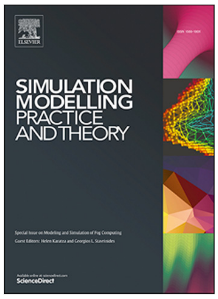
Journal Pre-proof



Exploring cost-efficient bundling in a multi-cloud environment

Chatzithanasis Georgios, Filiopoulou Evangelia, Michalakelis Christos, Nikolaidou Maria

PII: S1569-190X(21)00054-X
DOI: https://doi.org/10.1016/j.simpat.2021.102338
Reference: SIMPAT 102338

To appear in: *Simulation Modelling Practice and Theory*

Please cite this article as: C. Georgios, F. Evangelia, M. Christos et al., Exploring cost-efficient bundling in a multi-cloud environment, *Simulation Modelling Practice and Theory* (2021), doi: https://doi.org/10.1016/j.simpat.2021.102338.

Exploring Cost-Efficient Bundling in a Multi-Cloud Environment

Chatzithanasis Georgios[a], Filiopoulou Evangelia[b], Michalakelis Christos[c], Nikolaidou Maria[d]

Harokopion University of Athens — Department of Informatics and Telematics

[a]*geo.hatz@hua.gr*
[b]*evangelf@hua.gr*
[c]*michalak@hua.gr*
[d]*mara@hua.gr*

Abstract

When exploring hybrid cloud solutions, the provision of cloud computing resources by more than one provider, may result in important gains, such as better risk management, improvement of operational efficiency and avoidance of the lock-in effect by a single vendor. To this end, many enterprises decide to rely on hybrid strategies and multiple cloud providers. With the strategic distribution of the cloud management and resources between multiple providers, multi-cloud adoption can achieve greater efficiency, taking advantage of the economies of scale, and establish an optimal performance/cost balance of the multi-cloud service. Into this context, the paper explores the pricing strategies of 23 different providers for IaaS and estimates the efficiency of different multi-provider service bundles, utilizing DEA methodology. The results are quite encouraging, since the efficiency of multi-cloud solutions is increasing, while cost savings may be accomplished. They provide a measurable driver for clients exploring hybrid solutions of how to mix different IaaS services of different providers.

Preprint submitted to Simulation Modelling Practice and Theory *April 20, 2021*

Keywords: Hybrid Cloud Management, Cloud Pricing, Multicloud, DEA, Cloud Efficiency

1. Introduction

A hybrid cloud is defined as the combination of a public and a private cloud, using a mix of on-premises private cloud and third-party public cloud services, with appropriate orchestration between the two platforms [1]. Applications in a hybrid cloud implementation have some of their components in a private cloud, as for example storage and their database, while the application server can operate on the public cloud. A hybrid cloud offers increased availability and risk management and can act as the "fallback plan" in the cases that demand increases and the private cloud fails to cope. Hybrid cloud offers flexibility but also raises questions about how the public cloud implementation will be handled. Therefore, for the public cloud part of a hybrid approach multi-cloud solutions should be evaluated.

The multi-cloud environment is considered as the use of multiple cloud computing and storage services in a single heterogeneous architecture. It refers to a mixture of Infrastructure, Platform or Software as a Service (IaaS, PaaS, or SaaS) [1]. Most organizations already operate in a multi-cloud environment and most of the ones that do not currently use multi-cloud, plan to use multiple hybrid clouds [2]. According to a Gartner survey [2], the multi-cloud IT environment has already dominated the technology market and users use higher- quality services, thus optimizing service cost without compromising performance.

A multi-cloud environment is quite flexible and quickly adopts the optimal technologies for any task and by incorporating multiple clouds into a company's IT strategy, business requirements are lined up with the best cloud-hosting provider for each individual task. In addition, a competitive market is formed, offering optimal pricing for different resource capacities. Businesses compare different cloud solutions and secure the best available rates, based on their specific IT needs. Furthermore, avoiding lock-in by a single vendor is a core business requirement, or a way to achieve greater portability for their applications, together with risk management. With the achieved portability, an organization can easily move applications to another framework or platform. By carefully evaluating the expectations and potential pitfalls and having a strong negotiating position, which makes it easy to switch from one cloud provider to another, enterprises can exploit the cloud advantages and get the optimal multi-cloud service in a reasonable price [3, 4].

The most compelling reason for adopting multi-cloud services can be cost saving and efficiency is the key to this. Following the above considerations, a multi-cloud approach is proposed, which estimates the efficiency of multi-cloud services and highlights the efficient and reliable multi-cloud services at a better price. Based on efficiency, users can choose not only the optimal but also the most cost-efficient multi-cloud solution. Multi-cloud Infrastructure - as - a - Service (IaaS) solutions are composed by combining efficient single IaaS services and their efficiency is estimated by adopting Data Envelopment Analysis (DEA).

The rest of the paper is organized as follows: Section II presents the related work, while Section III describes the motivation of the proposed strategy. Section IV introduces a general description of the DEA method and thoroughly defines the four phases of the proposed methodology. Section V presents the evaluation results and finally Section VI concludes the paper and suggests future work.

2. Literature Review

In this section some of the key aspects of this publication are brought to light. From the multi-cloud approach, to the bundling of services and, finally, to the multi-attribute decision-making problem, which focuses on the calculation of the efficiency of IaaS cloud services [2].

2.1. A Multi-Cloud Approach

Cloud Computing has become a major technology that shapes the future of applications' deployment and the IT infrastructure enterprises implement. The rapidly changing market environment and competition is a key factor to the cloud computing adoption and diffusion. A scalable single-cloud approach seems the most alluring option for many enterprise users, as it is the easiest way for billing and supporting a company's cloud infrastructure. However, the concept of a single-cloud approach can result to a failure due to service unavailability [3]. The multi-cloud approach, that can solve such issues comes with its own drawbacks and among them is the security offered by each cloud provider and the inter-operability [4].

A number of relevant studies have been conducted and the results are quite promising. Among them, the work performed in [5] indicated that secure and reliable storage could be achieved, harvesting the multi-cloud potential. The approach improves the perceived data availability and in most of the cases it results in a significant reduction of latency. Other studies have examined the security of the approach and raised concerns about the security [6], while proposing interesting solutions for this problem. The common result of the studies is that the multi-cloud approach is continuously gaining ground and thus needs further examination regarding its merits and drawbacks.

2.2. Bundling of Services

Bundling is not a recently adopted strategy, since many economic sectors already use it for their operational needs. According to [7] the term tie-in sale is the primary form of a bundle definition from the early 60s. The concept of bundling has been found in academic research and the bundling of services and products always played a vital role in companies and their business strategies. The ultimate purpose is the cost efficiency and customer satisfaction [8].

As described in [9], bundling in an oligopolistic environment can be a strategy that can discourage new competitors to enter a market. This is because the combination of products or services becomes an obstacle that a rival company has to face and overcome with great impact in its resources.

A characteristic example that highlights the value of bundling in technology is the telecommunications industry [10]. Mobile value added services have become a very important factor to generate income and attract new customers and also to sustain that has already gained. Services like: SMS, MMS and data access in the past were considered value-added services. As technology advances, value added services have changed and more recent examples are: live streaming, gaming, online banking. As far as cloud computing is concerned, companies that operate in the sector -like Flexiant- have concluded that pre-configured bundles of infrastructure and services facilitate decision for purchasing a cloud solution [11]. A properly implemented bundling is also the key aspect for companies that aspire to differentiate from their competitors in the market.

2.3. Efficiency and Cost

Considering the efficiency and cost, the work performed in [12], based on the experimental testing and modeling, tried to evaluate the requirements in resources with cost and performance being the key factors.

In[13] a DEA (Data Envelopment Analysis) methodology was applied as an approach to measure the effective resource performance of cloud computing infrastructures, evaluating IaaS. DEA is a very popular methodology that has the ability to benchmark multiple inputs and outputs as well as computational ease. This is mainly because DEA has the ability to express the ecosystem under question as a linear problem.

In addition, a non parametric method that evaluates relative efficiency based on DEA was conducted by [14]. It ranked cloud services into different efficiency levels and indicated solutions towards performance improvement. The evaluation of the performance was based on a small group of IaaS services.

In [15] the authors calculated the IaaS efficiency, adopting DEA method. The proposed DEA model was described using functional and non-functional attributes and the impact of the corresponding attributes on the efficiency was examined. IaaS packages combing non-functional attributes promoted efficiency and profit margin.

3. Motivation

The primary objective of this work is to explore whether a multi-cloud approach is an efficient and cost effective way for enterprises willing to adopt cloud computing solutions. The multi-cloud approach should either a part of a hybrid cloud (private and public), though it could also act as a standalone strategy. [4].

According to [16] the global Infrastructure as a Service (IaaS) market grew 37.3% in 2019 to total $44.5 billion, up from $32.4 billion in 2018. Amazon retained the first position in the IaaS market in 2019, followed by Microsoft, Alibaba, Google and Tencent. The above research concluded that the IaaS market is very competitive and subject to constant alterations. The abundance of IaaS providers gives the option not only to compare prices and services between them, but also offer the ability to combine them into a cost effective bundle. Coronavirus pandemic has also enhanced cloud usage and investing, compelling enterprises to move their applications to the public cloud as a result of the COVID-19. This resulted to the realization of potential of public cloud and the broadening of the horizons of many enterprises [16]. For these reasons, multi-cloud model is expected to be further adopted and therefore a methodology to find the most cost effective is imminent.

With the large variety of providers, the research performed in this paper evaluates different combinations of multi-cloud implementations in the form of bundled IaaS. The purpose is to create an accurate estimation about efficiency and possible cost reduction. The results will give researchers and enterprises tangible conclusions based on real data, in order to use them in their estimations and multi-cloud implementations.

Following the work presented in [15], we propose the adoption of DEA methodology to explore the efficiency of multi-cloud solutions, since it has proven the most accurate for decision-making in the world of technology [2]. Multi-cloud hybrid strategy provides remarkable agility and cost efficiency, therefore a model based on Data Envelopment Analysis can point out the contribution of multi-cloud strategy to the efficiency of IaaS services.

Compared to the work presented in [15], the methodology proposed to explore multi-cloud efficiency differs. Both are based in DEA, thought the combination of input data and the methodology steps proposed are different. Though the work presented in [15] emphasized non-functional characteristics, the evaluation of IaaS packages explored in the current study is based on the essential characteristics of processing, memory and storage costs, as the purpose of the proposed model is to create a data set with the combination of single IaaS packages into multi-cloud bundles. Besides the differences on the input data set, our methodology was extended with additional phases to create multi-cloud bundles based on single IaaS packages rated high in efficiency and explore multi Iaas package combinations. In contrast to [15], DEA is applied recurrently to identify multi-cloud efficient bundles.

The aim is to provide results that propose value-for-money bundles which can help enterprises to use the right options, thus enjoying financial savings and benefit by an uninterrupted, scalable and an efficient IT infrastructure.

4. Methodology

The examination of the efficiency of a single provider of a service is quite interesting but the proposed consideration in terms of the efficiency of multi-cloud services is further challenging and demanding. Towards this direction, the proposed strategy composes multi-cloud services and estimates the relative efficiency, based on a DEA approach. A detailed description of the adopted methodology is presented in the following sections.

4.1. Data Envelopment Analysis

Data Envelopment Analysis (DEA) is a methodology for performance evaluation and a benchmarking technique which measures the relative efficiency of a set of decision-making units (DMUs). This technique is non-parametric, since it is entirely based on the input and output data. It is a significant tool for assessing the comparative efficiencies of decision-making units (DMUs), especially when the presence of multiple inputs and outputs makes comparison with other methods a complicated procedure.

Data Envelopment Analysis (DEA), a linear based multi-criteria decision making strategy, introduces a process for calculating relative efficiency, incorporating multiple inputs and outputs and without specifying the weights for the chosen inputs and outputs [17]. Each unit (DMU) is compared with only the "optimal" units and the most efficient units are highlighted.[18]

The calculation of relative efficiency where multiple inputs and outputs exist was addressed by Farell [19], based on the definition of a hypothetical efficient unit, as a weighted average of efficient units, to act as a comparator for an inefficient unit. A typical measure for relative efficiency is presented in Equation (1) for a DMU j with m outputs and n inputs.

$$efficiency = \frac{weightedsumoutputs}{weightedsumofinputs} \tag{1}$$

Which can be also written as presented in Equation (2) for a DMU with m outputs and n inputs.

$$\text{Efficiency of DMU } j = \frac{u_1 y_{1j} + u_2 y_{2j} + \ldots + u_m y_{mj}}{v_1 x_{1j} + v_2 x_{2j} + \ldots + v_n x_{nj}} \quad (2)$$

Where u_r is the weight given to output r, v_i is the weight given to input i, y_{rj} is the amount of output r from DMU j and x_{ij} is the amount of input i in DMU j.

According to Equation (2) the calculation of the relative efficiency of a DMU with multiple inputs and outputs requires a common set of weights. However, DMUs may consider their inputs and outputs in a different way and the application of an agreed common set of weights is difficult. Thus, the measure of efficiency based on the assumption that a single common set of weight is required is unsatisfactory. However, DEA values inputs and outputs differently and lets each unit adopt a set of weights that show it in the most favourable light in comparison with the other units. [2] [20].

Efficiency h_0 of a specific target DMU j_0 can be obtained as a solution to the system of Equations M1. Maximise the efficiency of unit j0, subject to the efficiency of all units being less or equal to 1.

$$\left.\begin{array}{l} \text{Max } h_0 = \frac{\sum_r u_r y_{rj_0}}{\sum_i v_i x_{ij_0}} \\ \text{subject to:} \\ \frac{\sum_r u_r y_{rj}}{\sum_i v_i x_{ij}} \leq 1 \text{ for each unit j} \\ u_r, v_i \geq \epsilon \text{ r=1,2,} \ldots, m \; i = 1, 2, \ldots, n \end{array}\right\} \text{M1}$$

The variables of the above equation M1 are the weights (u,v) which are the most favorable to unit j_0 as, meaning that DEA determines the weights compared to the other k-1 DMUs, to calculate the efficiency h_0 based on the assumption that more of outputs and less of inputs are desirable, so so that DMU j_0 seems as efficient as possible. Each unit is allowed to adopt a set of weights in the most favorable light in comparison to the other units. The weights (u, v) are constrained to be greater than, or equal to, some small positive quantity E, in order to avoid any input or output being totally ignored in determining the efficiency [18] [20]. The relative efficiency of each DMU cannot be more than 100% efficient when the same weights are adopted by each DMU, meaning that the efficiency is constrained to be lower than or equal to 1.

Model M1 is a fractional linear program and it first needs to be converted into a linear form, so that the methods of linear programming can be applied. The linearization process is relatively straightforward and the linear version of the constraints of M1 is shown in equation M2 [20] [21]. The relative efficiency of the target unit can be obtained by solving equation M2 .

$$\left.\begin{array}{l} \text{Max } h_0 = \sum_r u_r y_{rj_0} \\ \text{subject to:} \\ \sum_i v_i x_{ij_0} = 1 \\ \sum_i u_r y_{rj} - \sum_i v_i x_{ij} \leq 0 \text{ (for each DMU) } j = 1, 2, \ldots, k \\ u_r, v_i \geq \epsilon \; r = 1, 2, \ldots, m \; i = 1, 2, \ldots, n \end{array}\right\} \text{M2}$$

The proposed simulation model is based on the aforementioned model of DEA and the corresponding DMUs, inputs and outputs that compose the DEA model are considered to be the parameters of the simulation. In addition, Maxdea software was adopted for the simulation of the proposed methodology [22]. It is an easy-to-use but powerful and professional tool and has been widely used to measure performance in many research areas [23] [24].

4.2. Multi-Cloud Efficiency using DEA

The proposed model aims to calculate the relative efficiency rates of multi-cloud services. Towards this direction the current methodology creates a multi-cloud data set based on the most efficient single provider bundles that are described by their functional features (cpu, memory,storage) and their price.

The multi-cloud services are categorized into three groups:

- **Computation Optimized Instances (COI):** This category includes bundles ideal for compute bound applications, which benefit from high performance processors. Instances belonging to this category are well suited for batch processing workloads, media transcoding, high performance web servers, high performance computing (HPC) and as server engines.
- **Memory Optimized Instances (MOI):** It includes instances which are designed to deliver fast performance for workloads that process large data sets in memory. Instances of this group can be used as gaming servers.
- **Storage Optimized Instances (SOI):** SOI Refers to instances designed for workloads that require high, sequential read and write access to very large data sets on local storage. They are optimized to deliver

tens of thousands of low-latency, random I/O operations per second (IOPS) to applications [25].

The above categories are further divided into 3 subcategories, based on the size of their main feature [25]. For example, computation optimized bundles are classified into Small(S), Medium(M) and Large(L) bundles, based on the number of vCPUs. Memory and storage bundles are also sub-categorized, based on the size of the memory and storage capacity.

The proposed simulation model is based on the following phases:

Phase I:

- Collection of IaaS bundles
- Categorization of IaaS bundles into categories based on computation, memory and storage characteristics (COI, MOI,SOI).
- Each category is divided into three (3) subcategories (S/M/L COI, S/M,L S/M/L SOI, S/M/L MOI), based on the size of their main feature (CPU, RAM, storage). For example S-COI refers to small compute optimized instances.

Phase II:

- DEA method is applied to each subcategory (S/M/L COI, S/M/L S/M/L SOI, S/M/L MOI). The relative efficiency is estimated and the IaaS bundles are ranked by the efficiency score. Then, the five most efficient bundles are chosen. The chosen bundles hold the highest efficiency scores and are mixed for the composition of multi-cloud IaaS services.

Phase III:

- COI (S/M/L), MOI (S/M/L) and SOI (S/M/L) multi-cloud groups are composed based on the five bundles with the highest efficiency scores.

Phase IV:

- DEA is applied to each multi-cloud subcategory (S/M/L COI, S/M/L MOI and S/M/L SOI).

Figure 1 sums up the phases of the methodology. Each phase is presented in details, in the following sections.

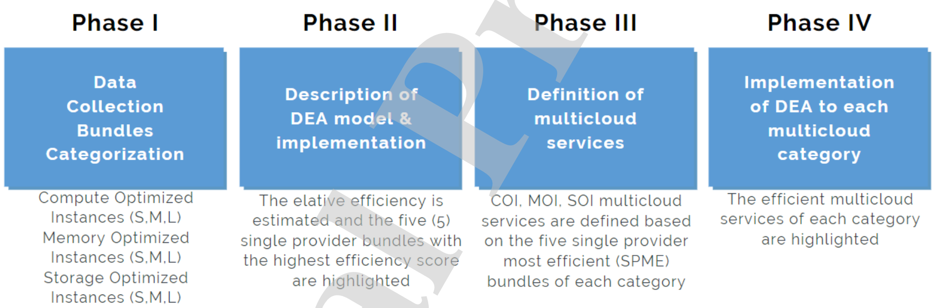


Figure 1: Phases of the proposed DEA Methodology

4.2.1. Phase I - Data Collection

The proposed methodology is based on IaaS cloud services. Data were collected from the Cloudorado database [26], a price comparison service of cloud computing that focuses on IaaS providers. Cloudorado can also be referred to as a price calculator for multiple cloud hosting providers, since the comparison is performed by calculating price for individually set of needs. Users can set cloud server requirements such as CPU, storage, memory, Operating System (OS) and get a price from numerous cloud providers. Cloudorado [26] helps users to make data-driven IaaS purchase decisions based on the pricing of cloud environment.

After their collection IaaS bundles were classified into the three categories as described above (COI, MOI,SOI). The categorization of the collected bundles is introduced by Amazon [25]. Amazon Web Services (AWS) are broadly adopted, therefore the proposed classification is considered to be reliable for the evaluation of the methodology [27]. In addition, each category is divided into three (3) subcategories (Small, Medium, Large), based on the size of their main feature (CPU, RAM, storage). The criteria for sub-categorization are presented in Table 1.

Table 1: **Optimized Instances**

COI (vCPUs)	**MOI (GB)**	**SOI (GB)**
Small: 1 - 4	Small: 8 - 16	Small: 100 - 500
Medium: 6 - 12	Medium: 32 - 64	Medium: 500 - 2,000
Large: 16-32	Large: 120 - 256	Large: 2,000 - 10,000

A total number of 409 price bundles were collected, coming from 23 providers. The number of compute, memory and storage optimized instances is 205, 104, 100 respectively. Data collection was based on requirements that were not addressed by all providers, thus the number of the collected IaaS bundles of each provider differs. Table 2 presents the cloud providers and the corresponding number of COI, MOI and SOI instances.

Table 2: **Cloud providers and number of instances per category**

Cloud Providers	COI	MOI	SOI
Microsoft Azure	10	6	9
Amazon	12	7	9
Google	12	6	8
Cloudsigma	9	5	6
Atlantic.net	7	4	2
M5	12	5	6
Elastichosts	9	4	4
Bitrefinery	9	4	4
Storm	9	6	8
Rackspace	12	6	9
E24cloud.com	9	3	4
Joynet	6	5	4
Stratogen	12	5	6
Eapps	9	4	2
Data dimension	12	6	5
Cloudware	6	3	2
Zippycloud	11	6	3
Exoscale	5	3	0
Vps.net	5	2	2
Dreamhost	6	3	2
Zettagrid	9	4	3
Cloudsolutions	12	5	6
Gigenet	2	2	0

4.2.2. Phase II-DEA implementation-Efficient bundles

Following the data collection and as the next phase, the methodology proceeds with the estimation of the relative efficiency rates, adopting DEA. DMUs, inputs and outputs compose an input-oriented DEA model. Cloud bundles are designated as Decision Making Units (DMUs). Each DMU is defined by the provider name and a given serial number. For example, in the Compute Optimized, instances category 'azure3' corresponds to the third IaaS instance of Microsoft Azure, as extracted from Cloudorado.

Price for an annual subscription is defined as the input of the proposed model. It is a multidimensional factor and there is usually a strong relation between price and efficiency. *Compute Power* (CPU cores), *Storage capacity* (GB) and *Memory* (GB) define the outputs of the model. Cloudorado offers a few more features, such as *Time On, Transfer In, Transfer Out* and the option that CPUs, RAM and storage can be distributed among more than one physical server, participating in the price bundling of Cloudorado. However, these characteristics are not considered as output parameters into the context of this work. The *Transfer In*, which refers to the number of bytes received by the server from the internet per month, does not observably contribute to the shaping of the pricing bundles and many providers (e.g. Amazon) charge only for the outgoing traffic, while others consider it as a small amount in the total price of services. In addition, the *Transfer Out*, describes the number of bytes sent by server to Internet per month, does not contribute to pricing [29]. Therefore and with no loss of generality, the *Transfer In* attribute was considered to be at a fixed level of 1GB and the *Transfer Out* at 10GB, per month. As far as *Time On* is concerned, it was set at a level of 100%

availability per day and the default offered value of non-distributed resources was also considered. In addition, Linux was chosen as the operating system, since it is an open-source operating system, without extra charge for the user [28].

Based on the current categorization the initial data set is divided into nine (9) categories and the proposed DEA-oriented model is applied for each category. As mentioned above, the software MaxDEA [22] is used for the implementation of DEA. The relative efficiency is estimated, the bundles are ranked by the efficient score and five (5) bundles with the highest efficiency scores are chosen, in order to be mixed for the composition of multi-cloud IaaS services.

4.2.3. Phase III - Multi-cloud Service Composition

In Phase III a new data set is formed, based on the five bundles with the highest efficiency scores, named as **S**ingle **P**rovider **M**ore **E**fficient (SPME) bundles. The proposed model aims to examine the contribution of multi-cloud strategy to cloud efficiency, therefore for more accurate results the five (5) chosen SPMEs are derived from different providers.

In the multi-cloud data set three patterns of services are designated. The first includes the possible combinations of three (3) out of five (5) SPMEs bundles, offering ten (10) possible combinations. The second pattern aims to cover every possible variation to the first multi-cloud data set, by including duplicate SMPE bundles of the same provider, for example SPME1,SPME1,SPME2. The possible combinations are totally twenty (20). The third pattern includes five (5) groups of three bundles, all from a single provider, for example SPME1,SPME1,SPME1, for comparison reasons.

The new multi-cloud data set includes thirty-five (35) multi-cloud bundles for each category (S/M/L COI, S/M/L MOI and S/M/L SOI).

As mentioned above, multi-cloud data set determination is based on the possible combinations of three (3) out of five (5) SPMEs bundles. The approach for composing the multi-cloud data set based on the possible combinations of the five (5) SPMEs bundles was initially examined. Later, the data set was formed based on the potential combinations of four (4) out of five (5) SPMEs bundles. However, the outcome revealed that the last two approaches did not result to a larger number of multi-cloud efficient bundles, therefore the approach based on the combinations of three (3) out of five (5) SPMEs bundles was finally chosen.

A simple example of a multi-cloud service of the medium compute-oriented category, based on three different single provider services, is presented in Table 3. The name of the multi-cloud service is 'cpu-medium1' and is described by medium compute instances derived from Azure, Storm and Dreamhost. Adding the size of the individual resources (vCPUs,RAM, storage) and the price of each instance, the final size and the final price are determined. For example, cpu-medium1 instance offers 24 vCPUs and its final price is 615$.

4.2.4. Phase VI: DEA implementation in multi-cloud data set

Multi-cloud bundles are designated as Decision Making Units (DMUs) and each DMU is defined by its category size and a given serial number. For example, multi-cloud instance 'cpu-medium1' represents the first multi-cloud instance that fits in the multi-cloud compute-medium category (M-COI).

Multi-cloud compute Power (vCPUs), storage capacity (GB) and memory(GB) define the outputs of the DEA model, whereas multi-cloud service

Table 3: **Sample of multi-cloud data set**

Cloud Services	vCPUs	Memory (GB)	Storage (GB)	Price ($)
Azure5	8	32	500	158
Storm6	8	32	1500	351
DreamHost5	8	16	280	106
cpu-medium1	**24**	**80**	**2,280**	**615**

price ($) constitutes the input of the model. Figure 2 depicts the proposed model multi-cloud model. The methodology collects and categorizes single provides services described by CPU, Memory Storage, and price. DEA is applied, pointing out the most efficient single Iaas packages. Finally, DEA is applied recurrently to identify multi-cloud efficient bundles, combining different number of IaaS packages.

The proposed multi-cloud DEA-model is applied to the data over each multi-cloud category (S/M/L COI, S/M/L MOI, S/M/L SOI), using the software MaxDEA [22] and the multi-cloud relative efficiency is finally estimated.

5. Results - Discussion

As mentioned above, the data set is categorized into three categories (COI, MOI and SOI) and each category is further divided into three subcategories, Small (S), Medium (M) and Large (L). Multi-cloud data set determination is based on the possible combinations of the five (5) top rated

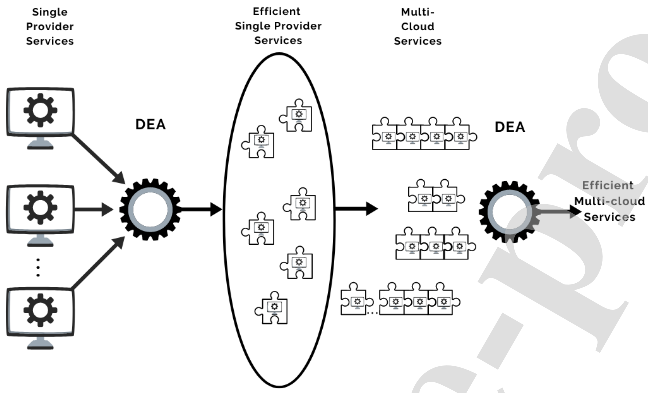


Figure 2: Multi-cloud DEA proposed model

SPMEs bundles in terms of efficiency is each subcategory. It should be noted that usually two of them were rated as efficient by the DEA approach, when examined as SPMEs. The proposed DEA model is applied over each category and the efficient multi-cloud bundles are pointed out.

The approach of composing the multi-cloud data set, based on the possible combinations of up to five (5) different providers was initially examined. As the number of efficient bundles consisting of more than 3 different providers was negligible, a data set based on the potential combinations of three (3) out of five (5) SPMEs bundles was chosen. As the outcome revealed, the five (5) or four (4) out of five (5) combinations did not result to a larger number of multi-cloud efficient bundles that the three (3) out of five (5), therefore the results presented in the following are based on the combinations of three (3) out of five (5) SPMEs bundles.

Table 4 presents the percentage of single-provider and multi-provider bundles characterized as efficient per subcategory, as well as the total percentage of efficient bundles. The groups of single provider services includes a cluster of three bundles, all from a single provider, for example SPME1,SPME1,SPME1. In addition, Figure 3 graphically illustrates the results.

Table 4: **Multi-cloud efficient bundles**

	COI S(%)	**COI M**(%)	**COI L**(%)	**MOI S**(%)	**MOI M**(%)	**MOI L**(%)	**SOI S**(%)	**SOI M**(%)	**SOI L**(%)
Groups of Single provider services	5.71	5.71	2.85	2.85	5.71	2.85	5.71	5.71	5.71
Multi-cloud Services	28.54	13.44	14.28	25.71	20.5	22.85	11.11	20.68	17.14
Total	34.25	19.15	17.13	28.57	26.21	25.71	16.82	26.39	22.85

In all cases the efficient multi-provider bundles outnumber the efficient groups of single provider ones. For example, in COI Small subcategory the rate of the efficient multi-cloud services equals to 28,5% whereas the rate of the efficient group of single provider services equals to 5,71%. The conclusion that multi-cloud strategy boosts efficiency is confirmed for each subcategory. Efficient multi-provider bundles consist of both efficient and inefficient single provider solutions. While an efficient solution may contribute to the efficiency of a multi-provider bundle, there is no indication that this determines the

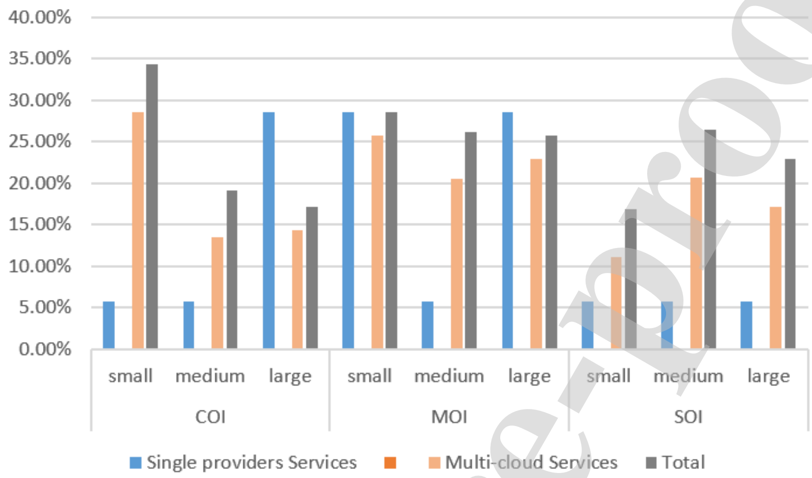


Figure 3: Results of DEA model per category

efficiency of the bundle in any way. Efficient multi-provider bundles consists of either tow or three providers in the same percentage.

The same conclusion is confirmed for all category (COI, MOI, SOI), but each category is affected by the multi-provider strategy to a different extent. MOI and COI categories present the most significant increase in the number of efficient multi-provider bundles, whereas the increase in the SOI group is noticeable but at a smaller scale. As mentioned in [29], storage price is decreasing but evidence indicate that this reduction may not as elastic as the cost of computing power. Figure 4 presents the contribution of multi-provider strategy to the cloud efficiency. The rate of efficient multi-cloud services in COI, MOI and SOI categories equals to 23.81%, 27.82% and 15.24% respectively. Furthermore, the rate of efficient bundles from a single provider has a minor contribution to the total number of efficient bundles per category.

More specifically, COI multi-cloud bundles are almost 14% more efficient that equivalent single provider services, whereas MOI multi-cloud bundles are 19.2% more efficient than similar single provider services. Finally SOI multi-cloud bundles 10.6% more efficient than corresponding single provider services.

Figure 4: Efficient bundles per category (COI/MOI/SOI)

By inspecting the results, it becomes obvious that the efficient multi-cloud bundles hold an overwhelming proportion in the total number of the efficient services. For example, the total number of efficient bundles in COI/Small category is comprised of 83.32% efficient multi-cloud bundles and 16,67% efficient groups bundles from a single provider. Figure 5 illustrates the corresponding proportions in each category.

Multi-cloud solutions also achieve significant cost saving. Table 5 illustrates the price reduction for each category. It can be observed that the final price of efficient groups of single provider bundles, is higher than the price

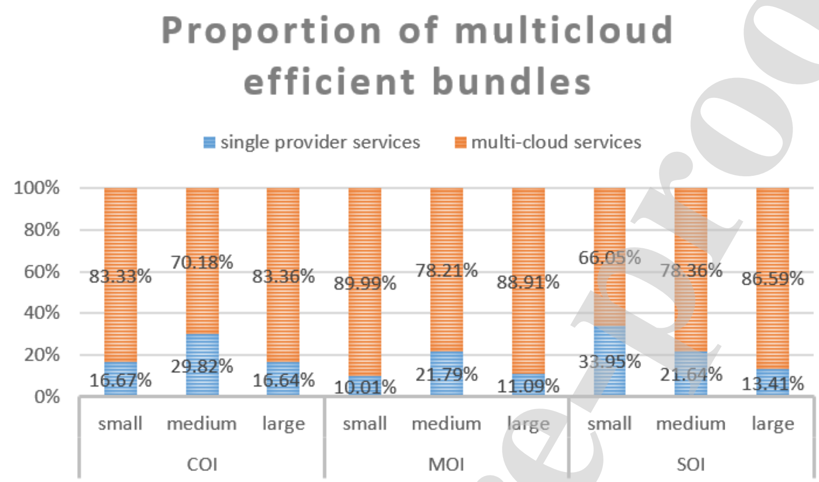


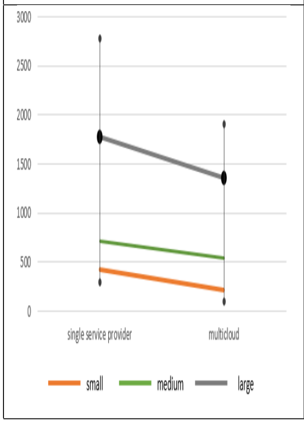
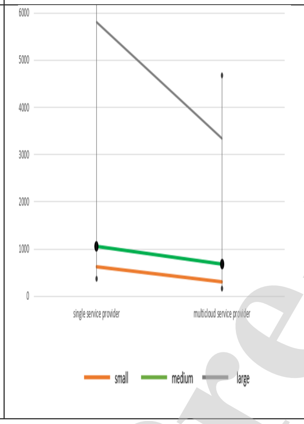
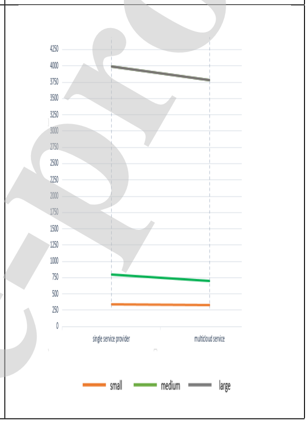
Figure 5: Proportion of efficient multi-cloud bundles in the total number of efficient bundles

of the efficient multi-cloud bundles and the specific conclusion is validated in each category (S/M/L COI, S/M/L MOI, S/M/L SOI). The price reduction appears to be significant in large categories, rather than in small and medium categories.

The results of the aforementioned experiment validate that multi-cloud strategy contributes significantly to cloud efficiency. Multi-cloud services deliver cost-efficient COI, MOI and SOI services. Thus, it is worth exploring the adoption of multi-cloud solutions utilizing existing technology, though these solutions may become more complex, than simply rely on a single provider.

Multi-cloud environment promotes cloud efficiency, however COI and MOI groups seem to be favored more than the SOI groups. As claimed by [29], the price per Byte of storage the last thirty (30) years is dropping but

Table 5: Price reduction per category

COI	MOI	SOI
		

evidence indicate that this reduction -according to simulations- may come to an end. Pricing models of current commercial cloud storage services are not suitable for long-term storage. Consequently, the findings of the above research [29] regarding the cost of cloud storage are also confirmed by the DEA analysis conducted in this work, indicating that the SOI category is less efficient in multi-cloud bundles.

A multi-cloud approach is strongly linked with cost reduction and can create even more considerable cost optimizations. There are fewer risks of vendor lock-in and users are free to choose resources based on the the appropriate mix of resources for their application (COI,SOI,MOI), promoting agility and competition in the most cost-efficient way.The cost saving is significant in large instances, indicating that the most profitable approach is for companies that demand high performance computing resources. According to the results, companies that will purchase large multi-cloud bundles will achieve optimal efficiency and significant cost savings.

6. Conclusions

Enterprises adopt the cloud environment usually opting for a single-cloud approach. As their business demands grow and fully embrace the cloud and its benefits, moving towards a hybrid approach is a common consequence, for a number of reasons. Hybrid cloud offers flexibility, raising questions on the same time about how the public cloud implementation will be handled and, therefore, a multi-cloud approach needs to be evaluated. A multi-cloud approach appears to be more complicated but after a thorough examination a multi-cloud aspect turned out to offer an extended range of benefits, promoting efficiency and cost reduction.

Towards this direction, the current work extends previous work [15] that explored the efficiency of single provider IaaS services. Both adopted DEA, however the input data and the proposed methodology steps differs. In [15] authors focused on functional and non-functional attributes whereas the current methodology focuses on IaaS services described by functional attributes (compute processing units, memory, storage) and price and creates a multi-cloud data set based on the most efficient single provider services. The multi-cloud data set is categorized into three categories and nine subcategories (S/M/L COI, S/M/L SOI,S/M/L SOI) and finally DEA is applied for each category exploring the multi-cloud efficiency of each group. In contrast to [15], DEA is applied recurrently to identify multi-cloud efficient bundles.

Based on the results, it is evident that multi-cloud strategy boosts efficiency, since the rate of efficient multi-cloud bundles is higher than the rate of efficient groups of single provider. For example, the total rate of efficient bundles in S-COI category is 34.25% including 28.54% efficient multi-cloud bundles and 5.71% efficient groups of single provider services. The same conclusion is confirmed for all categories and subcategories, but each category/subcategory is affected by the multi-provider strategy to a different extent. MOI and COI categories present the most significant increase in the number of efficient multi-provider bundles, whereas the increase in the SOI group is noticeable but at a smaller scale. More specifically, COI multi-cloud bundles are almost 14% more efficient that equivalent single provider services, whereas MOI multi-cloud bundles are 19.2% more efficient than similar single provider services. Finally SOI multi-cloud bundles are 10.6% more efficient than corresponding single provider services.

In addition, an important cost saving is affirmed, especially in bundles that are bound to be used to applications with high computational requirements. The proposed multi-cloud methodology reveals that businesses can select solutions that offer minimized monetary cost and guaranteed efficiency. It highlights multi-cloud solutions that provides the possible utilization of computing resources, improving cost visibility and enabling new and efficient solutions to perform different tasks.

As in most cases, there is the limitation in the proposed model, as qualitative characteristics of multi-cloud bundles were not explores. The proposed methodology refers to multi-cloud instances that are exclusively described by functional attributes. This considered to be a good direction for future research in order to examine the role of non-functional attributes to multi-cloud efficiency. A techno-economic analysis based on cloud brokerage services may also be examined. Efficient multi-cloud services can be offered to users by cloud broker, based on the proposed model results. Therefore, the role of cloud brokerage services and their profits in multi-cloud environment can be an interesting direction of the current work. This would greatly benefit buyers that would combine the most cost-effective bundles into one easier to use multi-cloud infrastructure harvesting the full potential of the combination.

References

[1] What is a hybrid cloud?, https://azure.microsoft.com/en-us/overview/what-is-hybrid-cloud-computing/, accessed: 24.10.2020.

[2] A. Charnes, W. Cooper, E. Rhodes, Measuring the efficiency of decision making units, European Journal of Operational Research 2 (6) (1978) 429 – 444. doi:https://doi.org/10.1016/0377-2217(78)90138-8. URL http://www.sciencedirect.com/science/article/pii/0377221778901388

[3] D. Sitaram, G. Manjunath, Chapter 2 - infrastructure as a service, in: D. Sitaram, G. Manjunath (Eds.), Moving To The Cloud, Syngress, Boston, 2012, pp. 23 – 71. doi:https://doi.org/10.1016/B978-1-59749-725-1.00002-0. URL http://www.sciencedirect.com/science/article/pii/B9781597497251000020

[4] D. B. Boupo, Hybrid Cloud : An Enterprise Solution 15 (4) (2017).

[5] R. Santhosh, J. Pradeep, Secure , Reliable and High Performance Multicloud Storage, International Journal of Distributed and Cloud Computing 4 (2016).

[6] J.-M. Bohli, N. Gruschka, M. Jensen, L. Lo Iacono, N. Marnau, Security and privacy-enhancing multicloud architectures, Dependable and Secure Computing, IEEE Transactions on 10 (2013) 212–224. doi:10.1109/TDSC.2013.6.

[7] M. L. Burstein, The Economics of Tie-in Sales, Palgrave Macmillan UK, London, 1988, pp. 145–155.

[8] H. Simon, G. Wuebker, Bundling - A Powerful Method to Better Exploit Profit Potential, Springer Berlin Heidelberg, Berlin, Heidelberg, 1999, pp. 7–28.

[9] B. Balebuff, Bundling as an Entry Barrier, The Quarterly Journal of Economics 4 (2) (2014) 347–361.

[10] A. Kar, P. Singh, A model for bundling mobile value added services using neural networks, Int. J. Applied Decision Sciences 5 (1) (2012) 47–63.

[11] T. Lucas, To bundle or not to bundle cloud services?, https://www.flexiant.com/2014/04/08/to-bundle-or-not-to-bundle-cloud-services/, accessed: 20.10.2020 (2014).

[12] P. Brebner, A. Liu, Performance and cost assessment of cloud services, in: E. M. Maximilien, G. Rossi, S.-T. Yuan, H. Ludwig, M. Fantinato

(Eds.), Service-Oriented Computing, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 39–50.

[13] L. M. D. Souza, Performance Evaluation Methodology for Cloud Computing using Data Envelopment Analysis, The Fourteenth International Conference on Network (c) (2015) 58–64.

[14] C. Xu, Y. Ma, X. Wang, A non-parametric data envelopment analysis approach for cloud services evaluation, in: F. Toumani, B. Pernici, D. Grigori, D. Benslimane, J. Mendling, N. Ben Hadj-Alouane, B. Blake, O. Perrin, I. Saleh Moustafa, S. Bhiri (Eds.), Service-Oriented Computing - ICSOC 2014 Workshops, Springer International Publishing, Cham, 2015, pp. 250–255.

[15] E. Filiopoulou, P. Mitropoulou, N. Lionis, C. Michalakelis, On the efficiency of cloud providers: A DEA approach incorporating categorical variables, IEEE Transactions on Cloud Computing 7161 (c) (2018) 1–16. doi:10.1109/TCC.2018.2850889.

[16] K. Costello, Gartner says worldwide iaas public cloud services market grew 37.3% in 2019, https://www.gartner.com/en/newsroom/press-releases/2020-08-10-gartner-says-worldwide-iaas-public-cloud-services-market-grew-37-point-3-percent-in-2019/, accessed: 20.10.2020 (2020).

[17] M. Braglia, A. Petroni, Evaluating and selecting investments in industrial robots, International Journal of Production Research 37 (18) (1999) 4157–4178.

[18] H. D. Sherman, J. Zhu, Data envelopment analysis explained, Service Productivity Management: Improving Service Performance using Data Envelopment Analysis (DEA) (2006) 49–89.

[19] R. D. Banker, A. Charnes, W. W. Cooper, Some models for estimating technical and scale inefficiencies in data envelopment analysis, Management science 30 (9) (1984) 1078–1092.

[20] Emrouznejad, Dea tutorial, https:http://deazone.com/en/resources/tutorial./, accessed: 20.10.2020 (2020).

[21] A. Charnes, W. W. Cooper, E. Rhodes, Evaluating program and managerial efficiency: an application of data envelopment analysis to program follow through, Management science 27 (6) (1981) 668–697.

[22] Maxdea software, http://maxdea.com/MaxDEA.htm/, accessed: 05.10.2020 (2020).

[23] Y. Hai, M. Deping, T. Ting, L. Wenjing, Annual index analysis on r&d input-output efficiency of faculties in huazhong normal university, in: Advances in Electrical Engineering and Electrical Machines, Springer, 2011, pp. 675–681.

[24] A. Iliyasu, Z. A. Mohamed, R. Terano, Data envelopment analysis models and software packages for academic purposes, Pertanika Journal of Scholarly Research Reviews 1 (1) (2015).

[25] Amazon web services, https://aws.amazon.com/, accessed: 20.10.2020.

[26] Cloudorado, https:/www.cloudorado.com/, accessed: 1.10.2020.

[27] F. Richter, Amazon leads $100 billion cloud market, https://www.statista.com/chart/18819/worldwide-market-share-of-leading-cloud-infrastructure-service-providers/, accessed: 05.10.2020 (2014).

[28] P. Mitropoulou, E. Filiopoulou, C. Michalakelis, M. Nikolaidou, Pricing cloud iaas services based on a hedonic price index, Computing 98 (11) (2016) 1075–1089.

[29] D. S. H. Rosenthal, D. C. Rosenthal, E. L. Miller, I. F. Adams, W. Mark, The Economics of Long-Term Digital Storage, UNESCO - United Nations Educational, Scientific and Cultural Organization (2010) 1–17.