

# **A New Algorithm for Improving Latency in Distributed Data Center for Logistics Information System over Cloud**

Shivani Dubey  
Research Scholar  
Ansal University, Gurgaun

Sunayana Jain  
Associate Professor  
IMS, Ghaziabad

## **ABSTRACT**

The cloud based distributed data center uses virtualization technology to share the resources to the outside world through a virtual machine. Cloud administrator selects the data center to access virtual machines by using administrative and dynamic policies. Every data center has multiple virtual machines. Selection of data center is an important task which affects on performance as well as cost effectiveness of the data center. This problem can be solved by centralized as well as distributed data center. In Logistics Company, Centralized data center faces bottleneck in operations like virtual machine migration, creation, deletion, and needs to contact central administrator which can increase the negligible amount of network traffic. The paper presents comparison of distributed and centralized data center and strategies of distributed data center for reducing the latency and cost of selection of data center over the cloud by proposing an algorithm distributed service broker policy (DSBP) for logistics information system.

## **Keywords**

Centralized & Distributed Data Center, Distributed data Center for Logistics, Selection and Scheduling policy Algorithms, DSBP

## **1. INTRODUCTION**

Data center is group of computer servers for the remote storage, processing and distribution of large amounts of data. Cloud service providers provide services using large scale with cost effectiveness in cloud environment [1]. The cloud resources are distributed and interconnected in geographic environment over a wide area network. The latency in communication between the different centers is not more far significant in centralized data center [2]. This paper presents lot of effectiveness of distributed data center and analyzes the importance of various selection algorithms to minimize cost and the maximum latency in communication between the virtual machines allocated for a user request.

## **2. CENTRALIZED AND DISTRIBUTED DATA CENTER APPROACH**

In centralized data center, the cloud provider is connected one or more data centers located into small geographical area because the distance between the end user and the data center can potentially be large. In central data center the central

cloud administrator allocates the virtual machine of the data center using centralized policy. Then the data center providers do not need worry about resource utilization. The centralized data center reduces the work load of administrator to handle the data center. In centralized data center, the centralized approach can be executed by central manager and local manager of data center. The centralized approach can be executed by central manager and local manager of data center. There are many operations like creation, deletion, virtual machine migration which needs to contact the central manager where large amount of network traffic is increased [3].

After the getting this type of problem distributed data center is developed in which the first central manager decides which data center will serve this request and sends it to local manager to decide where it is going to deploy. After that the local manager chooses which physical machine in a data center will be executed that virtual machine. Other hand distributed data center has lack of making new policies to maintain the performance and decrease the maximum latency for selection of data center [3]. A customer requests can be serviced from nearest location to them in distributed data center. It is beneficial to reduce network capacity needs for high bandwidth applications [4]. Distributed data center also reduces the latency of access have small variation in long path lengths and going through multiple service providers [5].

## **3. DISTRIBUTED DATA CENTER OVER CLOUD**

The large number of service requests increases the need of data centers and also scale of data centers and their latency increments so that an efficient computing resource management approaches for distributed over cloud is essential to service providers. Distributed data center is geographically distributed, interconnected, fully automated with high speed and low latency network. Distributed data center provides a distributed set of machines that are running at different locations, connected to a single network. By using flexibility, scheduling and automation techniques, this resource pool can provide users with service and access on demand, making a better resources [6]. Supply chain is distributed in nature and logistics firm wants to grow its supply and distribution, there should be need to scale IT services of supply chain at big level. Distributed Datacenter provide better bandwidth and traffic for logistics partners in cloud [7].

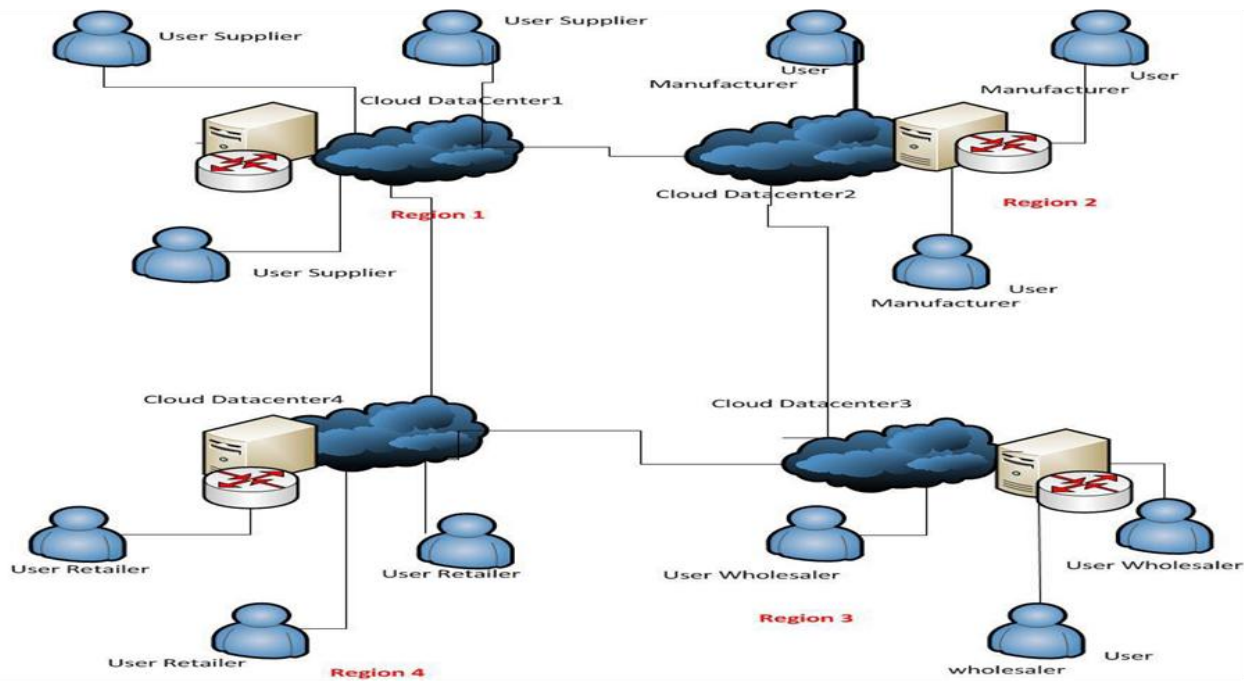


Fig 1: Distributed Data Center for Logistics Information System over Cloud

### 3.1 Users latency

Responses times for the end user are very important in certain applications, so having data centers near to users is also important and the ability to send traffic to different regions helps simplify this.

## 4. DISTRIBUTED DATA CENTER FOR LOGISTICS INFORMATION SYSTEM OVER CLOUD

Information is lifeblood of logistics and distribution system in logistics information system. It holds the whole system and coordinates all the components of logistics operation like, planning, coordination and operation [7]. Global logistics information infrastructure serves hundreds of thousands of customers in more than hundred countries. The logistics companies are expanding to help their customer achieve lower latency and higher throughput and to ensure that their data resides only in the region they specify [8]. Geographic distance from end users of cloud hosted applications translates to increased latency, which is a key challenge for some other applications and services wishing to leverage cloud computing. Latency from local data centers that are near to end users as do services like content distribution networks that have target to push static content towards the users at the edge of the network. While providers do indeed build distributed data center for in geographic environment needed to achieve adequate disaster tolerance. There are some important features of Distributed data center:

- Cloud service providers without spending large capital
- Keeps the data near to the different part like, low latency and high performance
- Enable always on cloud services
- Network function virtualization
- Internet of things

## 5. LATEST SELECTION ALGORITHMS APPROACHES FOR DISTIBUTED DATA CENTER

### 5.1. CloudGPS(Global Positioning System):

In 2013, Cong Ding, Yang Chen, Tianyin Xu, and Xiaoming Fu defined that Cloud computing provides an elastic and stable infrastructure for hosting online applications like web searching, emailing, instant messaging, online social networking and online gaming. Cloud infrastructure providers allocate computation nodes from one data center to different data centers. Distributed system has different solutions for minimize the user perceived latency and increase reliability for reducing service outage. The selection method provides a load balancing among service hosting nodes in order to effective the cloud service provider's investment. There are so many methods for selection, first is to grant the choice rights to the users so that user will choose its nearest server for obtaining the lowest access latency. Selection does not consider server workloads by some over capacity of servers because of certain distribution of users online and offline behavior. The second method is that cloud service providers or cloud infrastructure provider handle the server selection by using a centralized hierarchical stable marriage algorithm. The centralized architecture causes an extreme large overhead delay and makes systems less responsive to change in client request so it is not suitable for outsourcing system to provide services for all the commercial clouds. These both problems are solved by some outsourcing mapping systems, for example DONAR (decentralized server selection for cloud services) provides both proximity and server load in their selection policy. In this paper authors proposed CloudGPS, a friendly server selection scheme to address the above problems and server selection algorithm, which have three important metrics: network performance, server workload balance and inter domain transit traffic [9]. Long latency between clients and servers creates poor user experience and network performance, imbalanced workload of servers

increases the risk of server breaking down and inters domain transit traffic produces unnecessary ISP (Internet service providers) operational cost. The main goal of this paper is to minimize the network cost, balance client requests across servers and reduce inter domain transit traffic. The authors introduce transit traffic penalty coefficient to reflect the internet service provider's economic profit. They measured the selection accuracy of CloudGPS in closest server selection (CSS), i.e., the overhead of all the servers are unlimited. They randomly selected a number of nodes from dataset as the cloud node, and measure the stretch of using estimation to select the closest server. The stretch is defined as the distance to the closest cloud server cluster selected based on estimation, divided by the distance to the actual closest server cluster [10]. When the number of servers is 50, the average stretch of CloudGPS is 1.35, much smaller than of CloudGPS without SM (16.04) and round robin (16.03). CloudGPS's stretch for 140 servers is only 2.78, much smaller than CloudGPS without selective measurement SM (24.61) and round robin (19.97). The average stretch of CloudGPS without inter-domain transit traffic penalty is 1.347, and of CloudGPS is 1.350, respectively, when the number of servers is 40, they are almost the same. From the analysis above, they gave conclusion that comparing with the reduced measurement cost and inter-domain transit traffic, the performance is significant. CloudGPS is scalable to deal with the explosively-increasing numbers of clouds as well as user clients based on its DEM component. DEM (Distance estimation module) takes the advantages of two different kinds of NC (Network Coordinate) techniques to position both the clouds and users, achieving the measurement cost reduction from  $O(N)$  to  $O(1)$  for clouds with  $N$  server clusters. CloudGPS is ISP-friendly that effectively reduces inter-domain transit traffic leading to low ISP operational costs and improve end users' quality of service based on its MM (Mapping module) component which makes a balance between the closest server selection and the inter-domain transit traffic, in the limitation of servers' capacity.

## 5.2. LASTor (A Low Latency AS-Aware

**Tor):** In 2012, Masoud Akhondi, Curtis Yu, and Harsha V. Madhyastha defined the path selection which is established in Tor to Internet routing, anonymity guarantees can breakdown, where an autonomous system can correlate traffic across the entry an exist segment of a circuit. The authors also show both of these shortcomings in Tor can be addressed with only client server implementations. They designed and implement LASTor's path selection algorithm by which user can choose an appropriate tradeoff between latency and anonymity by specifying a value between 0 (lowest latency) and 1 (high anonymity) for a single parameter. In last they developed an efficient and accurate algorithm to indentify paths on which an autonomous system can correlate traffic between the entry and exist segments. For developing LASTor, authors made three main contributions, in first, latency gains are possible by solely accounting for inferred geographic locations of relays rather than needing up to date latency information of internet paths. They implemented the Weighted Shortest Path (WSP) algorithm which chooses paths with a preference for shorter paths. With the implementation of Weighted Shortest Path (WSP), an adversary can increase the probabilistically of a relay under his control being on the chosen path by simply setting up a large number of relays in the same location , which is near to the direct line between the source and the destination. They implemented ASTor to execute WSP on a graph of the Tor Network where nearby relays are clustered

together which increases the onus on an adversary to develop relays in different locations to ensure a high probability for the chosen path traversing a relay under his control. The side-effect of this process is that the run time of WSP is significantly reduced. They gave an example for demonstrate benefits of LASTor to improve latency by using it to visit the top 200 websites from 50 geographically distributed planet lab nodes. They also implemented all of the algorithms developed thus far to improve path latency for making path selection tunable and to incorporate AS (Autonomous System) awareness into path selection in LASTor Tor client. In the implementation, they defined that to select a path for the specific destination, LAST or executes the tunable AS aware WSP algorithm with the following steps [11]:

- For initialization, the LASTor client clusters all available relays and using the value for specified in its input configuration and it chooses three entry guards at random from 20% to 80% closest relay clusters to the client.
- LASTor resolves the destination's hostname on a distributed set of nodes that service requests to perform Domain Name Service (DNS) lookups. All requests are submitted by one of the circuits established upon initialization of the client.
- LASTor calculates the AS sets for the paths from the client to the entry guards and from all exit relays to the destination and mapping each candidate exit relay to the nearest among the IP address obtained for the destination.
- LASTor also computes the end to end distance on every candidate path through three clusters which satisfy the check of the AS sets for the entry and exit segments being disjoint. One cluster level path is selected with the probability of a path being chosen dependent on the end to end distance on it and the input value of the circuit to the destination via one relay selected at random from each of the clusters on the chosen cluster level path. After this the authors showed that it is necessary to carefully select entry guards and account for replicated destinations. They developed a space and time efficient technique for enabling LASTor to reliably detect the possible presence of snooping ASes on any path. They made path selection in LASTor tunable so that a user can easily choose an appropriate tradeoff between latency and anonymity [11].

## 5.3. Network Latency Profiling and

**Redundancy:** In 2014, Minseok Kwon, Zuocho Douy, Wendi Heinzlmany, Tolga Soyatay, He Bay, Jiye Shiz defined that the effects of profiting and redundancy on latency when a client has a choice of multiple servers to connect the base on network measurements. Network profilers in existing server selection approaches and mobile computing systems characterize the quality of network connections by measuring their performance such as average latencies and throughput and their standard deviations. They measured latencies to servers as samples when different sizes of data are sent and create models to estimate latencies for an arbitrary data size. By using estimated latencies at given data size they compared the latencies o server selection strategies with and without profiling or redundancy as the number of servers changes and data is divided into multiple chunks being sent to the servers. The contribution of authors is to characterize latency

behaviors when profiling or redundant data transfer is adopted for server selection using network measurements. The key findings are summarized as first is if data are evenly distributed across servers, latencies decrease initially as more servers take the burden, but eventually increase as more and more servers are involved. Second, if data are distributed greedily use the server with short response time first utilizing the average latencies to servers, latencies decrease and stabilize at the lowest point. Third, when data are evenly distributed to random servers without profiling, latencies do not change much even when more servers participate and forth, if data are distributed to random servers, redundantly that too but without profiling, latencies indeed decrease as more servers join and latencies further decrease with higher redundancy. This paper discussed server selection and data partitioning algorithms using their measured latencies. It is based on measured latencies to servers, they tested the random, fixed and greedy data portioning and server selection algorithms. Their results indicate the latencies can be reduced using the fixed algorithm compared to the random strategy and can be nearly optimized as data are partitioned and sent to servers greedily. The random algorithm can reduce latency significantly if data are sent redundantly. This implies the potential of dynamic profiling in which the greedy strategy is used for long lived regular traffic while random strategy with redundancy is used for short lived traffic.

The above algorithms present different solution for selection of servers, path and data center but there is some lack of process like how users can access applications and services which float around the cloud? In distributed data center, applications can be moved from one physical location to another in large geographic region. There are so many issues are raised like, latency, accuracy, scalability and cost of virtualization in distributed data center over cloud. The locations of data centers continue multiply due to the growth, bandwidth, space, power and latency requirement, while this has a big target for any logistics company to enable the better network infrastructure to connect the different location together. This network should have transparent platform for reliable delivery of services to distributed users and applications. Thus we are presenting a flow chart of distributed service broker policy (DSBP) algorithm in which discrete recourses are allocated for applications rather than completely isolated systems [12].

## 6. PROPOSED DISTRIBUTED SERVICE BROKER POLICY ALGORITHM (DSBP)

DSBP allows us to build data centers around the applications rather than the network. All the parts of data center like IP, network, storage and applications helps data centers to distribute information at multiple locations. DSBP plays an important role in:

- Make a bridge among the geographical distances according to business nature
- Improve the different process of resource sharing like, data migration, computation migration and process migration
- It has transparency for presenting hidden details
- Improvement of data protection like, data loss and corruption
- Perform as a remote replication for distribution of contents
- Protect security concerns like confidentiality to unauthorized individuals

- Protect many business applications like, disaster recovery infrastructure, real time disaster recovery solutions

### 6.1. Flow Chart of Distributed Service Broker Policy (DSBP)

We are presenting distributed service broker policy algorithm. In this technique we will use distributed datacenter where each region has its own datacenter and will see that how the response time will reduce when using datacenter selection policy based that user base request will executed on nearest datacenter. So the nearest datacenter which is highest in the proximity list response the user base request or job scheduling.

There are following steps of DSBP algorithm:

- 1) Select the region
- 2) Calculate number of Data Center in selected region
- 3) If there is single Data Center then send the request to that specific Data Center
- 4) If there are multiple Data Center, select the nearest Data Center with minimum communication delay and maximum usable bandwidth between user base (client) and nearest data center in selected region
- 5) Find out the upcoming request in data center
- 6) Send the request to selected nearest data center
- 7) Analyze the result

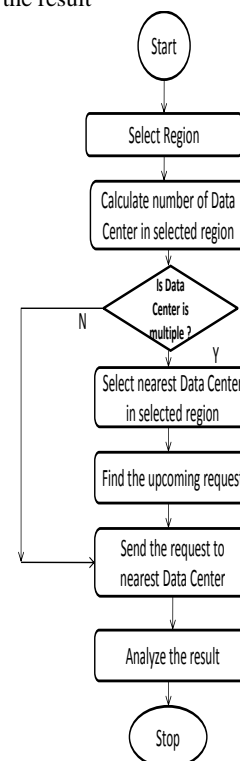


Figure 2: Flow chart of DSBP algorithm

## 7. RESULT ANALYSIS

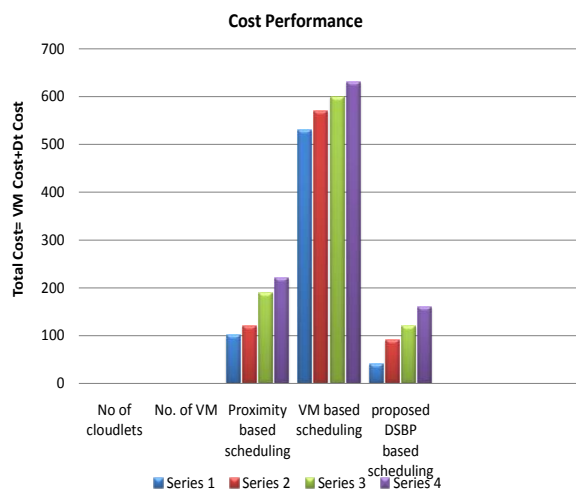
The proposed distributed service broker policy shall be implemented using simulation Cloud-Analyst. We define that how the response time will reduce when using data center selection policy based on user base request which will be executed on nearest data center. So the nearest data center which is highest in the proximity list response the user base

request. Same way user base request from different data center created with their respected ids. Each user base shall contain n number of users so each region create n number of user requests, these request shall be handled by its local data center which reside on their respective region. Table.1. presents the average response time (ART) of all user base of both centralized data center (CDC) and proposed distributed data center (DDC).

**Table.1: ART performance in centralized and distributed data center**

Data Center (DC)	Region	VM in DC	VM Memory	VM Bandwidth	VM Image size	User Base	Request	Avg. response time
CDC	0	1	512 Mb	1000	10000	1000	50	334.85
Proposed DDC	5	5	512 Mb	1000	10000	1000	50	50.2

The proposed DSBP algorithm also gives selection of least costly data center while in comparison to other service broker policy or datacenter selection scheduling of centralized data center. In the given graph Figure.3, the number of VM and No of cloudlet on x-axis while cost of different datacenters on y-axis.



**Figure 3: Proposed Cost performance graph using different data center selection scheduling Algorithm**

## 8. CONCLUSION

In this paper DSBP is the improvement of service broker policy which will drastically minimizes the response time and minimum cost of selection of data center observed by logistics user which leads to improvement of service request timing. Logistics Management is in initial phase for adopting cloud computing for its IT services. In IT information sharing is basic part which it uses. There are various modules and phases required for completely adopting the cloud. We are working on first module which is selection module and give the solution to its basic requirement. In cost service broker policy we are proposing a DSBP for cost effective data center scheduling algorithm which give selection of datacenter

which has least cost in terms of virtual machine cost and data transfer cost. Proposed algorithm will give effective improvement in comparison to proximity based and virtual machine based which already defined service broker policy and selection algorithms.

## 9. REFERENCES

- [1] Devyaniba Chudasama, Namisha Trivedi, Richa Sinha, Cost effective selection of Data Center by Proximity based routing policy for Service Brokering in Cloud Environment, International Journal of Computer Technology & Applications, Nov-Cec 2012, vol.3(6), 2057-2059.
- [2] Mansoor Alicherry, T.V.Lakshman, Network Aware Resource Allocation in Distributed Clouds, INFOCOM, 2012 Proceedings IEEE, 963-971.
- [3] Maurice Bolhuis, A Comparison between Centralized and Distributed Cloud Storage Data Center Topologies, 19<sup>th</sup> Twente Student Conference on IT, June 24,2013, Enschede, The Netherlands.
- [4] A. Gottlieb, Beware the network cost gotchas of cloud computing, Cloud Computing Journal, June 2011.
- [5] Leighton, Improving performance on the Internet, Commun. ACM, 52, February 2009ngton.
- [6] Cloud Architecture and Datacenter Design in Distributed Computing: Clusters, Grids and Clouds, Kai Hwang, Geoffrey Fox, and Jack Dongarra, May 2, 2010
- [7] Dubey Shivani, Jain Sunayana, Logistics Information System and Cloud Computing, journal-of-computer-science, journal-of-computer-science, 28 Feb, 2014
- [8] Zhen Yu, Gui-Cheng Shen, Bing-Wu Liu, research on Architecture of Logistics Public Information Platform, Intelligent Information Management, 2012, 4, 396-400
- [9] Cong Ding, Yang Chen, Tianyin Xu, and Xiaoming Fu, CloudGPS: A Scalable and ISP-Friendly Server Selection Scheme in Cloud Computing Environments, Proceeding of the 20th IEEE/ACM International Workshop on Quality of Service, Coeimbra, Portugal, 2012
- [10] R. Zhang, C. Tang, Y.C. Hu, S. Fahmy, and X. Lin. Impact of the inaccuracy of distance prediction algorithms on internet applications-an analytical and comparative study. In IEEE INFOCOM, 2006.
- [11] Masoud Akhoondi, Curtis Yu, and Harsha V. Madhyastha, LASTor: A Low-Latency AS-Aware Tor Client, Proceeding of the IEEE Symposium on Security and Privacy, IEEE Computer Society, Washington, DC, USA, 2012
- [12] Minseok Kwon, Zuochoao Douy, Wendi Heinzelman, Tolga Soyatay, He Bay, Jiye Shiz, Use of Network Latency Profiling and Redundancy for Cloud Server Selection, Proceeding of IEEE Cloud, June-July 2014