REGULAR PAPER

Cloud services cost comparison: a clustering analysis framework

George Fragiadakis[1] [iD] · **Vasiliki Liagkou[1]** [iD] · **Evangelia Filiopoulou[1]** [iD] ·
Dimitris Fragkakis[1] · **Christos Michalakelis[1]** [iD] · **Mara Nikolaidou[1]** [iD]

Abstract

As cloud computing has grown in popularity, several different models and deployment strategies have emerged, fulfilling specific needs of different users. Thus cloud service selection is a major task that combines business and technical aspects. Several cloud providers categorize their bundles to help users find the best fit. However, the policy behind this categorization is unclear, making it difficult to be consistent among providers. Therefore, this work explores the complication of selecting the optimal cloud service among numerous and comparable solutions and introduces a selection framework for cloud services based on clustering analysis, providing an overall size categorization of cloud services derived from the cloud market and making the categorization policy explicit and homogeneous across all providers. Finally, a real-word case study is presented, highlighting the application of the proposed framework.

Vasiliki Liagkou, Evangelia Filiopoulou, Dimitris Fragkakis, Christos Michalakelis and Mara Nikolaidou have contributed equally to this work.

✉ George Fragiadakis
  gfragi@hua.gr

  Vasiliki Liagkou
  vliagkou@hua.gr

  Evangelia Filiopoulou
  evangelf@hua.gr

  Dimitris Fragkakis
  itp20144@hua.gr

  Christos Michalakelis
  michalak@hua.gr

  Mara Nikolaidou
  mara@hua.gr

[1] Department of Informatics and Telematics, Harokopio University of Athens, Omirou 9, 17778 Tavros, Greece

🖄 Springer

1 Introduction

The digital world offers the tools and the technology to help businesses grow globally and promote their services outside their geographical boundaries. A digital solution is a challenging business trend that enables companies to interact with customers, collaborate with employees, store vast volumes of data more effectively, and provide better information management [1]. Towards this direction, cloud computing empowers businesses to deploy faster and flexible digital solutions, with improved manageability and less maintenance, and enables digital solution designers to regulate resources based on fluctuating business demand [2]. Based on Gartner, more than 85% of organizations will embrace a cloud-first principle by 2025 and will not be able to fully execute on their digital strategies without the use of cloud-native architectures and technologies and 95% of new digital workloads being deployed on cloud-native platforms[3].

Cloud environments offer to cloud users the prospect to design more reliable solutions that utilize more resources than ever before [4]. A large number of similar or equivalent resources are provided by different cloud providers and cloud users can select suitable resources and deploy them for cloud workflow applications. [5]. However, the large amount of cloud services makes the cloud service selection stressful for users, since they have to find the optimal solution that fulfills functionality, followed by qualitative needs and cost constrains. Towards this direction, a cloud selection framework, based on a clustering analysis is proposed, offering guidance through the digital solution design and proposing a cloud service categorization based on the functional features. The proposed approach gathers cloud services from cloud market, described by the common functional attributes (CPU, RAM, storage capacity and type) and applies clustering analysis based on these key characteristics. The aforementioned characteristics are chosen, since they are the common characteristics, regardless the specific technology (IaaS, CaaS or PaaS).

The rest of the paper is organized as follows: Sect. 2 presents background details, whereas Sect. 3 introduces the proposed cloud selection framework. Section 4 presents the clustering analysis and describes the implementation of clustering in cloud services. In addition the challenging discoveries of the clustering are highlighted. Section 5 describes a case study that applies the proposed approach and finally Sect. 6 concludes the paper.

2 Background

Since businesses have embraced cloud environment, various applications have been moved or deployed to the cloud. Cloud providers offer a wide selection of cloud services each one optimized to fit different use cases and different budget limitations.

To this end, the body of the current section includes studies that offer decisional guidance for cloud service selection in terms of cost but also includes works that have implemented clustering analysis in cloud computing environment.

2.1 Choosing a cloud solution

Several approaches have been proposed to solve the service selection problem. In [6], authors discussed the problem of choosing the most suitable microservice architecture for a web application was addressed. Authors introduced three different approaches; a monolithic architecture, a microservice architecture operated by the cloud customer and finally a microservice architecture operated by the cloud provider. In addition, the cost was calculated for each architecture and based on the results microservices reduced infrastructure costs in compared to standard monolithic architecture. Moreover, the use of services specifically designed to deploy and scale microservices reduced infrastructure costs by 70%.

Moreover, a decision-making approach among Function as a Service (FaaS), Platform as Service (PaaS) and Container as a Service (CaaS) in terms of cost and effectiveness was discussed in [7] by proposing a simulation framework. The paper concluded that scaling, function configurations, dependent services, network latency influenced cost and performance.

In [8] the primary problem of selecting the optimal VM regarding the cost and the performance metrics for a given workload and user requirements, was introduced. Authors presented PaRIS, a data-driven system that estimated accurate performance for minimal data and predicted performance for different user requirements. In addition cost was calculated for numerous VM types across various cloud providers. In addition, a mathematical decision model was developed in [9]. The proposed model determined the selection of cloud computing services offered by different providers, taking into considerations integrity, confidentiality and availability risks and costs.

2.2 Clustering analysis for cloud services

Clustering analysis as a data mining function places data elements in similar groups [10]. It can be used as a stand-alone tool for solving problems related to data grouping and also as a pre-processing step for the implementation of various techniques and algorithms. It holds a guiding role in many areas of scientific research such as geo-information science [11, 12], medical research [13–15] education area [16, 17] and telecommunications field [18–20].

There are several papers in cloud computing fields that have adopted clustering analysis. Acknowledging that cloud services are heterogeneous in terms of quality attributes, clustering analysis was adopted for resource allocation [21–24], and also for examining security issues in cloud environment, adopting Hidden Markov Model and using clustering techniques [25].

Furthermore, clustering analysis was used to categorize microservices based on runtime performance [26]. Authors used graph clustering to characterize the call graph dependency structure and runtime performance of production microservices at Alibaba

clusters showing that the distribution of microservices execution time is heavy-tailed. Moreover, in [27] authors proposed a data cleaning approach that removes outliers or missing values, based on clustering analysis. The proposed approach was adopted in cloud computing and Big Data technologies.

In [28] authors proposed a Tukey-HSD based clustering analysis applied in a set of VMs. The proposed scheme formed clusters based on two key-parameters, CPU and RAM utilization. In addition, in [29] a cloud computing-based analysis on massive data of power utilization was presented. An improved k-means algorithm was proposed based on mapReduce model, improving the efficiency of the clusters.

However, the preceding work might be reused on various and varied market analyses in general. Clustering analysis is a general technique that can be used to group data elements based on their similarity, and it can be applied to a wide range of problems. However, the specific problem, data, and domain can affect the choice of clustering algorithm, distance metric, and evaluation criteria used in the analysis. Therefore, while the general idea of clustering can be applied to different problems, the specific implementation may need to be tailored to the specific problem at hand. Additionally, as technology and research are constantly evolving, new methods and techniques may be developed that may be more suitable for certain problems than existing methods. Therefore, it may not be possible to directly repurpose past work on clustering without considering the specific context and problem.

2.3 Framing the challenges

The selection of a cloud service is a significant task that combines business and technical aspects and is often a focus of research. When the boundaries among cloud models become blurry and confusing, users require guidance. Then, cost can be a major decision factor for them. Summarizing the aforementioned related literature, it is evident that cost holds is a key factor in the cloud service selection process. In [6] authors explored the differences between a monolithic architecture and a microservice architecture in terms of cost, whereas authors in [7] explored Function as s Service (FaaS), Platform as Service (PaaS) and Container as a Service (CaaS) in terms of cost and effectiveness. In addition in [8] authors proposed a performance benchmarking for selecting the optimal VM based on users' requirements and cost.

However, the current work proposes a different approach, offering an overall categorization of cloud market as well as a cost-oriented evaluation regardless of cloud technology (IaaS, PaaS, CaaS). The most significant market participants, such as Amazon, have segmented their offerings (virtual machines or instances) into categories like general purpose, memory optimized, storage optimized and compute optimized. Smaller niche providers, on the other hand, do not offer their products in this manner. As a result, a categorization of a wide range of instance types optimized to fit different use cases along with a cost evaluation is proposed, addressing a gap in the cloud market and in research. The current work serves as a primary decision point, informing users of the costs of the potential cloud service solutions that meet their needs. Unlike previous works such as [8] cloud service performance is not considered. However the discoveries of the proposed approach indicate potential cloud service solutions and

the associated costs. Acknowledging that cloud service performance benchmarking is important in the cloud service selection process, the proposed approach focuses on the first step in a cloud decision-making process, allowing users to avoid wasting time exploring solutions that may be withdrawn due to cost constraints.

3 A selection framework for cloud services

Building an information system is a difficult task. A solution architect is a project manager, researcher, designer, and business analyst all rolled into one, examining and comprehending the project from various perspectives [30]. He manages the project from a business standpoint as well as from the perspective of a software engineer, selecting the appropriate tools and platforms.

The increasing volume of services is challenging, therefore cloud platform selection decision is crucial and confusing for businesses. There are several advantages to select the most suitable cloud service, including the most evident of saving money. The solution architect has a vague concept of the needed resources and, more importantly, the associated cost of this solution while designing it. However, he may end up purchasing an ineffective cloud solution and either saving less money, overpaying for wasted resources, or having insufficient resources [2]. In this context, the current work presents a transparent decision-flow approach, based on clustering analysis, that offers an overall size-categorization of cloud services, derived from cloud market. Since, many business decisions require knowledge of the final cost, the price of each cloud bundle group is highlighted, guiding to the final decision.

The current work addresses the complication in choosing the optimal cloud service among numerous and comparable solutions. Figure 2 presents the decision flow diagram of the solution architect. It includes both the decision flow which is time-consuming, represented by the red line and the proposed approach which places an emphasis on making decisions as quickly and as immediately as possible and is depicted by the blue line.

More specifically, the "Conventional Approach" decision flow (red line) entails the process of identifying the best option for you from a huge variety of packages and cloud providers. This takes time, and not everyone has the resources to consider all of the options. The "Proposed Approach" (blue line), on the other hand, is the technique that the authors are providing in this study, with the major outcome being equally classified resource clusters and their prices based on basic cloud computing features, referred to as categories from now on.

When considering the following elements of the cloud environment, it is clear why the conventional procedure is time intensive. According to [31], the cloud service providers are numerous, regardless of the cloud technology (IaaS, PaaS, CaaS) that they supply. Furthermore, the compilation of bundles ranges from large to little players of the market. As a result, the final solution is difficult to discern during the design stage of the digital solution.

As previously stated, several cloud providers, namely the market's major players, categorize the bundles they offer. They use this to help prospective clients find the greatest fit for their needs. The policy underlying this categorization is unclear, and

it is difficult to be consistent among the providers who offer such a classification. As a result, the proposed approach offered by this study is centered on making the categorization policy explicit and homogeneous among all providers investigated. It is obvious that the stated technique could be employed at any moment and for any comparable dataset (dataset is the input for the clustering framework as illustrated on the green line in Fig. 2) and the homogeneity and transparency would remain the same. The key reason for this is because classification occurs regardless of the provider, but only their bundles are input for this approach, and only the fundamental characteristics of cloud technology are picked to be categorized. The greater the number of sources, the more accurate the results.

So another significant advantage of the study's findings is that it standardizes policy classification across all providers. As a direct consequence of this, the interested party begins by evaluating the cost of the product based on its key characteristics. After that he is able to differentiate in terms of quality among the available options by focusing purely on non-functional factors.

Clustering the cost of the offerings from all providers into uniformly sized clusters can potentially help to narrow down the options and make the selection process more manageable. By grouping the offerings into clusters based on cost, you can quickly identify which providers offer the most affordable solutions. The interested party may then concentrate on the assessed offerings from those providers, which can save time and effort compared to evaluating all offerings from all providers.

It is important to note that, while clustering by cost can be a useful approach, it may not be sufficient to make a final decision. Other factors such as performance, compliance requirements, service level agreements, and security features may also be important and should be taken into account when making the final decision. Additionally, it's important to validate the clusters by also comparing them with other performance factors like response time, uptime, and availability. However the proposed approached is the first level of a whole cloud services decision process. After the elimination of the bundles, cloud architects would be able to focus more intently on the proposed final decision and be compelled to conduct performance benchmarking on specific cloud architectures.

So it is possible to narrow down your options and find the most affordable solutions by grouping the prices of all providers' offerings into uniformly sized clusters. But it should be combined with other evaluation criteria and take into account the specific requirements of your workload or application to make the final decision.

In order to get into more depth about the clustering framework given in this study, the most significant clustering characteristics need to be chosen. It is important to make the right choice since the computed clusters will be built based on these features. Therefore, central processing unit, random access memory, storage capacity, and storage speed are chosen [32], because these are the common characteristics, regardless the specific technology (IaaS, CaaS or PaaS). In addition, they have the greatest impact on cloud pricing according to [33]. It is important to note that the cost of the network is not included because the pricing for cloud network services is fixed and does not come in any predefined packages.

Consequently, after the clustering analysis, cloud services are categorized based on the CPU, RAM and storage capacity and type and the corresponding cost of each

cluster is calculated. Thus, the solution architect based on the formed groups can determine which cluster his solution will fall into. Then, the research to the suppliers will be more efficient since he already knows the bundle range he wants to purchase.

The choice of a clustering method, according to [34], has a direct impact on the clustering outcomes. Since there are so many alternative clustering methods in the literature, it is critical to carefully study the features of the underlying problem before selecting a suitable technique. The provided dataset of *n* items will be split into a collection of *k* clusters based on the computation of clustering metrics.

Once the appropriate algorithm and number of *k* clusters are selected, a model will be produced that may be applied in the future for comparable datasets. The final result of this procedure will be the clusters of the selected characteristics, which may be utilized for further investigation of the problem in question.

4 Clustering analysis

As shown in Fig. 1, the clustering framework constitutes a significant component of the approach that has been proposed. It is possible to break it down into the steps that are listed below, as shown in Fig. 2.[1] These steps will be discussed in greater detail in the sections that follow.

- Data collection: The cloud bundles collection is acquired from six major cloud providers for both IaaS and CaaS services. As part of feature selection, authors chose the fundamental characteristics for target in clustering analysis based on price indices provided in [35].
- Clustering analysis and Evaluation: Various clustering approaches are investigated and compared using common metric values. Finally, the k-means approach is identified as the most suited.
- Clustering Results: The categories were retrieved, and the output of a model may be employed in future identical datasets.
- Sizing is determined by the clustering findings.
- Data labeling: The implementation of the created machine learning model on the cloud bundles dataset.

4.1 Data collection and preparation

In the cloud industry today, there are several cloud providers. The collected cloud bundles dataset [2,3] contains approximately 1229 bundles from six major cloud providers: Google [36], Amazon[37], Microsoft[38], IBM[39], Alibaba Cloud [40] and Digital Ocean[41], manually collected by authors using each service's official resources calculator on their respective website. Each row covers the key components of each bundle as well as the pricing. The bundles are classified according to CPU, RAM, storage size, and disk type. However, some concerns arose during the data preparation

---

[1] The figure has been designed using resources from https://www.Flaticon.com.

[2] https://github.com/gfragi/cloudCasestudy/blob/master/datasets/caas_data.csv.

[3] https://github.com/gfragi/cloudCasestudy/blob/master/datasets/iaas_data.csv.
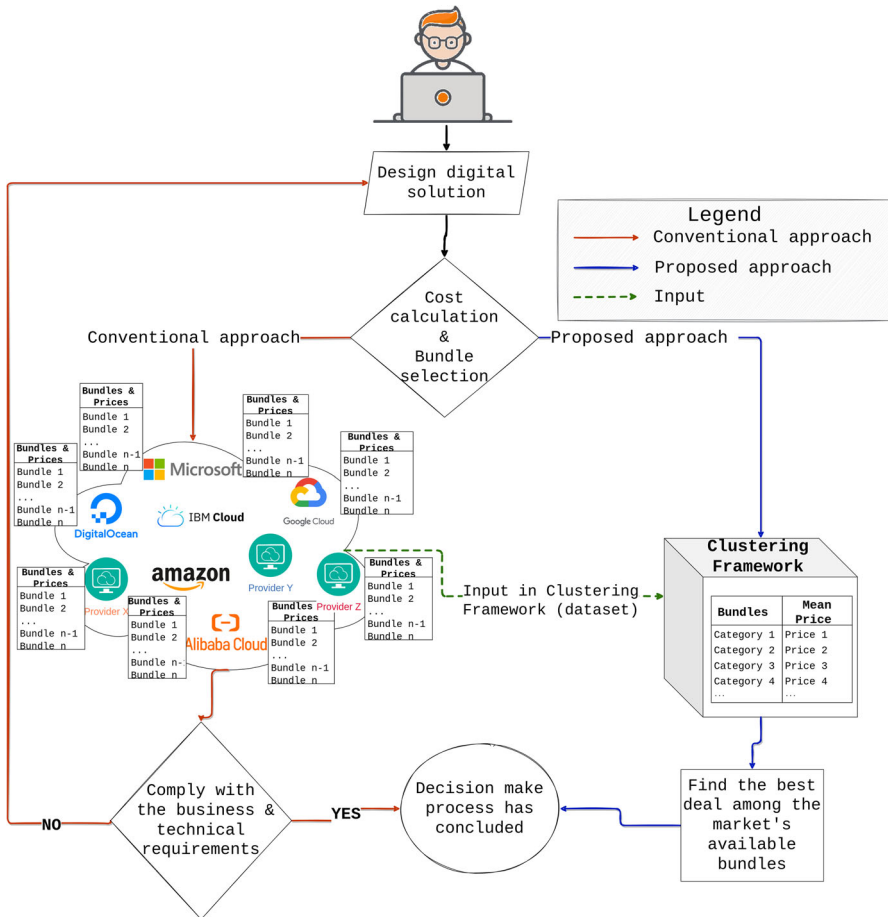
Fig. 1 Decision making process

process that required to be addressed before calculating the cost of each cluster, as a result of the grouping computation.

The provider's obtained data comprises cloud packages that meet a variety of service demands, resulting in a wide range of RAM and CPU combinations being given. Storage, too, provides a choice of storage capacity and disk type possibilities. This also makes estimating the cost of the chosen service challenging, as expenses vary depending on the resources given.

The above-mentioned restrictions were overcome by employing clustering analysis and categorizing these four essential traits into two unique offered bundled resources. A cloud resource is defined by its CPU and RAM combination, while another is defined by its storage capacity and disk type. A cloud bundles dataset is constructed by employing the distinctive values from the six providers as the average for each class in the clustering process.
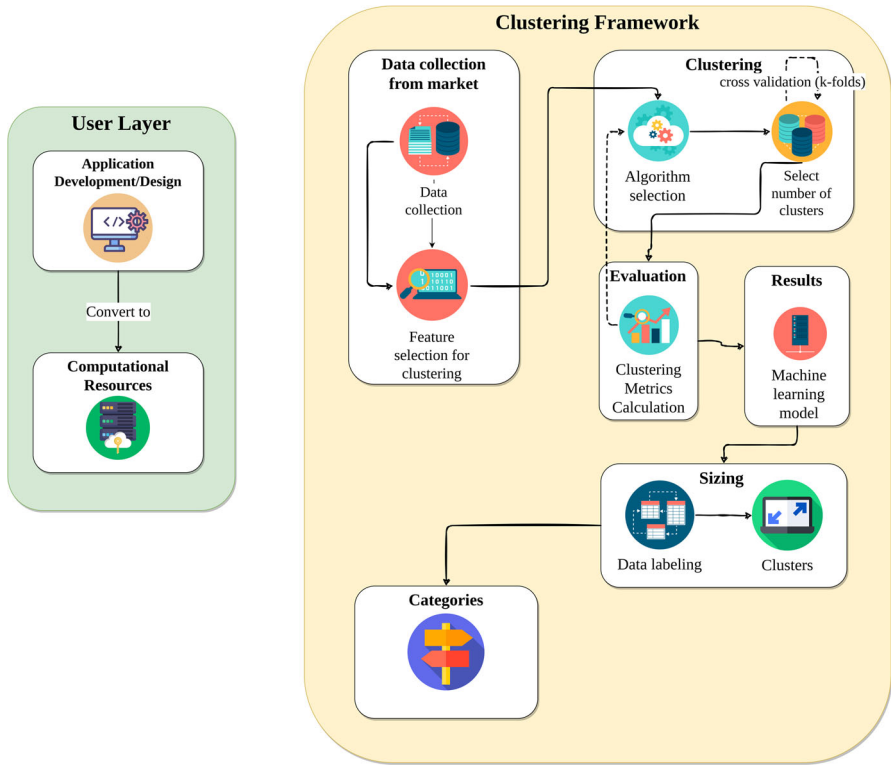
Fig. 2 Clustering framework

4.1.1 Implementation

The Python program[4] that was constructed for this study makes use of open source Python components and is available for download under an open source license. In addition to libraries such as numpy, panda, and plotly, the low-code machine learning package PyCaret [42] from the Python programming language is used.

4.2 Clustering

A significant number of data points in a data collection necessitates clustering the data into smaller categories. Clustering is a technique to machine learning that combines together data elements. Given a set of data points, cluster analysis is a technique for classifying each data point into a certain category. The features and qualities of data points within the same group should be comparable, whilst those of data points in other groups should be markedly unlike. This represents an instance of unsupervised learning. Unsupervised learning, in which the training data set is unlabeled and the objective

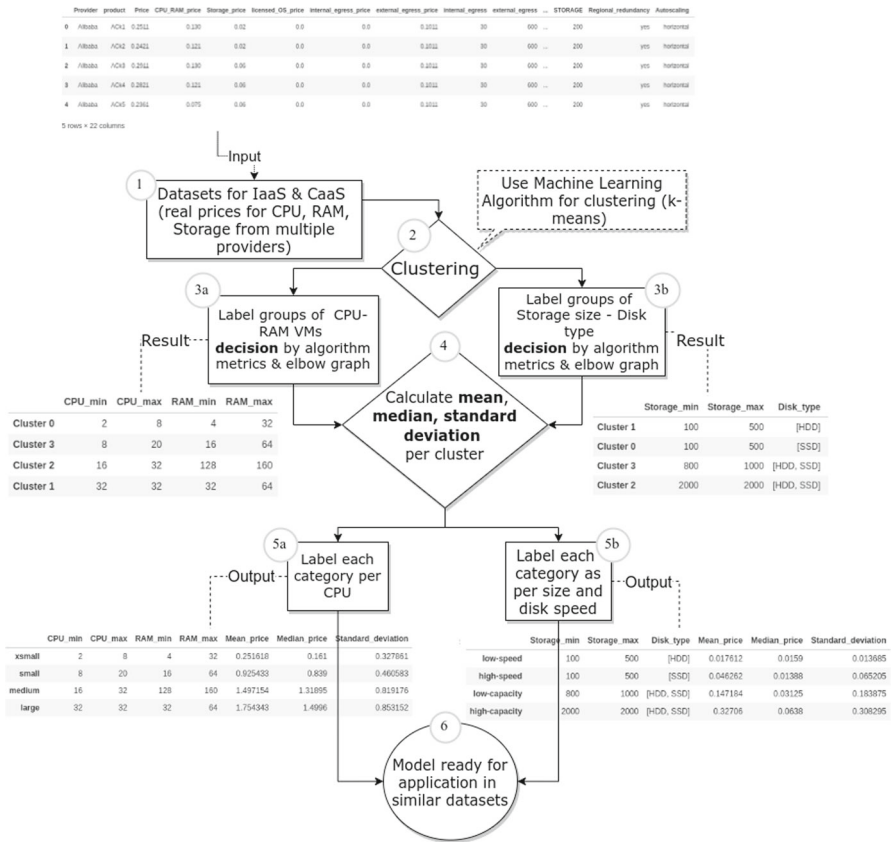---

[4] https://github.com/gfragi/cloudCasestudy.git.

Fig. 3 The process for clustering the dataset in this research

is to uncover underlying similarities, results in a somewhat compact representation of the data [43].

Using a clustering algorithm to classify the gathered data is the method employed in this study. This method classifies unlabeled data by recognizing groups of objects whose average distances between members are less than the average distances between members of other clusters. At [44] indicates that cluster analysis encompasses a variety of methodologies and algorithms for classifying objects of similar categories. Choosing the proper clustering technique and determining the ideal number of clusters in a dataset are essential concerns in clustering.

One may look at Fig. 3 to gain an overview of the general procedures involved in the clustering approach and how it was used to this research. After collecting the data as mentioned in 4.1, the first stage, as shown in Fig. 3, is to choose just the desired features (so to exclude others such ingress, egress etc). In the second phase, the clustering approach was used twice after picking the criteria (CPU, RAM, Storage capacity, and disk type) for which the grouping will be done. It was used the first time for the CPU and RAM, and the second time for the storage capacity and

Table 1 Results of clustering performance indicators

	dbscan	k-means	sc
Calinski–Harabasz	1557,6	4675	2470,6
Silhouette	0,751	0,673	0,75

disk type. Because the two types of resources had different properties, the clustering approach has to be implemented independently for compute and storage resources. It's possible that compute needs won't expand at the same rate. Decoupling storage from compute enables distinct cost management for storage and compute, as well as the implementation of various cost optimization features to achieve the goal of reducing overall costs [45]. The selection of the approach, as well as the number of groups, was determined by the corresponding metrics, as will be explained in more detail below. The original table was then labeled based on the results of the grouping (steps 3a and 3b at Fig. 3), which made it possible to calculate the mean, median, and standard deviation for each of the groups (step 4 at Fig. 3). This was done after the original table had been labeled (steps 5a and 5b at Fig. 3) based on the results of the grouping. Finally, Tables 2 and 3 were constructed, and a model was produced in step 6 that may be employed in further study involving data that is equivalent to the current one.

4.2.1 Clustering algorithm selection

The clustering algorithm and optimal number of clusters are chosen in the current research based on the results of a trial that evaluates a range of clustering performance measures.

According to [46], the silhouette coefficient is the most used method for combining cohesion and separation metrics into a single value (the range is [−1, 1], with the closer to 1 the better). However, according to the authors of [46], the Calinski-Harabasz coefficient (CH), also known as the variance ratio criteria, is a measure based on the internal dispersion of clusters and the dispersion between clusters (the higher the number, the better the performance).

Initially, Density-Based Spatial Clustering (DBSCAN), k-means, and Spectral Clustering are used as clustering approaches (SC). The Table 1 displays the estimated values of the clustering performance indicators which aid in algorithm selection.

Despite the fact that the silhouette indication of DBSCAN appears to provide the best results when compared to the other algorithms, the metric Calisnki-Harabasz coefficient differentiates k-means as the best approach, as shown in Fig. 4.

4.2.2 Number of clusters determination

As seen in Fig. 5, the number of clusters is computed using the Distortion Score Elbow. For a range of k values (e.g., 1 to 10), the below method clusters the dataset using k-means and computes an average score for each cluster for each value of k. According to [47], the distortion score is calculated by default as the sum of the square distances between each point and its assigned center. According to Fig. 5, the optimal number of clusters is four.
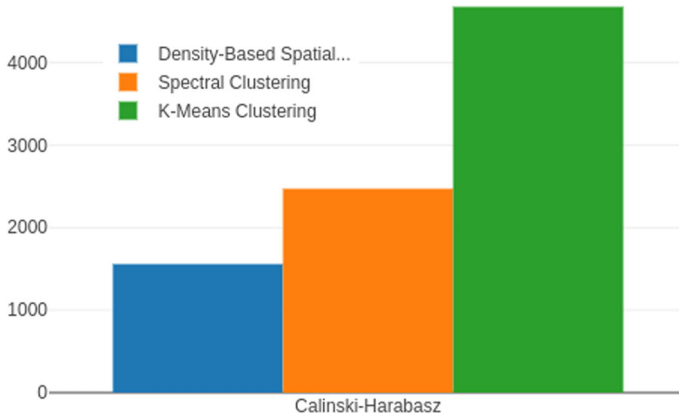
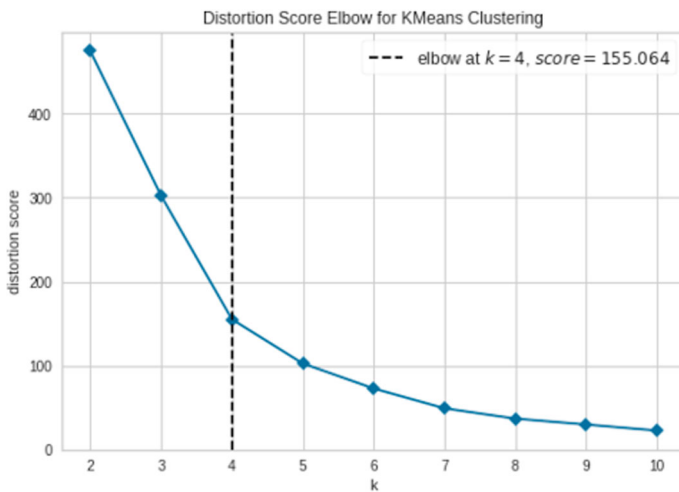**Fig. 4** Calisnki–Harabasz coefficients compared graphically

Fig. 5 Distortion Score Elbow for k-means clustering

4.2.3 Clustering findings

The dataset is grouped using clustering method, which divides it into four categories based on CPU and RAM characteristics. A deeper look at the CPU-RAM clusters reveals that the dataset has been split based on the size of the CPU-RAM. As a result, the newly created clusters have been given the names $xsmall$, $small$, $medium$, and $large$, as seen in Table 2 in more detail.

Furthermore, for the Storage capacity and disk type combinations, the clustering approach is employed once again, and the dataset is separated into four clusters depending on hard disk speed and capacity. As illustrated in Table 3, these clusters are classified $high - speed$, $low - capacity$, $low - capacity$, and $high - capacity$.

Table 2 CPU and RAM combinations clusters

	CPU min	CPU max	RAM min	RAM max
Xsmall	2	8	4	32
Small	8	16	16	64
Medium	16	32	128	160
Large	32	32	32	64

Table 3 Storage and disk type combinations clusters

	Storage min	Storage max	Disk type
Low-speed	100	500	[HDD]
High-speed	100	500	[SSD]
Low-capacity	800	1000	[HDD, SSD]
High-capacity	1000	2000	[HDD, SSD]

	Provider	Price	CPU_RAM_price	Storage_price	CPU	RAM	STORAGE	Disk_type	technology	Cluster
0	Alibaba	0.2511	0.130	0.02	4	8	200	HDD	iaas	Cluster 3
1	Alibaba	0.2421	0.121	0.02	4	8	200	HDD	iaas	Cluster 3
2	Alibaba	0.2911	0.130	0.06	4	8	200	SSD	iaas	Cluster 2
3	Alibaba	0.2821	0.121	0.06	4	8	200	SSD	iaas	Cluster 2
4	Alibaba	0.2361	0.075	0.06	4	8	200	SSD	iaas	Cluster 2

(a) labeled by CPU & RAM

	Provider	Price	CPU_RAM_price	Storage_price	CPU	RAM	STORAGE	Disk_type	technology	Cluster
0	Alibaba	0.2511	0.130	0.02	4	8	200	HDD	iaas	Cluster 0
1	Alibaba	0.2421	0.121	0.02	4	8	200	HDD	iaas	Cluster 0
2	Alibaba	0.2911	0.130	0.06	4	8	200	SSD	iaas	Cluster 1
3	Alibaba	0.2821	0.121	0.06	4	8	200	SSD	iaas	Cluster 1
4	Alibaba	0.2361	0.075	0.06	4	8	200	SSD	iaas	Cluster 1

(b) labeled by storage capacity and disk type

Fig. 6 Cloud providers' dataset labeled

4.2.4 Data labeling

The labeling of the collection of data is done point by point using the derived machine learning model based on the clustering findings. As a consequence, mean values and standard deviations for each group may be calculated. The labeled datasets for CPU-RAM and storage clustering findings are displayed in Fig. 6a and 6b, respectively.

Table 4 CPU-RAM clusters' average cost per hour

	Mean price ($/h)	Standard deviation
Xsmall	0,25	0,33
Small	0,92	0,46
Medium	1,49	0,82
Large	1,75	0,85

Table 5 Cost per GB per hour on average for storage capacity and disk type combinations

	Mean price (GB/h)	Standard deviation
Low-speed	0,01761	0,013
High-speed	0,04626	0,065
Low-capacity	0,1471	0,18
High-capacity	0,327	0,31

4.3 Average cost calculation

At this point, Table 4 shows the computation of each group's mean price and standard deviation for the CPU-RAM combination based on the labeled dataset of cloud providers given above.

As demonstrated in Table 5, calculating the average value of each category after labeling based on storage capacity and disk type is straightforward.

4.4 Clustering discoveries

The clustering analysis was performed on an extended set of price bundles from six various cloud providers, including the leading ones. The findings provide insight into the various cloud provider's pricing schemes for computing resources.

Per Table 2, it is noticed that the CPU-RAM value ranges in the resulted clusters correlate with the compute engine categories set by leading cloud providers. For example, comparing the Google's compute engine types with the resulted clusters, it reveals that the *xsmall* and *small* clusters correspond to *General Purpose* machine types and the *medium* and *large* clusters correspond to *Memory Optimized* and *CPU Optimized* types respectively. This demonstrates that the outcome of this research is aligned with the provider's policy.

Furthermore, according to the Tables 2 and 4 it worth noting that the price growth rate from *xsmall* to *small* CPU-RAM clustering group is 368% although the maximum capacity growth rate is 200% for the same groups. The same pattern does not hold true for the remaining CPU-RAM cluster levels, for example, the price growth rate from *small* to *medium* is 162% despite the fact the maximum capacity growth rate is 200%. It is also worth noting that neither storage billing follows the same pattern. More specifically, based on Tables 3 and 5, the price growth rate from low-capacity to high-capacity is roughly the same with the corresponding maximum storage capac-

ity. However, the high-speed pricing is about three times more expensive than the low-speed price for the same storage capacity.

Because there are no overlaps, the number of categories in both computational resources and storage on this study is justifiable when compared to the costs paid each category. They also reflect the costs incurred by the specific resource selection from any provider. The average cost provided here is consistent with the provider's pricing of bundles per category. For example, if you use Microsoft's calculator [38] and select the resources described in the *xsmall* bundle, as shown in Table 2, the associated pricing is aligned between the values in Table 4 and the Microsoft's bundle price.

5 Case study

Study in Greece (SiG) [48] is Greece's national agency for the internationalization of Greek Higher Education. The Study in Greece portal hosts and maintains all the pertinent information related to the Greek academic world, study options for international students and student life in Greece.

The case study presents the design of an online application system, developed by SiG's team, which will be used to submit to academic programs. The team through the design process searched for the optimal cloud solution in terms of cost between IaaS and CaaS technologies, adopting the proposed approach. The business coordinator requested a controlled budget for the solution architect's pick. Simultaneously, the key technical assistance provided from the upper level was the flexibility in developing the solution in the future, since predictions show that this particular system will resonate with users. As a result, the SiG solution architect should consider these business needs while selecting appropriate resources and cloud technologies.

The case study covered in this paper is a borderline application with simple requirements. Users of the platform are filtering available data or posting new events (posts). This implies that users are running simple queries and posting to the database, as well as logging in and updating their postings on the site. The authors' decision to use their technique in such a simplified case study does not imply that the proposed strategy cannot be applied to more demanding and complicated situations. The authors' focus is to demonstrate in a straightforward manner how their approach to first-level decision-making in terms of cost works. This may be used to demonstrate the benefits of such an isolated approach to architects by going further in their final selection by analyzing other criteria, such as performance, in limited alternatives.

More specifically, there are three separate user categories that will engage with this platform: applicants, program secretaries, and evaluators. In order to achieve this goal, there are a great deal of services and features, both front-end and back-end, that need to be built (such as user interface, system to store the data, etc.). Additionally, the number of academic institutions that are going to be covered by the programs that will be a part of this platform will gradually rise until they reach the point where they cover the whole region. This indicates that there is a requirement for both the continued development of the technology and its use. Because of this, a growth in the needs for the platform's resources is unavoidable and is directly related to the rise in the number of individuals making use of the platform.

Adopting the offered technique on his final decision might result in two advantages on the final report, which will be presented to the business coordinator for approval. First, because he could provide an exact cost estimate regardless of the supplier, he could present more accurate arguments for choosing one technology over the other depending on the users. Second, he may suggest a gradual spending fee rather than a one-time expense, as long as the number of customers grows. Furthermore, this might be a compelling argument for his team to profit from employing an expert on CaaS systems, as this cost cannot be considered when implementing cutting-edge technology.

5.1 Case study architecture

When deciding on the most suitable cloud service environment, the decision maker needs to take into account the application resource requirements. These requirements will be incorporated into the recommended strategy by this research with an emphasis on finding the most cost-effective approach, taking into account prices currently available on the market. Thus, it is necessary to split the application into its services and assign each service to the corresponding infrastructure resources. At their most fundamental level, the resources needs may be distilled down to computational power, memory and storage space.

The application consists of four main microservices, including a Frontend service (S_1), a Backend service ($S2$), a Database service ($S3$) and a Shared memory service ($S4$) as illustrated in Fig. 7. The Frontend service communicates on behalf of the clients and provides the path for exchanging messages between the clients and the Backend service. The Backend service implements the application business logic by processing the client requests. The Database service stores the application data while the Shared memory service is serving a shared memory layer in between the service components.

The aforementioned services have been developed using the following software platforms. The NGINX server [49] is used for the Frontend service, the NodeJS platform [50] is used as Backend server runtime and the PostgreSQL database [51] and the Redis cluster [52] are used for the Database and the Shared memory services respectively.

5.2 Deploying on different cloud technologies

The following section presents a baseline architecture of deploying the case study application in IaaS and CaaS. Also, a mathematical formulation has been defined to calculate the total compute requirements for the characteristics used in clustering analysis, CPU, RAM and Storage, on each cloud service.

The authors selected to take into consideration the CPU, RAM and Storage capacity because these are the common characteristics between IaaS and CaaS which have the most significant influence on cloud prices [35]. Furthermore, the network cost is not included because the cloud network pricing is concrete and is not grouped into price bundles.

Fig. 7 Architectural design
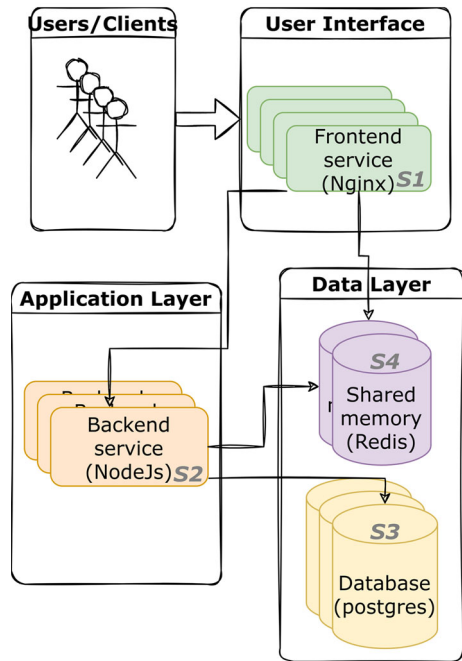
5.2.1 Deploying on IaaS

The architecture of deploying the SiG application system based on traditional VM based technology using IaaS cloud solution is given in Fig. 8. In IaaS, the Frontend service (S_1) is collocated with the Shared Memory service (S_4) in the VM machine, called VM_{FE}. The rest of two services, Backend service and Database service are splitted in separate VM machines, called VM_{BE} and VM_{DB} accordingly. In all VM types, Ubuntu server 20.04 LTS 64-bit [53] is used as the host operating system. The hypervisor is hosted and maintained by the cloud provider. All the VM types can be scaled independently to meet the workload demands or to attain high availability.

5.2.2 Deploying on CaaS

The SiG application system deployment architecture in container based technology using CaaS cloud solution is given in Fig. 9. In this deployment, all the aforementioned services are containerized and run on a container orchestrator platform across distributed cluster of virtual machines, hosted and managed by the cloud provider [54]. Containers, in contrast to VMs, fundamentally support horizontal and vertical scalability on a service level. In container technology multiple identical containerized services can run within the same machine and expand when needed. For the rest of the paper, the VM units in IaaS and the service units in CaaS are called instances depending on which cloud solution are referring to.
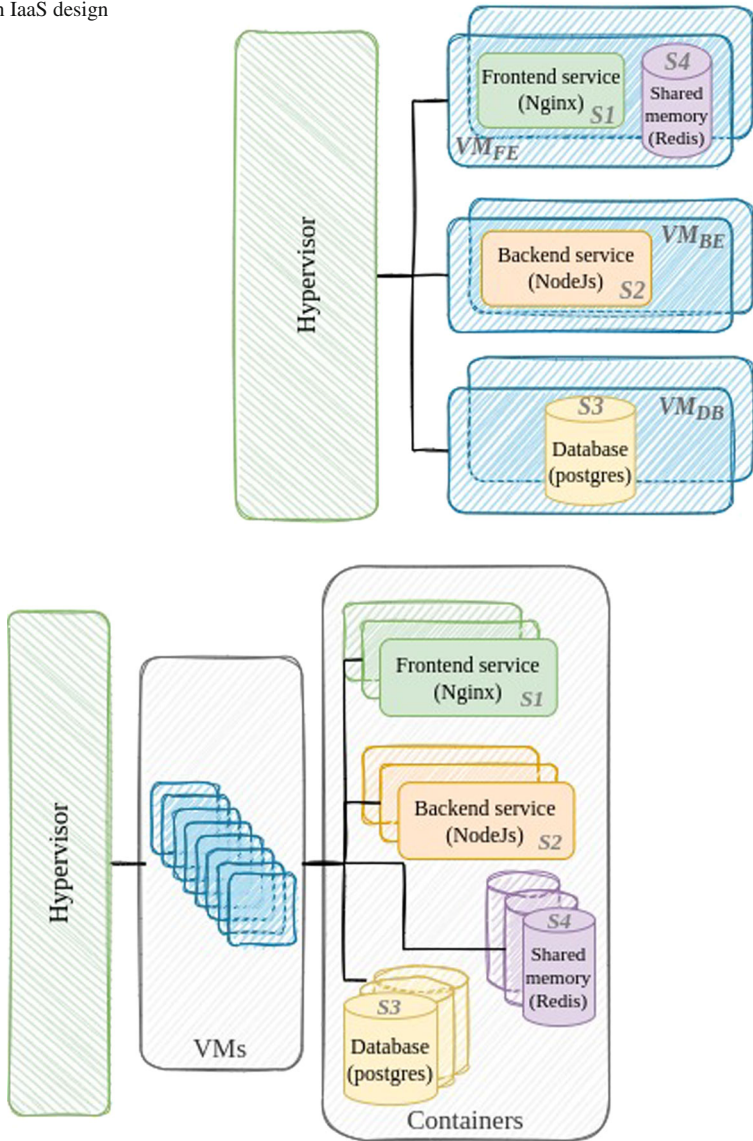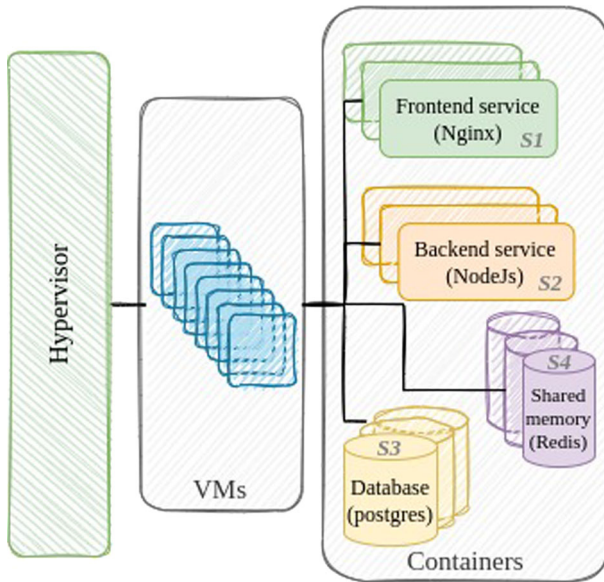
Fig. 8 Scaling in IaaS design

Fig. 9 Scaling in CaaS design

5.3 User-based resource calculation

The total compute requirements for the application are described by the three different equations, (1), (2) and (3) for CPU, RAM and Storage capacity accordingly based on the number of the desired users, denoted as $application_{users}$. The variable $instance_{users}$ represents the number of maximum users an instance supports. Whenever the load $application_{users}$ exceeds the $instance_{users}$, the instance scales out in

each cloud model.

$$Total\_CPU = \frac{application_{users}}{instance_{users}} * CPU[vCores] \tag{1}$$

$$Total\_RAM = \frac{application_{users}}{instance_{users}} * RAM[GB] \tag{2}$$

$$Total\_Storage = \frac{application_{users}}{instance_{users}} * Capacity[GB] \tag{3}$$

The equations are common in both cloud model scenarios. Each equation calculates the resources each instance requires in the corresponding cloud solution. For example, the resources for VM_{FE} in IaaS is the sum of the required resources for the S_1 service, the S_4 service and the operating system resources [53]. The authors need to mention here that the compute requirements for software components themselves are the same either the application is VM-based or containerized. The resources and costs between cloud models are calculated when only one service scales based on another without modifying software requirements.

The compute resources and maximum number of supported users for each instance using the software platforms described in Sect. 5.1 are based on their vendors' guidelines and are given in Table 6. They rely on the corresponding vendors' best practice technical recommendations. They have been provided as guidelines by vendors based on their internal testing so as to help the capacity planning information.

It is crucial to emphasize that the cost model for IaaS and CaaS is based on compute capacity that has been allocated rather than capacity that has been used. The authors took this decision because it would result in the highest solution cost. Therefore, the expense of deploying this solution shouldn't come as a surprise to the solution architect. More specifically the cost for a certain application service depends on the capacity which has been allocated for this service although it may not be utilized at all times. Based on that, the current work takes into consideration the compute capacity that needs to be provisioned as best practice by the vendors for each service, as illustrated in Table 6. Some usage testing would be done by service providers to come up with these best practices, which will help users come up with their own solutions. The method that was used to find these best practices was outside the scope of this study. Additionally, there are more parameters other than number of users that affect the compute requirements for each service. Nevertheless, load and performance tests are needed so as to observe how much they affect the resource needs which are also outside the scope of this study. However in CaaS, some providers offer the option to pay only for the compute resources the application is using but this is not considered in the current study.

5.3.1 Creating a user dataset

The hardware requirements of the proposed cloud-based architecture are determined by the number of users anticipated. The number of users for evaluating the proposed cloud architecture spans from 10 to 80.000. Four user sub-ranges are established based on the maximum number of each service, as shown in Table 6, (user groups: 10–1024,

Table 6 Best practices for each service based on user demand (virtual machine or container)

	Service	Max users	CPU [vCores]	RAM [GB]	STORAGE [GB]
IaaS	VM_{FE} - $Nginx$, $Redis$	1024	13	17	267,5
	VM_{BE} - $NodeJS$	65,535	9	33	602,5
	VM_{DB} - $PostgreSQL$	5000	9	25	102,5
CaaS	$S1$ - $Nginx$	1024	8	8	255
	$S2$ - $NodeJS$	65,535	8	32	600
	$S3$ - $PostgreSQL$	5000	8	24	100
	$S4$ - $Redis$	10,000	4	8	10

1025–5000, 5001–10,000, 10001–80000). Then, from each user group, 100 distinct sample numbers are picked at random, and lastly, 400 samples of the number of people who will use the cloud-based application service are applied to the mathematical formula using a random number generator. Finally, as seen in Fig. 10, the formed users' data-set is subjected to the aforementioned mathematical equations in order to determine the necessary hardware resources.

5.3.2 Cost calculation based on users

In the case study, the cost of deploying a cloud-based application is calculated by deploying the IaaS model and the CaaS model. The results of the clustering analysis for the CPU and RAM combination are presented in Table 2. These results are used to determine the relevant combination of resources per instance, which is described in the same table, based on the initial requirements of the cloud-based application. Therefore, in the case that a bundle is selected on the basis of the requirements, the appropriate average cost from Table 4 is multiplied by the overall number of instances based on the Eq. 4.

$$CPU\_RAM\_Cost = Mean\_price * Total\_instances \qquad (4)$$

As a result, the IaaS design must be implemented using a mix of CPU and RAM that is mapped on the cluster identified as *small*, and for the CaaS design on the cluster labelled as *xsmall*. This means that the overall cost is computed by multiplying the number of instances required by the average price of the relevant cluster for CPU and RAM cost.

Using the same approach, storage space costs are determined by multiplying the average value and the related storage space by Eq. 5. As a result of the preceding procedure, it is feasible to calculate the total cost of Storage based on the total number of users. Figure 12 represents the findings of this estimation.

$$Storage\_Cost = Mean\_price * Storage\_capacity \qquad (5)$$

5.4 Case study results

The results after applying the clustering method in the cloud bundles dataset revealed important findings about the pricing policies of the providers based on the size of the bundles. Moreover, the proposed approach introduce an interesting comparison analysis between IaaS and CaaS resources and the corresponding costs.

Figure 13a illustrates the price evolution of CPU-RAM based on the supported number of users after applying the clustering approach in the IaaS and CaaS deployments. By incorporating the linear trend, it is demonstrated that the growth rate of change in proportion to users in IaaS is roughly four times more than in CaaS (slopes are $1 \cdot 10^{-3}$ and $3 \cdot 10^{-4}$ respectively) on compute resources. This is consequence of the clustering analysis results. Based on Tables 6 and 2, the CaaS instances fit to *xsmall*

	users	component	instances	RAM	CPU	STORAGE	Disk_type	technology
0	576	VMfe	1	17	13	267.5	SSD	iaas
1	576	VMbe	1	33	9	602.5	HDD	iaas
2	576	VMdb	1	25	9	102.5	SSD	iaas
3	960	VMfe	1	17	13	267.5	SSD	iaas
4	960	VMbe	1	33	9	602.5	HDD	iaas
...	...	...	...	...	...	...	...	...
2795	26324	S4	3	24	12	30.0	HDD	caas
2796	55427	S1	55	440	440	14025.0	SSD	caas
2797	55427	S2	1	32	8	600.0	HDD	caas
2798	55427	S3	12	288	96	1200.0	SSD	caas
2799	55427	S4	6	48	24	60.0	HDD	caas

Fig. 10 Users dataset

	users	component	instances	RAM	CPU	Storage Price	Disk_type	technology	Cluster
0	576	VMfe	1	17	13	12.375094	SSD	iaas	low-capacity
1	576	VMbe	1	33	9	10.611048	HDD	iaas	high-speed
2	576	VMdb	1	25	9	4.741859	SSD	iaas	low-capacity
3	960	VMfe	1	17	13	12.375094	SSD	iaas	low-capacity
4	960	VMbe	1	33	9	10.611048	HDD	iaas	high-speed
...	...	...	...	...	...	...	...	...	...
2795	26324	S4	3	24	12	0.528351	HDD	caas	high-speed
2796	55427	S1	55	440	440	4587.016500	SSD	caas	low-speed
2797	55427	S2	1	32	8	10.567018	HDD	caas	high-speed
2798	55427	S3	12	288	96	176.621229	SSD	caas	high-capacity
2799	55427	S4	6	48	24	1.056702	HDD	caas	high-speed

Fig. 11 Scaling services as users evolve

group. In IaaS they level up to the next resource group instead which is the *small* one. The average price for the latter is nearly four times, and more precisely 3.68, higher than the former's based on the clustering results. Per Fig. 13a this rate is influenced in the CPU-RAM price evolution for the two cloud models.

Though, the price evolution of storage, as shown in Fig. 13b, is nearly identical between IaaS and CaaS. One aspect of this is that the storage capacity for Shared Memory service, which unnecessarily scale in VM_{FE}, is very low compared to the Frontend service's storage needs. Based on Table 3, this growth is not large enough to move the IaaS instances to the next storage clustering group. So, the same price level is applied for storage in both cloud models. Another aspect of the results is that CPU and RAM characteristics are the ones that contributing most to the shaping of the final price compared to storage [35].

Concerning the required compute resources in each cloud model, it is essential to observe how the application services scale as users evolve as illustrated in Fig. 11.

After the initial deployment, the application needs to be scaled up, to meet the user demand, according to the requirement of a specific service. The "jumps" in Fig. 11 represent the thresholds where the service scales out to an additional instance to serve the increasing load. Since the resource requirements differ from service to service, as given in Table 6, the $S1$ service scales more rapidly than the rest. This difference creates a scaling problem in IaaS model since all collocated services with $S1$ in the VM host are scaling in parallel although their operations exhibit different workloads. In CaaS, the scaling is more granular since each microservice is deployed and scaled independently to meet the workload needs without impacting the rest.

Furthermore, the CaaS implementation allows fitting multiple containers on a single host avoiding the multiplied OS resources when scaling. As seen in Fig. 14, this increases the anticipated infrastructure resources for IaaS compared to CaaS. More specifically, memory and CPU differ by over 30%, while storage differs by around 10% across the two cloud systems.

5.5 Case study outcome

According to case study results, deploying an information system to IaaS or CaaS cloud technology without changing the software resource requirements resulted in instances moving to a lower cluster level when in CaaS while also allowing them to function on cheaper devices. However, for a small portion of user usage, IaaS may be a better solution than CaaS because the latter may conceal other expenses that were not included in the proposed strategy, such as extra fees for managed orchestration services or even extra expenditures to acquire containerization skills.

It is demonstrated that the cost of the solution plays a key role in the selection of the solution and the proposed approach simplifies the decision-making process. The decision maker gets going with a first level decision by choosing a specific cloud service based on cost limited to a specific cluster group. Afterwards he may proceed with second level decisions such as by choosing a specific bundle within this cluster group or considering different selection criteria other than cost for a low user base.

In the case study given, the solution architect designed the information system solution and then reported the pros and disadvantages of both deployments to the business coordinator. Because of the lower operational costs as users increased, it was agreed that the CaaS cloud technology would be used to deploy the aforementioned information system. The second decision was to add a CaaS cloud system expert to the team, which would be beneficial in the long run as the platform's popularity grew. All of this occurred before to the development's start, making the design plan solid and future-proof.

6 Conclusions

Cloud providers offer numerous services with different pricing schemes, aiming to fulfill the constantly increasing client requirements. Each cloud service model offers different levels of control and management. Understanding the pricing among the

CPU & RAM price IaaS vs Caas
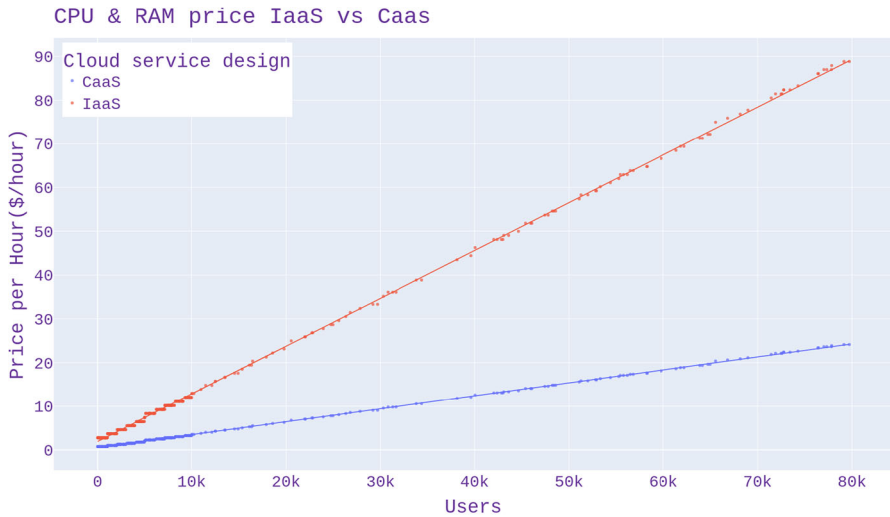
(a) CPU and RAM price growth by user

Storage Price IaaS vs CaaS
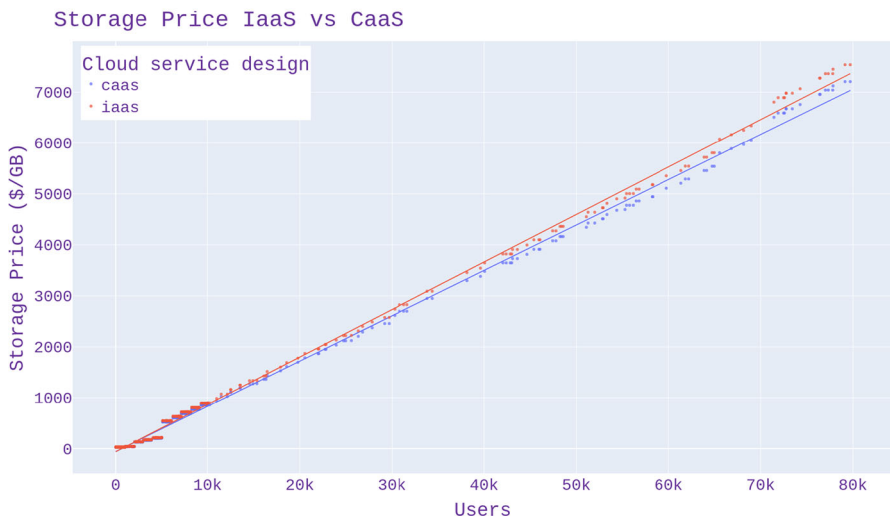
(b) The growth of storage prices by user

Fig. 12 Cost of storage

cloud services can help the decision makers to choose the most suitable option with his needs which proves a challenge for the companies.

This study addresses the fundamental problem of accurately and economically choosing the optimal service for a digital solution. The current work presents a decision-making approach, based on clustering analysis on a collection of cloud price bundles captured from six major cloud providers. These bundles are described by CPU, memory and storage which are the main functional characteristics having the
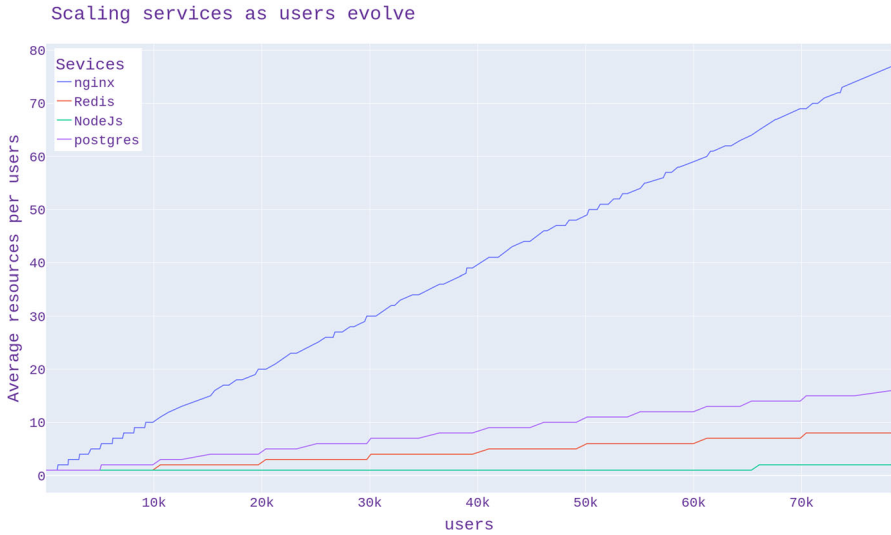
Fig. 13 Price evolution of IaaS and CaaS based on number of users
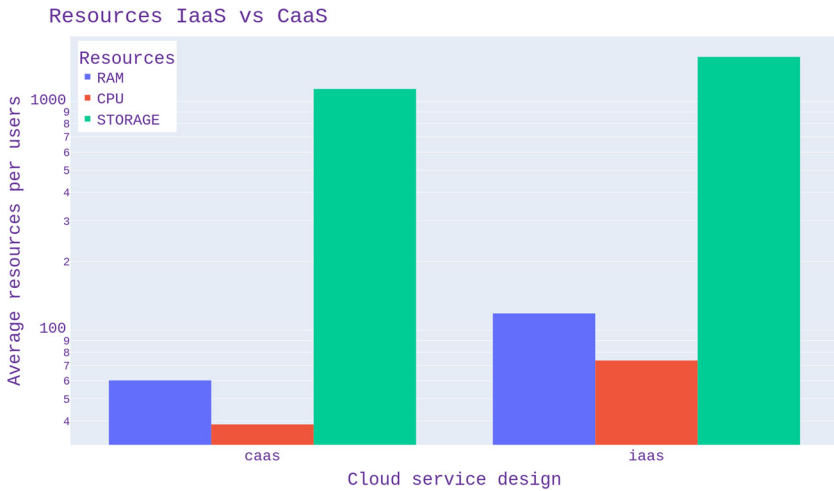
Fig. 14 Comparison between IaaS and CaaS resources

most significant influence on cloud prices. The bundles were categorized into four cluster groups based on their size and for each group an average cost is determined across all providers. The results standardize the pricing policy classification on different cloud computing services across all providers and provides valuable insights to decision makers. The proposed approach integrates the cost parameter in an organized way in the decision-making process. It demonstrates that the cost is a key parameter in this process and will be the first one to consider in the cloud service selection. Among to the important findings that were revealed about the pricing policies of the providers,

the grouping results, based on the size of the bundle, agree with the equivalent grouping that the leading service providers have in their pricing policy.

Furthermore, the cluster analysis findings were utilized to calculate the overall computing expense of deploying a case study application in different cloud services and compare them afterwards. By advancing the results, the decision maker is able to make a first level decision based on cost and consider the rest variables in a second level step.

To the limitations of this study, the first is that it considers compute costs only for CPU, RAM, and storage. Additional costs that are factored into cloud prices, such as costs for operation, security, etc., are excluded from this work. Furthermore, the two deployment models were compared based on the provisioned capacity, which doesn't take resource utilization into account. The choice of the presented case study may be considered simplistic as the resource requirements scale linearly with workload demand. So, the fact that random requests are used to test the dataset is another limitation of this study. A more accurate trace might be used in a later composition in its place.

Further extensions and future research directions would be to include non-functional characteristics in clustering analysis. Based on this methodology, an online service might be created where the input dataset would aggregate prices in real time from as many sources as possible to enable real-time alignment of cluster average costs. Moreover, the same dataset of services' price bundles could be applied to different decision making methodologies.

Declarations

Conflict of interest All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

References

1. Kosimov J, Ruziboyeva G (2022) The role of the digital economy in the world. Sci Prog 3(2):435–441
2. Singh S, Chana I (2013) Cloud based development issues: a methodical analysis. Int J Cloud Comput Serv Sci 2(1):73
3. GSartner: gartner says cloud will be the centerpiece of new digital experiences. https://tinyurl.com/mrvahc52

4. Moghaddam FF, Rohani MB, Ahmadi M, Khodadadi T, Madadipouya K (2015) Cloud computing: Vision, architecture and characteristics. In: *2015 IEEE 6th control and system graduate research colloquium (ICSGRC)*, pp 1–6 (2015). IEEE

5. Liu X, Chen J, Yang Y (2012) Temporal QoS management in scientific cloud workflow systems. Elsevier, Amsterdam

6. Villamizar M, Garces O, Ochoa L, Castro H, Salamanca L, Verano M, Casallas R, Gil S, Valencia C, Zambrano A, et al (2016) Infrastructure cost comparison of running web applications in the cloud using aws lambda and monolithic and microservice architectures. In: *2016 16th IEEE/ACM international symposium on cluster, cloud and grid computing (CCGrid)*, pp 179–182. IEEE

7. Manner J (2019) Towards performance and cost simulation in function as a service. Proceedings of the of ZEUS

8. Yadwadkar NJ, Hariharan B, Gonzalez JE, Smith B, Katz RH (2017) Selecting the best vm across multiple public clouds: a data-driven performance modeling approach. In: Proceedings of the 2017 symposium on cloud computing, pp 452–465

9. Martens B, Teuteberg F (2012) Decision-making in cloud computing environments: a cost and risk based approach. Inf Syst Front 14(4):871–893

10. Lee RC (1981) Clustering analysis and its applications. Advances in information systems science. Springer, Berlin, pp 169–292

11. Deng M, Liu Q, Wang J, Shi Y (2013) A general method of spatio-temporal clustering analysis. Sci China Inf Sci 56(10):1–14

12. Barnes A, Islam MM, Toma L (2013) Heterogeneity in climate change risk perception amongst dairy farmers: a latent class clustering analysis. Appl Geogr 41:105–115

13. Tseng K-K, Li J, Tang Y-J, Yang C-W, Lin F-Y, Zhao Z (2020) Clustering analysis of aging diseases and chronic habits with multivariate time series electrocardiogram and medical records. Front Aging Neurosci 12:95

14. Bruisten SM, Schouls L (2009) Molecular typing and clustering analysis as a tool for epidemiology of infectious diseases. Modern infectious disease epidemiology. Springer, Berlin, pp 117–141

15. Xia K, Wu Y, Ren X, Jin Y (2013) Research in clustering algorithm for diseases analysis. J Netw 8(7):1632

16. Mohamed Nafuri AF, Sani NS, Zainudin NFA, Rahman AHA, Aliff M (2022) Clustering analysis for classifying student academic performance in higher education. Appl Sci 12(19):9467

17. Huang L (2016) Teaching management data clustering analysis and implementation on ideological and political education of college students. In: 2016 International conference on smart grid and electrical automation (ICSGEA), pp 308–311. IEEE

18. Hong R, Zheng Y, Wu Y-R (2009) Clustering analysis of telecommunication customers. J China Univ Posts Telecommun 16(2):114–128

19. Velmurugan T (2014) Performance based analysis between k-means and fuzzy c-means clustering algorithms for connection oriented telecommunication data. Appl Soft Comput 19:134–146

20. Qiuru C, Ye L, Haixu X, Yijun L, Guangping Z (2012) Telecom customer segmentation based on cluster analysis. In: 2012 International conference on computer science and information processing (CSIP), pp 1179–1182. IEEE

21. Shahidinejad A, Ghobaei-Arani M, Masdari M (2021) Resource provisioning using workload clustering in cloud computing environment: a hybrid approach. Clust Comput 24(1):319–342

22. Singh S, Chana I, Buyya R (2017) Star: Sla-aware autonomic management of cloud resources. IEEE Trans Cloud Comput 8(4):1040–1053

23. Wei J, Zeng X-f (2019) Optimal computing resource allocation algorithm in cloud computing based on hybrid differential parallel scheduling. Clust Comput 22(3):7577–7583

24. Shang Q (2021) A dynamic resource allocation algorithm in cloud computing based on workflow and resource clustering. J Int Technol 22(2):403–411

25. Kumar P, Sehgal V, Shah K, Shukla SSP, Chauhan DS, et al (2011) A novel approach for security in cloud computing using hidden markov model and clustering. In: 2011 World congress on information and communication technologies, pp 810–815. IEEE

26. Luo S, Xu H, Lu C, Ye K, Xu G, Zhang L, Ding Y, He J, Xu C (2021) Characterizing microservice dependency and performance: Alibaba trace analysis. In: Proceedings of the ACM symposium on cloud computing, pp 412–426

27. Rahul K, Banyal RK (2019) Data cleaning mechanism for big data and cloud computing. In: 2019 6th international conference on computing for sustainable global development (INDIACom), pp 195–198

28. Mudgill V, Aujla GS, Kumar N, Obaidat MS, Prodan R (2018) Dlopc: data locality independency-aware vm clustering in cloud computing. In: 2018 IEEE globecom workshops (GC Wkshps), pp 1–6. IEEE
29. Zhang C, Gong Y, Zhang H (2019) Research on massive data analysis of power utilization based on improved k-means algorithm and cloud computing. In: 2019 international conference on intelligent computing, automation and systems (ICICAS), pp 239–242. IEEE
30. Krishnamurthy R (2017) Breezing my way as a solution architect: a retrospective on skill development and use. IEEE Softw 34(3):9–13
31. Peterson R. Cloud Service Providers (2022 Update). https://dgtlinfra.com/top-10-cloud-service-providers-2022/
32. Ren L, Zhang L, Wang L, Tao F, Chai X (2017) Cloud manufacturing: key characteristics and applications. Int J Comput Integr Manuf 30(6):501–515. https://doi.org/10.1080/0951192X.2014.902105
33. Mitropoulou P, Filiopoulou E, Michalakelis C, Nikolaidou M (2016) Pricing cloud IAAS services based on a hedonic price index. Computing 98(11):1075–1089
34. Wegmann M, Zipperling D, Hillenbrand J, Fleischer J (2021) A review of systematic selection of clustering algorithms and their evaluation. ArXiv:2106.12792
35. Liagkou V, Fragiadakis G, Filiopoulou E, Michalakelis C, Kamalakis T, Nikolaidou M (2022) A pricing model for container-as-a-service, based on hedonic indices. Simul Model Pract Theory 115:102441
36. Google: Google Kubernetes Engine. https://cloud.google.com/kubernetes-engine
37. Amazon: Amazon Elastic Container Service (Amazon ECS). https://aws.amazon.com/ecs/
38. Microsoft: Azure Kubernetes Service (AKS). https://azure.microsoft.com/en-us/products/kubernetes-service/
39. IBM: IBM Cloud Kubernetes Service. https://www.ibm.com/cloud/kubernetes-service
40. Alibaba: Alibaba Cloud Container Service for Kubernetes (ACK). https://www.alibabacloud.com/product/kubernetes?spm=a3c0i.11270126.6791778070.449.6c865f934LnCGu
41. DigitalOcean: Business-ready Kubernetes. https://www.digitalocean.com/products/kubernetes/
42. Ali M (2020) PyCaret: An Open Source, Low-code Machine Learning Library in Python. PyCaret version 1.0.0. https://www.pycaret.org
43. Alonso-Betanzos A, Bolón-Canedo V (2018) Big-data analysis, cluster analysis, and machine-learning approaches. Sex-Specif Anal Cardiovasc Funct 607–626
44. Frades I, Matthiesen R (2010) Overview on techniques in cluster analysis. Bioinf Methods Clin Res, 81–107
45. AWS (2022) Best Practice 12.1 - Decouple storage from compute. https://docs.aws.amazon.com/wellarchitected/latest/analytics-lens/best-practice-12-1.html
46. Palacio-Niño J-O, Berzal F (2019) Evaluation metrics for unsupervised learning algorithms. arxiv:1905.05667
47. Yellowbrick: Elbow Method. https://www.scikit-yb.org/en/latest/api/cluster/elbow.html
48. SIG: Study in Greece: The gateway to the Greek academic world. https://studyingreece.edu.gr/
49. Nginx: NGINX controller tech specs. https://tinyurl.com/3jt5rtyj
50. Nodejs: Node.js cross-platform JavaScript runtime environment. https://nodejs.org/en/
51. Postgresql: PostgreSQL: open source relational database. https://www.postgresql.org/
52. Redis: Redis: in-memory data store. https://redis.io/
53. Ubuntu: Ubuntu: Linux distribution. https://ubuntu.com/
54. Watada J, Roy A, Kadikar R, Pham H, Xu B (2019) Emerging trends, techniques and open issues of containerization: a review. IEEE Access 7:152443–152472