IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received October 23, 2019, accepted November 20, 2019, date of publication November 27, 2019,
date of current version December 12, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2956169

A Dynamic Virtual Datacenter Selection Strategy for Integrated Cloud Service Platform Construction With Multiclouds

BO HUANG [1], CHANGLAI DU [2], MENGTING SUN [3], AND XIAOQUN YUAN [3]

[1]Department of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
[2]Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA
[3]School of Information Management, Wuhan University, Wuhan 430072, China

Corresponding author: Xiaoqun Yuan (yuan20030308@whu.weu.cn)

This work was supported in part by the National Science Foundation of China under Agreement 71673208, and in part by the Nanjing Qixia District's Science and Technology Plan under Agreement GC201805.

ABSTRACT The pay-as-you-go model and network virtualization of cloud computing allow micro and small content businesses (MSCBs) who construct their integrated cloud service platforms (ICSPs) with virtual datacenters (VDCs) to serve their end users with low service latency and construction cost. However, designing a flexible VDC selection strategy to meet the demands of MSCBs is a challenging task. To address this problem, a dynamic VDC selection strategy is designed for MSCBs to construct their ICSPs flexibly with the VDCs from different clouds. To this end, three dynamic landmark selection metrics are proposed and applied to construct the network coordinates. Then, a dynamic VDC selection algorithm is presented to determine the locations and service resources of VDCs, which are purchased from different clouds by MSCBs to construct their ICSPs. Based on our VDC selection strategy, a simulator is developed based on our designed experimental framework to evaluate our VDC selection strategy. The experimental results show that compared with previous server placement strategies, our strategy can actively and effectively determine VDCs' locations and allocate service resources for each VDC with less computing consumption and is a practical VDC selection strategy for cloud construction in a multicloud environment.

INDEX TERMS Cloud construction, geo-distributed clouds, virtual datacenter selection, resource allocation, service node location.

I. INTRODUCTION

Cloud computing has transformed the entire Internet service industry due to its market-oriented resource allocation model [1], [2], which significantly reduces Internet service providers' cost to construct their cloud-based service platforms for their Internet services, especially for many latency-sensitive Internet services, such as news, video and publishing services. Due to the large scale of their end users, these content service providers, especially micro and small content businesses (MSCBs), always construct their service platforms by implementing the pay-as-you-go model and network virtualization technologies to overcome the shortage of funding and provide low latency service to their end users [3]. To this end, they purchase resources from other geo-distributed clouds, named virtual datacenters (VDCs) in this paper, which have been created as virtual machine clusters (e.g., Amazon Virtual Cluster)[1] by these cloud providers [4]. These VDCs are integrated into MSCBs' service platforms whose services are deployed by applying network virtualization technologies. Then, suitable virtual machines (VMs) of their VDCs near end users are used to respond to requests. Obviously, VDC selection is the first step and the pillar for MSCBs in constructing their virtual service platforms, named integrated cloud service platforms (ICSPs), and it directly and significantly affects the ICSPs' construction costs and their content service delays. Thus, an efficient VDC selection strategy is crucial to achieving ICSP low latency content service and the success of investment in ICSP construction.

However, most data center selection works have focused on the data center allocation of existing geo-distributed data centers and data center placement for large scale

The associate editor coordinating the review of this manuscript and approving it for publication was Ibrar Yaqoob.

[1]http://docs.aws.amazon.com/redshift/latest/mgmt/managing-clusters-vpc.html, online; 2019-11-20

networks [5]–[7]. The former aims to improve the service performance of geo-distributed data centers, such as reducing response time and balancing a load of different data centers by choosing suitable data centers from one service platform to serve end users [5], [8]. The general method formulates the data center selection as different optimal problems with the aim of selecting M service nodes from N potential sites, which means that servers are selected from a candidate node pool. However, VDC selection aims to construct ICSPs, which means that there is no existing candidate node pool. In addition, due to the large scale of networks and the rapid development of mobile terminals, VDC selection should locate data centers anywhere they need to serve mobile terminals. Obviously, these data center allocation models cannot yield their expected performances for the scenario of VDC selection in ICSP construction. In contrast to these works, the latter aims to construct geo-distributed data center-based service platforms for large-scale network services. Most works implement clustering technologies to locate the sites of data centers and actively allocate the resources for each located site [9]. By implementing clustering technologies, they can locate the selected data center anywhere to guarantee their QoS, which matches the large-scale geo-distribution of end users. However, these solutions aim to construct their physical data center while the resources of VDCs are purchased from other geo-distributed clouds. Therefore, these solutions cannot yield their expected results for VDC selection of ICSP construction.

Obviously, due to the large-scale geo-distribution and mobility of end users and the financial restraint of MSCB, previous works on data center selection are not suitable for the scenario of VDC selection of ICSP construction, and there are challenges in addressing this problem. One obvious challenge lies in Internet information achievement, which aims to estimate the network delays of different Internet devices, such as end users and VDCs. Generally, both geographic coordinates and network coordinates can be used to estimate the network delays, but neither of them meets the requirement of effective VDC search due to the inaccuracy of geographic coordinates and the time complexity of network coordinate construction [10]. Another challenge is how to design a lightweight VDC selection algorithm for MSCBs to actively construct their ICSPs. To make more profit with limited capital, MSCBs tend to adapt their selected VDCs according to the change of end users and candidate clouds, which means that MSCBs should choose suitable candidate clouds pervasively with low time complexity. Unfortunately, it is difficult for existing server placement strategies to do so. First, approximate server placement algorithms for facility location achieve their server placement strategies passively. Regarding the cluster-based algorithms, although these algorithms can actively identify suitable ICSPs' VDCs, their time complexity or the inaccuracy of the generated server placement strategies cannot satisfy the VDC selection of ICSPs' construction.

To address these challenges, a practical VDC selection strategy is proposed for ICSP construction in a multicloud environment in this paper, where the server placement framework of NetClust is applied to reduce the measurement cost and locate VDCs pervasively. To reduce the algorithm complexity of the GNP-based network coordinate construction, three landmark selection metrics are first presented. These metrics may help to reduce the computing consumption sharply with the excepted estimation accuracy of network coordinate construction. Then, a fast VDC selection algorithm is designed to reduce the time complexity of NetClust's k-means clustering-based server selection algorithm. Based on our VDC selection strategy and ICSP framework, a simulation framework for VDC selection is designed, and the corresponding simulator is developed, which is used to imitate the cloud selection process of ICSP construction and verify the effectiveness of our VDC selection strategy. The experimental results show that our VDC selection strategy can significantly reduce the computing resource consumption and locate end users to "optimal" VDCs. Thus, it is a practical VDC selection strategy for the scenario where MSCBs construct their cloud service platform with resources from other clouds to serve their end users. The major contributions of this paper are summarized as follows.

1) Three landmark selection metrics are presented and used to construct low-dimensional GNP-based network coordinates to gain global network information with low computing consumption. Compared with the network coordinate construction method of NetClust, our low-dimensional network coordinates construction method reduces the computation consumption sharply without degrading estimation accuracy. In addition, it is also helpful to significantly save the storage resources of ISCP construction for MSCBs.

2) A fast VDC selection algorithm is proposed to achieve a VDC selection strategy for our ICSP in a geo-distributed cloud environment. Our VDC selection algorithm integrates the hierarchical clustering algorithm and the fast density peaks finding algorithm, which can select the suitable VDCs from clouds and efficiently allocate the appropriate network resources for each selected VDC.

3) A simulation framework for the VDC selection strategy is designed, and the corresponding simulator is developed to verify the performance of our VDC selection strategy with real-world data from Zhejiang Publishing United Group.[2] The results show that 1) the network coordinates constructed based on our landmark selection metrics achieve lower computational and network probing costs with assured estimation accuracy, and 2) the VDC selection strategy significantly reduces the computing consumption compared with existing cluster-based server placement strategies.

The rest of this paper is organized as follows. Section II reviews related works on resource allocation of cloud computing and server placement. Section III presents the system model of ICSPs and our practical VDC selection strategy. Section IV first investigates the details of our VDC selection strategy: the landmark selection for network coordinate

[2]http://www.zjcb.com/, online; 2019-11-20

construction and the fast VDC selection algorithm. Then, our practical VDC selection strategy with the corresponding algorithm is proposed. Section V evaluates the performance of the proposed algorithm via simulation experiments. Finally, Section VI concludes this paper.

II. RELATED WORKS

This section describes the preliminaries and related works on resource allocation of cloud computing and service node placement.

A. RESOURCE ALLOCATION OF CLOUD COMPUTING

Essentially, data center selection in existing service platforms aims to optimize the service performance of existing service platforms, such as balancing their service load and maximizing their service profits by choosing a suitable data center to serve end users [11], [12]. Therefore, it can be classed to resource allocation of cloud computing. In fact, resource allocation is a popular topic for cloud computing, and many significant works have focused on it.

X. Wang et al. proposed a cloud service selection model adopting cloud service brokers to assist users in efficiently selecting their preferred cloud services. Then, a dynamic cloud service selection strategy was proposed based on this model [13]. R. Liu et al. proposed a mobility-aware framework named MACSS for mobile cloud streaming services. This framework provided dynamic and optimized server selection functions to support user mobility [14]. A. Selvaraj et al. presented an optimizing VM selection solution called analogous particle swarm optimization (APSO) under a cloud computing environment, which implemented a swarm intelligence approach [15]. A. Ponraj proposed a VM placement algorithm to minimize the computing time and data transferring time, which considered computation resources, quality of service (QoS) metrics, virtual machine status, and I/O data with priority-based probability queuing model [16]. W. Shi et al. proposed an efficient online auction mechanism to address virtual cluster (VC) allocation and designed a novel online algorithm for dynamic VC provisioning and pricing, achieving truthfulness, individual rationality, and computational efficiency in social welfare [4]. S. B. Melhem et al. proposed a host load detection algorithm to find the future overutilized/underutilized host states to avoid immediate VM migration. Then, they proposed a VM placement algorithm to determine the set of candidate hosts to receive migrated VMs to reduce their VM migrations in the near future [17]. Xu *et al.* [18] developed a lightweight interference-aware VM live migration strategy based on the relationship between VM performance interference and key factors of workloads. Y. Liu et al. focused on VM placement and proposed a large number of scheduling algorithms for VM placement to balance the performance and cost of data center [19]. Tian *et al.* [20] presented a live VM allocation algorithm and a lightweight simulation system for the algorithm.

These works cover all aspects of cloud computing services such as VM provision, VM migration and VM clustering among datacenters, and routing algorithms between datacenters and end users. The goal of these works was to achieve resource utilization and to make more profit for cloud providers. All of them were based on the hypothesis that the cloud systems have been constructed and could serve end users, which is not suitable for the scenarios of the VDC selection that aims to construct ICSPs.

B. SERVICE NODE PLACEMENT

In fact, service node placement is always a topic of distributed systems and has been widely concerned, from mirror site selection to edge service node placement [21]–[23]. Among them, S. Jamin et al. studied the performances of different placement strategies and found that increasing the number of mirror sites was effective in reducing client download time and server load [21]. K. Xu et al. proposed joint optimizing approaches for replica server placement, content caching in selected servers, and content request load assignment among the servers to minimize the ratio of unserved content request load when the network resources and server capacity were both limited [22]. X. Yuan et al. proposed a server placement model for peer-to-peer (P2P) live streaming systems that considered the Internet service provider friendship and peers' contribution [24]. H. Yin et al. focused on the server placement of large-scale streaming systems and designed a cluster-based server placement framework to deploy suitable servers [9]. C. Dong et al. also focused on the service node placement of live streaming and proposed an online algorithm to save operational costs for CSLSPs by jointly and dynamically choosing the correct data centers for broadcasters and viewers [7]. Y. Zhang and M. Tatipamula focused on server placement for social networks and proposed three scalable server placement strategies to select server locations among all possible locations with less cost [25]. Q. Zhang et al. presented a framework for dynamic service placement based on control- and game-theoretic models and proposed a coordination mechanism to maximize the social welfare of the system [26]. T. Do and Y. Kim proposed an optimization model and topology-aware resource-efficient placement algorithm (TARE), which considered the different requirements of various high availability clusters and could be employed to optimally deploy high availability clusters with different redundancy configurations over geo-distributed cloud infrastructures [27]. H. Xiang et al. focused on server placement and proposed a cluster-based flexible server placement to allocate resources for edge service nodes [10].

Although the methods of these works are different from each other, most of them can be classified into two categories. One formulates the service node placement as a facility location problem with the aim of selecting M service nodes from N potential sites [21], [22], [24]. It passively selects its service nodes from a candidate node pool, constraining its scalability by the candidate node pool. Thus, it is suitable for small-scale service systems or service systems with fixed distributed end users, e.g., wired end users. To address this problem, clustering-based solutions have been proposed to select and deploy service nodes [9], [10]. By applying clustering technologies, they can locate service nodes anywhere. Therefore, this active service node placement solution is suitable for large-scale service platform construction, such as the platform construction of data center-based clouds and large-scale live streaming. However, due to the financial constraint, most MSCBs attempt to construct their ICSPs with fewer costs. In addition, the popularization of mobile end users makes service node placement locate its service nodes at arbitrary sites. Obviously, although these clustering-based service node placement solutions can locate service nodes at arbitrary sites, it is difficult for them to dynamically adjust their placement strategies according to the change in mobile users. Thus, it is necessary to design a lightweight dynamic service node placement strategy to satisfy these features of ICSP construction for MSCBs.

In addition to optimizing the service performance of existing service performance, data center selection is also applied to construct service platforms, which is similar to the service node placement in distributed systems.

III. OVERVIEW OF INTEGRATED CLOUD SERVICE PLATFORM

MSCBs construct their ICSPs by integrating the resources of VDCs and use "near" VDCs to serve end users with high service quality, as shown in Fig. 1. An ICSP consists of three major components: 1) *VDCs*; 2) *Controller* and 3) *end users*. VDCs, selected from geo-distributed clouds based on the end user requirements, are the resources of ICSPs and provide end users with Internet service, which consists of all kinds of resources from the selected clouds. In each VDC, many VMs may be deployed according to the requirement of ICSPs. *Controller* is the central control unit of ICSPs, which responds to construct ICSPs, and manages and monitors the operation of ICSPs. *End users* obtain all types of Internet services from ICSPs. To integrate VDC resources into ICSPs and provide services to end users, *controller* determines whether the ICSP construction should (re)start according to the constraints of Eq. (a) described in Subsection IV. A. Then it applies the VDC selection strategy to select VDCs from geo-distributed candidate clouds and allocates appropriate resources to these VDCs based on the information of end users, such as the number of end users. Finally, the resources of VDCs are integrated into the ICSP, which manages and schedules to meet the requirements of end users.

VDC selection is the core process of ICSP construction and directly influences the service performance of ICSPs. For example, in Fig. 1, the strategy in which Cloud 1, Cloud 2 and Cloud 3 are chosen to construct ICSP could achieve the minimal service delay when the unit resource of all clouds is at the same price because these VDCs are nearest to end users. Thus, it is necessary for VDC selection strategy design to use global Internet information, such as the location of each end user and candidate clouds. Furthermore, limited capital requires resource consumption and flexibility to be considered when a VDC selection strategy is designed. To address these challenges, we take advantage of the geo-distributed feature of candidate clouds and use them as probes to collect the end users' information, which means that MSCBs should reach an agreement with their candidate geo-distributed clouds to use their resources to collect the end users' information whenever necessary. That is, once the *Controller* determines to (re)start the ICSP construction, it applies VDC selection strategies to select VDCs from geo-distributed candidate clouds and allocates appropriate resources for these VDCs based on the information of end users, such as the number of end users. Our VDC selection framework is designed, as shown in Fig. 2. It consists of four components: *end users*, *VDCs*, *candidate geo-distributed clouds* and a *controller*, as shown in Fig. 2(a). Among them, end users are the consumers of ICSPs, and VDCs are the virtual data centers to construct ICSPs. Of course, these VDCs are selected from candidate geo-distributed clouds according to the VDC selection strategy generated by the controller. Therefore, candidate geo-distributed clouds are the resource source through which MSCBs construct their ICSPs, and controller is the core component where the VDC selection strategy is generated.

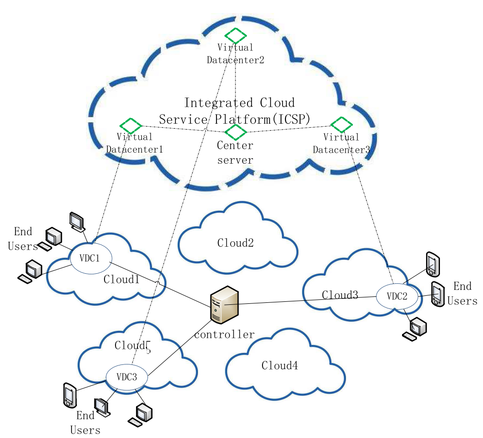
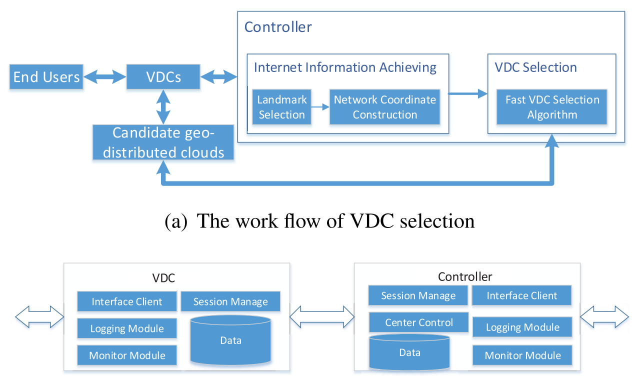


FIGURE 1. The system of integrated cloud service platform.



(a) The work flow of VDC selection

(b) The interaction of VDCs and controller

FIGURE 2. The framework of VDC selection.

Fig. 2(a) also shows that the VDC selection strategy is a dynamic process in which *VDCs* and the *controller* should cooperate with each other. In fact, they cooperate through their Internet interface: the client-side interface running in *VDC*, and the server-side interface running in the *controller*, as shown in Fig. 2(b). In Fig. 2(b), the *controller* includes a session management module, a central control module, a monitor module, a logging module and an interface server, while the *VDC* includes a session management module, a logging module, a monitor module and an interface client. In the *controller*, the session management module manages the operation of other modules. The central control module implements the function of network coordinate construction and VDC selection. The monitor module processes the information stored in the data and determines whether our VDC selection strategy should be started based on the collected information. The interface server maintains the connection to the *VDC* sides. The logging module collects the information of ICSP operation and the network environment conditions, preprocesses it and stores it into the data. Similar to the *controller* side, the session management module of the *VDC* manages and schedules the operation of the other modules. The monitor module processes the information stored in the data and determines the operation condition. The logging module collects the information of ping measurement results, and the interface client maintains the connection between the *VDC* sides and the end users. Based on this framework, the workflow of our VDC selection can be described as shown in Fig. 2(a): when the VDC selection strategy is triggered by some condition, such as the overload of VDC(s) and(or) poor QoS, the monitor module of *controller* sends this message to the center control module. Then, the lists of candidate VDCs' IPs are sent to all candidate geo-distributed clouds through the interface server. Once the list is received, each VDC measures the delays to all destinations and feeds the result back to the *controller* through its interface client to select the landmarks for the network coordinate system by applying the proposed landmark selection metrics. Then, the IP lists of all end users and candidate geo-distributed clouds are sent to these landmarks to gain their measurement data for the network coordinate system construction. This measurement data is fed back to the *controller* and mapped to the constructed network coordinate system to construct the network coordinates of all end users and VDCs, through which the global Internet information is achieved. Finally, the center control module uses our fast VDC selection algorithm to the constructed network coordinate to determine the locations of selected VDCs and allocate the resources for each selected VDC.

IV. VDCs SELECTION STRATEGY

This section aims to describe our VDC selection strategy in detail. To this end, the VDC selection is first formulated as an optimization problem with the aim of minimizing the placement cost and the service delay. Then, key techniques, that is, obtaining Internet information and the VDC selection algorithm, which affect the performance of the VDC selection strategy, are introduced.

A. PROBLEM FORMULATION

MSCBs apply a VDC selection strategy to determine the location and resource requirement of each VDC and use these VDSc to construct their ICSPs to serve their end users with guaranteed service quality, as shown in Fig. 1. Therefore, the ICSP' QoS is the key factor to be considered in the VDC selection strategy design.

As is known, service delay, which is always measured by TTL, is the main metric of many applications, especially for these time-sensitive applications, e.g., live streaming and online games. Thus, in this paper, TTL is used as the quality of the ICSP service. Additionally, because these resources are purchased from the corresponding candidate geo-distributed clouds, most MSCBs have limited money for constructing their ISCPs. Thus, the deployment cost can be considered as another metric of VDC placement.

Consider a space V consisting of all end users and candidate geo-distributed clouds, and both end users and candidate geo-distributed clouds are considered as points in V, and the number of end users is N. Then, the VDC placement of ICSPs can be formulated as a clustering-based multiconstrained optimization problem.

$$min: k_1 \sum_{i=1}^{N}\sum_{j=1}^{N} r_j y_{ij} c_i x_i + k_2 \sum_{i=1}^{N}\sum_{j=1}^{N} L_{ij} y_{ij}$$
$$s.t. \begin{cases} x_i, yij \in \{1, 1\} \\ U_{ij} = 2 * u_{ij} \\ U_{ij} \geq D_{thr} \end{cases} \tag{1}$$

where k_1 and k_2 are the weights of the optimization object and $k_1+k_2 = 1$. That is, once one MSCB prefers deployment cost rather than service delay, he may increase k_1. The larger k_1, the more he concern the deployment cost. $k_1 = 1$ means that the VDC placement aims to minimize its deployment cost with guaranteed service quality. U_{ij} denotes the service delay between the ideal VDC x_i and an end user x_j. c_i denotes the price of unit resources of the ideal VDC x_i and r_j denotes the resource demand of the end user x_j. Binary variables x_i and y_{ij} indicate whether the point x_i is selected as an ideal VDC, and whether the point x_j can be served by the ideal VDC x_i. That is, $x_i = 1$ indicates that point x_i is selected as an ideal VDC. Similarly, $y_{ij} = 1$ indicates that the point x_j can be served by the ideal VDC x_i. Constraint $U_{ij} = 2 * u_{ij}$ denotes the service delay between point i and edge node j, which is determined by the distance u_{ij} between point i and edge node j, and constraint $U_{ij} \leq D_{thr}$ denotes that the service delay should satisfy a given service quality D_{thr}.

Obviously, Eq. 1 is NP hard. According to Section II.B, the mobility of end users makes it passively suitable for selecting VDCs from a candidate node pool, while clustering-based solutions such as NetClust may locate service nodes anywhere. Although this solution is suitable for large-scale service platforms such as data center-based clouds, it is not flexible enough to change end users due to the computing complexity of the clustering algorithm and network coordinate construction. However, due to the limited capital, MCSBs tend to change their VDC placement strategy according to the change in end users and the resource price of candidate clouds to reduce their deployment cost. In other words, the VDC placement strategy should not only locate its VDCs at arbitrary sites but also be flexible enough to adapt the change in end users and candidate clouds.

To address these features, we implement the placement framework of NetClust to design our lightweight VDC placement strategy, which consists of obtaining Internet information and VDC Selection. The former aims to obtain Internet information, including the location of end users, candidate geo-distributed clouds and VDCs and the service delay between each candidate geo-distributed cloud or VDC and each end user. In Eq. 1, the service delay of two points determined by their distance indicates that the Internet can be mapped to a network space, which can be described as a certain network coordinate. Then, this information of end users, candidate geo-distributed clouds and VDCs can be described by their corresponding point in this network coordinate. By this network coordinate, the cluster-based VDC placement algorithm can be mapped to the clustering algorithm for these points in this network coordinate space.

B. OBTAINING INTERNET INFORMATION

The VDC selection strategy uses a technique for obtaining Internet information to obtain the Internet information of the ICSPs' service footprints. In fact, geographic coordinate [28] and network coordinate [29], [30] construction are currently always implemented to obtain Internet information due to all types of terminals. Generally, network coordinate construction, including distributed network coordinate [29] and global network coordinate [30], is the most popular method for generating the Internet topology of terminals. Considering the accuracy and robustness among different network coordinate construction methods, the global network coordinate is implemented to formulate the Internet topology information of candidate clouds and end users [30]. The construction process can be described as follows [9]:

Let V be the Euclidean network coordinate space to be constructed, g^s_h denote the coordinates of a host h, and f^s denote the distance function that operates on these coordinates. Then, the distance between hosts h_1 and h_2 can be calculated as $f^s(g^s_{h_1}, g^s_{h_2}) = d^s_{h_1 h_2}$. For n landmarks $La_1, \ldots, La_n$, the pairwise distance is $u_{La_i La_j}$ between La_i and La_j, where $i, j \in \{1, \ldots, n\}$. The error of the calculating coordinates is minimized by $f_{err}(g^s_{La_1}, \ldots, g^s_{La_n}) = \sum_{La_i, La_j \in \{La_1, \ldots, La_n\} | i>j} \varepsilon(u_{La_i, La_j}, u^s_{La_i, La_j})$, where $\varepsilon(u_{La_i, La_j}, u^s_{La_i, La_j}) = (u_{La_i La_j} - u^s_{La_i La_j})^2$. Minimizing this function, the optimal coordinates for the landmarks $g^s_{La_1}, \ldots, g^s_{La_n}$ can be obtained. Similarly, the coordinate for any host h can be achieved by $f_err(g^s_h) = \sum_{La_i \in \{La_1, \ldots La_n\}} \varepsilon(u_{hLa_i}, u^s_{hLa_i})$.

In this paper, we take all candidate clouds as candidate landmarks and use the measured inputs of these candidate clouds to select the coordinate landmarks according to the proposed metrics. Then, the coordinate value of each end user and all candidate clouds are calculated based on the delay to each landmark. Generally, one network coordinate system consists of the coordinate origin, axes, and coordinate units, which are determined by its landmarks. Furthermore, the landmarks are also probing hosts, which measure the necessary network delay of network coordinate construction to all end users and candidate clouds' IPs. Finally, they are the foundation of the location of the network coordinate for each client and candidate cloud. Thus, the first and most important step is to select the coordinate landmarks for GNP network coordinate construction, which significantly influences the precision of the constructed network coordinate and the computing complexity of network space mapping.

However, few works have focused on how to select landmarks. P. Francis et al. formulated the probing host placement problem as a graph optimization problem [31]. H. Zhang et al. proposed a method to select landmarks from probing servers based on three metrics: the max separation metric, the N-cluster-medians metric, and the N-medians metric [30]. M. Rabinovich et al. found that suitable triangle landmark selection achieved lower computational and network probing costs with assured estimation accuracy [32]. Based on these previous works, we propose three landmark selection metrics: the max separation metric, the landmark-dimension number metric and the landmark location metric.

1) THE MAX SEPARATION METRIC

Assume V is an L-dimensional Euclidean space composed of network coordinates. L is the dimension of V. N denotes the set of coordinate points in space V. i and j represent two arbitrary hosts in space V, x_i, x_j are the coordinate values for point i, j. $d(i, j)$ denotes the distance function of x_i, x_j. Then, we have the following:

$$\mathrm{d(i, j)} = |\mathrm{x_i} - \mathrm{x_j}| = \sqrt{\sum_{l=1}^{L} (x_{il} - x_{jl})^2} \tag{2}$$

where x_{il} represents the coordinate value of the lth dimension of host i.

Let δ be the maximum measurement error of the Internet measurement. That is, for two arbitrary hosts i and j, the following equation always holds:

$$\delta_{\mathrm{ij}} = |\mathrm{d(i, j)_{real}} - \mathrm{d(i, j)_{measure}}| \leq \delta \tag{3}$$

Clearly, to guarantee the accuracy of the measurement, the distance between two arbitrary landmarks must be large enough.

2) THE LANDMARK-DIMENSION NUMBER METRIC

Generally, when constructing the network coordinate with coordinate landmarks to build the network coordinate system, the following relationship between the landmark number and the coordinate dimension always exists [30]:

$$N \geq d + 1 \tag{4}$$

where N denotes the landmark number and d presents the dimension of coordinate.

3) THE LANDMARKS LOCATION METRIC

Constructing the GNP network coordinate achieves the integrated Internet information, which helps to find the server deployment locations and allocate the service resources among these locations. Thus, the landmarks must be deployed in sites where more Internet information can be conveniently collected. In fact, a point-of-presence (PoP) is the local access point for one Internet service provider (ISP) connecting to other ISPs. It consists of high-speed telecommunication equipment and technologies that enable users to connect to the Internet via their ISP. Therefore, it is an information convergent point of the Internet and suitable for coordinating landmark deployment.

Based on our landmark selection metrics, we select three landmarks from all candidate clouds as the coordinate landmarks and design the Internet information achieving strategy as follows:

1) The *controller* sends the IP list of candidate *VDCs* to the candidate *VDCs*;

2) Each candidate VDC uses the lightweight ping to measure network delays to these IPs, preprocesses these measured results, and sends them back to the *controller*.

3) The *controller* applies our coordinate landmark selection metrics to the measured data and chooses the optimal three VDCs as the coordinate landmarks;

4) Based on the three selected landmarks, the GNP network coordinate system can be constructed with the ping measurement delay between each pair of *VDCs*.

5) The *controller* sends the IP list of the coordinate landmarks to all candidate VDCs and end users;

6) Candidate VDCs and end users use the lightweight ping to measure the network delay based on these IPs, preprocesses this measure data and sends it back to *controller* as response;

7) At last, *end users* and *VDCs* are mapped to our network coordinate system based on their measurement delay to our selected landmarks, through which the Internet information can be obtained.

Obviously, this strategy for obtaining Internet information is suitable for VDC selection due to changes in end users or Internet conditions. That is, the ICSPs of MSCBs are constructed and operated, and MSCBs change their VDCs of ICSPs to adapt to the change in end users or Internet conditions. Therefore, *controller* can make a suitable choice to select few suitable candidate clouds as their candidate VDCs from their candidate cloud lists, as shown in Step 1 and Step 2. In fact, Step 1) and Step 2) of this network coordinate construction aims to avoid the large time consumption of ping processing because there are too many different clouds to provide resources for VDC selection. In contrast, when MSCBs start their VDC selection for the first time ICSP construction because time consumption is not the main factor, existing VDC selection strategies such as Net Clust may also be applied.

C. VDCs SELECTION

Based on the Internet information obtained by the network coordinate, a fast VDC selection algorithm is proposed to obtain the VDC selection strategy for the ICSP.

Generally, existing server placement strategies can be divided into active selection strategies and passive selection strategies. The former designs a selection sever strategy by choosing M servers from a candidate server pool passively, while the latter applies k-means clustering to divide all end users into several groups. Then, the placement strategy can be determined according to Eq. 1. For example, $k_1 = k_2 = 0.5$. That is, the VDC selection strategy aims to minimize the sum of the deployment cost and the response time. Obviously, the candidate server pool of the passive selection strategy limits the service footprint of ICSP, which is not suitable for the goal of enlarging the service footprint of ICSP. For the k-means clustering-based strategies, because the clustering time of k-means clustering is influenced by the selection of initial centroids, their time complexity fluctuates significantly due to the randomness of initial centroid selection, which leads to more computing resources reserved for ICSP construction. Obviously, it is not suitable for the scenario of the ICSP construction of MSCBs.

To overcome these problems, we propose a fast VDC selection algorithm, as shown in Algorithm 1, which is based on the hierarchical clustering algorithm and incorporates with the fast density peaks finding algorithm [33], [34]. In Algorithm 1, the fast density peaks finding algorithm, which assumes that a clustering center has a higher local density than other points around it, is applied to determine subclusters and their centroids. Then, the hierarchical clustering algorithm is used to determine the suitable VDCs. Based on these selected VDCs, all end users are classified into the selected VDCs according to some metrics, such as minimum deployment cost or minimum service latency. Then, the network resources are allocated to these VDCs according to the demand of end users. Thus, our fast VDC selection algorithm consists of achieving the subcluster centroid and clustering and resource allocation.

In Algorithm 1, achieving the subcluster centroid aims to quickly classify end users into some subclusters and obtain the centroid of each subcluster, including Step 1, Step 2 and Step 3. For each point p_i, let ρ_i denote the local density of p_i, δ_i be the minimum distance from the nearest high dense point. Then, the local density p_i can be calculated as follows:

$$\rho_{\rm i} = \sum_{l=1}^{N} X(d_{ij} - d_c) \tag{5}$$

where d_{ij} is the distance between point p_i to p_j, d_c is the cutoff distance, which we determine this value as 0.02. Then, δ_i is

Algorithm 1 Fast VDC Selection Algorithm

Input V, the end users space composed of network coordinate with the dimension d;
m, the number of selected VDCs;
Cen_i, the centroid of Subcluster i;
Start 1) For each point p_i, calculate its ρ_i and δ_i according to Equations (5) and (6), respectively;
2) Calculate the β_i based on Equation (7) and choose the centroids of the selected subclusters;
3) For each point p_i, determine the Cen_i which is nearest to p_i after computing the distance: $Dis(Cen_i, j) = arg_j min_{i=1}^{m} |Cen_i - p_j|$, update the Cen_i with $(n * Cen_i + p_j)/(n+1)$;
4) Calculate the average subcluster linkage and construct the hierarchical clustering tree;
5) Determine the clustering result and the centroid of each cluster based on the hierarchical clustering tree and the number of VDCs;
6) Choose the coordinate value of the suitable VDC for each cluster based on the deployment cost and (or) the user experience metrics;
7) Allocate the service resource among these VDCs based on the number of points in each cluster and achieve the logical VDC selection strategy;
8) Map the coordinate value of these logical VDCs to IP;
9) Parse each IP to its corresponding physical candidate cloud and obtain our VDC selection strategy for geo-distributed clouds.

defined as

$$\delta_i = \begin{cases} min(d_{ij}), & \text{if } \ni j, \text{s.t.} \rho_j > \rho_i \\ max(d_{ij}), & \text{if } \ni j, \ \text{s.t.} \rho_j < \rho_i \end{cases} \tag{6}$$

Equation (5) describes the number of end users who are closer than d_c to p_i, and Equation (6) presents the locally or globally high density of end users. Based on Equations (5) and (6), the probability of end users as centroids β_i can be described as:

$$\beta_{\mathrm{i}} = max_{j \in N}(\rho_j \delta_j) \tag{7}$$

Then, we obtain the probability of end users as centroids, from β_i list, we select the centroids of subclusters.

Based on these subclusters, clustering applies the improved hierarchical clustering algorithm to locate the appropriate VDCs, which consists of Step 4, Step 5 and Step 6. In our algorithm, the average linkage metric is used to obtain the distances for each subcluster pair. Based on these distances, we construct the hierarchical clustering tree for our VDC selection. By merging the hierarchical clustering tree, we can obtain the centroids of the selected VDCs.

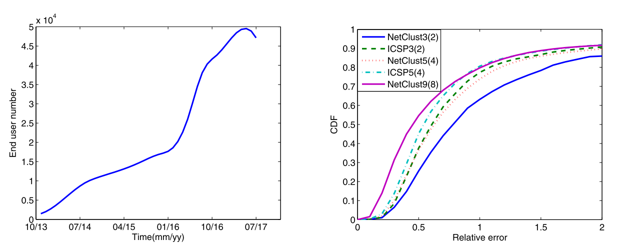
Finally, resource allocation allocates resources for each selected VDC according to its deployment cost and/or its service quality to obtain the logical VDC selection strategy of ICSP, including Step 7, Step 8 and Step 9 of Algorithm 1. Furthermore, resource allocation also maps this logical strategy to the physical Internet environment and chooses suitable clouds from all candidate clouds as the VDCs of ICSP. To this end, the centroids of clusters are taken as the service points, which denote the network coordinate of the corresponding candidate clouds. Then, their deployment cost and distances to all candidate clouds' coordinates are calculated, and the clouds with the minimal deployment cost and (or) distance are chosen as the corresponding logical VDCs. After that, we classify all network coordinates into these logical VDCs to obtain their demanded resources and deployment cost based on our network coordinate, which means we achieve the logical VDC strategy. Finally, these logical VDCs are mapped to IP and parsed to their physical deployment sites to obtain our VDC selection strategy.

V. EXPERIMENTAL DESIGN AND PERFORMANCE ANALYSIS

In this section, one simulator is designed based on our ICSP framework and VDC selection strategy. Then, this simulator is applied to evaluate the performance of our VDC selection strategy with practical trace data.

A. EXPERIMENTAL DESIGN

Based on the trace data and our framework of VDC selection, as shown in Fig. 2, the corresponding simulator is designed to simulate and evaluate our VDC selection strategy. This simulator consists of one *controller* and 30 candidate clouds (candidate VDCs), which implement the same workflow of our VDC selection mentioned in Section III. To evaluate the performance of the proposed algorithm under a real network environment, our inputs are carefully designed based on practical content service data. These data are collected from one content service platform of the Zhejiang Publishing United Group, which is an integrated content service platform and provides both traditional content service, digital content service and online B2C e-commerce service to end users, as shown in Fig. 3(a). In Fig. 3, the blue curve plots the change in the number of end users of this service platform from October 2013 to July 2017. Then, the distances from each candidate VDC to these end users and other candidate VDCs are calculated by its logging module according to



(a) The end user number of content service platform (b) The relative error of different coordinate landmark selection metrics

FIGURE 3. The user number and the relative error of different coordinate landmark selection metrics. (a) The end user number of content service platform. (b) The relative error of different coordinate landmark selection metrics.

their response delay and sent to the *controller*. Based on the received response delay, *controller* constructs the network coordinates and selects the VDCs and allocates the resources for each selected VDC.

To investigate the performance of our VDC selection strategy, both Zhang's landmark selection metrics [30] and our proposed landmark selection metrics are applied to landmark selection, and the VDC selection algorithms of the adaptive cloudlet placement method (ACPM) [23], tentacles [10] and our VDC selection (ICSP) are implemented to the constructed network coordinate to determine the location of selected VDCs and allocate the resources for each selected VDC. Then, the VDC selection strategies of ICSP and NetClust are applied to the 4-dimensional GNP network coordinate and plane network coordinates constructed based on Zhang's landmark selection metrics. In addition, the end user number and their corresponding response delays with 30 candidate VDCs in October 2015 are applied to another experiment to evaluate the performance fluctuation affected by the number of selected VDCs, whose number increases from 1 to 25 with Step 1 in Algorithm 1.

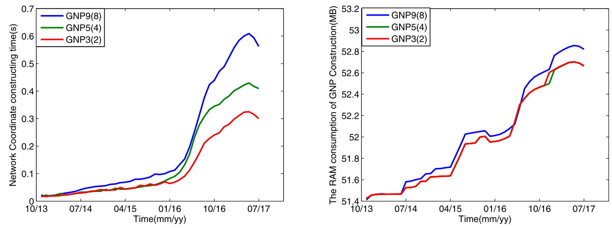
B. PERFORMANCE VERIFICATION

In this section, the developed simulator is used to investigate our VDC selection strategy and verify its performance, including the landmark selection and VDC selection.

1) LANDMARK SELECTION

Similar to the work in [30], the directional relative error is also used to verify the performance of our landmark selection metrics, shown as Fig. 3(b). Fig. 3(b) compares the directional relative error between our landmark selection metrics and Zhang's landmark selection metrics. The curve of NetClust5(4) means the directional relative error curve of 5-land marks and 4-dimensional GNP network coordinates, which is applied in NetClust. ICSP5(4) denotes the directional relative error curve of 5 land marks with a 4-dimensional GNP network coordinate based on our metrics. Similarly, the directional relative error of NetClust9(8), NetClust3(2) and ICSP3(2) are achieved.

Fig. 3(b) shows that the accuracy of ICSP5(4) is worse than that of NetClust9(8) but much better than that of NetClust5(4). Although at a low relative error level, the accuracy of the ICSP5(4) is worse than that of NetClust9(8), and when the relative error reaches 0.7, there is no distinct difference between them. Fig. 3(b) also shows that the accuracy of NetClust5(4) is always much worse than that of ICSP5(4). The accuracy of ICSP3(2) is not worse than that of NetClust5(4) and much better than that of NetClust3(2). When the value of relative error reaches 0.35, the accuracy of ICSP3(2) is higher than NetClust5(4). Generally, NetClust5(4) shows that the measurement delay of these 5 selected candidate VDCs is required to construct a 4-dimensional GNP network coordinate, while ICSP3(2) means that 3 selected candidate VDC measurement data are used to construct a GNP plane network coordinate. This means that the VDC selection strategy is carried out in a 4-dimensional space for NetClust while applied to a plane coordinate for our ICSP. Obviously, the network coordinate constructed based on our landmark selection metrics reduces both the computational and network probing cost while the estimation accuracy is sustainable. It is feasible for our VDC selection strategy to construct plane network coordinates by applying our landmark selection metrics.



(a) VS. Construction Time (b) VS. CPU Consumption

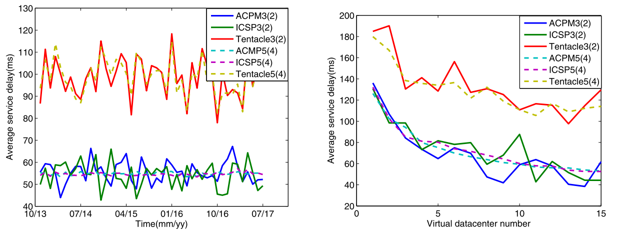
FIGURE 4. **The relative error of different coordinate landmark selection metrics. (a) VS. Construction time. (b) VS. CPU consumption.**

Fig. 4 shows the resource consumption of different dimensional GNP construction with the change of the end users, including GNP construction time and the maximum occupied RAM. In Fig. 4, the blue curve, green and red curves denote the resource consumption of 8-dimensional, 4-dimensional and plane GNP network coordinate construction, respectively. Fig. 4(a) plots the curves of different dimensional GNP construction times with the change of the end users, and Fig. 4(b) shows that the maximum RAM consumption of different dimensional network coordinate construction varies with the change of end users. From Fig. 4(a), it is clear that the construction time increases with the increasing dimensions of the GNP. Fig. 4(a) shows that the 8-dimensional GNP requires the most time to construct its network coordinate, while the construction time consumption of the plane network coordinate is the least. Fig. 4(b) shows that the maximum occupied RAM of the 4-dimensional and plane GNP network coordinate construction are similar to each other and much less than that of the 8-dimensional GNP. Furthermore, Fig. 4(b) shows that the maximum occupied RAM and construction time of these three-dimensional network coordinates are no more than 53*MB* and 0.7*s*, respectively, which means that the network coordinate construction is not the key factor for MSCBs to construct their ICSPs.

2) THE PERFORMANCES OF DIFFERENT SELECTION ALGORITHMS

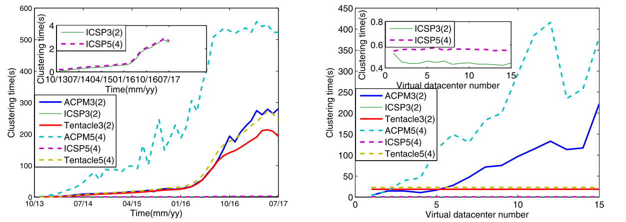
To investigate the performance of our VDC selection algorithm, the VDC selection algorithms of NetClust, ICSP and tentacles are applied to the same two network coordinate systems, 4-dimensional and plane GNP network coordinate.

Fig. 5 illustrates the service quality performance of NetClust, ICSP and tentacles with the change of the end users and clusters number under different network coordinates. The curve of NetClust5(4) shows the relationship between the



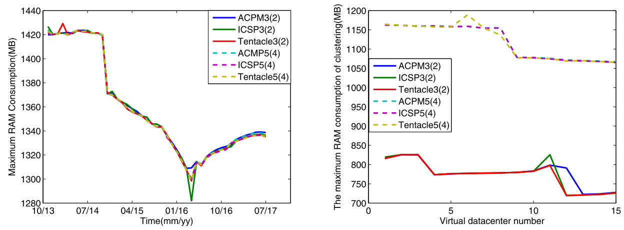
(a) The service delay curve for different end users (b) The service delay curve for different VDCs

FIGURE 5. **The service quality performance of different VDC selection strategies.**



(a) The clustering time for different end users (b) The clustering time for different VDCs

FIGURE 6. **The clustering performance of different VDC selection strategies.**



(a) VS. different dimensional network coordinate (b) VS. different VDCs

FIGURE 7. **The RAM occupation performance of different VDC selection strategies.**

service performance and end users' number (cluster number) when the NetClust VDC selection strategy is applied to the 4-dimensional network coordinate. The curves of ICSP5(4) and tentacles5(4) demonstrate that the service performance fluctuates with the end users' number (cluster number) when our selection algorithm and the tentacles selection algorithm are applied to the same network coordinate. Fig. 5(a) plots the service delay curve for different end users when VDC selection strategies of our ICSP and NetClust are applied to the 4-dimensional network coordinate and the plane coordinate. Fig. 5(b) shows the relationship between the service delay of these three VDC selection algorithms and the number of selected VDCs. The number of selected VDC is 5 in Fig. 5(a), and the number of end users is 13, 577 in Fig. 5(b). From Fig. 5, it is clear that there is no significant difference in service performance between our VDC selection strategy and the server placement strategy of NetClust, and both service performances are much better than that of tentacles. In addition, Fig. 5 also shows that for each VDC selection strategy, the higher the dimension of the network coordinate is, the smoother the service performance curve is. Finally, the number of selected VDC has a much greater impact on service performance than the number of users. Fig. 5(a) shows that the average service delay fluctuates from 42*ms* to 66*ms* for NetClust and ICSP, respectively, and this value fluctuates within 77*ms* and 120*ms* when the number of end users increases from 1, 485 to 49, 545. However, this value decreases from 140*ms* to less than 43*ms* for NetClust and ICSP and 200*ms* to less than 100*ms* for tentacles, in Fig. 5(b). The main reason is that these three selection algorithms allocate the resources for each selected VDC based on its service scopes. This means that for one selected node, no matter how many end users there are, the average service delay of this selected node is unchanged when its service scopes of centroid coordinate do not change. However, with the selected VDC number increasing, the service scope of each selected VDC decreases, which means that the average service delay of each selected VDC decreases.

These experiments are repeated 10 times, and the average value of the results is used to investigate the clustering time and the maximum RAM occupation of these three algorithms under different network coordinates, as shown in Fig. 6 and 7. In Fig. 6, the curves of NetClust5(4), ICSP5(4) and tentacles5(4) plot that clustering elapsed time fluctuates with end users' number (cluster number) when these three selection algorithms are applied to the same 4-dimensional network coordinate. Fig. 6(a) shows that the clustering elapsed times of different selection strategies fluctuate with different numbers of end users, while Fig. 6(b) discloses the relationship between the selection strategies' clustering elapsed time and the number of selected nodes. To show the clustering elapsed time difference between ICSP5(4) and ICSP 3(2), one small figure in Fig. 6(a) and Fig. 6(b) are further plotted. Comparing Fig. 6(a) with Fig. 6(b), it is clear that ICSP3(2) has the shortest elapsed time to obtain VDCs, while NetClust5(4) has the longest elapsed time under the same conditions. Fig. 6(a) shows that the value of ICSP3(2) is lower than the others, while that of NetClust5(4) is much larger than any others in the case of the same number of end users. This rule holds in Fig. 6(b) when the number of selected VDCs is the same. In addition, it is clear that the clustering elapsed time of NetClust5(4) and NetClust3(2) increases sharply with the number of end users and the selected VDCs increasing in Fig. 6(a) and Fig. 6(b). Different from the selection strategy of NetClust, both tentacles and our selection are only positively related to the number of end users. Fig. 6(a) shows that as the number of end users increases, the value of the elapsed time increases for both ICSP5(4) and ICSP3(2). However, in Fig. 6(b), it is clear that this value of ICSP3(2), ICSP5(4), tentacles5(4) and tentacles5(4) are almost invariant. The subfigure of Fig. 6(b) further shows that this value of ICSP5(4) and ICSP3(2) fluctuates within a small scale with increasing VDC number.

Fig. 7 illustrates the maximum RAM occupation of NetClust, ICSP and tentacles with the change of the end users and clusters number under different network coordinates. The curves of NetClust5(4), ICSP5(4) and tentacles5(4) show the relationship between the maximum RAM occupation and end user number (cluster number) when these three VDC selection algorithms are applied to the 4-dimensional network coordinate, while the curves of NetClust5(4), ICSP5(4) and tentacles5(4) illustrate the maximum occupied RAM fluctuation with the change in end users and the VDC number when they are applied to the plane network coordinate. Fig. 7(a) plots the maximum occupied RAM curve for different end users when VDC selection strategies of our ICSP and NetClust are applied to the 4-dimensional network coordinate and the plane coordinate, respectively. Fig. 7(b) shows the relationship between the maximum RAM occupation of these three VDC selection algorithms and the number of selected VDCs. Similar to Fig. 5, the number of selected VDC is 5 in Fig. 7(a), and the number of end users is 13577 in Fig. 7(b). From Fig. 7, it is clear that there is no significant difference in maximum RAM occupation among these three selection algorithms when the VDC number is fixed, while the dimension of the network coordinate may significantly affect the maximum RAM occupation of VDC selection algorithms when the number of end users is fixed. Fig. 7(a) shows that the maximum RAM occupation curves of these three algorithms almost coincide. In Fig. 7(b), there are two significantly different curves for different dimensions. The maximum RAM occupation curves of each selection algorithm for the 4-dimensional network coordinate are much greater than those of the plane network coordinate. However, for each type of network coordinate, there is no significant difference in maximum RAM occupation among these three selection algorithms.

Integrating the performance of the landmark selection metrics and VDC selection algorithms, our solution has the advantage of lower computing resource consumption due to the lower complexity of our node selection algorithm and suitable landmark selection metrics. The lower complexity of our VDC selection algorithm also reduces the computing resource consumption during the optimal VDC determination. At the same time, our landmark selection metrics allow us to obtain enough accurate network information with fewer dimensional network coordinates, further reducing the computing resources consumption. Although the dimension of the network coordinate is reduced to 2, the accuracy of our proposed ICSP3(2) is higher than NetClust5(4), as shown in Fig. 4. In addition, the results in Fig. 5 and Fig. 6 show that although the average service delay of ICSP3(2) is slightly lower than that of NetClust5(4), the elapsed time of ICSP3(2) is much lower than that of NetClust5(4). Thus, our solution can select the suitable VDCs from the whole network coordinate by constructing a plane network coordinate and applying a lightweight VDC selection algorithm, which means that our solution could be carried out online. Additionally, *controller* monitors the ICSP operation and Internet environment. Whenever the starting of the VDC selection strategy is triggered, for example, when there is an overload of one VDC and/or a change of many end users, *controller* will start our VDC selection strategy by sending the IP list of all end users and other candidate VDCs to candidate VDCs. Thus, our solution can adjust the change in the Internet environment and its operation condition, which is a flexible solution for micro- and small-content service companies to construct their ICSP in a multicloud environment.

C. COMPUTING COMPLEXITY ANALYSIS OF DIFFERENT STRATEGIES

Fig. 6 and 7 can be explained by the different time complexities of these two VDC selection strategies. In our algorithm, Step 1 has the time complexity of $O(N^2 + N)$, Step 2 and Step 3 have the time complexity of $O(N)$, so the time complexity of obtaining the subcluster centroid is $O(N^2)$. Clustering applies a hierarchical clustering algorithm, which is based on fuzzy graph connectedness and has a time complexity of $O(N)$ when the number of clusters is far smaller than the number of end users ($30 \ll 500$). Regarding resource allocation, Step 7 has a time complexity of $O(N)$. Step 8 maps the network coordinate value of each selected VDC to its corresponding IP. Step 9 parses these IPs to their candidate clouds. Obviously, the time complexity of these two steps is $O(K)$, where K is the cluster number. Thus, the time complexity of resource allocation is $O(K)$. Therefore, the complexity of our selection algorithms is $O(N^2 + N + K)$, the same as $O(N^2)$. However, as mentioned in [9], the computing complexity of our algorithm is $O(N^2)$, while this of NetClust is $O(N^3)$, resulting in the elapsed time of NetClust5(4) and NetClust3(2) being larger than ICSP5(4) and ICSP3(2), respectively, when the number of end users and the VDCs are the same. It also explains the phenomenon that both the elapsed time of these two node selection strategies are increasing as the number of end users increases. The phenomenon that the elapsed time of the NetClust's node selection strategy is positively related to the number of VDCs could be explained by incorrect selection of the initial nodes. Their node selection algorithm is based on K-Means, whose time complexity depends on the iteration number, the cluster number and the end users' number. Thus, incorrect selection of the initial nodes may result in a sharp increase in the iteration number, leading to an increase in the elapsed time.

In addition, our proposed landmark selection metrics are also helpful in improving the performance of our VDC selection strategy. As mentioned in [30], the computing complexity of landmarks' coordinates construction is $O(N^2D)$, and each end host coordinates is $O(ND)$, where D is the dimensionality of network coordinate space and N is the end users' account. This means that the computational complexity of coordinate construction is the linear function of the number of landmarks for a given host's account. Fig. 3(b) shows that the errors of ICSP5(4) and ICSP3(2) are similar to those of NetClust9(8) and NetClust5(4), respectively and are much smaller than the same number of landmarks selected by the metrics of [30]. This means that to achieve a similar precision network information, our landmark selection metrics may significantly reduce the elapsed time of network coordination construction for the same hosts, as shown in Fig. 4. In addition to saving computing resources, our landmark selection metrics are also helpful in reducing storage resources. Our landmark selection metrics may achieve a similar error of network coordination construction with fewer landmark numbers, resulting in the dimensionality of network coordinate space constructed by our landmark selection metrics being much fewer than that constructed by [30]. That is, for each end user and candidate cloud, the resources used to store its Internet information with our strategy are less than those with NetClust. Because this network coordinate information is the input of VDC selection algorithms to generate the final placement strategy, our strategy can save the storage resources of ISCP construction for MSCBs significantly.

VI. CONCLUSION

This paper aims to address the challenges of VDC selection from different geo-distributed clouds, which should be implemented by micro- and small-content service companies when they use these selected VDCs to construct their ICSPs. To this end, a practical service VDC strategy was proposed for ICSP construction in a multicloud environment. To avoid the high time complexity of NetClust, three landmark selection metrics were presented to significantly reduce the dimensions of network coordinates without degrading accuracy, leading to the sharp reduction of computing resource consumption for the following network coordinate construction and VDC selection. Then, a fast VDC selection algorithm was designed to reduce the computing resource consumption effectively with the time complexity of $O(N^2)$. The experimental results show that our landmark selection metrics can reduce the dimensions of network coordinates with assured accuracy. The experimental results also demonstrate that our VDC selection algorithm not only reduces the computing resources assumption but also locates end users to "optimal" VDCs. It is suitable for the scenario where micro and small companies construct their ICSPs by using other cloud resources.

ACKNOWLEDGMENT

The authors would like to thank Zhejiang Publishing United Group, who provides the part of our experiment data.

REFERENCES

[1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "Above the clouds: A Berkeley view of cloud computing," *Commun. ACM*, vol. 53, no. 4, pp. 50–58, 2010.

[2] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility," *Future Generat. Comput. Syst.*, vol. 25, no. 6, pp. 599–616, 2009.

[3] P. Gupta, A. Seetharaman, and J. R. Raj, "The usage and adoption of cloud computing by small and medium businesses," *Int. J. Inf. Manage.*, vol. 33, no. 5, pp. 861–874, Oct. 2013.

[4] W. Shi, C. Wu, and Z. Li, "An online mechanism for dynamic virtual cluster provisioning in geo-distributed clouds," in *Proc. IEEE INFOCOM*, Apr. 2016, pp. 1–9.

[5] W. Huang, M. Tacca, N. So, M. Razo, and A. Fumagalli, "Cooperative data center selection for optimal service performance: An ILP formulation," in *Proc. IEEE 10th Int. Symp. Parallel Distrib. Process. Appl.*, Jul. 2012, pp. 523–526.

[6] S. Yan, X. Wang, M. Razo, M. Tacca, and A. Fumagalli, "Data center selection: A probability based approach," in *Proc. 16th Int. Conf. Transparent Opt. Netw. (ICTON)*, 2014, pp. 1–5.

[7] C. Dong, W. Wen, T. Xu, and X. Yang, "Joint optimization of data-center selection and video-streaming distribution for crowdsourced live streaming in a geo-distributed cloud platform," *IEEE Trans. Netw. Service Manag.*, vol. 16, no. 2, pp. 729–742, Jun. 2019.

[8] Y. C. Hounnontin and Z. Dziong, "A framework for WAN optimization and data center selection using NaaS," in *Proc. IEEE 19th Int. Conf. High Perform. Comput. Commun. Workshops (HPCCWS)*, Dec. 2017, pp. 16–21.

[9] H. Yin, X. Zhang, T. Zhan, Y. Zhang, G. Min, and D. O. Wu, "NetClust: A framework for scalable and Pareto-optimal media server placement," *IEEE Trans. Multimedia*, vol. 15, no. 8, pp. 2114–2124, Dec. 2013.

[10] H. Xiang, X. Zhang, H. H. Liu, Y. Luo, C. Tian, S. Zhao, and F. Liu, "Edge provisioning with flexible server placement," *IEEE Trans. Parallel Distrib. Syst.*, vol. 28, no. 4, pp. 1031–1045, Apr. 2017.

[11] P. M. Rekha and M. Dakshayini, "Cost based data center selection policy for large scale networks," in *Proc. Int. Conf. Comput. Power, Energy, Inf. Commun. (ICCPEIC)*, Apr. 2014, pp. 18–23.

[12] J. Xu, J. Tang, B. Mumey, W. Zhang, K. Kwiat, and C. Kamhoua, "Cost-efficient virtual server provisioning and selection in distributed data centers," in *Proc. IEEE Next Gener. Netw. Symp. (ICC)*, Jun. 2015, pp. 5466–5472.

[13] X. Wang, J. Cao, and Y. Xiang, "Dynamic cloud service selection using an adaptive learning mechanism in multi-cloud computing," *J. Syst. Softw.*, vol. 100, no. 2, pp. 195–210, 2015.

[14] R. Liu, X. Yuan, J. Xu, J. chen, Y. Zeng, M. Cao, J. Liu, L. Xu, and Q. Fang, "A novel server selection approach for mobile cloud streaming service," *Simul. Model. Pract. Theory*, vol. 50, no. 1, pp. 72–82, Jan. 2015.

[15] A. Selvaraj, R. Patan, A. H. Gandomi, G. G. Deverajan, and M. Pushparaj, "Optimal virtual machine selection for anomaly detection using a swarm intelligence approach," *Appl. Soft Comput.*, vol. 84, no. 11, 2019, Art. no. 105686.

[16] A. Ponraj, "Optimistic virtual machine placement in cloud data centers using queuing approach," *Future Gener. Comput. Syst.*, vol. 93, no. 4, pp. 338–344, Apr. 2019.

[17] S. B. Melhem, A. Agarwal, N. Goel, and M. Zaman, "Markov prediction model for host load detection and VM placement in live migration," *IEEE Access*, vol. 6, pp. 7190–7205, 2017.

[18] F. Xu, F. Liu, L. Liu, H. Jin, B. Li, and B. Li, "iAware: Making live migration of virtual machines interference-aware in the cloud," *IEEE Trans. Comput.*, vol. 63, no. 12, pp. 3012–3025, Dec. 2014.

[19] Y. Liu, K. Wang, L. Ge, L. Ye, and J. Cheng, "Adaptive evaluation of virtual machine placement and migration scheduling algorithms using stochastic Petri nets," *IEEE Access*, vol. 7, pp. 79810–79824, 2019.

[20] W. Tian, Y. Zhao, M. Xu, Y. Zhong, and X. Sun, "A toolkit for modeling and simulation of real-time virtual machine allocation in a cloud data center," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 1, pp. 153–161, Jan. 2015.

[21] S. Jamin, C. Jin, A. R. Kurc, D. Raz, and Y. Shavitt, "Constrained mirror placement on the Internet," in *Proc. IEEE INFOCOM*, Apr. 2001, pp. 31–40.

[22] K. Xu, X. Li, S. K. Bose, and G. Shen, "Joint replica server placement, content caching, and request load assignment in content delivery networks," *IEEE Access*, vol. 6, pp. 17968–17981, Mar. 2018.

[23] H. Xiang, X. Xu, H. Zheng, S. Liu, T. Wu, W. Dou, and S. Yu, "An adaptive cloudlet placement method for mobile applications over GPS big data," in *Proc. IEEE GLOBECOM*, Washington, DC, USA, Dec. 2016, pp. 1–6.

[24] X. Yuan, H. Yin, G. Min, X. Liu, W. Hui, and G. Zhu, "A suitable server placement for peer-to-peer live streaming," *J. Supercomputing*, vol. 64, no. 3, pp. 1092–1107, 2013.

[25] Y. Zhang and M. Tatipamula, "The freshman handbook: A hint for the server placement of social networks," in *Proc. 20th Int. Conf. Companion World Wide Web*, 2011, pp. 173–174.

[26] Q. Zhang, Q. Zhu, M. F. Zhani, R. Boutaba, and J. L. Hellerste, "Dynamic service placement in geographically distributed clouds," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 12, pp. 762–772, Dec. 2013.

[27] T.-X. Do and Y. Kim, "Topology-aware resource-efficient placement for high availability clusters over geo-distributed cloud infrastructure," *IEEE Access*, vol. 7, pp. 107234–107246, 2019.

[28] B. Wong, A. Slivkins, and E. G. Sirer, "Meridian: A lightweight network location service without virtual coordinates," in *Proc. Conf. Appl. Technol. Archit. Protocols Comput. Commun.*, 2005, pp. 85–96.

[29] F. Dabek, R. Cox, F. Kaashoek, and R. Morris, "Vivaldi: A decentralized network coordinate system," in *Proc. Conf. Appl. Technol. Archit. Protocols Comput. Commun.*, vol. 34, 2004, pp. 15–16.

[30] T. S. E. Ng and H. Zhang, "Predicting Internet network distance with coordinates-based approaches," in *Proc. IEEE INFOCOM*, Jun. 2002, pp. 170–179.

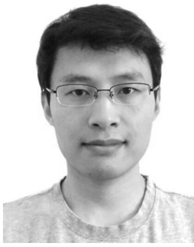
[31] P. Francis, S. Jamin, L. Zhang, D. F. Gryniewicz, Y. Jin, and V. Paxson, "An architecture for a global Internet host distance estimation service," in *Proc. IEEE INFOCOM*, Mar. 1999, pp. 210–217.

[32] Z. Wen and M. Rabinovich, "Network distance estimation with dynamic landmark triangles," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 36, no. 1, pp. 433–434, 2008.

[33] Y. Dong, Y. Zhuang, K. Chen, and X. Tai, "A hierarchical clustering algorithm based on fuzzy graph connectedness," *Fuzzy Sets Syst.*, vol. 157, no. 13, pp. 1760–1774, 2006.

[34] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492–1496, Jun. 2014.

BO HUANG received the Ph.D. degree from the Department of Electronics and Information Engineering, Huazhong University of Science and Technology, China, in 2011. From April 2002 to January 2006, he joined the R&D Center, ZTE Company, Nanjing, China, where he was involved in the 3G Core Net development. He is currently a Faculty Member with the College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing. His research interests include mobile edge computing and industrial communication networks.



CHANGLAI DU received the Ph.D. degree from the Computer Science Department, Virginia Tech University. His current research interests include security enhancements and vulnerability analysis of wireless systems, sensing capability of wireless signals and its application in security enhancements, blockchain security and applications, and mobile and the IoT security.

MENGTING SUN is currently pursuing the degree with the School of Information Management, Wuhan University, Wuhan, China. Her research interest includes Edge Cloud Comp.

XIAOQUN YUAN received the Ph.D. degree from the Department of Electronics and Information Engineering, Huazhong University of Science and Technology, China, in 2011. He is currently an Associate Professor of information management with the School of Information Management, Wuhan University. His current research interests include overlay networks, peer-to-peer systems, live streaming, source optimization, and digital publishing.