ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 3, No. 5, 235-247 (2018)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Towards an Efficient Federated Cloud Service Selection to Support Workflow Big Data Requirements

Mohamed Adel Serhani[1*], Hadeel Al Kassabi[2,1], Ikbal Taleb[2]

[1]*Information Systems and Security Department, College of Information Technology, 15551, UAE*

[2]*Concordia Institute for Information System Engineering, Concordia University, H3G 1M8, Canada*

ARTICLE INFO

Article history:
Received: 14 August, 2018
Accepted: 27 September, 2018
Online: 08 October, 2018

Keywords:
Big Data
Cloud
Service selection
Workflow
QoS

ABSTRACT

Cloud Computing is considered nowadays an attractive solution to serve the Big Data storage, processing, and analytics needs. Given the high complexity of Big Data workflows and their contingent requirements, a single cloud provider might not be able alone to satisfy these needs. A multitude of cloud providers that offer myriad of cloud services and resources can be selected. However, such selection is not straightforward since it has to deal with the scaling of Big Data requirements, and the dynamic cloud resources fluctuation. This work proposes a novel cloud service selection approach which evaluates Big Data requirements, matches them in real time to most suitable cloud services, after which suggests the best matching services satisfying various Big Data processing requests. Our proposed selection scheme is performed throughout three phases: 1) capture Big Data workflow requirements using a Big Data task profile and map these to a set of QoS attributes, and prioritize cloud service providers (CSPs) that best fulfil these requirements, 2) rely on the pool of selected providers by phase 1 to then choose the suitable cloud services from a single provider to satisfy the Big Data task requirements, and 3) implement multiple providers selection to better satisfy requirements of Big Data workflow composed of multiples tasks. To cope with the multi-criteria selection problem, we extended the Analytic Hierarchy Process (AHP) to better provide more accurate rankings. We develop a set of experimental scenarios to evaluate our 3-phase selection schemes while verifying key properties such as scalability and selection accuracy. We also compared our selection approach to well-known selection schemes in the literature. The obtained results demonstrate that our approach perform very well compared to the other approaches and efficiently select the most suitable cloud services that guarantee Big Data tasks and workflow QoS requirements.

1. Introduction

Cloud Computing is a promising potential venue for processing Big Data tasks as it provides on-demand resources for managing and delivering efficient computation, storage, and cost-effective services. However, managing and handling Big Data implicates many challenges across several levels, among which are the difficulty of handling the dynamicity of the environment resources, the dataflow control throughout the service compositions, and guaranteeing functional and performance quality. Therefore, abundant Cloud Service Providers (CSPs) offering comparable services and functionalities proliferate in the market to meet the growing challenging demands. Subsequently, the selection of the most appropriate cloud provider is recognized to be a challenging task for users. Not only appropriate in terms of functionality provisioned, but also satisfying properties required by the user such as specific levels of quality of service and reputation, especially with the exaggerated cloud providers' marketing claims of guaranteed QoS levels.

Hence, providing an automatic and modest means for selecting a cloud provider which will enable Big Data tasks and guarantee a high level of Quality of Cloud Service (QoCS) is a necessity. Moreover, modeling and evaluation of trust among competing cloud providers enables wider, safer and more efficient use of Cloud Computing.

Therefore, it is necessary to propose a comprehensive, adaptive and dynamic trust model to assess the cloud provider Quality of Service prior to making selection decisions.

*Mohamed Adel Serhani, Email: serhanim@uaeu.ac.ae

A large number of CSPs are available today. Most pf CSPs offers a myriad of services, for instance, Amazon Web Service (AWS) offers 674 varying services which are classified according to locality, Quality Of Service, and cost [1]. Automating the service selection to not only rely of simple criterion such as cost, availability, and processing power, but to consider service quality agreement is crucial. Current CPS selection approaches support straightforward monitoring schemes and do not provide a comprehensive ranking and selection mechanism. For instance, CloudHarmony [2] supports up-to-date benchmark results that do not consider the price while Cloudorado [3] supports price measurement, however neglects other dynamic QoS properties.

Selecting the best CSP reveals twofold objectives, and adds value to both CSPs and Big Data users as well as applications. CSPs provision services that attract clients' interest and support their processing and storage needs. However, users must ensure that services they were offered meet their expectation in terms of quality and price.

Difficulties linked to CSP selection to handle Big Data tasks include for example the following: 1) The limited support for Big Data users in describing their various QoS needs of different Big Data tasks. 2) The difficulty to search in a high dimensional database or repository of CSPs. 3) The challenge to consider the continuous variations in the QoS needs and the Big Data related requirements. And 4) The limited support for mapping Big Data task quality requirements to the underlying cloud services and resources quality characteristics. By doing so, we can guarantee an end-to-end quality support from the top-down Big Data quality consideration to cloud services and resources quality enforcement.

Our main objective in this work is to build a full-fledged approach that supports Big Data value chain with the best cloud services and resources that are trustworthy, automatically scale, and support complex and varying Big Data quality requirements. This is possible with the development of a comprehensive cloud services selection model that fulfills the needs of a Big Data job with the efficient supporting cloud services. Our solution will impose QoS of Big Data processes through dynamic provisioning of cloud services by one or multiple CSPs that will ensure high quality cloud services and fulfill crucial Big Data needs. we propose in this paper a selection approach which includes three phases as follows: our first selection scheme, eliminates CSPs that cannot support the QoS requirements of a Big Data job, which decreases the next selection stage search scope. Consecutively, our second selection stage extends the Analytic Hierarchy Process (AHP) approach to provide selection based on ranking cloud services using various attributes such as Big Data job characteristics, Big Data task profile (BDTP), Quality of Service and considering the continuous changes in cloud services and resources.

The third phase consists of selecting cloud services among different cloud providers, this happens mainly if none of the cloud providers can support the BDTP solely. In addition, if the Big Data job is possibly split into smaller jobs, during the three selection phases, our approach maps the upper quality requirements of the Big Data job to lower level matching quality characteristics of cloud services.

2. Related Work

Cloud service selection attracted the attention of researchers because of its crucial role in satisfying both the users' and providers' objectives having high quality service while optimizing resource allocation and costs. They proposed various approaches to handle and manage the cloud service selection problem. In this section we outline and classify these approaches and emphasize on their strengths and weaknesses.

A broker-based system is described in [4] where the authors proposed a multi-attribute negotiation to select services for the cloud consumer. The quality data is collected during predefined intervals and analyzed to detect any quality degradation, thus allowing the service provider to allocate additional resources if needed to satisfy the SLA requirements. Another broker-based framework was proposed to monitor SLAs of federated clouds [5] with monitored quality attributes measured periodically and checked against defined thresholds. Additionally, in [6], the authors proposed a centralized broker with a single portal for cloud services, CSP, and cloud service users. The authors in [7] proposed a distributed service composition framework for mobile applications. The framework is adaptive, context-aware and considers user's QoS preferences. However, this framework is not suitable for for cloud service selection due to heterogeneity and dynamicity nature of the cloud environments.

The authors in [8] proposed a broker–based cloud service selection framework which uses an ontology for web service semantic descriptions named OWL-S [9]. In this framework, services are ranked based on a defined scoring methodology. First, the services are described using logic-based rules expressing complex constraints to be matched to a group of broker services. Another service selection system was proposed in [10] where the authors proposed a declarative ontology-based recommendation system called '*CloudRecommender'* that maps the user requirements and service configuration. The objective of the system is to automate the service selection process, and a prototype was tested with real-world cloud providers Amazon, Azure, and GoGrid, which demonstrated the feasibility of the system.

In [11], a declarative web service composition system using tools to build state charts, data conversion rules, and provider selection policies was proposed. The system also facilitates translation of specifications to XML files to allow de-centralized service composition using peer-to-peer inter-connected software components. In addition, the authors in [12] proposed a storage service selection system based on an XML schema to describe the capabilities, such as features and performance.

Optimizing the performance is a significant issue in Cloud Computing environments. In other words, better resource consumption and enhanced application performance will be achieved when embracing the appropriate optimization techniques [13]. For example, minimizing the cost or maximizing one or more performance quality attributes. In [14], a formal model was proposed for cloud service selection where the objective is to not only the cost but also the risks (e.g., cost of coordination and cost of maintenance). In this evaluation, the model studies different cost factors, such as coordination, IT service, maintenance, and risk taking. Furthermore, the risks are denoted in terms of integrity, confidentiality, and availability.

The authors in [15] proposed a QoS-aware cloud service selection to provide SaaS developers with the optimized set of

composed services to attend multiple users having different QoS level requirements. They used cost, response time, availability, and throughput as different QoS attributes. The ranking of services is evaluated using integer programming, skyline, and a greedy algorithm providing a near-optimal solution.

Different optimization techniques were adopted for cloud service selection in the literature. One of which were proposed in [16], which used a probabilistic and Bayesian network model. The authors modeled the discovery of cloud service as a directed acyclic graph DAG to represent the various entities in the system. In [18], the authors model cloud service selection as a multi-objective p-median problem according to pre-defined optimization objectives. Their objectives are to optimize the QoS, the number of provisioned services, the service costs, and network transmission costs simultaneously in the given continuous periods. The model also supports the dynamic changing users' requirements over time. Similarly in [17], the authors suggested a service selection model based on combining fuzzy-set multiple attribute decision making and VIKOR. Nevertheless, the discrepancies among user requirements and the providers were not addressed.

The authors in [19] incorporated the IaaS, PaaS, and SaaS service subjective quality attributes based on user preference and applied fuzzy rules based on training samples for evaluation of cloud services quality. A resource management framework is proposed in [20] using a feedback fuzzy logic controller for QoS-based resource management to dynamically adapt to workload needs and abide by SLA constraints. Also, fuzzy logic was adopted in [21] to allow for a qualitative specification of elasticity rules in cloud-based software for autonomic resource provisioning during application execution. A CSP ranking model was proposed in [22] based on user experience, and service quality using an intuitionistic fuzzy group decision making for both quantifiable and non-quantifiable quality attributes to help users select the best CSP conferring to their requirements.

Another cloud service recommendation system was presented in [23] with a selection based on similarity and clustering according to user QoS requirements for SaaS, including cost, response time, availability, and throughput. The users are clustered according to their QoS requirements and are ranked based on multiple aggregation QoS utility functions. Their approach is composed of different phases, starting with clustering the customers and identifying the QoS features, then mapping them onto the QoS space of services, clustering the services, ranking them, and finally finding the solution of service composition using Mixed Integer Programming technology.

Additionally, Multiple Criteria Decision Making (MCDM) models and fuzzy synthetic decision were commonly used in combination for service selection. In [24], fuzzy synthetic decision was applied for selecting cloud providers taking into consideration user requirements. Furthermore, the authors in [25] adopted fuzzy-set theory to evaluate cloud providers trust based on quality attributes related to IaaS. Also in [26], the authors proposed a framework for QoS attributes-based cloud service ranking by applying AHP techniques. A case study was presented to evaluate their framework. Yet, this work was limited to using the measurable QoS attributes of CSMIC rather than including the non-measurable QoS criteria as well [17]. Other works used AHP approach for coud service selection, such as in [1], where the authors adopted MCDM method using AHP to select CPs based on real-time IaaS quality of service. Similarly, The authors in [27] distributed cloud resource management based on SLA and QoS attributes. They adopted AHP to cope with the cloud environment changes during the resource selection process. However, both works exhibit the limitation of only considering the QoS of the cloud services as their selection basis.

Web services frequently undergo dynamic changes in the environment such as overloaded resources. Hence, the authors in [28] proposed a multi-dimensional model, named AgFlow, for component services selection according to QoS requirements of price, availability, reliability, and reputation. The model optimizes the composite service QoS required by the user and revises the execution plan conforming with resource performance dynamic changes. The authors in [29] proposed an SLA renegotiation mechanism to support and maintain QoS requirements in cloud-based systems. They use historical monitoring information including service statuses such as availability, performance, and scalability to predict SLA violations.

Few existing cloud federation projects are based on brokering technologies for multi-cloud composed services. Hence, more research needs to be done towards a standardized methodology for handling interoperability and standard interfaces of interconnected clouds [30]. Trustworthiness evaluation models among different cloud providers were proposed and focus on a fully distributed reputation-based trust framework for federated Cloud Computing entities in cloud federation. In this model, trust values are distributed at each cloud allowing them to make service selection independently [31]. Trust modeling was also tackled in federated and interconnected cloud environments where both consumers and different cloud providers need to trust each other to cooperate [32].

The literature is missing a comprehensive selection model that incorporates all cloud service layers, dimensions, and components in a multi-dimensional model that satisfies service selection for such constrained Big Data applications. Additionally, among the several methods used to determine the user's QoS preference, none exhibits the flexibility to make it responsive to the user's point-of-view as well as comprehends the specific characteristics related to Big Data applications. Accordingly, service selection models are to take into consideration the subsequent requirements: 1) Transparency for stakeholders (such as, customers, CPs, and service brokers), 2) Simple interface that is user friendly, easy to, configure, control and integrate 3) Maintainable and self-adapting to service layers, such as, SaaS, IaaS, and PaaS, and 4) Require low communication overhead by using low number and lightweight messages between stakeholders.

We aim in this work to build a complete, flexible, and QoS driven solution to assess different CSPs' services' capabilities of handling various Big Data tasks. Hence, we develop a three-phase cloud service selection scheme that considers the task complexity and the dynamicity of cloud resource and services. The first step in the selection process consists of apprehending required Big Data quality of service, define and endorse these requirements using the proposed Big Data Task Profile (BDTP). It adopts three selection phases to assess in real-time the CPs QoS and their corresponding services and choose only those that match these requirements.

3. Big Data Task Profile

We explain in this section the main elements of our Big Data specification model as depicted in Figure 1. For every different Big Data task, we model the related profile categories. Additionally, we model a set of attributes and characteristics classifications for

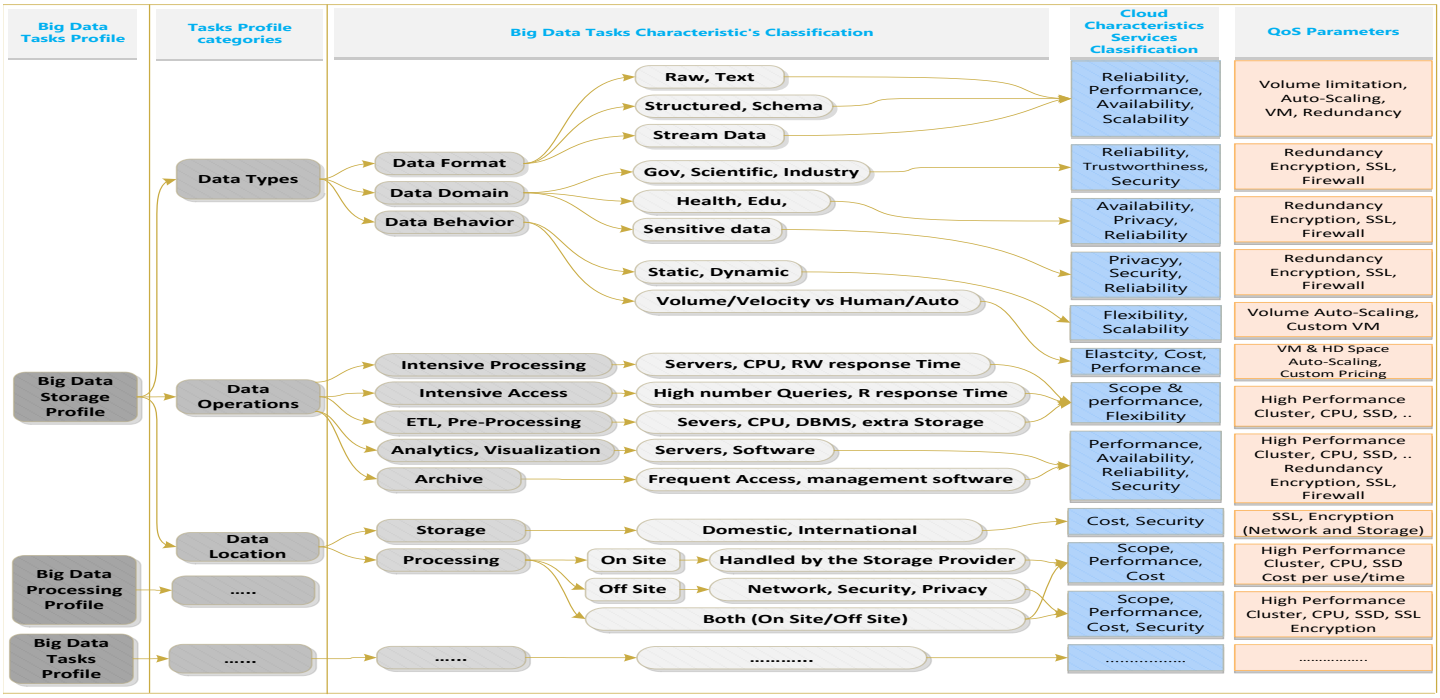


Figure 1. Profile-based Specification Model for BD Tasks

each category. Furthermore, we map the Big Data characteristics to its corresponding cloud attribute and services.

3.1. Big Data Task Profile (BDTP) Specification

The BDTP specifies the main Big Data task requirements that need to be satisfied, and it is modeled as a set of triples: ***R= {DT, DO, DL};*** where , ***DT*** refers to Data Type, and ***DO*** refers to Data Operation, and ***DL*** refers to Data Location. A Big Data request profiled based on BDTP, which defines the requirements and the most appropriate quality specifications that meet a certain Big Data task (such as, Big Data storage). For instance, Storage Profile specifies the following requirements:

Data Types Specifications:

a) Format: structured, unstructured, semi-structured, or stream data.
b) Domain: government, health, smart cities, etc.
c) Behavior: volume static vs dynamic scale, and velocity.

Stored Data Possible Operations:

a) Intensive processing
b) Intensive access
c) Extract Transform Load
d) Analytics and visualization.
e) Archive and backups only

Data Storage Location:

a) Storage Preference
   - Local cloud service provider
   - Geographically disperse site: this involves considering the following properties: network bandwidth, and security of data.

b) Data processing location:
   - On site: security and cost requirements (high or low).
   - Off site: network, security, cost, and servers requirements

Table 1: Sample Profiles for Big Data Tasks

TaBig Data Tasks	Related Cloud Services	Needed Resources	Cloud Services Classification	QoS Parameters
Generation and Collection	PaaS, DaaS	A, C	d, c, e	1, 2
Preprocessing	LaaS	A, B, C, D	a, c, e, f	1, 2, 3, 4
Processing	PaaS, SaaS	A, B, C, D	a, c, e, f	1, 2, 3, 4
Analytics	SaaS	C, D	i, f	1, 3, 4
Visualization	SaaS	C, D	i, f	1, 3, 4
Storage	DaaS	A, C	a, b, c, d, e	1, 2
Transport	LaaS	A	a, b, c, d, e	2
A. Networks B. Servers C. Storage D. Applications E. Infrastructure	1. Storage RW Speed (SSD/HDD) 2. Network Speed Mb/s & Latency 3. CPU Speed, Core, Count 4. RAM size		a. Performance b. Security c. Reliability d. Availability e. Scalability f. Transformation g. Hetriogeneity h. Privacy i. Governance	

Figure 2 illustrates the events issuing succession that deal with a Big Data request. Once a request is received, the best suitable BDTP is selected from the stored profile, in addition, the requirement is normalized to generate a profile *R*. Then the profile is linked with the user's quality of service requirement to produce an updated profile *R'* which will assist in the 3-phase selection. In the first selection stage we generate a list of CSPs ***CPList*** that is used for the second selection phase to generate another list of cloud services *CSList*.

The tipples ***R= {DT, DO, DL}*** represent the BDTP-based user requirements after mapping the appropriate BDTP profile to be

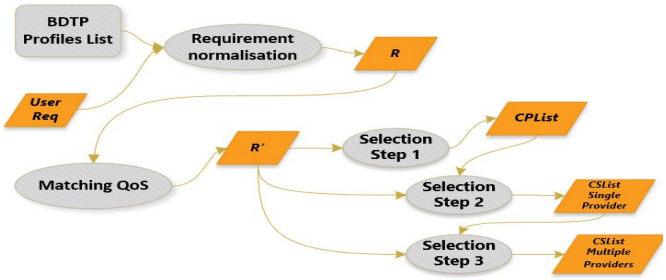


Figure 2. Two-step CSP selection based on BDTP

used. Afterwards, the profiled requirement R are translated into $\boldsymbol{R'= \{QoS_0, QoS_1,\ldots,QoS_i\}}$, for usage in the first selection stage. The later produces a list of CSPs fulfilling $\boldsymbol{R'}$ and noted as: $\boldsymbol{CPList = \{CP_1, CP_2,\ldots,CP_i\}}$. Moreover, the next selection stage based on cloud services is initiated on the same $\boldsymbol{R'}$ to produce a set of cloud services list noted as: $\boldsymbol{CSList = \{CS_1, CS_2,\ldots, CS_i\}}$. The third selection step, retrieves the *CSList* from the second selection phase and look for other cloud services from different providers that can satisfy the request.

3.2. Big Data Workflow Profile (BDTP) Specification

In this section, we describe a simple workflow applied in a case where a patient needs to be continuously monitored to predict epileptic seizures before they actually occur. The monitoring process involves placing multi-channel wireless sensors on the patient's scalp to record EEG signals and continuously stream the sensory data to a smartphone. This process does not restrict the patient's movements. The continuous recorded sensor data, such as 1 GB of data per hour of monitoring is considered a Big Data. However, smartphones lack the capabilities to handle this Big Data, whereas Cloud Computing technologies can efficiently enable acquiring, processing, analyzing, and visualization data generated form monitoring. Figure 3 describes the epilepsy monitoring workflow, where task t_1 is the data acquisition task that is responsible for collecting the EEG data is from the scalp by sensor electrodes then transfers the signals to be preprocessed to computing environment or to temporary storage t_2, which is storing the raw EEG signals. Task t_3 performs data cleansing and filtering processes to eliminate undesirable and noisy signals. Task t_4, is the data analysis task where the EEG data is analyzed to mine meaningful information to provision diagnosis and help decision-making. Finally, t_5 is the task responsible for storing the results.

In this workflow, a task is modeled as a tuple $t\langle tn, in, out\rangle$, where, tn is the task name and in and out are the input and the output data set respectively. Task dependency is modeled in $\boldsymbol{E} = \{(t_i, t_j) | t_i, t_j \in \boldsymbol{T}\}$, where t_j *is* dependent on t_i when t_j is invoked after the t_i is completed. The data flow is modeled by tracking the task input and output states. For each task t_i, we keep information about the data parameters, type and format.

3.3. Matching the BDTP to Cloud Service QoS

As we define $\boldsymbol{R= \{DT, DO, DL\}}$ to be a triple including Data Types, Data Operations and Data Location, we map each request's parameters from high level task specification to a low-level cloud

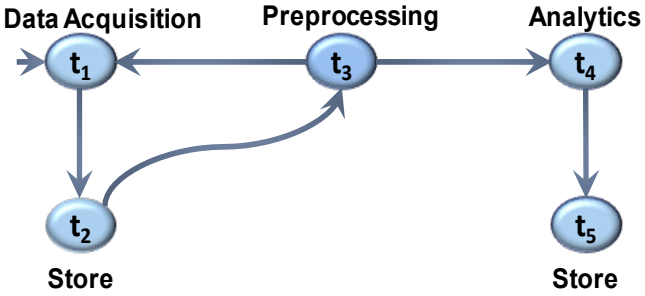


Figure 3. Workflow Example

service's QoS attributes having values and ranges that satisfy each requirement of the BDTP. For each selection phase, the matching process engenders a predefined profile. The QoS Profile is continuously revised to incorporate customer's request needs even after mapping and adjustments of quality attributes. Table 1 illustrates the matching scheme of Big Data tasks to cloud services QoS attributes.

3.4. Web-based Application for Collecting of Big Data Workflow QoS Requirements

In this section, we describe a web-based application we developed for collecting Big Data workflow QoS preferences from the user and generating a quality specification profile, which will be used as basis for task and workflow quality-based trust assessment as shown in Figure 4. This GUI application, collects the quality specification that illustrates the main requirements of a Big Data workflow and its composed tasks. Some of the workflow quality requirements are application domain, data type, operations and location. Furthermore, the application collects the required quality information for every composed task in the workflow, such as quality dimension, quality attributes and the weight values required for the overall trust score calculation. In addition, output data quality is specified for each task along with the weights preferred by the user. Finally, a complete workflow quality profile is generated that enumerates the most suitable requirements and specifications, which fits each Big Data task, such as Big Data preprocessing.

4. Cloud Service Selection Problem Formulation

One of the multi-criteria decision making methods is the Analytic Hierarchy Process (AHP) which is often used for such problems. It adopts a pairwise comparison approach that generates a preferences set mapped to different alternatives [33]. The advantage of AHP methodology is that it allows converting the subjective properties into objective measurements so they can be included in the decision-making, and hence permits the aggregation of numerical measurements and non-numerical evaluation. Additionally, it integrates the user's preference through getting the relative importance of the attributes (criteria) according to the user perception [1]. Accordingly, the quality attributes are represented as a hierarchal relationship, that matches the decision makers form of thinking [34]. Our recommended cloud service selection hierarchy is shown in Figure 5. This hierarchy clearly fits the mapping structure of Big Data to cloud services.

The AHP is intended to pairwise compare all different alternatives which are the quality attributes in our case. Therefore, the more quality attributes are considered, the larger the comparison matrix becomes and the higher number of comparison

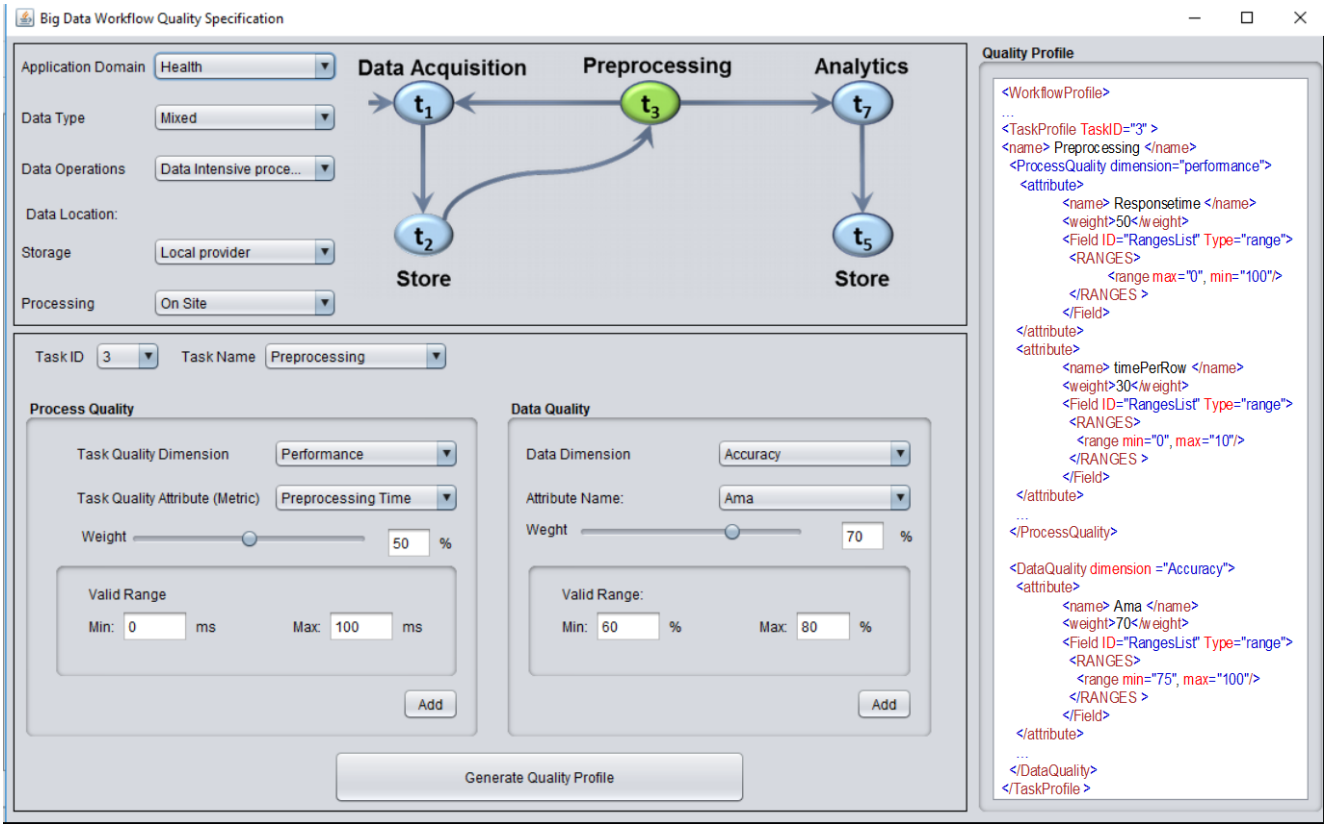


Figure 4. User interface for the collection of Big Data QoCS requirements

will be performed. Hence, we suggest to modify the original AHP approach as in [13].

The idea is to simplify the techniques to avoid the pairwise comparison by normalizing the quality attributes comparison matrix using geometric means which will decrease the required processing to reach a selection decision. Nevertheless, this modification will result in a converged weight matrix as a reason for adopting the geometric mean normalization and hence having a close attribute weight values. Eventually, the attribute priorities will diminish and will not satisfy the objective of this method. To solve this problem, we propose using the simple mean instead of geometric mean for normalization and calculating the attribute weights that matches the user priorities. We followed three steps in our selection approach given as:

4.1. Step1: Hierarchy Model Construction

We adopt the following definitions in our selection model [35]:

Definition 1: *The goal of decision problem which is the main objective and motivation. Here, the goal is the cloud service selection that best matches Big Data task profile conferring to the customer preference.*

Definition 2: *The alternatives which are represented with a set of various options open to the users to be considered in the decision. In our case, they are the group of available cloud services supplied by different cloud providers and matches the recommended BDTP.*

Definition 3: *The criteria of a decision problem. In this case, they are the quality attributes evaluated and upon which the comparison between alternatives is based on to eventually reach a decision. Specifically, they are the QoS attributes provided by the BDTP. The cloud services (alternatives) will be evaluated in comparison to the quality attributes (criteria) for measuring the matching level of the goal of the problem.*

The QoS attributes (criteria) for our decision-making problem are depicted in Figure 1 where they are quantified and qualified using the BDTP by assigning acceptance threshold values or ranges of values [35].

$$\mathrm{CP} = \{cp_i | i = 1,2,3, \dots n\} \tag{1}$$

$$\mathrm{S} = \{s_i | i = 1,2,3, \dots n\} \tag{2}$$

where $\forall\ s_i \in S$ is offered by one $cp_i \in CP$

$$\mathrm{A} = \{a_j | j = 1,2,3, \dots m\} \tag{3}$$

$$\mathrm{P} = \begin{bmatrix} p_{11} & \cdots & p_{n1} \\ \vdots & \ddots & \vdots \\ p_{1m} & \cdots & p_{\mathrm{nm}} \end{bmatrix} \tag{4}$$

where s_1, s_2 … s_n are the existing n alternative cloud services provided to the user. These services may be offered by various providers. a_1, a_2,..., a_m are the QoS attributes (criteria) from the BDTP mapped to the Big Data task required, for example: storage size, processing power, speed, availability, and reliability. p_{ij} is the performance of the i^{th} alternative s with respect to the j^{th} attribute.

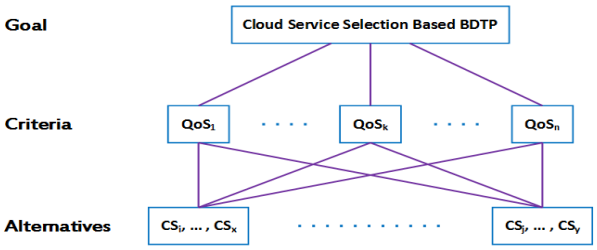


Figure 5. Cloud service selection hierarchy model

4.2. Step2: Attributes Weights and Ranking

AHP scheme consists of mapping each property to a rank or a priority level compared to other criteria applied in different evaluations. Then, an importance level is given by a user for each property opposed to all others [35]. This is performed after building a pairwise comparison matrix using a weighbridge of level of importance. An attribute can be compared to itself and the related importance is set to 1. Therefore, the matrix diagonals are all set to 1 [34]. The importance level is within the range between 1 to 9, where 1 refers to the lowest importance attribute and 9 refers to the most important attribute having the highest value.

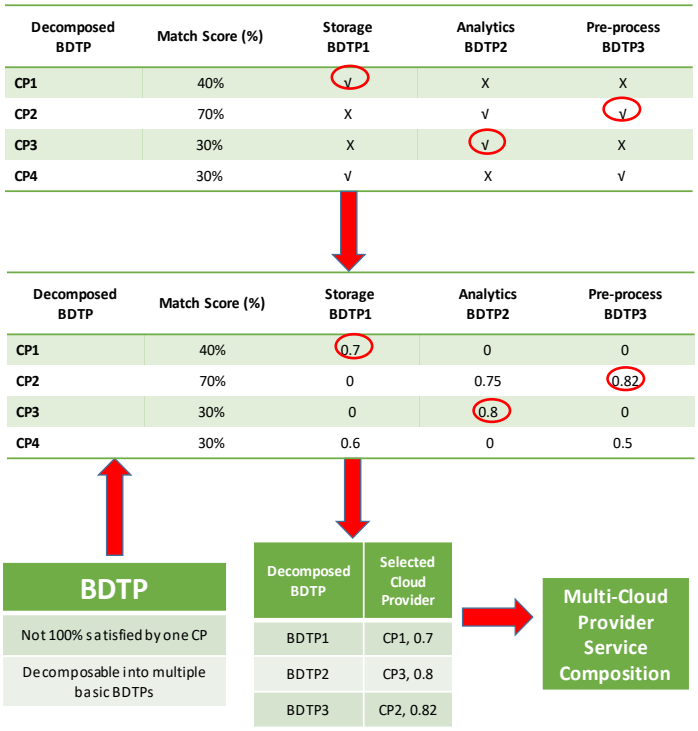
For m attributes, our pairwise comparison of attribute i with attribute j we get a square matrix $A_{M \times M}$ where r_{ij} designates the comparative importance of attribute i with respect to attribute j. This matrix has diagonal values assigned to 1. s.t. $r_{ij} = 1$ when $i = j$. Moreover, it contains reciprocal values across the diagonal, the ratio is inverted s.t. $r_{ji} = 1/r_{ij}$.

$$R = \begin{bmatrix} 1 & r_{12} & \dots r_{1n} \\ 1/r_{12} & 1 & \dots r_{2n} \\ \vdots & \vdots & 1 \end{bmatrix} \quad (5)$$

Then, we define a normalized weight w_i for each attribute based on the geometric mean of the i^{th} row. We choose the geometric mean methodology as an extended version of AHP for its simplicity, easiness of calculating the maximum Eigen value, and for decreasing the inconsistencies of judgment using $GM_i = \left(\prod_{j=1}^{m} r_{ij}\right)^{1/m}$ [34]. After that, the geometric means are normalized for all rows in the matrix using $w_i = GM_i / \sum_{i=1}^{m} GM_i$. Nevertheless, we get equal weights which disallow differentiation between attributes importance. Thus, we suggest to apply the normalized mean values for each row as follows:

$$M_i = \sum_{j=1}^{m} r_{ij} \quad (6)$$

$$w_i = M_i / \sum_{i=1}^{m} M_i \quad (7)$$



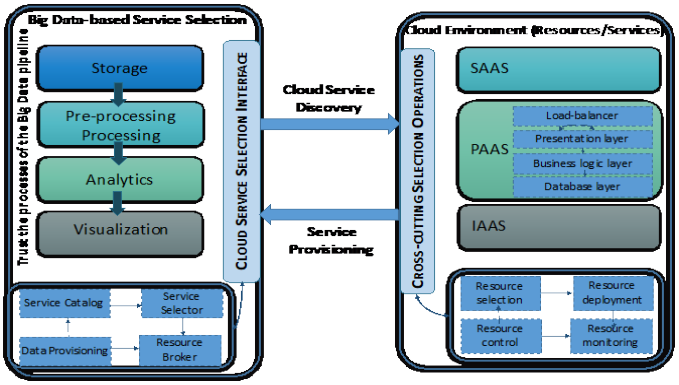


Figure 6. Big Data workflow quality enforcement based on cloud service selection

Table 2: Multi-provider cloud service selection m-pcss

4.1. Step 3: Calculate the Ranking Score of All Alternatives

To generate the rating scores for each cloud service (alternative), we use Simple Additive Weighting method by multiplying weights obtained from eq. 7 w_j of each attribute j with its corresponding performance value in Matrix P from eq. 4. Then summing all resulted values as in:

$$Score_i = \sum_{j=1}^{m} w_j \times (m_{ij})_{normal} \quad (8)$$

Where $(m_{ij})_{normal}$ is the normalized value of m_{ij} and $Score_i$ is the overall rating score of the alternative cloud service S_i. Finally, we select the cloud service (alternative) that has the highest score value:

$$S_{bestscore} = \max_{1 \le i \le n} Score_i \quad (9)$$

5. Model for Cloud Service Selection

We here describe our cloud service selection model to fulfill the quality of Big Data workflow over federated clouds. Figure 6 overviews how various Big Data processes, including storage, processing, and analytics can be provisioned with the cloud services and resources efficiently and with high quality. It details the main components involved in cloud service discovery and provisioning for Big Data value chain. Such components used for selection include service catalog, service broker, and service selector. However, components involved in cloud service provisioning in response to cloud service selection requests include resource selection, deployment, control, and monitoring.

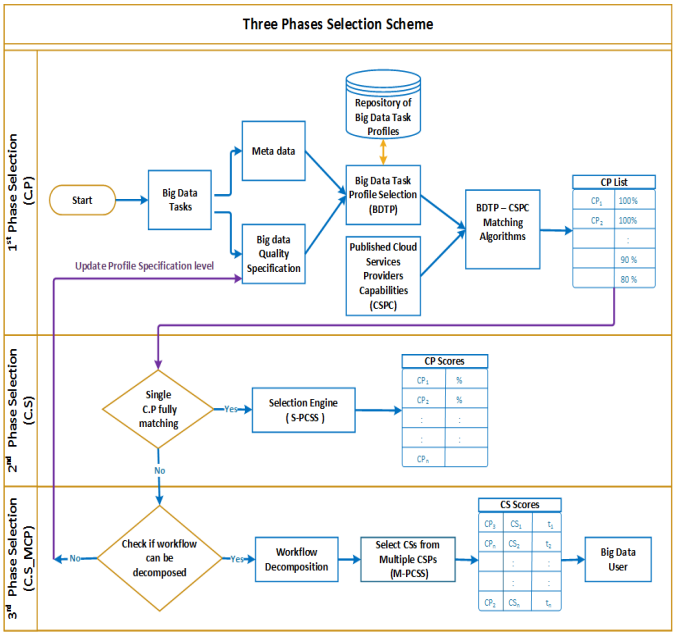


Figure 7. Three phases cloud service selection model

5.1. Cloud Service Selection

As soon as a service request is issued to support Big Data processing and storage while guaranteeing certain QoS, cloud resources are reserved to deploy and process Big Data workflow over the cloud infrastructure. Then, the workflow execution is monitored to detect if any performance degradation occurred and respond with the appropriate adaptation actions to maintain high quality service provisioning.

Figure 7, describes the selection scheme which is implemented in three phases: the first phase involves choosing the most suitable CSPs that conform to the Big Data workflow requirements, however the second phase involves choosing among CSPs the services that fulfill the Big Data Task profile (BDTP). The third phase selection consists of conducting further selection strategy to choose services from different CSPs that satisfy different tasks of a single workflow and maximize the overall quality of the workflow. In the following, we describe in detail each of the three selection phases:

CSP selection phase: Big Data workflows described as an aggregation of tasks present a set of quality requirements, such as, trust, in addition, to extra information known as metadata, such as, type of data, and its characteristics. The Big Data task profile selection component takes as input the metadata and the Big Data quality specification to find and retrieve the closest suitable profile from the Big Data profile repository that responds to the task(s) quality requirements. Both selected profile and published cloud provider's competencies are used to trigger the execution of the CP-Profile matching algorithm which matches the BDTP profile to the CSP published competencies. A list containing scored CSPs is generated by this algorithm. A score granted to each provider refers the ratio of which the CSP is capable to accomplish the Big Data task(s) given the set of quality requirements.

CS selection phase with single provider: the second selection phase is initiated to choose the corresponding cloud services from the list of phase 1 selected CSPs according to two stages:

Stage 1: A single provider cloud service selection algorithm (S_PCSS) is performed if a specific cloud provider completely matches the QoS of the Big Data task. The output of this algorithm is a list of CSPs with their measured scores. Here, we provide an extension of the AHP Method to use a simple mean instead of geometric mean to measure the attribute weight. This leads to variation in the generated weight values for each attribute that matches the pairwise importance levels given by the user.

Stage 2: A process of decomposing Big Data workflow into tasks is triggered if no single CSP is able to fulfil the QoS of the BDTP. Tasks of the workflow should be independent and can be processed impartially. If a workflow cannot be decomposed into undependably executable tasks, a loopback to previous phase will allow reviewing the profile specification to meet the selection measures.

CS selection phase with multiple providers: the third selection phase. Once a workflow can be decomposed into a set of tasks, the multi-provider cloud service selection algorithm is implemented to cope with multiple service selection from various cloud providers to maintain the quality of aggregated workflow tasks. Table II depicts an example of BDTP decomposition into three independent profiles for storage, pre-processing, and analytics. A score is calculated for each CSP with regards to each profile and cloud providers that have the highest score are selected to handle each profile independently.

5.2. Selection Algorithms

According to the scheme described in Figure 7, we have developed three consecutive algorithms to support the three phases selection as follows:

The BDTP-CSPC algorithm: maps the BDTP with each CSP Capabilities (CSPC), for example, availability and cost. The selection is performed according to the providers' capabilities satisfaction without considering customer favoured priorities. Figure 8 describes the algorithm which requires the list of CSPs, the list of required quality attributes (profile) and the list of published quality attributes for each cloud provider. Then performs one-to-one matching of each pair of attributes (profile-published) and outputs a list of scored CSPs which completely match the BDTP. Each CSP is linked to a set of provided quality characteristics. The algorithm performs an evaluation of each CSP matching score based on the percentage of fulfilled quality attributes required by the BDTP. The BDTP-CSPC matching

Algorithm 1 BDTP-CSPC matching algorithm

Input: $CPList$ //List of CPs,
$profileAttrList$ //List of required quality attributes (profile),
$attrListPerCP$ //List of published quality attributes of each cloud provider

Output: $cpScoresList$ //Matching Score for each CP

```
1: procedure CPMatchScore(CPList, attrListPerCP, profileAttrList)
2:     for all cp ∈ CPList do
3:         for all attr ∈ profileAttrList do
4:             cpAttrValue ← find(attr, attrListPerCP[cp])
5:             if (matchRequirement (cpAttrValue, attr))
6:                 count ← count + 1
7:             endif
8:             score[cpScoresList] ← count/size(profileAttrList)
9:         end for
10:    end for
11:    return cpScoresList[cp]
12: end procedure=0
```

Figure 8. BDTP-CSPC matching algorithm

Algorithm 2 Single Provider Cloud Service Selection (S_PCSS) algorithm

Input: *serviceList*, //List of CSs
profileAttrList, //List of required quality attributes (BDTP)
srvcAttrValue, //List of published quality attributes of each CS

Output: *srvcScoresList* //Matching Score for each CS

```
procedure EvaluateServiceAHP_S_CP(serviceList, profileAttrList, srvcAttrValue)
    R ← buildComparativeImportanceMatrix(BDTP, profileAttrList)
    M ← R.size()
    sumMN ← 0
    for s_i ← 1, M do // each service (row)
        MN_i ← 1
        for attr_j ← 1, M do // each attribute (col)
            MN[s_i] ← MN[s_i] + R[s_i][attr_j]
        end for
        sumMN = sumMN + MN[s_i]
    end for
    for i ← 1, M do // each service (row)
        weights[i] = MN[i]/sumMN // weight for each quality attribute
    end for
    for i ← 1, serviceList.size() do
        score_i ← 0
        for j ← 1, srvcAttrValue.size() do
            score_i ← score_i + srvcAttrValue[i][j] * weights[j]
        end for
        srvcScoresList[i] ← score_i
    end for
end procedure=0
```

Figure 9. Cloud Service Slection Algorithm - Single Provider

algorithm is scalable with the proliferation of cloud providers and quality attributes considered.

The S_PCSS algorithm: handles the second stage selection mechanism that considers thorough information about the attributes described in the BDTP to provide ranking values of the cloud services offered by the selected CSPs by the BDTP-CSPC algorithm. We adopted AHP and MADM to implement our selection strategy of cloud services. Figure 9 explains the single selection algorithm that uses a list of cloud services, the list of required quality attributes (BDTP), and the list of published quality attributes for each cloud service. Then, it generates a comparative matrix identifying the priority level of each published quality attribute in comparison to other quality attributes existing in the BDTP. Afterwards, this matrix is used to calculate and return a list of ranked cloud services with the highest scores and satisfy the Big Data task profile.

Algorithm 3 Multi-Provider Cloud Service Selection (M_PCSS) algorithm

Input: *cpList*, //List of CPs
serviceList, //List of CSs
profileAttrList, //List of required quality attributes (BDTP)
srvcAttrValue, //List of published quality attributes of each CS
cpScoresList, //List of CPs scores

Output: *srvcScoresList* //Matching Score for each CS

```
procedure EvaluateServiceAHP_M_CP(cpList, serviceList, profileAttrList, srvcAttrValue, cpScoresList)
    for cp_i ← 1, N do // each cloud service provider
        cpServiceScore[cp_i] ← EvaluateServiceAHP_S_CP(serviceList, profileAttrList, srvcAttrValue[cp_i]))
        for attr_j ← 1, M do
            cpServiceScore[cp_i] ← cpServiceScore[cp_i] × cpScoresList[cp_i]
        end for
    end for
    for attr_j ← 1, M do // each attribute (col)
        maxAttr ← 1
        for cp_i ← 2, N do // each cp (row)
            if (cpServiceScore[cp_i][attr_j] ≥ cpServiceScore[maxAttr][attr_j])
                maxAttr ← cp_i
        end for
        srvcScoresList[attr_j] ← tuple(maxAttr, cpServiceScore[maxAttr][attr_j
    end for
    return srvcScoresList
end procedure=0
```

Figure 10. Cloud Service Slection Algorithm - Multi- Provider

The M_PCSS algorithm: this algorithm handles the third stage selection where none of the CSPs fully supporting the Big Data workflow. In this situation, the workflow is decomposed into single independent tasks which will be processed by different cloud providers. Figure 10 describes the M_PCSS algorithm, the later takes as input the list of cloud providers, their offered cloud services and their calculated scores as well as the list of required quality attributes (BDTP), and the list of published quality attributes for each cloud service. It first applies the ***S_PCSS*** algorithm to receive the cloud service scores within each cloud provider. Then it finds the best matching services having the highest score among all cloud providers. Additionally, this algorithm favors the cloud provider that provides more services to minimize the communication and cost overhead due to data transfer and processing distribution. This is achieved by multiplying the cloud provider score to the service score to reach a final cloud service score.

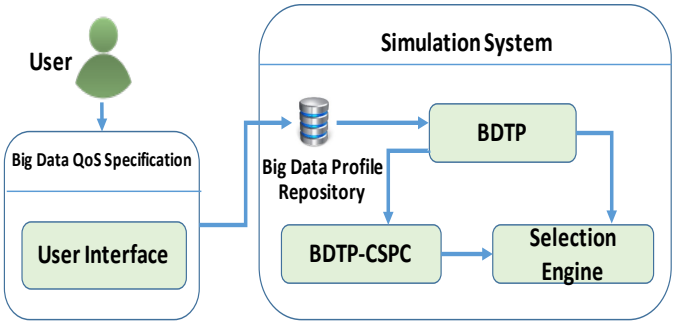


Figure 11. Simulator Components

6. Evaluation of Cloud Service Selection

This section details the experiments we conducted to assess the three-phase selection approach using various experimental scenarios.

6.1. Environment Setting

The setting and the simulation parameters we have used to conduct the experiments are described hereafter:

Setting and simulation parameters
Desktop: CPU Intel Core TM i7-3770K @ 3.40 GHz and Turbo Boost, DDR3 RAM 32GB, HD 1TB, and OS 64-bit.
Number of CSPs: 1 - 100.
Number of services provided by each CSP: 1 - 100.
QoS attributes: data size, distance, cost, response time, availability, and scalability.

6.2. Simulator

Figure 11 depicts the main modules of the JAVA simulator we have developed to implement the selection algorithms we have developed to support the three selection phases of cloud service providers and their related cloud services based on the BDTP and

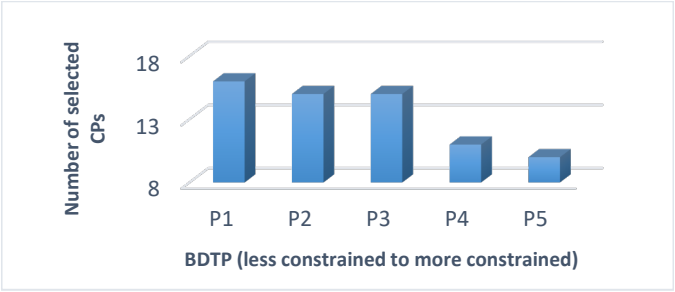


Figure 12. CSP profile-based matching with different constrains levels

the AHP method. The simulator comprises five main components as follows:

BDTP component: this module classifies the Big Data task requests into three categories: data type, data operation and data location. It also sets the acceptance level (minimum, maximum, threshold), for each quality property and eventually normalizes the performance scores.

BDTP-CSPC component: integrates the full implementation of BDTP-CSPC selection algorithm we described above. This module measures a score for each cloud provider that matches the BDTP. CSPs scoring 100% are nominated to the second phase selection Engine.

Selection Engine: integrates the implementation of the S_PCSS algorithm. The later uses the BDTP and the selected CSPs nominated in the first phase, then implements AHP to rank and retrieve the set of cloud services from the list of CSPs that fulfil Big Data task. Moreover, the selection engine implements the M_PCSS selection algorithm to incorporate the implementation of selecting cloud services from different CSPs while calculating cloud services scores for each cloud provider. Afterwards, it selects the best matching cloud service with the highest score among all cloud providers.

Big Data QoS specification: it supports and guides users through an interface to specify the Big Data task quality attributes as depicted in Figure 4 above.

Big Data profile repository: serves as repository of Big Data task profiles. It is accessed to retrieve the appropriate profile when a Big Data task request is issued and a selection of suitable CSP and services need to take place to respond to the initiated request.

In addition, to the above implemented entities, the simulator generates multiple CSPs offering multiple cloud services having various QoS attributes performance levels to produce a CSP list that serves the selection algorithms. Other implemented modules include, communication interfaces, scoring schemes implementation, invocation interfaces, and storage management interfaces.

6.3. Experimental Scenarios

In this sub-section, we detail the various scenarios we have chosen to assess our 3-phase selection model and the related implemented algorithms. Scenarios were selected to validate three main properties: CSP selection accuracy, model scalability, and communication overhead.

In the following, we explain the developed scenarios to help evaluating our 3-Phase selection model.

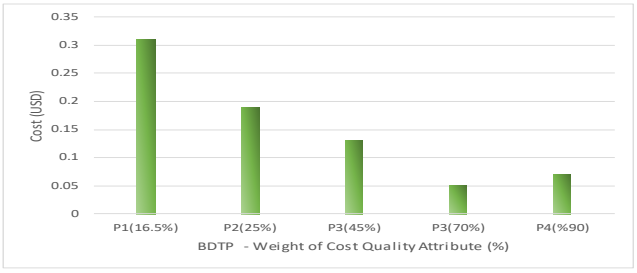


Figure 13. CSP selected with different levels of BDTP strictness (Cost)

Figure 14. CSP selected with different levels of BDTP strictness (Response Time)

Scenario 1: evaluates the accuracy of the the first phase selection in terms of retrieving different Big Data task profiles while fixing the number of cloud providers to 20 CSPs. Figure 12 demonstrates that the less the number of selected CSPs the more the BDTP becomes constrained (e.g. includes extensive quality constraint to consider and evaluate).

Scenario 2: evaluates the accuracy of the the second phase selection based AHP while varying profiles and fixing the number of cloud providers. This will also retroactively validate the first selection results. Figure 13, demonstrates that the more constrained the BDTP is, which will add more weight on the cost quality attribute, the more the recommended CS provides a better cost. In the same manner, Figure 14, stresses the same results but

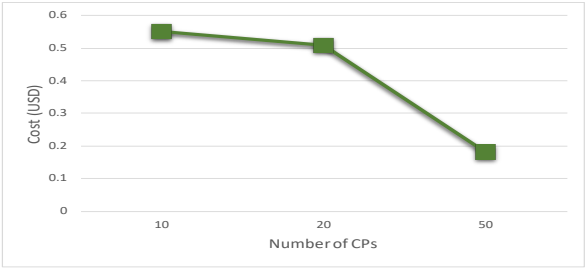


Figure 15. Cost of selected CSP with an increasing number of available CSPs

now with the response time quality attribute.

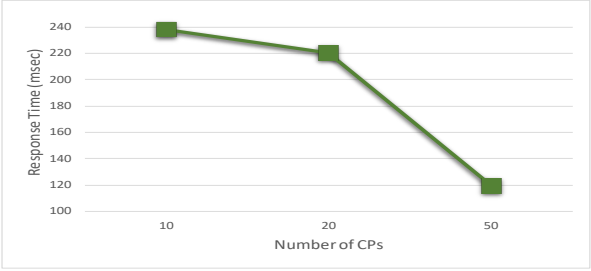


Figure 16. Response Time of selected CSP with an increasing number of available CSPs

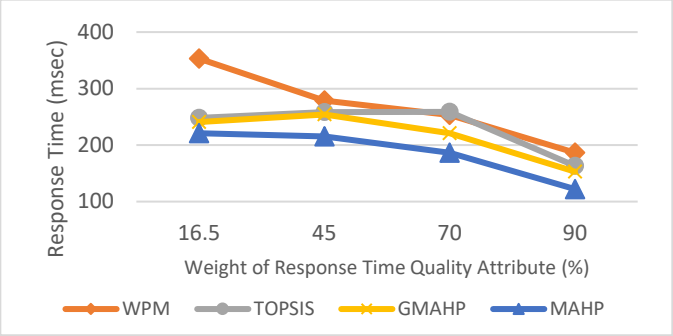


Figure 17. MAHP model behavior with different QoS attribute weight values benchmarked with other models.

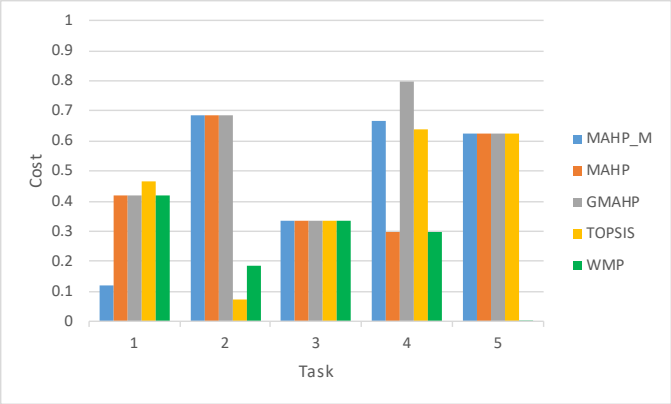


Figure 18. Cost of selected task per algorithm.

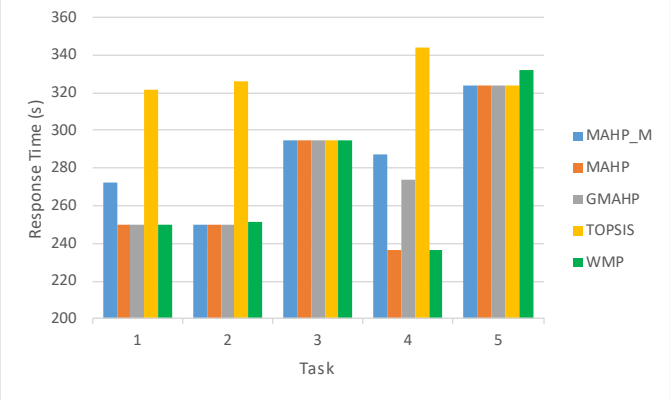


Figure 19. Response time of selected task per algorithm.

Scenario 3: evaluates the scalability of the selection model while increasing the number of CSPs and measuring the QoS cost and response time of the selected CSs after executing the 3-Phase selections respectively. Figure 15 and Figure 16, demonstrate that our 3-phase selection scheme scales perfectly as elucidated through a decrease in the cost and the response time respectively as the number of cloud providers increase. This is because more options are available to select among them which leads to better QoS fulfilment.

Scenario 4: we compare our 3-phase selection scheme with other MADM selection schemes. Our model used the *simplified* AHP using mean values of pairwise comparison matrix (MAHP). However, the other models used are Weighted Product Method (WPM), Technique for Order Preference by Similarity to Ideal

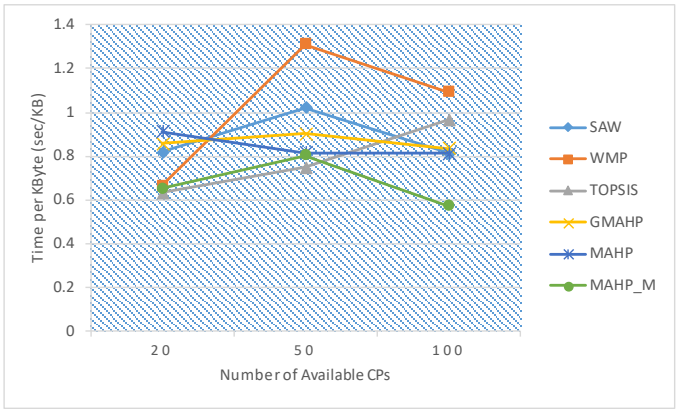


Figure 20. Average response time of selected services using (MAHP) and (MAHP_M) benchmarked with other models.

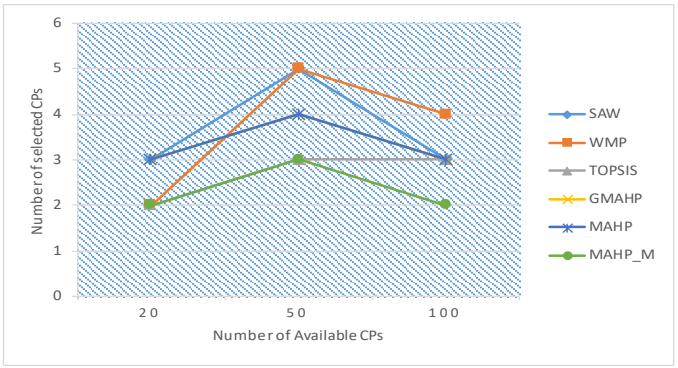


Figure 21. Number of differnet CSPs using (MAHP) and (MAHP_M) benchmarked with other models .

Solution (TOPSIS), and the modified Geometric Means AHP (GMAHP) to make cloud service selection. Comparison was conducted based on the response time quality attribute. Figure 17 demonstrates that MAHP gives better results compared to all other models, it provisions lower response times for all levels of selected quality attribute weights.

Scenario 5: we compare our 3-phase selection algorithm to other MADM selection methods by showing the cost and response time for each task composed in the workflow. As depicted in Figure 18 and Figure 19, the (MAHP) provisions lower task cost and response time respectively, and gives similar results as (GMAHP) and (TOPSIS). However, our *modified* AHP (MAHP_M) method provisions higher cost and response time per task than the (MAHP) since it gives higher preferences to selection of services from an existing cloud provider to minimize the communication and data transfer overhead.

Scenario 6: we compare the average response time of the workflow composed services using different selection schemes. The measured response time includes the communication and transfer overhead due to using services from different CSPs. As depicted in Figure 20, our (MAHP_M) method has the lowest response time because it minimizes the number of CSP among which services are selected, hence, minimizes the time wasted in communication overhead. In addition, Figure 21, shows the number of different CSPs providing the selected services and it shows that our (MAHP_M) method has the lowest number of CSPs and hence has the lowest overhead.

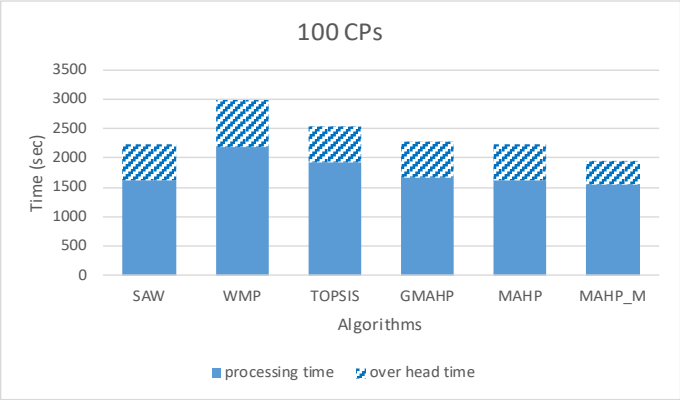


Figure 22. Workflow overhead and processing time of the MAHP and MAHP_M models benchmarked with other models.

Scenario 7: we compare the communication and data transfer overhead due to using different cloud providers. In this scenario, we used 100 CSPs and measured the total workflow execution time and the overhead time when using different selection methods. As shown in Figure 22, our (MAHP_M) method has the least overhead and accordingly total time amongst the rest of the methods. This is because our (MAHP_M) favors services that belong to already selected CSPs to minimize the overhead.

7. Conclusion

Big Data has emerged as a new paradigm for handling gigantic data and get valuable insights out of it. The special characteristics of Big Data reveals new requirements in terms of guaranteeing high performance and high quality of various Big Data processes (e.g. processing, storage, and analytics). The cloud infrastructure and resources are considered a perfect source of resources and services to support Big Data specific quality requirements. Selecting among myriad of cloud service providers the appropriate services and resources that meet these requirements is challenging given the diversity and the complexity of Big Data workflows.

In this paper, we proposed an efficient federated cloud service selection to support workflow Big Data requirements. It is a 3-phase selection scheme which is implemented through three phases. In the first selection phase, it captured the Big Data QoS requirements through the BDTP. However, in the second selection phase, a scored list of cloud services that satisfies the BDTP is generated. Finally, the third selection phase goes further and scored cloud services from different CSPs to better match the workflow quality requirements.

The main contributions of our selection scheme is the integration of a BDTP that ensures the QoS of Big Data tasks and is considered as a reference model for the three successive selection phases. In addition, revising the profile is advisable to have an efficient selection decision. We proposed a further contribution by extending the AHP method by adopting the mean values of pairwise comparison matrix alternative than using the geometric mean. The later shown weakness in producing a weight matrix with equal values of weights for all attributes. The last contribution is supporting workflow key requirements through the selection of multiple cloud services form multiple CSPs which maximized the Big Data complex workflow requirement fulfilment.

We conducted extensive experimentation to evaluate different properties of our 3-phase selection scheme. The results we have obtained proved that our selection model: integrated well the BDTP and guaranteed Big Data QoS requirements, scaled with the growing number of CSPs, performed better than the other MADM schemes such as TOPSIS, WPM, and the SAW, and enforced QoS requirement of Big Data workflows through varying cloud services from multiple CSPs.

For future work, we plan to have an extension for our selection scheme with more scenarios and complex Big Data workflows where other properties such as data security and privacy can also be considered. Furthermore, we are considering to assess our selection scheme against various selection techniques where we use an existing cloud environment.

References

[1] M. Zhang, R. Ranjan, M. Menzel, S. Nepal, P. Strazdins and L. Wang, "A Cloud Infrastructure Service Recommendation System for Optimizing Real-time QoS provisioning Constraints," IEEE Systems Journal, pp. 1-11, 2015. arXiv preprint arXiv:1504.01828

[2] "Cloudharmony," [Online]. Available: http://cloudharmony.com/. [Accessed 5 2017].

[3] "Cloud Computing Comparison Engine," [Online]. Available: http://www.cloudorado.com/. [Accessed 5 2017].

[4] E. Badidi, "A Cloud Service Broker for SLA-based SaaS provisioning," in International Conference on Information Society (i-Society), IEEE, 2013. https://doi.org/10.5121/ijccsa.2016.6201

[5] A. Al Falasi, M. A. Serhani and R. Dssouli, "A Model for Multi-levels SLA Monitoring in Federated," in IEEE 10th International Conference on Autonomic and Trusted Computing (UIC/ATC) Ubiquitous Intelligence and Computing, 2013. https://doi.org/10.1109/uic-atc.2013.14

[6] J. Lee, J. Kim, D.-J. Kang, N. Kim and S. Jung, "Cloud service broker portal: Main entry point for multi-cloud service providers and consumers," in IEEE 16th International Conference on Advanced Communication Technology, 2014. https://doi.org/10.1109/icact.2014.6779131

[7] Q. Z. Sheng, B. Benatallah, Z. Maamar, M. Dumas and A. H. Ngu, "Enabling personalized composition and adaptive provisioning of web services," in International Conference on Advanced Information Systems Engineering, 2004. https://doi.org/10.1007/978-3-540-25975-6_24

[8] L. D. Ngan and R. Kanagasabai, "Owl-s based semantic cloud service broker," in IEEE 19th International Conference on Web Services (ICWS), IEEE, 2012. https://doi.org/10.1109/icws.2012.103

[9] B. C. Grau, I. Horrocks, B. Motik, B. Parsia, P. Patel-Schneider and U. Sattler, "OWL 2: The next step for OWL," Web Semantics: Science, Services and Agents on the World Wide We, vol. 6, no. 4, pp. 309-322, 2008. https://doi.org/10.2139/ssrn.3199412

[10] M. Zhang, R. Ranjan, S. Nepal, M. Menzel and A. Haller, "A declarative recommender system for cloud infrastructure services selection," in International Conference on Grid Economics and Business Models, Springer, 2012. https://doi.org/10.1007/978-3-642-35194-5_8

[11] B. Benatallah, M. Dumas, Q. Z. Sheng and A. H. Ngu, "Declarative composition and peer-to-peer provisioning of dynamic web services," in 18th International Conference on Data Engineering. Proceedings., 2002. https://doi.org/10.1109/icde.2002.994738

[12] A. Ruiz-Alvarez and M. Humphrey, "An automated approach to cloud storage service selection," in the 2nd international workshop on Scientific cloud computing, ACM, 2011. https://doi.org/10.1145/1996109.1996117

[13] F. Pop and V. Cristea, "The Art of Scheduling for Big Data Science," in Big Data: Algorithms, Analytics, and Applications, Chapman and Hall, 2015, p. 105–120.

[14] B. Martens and F. Teuteberg, "Decision-making in cloud computing environments: A cost and risk based approach," Information Systems Frontiers, vol. 14, no. 4, pp. 871-893, 2012. https://doi.org/10.1007/s10796-011-9317-x

[15] Q. He, J. Han, Y. Yang, J. Grundy and H. Jin, "QoS-driven service selection for multi-tenant SaaS," in IEEE 5th international conference on Cloud Computing (Cloud), 2012. https://doi.org/10.1109/cloud.2012.125

[16] A. O. Akinwunmi, E. A. Olajubu and G. A. Aderounmu, "A trustworhty model for reliable cloud service discovery," International Journal of Computer Applications, vol. 87, no. 16, 2014. https://doi.org/10.5120/15293-3962

[17] H. M. Alabool and A. K. Mahmood, "Trust -Based Service Selection in Public Cloud Computing Using Fuzzy Modified VIKOR Method," Australian Journal of Basic and Applied Sciences, vol. 7, no. 9, pp. 211-220, 2013. ISSN 1991-8178

[18] H. K. Chen, C. Y. Lin and J. H. Chen, "A multi-objective evolutionary approach for cloud service provider selection problems with dynamic demands," in European Conference on the Applications of Evolutionary Computation, Springer. https://doi.org/10.1007/978-3-662-45523-4_68

[19] S. Telenyk, P. Bidyuk, E. Zharikov and M. Yasochka, "Assessment of cloud service provider quality metrics," in International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo), IEEE, 2017. https://doi.org/10.1109/ukrmico.2017.8095422

[20] C. Anglano, M. Canonico and M. Guazzone, "FC2Q: exploiting fuzzy control in server consolidation for cloud applications with SLA constraints," Concurrency and Computation: Practice and Experience, vol. 27, no. 17, pp. 4491-4514, 2015. https://doi.org/10.1002/cpe.3410

[21] P. Jamshidi, A. Ahmad and C. Pahl, "Autonomic resource provisioning for cloud-based software," in In Proceedings of the 9th International Symposium on Software Engineering for Adaptive and Self-Managing Systems, ACM, 2014. https://doi.org/10.1145/2593929.2593940

[22] S. S. Wagle, M. Guzek and P. Bouvry, "Cloud service providers ranking based on service delivery and consumer experience," in IEEE 4th International Conference in Cloud Networking (CloudNet), IEEE, 2015. https://doi.org/10.1109/cloudnet.2015.7335308

[23] Y. Wang, Q. He and Y. Yang, "Qos-aware service recommendation for multi-tenant saas on the cloud," in EEE International Conference on Services Computing (SCC), 2015. https://doi.org/10.1109/scc.2015.33

[24] S. Wang, Z. LiuQibo , S. Zou and F. Yang, "Towards an accurate evaluation of quality of cloud service in service-oriented cloud computing," Journal of Intelligent Manufacturing, vol. 25, no. 2, p. 283–291, 2014. https://doi.org/10.1007/s10845-012-0661-6

[25] M. Alhamad, T. Dillon and E. Chang, "A trust-evaluation metric for cloud applications," International Journal of Machine Learning and Computing, vol. 1, no. 4, pp. 416-421, 2011. https://doi.org/10.7763/ijmlc.2011.v1.62

[26] S. K. Garg, S. Versteeg and R. Buyya, "A framework for ranking of cloud computing services," Future Generation Computer Systems, vol. 29, no. 4, pp. 1012-1023, 2013. https://doi.org/10.1016/j.future.2012.06.006

[27] S. Khaddaj, "Cloud computing: service provisioning and user requirements," in 11th IEEE International Symposium on Distributed Computing and Applications to Business, Engineering & Science (DCABES), 2012. https://doi.org/10.1109/dcabes.2012.76

[28] L. Zeng, B. Benatallah, A. H. Ngu, M. Dumas, J. Kalagnanam and H. Chang, "Qos-aware middleware for web services composition," IEEE Transactions on software engineering, vol. 30, no. 5, pp. 311-327, 2004. https://doi.org/10.1109/tse.2004.11

[29] A. F. M. Hani, I. V. Paputungan and M. F. Hassan, "Renegotiation in service level agreement management for a cloud-based system," ACM Computing Surveys (CSUR), vol. 47, no. 3, p. 51, 2015. https://doi.org/10.1007/978-3-642-40861-8_9

[30] A. N. Toosi, R. N. Calheiros and R. Buyya, "Interconnected cloud computing environments: Challenges, taxonomy, and survey," ACM Computing Surveys (CSUR), vol. 47, no. 1, p. 7, 2014. https://doi.org/10.1145/2593512

[31] J. Abawajy, "Determining Service Trustworthiness in Intercloud Computing Environments," in the 10th International Symposium on Pervasive Systems, Algorithms, and Networks (ISPAN'09), 2009. https://doi.org/10.1109/i-span.2009.155

[32] D. Bernstein and D. Vij, "Intercloud security considerations," in IEEE Second International Conference on Cloud Computing Technology and Science (CloudCom), 2010. https://doi.org/10.1109/cloudcom.2010.82

[33] M. Whaiduzzaman, A. Gani, N. Badrul Anuar, M. Nazmul Haque and I. Tanzeena Haque, "Cloud service selection using multicriteria decision analysis," The Scientific World Journal, 2014. https://doi.org/10.1155/2014/459375

[34] R. V. Rao, "Improved multiple attribute decision making methods," Decision Making in Manufacturing Environment Using Graph Theory and Fuzzy Multiple Attribute Decision Making Methods, pp. 7-39, 2013. https://doi.org/10.1007/978-1-4471-4375-8_2

[35] S. Klutho, "Mathematical Decision Making: An Overview of the Analytic Hierarchy Process," 2013. [Online]. Available: www. whitman. edu/mathematics/SeniorProjectArchive/2013/Klutho. pdf. [Accessed 25 5 2017].