European Journal of Operational Research
Available online 10 November 2018
In Press, Corrected Proof

Production, Manufacturing, Transportation and Logistics

Under what conditions can an application service firm with in-house computing benefit from cloudbursting?

Chen, Li-Ming [a], Chang, Wei-Lun [b]

Show more

https://doi.org/10.1016/j.ejor.2018.11.016

Highlights

- Cloudbursting advantage and disadvantage for an application service firm.
- Analysis of the best cloud choice under non-competition and competition systems.
- Nash equilibrium in competition systems.
- Interface between operations management and information management.

Abstract

Cloudbursting, a hybrid cloud computing model, helps firms supplement their internal computing capacity from a private cloud by using external computing resources from a public cloud to meet increased demand. This paper examines whether cloudbursting benefits an application service firm that uses only its in-house capacity. Cloudbursting provides computing scalability and cost-effectiveness, but poses potential risks from data leakage when bursting into a public cloud. A private cloud reduces such risks, however, is constrained by resource limitations. We develop quantitative models under both non-competitive and competitive systems, and then determine the best choice between cloudbursting and a private cloud. Overall, we show that a profit-maximizing firm will benefit from migrating to cloudbursting if risk is considerably low and will maintain a private cloud if risk is considerably high. Interestingly, one exception occurs in a competition system whereby two competing firms access external resources as needed from the same public cloud. In this case, one firm will counterintuitively remain in a private cloud even though risk is considerably low while its competitor will migrate to cloudbursting. The numerical study conducts a sensitivity analysis to link a firm's profit performance with its best cloud choice.

Keywords

OR in service industries; Service supply chain; Game theory

1. Introduction

Cloud computing continues to create buzz in today's business landscape. Thanks to the dizzying evolution of cloud deployment technology, a hybrid application deployment mode, *cloudbursting,* emerges whereby an application first uses a firm's internal resources from its in-house IT data center or a private cloud, and then bursts into external resources from a public cloud, whenever processing workloads exceed its internal computing capacity. Thus, cloudbursting offers a solution for firms that provide computing applications based only on their internal computing resources but need additional computing resources when demand spikes.

To illustrate cloudbursting implementation, we use the case of rendering, often called image synthesis. It is a computerized process of generating an image from a 2D or 3D model. A scene file of an animation contains complex objects reflecting geometry, viewpoint, texture, lighting, color, shade, and various visual specifications. The result of animation rendering is to convert 2D or 3D models into digital images for an animation video. Unfortunately, to consolidate the animation rendering work, animation studios require intensive computing resources. When rendering workloads spike and the internal computing resources cannot keep pace handling workloads, studios must either expand their own data center or miss their deadline and risk losing clients. In this case, cloudbursting enables studios to quickly move workloads to a public cloud when their internal resources hit computing limits. Amazon Web Service (AWS), a public cloud provider, is well known in the rendering industry to lease external resources to many studios (Vecchiola, Chu, & Buyya, 2009). Dell offers another cloudbursting service to successfully fulfill business needs. For instance, over the 2013 Black Friday–Cyber Monday weekend (a four-day U.S. Thanksgiving holiday weekend), Dell Inc. contracted with Microsoft Azure, a cloud computing platform that provides external resources, to move online shoppers' processing from its in-house IT to Microsoft Azure's public cloud when needed. Compared to the four-day shopping period in the previous year, Dell.com received more than 920 million hits, with more than 140 million of these hits via Microsoft Azure (Microsoft Corp., 2014). The increased hits illustrate how cloudbursting enabled Dell to gain more on-line orders and profits; since then Dell Inc. has collaborated with Microsoft Azure.

Cloudbursting is not only valued for matching business opportunity with unlimited computing resources, but is considered a strategic computing model that can save firms from investing in IT infrastructure to extend their private computing capacity along with additional IT personnel and training costs (Armbrust et al., 2010, Lee and Lodree, 2010, Mattess et al., 2011). This savings are observed in the biotechnology industry where many startups have preferred to spend large segments of their budgets on drug research than on IT computing infrastructure. However, today's drug development requires heavy computing work to analyze data collected from clinical experiments so cloudbursting offers an elegant solution for such data analysis.

Based upon these aforementioned practices, computing scalability and cost effectiveness are what drive firms with in-house computing to contemplate migrating to cloudbursting. However in doing so, these firms face security concerns since the computing environment in a public cloud is under *multi-tenancy* whereby malicious *tenants* could intentionally hack others' data (Levitin et al., 2018, Alani, 2014, Nalinipriya et al., 2016, Godfrey and Zulkernine, 2014, Marston et al., 2011, Ren et al., 2012). Equifax, which announced in September 2017 that its database had been hacked, highlights the recent spate of hacking into large company databases. Equifax, one of three major consumer credit reporting agencies whose databases had been penetrated illegally, reported the loss of 209,000 credit card numbers and 182,000 personal identity losses, including social security numbers (Bernard, Hsu, Perlroth, & Lieber, 2017). Similarly, U.S. retail giant Target reported that malicious tenants had hacked into its database in November 2013, resulting in the data breach of personal and credit card information of up to 110 million individuals (Babcock, 2014).

This tradeoff between the benefits of resource scalability and cost effectiveness versus damage from data breaches stimulated us to better assess if an application service firm with its in-house computing capacity should stay in a private cloud or adopt cloudbursting. Note that with technological advancements, movements between different cloud platforms and latency issues are not our research concern. The goal of this study is to strategically explore this question: Under what conditions can an application service firm with in-house computing benefit from cloudbursting? To answer this question, we develop quantitative models and then apply them to the problem of determining this firm's choice between a private cloud and cloudbursting.

The modeling framework is a two-stage service supply chain. An application service firm, called a downstream player, provides the application to users, which needs computing resources. The public cloud such as AWS or Windows Azure, called an upstream player, is the provider of the external computing resources. An application service firm can distribute resources to handle workloads in two ways. The first way is via internal resources only and the second way is firstly via internal resources and then through external resources for unmet workloads. The resource distribution to meet workloads is similar to those examined in conventional supply chain studies whereby the physical products move from upstream and downstream to end customers. Demirkan and Cheng (2008) and Demirkan, Cheng, and Bandyopadhyay (2010) have utilized this concept to analyze the application service supply chain by developing the newsvendor model to determine an application service provider's (ASP's) resource capacity from an application infrastructure provider (AIP). Our quantitative models adopt the newsvendor-typed model structure and the application service firm is the focal firm in the discussion.

We first develop the model for the non-competition system. Later, with similar applications offered by various thriving application service firms, we consider two firms under competition. In the competition system, two firms might access external resources as needed either from different public cloud providers or from the same public cloud provider. Thus, we develop the model for two firms accessing external resources from separate public cloud providers and then we develop the model for two firms accessing external resources from the same public cloud provider. In each of three models, we formulate a firm's (or a competing firm's) profit functions with respect to two cloud types: private and cloudbursting. The best cloud choice is based on the firms' financial performance. To the best of our knowledge, this is the first study to assess whether cloudbursting is better for a profit-maximizing application service firm with in-house computing.

Within this framework we generate the following contributions. Firstly, this work analyzes whether migrating to cloudbursting is worthwhile for an application service firm with in-house computing in both non-competition and competition systems. Overall, our results reinforce the general understanding that a firm should maintain a private cloud if security risk is considerably high and that it should use cloudbursting if risk is considerably low. However, we surprisingly find that, even under low risk, this firm under competition may counterintuitively maintain a private cloud while its competitor migrates to cloudbursting. The analysis reveals that this competing firm's best cloud choice is not affected simply by the risk level of accessing external resources, but also by the magnitude of *demand interchangeability* — defined as the shift of application demand from a competing firm to another, and the acquisition channel of external resources between competing firms (i.e., external resources from separate public cloud providers or from a common public cloud provider). Further, this work conducts a numerical study through a sensitivity analysis to link a firm's profit performance with its best cloud deployment decision in both non-competition and competition systems.

The remaining sections are organized as follows. We first review the relevant literature. Then, we consider a firm's best cloud deployment strategy under two scenarios: non-competition and competition, followed by a numerical study. Lastly, we summarize our findings and offer key managerial insights.

2. Literature review

We classify the relevant literature into two streams. The first provides an overview of cloud computing and presents quantitative models for computing resource allocation. The second is from the operations management context on sourcing issues based on the newsvendor problem.

Cloud computing is a distribution technology that enhances flexibility, scalability, and quality of cloud service (Iyer and Henderson, 2012, Benlian and Hess, 2011, Sultan, 2011, Venters and Whitley, 2012, Vithayathil, 2018, Lang et al., 2018). Four types of clouds are identified by the U.S. National Institute of Standards and Technology (NIST): private cloud, community cloud, public cloud, and hybrid cloud (Pardeshi, 2014). Cloudbursting belongs to a hybrid cloud. Models of cloud computing are discussed in the literature. For instance, Levitin et al. (2018) optimize users' data protection policy in a cloud computing environment subject to co-residence security risks. Chen and Wu (2013) examine the impact of cloud computing on market structure, firm profitability, and consumer welfare; their model captures the benefit of unlimited capacity and the barrier from security concerns. The impact of security risk is also studied in Nicolaou et al., 2012, Furuncu and Sogukpinar, 2015, August et al., 2014, Oliveira et al., 2014, and Schwarz, Jayatilaka, Hirschheim, and Goles (2009). Aware of the security concerns of accessing external resources, albeit the benefits from unlimited capacity, we consider this benefit-risk conundrum in our three models.

Cloud models from the cloud providers' perspective are quite popular. Püschel, Schryen, Hristova, and Neumann (2015) introduce policy-based service admission control models to maximize the revenue of cloud providers. In a similar vein, revenue management through pricing mechanisms for cloud providers is investigated in Keskin and Taskin, 2015, Naldi and Mastroeni, 2016, Winkler and Brown, 2013, and Ding, Xia, Wang, Wu, and Zhang (2017). In addition, Dorsch and Häckel (2014) propose an optimization model to analyze the capacity-planning problem of a cloud service vendor aimed to handle fluctuated demand. They consider a situation where a vendor outsources specific segments of a business process to external providers, whereby the vendor can choose among different plans: dedicated, fixed or elastic capacity, or a combination thereof; these capacities are discussed in our study's examination of the benefits of maintaining a private cloud or deploying cloudbursting.

In terms of the model structure, our work most relates to that of Anselmi, Ardagna, and Passacantando (2014) and Ardagna, Panicucci, and Passacantando (2013)). They propose a service provisioning problem for a platform as a service provider (PaaS) who provides computing resources to software as a service providers (SaaSs), as a Generalized Nash Equilibrium problem. More studies related to the Generalized Nash Equilibrium problem are reviewed in Bigi et al., 2013, Facchinei and Kanzow, 2010, and Cavazzuti, Pappalardo, and Passacantando (2002). Although various studies on cloud modeling are proposed, most are initiated from the resource providers' perspective. Our work enriches the existing literature by developing models for the service provider and identifying the Nash Equilibrium to help determine the service provider's cloud choice between cloudbursting and a private cloud.

Cloudbursting involves resource allocation from various sources, and the operations management literature, the second stream of literature we examine, offers abundant allocation research in this area, including sourcing strategies for managing inventory. In particular, inventory is sourced either from two suppliers, referred to as *dual sourcing*, or from the same supplier but with different timings, referred to as a *recourse option*. Numerous papers analyze the tradeoffs between sourcing from a low-cost but risky supplier and from a high-priced but reliable supplier (Tomlin, 2006, Tomlin, 2009a, Tomlin, 2009b, Wang et al., 2010). The newsvendor problem is a widely applicable tool used to manage dual sourced inventory. It can help determine the best order amount from both sources to match uncertain demand so that the overall cost (profit) is minimized (maximized). Reviews of this problem can be founded in Silver, Pyke, and Peterson (1998) and Qin, Wang, Vakharia, Chen, and Seref (2011).

The recourse option is another sourcing method that a downstream player uses to place an order in advance and the upstream supplier provides its recourse option to place another order at an emergency purchase price that is higher than the regular price when demand exceeds supply. The value of the emergency supply is similar to that of the external resources which handle unmet demand in cloudbursting. There exists a burgeoning literature focusing on the newsvendor under recourse option including Vipin and Amit, 2017, Agrawal and Seshadri, 2000, Brito and de Almeida, 2012, Khouja, 1996, Lee and Lodree, 2010, and Lodree (2007). Inspired by the newsvendor problems under dual sourcing and the recourse option, our work applies this widely popular newsvendor model to analyze application service firms' optimal cloud choices.

3. Model setting

Consider an application service firm that provides the application service to end users and such service workloads are handled by computing resources. Since each use of the service could require multiple resources, total workloads from all end users are quantified by the computing resources needed, denoted as *application demand* (or *demand*). For instance, if five end users are using the service and each service consumes 2 units of resources, the application demand is 10 units. Let application demand d, a random variable in a time interval with a positive support $[0, \bar{\xi}]$, a probability density function $f(\cdot)$, and a cumulative density function $F(\cdot)$, take the stochastic form. Explicitly, realized demand ξ is in the interval $[0, \bar{\xi}]$ with $\int_0^{\bar{\xi}} f(\xi)\, d\xi = 1$ and $F(a) = \int_0^a f(\xi)\, d\xi$. The length of such a time interval varies with different application services. For example, rendering workloads use resources continually thus the appropriate time interval can be a minute. However, for on-line shopping, each resource is occupied by shoppers, on average, for several minutes. Thus the appropriate time interval is a second. The service firm adopts the *pay-as-you-go* mode to charge a unit resource with the service fee p paid by the end users. Resource acquisition comes from two sources: internal (a private cloud) and external (leased from a public cloud provider). Internal resource units have a marginal cost w^P that is less than p; such resources have no risk concerns. Note that this cost is composed of infrastructure costs (hardware and software), labor costs (staff to operate the private cloud), energy costs (electricity), and service costs (routine operations) (Martens and Teuteberg, 2012, Walterbusch et al., 2013). External resources have a marginal cost w^C, which includes rental fees for external computing from public cloud providers and integration costs for internal systems. When application demand is computed through external resources, risk rises. Thus, a penalty cost π for each leased external resource represents the business loss if a risk actually becomes a data breach.

3.1. Non-competition system

With resources distributed from two sources, we first portray the model for a monopolistic application service firm that adopts a private cloud alone, labeled P, in which this firm has internal resources, q units, which incur a capital expenditure $w^P q$. Later, application demand ξ is realized, which generates revenue $p\min[\xi, q]$. To sum up, the firm with deployment P reaps the expected profit

$$\Pi(q|P) = p\mathbf{E}\min[d, q] - w^P q. \tag{1}$$

Next, we consider a monopolistic application service firm that will burst into a public cloud when demand spikes and internal resources cannot handle workloads (i.e., cloudbursting), labeled C. This cloudbursting cloud deployment compensates for the deficiency of deployment P by fulfilling application demand via internal resources first and bursting into a public cloud when demand exceeds its in-house capacity. The sequence of events to build deployment C is as follows. First, a firm owns internal resources q prior to demand realization, with a total cost of $w^P q$. When demand ξ arrives, the firm first allocates its internal resources to serve application demand. If the internal resources cannot fulfill demand, a firm bursts into a public cloud to handle unmet demand $[\xi - q]^+$ so that the firm generates revenue $p\min[\xi, q]$ from its internal resources and creates additional profits $(p - w^C)[\xi - q]^+$ from external resources. Through cloudbursting, the firm has risk incidence $\rho \in (0, 1)$, resulting in a total penalty cost $\pi[\xi - q]^+$. In other words, each external resource generates expected profit $(p - w^C - \rho\pi)$. Accordingly, a firm in deployment C has the expected profit

$$\Pi(q|C) = p\mathbf{E}\min[d, q] - w^P q + (p - w^C - \rho\pi)\,\mathbf{E}[d - q]^+. \tag{2}$$

We depict deployment P and C in Fig. 1 with summarized notations in Table 1.

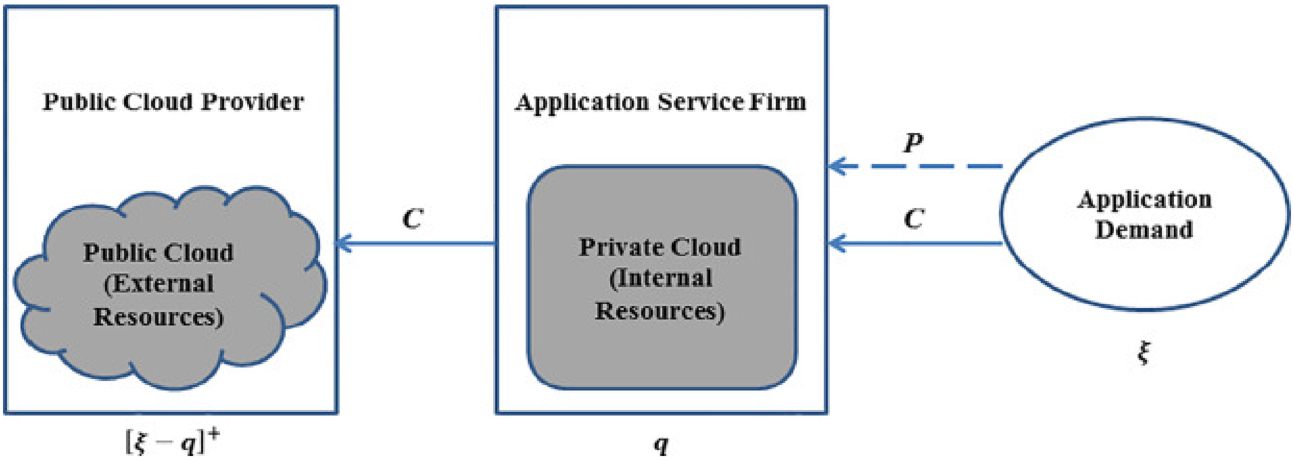


Fig. 1. Deployment P and C.

Table 1. Summary of notations.

Notation	Description
d	Random demand with support $[0, \bar{\xi}]$, pdf $f(\cdot)$ and cdf $F(\cdot)$
w^P	Marginal resource cost from a private cloud
w^C	Marginal resource cost from a public cloud
p	Marginal service fee
π	Marginal penalty cost
q	Amount of internal resources
ρ	$\in (0, 1)$, risk incidence

Overall, deployment C has a scalable competence to moderate its usage from in-house computing capacity and its leased public cloud. Nonetheless, external resources bring risk from which an application service firm's best deployment choice is based on a measured balance between scalability and risk. To assess this, we define a threshold as

$$\overline{\rho} \equiv \frac{p-w^C}{\pi}, \tag{3}$$

which represents the risk tolerance. With this threshold, the resource decision on both deployment models is illustrated in Lemma 1. We further describe the optimal deployment strategy in Proposition 1.

Lemma 1

An application service firm faces the following decisions on resource allocation:

(a) *In deployment P, the optimal resource amount from a private cloud is set to* $q^P = F^{-1}\left(\frac{p-w^P}{p}\right)$.

(b) *In deployment C, the optimal resource amount from a private cloud is set to* $q^C = F^{-1}\left(\frac{w^C+\rho\pi-w^P}{w^C+\rho\pi}\right)$, *and this firm bursts into a public cloud for unmet application demand.*

Proposition 1

(a) When $\rho < \overline{\rho}$, *an application service firm migrates to cloudbursting; (b) Otherwise, this firm uses its private cloud alone.*

Based on Lemma 1 and Proposition 1, the firm's cloud choice (private or cloudbursting) is influenced by parameters related to external resources. When either the marginal penalty cost or the marginal cost of external cost is considerably low, a firm has a larger risk tolerance so that cloudbursting becomes preferable when risk due to leased external resources is considerably low and further this firm requires less internal resources than it needs from a private cloud because excessive demand can prompt it to burst into a public cloud. However, when such a penalty cost or marginal cost is high (i.e., considerably high risk), a firm tends to favor a private cloud and interestingly uses less internal resources than it would from cloudbursting. We also find that the threshold $\overline{\rho}$ (e.g., the risk tolerance) for evaluating this firm's cloud choice is irrelevant to the application demand distribution.

4. Competition system

Move to study that firms with similar application service compete on the market. This section considers a duopolistic system from which two firms, i and j, compete for application demand. We use subscript $n \in \{i, j\}$ to denote firm n with application demand d_n, a random variable with the positive support $\xi_n \in [0, \overline{\xi}]$, probability density function $f_n(\cdot)$, and cumulative density function $F_n(\cdot)$. In addition, firm n's unit service fee, marginal internal resource cost, marginal penalty cost from external resources, and amount of internal resources are represented as p_n, w_n^P, π_n, and q_n, respectively. In most regards, each firm's model setting in deployment P and C is identical to the monopoly case except for *demand interchangeability* between firms, defined as the migration of demand from one firm to another (for more on such interchangeability under competition, see Lippman and McCardle (1997)). In other words, the interchangeability of application demand from one firm to another is the main difference between the duopoly and monopoly models.

Two factors drive demand interchangeability. The first one comes from the constraint of scaling up resources when demand unexpectedly surges in deployment P. As a result, partial unmet demand shifts service to the competing firm. Such a shift is characterized by a proportion ν of unmet demand where $0 \leq \nu \leq 1$. In this work, we simply and reasonably assume $\nu \equiv 1$; that is, firm n's unmet demand, quantified by $[\xi_n - q_n]^+$, will reallocate the application service to the competing firm. The second factor comes from the risk impact when a firm bursts into a public cloud in deployment C. Once risk is realized, partial application demand, which was originally served by external resources, shifts service to another firm although a service fee has been paid to the original application service firm. The proportion of demand shift between firms due to this factor is measured by a proportion ϕ, where $0 \leq \phi \leq 1$ so that firm n reallocates demand $\phi[\xi_n - q_n]^+$ to its competitor. In the later, demand interchangeability specifically refers to this second factor.

We now define the cloud architecture for two firms under competition. Let (x, y) be a firm's cloud structure where firm i has deployment x and firm j has deployment y. With each firm's two choices on cloud deployment, there are four possible cloud structures: $(x, y) \in \{(P, P), (C, P), (P, C), (C, C)\}$. Firms i and j's profit functions are $(\Pi_i(\mathbf{q}|x, y), \Pi_j(\mathbf{q}|x, y))$, respectively when firm i adopts deployment x and firm j has deployment y, where $\mathbf{q} = (q_i, q_j)$. With profit functions, we are interested in the equilibrium of cloud deployment for the competing firms (i.e., the best cloud deployments for each firm. To solve this, we organize a two-stage game. In the first stage, we must identify the existence of $\mathbf{q}^* = \left(q_i^*, q_j^*\right)$ for each cloud structure and this existence indicates a pure strategy Nash equilibrium if neither firm can improve its profit by unilaterally changing its internal resources; that is, $\Pi_i\left(\left(q_i^*, q_j^*\right)|x, y\right) \geq \Pi_i\left(\left(q_i, q_j^*\right)|x, y\right)$ and $\Pi_j\left(\left(q_i^*, q_j^*\right)|x, y\right) \geq \Pi_j\left((q_i^*, q_j)|x, y\right)$ for all $q_n \geq 0$, where $n \in \{i, j\}$. Note that each cloud structure has a corresponding $\mathbf{q}^*$. We next move to the second stage by solving a simple game, illustrated in Fig. 2, to identify the best cloud deployment for each firm under competition.

		firm j	
		P	C
firm i	P	$\Pi_i(\mathbf{q}^*|P,P)$ $\Pi_j(\mathbf{q}^*|P,P)$	$\Pi_i(\mathbf{q}^*|P,C)$ $\Pi_j(\mathbf{q}^*|P,C)$
	C	$\Pi_i(\mathbf{q}^*|C,P)$ $\Pi_j(\mathbf{q}^*|C,P)$	$\Pi_i(\mathbf{q}^*|C,C)$ $\Pi_j(\mathbf{q}^*|C,C)$

Fig. 2. Cloud deployment game in the second stage.

With the above articulation of the necessary denotation and deployment structure, we now start to quantitatively develop models under competition. We first consider the competition system in which both firms, in cloud structure (C, C) under the first stage, access external resources from separate public cloud providers. Since application service firms' choices of public cloud providers for leased external resources are growing, we use VMware, the industry-leading virtualization software company as an example to explain why we develop this competition system. VMware launched the vCloud hybrid cloud service in 2013 to compete with AWS (Leong, 2013). We then study the competition system in which both firms, in cloud structure (C, C) under the first stage, access external resources from a common public cloud provider. The aforementioned rendering industry performs a real-life practice that many film studios are known to burst into the same public cloud — AWS during peak time.

4.1. System in which firms access external resources as needed from separate public cloud providers

For competing application service firms that access external resources as needed from separate public cloud providers, let ρ_n denote the risk incidence and $\mathbf{w}_n^C$ the marginal resource cost when firm n bursts into its public cloud, where $n \in \{i, j\}$. Let $\Pi_n^1(\mathbf{q}|x,y)$ be firm n's expected profit under structure (x, y). Next, we formally develop the models, starting from structure (P, P). When both firms deploy only a private cloud, they assume no risk. Thus, a firm's *stealth* of demand comes from another firm's insufficiency of internal resources (the aforementioned first factor of demand interchangeability). Explicitly, when firm j has internal resources q_j, realized application demand ξ_j impedes the fulfillment of amount $[\xi_j - q_j]^+$, which then migrates to its competitor and so that firm i has potential application demand $\xi_i + [\xi_j - q_j]^+$. Nevertheless, the actual amount of firm i's stealth on demand depends on the availability of its internal resources as it first serves its own application demand ξ_i. To sum up, firm i computes demand $\mathbf{min}\left[q_i, \xi_i + (\xi_j - q_j)^+\right]$ using its own resources. Accordingly, expected profits of the two firms in (P, P) are

$$\begin{aligned} \Pi_i^1(\mathbf{q}|P,P) &= p_i \mathbf{E}\mathbf{min}\left[q_i, d_i + (d_j - q_j)^+\right] - w_i^P q_i \\ \Pi_j^1(\mathbf{q}|P,P) &= p_j \mathbf{E}\mathbf{min}\left[q_j, d_j + (d_i - q_i)^+\right] - w_j^P q_j. \end{aligned} \tag{4}$$

Next, for structure (C, P) (e.g., firm i with cloudbursting and firm j with a private cloud), firm j has the likelihood not to fulfill all application demand because of insufficient resources and thus firm i can benefit by stealing demand from firm j. Thus, firm i can satisfy demand $\mathbf{min}\left[q_i, \xi_i + (\xi_j - q_j)^+\right]$ through internal resources and demand $\left[\xi_i + (\xi_j - q_j)^+ - q_i\right]^+$ by bursting into a public cloud. Firm i's marginal profit from external resources is $p_i - w_i^C - \pi_i$ in the presence of risk, and equals $p_i - w_i^C$ in the absence of risk. Although firm i gains demand from firm j's demand migration, bursting into a public cloud puts firm i at risk of losing a proportion ϕ of its application demand to firm j. Thus, firm j's private cloud supplies demand $\mathbf{min}\left[q_j, \xi_j + \phi(\xi_i - q_i)^+\right]$ when firm i faces risk, and it supports only demand $\min[q_j, \xi_j]$ when firm i has no such risk. Expected profits of two firms in structure (C, P) are expressed as

$$\begin{aligned} \Pi_i^1(\mathbf{q}|C,P) &= \rho_i \left\{ p_i \mathbf{E}\mathbf{min}\left[q_i, d_i + (d_j - q_j)^+\right] \right. \\ &\quad \left. + \left(p_i - w_i^C - \pi_i\right) \mathbf{E}\left[d_i + (d_j - q_j)^+ - q_i\right]^+ \right\} \\ &\quad + (1 - \rho_i) \left\{ p_i \mathbf{E}\mathbf{min}\left[q_i, d_i + (d_j - q_j)^+\right] \right. \\ &\quad \left. + \left(p_i - w_i^C\right) \mathbf{E}\left[d_i + (d_j - q_j)^+ - q_i\right]^+ \right\} - w_i^P q_i. \\ \Pi_j^1(\mathbf{q}|C,P) &= \rho_i p_j \mathbf{E}\mathbf{min}\left[q_j, d_j + \phi(d_i - q_i)^+\right] \\ &\quad + (1 - \rho_i) p_j \mathbf{E}\mathbf{min}\left[q_j, d_j\right] - w_j^P q_j. \end{aligned} \tag{5}$$

Symmetrically, in structure (P, C) in which firm i deploys a private cloud and firm j adopts the cloudbursting model, expected profits are

$$
\begin{aligned}
\Pi_i^1(\mathbf{q}|P,C) &= \rho_j p_i \mathbf{E}\min\left[q_i, d_i + \phi(d_j - q_j)^+\right] \\
&+ (1-\rho_j)\, p_i \mathbf{E}\min\left[q_i, d_i\right] - w_i^P q_i \\
\Pi_j^1(\mathbf{q}|P,C) &= \rho_j \{p_j \mathbf{E}\min\left[q_j, d_j + (d_i - q_i)^+\right] \\
&+ \left(p_j - w_j^C - \pi_j\right) \mathbf{E}\left[d_j + (d_i - q_i)^+ - q_j\right]^+ \Big\} \\
&+ (1-\rho_j)\ \{p_j \mathbf{E}\min\left[q_j, d_j + (d_i - q_i)^+\right] \\
&+ \left(p_j - w_j^C\right) \mathbf{E}\left[d_j + (d_i - q_i)^+ - q_j\right]^+ \Big\} - w_j^P q_j.
\end{aligned} \tag{6}
$$

The last structure remaining is (C, C) under which both firms burst into a public cloud with different providers. When firm i encounters risk while firm j does not, this case has a probability $\rho_i(1-\rho_j)$ under which firm i serves demand min[q_i, ξ_i] through its internal resources, and demand $[\xi_i - q_i]^+$ via external resources with a marginal profit $p_i - w_i^C - \pi_i$. When firm i does not encounter risk while firm j faces risk with a probability $(1-\rho_i)\rho_j$, firm i supplies demand $\min\left[q_i, \xi_i + \phi(\xi_j - q_j)^+\right]$ from its internal resources and demand $\left[\xi_i + \phi(\xi_j - q_j)^+ - q_i\right]^+$ by bursting into a public cloud with the marginal profit $p_i - w_i^C$. Once both firms simultaneously face risk with a probability $\rho_i\rho_j$, firm i serves demand $\min\left[q_i, \xi_i + \phi(\xi_j - q_j)^+\right]$ from its internal resources and demand $\left[\xi_i + \phi(\xi_j - q_j)^+ - q_i\right]^+$ from external resources with the marginal profit $p_i - w_i^C - \pi_i$. Lastly, when both firms face no risk with a remaining probability $(1-\rho_i)(1-\rho_j)$, firm i serves demand min[q_i, ξ_i] from its internal resources, and handles unmet demand $[\xi_i - q_i]^+$ from external resources with a marginal profit $p_i - w_i^C$. The expected profits in structure (C, C) are expressed as

$$
\begin{aligned}
\Pi_i^1(\mathbf{q}|C,C) &= \rho_i\rho_j \{p_i \mathbf{E}\min\left[q_i, d_i + \phi(d_j - q_j)^+\right] \\
&+ (p_i - w_i^C - \pi_i)\, \mathbf{E}\left[d_i + \phi(d_j - q_j)^+ - q_i\right]^+ \Big\} \\
&+ (1-\rho_i)\,\rho_j \{p_i \mathbf{E}\min\left[q_i, d_i + \phi(d_j - q_j)^+\right] \\
&+ (p_i - w_i^C)\, \mathbf{E}\left[d_i + \phi(d_j - q_j)^+ - q_i\right]^+ \Big\} \\
&+ \rho_i (1-\rho_j)\, \{p_i \mathbf{E}\min\left[q_i, d_i\right] \\
&+ (p_i - w_i^C - \pi_i)\, \mathbf{E}\left[d_i - q_i\right]^+ \} \\
&+ (1-\rho_i)(1-\rho_j)\, \{p_i \mathbf{E}\min\left[q_i, d_i\right] \\
&+ (p_i - w_i^C)\, \mathbf{E}\left[d_i - q_i\right]^+ \} - w_i^P q_i \\
\Pi_j^1(\mathbf{q}|C,C) &= \rho_i\rho_j \{p_j \mathbf{E}\min\left[q_j, d_j + \phi(d_i - q_i)^+\right] \\
&+ \left(p_j - w_j^C - \pi_j\right) \mathbf{E}\left[d_j + \phi(d_i - q_i)^+ - q_j\right]^+ \Big\} \\
&+ \rho_i (1-\rho_j)\, \{p_j \mathbf{E}\min\left[q_j, d_j + \phi(d_i - q_i)^+\right] \\
&+ \left(p_j - w_j^C\right) \mathbf{E}\left[d_j + \phi(d_i - q_i)^+ - q_j\right]^+ \Big\} \\
&+ (1-\rho_i)\,\rho_j \{p_j \mathbf{E}\min\left[q_j, d_j\right] \\
&+ \left(p_j - w_j^C - \pi_j\right) \mathbf{E}\left[d_j - q_j\right]^+ \Big\} \\
&+ (1-\rho_i)(1-\rho_j)\, \{p_j \mathbf{E}\min\left[q_i, d_i\right] \\
&+ \left(p_j - w_j^C\right) \mathbf{E}\left[d_j - q_j\right]^+ \Big\} - w_j^P q_j.
\end{aligned} \tag{7}
$$

Based on the two-stage game, the analysis in the first stage shows that each of the four cloud structures has a pure-strategy, unique Nash equilibrium $\mathbf{q}^* = \left(q_i^*, q_j^*\right)$, formally presented as Proposition 2. (For more, see Lippman and McCardle (1997).)

Proposition 2

A pure-strategy, unique Nash equilibrium $\left(q_i^*, q_j^*\right)$ *exists in each cloud structure.*

The second stage follows to analyze competing firms' best cloud deployment. We define a threshold as $\overline{\rho}_n \equiv \frac{p_n - w_n^C}{\pi_n}$ for $n \in \{i, j\}$; the results are listed in Theorem 1.

Theorem 1

For the system in which two competing firms access external resources as needed from separate public cloud providers, each firm's best strategy is:

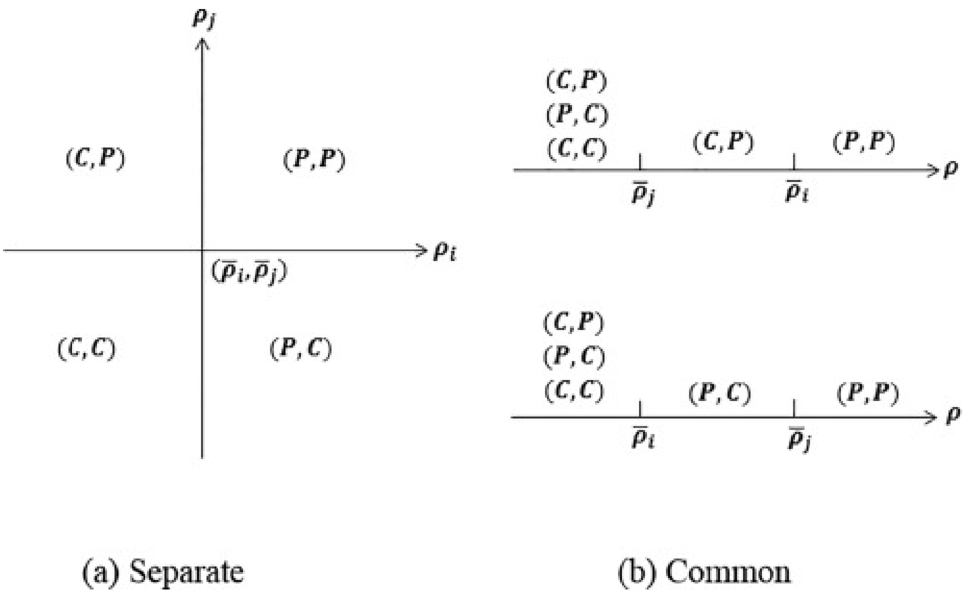
(a) *When* $\rho_i < \overline{\rho}_i$ *and* $\rho_j < \overline{\rho}_j$, *both firms deploy cloudbursting;*

(b) *when* $\rho_i \geq \overline{\rho}_i$ *and* $\rho_j \geq \overline{\rho}_j$, *both firms maintain a private cloud;*

(c) *when* $\rho_i \geq \overline{\rho}_i$ *and* $\rho_j < \overline{\rho}_j$, *firm i maintains a private cloud and firm j deploys cloudbursting;*

(d) *otherwise, firm i deploys cloudbursting and firm j maintains a private cloud.*

As seen, each firm makes its best cloud deployment decision *independently* based upon its risk level and its risk tolerance (i.e., $\overline{\rho}_n$). That is, cloudbursting is preferable either (1) when a firm encounters relatively less risk from leased external resources, or (2) a firm is responsible either for a lower penalty cost while exposed to risk or for a lower marginal cost of external resources (i.e., lower π_n or w_n^C). Otherwise, a firm benefits from a private cloud (Fig. 3(a)). Notice that these results are identical to those in the monopolistic system.



Download : Download high-res image (117KB)

Download : Download full-size image

Fig. 3. Equilibria of deployment for two competing application service firms.

4.2. System in which firms access external resources as needed from a common public cloud provider

We proceed to consider that two firms access external resources as needed from a common public cloud provider. In this system, the risk incidence of two firms is $\rho = \rho_i = \rho_j$ and the marginal cost of external resources is the same, $w^C = w_i^C = w_j^C$. Let $\Pi_n^2(\mathbf{q}|x, y)$ be firm *n*'s expected profit under cloud structure (*x*, *y*), where $n \in \{i, j\}$. In structures (*P*, *P*), (*C*, *P*), and (*P*, *C*), the two firms' expected profits are the same as when the two firms consider to access external resources from separate public cloud providers so that

$$\Pi_n^2(\mathbf{q}|x, y) = \Pi_n^1(\mathbf{q}|x, y), \tag{8}$$

where $(x, y) = \{(P, P), (C, P), (P, C)\}$ and $n \in \{i, j\}$. However, in structure (*C*, *C*), the two firms' expected profits are different. This is because the occurrence of risk lets both firms that burst into the same public cloud simultaneously face data breaches while there is likelihood that one firm will encounter data breaches but another will not if they burst into different public cloud providers for external resources. Hence, when risk occurs, firm *i* computes demand $\mathbf{min}\left[q_i, \xi_i + \phi(\xi_j - q_j)^+\right]$ through internal resources and leases external resources for demand $\left[\xi_i + \phi(\xi_j - q_j)^+ - q_i\right]^+$ with a marginal profit $p_i - w^C - \pi_i$. Conversely, in the absence of risk, firm *i* fulfills only demand $\min[q_i, \xi_i]$ through its internal resources and computes the rest of its demand, $[\xi_i - q_i]^+$, via external resources with a marginal profit, $p_i - w^C$. Likewise, the same model development applies to firm *j*. Expected profits of the two firms in structure (*C*, *C*) are

$$\begin{aligned} \Pi_i^2(\mathbf{q}|C, C) &= \rho\{p_i \mathbf{E}\mathbf{min}\left[q_i, d_i + \phi(d_j - q_j)^+\right] \\ &+ (p_i - w^C - \pi_i)\,\mathbf{E}\left[d_i + \phi(d_j - q_j)^+ - q_i\right]^+\} \\ &+ (1-\rho)\{p_i \mathbf{E}\mathbf{min}[q_i, d_i] \\ &+ (p_i - w^C)\,\mathbf{E}[d_i - q_i]^+\} - w_i^P q_i \\ \Pi^2{}_j(\mathbf{q}|C, C) &= \rho\{p_j \mathbf{E}\mathbf{min}\left[q_j, d_j + \phi(d_i - q_i)^+\right] \\ &+ (p_j - w^C - \pi_j)\,\mathbf{E}\left[d_j + \phi(d_i - q_i)^+ - q_j\right]^+\} \\ &+ (1-\rho)\{p_j \mathbf{E}\mathbf{min}[q_j, d_j] \\ &+ (p_j - w^C)\,\mathbf{E}[d_j - q_j]^+\} - w_j^P q_j. \end{aligned} \tag{9}$$

Similar to the system where two firms burst as needed into a public cloud with different providers, we first utilize the same technique used to prove Proposition 2 to verify the existence of a pure-strategy, unique Nash equilibrium $\left(q_i^*, q_j^*\right)$ in each cloud structure (detail omitted). Next we move to the second stage and describe the equilibrium of competing firms' deployment in Theorem 2.

Theorem 2

For the system in which two competing firms access external resources from a common public cloud provider, each firm's best strategy is:

(a) *When $\rho \geq max\{\overline{\rho}_i, \overline{\rho}_j\}$, both firms maintain a private cloud;*

(b) *when $\rho < min\{\overline{\rho}_i, \overline{\rho}_j\}$, either both firms deploy cloudbursting, or one firm deploys cloudbursting and the other maintains a private cloud* (e.g., *structure* (C, C), (C, P), (P, C));

(c) *when $\overline{\rho}_j \leq \rho < \overline{\rho}_i$, firm i deploys cloudbursting and firm j maintains a private cloud;*

(d) *when $\overline{\rho}_i \leq \rho < \overline{\rho}_j$, firm i deploys a private cloud and firm j deploys cloudbursting.*

To conclude Theorems 1 and 2, the results reveal that when the risk level of each firm's public cloud provider is above $\min\{\overline{\rho}_i, \overline{\rho}_j\}$, both firms will *independently* choose the cloudbursting model or a private cloud based on whether their risk level from external resources exceeds the identified threshold. Such independence is irrelevant to each firm's source of external resources. However, with a risk level below $\min\{\overline{\rho}_i, \overline{\rho}_j\}$, each firm's threshold weakens so as to identify whether it should deploy cloudbursting or maintain a private cloud. As found in Theorem 2(b), when risk is considerably low, two competing firms may either (i) both deploy cloudbursting, the result being identical to the that of two firms that consider accessing external resources from separate public cloud providers, or (ii) deploy different cloud models (i.e., one firm chooses a private cloud and the other cloudbursting) (Fig. 3(b)). This interesting outcome of cloud choices shown in (ii) can be rationally explained as follows. We know that a private cloud reduces risk and a cloudbursting model offers a firm more business opportunity by scaling up capacity. With the simultaneous risk incidence, the competing firms might consider mitigating such possible concurrent risk by forming a hedge-like equilibrium where one firm picks the scalability advantage and another firm picks the remaining risk-free advantage.

As for which exact cloud structure from (C, C), (C, P), (P, C) the competing firms that access external resources as needed from the same public cloud will deploy when risk remains low, the model complexity under competition results in the analytical difficulty. However, based on our pilot test, we find that not only do the thresholds of risk identified for both competing firms affect the cloud structure, but the proportion of demand migration between competing firms ϕ also has a decisive influence over firms' best choice of cloud deployment. To gain more insight regarding the impact of this parameter, we first investigate a special case and then conduct a numerical study for a generalized conclusion of the competing firms' best cloud choices.

4.3. A special case: Two competing application service firms with equal market dominance

We assume that both firms have the same level of market-dominance so we equally split demand. In this sense, we set $p = p_i = p_j$; the demand fluctuation between the competing firms is assumed to be constant so we set $\xi_i = \xi_j$ (i.e., $d_i = d_j$) in the models aforementioned. In this special case, each firm's best cloud choice with respect to both competition systems — common or separate cloud providers for external resources as needed — are analyzed. We define a threshold for a firm's proportion of demand migration when risk appears as $\overline{\phi}_n \equiv \frac{-p+w_n^C+\rho_n\pi_n}{\rho_n(p-w_n^C-\pi_n)}$, where $n \in \{i, j\}$; the threshold is assumed to be well defined as $\overline{\phi}_n \in (0,1)$. When competing firms consider bursting into a public cloud, we once again simply set risk $\rho = \rho_i = \rho_j$ and a marginal cost of external resources $w^C = w_i^C = w_i^C$ under which $\overline{\phi}_n \equiv \frac{-p+w^C+\rho\pi_n}{\rho(p-w^C-\pi_n)}$. Results of two competing firms' best cloud deployments in which demand is equally split are described in Theorem 3 and graphically in Fig. 4.

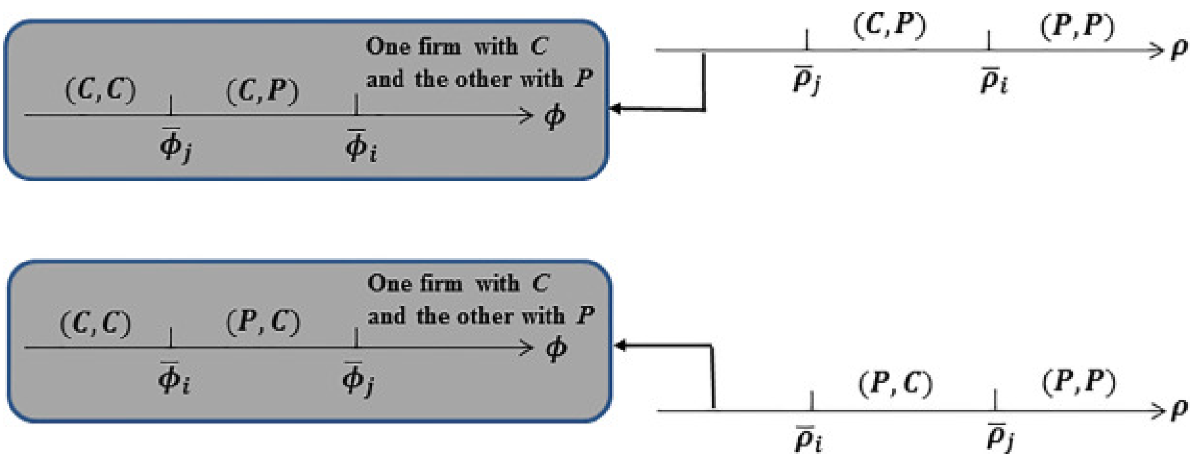


Fig. 4. Equilibria for firms' deployment when two competing firms burst as needed into a common public cloud provider and these two firms have equal market dominance.

Theorem 3

Suppose that two competing application service firms have equal market dominance. In the competition system where both firms as needed burst into separate public cloud providers, the equilibrium of each firm's cloud deployment choice is identical to that in Theorem 1. In the competition system where both firms as needed burst into a common public cloud provider, the equilibrium for each firm's cloud deployment choice is:

(a) *When $\rho \geq max\{\overline{\rho}_i, \overline{\rho}_j\}$, both firms maintain a private cloud;*

(b) *when $\overline{\rho}_j \leq \rho < \overline{\rho}_i$, firm i deploys cloudbursting and firm j maintains a private cloud;*

(c) *when $\overline{\rho}_i \leq \rho < \overline{\rho}_j$, firm i maintains a private cloud and firm j deploys cloudbursting;*

(d) *when $\rho < min\{\overline{\rho}_i, \overline{\rho}_j\}$ and $\phi < min\{\overline{\phi}_i, \overline{\phi}_j\}$, both firms deploy cloudbursting;*

(e) *when $\rho < min\{\overline{\rho}_i, \overline{\rho}_j\}$ and $\overline{\phi}_j \leq \phi < \overline{\phi}_i$, firm i deploys cloudbursting and firm j maintains a private cloud;*

(f) *when $\rho < min\{\overline{\rho}_i, \overline{\rho}_j\}$ and $\overline{\phi}_i \leq \phi < \overline{\phi}_j$, firm i maintains a private cloud and firm j deploys cloudbursting;*

(g) *when $\rho < min\{\overline{\rho}_i, \overline{\rho}_j\}$ and $\phi \geq max\{\overline{\phi}_i, \overline{\phi}_j\}$, one firm deploys cloudbursting and the other maintains a private cloud.*

When both firms have equivalent market dominance, both firms will benefit from maintaining a private cloud if risk is high. Nonetheless, when risk is considerably low and both firms access external resources as needed from the same cloud provider, one firm may surprisingly prefer a private cloud, which is clearly driven by risk and demand interchangeability (see Theorem 3(d)–3(g)). Thresholds $(\overline{\rho}_n, \overline{\phi}_n)$ are identified through parameters related to external resources: the level of risk, the marginal cost, and the unit penalty cost. The following reflects the *attractiveness* of demand interchangeability due to a firm's marginal penalty cost. When firms' penalty costs caused by external resources are considerably low, demand interchangeability due to risk occurrence becomes *attractive* since external resources can handle such a demand shift by gaining more revenue with less penalty burden (i.e., increased attractiveness of demand interchangeability is represented by increased $\overline{\phi}_n$) so that firms prefer cloudbursting (Theorem 3(d)). When risk remains low but a firm's increased penalty cost reduces attractiveness of demand interchangeability, this firm remains with a private cloud while the other firm with the still low penalty cost deploys cloudbursting (Theorem 3(e),(f)). When the attractiveness of demand interchangeability decreases for both firms with the higher penalty cost and considerably low risk, then to prevent the concurrent, severe damage from accessing external resources once risk occurs, both firms choose to deploy different clouds with one firm remaining in a risk-free private cloud and another firm choosing public cloudbursting (Theorem 3(g)); by doing so, risk is diffused.

5. Numerical study

This section conducts a sensitivity analysis to uncover more insights relevant to a firm's best cloud deployment choice with corresponding profit performance. Demand uncertainty for firm n, where $n \in \{i, j\}$, is captured by a truncated normal distribution with mean μ_n and variance ${\sigma_n}^2$. The truncation occurs at $\pm\mu_n$ of the mean value; that is, realized firm n's demand ξ_n is bounded in the interval $[0, 2\mu_n]$. The parameter setting below serves as a baseline.

p_n	w_n^C	w_n^P	π_n	ϕ	μ_n	(σ_i, σ_j)	ρ_n
6	2	3	12	0.3	50	(10, 5)	0.3

We examine three parameters related to the access of external resources: the level of risk, the marginal cost of external resources, and the unit penalty cost, which have been validated so as to relate to $(\overline{\rho}_n, \overline{\phi}_n)$ from our analytical work. The results under the non-competition system show that a monopolistic firm is incentivized to favor cloudbursting when either one of these three parameters is reduced (see firm i profits only (Fig. 5)). This is consistent with our analytical outcome in Theorem 1 where the reduction of these parameters (e.g., a larger threshold $\overline{\rho}$) results in a larger risk tolerance and the choice of cloudbursting. Once this monopolistic firm migrates to cloudbursting, the more reduction in either one of these parameters benefits its profit. Conversely, a monopolistic firm's profit performance decreases when either one of these three parameters increases, and then it prefers a private cloud.

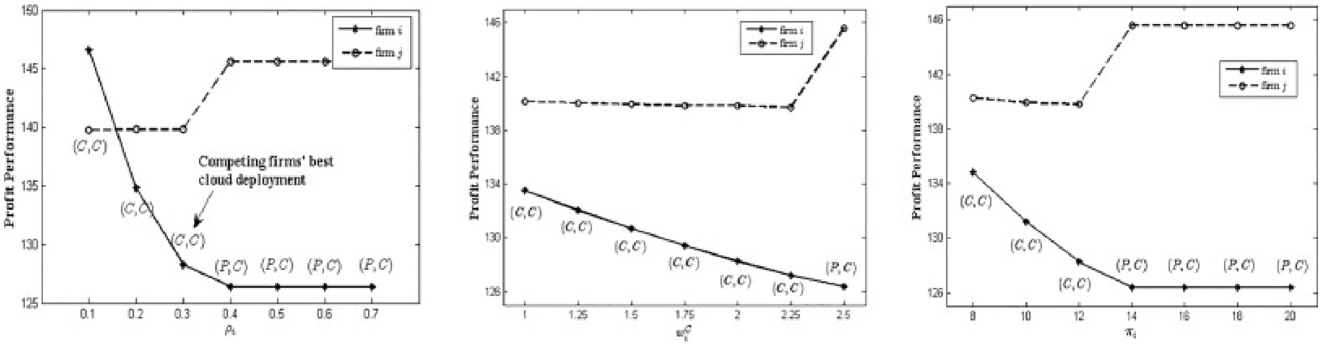
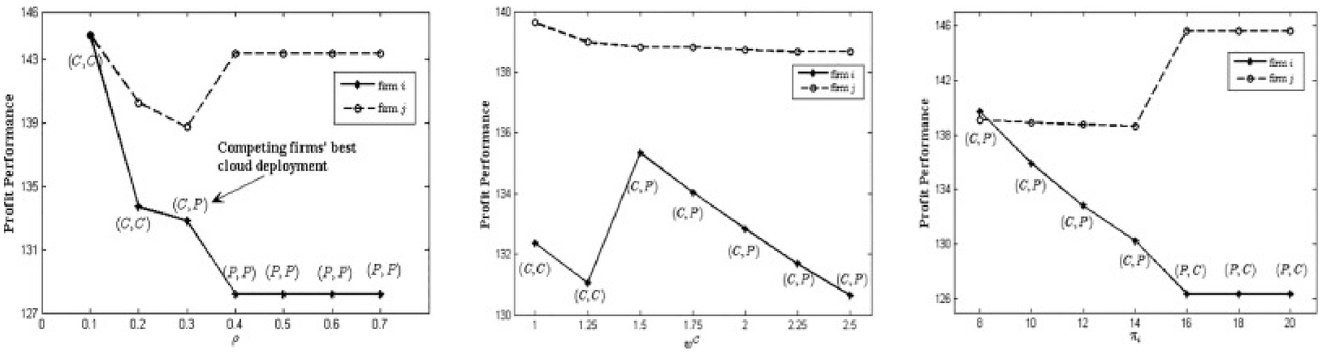


Fig. 5. Sensitivity analysis under competition in which two firms burst as needed into separate public cloud providers.

Move to the competition systems. For the system in which firm *i* is competing against firm *j* and the two firms burst as needed into separate clouds, firm *i* responds as it did under monopoly — by shifting to cloudbursting when the three parameters decline (Fig. 5) and such a shift enhances firm *i*'s profit. Simultaneously, its competitor *j* loses profits due to its transition from a private cloud to cloudbursting. However, when these parameters are sufficiently small and firm *i* has maintained cloudbursting, the further reduction in these parameters interestingly benefits not just firm *i*'s profit but also the competing firm's profit.

For the systems in which two firms burst as needed into a common public cloud provider, their best cloud choices with respect to parameters are complex. When the risk level is sufficiently large ($\rho \geq 0.4$), both firms remain in a private cloud. However, the decreased risk level incentivizes both firms to migrate to cloudbursting. When risk drops to 0.3, one competing firm takes the lead by migrating to cloudbursting. Once such risk becomes sufficiently small ($\rho \leq 0.2$), both competing firms take advantage of the scalability by deploying cloudbursting. Further, we find that a firm's leading migration to cloudbursting due to the low risk level improves its profit. This migration initially hurts its competitor's profit but gradually benefits its profit once risk declines more. The sensitivity analysis with respect to the marginal cost of external resources shows that, albeit low risk, a firm does not necessarily adopt cloudbursting. This result is different from the system where firms excess external resources as needed from different public cloud providers. A sufficiently small value (e.g., $w^C = 1$ or 1.25) on the marginal cost of external resources incentivizes both competing firms to choose cloudbursting. The increase on this parameter (e.g., w^C from 1.5 to 2.5) results in one competing firm favoring a private cloud and another competing firm to choose cloudbursting. Such different cloud choices between firms are also seen when the unit penalty cost becomes a concern for firms (i.e., π_i from 8 to 12 and $\pi_j = 12$) even though risk is low. When a firm's penalty cost becomes too high, which is not just a concern but a threat to cloudbursting, it then stays in a private cloud and its competing firm with the relatively low unit penalty cost deploys cloudbursting (i.e., π_i from 16 to 20). For both the marginal cost of external resources and the unit penalty cost, our results reveal that a competing firm's migration to cloudbursting benefits its profit but simultaneously hurts its competitor's profit.

We now sum up our numerical findings under two competition systems. Besides the parameter setting used to show the graphical outcomes in Figs. 5 and 6, we have examined a wide range of parameter combinations. From our results, we present two observations. First, a firm's cloud choice heavily depends on three parameters related to external resources: the level of risk, the marginal cost of external resources, and the unit penalty cost. When competing firms' external resources as needed are from different public clouds, then a firm deploys cloudbursting when its data breach risk is considerably low and this firm maintains a private cloud when risk is considerably high. When competing firms' access of external resources as needed is from the same public cloud, two firms under low risk choose different deployments (one for cloudbursting, the other private) when either the marginal cost of external resources or the unit penalty cost is considerably high. However when the three parameters related to external resources are sufficiently small, two competing firms accessing external resources as needed from the same public cloud provider both deploy cloudbursting. Second, a competing firm's migration to cloudbursting improves its profit performance and simultaneously hurts its competitor's profit.



Download : Download high-res image (235KB) Download : Download full-size image

Fig. 6. Sensitivity analysis under competition in which two firms burst as needed into a common public cloud provider.

6. Conclusion

Cloudbursting, a hybrid model used to burst into a pubic cloud as needed, provides the scalability and cost-effectiveness for an application service firm with in-house capacity to run its application. Thus, cloudbursting has drawn the attention of both practitioners and scholars in today's lightning-fast, highly competitive markets. This work examines the tradeoffs between two cloud types, private and cloudbursting, for a profit-maximizing application service firm.

We develop models for this firm based on three scenarios: a monopoly, and two types of duopoly, one in which firms burst as needed into separate public cloud providers, and the other duopoly in which firms burst as needed into a common public cloud provider. In each of system, we identify thresholds to illustrate whether a firm should remain in a private cloud or migrate to cloudbursting. The analysis under monopoly shows that a firm's best cloud choice is to remain in a private cloud if risk from leased external resources is considerably high while it prefers cloudbursting if risk is considerably low. Our study further extends the contribution to the competition systems. In the system of competing firms who access external resources as needed from separate public cloud providers, a competing firm prefers cloudbursting when its risk is considerably low. However, the desirability of cloudbursting fades with considerably high risk so that a competing firm prefers to maintain a private cloud.

Nonetheless, in the system of competing firms who access external resources as needed from a common public cloud provider, this firm has the likelihood to maintain a private cloud even though its risk remains low. Two crucial reasons explain this deployment choice. First, competing firms always encounter risk simultaneously and thus one firm might consider taking advantage of the risk-free private cloud. Second, besides the level of risk, the proportion of demand interchangeability plays another decisive role to impact a competing firm's cloud choice. Found in the special case and the numerical study, the attractiveness of this proportion is evaluated by parameters related to external resources: the level of risk, the marginal cost of external resources and the unit penalty cost. We interestingly find that competing firms with low security risk favor cloudbursting if the attractiveness of demand interchangeability is high (e.g., a sufficiently low marginal cost of external resources or a sufficiently low unit penalty cost). Nonetheless, when the attractiveness of demand interchangeability is low, a competing firm stays in a private cloud and the other adopts cloudbursting even though risk remains low. This observation is consistent with the analytical finding where two competing firms under low risk have the likelihood of choosing different cloud deployments when they burst as needed into the same public cloud for external resources.

The numerical study validates the analytical results and simultaneously provides more insights into firms' best cloud choices. More importantly, we provide a visual understanding of the linkage between a firm's (or competing firms') best choice of cloud and its (their) corresponding profit performance. When three examined parameters related to external resources are perceived to diminish, a firm in both non-competitive and competitive systems tends to migrate to cloudbursting. Further, a competing firm's migration to cloudbursting improves its profit performance and simultaneously hurts its competitor's profit.

To summarize, this work pioneers a unique method of analyzing a firm's optimal cloud deployment in both non-competitive and competitive systems, which benefits a cloud-based company facing a crucial decision: to maintain a private cloud or migrate to cloudbursting. To do so, we evaluate a firm's profit function. In essence, we have opened a door to a new analytic corridor that we hope others will explore.

Appendix. Supplementary materials

Download : **Download Word document (70KB)**

References

Agrawal and Seshadri, 2000 V. Agrawal, S. Seshadri

Impact of uncertainty and risk aversion on price and order quantity in the newsvendor problem
Manufacturing & Service Operations Management, 2 (4) (2000), pp. 410-423
View Record in Scopus Google Scholar

Alani, 2014 M.M. Alani
Securing the cloud: Threats, attacks and mitigation techniques
Journal of Advanced Computer Science and Technology,, 3 (2) (2014), pp. 202-213
CrossRef View Record in Scopus Google Scholar

Anselmi et al., 2014 J. Anselmi, D. Ardagna, M. Passacantando
Generalized Nash equilibria for SaaS/PaaS clouds
European Journal of Operational Research, 236 (1) (2014), pp. 326-339
Article Download PDF View Record in Scopus Google Scholar

Ardagna et al., 2013 D. Ardagna, B. Panicucci, M. Passacantando
Generalized Nash equilibria for the service provisioning problem in cloud systems
IEEE Transactions on Services Computing, 6 (4) (2013), pp. 429-442
View Record in Scopus Google Scholar

Armbrust et al., 2010 M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R.H. Katz, A. Konwinski, M. Zaharia
A View of Cloud Computing
Communications of the ACM, 53 (4) (2010), pp. 50-58
CrossRef View Record in Scopus Google Scholar

August et al., 2014 T. August, M.F. Niculescu, Shin H.
Cloud implications on software network Structure and security risks
Information Systems Research,, 25 (3) (2014), pp. 489-510
CrossRef View Record in Scopus Google Scholar

Babcock, 2014 Babcock, C. (2014). 9 worst cloud security threats, Accessed March 03, 2014, Retrieved from http://www.informationweek.com/cloud/infrastructure-as-a-service/9-worst-cloud-security-threats/d/d-id/1114085.
Google Scholar

Benlian and Hess, 2011 A. Benlian, T. Hess
Opportunities and risks of software-as-a-service: Findings from a survey of IT executives
Decision Support Systems, 52 (2011), pp. 232-246
Article Download PDF View Record in Scopus Google Scholar

Bernard et al., 2014 Bernard, T.S., Hsu, T., Perlroth, N., & Lieber, R. (2017). Equifax says cyberattack may have affected 143 million in the U.S., Retrieved from https://www.nytimes.com/2017/09/07/business/equifax-cyberattack.html.
Google Scholar

Bigi et al., 2013 G. Bigi, M. Castellani, M. Pappalardo, M. Passacantando
Existence and solution methods for equilibria

European Journal of Operational Research, 227 (1) (2013), pp. 1-11
Article Download PDF View Record in Scopus Google Scholar

Brito and de Almeida, 2012 A.J. Brito, A.T. de Almeida
Modeling a multi-attribute utility newsvendor with partial backlogging
European Journal of Operational Research, 220 (3) (2012), pp. 820-830
Article Download PDF View Record in Scopus Google Scholar

Cavazzuti et al., 2002 E. Cavazzuti, M. Pappalardo, M. Passacantando
Nash equilibria, variational inequalities, and dynamical systems
Journal of Optimization Theory and Applications, 114 (3) (2002), pp. 491-506
View Record in Scopus Google Scholar

Chen and Wu, 2013 Chen P.Y., Wu S.Y.
The impact and implications of on-demand services on market structure
Information Systems Research, 24 (3) (2013), pp. 750-767
CrossRef View Record in Scopus Google Scholar

Demirkan and Cheng, 2008 H. Demirkan, Cheng H.K.
The risk and information sharing of application services supply chain
European Journal of Operational Research, 187 (3) (2008), pp. 765-784
Article Download PDF View Record in Scopus Google Scholar

Demirkan et al., 2010 H. Demirkan, Cheng H.K., S. Bandyopadhyay
Coordination strategies in an SaaS supply chain
Journal of Management Information Systems, 26 (4) (2010), pp. 119-143
View Record in Scopus Google Scholar

Ding et al., 2017 Ding S., Xia C., Wang C., Wu D., Zhang Y.
Multi-objective optimization based ranking prediction for cloud service recommendation
Decision Support Systems, 101 (2017), pp. 106-114
Article Download PDF View Record in Scopus Google Scholar

Dorsch and Häckel, 2014 D. Dorsch, B. Häckel
Combining models of capacity supply to handle volatile demand: The economic impact of surplus capacity in cloud service environments
Decision Support Systems, 58 (2014), pp. 3-14
View Record in Scopus Google Scholar

Facchinei and Kanzow, 2010 F. Facchinei, C. Kanzow
Generalized Nash equilibrium problems
Annals of Operations Research, 175 (1) (2010), pp. 177-211
CrossRef View Record in Scopus Google Scholar

Furuncu and Sogukpinar, 2015 E. Furuncu, I. Sogukpinar
Scalable risk assessment method for cloud computing using game theory (CCRAM)
Computer Standards & Interfaces, 38 (2015), pp. 44-50
Article Download PDF View Record in Scopus Google Scholar

Godfrey and Zulkernine, 2014 M. Godfrey, M. Zulkernine
Preventing cache-based side-channel attacks in a cloud environment
IEEE Transactions on Cloud Computing,, 2 (4) (2014), pp. 395-408
View Record in Scopus Google Scholar

Iyer and Henderson, 2012 B. Iyer, J.C. Henderson
Business value from clouds: Learning from users
MIS Quarterly Executive, 11 (1) (2012), pp. 51-60
CrossRef View Record in Scopus Google Scholar

Keskin and Taskin, 2015 T. Keskin, N. Taskin
Strategic pricing of horizontally differentiated services with switching costs: A pricing model for cloud computing
International Journal of Electronic Commerce, 19 (3) (2015), pp. 34-53
CrossRef View Record in Scopus Google Scholar

Khouja, 1996 M. Khouja
A note on the newsboy problem with an emergency supply option
The Journal of the Operational Research Society, 47 (12) (1996), pp. 1530-1534
View Record in Scopus Google Scholar

Lang et al., 2018 Lang M., M. Wiesche, H. Krcmar
Criteria for selecting cloud service providers: A Delphi study of quality-of-service attributes
Information and Management, 55 (2018), pp. 746-758
Article Download PDF View Record in Scopus Google Scholar

Lee and Lodree, 2010 Lee H., E.J. Lodree
Modeling customer impatience in a newsboy problem with time-sensitive shortages
European Journal of Operational Research, 205 (3) (2010), pp. 595-603
Article Download PDF View Record in Scopus Google Scholar

Leong, 2013 Leong, L. (2013). VMware joins the cloud wars with vCloud Hybrid Service, accessed May 21, 2013, Retrieved from http://blogs.gartner.com/lydia_leong/2013/05/21/vmware-joins-the-cloud-wars-with-vcloud-hybrid-service/.
Google Scholar

Levitin et al., 2018 Levitin G., Xing L., Dai Y.
Co-residence based data vulnerability vs. security in cloud computing system with random server assignment
European Journal of Operational Research, 267 (2) (2018), pp. 676-686
Article Download PDF View Record in Scopus Google Scholar

Lippman and McCardle, 1997 S.A. Lippman, K.F. McCardle
The competitive newsboy
Operations Research, 45 (1) (1997), pp. 54-65
CrossRef View Record in Scopus Google Scholar

Lodree, 2007 E.J. Lodree
Advanced supply chain planning with mixtures of backorders, lost sales, and lost contract
European Journal of Operational Research, 181 (1) (2007), pp. 168-183
Article Download PDF View Record in Scopus Google Scholar

Marston et al., 2011 S. Marston, Li Z., S. Bandyopadhyay, Zhang J., A. Ghalsasi
Cloud computing—the business perspective
Decision Support Systems, 51 (1) (2011), pp. 176-189
Article Download PDF View Record in Scopus Google Scholar

Martens and Teuteberg, 2012 B. Martens, F. Teuteberg
Decision-making in cloud computing environments: A cost and risk based approach
Information Systems Frontiers, 14 (4) (2012), pp. 871-893
CrossRef View Record in Scopus Google Scholar

Mattess et al., 2011 M. Mattess, C. Vecchiola, S.K. Garg, R. Buyya
Cloud bursting: Managing peak loads by leasing public cloud services
Cloud Computing: Methodology, Systems, and Applications, CRC Press (2011)
Google Scholar

Microsoft Corp 2014 Microsoft Corp. (2014). Dell uses cloud solution to manage online traffic, Microsoft Corp. customer deals during the 2013 holiday season, accessed February 20, 2014. Retrieved from https://www.microsoft.com/en-us/download/details.aspx?id=53126.
Google Scholar

Naldi and Mastroeni, 2016 M. Naldi, L. Mastroeni
Economic decision criteria for the migration to cloud storage
European Journal of Information Systems, 25 (1) (2016), pp. 16-28
CrossRef View Record in Scopus Google Scholar

Nalinipriya et al., 2016 G. Nalinipriya, P.J. Varalakshmi, K.G. Maheswari, R. Anita
An extensive survey on co-resident attack in dynamic cloud computing environment
International Journal of Applied Engineering Research,, 11 (5) (2016), pp. 3019-3023
View Record in Scopus Google Scholar

Nicolaou et al., 2012 C.A. Nicolaou, A.I. Nocolaou, G.D. Nocolaou
Auditing in the cloud: Challenges and opportunities
The CPA Journal, 82 (1) (2012), pp. 66-70
View Record in Scopus Google Scholar

Oliveira et al., 2014 T. Oliveira, M. Thomas, M. Espadanal
Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors
Information and Management, 51 (5) (2014), pp. 497-510
Article Download PDF View Record in Scopus Google Scholar

Pardeshi, 2014 V.H. Pardeshi

Cloud computing for higher education institutes: Architecture, strategy and recommendations for effective adaptation
Procedia Economics and Finance, 11 (2014), pp. 589-599
Article Download PDF View Record in Scopus Google Scholar

Püschel et al., 2015 T. Püschel, G. Schryen, D. Hristova, D. Neumann
Revenue management for cloud computing providers: Decision models for service admission control under non-probabilistic uncertainty
European Journal of Operational Research, 244 (2) (2015), pp. 637-647
Article Download PDF View Record in Scopus Google Scholar

Qin et al., 2011 Qin Y., Wang R., A.J. Vakharia, Chen Y., M.M.H. Seref
The newsvendor problem: Review and directions for future research
European Journal of Operational Research, 213 (2) (2011), pp. 361-374
Article Download PDF View Record in Scopus Google Scholar

Ren et al., 2012 Ren K., Wang C., Wang Q.
Security challenges for the public cloud
IEEE Internet Computing,, 16 (1) (2012), pp. 69-73
CrossRef View Record in Scopus Google Scholar

Schwarz et al., 2009 A. Schwarz, B. Jayatilaka, R. Hirschheim, T. Goles
A conjoint approach to understanding IT application services outsourcing
Journal of the Association for Information Systems, 10 (10) (2009), pp. 748-781
View Record in Scopus Google Scholar

Silver et al., 1998 E.A. Silver, D.F. Pyke, R. Peterson
Inventory management and production planning and scheduling
(3rd ed.), John Wiley & Sons, New York (1998)
Google Scholar

Sultan, 2011 N. Sultan
Reaching for the 'cloud': How SMEs can manage
International Journal of Information Management, 31 (3) (2011), pp. 272-278
Article Download PDF View Record in Scopus Google Scholar

Tomlin, 2006 B.T. Tomlin
On the value of mitigation and contingency strategies for managing supply chain disruption risks
Management Science, 52 (5) (2006), pp. 639-657
CrossRef View Record in Scopus Google Scholar

Tomlin, 2009a B.T. Tomlin
Disruption-management strategies for short life-cycle products
Naval Research Logistics, 56 (4) (2009), pp. 318-347
CrossRef View Record in Scopus Google Scholar

Tomlin, 2009b B.T. Tomlin

Impact of supply learning when suppliers are unreliable
Manufacturing and Service Operations Management, 11 (2) (2009), pp. 192-209
CrossRef View Record in Scopus Google Scholar

Vecchiola et al., 2009 C. Vecchiola, Chu X., R. Buyya
Aneka: A software platform for .NET-based cloud computing
High Speed and Large Scale Scientific Computing, 18 (2009), pp. 267-295
View Record in Scopus Google Scholar

Venters and Whitley, 2012 W. Venters, E. Whitley
A critical review of cloud computing: Researching desires and realities
Journal of Information Technology, 27 (2012), pp. 179-197
CrossRef View Record in Scopus Google Scholar

Vipin and Amit, 2017 B. Vipin, R.K. Amit
Loss aversion and rationality in the newsvendor problem under recourse option
European Journal of Operational Research, 261 (2) (2017), pp. 563-571
Article Download PDF View Record in Scopus Google Scholar

Vithayathil, 2018 J. Vithayathil
Will cloud computing make the information technology (IT) department obsolete?
Information Systems Journal, 28 (4) (2018), pp. 634-649
CrossRef View Record in Scopus Google Scholar

Walterbusch et al., 2013 M. Walterbusch, B. Martens, F. Teuteberg
Evaluating cloud computing services from a total cost of ownership perspective
Management Research Review, 36 (6) (2013), pp. 613-638
CrossRef View Record in Scopus Google Scholar

Wang et al., 2010 Wang Y., W. Gilland, B. Tomlin
Mitigating supply risk: Dual sourcing or process improvement?
Manufacturing & Service Operations Management, 12 (3) (2010), pp. 489-510
CrossRef View Record in Scopus Google Scholar

Winkler and Brown, 2013 T.J. Winkler, C.V. Brown
Horizontal allocation of decision rights for on-premise applications and software-as-a-service
Journal of Management Information Systems, 30 (3) (2013), pp. 13-48
CrossRef Google Scholar

View Abstract