloud providers play the game cooperatively and infinitely, the long-term profit of cloud provider i is:

$$\sum_{t=0}^{\infty}[P_i^c \cdot \lambda_i^t] = P_i^c + \frac{\lambda_i \cdot \left(1 - \lambda_i^{T_i}\right)}{1 - \lambda_i} \cdot P_i^c + \frac{\lambda_i^{T_i+1}}{1 - \lambda_i} \cdot P_i^c. \quad (49)$$

However, if cloud provider i deviates from the cooperation in the current stage, then all the cloud providers will operate at Nash equilibrium for T_i game plays. Then, we have the other two cases:

Case 2. First, cloud provider i becomes cooperative after T_i

rounds. Therefore, long-term profit of cloud provider i is:

$$P_i^d + \sum_{t=1}^{T_i}[P_i^p \cdot \lambda_i^t] + \sum_{t=T_i+1}^{\infty}[P_i^c \cdot \lambda_i^t]$$

$$= P_i^d + \frac{\lambda_i \cdot \left(1 - \lambda_i^{T_i}\right)}{1 - \lambda_i} \cdot P_i^p + \frac{\lambda_i^{T_i+1}}{1 - \lambda_i} \cdot P_i^c. \quad (50)$$

Case 3. Second, cloud provider i may resist the punishment and continue to be non-cooperative. Then, long-term profit of cloud provider i is:

$$\sum_{t=0}^{\infty}[P_i(t) \cdot \lambda_i^t] = P_i^d + \sum_{t=1}^{\infty}[P_i^p \cdot \lambda_i^t] = P_i^d + \frac{\lambda_i}{1 - \lambda_i} \cdot P_i^p. \quad (51)$$

To promote cooperation, the utility obtained in case 1 should be the greatest among all three cases. The utility obtained in case 2 is larger than that obtained in case 3 in that $P_i^c \geq P_i^p$ by definition. Therefore, the sufficient and necessary condition is given by:

$$P_i^c + \frac{\lambda_i \cdot \left(1 - \lambda_i^{T_i}\right)}{1 - \lambda_i} \cdot P_i^c \geq P_i^d + \frac{\lambda_i \cdot \left(1 - \lambda_i^{T_i}\right)}{1 - \lambda_i} \cdot P_i^p, \quad (52)$$

which gives the conclusion in the theorem. $\square$

Therefore, in the special case with $T_i = \infty$ (i.e., the trigger strategy), the sufficient and necessary condition of cooperation among all the cloud providers is:

$$\lambda_i \geq \frac{P_i^d - P_i^c}{P_i^d - P_i^p}. \quad (53)$$

In other words, if and only if the above lower bound of λ_i is satisfied, the threatening implemented by the Striker strategy is effective.

DEFINITION 3. We define a threat to be *credible* if under the triggering condition, the threat-claimer would obtain no less utility than not carrying out the threat.

THEOREM 6. On the condition that Theorem 5 holds, the Striker strategy is credible and will always be adopted by cloud providers.

Due to page limit, we omit the proof which is similar to Theorem 5.

IV. PERFORMANCE EVALUATION

In this section, we present our evaluation results on the proposed game model and algorithms for competitive resource pricing.

A. Setup

We augment our evaluation with realistic electricity prices in the spot market [9]. We consider a cloud market with two cloud providers and one tenant user so as to obtain clear insights into competitive cloud resource pricing. In the evaluation, we first use the average hourly prices of the electricity markets in California and New York on February 29, 2012 (i.e., $Pr_1 = 26.4146\$/MWh$ and $Pr_2 = 29.7846\$/MWh$) so that we can observe the impact of price competition on
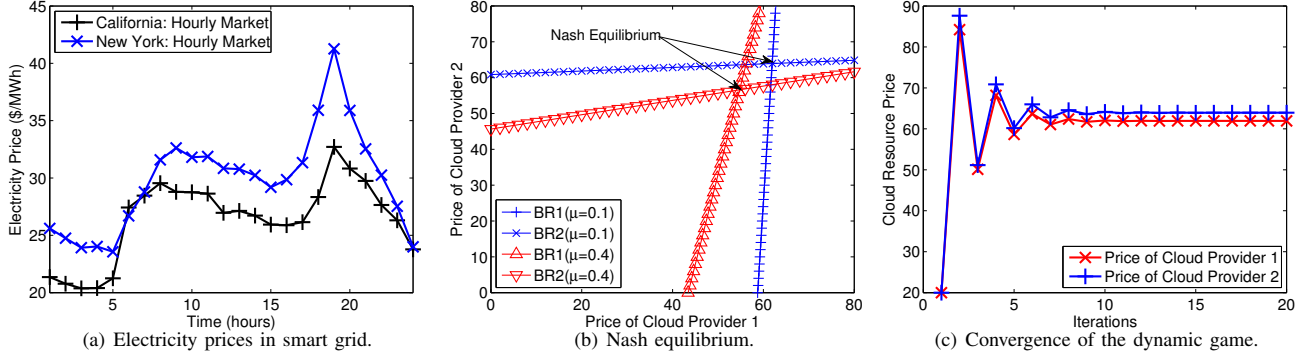
Fig. 1: Spot electricity markets, Nash equilibrium, and algorithm convergence. (a) The price variation in the hourly market in California and New York on February 29, 2012. (b) Illustration of Nash equilibrium calculated by best response functions. (c) The convergence of the cloud providers' prices to the Nash equilibrium in the dynamic game.

equilibrium instance prices. Then, we use the hourly electricity price time series (as shown in Fig. 1(a)) to examine the impact of electricity price dynamics on the spot resource prices in the cloud market. For the cloud profit model, we use $\alpha = 0.1$, $r = 5$, $E_i = 10$. For tenant surplus, we use the maximum incurred delay $L = 30000$, $\gamma_1 = 5000$, $\gamma_2 = 3000$, $\mu = 0.1$, and $K = 10$ by default.

B. Nash Equilibrium and Convergence of the Dynamic Game

Fig. 1(b) illustrates the calculation of the Nash equilibrium using best response functions. It is observed that higher resource substitutability leads to lower equilibrium instance prices due to more fierce competition. Fig. 1(c) shows the price dynamics under the dynamic game for $\delta_i = 0.8$. It is observed that the prices for both cloud providers converge to the equilibrium levels despite the initial stage of price fluctuations. Actually, the fluctuations are larger for greater δ_i, and the dynamic game even never converges for sufficient large δ_i.

C. Price Competition among Cloud Providers

Fig. 2(a) shows the impact of network delay on equilibrium prices. We find that the price of cloud provider 1 is lower due to the larger network delay, when $\gamma_2 < \gamma_1$, and vice versa for $\gamma_2 > \gamma_1$. This shows the competitive relationship between the two cloud providers for pricing scheme design. That is, lower service quality of one cloud provider may decrease its own prices but increase the resource prices of the other one at Nash equilibrium. This reflects the adaptation of the tenant's demands to the service quality of different cloud resource providers. Indeed, lower service quality will reduce the tenant demand, and cloud provider 2 will lower down the resource price so as to increase its own profit at Nash equilibrium. In Fig. 2(b), we use the same network delay for both cloud providers (i.e., $\gamma_i = 5000$) so as to observe the impact of cloud provider numbers on cloud resource prices. It is noticed that resource prices decrease with the increase in the number of cloud providers. This illustrates the formulation of an oligopoly market for cloud resource sharing. Fig. 2(c)

shows the dynamics of cloud resource prices in a smart grid environment. By comparing it with Fig. 1(a), we observe the positive correlation between cloud resource prices and electricity cost. This indicates the critical role of energy cost in cloud resource pricing.

D. Tradeoff between Cloud Profit and tenant Surplus

From Fig. 3(a) and Fig. 3(b), it is observed that there is a negative relationship between cloud profit and substitutability, but a positive relationship between tenant surplus and substitutability. That is, the cloud profit decreases, while the tenant surplus increases with the increase of substitutability. We observe similar relationships among cloud profit, tenant surplus, and electricity prices. This implies that there is a fundamental tradeoff between cloud profit and tenant surplus. Moreover, compared with Nash equilibrium, the cloud profit is higher and the tenant surplus is lower under the optimal pricing for cloud profit maximization. Fig. 3(c) shows the lower bounds of the discount factors for both cloud providers so that the punishment in trigger strategy is effective. It is observed that the lower bounds are higher for higher substitutability. At the same time, the lower bounds can be readily satisfied because in practice the discount factors are much larger.

V. RELATED WORK

Cloud resource pricing recently draws great attention to the research community [10]. In a realistic cloud market, both spot prices and usage-based prices exist for the convenience of tenant users. Wang *et al.* [7] optimize cloud revenue by dynamically partitioning the cloud capacity between the two pricing tiers. Niu *et al.* [11], [12] argue the necessity of cloud tenants to multiplex cloud resources among correlated tenant traffic. Most relevant, Xu *et al.* [13], [14] assume a demand distribution and achieve cloud revenue maximization by proposing a centralized optimization framework. Cao *et al.* [15] explore cloud profit maximization by building queueing models for optimal multiserver configuration. Despite the above recent studies on cloud resource pricing, they ignore

(a) The impact of substitutability.

(b) The impact of cloud provider number.

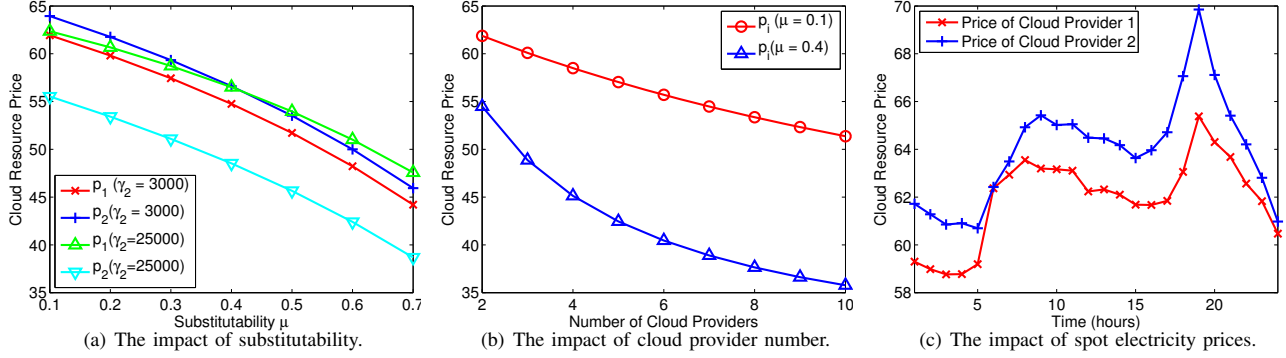(c) The impact of spot electricity prices.

Fig. 2: Cloud resource prices under cloud provider competition. (a) Price variation under different cloud resource substitutability and network delay. (b) Price variation under multiple cloud providers. (c) Price dynamics using electricity spot prices in hourly electricity spot markets on February 29, 2012.

(a) Cloud profit in smart grid.

(b) tenant surplus in smart grid.
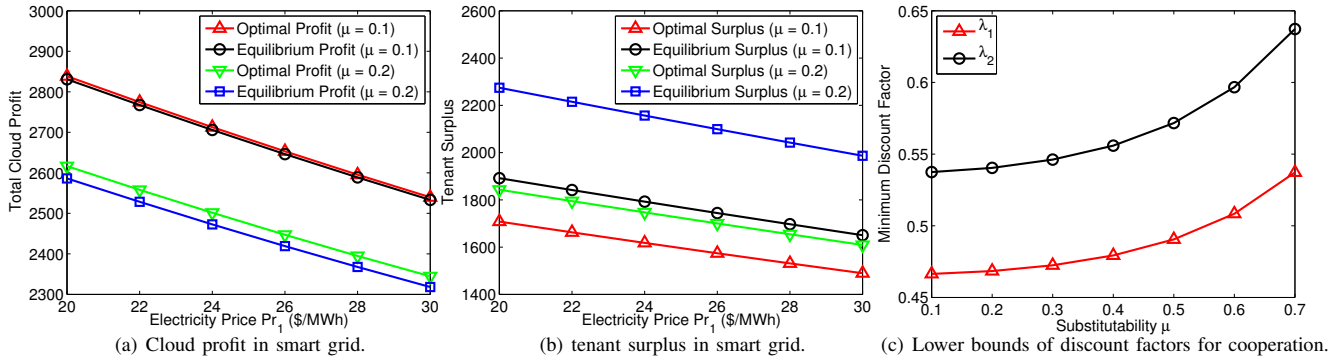
(c) Lower bounds of discount factors for cooperation.

Fig. 3: Cloud profit, tenant surplus, and the maintenance of total profit maximization. (a) Cloud profit comparison under different substitutability and electricity prices. (a) tenant surplus comparison under different substitutability and electricity prices. (c) Minimum discount factors to maintain cooperation for total profit maximization under trigger strategy.

the problem of the competitive nature among cloud providers and its critical impact on cloud profit and tenant demands.

VI. CONCLUDING REMARKS

In this paper, we explore the problem of competitive cloud resource pricing for cloud providers. Indeed, cloud providers compete for profit maximization in an oligopoly market. We realistically model the pricing scheme of the cloud provider by considering the impact of network delay, renewables, and electricity prices. We propose a noncooperative game to model such competitive resource pricing. We then conduct equilibrium analysis under the assumption of perfect information. To relax the assumption of perfect information, we propose the adoption of dynamic game to reach Nash equilibrium in a distributed manner by using local information only. The results revealed insightful observations for practical pricing scheme design. In the future, we would like to extend our model to the more general case of multiple tenant users.

REFERENCES

[1] Amazon EC2, 2013, http://aws.amazon.com/ec2/.
[2] Google App Engine, 2013, https://appengine.google.com/start.
[3] Windows Azure, 2013, http://www.windowsazure.com/.
[4] Dropbox, 2013, http://www.dropbox.com/.
[5] B. C. Tak, B. Urgaonkar, and A. Sivasubramaniam, "Cloudy with a Chance of Cost Savings," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 6, pp. 1223–1233, June 2013.
[6] M. Mihailescu and Y. M. Teo, "Dynamic Resource Pricing on Federated Clouds," in *Proc. of 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing*, 2010.
[7] W. Wang, B. Li, and B. Liang, "Towards Optimal Capacity Segmentation with Hybrid Cloud Pricing," in *Proc. of ICDCS*, June 2012.
[8] M. J. Osborne, *An Introduction to Game Theory*. Oxford University Press, 2004.
[9] United States Energy Information Administration, 2013, http://www.eia.gov/.
[10] H. Wang, Q. Jing, R. Chen, B. He, Z. Qian, and L. Zhou, "Distributed Systems Meet Economics: Pricing in the Cloud," in *Proc. of USENIX HotCloud*, June 2010.
[11] D. Niu, C. Feng, and B. Li, "A Theory of Cloud Bandwidth Pricing for Video-on-Demand Providers," in *Proc. of INFOCOM*, March 2012.
[12] ——, "Pricing Cloud Bandwidth Reservations under Demand Uncertainty," in *Proc. of SIGMETRICS*, June 2012.
[13] H. Xu and B. Li, "Maximizing Revenue with Dynamic Cloud Pricing: The Infinite Horizon Case," in *Proc. of IEEE ICC*, June 2012.
[14] ——, "A Study of Pricing for Cloud Resources," *ACM SIGMETRICS Performance Evaluation Review, Special Issue on Cloud Computing*, March 2013.
[15] J. Cao, K. Hwang, K. Li, and A. Y. Zomaya, "Optimal Multiserver Configuration for Profit Maximization in Cloud Computing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 6, pp. 1087–1096, June 2013.