Constrained Pricing for Cloud Resource Allocation

Makhlouf Hadji, Wajdi Louati, Djamal Zeghlache

Institut Telecom, Telecom SudParis, UMR CNRS 5157
9, Rue Charles Fourier, 91011, Evry, France
Email: {makhlouf.hadji, wajdi.louati, djamal.zeghlache}@it-sudparis.eu

Abstract—Constrained pricing in a cloud computing environment is addressed using game theory. The objective of the model and the game is to enable cloud providers to maximize their revenue while satisfying users by maximizing their utilities. The users' net utility is modeled as a function of resource demand with a corresponding price. The game consists in the cloud provider suggesting differentiated prices according to demand and users updating their requests in view of the proposed price. The objective is to determine the optimal suggested prices by the cloud-provider and the optimal user demands. A theoretical model based on Stackelberg game is proposed and a Stackelberg/Nash equilibrium solution is found. Performance results show that rejecting a certain number of users is not necessarily the best one and the solution is to deploy the right amount of resources to maintain good reputation while satisfying users' requirements.

Keywords—Cloud Computing; Pricing; Game Theory; Nash Equilibrium;

I. INTRODUCTION

Cloud computing is a large scale distributed solution for providing compute, storage and communications resources and services on demand through data centers. Cloud computing has advantages over traditional computing such as avoiding capital investments and operational expense for users, pay-as-you-go metered services and resource multiplexing and sharing for providers. These services are deployed via third-party companies (service providers) and accessed by users anytime [1]. The advantages of cloud computing have attracted a number of actors such as Amazon, Google and Microsoft to enter and invest this new market. Current cloud providers propose different pricing models. Amazon AWS charges customers by the number of virtual machines allocated over a time duration (for example, Amazon EC2 provides a virtual machine (VM) at a fraction of a dollar per hour) [2]. Details on pricing problems can be found in [3], [4], [5] with cloud providers offering services at various prices depending on users’ demands. In order to offset their costs and make benefits, cloud providers need to understand users’ behavior. Users are interested in minimizing their cost or maximizing their utility. Providers have to identify the best strategies according to users’ reactions and decisions.

The objective of this work is to model the behavior of users and providers using a game theory approach to identify the most appropriate strategies for the players. The model looks for a Nash equilibrium point (NEP) between the users’ demands and the providers’ offers [6]. The aim is also to determine the evolution of prices, at the Nash equilibrium point, in terms of the number of users involved and the amount of available resources. Results show that the cloud provider’s strategy to reject a certain number of users is not necessarily the best one and the solution is to deploy the right amount of resources indicated by the NEP to scale and maintain good reputation and revenue.

This paper focuses on modeling the interactions between cloud providers and the users as a non-cooperative and constrained game. Cloud providers present limited (finite size) resources to their users with a price per resource unit proposal. The price may differ from one geographic location to another. Users decide how to adjust their demands according to the proposed pricing. Users are considered as selfish and only interested in maximizing their own utilities.

Similar studies in previous work can be found in for example [5] presenting a game theoretic pricing mechanism for statistically guaranteed service in packet-switched networks. This mechanism is inspired by MacKie-Mason and Varian [7] who proposed a ’smart market’ mechanism. The proposed mechanism in [5] provides congestion control, differentiated quality of service and efficient resource allocation. They show that their suggested mechanisms offer better quality and lower price to users and that service providers can derive their new service and revenue models.

Our work is inspired by the pricing problem without constraints presented in [3] that presents an optimal policy to maximize the cloud provider’s revenue and an optimal vector of users’ demands to maximize their utilities. The authors presented the pricing problem for non-cooperative users and demonstrate implicitly the existence of a Nash/Stackelberg equilibrium but with no limited resources constraints.

In our work, we consider the problem of constrained pricing in a Stackelberg/Nash manner taking into account two key conditions or hypotheses. The first one is to consider finite resources (limited amount of resources) and the second
one consists of making differentiated pricing proposals to explore the revenue space or opportunities and select winning strategies. This differentiation can also be justified in view of some cloud service providers such as Amazon that offers IaaS services with different prices per geographic region [8].

The remainder of this paper is organized as follows: Section II introduces a pricing model based on Nash/Stackelberg game formulation. The objective is to maximize revenue for cloud providers while maximizing the users’ utilities under some constraints. Section III discusses the case of symmetric users where each has the same unit measurement of the utility factor. Section IV is dedicated to the evaluation of the proposed pricing model. Conclusions follow in Section V.

II. THE SYSTEM MODEL

This section describes and formulates the multi-cloud pricing problem as a mathematical model. Let \( j \) denote a cloud provider offering services to a finite number of users \( I = \{1, 2, \ldots, N\} \) that share the resources. The cloud service (e.g., virtual machines) offered by the provider is composed of a certain amount of memory, CPU, storage, I/O performance, bandwidth. We assume that the cloud provider \( j \) has limited resources denoted by \( B_j \).

This work focuses on the case where \( B_j \) is an aggregated criterion of different cloud resources. Distinguishing the nature of resources (compute, storage, networks) that are characterized by different constraints, requires the optimisation of a multi-criteria function over multiple constraints but is out of the scope of this paper.

The cloud provider \( j \) presents to the users a menu with prices per service and per geographic location. Let \( \mu_{i,l}^j \) denote the unit price of the service charged by the cloud provider \( j \) to the user \( i \) situated in the location \( l \in L \) where \( L \) represents the set of all considered locations. \( b_{i,l}^j \) is used to denote the amount of aggregated resources requested by the user \( i \) into a single resource demand. This amount can be expressed as a combination of values for storage, CPU, bandwidth and other resource characteristics. We note that the same hypothesis is considered in [9].

The objective of the provider \( j \) is to maximize its revenue function given by: 

\[
R_j(b, \mu) = \sum_{l \in L} \sum_{i = 1}^{N(l)} \mu_{i,l}^j b_{i,l}^j 
\]

where \( N(l) \) represents the number of users situated in the geographic region \( l \) and \( \sum_{l \in L} N(l) = N \).

On the other hand, the objective of the user \( i \) is to maximize the following net utility function given by:

\[
U_i(b, \mu) = \alpha_i \log \left( 1 + \frac{Mb_i^j}{\sum_{k \neq i} b_k^j} \right) - \mu_{i,l}^j b_{i,l}^j. (1)
\]

where \( \alpha_i \log \left( 1 + \frac{Mb_i^j}{\sum_{k \neq i} b_k^j} \right) \) is the utility of the demand \( b_i^l \) to user \( i \) and \( \mu_{i,l}^j b_{i,l}^j \) is the total price to pay. \( M \) is the job decoupling factor which is higher for cloud providers who can separate sufficiently the demands of different users \((M > 1)\) and \( \alpha_i > 0 \) is a constant which converts utility to currency. The constraint given by 

\[
\sum_{i = 1}^{N} b_i^j \leq B_j (2)
\]

allows to exceed the maximum available resource \((B_j)\) in cloud provider \( j \). Unless otherwise stated, we will use \( \tilde{b} \) to denote \( \sum_{i = 1}^{N} b_i^j \). We denote by \( \tilde{\alpha} = \sum_{i = 1}^{N} \alpha_i. \)

For the user the goal is to optimize function (1) subject to constraint (2). Given this behavior, the cloud provider \( j \) aims at maximizing its revenue by suggesting optimal prices. The cloud provider’s objective is to solve (3).

\[
\max_{\mu} R_j(b^*, \mu) = \sum_{l \in L} \sum_{i = 1}^{N(l)} \mu_{i,l}^j b_{i,l}^j (3)
\]

The central or main goal of this paper is to find an agreement between users and the cloud provider. This allows users to partake in the interaction by proposing their respective demands and this in turn enables the cloud provider to better manage its resources. This is a non-cooperative game, where each user is interested in optimizing her/his utility, that can be modeled as a Stackelberg game [3],[6], where the cloud providers set prices and users update their demands to maximize their utilities.

By solving the Stackelberg/Nash game, we have to compute Nash Equilibrium Point (NEP) where neither the cloud provider nor any user has incentive to deviate unilaterally from the NEP; as stated in the following definition.

**Definition 2.1:** (Nash Equilibrium Point) [6] Let \( \mu^* \) be a solution for the problem (3) of a cloud provider and \( (b^*) \) be a solution for the user’s problem (1) subject to the constraint (2). The couple \((\mu^*, (b^*))\) is a NEP for the game if for any couple \((\mu, b)\), we have 

\[
U_i((b_i)^*, (b_i) - \mu^*) \geq U_i((b_i), (b_i)^*), \forall i = 1, \ldots, N
\]

and

\[
R_j((b_j)^*, (\mu) - \mu^*) \geq R_j((b_j), (\mu)^*), \mu
\]

where \( b_{i,j}^* = (b_1^*, b_2^*, \ldots, b_{i-1}^*, b_{i+1}^*, \ldots, b_N^*) \).

The utility function chosen for a user \( i \) is 

\[
\alpha_i \log \left( 1 + \frac{Mb_i^j}{\sum_{k \neq i} b_k^j} \right), \text{ which is closely related to the utility function } \alpha_i \log(b_i^j) \text{ cited in [4] and [10]. If we use a utility function } \alpha_i \log \left( \frac{Mb_i^j}{\sum_{k \neq i} b_k^j} \right), \text{ then a user is forced to present a nonzero demand since its utility becomes } -\infty \text{ if } b_i^j = 0. \]

Then, one can see that our utility function allows users to decide whether to join the cloud provider or not. This utility function can be interpreted as the user’s quality of service in the presence of other users.
Proposition 2.1: There exists a Nash/Stackelberg equilibrium point \((b^*, \mu^*)\) of the game.

Proof: For sake of clarity, we distinguish two different cases: uniform prices and differentiated prices, detailed in II-A and II-B respectively.

A. Uniform prices case

In this case, we suppose that the cloud provider suggests uniform prices to users, so that \(\mu_{i,l}^* = \mu^*\). In other words, all users are situated in the same geographic location \(|\mathcal{L}| = 1\).

The original game is now equivalent to the case where one user 1 has the following objective function \(U_i\):

\[
U_i(b, \mu^*) = \alpha_i \log \left(1 + \frac{Mb_i^*}{\sum_{k \neq i} b_k^*}\right) - \mu^* b_i^* \tag{4}
\]

To optimize the function (4) over the vector of demands \(b\) and subject to the constraint given in (2), we use the Lagrangian method.

\[
L = \alpha_i \log \left(1 + \frac{Mb_i^*}{\sum_{k \neq i} b_k^*}\right) - \mu^* b_i^* - \lambda \left(\sum_i b_i^* - M\right) - \lambda
\]

\[
\lambda \text{ is the Lagrangian multiplier. Let us now calculate the derivative form of } L \text{ with respect to } b_i^*, \text{ and let it equal to zero.}
\]

\[
\frac{\partial L}{\partial b_i^*} = \frac{M \alpha_i}{b_i^* + (M - 1) b_i^*} - \mu^* - \lambda = 0
\]

This leads to

\[
\left(b_i^* + (M - 1) b_i^*\right)(\mu^* + \lambda) = M \alpha_i \tag{6}
\]

and then we have

\[
(M - 1) b_i^* = \frac{M \alpha_i}{\lambda + \mu^*} - \bar{b}
\]

which gives the following equality (note that \(M > 1\))

\[
b_i^* = \frac{M}{M - 1} \frac{\alpha_i}{\lambda + \mu^*} - \frac{1}{M - 1} \bar{b} \tag{7}
\]

By summing (7) over all users \(i = 1, \ldots, N\), we have

\[
\sum_{i=1}^{N} b_i^* = \frac{M}{M - 1} \sum_{i=1}^{N} \frac{\alpha_i}{\lambda + \mu^*} - \frac{1}{M - 1} \bar{b} \tag{8}
\]

This is equivalent to write

\[
\bar{b} = \frac{M}{M - 1} \frac{\bar{\alpha}}{\lambda + \mu^*} - \frac{N}{M - 1} \bar{b} \tag{9}
\]

By doing some arrangements, we obtain

\[
\bar{b} \left(1 + \frac{N}{M - 1}\right) = \frac{M}{M - 1} \frac{\bar{\alpha}}{\lambda + \mu^*}
\]

This helps us to conclude that

\[
\bar{b} = \frac{M}{M + N - 1} \frac{\bar{\alpha}}{\lambda + \mu^*} \tag{10}
\]

and by replacing it in (7), we obtain

\[
b_i^* = \frac{M}{M - 1} \left(\alpha_i - \frac{\bar{\alpha}}{M + N - 1}\right) \frac{1}{\lambda + \mu^*} \tag{11}
\]

In the other hand, we have

\[
\frac{\partial L}{\partial \lambda} = 0 \iff \bar{b} = B^j \tag{12}
\]

and from the equality (10):

\[
\lambda + \mu^* = B^j \frac{M \bar{\alpha}}{M + N - 1}\bar{b}
\]

For the values of \(\lambda \geq 0\), we conclude

\[
\lambda = \frac{M \bar{\alpha}}{M + N - 1} B^j - \mu^* \tag{13}
\]

and by replacing it in (7), we obtain

\[
b_i^* = \frac{M}{M - 1} \left(\alpha_i - \frac{\bar{\alpha}}{M + N - 1}\right) \frac{B^j(M + N - 1)}{M \bar{\alpha}} \tag{14}
\]

By simplifying (14), we easily find the following equality:

\[
b_i^* = \frac{B^j}{M - 1} \left(\alpha_i - \frac{\bar{\alpha}}{M + N - 1}\right) \frac{(M + N - 1)}{\bar{\alpha}} \tag{15}
\]

Finally, by developing (15) we obtain the following result:

\[
(b_i^*)^* = \frac{B^j}{M - 1} \left(\alpha_i(M + N - 1)\bar{\alpha} - 1\right) \tag{16}
\]

Maximization of \(R_j\) over \(\mu^* \geq 0\) dictates that the revenue-maximization price is given by

\[
(\mu^*)^* = \frac{M}{M + N - 1} \frac{\bar{\alpha}}{B^j} \tag{17}
\]

It can be shown by substituting (17) in the revenue formula (3) that the revenue of a cloud provider \(j\) at every NEP is fixed and given by

\[
R_j^* = \frac{M \bar{\alpha}}{M + N - 1} = \frac{MNa_{av}}{M + N - 1} \tag{18}
\]

where \(a_{av} = \frac{1}{N} \sum_{i=1}^{N} a_i = \frac{1}{N} \bar{\alpha}\).

The second part of the proof consists to consider users which are charged different prices.

B. Differentiated prices case

We consider different possible prices suggested by a cloud provider to different users. We assume thus that \(|\mathcal{L}| > 1\) and the proposed prices are different from a geographic location to another. The net utility function of a user \(i\) is then given by

\[
U_i(b, \mu) = \alpha_i \log \left(1 + \frac{Mb_i}{\sum_{k \neq i} b_k}\right) - \mu_i b_i \tag{19}
\]
Subject to:

\[ \sum_{i=1}^{N} b_i^j \leq B^j \]  

(20)

To optimize the new function (19) over the vector of demands \( b \) and subject to the constraint given in (20), we use the Lagrangian method.

\[ L = \alpha_i \log \left(1 + \frac{Mb_i^j}{\sum_{k \neq i} b_k^j}ight) - \mu_i b_i^j - \lambda \left(\sum_{i=1}^{N} b_i^j - B^j\right) \]

\( \lambda \) is the Lagrangian multiplier. Let us now calculate the derivative form of \( L \) with respect to \( b_i^j \), and let it equal to zero.

\[ \frac{\partial L}{\partial b_i^j} = \frac{M \alpha_i}{\sum_{k \neq i} b_k^j + M b_i^j} - \mu_i^j - \lambda = 0 \]

This leads to

\[ M \alpha_i = \left(\sum_{k \neq i} b_k^j + M b_i^j\right) \left(\mu_i^j + \lambda\right) \]  

(21)

From (21) we deduce

\[ b_i^j = \frac{\alpha_i}{\lambda + \mu_i^j} - \frac{\sum_{k \neq i} b_k^j}{M} \]  

(22)

Multiplying (22) by \( M \), we can easily deduce that

\[ \sum_{k \neq i} b_k^j = \frac{M \alpha_i}{\lambda + \mu_i^j} - M b_i^j \]  

(23)

By adding \( b_i^j \) in the two sides of (23), we obtain

\[ \sum_{k=1}^{N} b_k^j = \bar{b} = \frac{M \alpha_i}{\lambda + \mu_i^j} - (M - 1) b_i^j \]  

(24)

By summing the two sides of (24) over all users \( i = 1, \ldots, N \), we obtain

\[ N \bar{b} = M \sum_{i=1}^{N} \frac{\alpha_i}{\lambda + \mu_i^j} - (M - 1) \bar{b} \]  

(25)

In other terms, we find by dividing (25) by \( N \) \( (N \geq 1) \)

\[ \bar{b} = \frac{M}{M + N - 1} \sum_{i=1}^{N} \frac{\alpha_i}{\lambda + \mu_i^j} \]  

(26)

One can remark the equivalence between the two formulas (26) and (24), then we can write

\[ \frac{M}{M + N - 1} \sum_{k=1}^{N} \frac{\alpha_k}{\lambda + \mu_k^j} = M \alpha_i \frac{\alpha_i}{\lambda + \mu_i^j} - (M - 1)b_i^j \]  

(27)

Then it is not difficult to conclude that

\[ b_i^j(\mu_i^j) = \frac{M}{M - 1} \left[ \alpha_i \frac{1}{\lambda + \mu_i^j} - \frac{1}{M + N - 1} \sum_{k=1}^{N} \frac{\alpha_k}{\lambda + \mu_k^j} \right] \]  

(28)

For \( \lambda \geq 0 \) and from the equality (24) we can write

\[ (M - 1)b_i^j = \frac{M \alpha_i}{\lambda + \mu_i^j} - \bar{b} \]  

(29)

and as \( \bar{b} = B^j \) (see (12)), we replace it in (29) and we obtain

\[ (M - 1)b_i^j = \frac{M \alpha_i}{\lambda + \mu_i^j} - B^j \]  

(30)

By summing (30) over all users \( i = 1, \ldots, N \), we obtain

\[ (M - 1) \sum_{i=1}^{N} b_i^j = M \sum_{i=1}^{N} \frac{\alpha_i}{\lambda + \mu_i^j} - N B^j \]  

(31)

which is equivalent to write

\[ (M - 1) \bar{b} = M \sum_{i=1}^{N} \frac{\alpha_i}{\lambda + \mu_i^j} - N B^j \]  

(32)

As mentioned above \( \bar{b} = B^j \), then by some computation and by distinguishing the \( i \text{th} \) term, we find

\[ \frac{M + N - 1}{M} B^j = \sum_{i=1}^{N} \frac{\alpha_i}{\lambda + \mu_i^j} = \frac{\alpha_i}{\lambda + \mu_i^j} + \sum_{k \neq i} \frac{\alpha_k}{\lambda + \mu_k^j} \]  

(33)

Using (33), we can easily deduce

\[ \frac{\alpha_i}{\lambda + \mu_i^j} = \frac{M + N - 1}{M} B^j - \sum_{k \neq i} \frac{\alpha_k}{\lambda + \mu_k^j} \]  

(34)

In the other hand, we can also write

\[ \sum_{k \neq i} \frac{\alpha_k}{\lambda + \mu_k^j} = \frac{\sum_{k \neq i} \alpha_k \left( \prod_{I \neq i, I \neq k} (\lambda + \mu_i^j) \right)}{\left( \prod_{m \neq i, m \neq k} (\lambda + \mu_m^j) \right)} \]  

Then, by replacing this last equality in (34), we find

\[ \frac{\alpha_i}{\lambda + \mu_i^j} = \frac{M + N - 1}{M} B^j - \frac{\sum_{k \neq i} \alpha_k \left( \prod_{I \neq i, I \neq k} (\lambda + \mu_i^j) \right)}{\left( \prod_{m \neq i, m \neq k} (\lambda + \mu_m^j) \right)} \]  

(35)

For sake of simplicity and clarity, we put \( \gamma = \frac{M + N - 1}{M} B^j \). From (35) we deduce

\[ \left( \gamma - \frac{\alpha_i}{\lambda + \mu_i^j} \right) \left( \prod_{m \neq i, m \neq k} (\lambda + \mu_m^j) \right) - \sum_{k \neq i} \alpha_k \left( \prod_{I \neq i, I \neq k} (\lambda + \mu_i^j) \right) = 0 \]  

(36)
(λ + μ_i^t) \left( γ - \frac{α_i}{λ + μ_i^t} \right) \left( \prod_{m \neq i, m \neq t} (λ + μ_m^t) - \sum_{k \neq i} α_k \left( \prod_{l \neq i, l \neq k} (λ + μ_l^t) \right) \right) = 0 \quad (37)

The equality (36) can be rewritten as the equality (37). By summing (37) over all \( t = 1, \ldots, N \), \( t \neq i \), we obtain the equality (38).

After simplifications, we find the equality given by (39). From the equality (39), the term \( \prod_{m \neq i, m \neq t} (λ + μ_m^t) \) cannot be equal to zero, then we conclude that

\[
(λ + μ_i^t) \left( γ - \frac{α_i}{λ + μ_i^t} \right) = (N - 1)α_i \quad (40)
\]

which we can write as

\[
\left( γ - \frac{α_i}{λ + μ_i^t} \right) = (N - 1) \frac{α_i}{λ + μ_i^t} \quad (41)
\]

and use

\[
\frac{α_i}{λ + μ_i^t} = \frac{B_j}{M} + \frac{M - 1}{M} b_i^t
\]

The formula (41) can be rewritten as

\[
γ - \frac{B_j}{M} - \frac{M - 1}{M} b_i^t = (N - 1) \left( \frac{B_j}{M} + \frac{M - 1}{M} b_i^t \right)
\]

with \( t \neq i \).

After some computation and by replacing \( γ \) by its value, the last formula is simplified to

\[
(M - 1)B_j - (M - 1)b_i^t = (N - 1)(M - 1)b_i^t \quad (42)
\]

By summing (42) over all \( t = 1, \ldots, N \), we find

\[
(M - 1)B_j - (M - 1)b_i^t = \frac{(N - 1)(M - 1)}{N} B_j
\]

It is not difficult to deduce

\[
b_i^t* = \frac{B_j}{N} \quad (43)
\]

Note that from (30) and by putting \( λ = 0 \), we have also

\[
(M - 1)b_i^t = \frac{Mα_i}{μ_i^t} - B_j \quad (44)
\]

By replacing the obtained value of \( b_i^t* \) into the equality (44), we obtain the following equality:

\[
μ_i^t* = \frac{Mα_i}{(M - 1)b_i^t* + B_j} = \frac{MNα_i}{(M + N - 1)B_j} \quad (45)
\]

Finally, we can compute the revenue of a cloud provider by replacing the obtained results of (43) and (45) into the given formula (3). We obtain at the equilibrium:

\[
R_j^* = \frac{Mα}{M + N - 1} = \frac{MNα_{av}}{M + N - 1} \quad (46)
\]

For large values of \( N \), we have

\[
\lim_{N \to \infty} R_j^* = Mα_{av} \quad (47)
\]

Remark 2.1: It is not difficult to verify the uniqueness of the NEP of the game in the two above discussed cases. Note that for a fixed price \( μ \), we have

\[
U_i(b, μ^t) = α_i \log \left( 1 + \frac{M b_i^t}{\sum_{k \neq i} b_k^t} \right) - μ_i b_i^t
\]

the first derivative form of \( U_i \) with respect to the demand \( b_i^t \) is given by

\[
\frac{∂U_i}{∂b_i^t} = \frac{α_i M}{b + (M - 1)b_i^t} - μ_i
\]

then the second derivative is given by

\[
\frac{∂^2U_i}{∂(b_i^t)^2} = \frac{-α_i M^2}{(b + (M - 1)b_i^t)^2} \quad (48)
\]

One can notice that \( U_i \) is a concave function of \( b_i^t \) and the second derivative given in (48) is negative. This leads to conclude the uniqueness of the NEP.

III. Symmetric Users Case

In this section, we suppose that all users belong to the same region (\( |L| = 1 \)) and are symmetric (i.e. all users have the same factor \( α \)). Let \( α_i = α_j = α, \forall i, j = 1, \ldots, N \). The optimal price in this case is given by

\[
(μ^t)^* = \frac{M}{M + N - 1} \frac{α}{B_j} = \frac{M}{M + N - 1} \frac{Nα}{B_j} \quad (49)
\]

The optimal revenue for the cloud provider \( j \) can be written in the following form using (49):

\[
R_j(μ^*, b^*) = \frac{MαN}{M + N - 1} \quad (50)
\]

It is not difficult to observe that the equality (50) is increasing and concave in the number of users \( N \). Moreover, we have

\[
\lim_{N \to \infty} R_j(μ^*, b^*) = Mα
\]

In other words, this limit gives an upper bound on the revenue for a cloud provider \( j \). One can remark that the more users the cloud provider accommodates, the better revenue the cloud provider gains up to a multiplicative value of his reputation \( M \). In the other hand, if we suppose that \( R_0 \) is the threshold to guarantee by the cloud provider, then

\[
R_j \geq R_0 \iff \frac{MαN}{M + N - 1} \geq R_0
\]

363
\[
\sum_{i \neq j} (\lambda + \mu_i^j) \left( \gamma - \frac{\alpha_i}{\lambda + \mu_i^j} \right) \left( \prod_{m \neq i, m \neq j} (\lambda + \mu_m^i) \right) - (N - 1) \sum_{k \neq i} \alpha_k \left( \prod_{m \neq i, m \neq k} (\lambda + \mu_m^i) \right) = 0 \quad (38)
\]

\[
\sum_{i \neq j} \left[ (\lambda + \mu_i^j) \left( \gamma - \frac{\alpha_i}{\lambda + \mu_i^j} \right) - (N - 1)\alpha_j \right] \left( \prod_{m \neq i, m \neq l} (\lambda + \mu_m^i) \right) = 0 \quad (39)
\]

This form gives a lower bound on the number of users that can be accommodated by the cloud provider \(j\) to guarantee a certain threshold \(R_0\). After some computation, we obtain

\[
N \geq \frac{(M - 1) R_0}{\alpha M - R_0}, \quad (R_0 < M \alpha) \quad (51)
\]

IV. NUMERICAL RESULTS

This section evaluates the proposed cloud pricing model to provide better understanding of both cloud providers and users behavior to achieve highest payoffs.

It is important to note that the following experiments take into account the fact that the cloud provider’s reputation \(M\) is fixed in a time slot \(t\) and depends only on the number of rejected users at time slot \(t - 1\). \(M\) can nevertheless be modeled in a dynamic manner to characterize more precisely the provider’s reputation.

Figure 1 depicts the evolution of prices, at the Nash equilibrium point, charged by a cloud provider (in the uniform case) when the number of users increases and the finite amount of available resources grows. We fixed both \(N\) and \(B_j\) between 1 and 100. This section does not discuss the less business relevant case where a very small number of users is involved and a small amount of resources is available. Focussing on more realistic cases, with enough resources and users, the strategy that stands out from the results is that providers should keep the price for users under control when there are few users so as to increase demand by attracting more users. When an important customer base is sharing considerable amounts of resources, the provider can achieve better revenue by decreasing the proposed prices to users and can rely on the NEP envelop to set the optimal values. This outcome of the model must be tuned with more realistic user-provider relationships as it reflects only the optimal strategy derived from the model. The model simply indicates that for small customer bases the best approach is to continue attracting more users while the right pricing must be selected when more users join. This subtle game between users and providers needs to be investigated further. The lesson is that providers need large amounts of resources and many users for a more profitable service and this is the natural setting for cloud computing services and commercial offers today. Table I shows some numerical values for \(\mu^j_i\), \(N\), \(B^j\).

![Figure 1. The evolution of proposed price \(\mu^j_i\) at the Nash equilibrium point with the number of users \(N\) and the limited resources \(B^j\).](image)

An other experiment has thus been set-up to observe the evolution of the cloud provider’s revenue with the number of users \(N\). \(M = 10\), \(\alpha_{av} = 10\) and \(N\) is in the range from 1 to 1000. Note that the cloud provider’s revenue is given by (46) for example. Figure 2 depicts an exponential evolution of \(R_j\) when \(N\) is in the range from 1 to 100, a slight evolution when \(100 \leq N \leq 350\), and a stable (additional revenue tapering off) evolution for \(N \geq 350\). Results prove that \(R_j\) reaches a limit related to the upper bound given by (47) for \(R_j \approx 100\). This can be explained as follows: In the beginning of the interaction between users and the provider, the resources are free and approximately 350 users are served for a revenue neighbouring 100. Next, the additional revenue tapers off when the provider’s maximum capacity is reached. This may lead the cloud provider to play an obvious strategy which consists in rejecting new users \((N_{rejected} = 1000 - 350 = 650)\). We argue that the strategy to reject users...
is not necessarily the best one. In fact, rejecting users may affect the provider’s reputation represented in Section II by the factor \( M \). Figure 3 shows that when a provider affects its reputation (for instance from \( M = 7 \) to \( M = 5 \)) then the revenue decreases. The solution for the cloud provider is to deploy the right amount of resources to scale and maintain good reputation and good revenue while users are also satisfied. An additional solution for the cloud provider to scale its resources is to be involved in a cloud federation and thus avoid the need for more capital investments. Thus, outsourcing resources to the federated cloud providers can be preferable to overprovisioning a private data center when demand varies over time. In addition, federation also allows cloud providers to insource their resources to other providers if these are not being used.

![Figure 2. Cloud provider's revenue versus the number of users](image)

![Figure 3. Cloud provider's revenue comparison before and after dropping reputation](image)

V. CONCLUDING REMARKS

This paper explored the constrained pricing problem in order to maximize the revenue of a cloud provider dealing with a finite set of users located in different geographic locations. We formulated and modeled mathematically the problem as a Stackelberg game and presented an optimal pricing policy for different cases. In the case of limited resource capacity and for a large number of users, the obvious strategy to reject a certain number of users is not the best decision to select. The results show that this strategy will decrease the provider’s revenue and reputation. The best approaches consist of investing more when needed so as to keep a good ratio between the customer base and the amount of cloud computing resources. Getting involved in a federation is an attractive alternative for the providers and is the subject of our next investigations.

Our future research will address also competition among different cloud providers dealing with a finite set of users. This will be modeled as a game with multiple leaders and multiple followers. The notion of coalitions and federations should also be in the scope to identify the best strategies for cloud computing stake holders. The fact that compute, storage and network resources have different characteristics and constraints is also envisaged by introducing a criterion per resource type in the utility function leading to a multi-criteria and constraints optimisation.

ACKNOWLEDGMENT

The authors are grateful for the useful comments and recommendations provided by the reviewers.

REFERENCES


