International Journal of Future Generation Communication and Networking
Vol. 11, No. 6 (2018), pp.1-12
http://dx.doi.org/10.14257/ijfgcn.2018.11.6.01

Data Center Allocation during Fault Detection and Fault Tolerance in Cloud Network Based on Real Data Analysis

Debabrata Sarddar[1], Manas Kumar Sanyal[2], Sougata Chakraborty[3] and Priyajit Sen[4]

[1]Assistant Professor, Department of Computer Science and Engineering, University of Kalyani, Kalyani, Nadia, 741235, West Bengal, India
[2]Professor, Department of Business Administration, University of Kalyani, Kalyani, Nadia, 741235, West Bengal, India
[3]Senior System Engineer, IBM India Pvt. Ltd., Kolkata, West Bengal, India
[4]Assistant Professor, Department of Computer Application, B.P. Poddar Institute of Management and Technology, Kolkata, Indi
dsarddar1@gmail.com; manassanyal123@gmail.com; me.sougata.chakraborty@gmail.com; priyajit91@gmail.com

Abstract

Now-a-days, Data Center Allocation in the cloud network is a challenging work indeed when various cloud nodes are available in the network. One of them can be chosen as a data center within a specific region which may not be the best fit data center containing the overall good health. To deal with this problem, we have proposed a new algorithm that tells us how we can choose the data center among the set of all available data centers. We can choose a data center by checking few attributes like workload, storage, network speed, latency, traffic etc. of the cloud nodes based on real data analysis. The most advantageous and suitable node among all other nodes can be chosen as a data center. When any fault (system crash, power failure, hardware problems etc.) occurs in the data center then it is necessary to switch to another most suitable data center instead.

Keywords: *Data Center allocation, Fault Detection, Fault Tolerance, Local Backup, Centralized Backup*

1. Introduction

Data center contains servers, communication media and data storage devices inside it. It is a facility to combine the networked computers that is used to organize, create, process and access huge amount of data. They work as the principal storage space for different IT equipment like servers, storage subsystems *etc*.

2. Characteristics of Data Centers

2.1. Work Load

Workload is the attribute by which the current load of data centers can be measured. That means what is its actual load capacity, what is the load given to a data center and what is the remaining load can be measured with the help of this attribute. Here, load means number of users in millions currently associated with a data center.

Received (May 4, 2018), Review Result (July 5, 2018), Accepted (July 12, 2018)

ISSN: 2233-7857 IJFGCN

2.2. Storage

Storage is one of the precious attributes of data centers. It can measure the storage capacity of a data center. That means how much data a data center can hold and how much data is already stored within the data center and its remaining size also. Storage of data center is measured in Exabyte.

2.3. Traffic Growth

The network traffic is the flow of data packets across the network. It has been increasing rapidly day by day with the expansion of the number of users over the network. The phenomenon of dynamic growth of the network traffic worldwide is called as Traffic Growth.

2.4. Network Speed

Originally, the network bandwidth is termed as the network speed. Earlier, the unit of network speed was bits per second or bps and afterwards, it has been expressed as Kilobits per second or Kbps. In modern networking and communication system, network speed is measured in the millions of bits per second or megabits per second (Mbps). Billions of bits per second or gigabits per second (Gbps) is used in case of higher speed.

2.5. Latency

Latency is the time taken by a data packet for traversing in a network from one hop to another.

3. Services Offered by the Data Centers

Globally, the data centers do provide some services to the cloud users. In that case, we need to understand the term Data Center as a Service (DCaaS) beforehand. It is basically a provision for the clients to use the data center facilities and infrastructure. Mainly, the following services are offered by the data centers, such as

3.1. End User Services

End User Services are always extremely responsive to end user support environments. These services are flexible and efficient in nature. The data center providers deliver infrastructure to keep the end user support environment up and running without compromising security, compliance and control. Every data center provider organization’s objective is to meet the customer satisfaction by providing a faster resolution of any complex problem or failure with an improved productivity with a lower cost.

3.2. Application Services

As the name suggests, these services are purely the software-based services for the other organizations or the worldwide customers over the network. Sometimes, the application service provider organizations can provide the software-based services to develop integrate and manage their own network also. Examples can be cited for this service as web hosting and email hosting. Mobile environments and any hand-held devices could be under the inclusion of this service category.

3.3. Information Technology Services

As the major parts of these services, the following services are offered by the data center service provider organizations.

3.3.1. Data/ Information Protection or Security Services

Data or information security services are the key services provided by the data center provider organization ensuring the fact that the eavesdroppers or the sleeper cells do not steal information or data and spread harm over the network due to some self-interested profit-making activities and spoil the access of the other users.

3.3.2. Outsourcing Services

When an organization signs a contract with an outsourcing vendor company to provide IT services to the customers worldwide. These services can include data migration, data recovery, data backup, data storage enhancement, server upgradation and the overall data center maintenance *etc*.

3.3.3. Technical Support and Consulting Services

Practically, Technical support is the assistance to the customers to solve technical problems. These technical support services can include the problems with the servers, storage, network equipment, software used in the data center. Installation and configuration of the technical equipment are also laying with these services. Consulting services include making a new strategic decision or designing modern technology, redesigning existing technology or system integration with an existing system, data center location, migration, virtualization and automation *etc*.

3.3.4. Financing and Leasing Services

Financing is the act of acquiring capital for purchasing the infrastructure or services. Financing services are offered by the lending organizations to clients in the form of a loan which is paid back afterwards within a fixed time span with an applied interest rate as per the agreement. Leasing is the act of entering into a contractual agreement for a predetermined time frame to obtain provisional possession of the infrastructure or services in ex-change of fixed term compensation. In terms of data center services, financing service includes the upgradation of the entire data center facility and the leasing service comprises the leasing of data center equipment, servers and all other setups.

4. Fault Detection and Fault Tolerance

Normally, faults can be termed as the deviation from the expected behavior of a system which has the adverse impact on the functionality and the performance of that system. If a system behaves erroneously and fails to produce an expected outcome functionally, then the fault is categorically said as functional fault. Whenever the faulty nature of a system degrades the performance of that system then the fault is performance fault. Technical failures associated with the orchestration of the datacenters in cloud network are the classic example of functional fault whereas the latency is a perfect example of performance fault.

To check the expected behavior of the system, constant monitoring via manual inspection and automated process execution with the regular time interval is the key method to detect the faults. This phenomenon is called as fault detection.

If the fault is detected, then the additional mechanism is used to study the nature of the fault and track its root cause as well. This process is fault diagnosis.

Evidence generation is the ultimate step which is used to convince the system administrator or the concerned authority about the diagnosis result. Evidences could be generated from the system log and these evidences might be used for the debugging purpose. Evidences are kept as the reference for studying the frequency or the occurrence of the failures to ensure fault avoidance in future. In this way, the fault tolerant system could be originated in the later stage.

Above all, fault tolerance focuses on the reliability, availability and the robustness of any system. In this context, once any kind of fault (system crash, power failure, hardware problems *etc.*,) occurs in the data center then it is necessary to use another most suitable data center instead which would be expected to be a fault tolerant system.

5. Related Work

Modern data centers are growing in their scale and complexity. They are varying dynamically due to the addition and elimination of system components, execution environments, frequent updates and upgrades *etc*. The Failure Detection and Prediction mechanism can avoid failure with high accuracy [1].

To achieve robustness and reliability of cloud computing, failure should be handled effectively. The algorithm using Artificial Neural Network is proposed to detect faults, so that the gaps of already proposed algorithms can be filled by providing a suitable fault tolerant mode [2].

An empirical study of the network traffic in 10 data centers belonging to three different types of organizations, including university, enterprise, and cloud data centers have been made for designing more improved networks for data centers in cloud [3].

The realistic data of the Cloud Data Centers such as storage, workload, network traffic, latency *etc.*, is collected from CISCO white paper [4].

The technique to improve the fault-tolerance and reliability demonstrate that fault-tolerance and reliability against faulty and even malicious clouds in cloud computing can be achieved [5].

A survey on the challenges, techniques and implementation of Fault Tolerance in Cloud Computing have been made to identify different faults and techniques to design more efficient fault tolerant network [6].

Disaster such as power outage and natural hazard may hamper Data Center management. The technique of disaster recovery and classification of disaster tolerance will improve availability of Data Centers and to accelerate the employment of Disaster Recovery [7].

A survey has been made on the fault tolerance architecture in cloud computing to understand the fault and fault tolerance in cloud network [8].

Sometimes it is a critical problem is to select a set of servers and links in the physical data center of a cloud to satisfy the request in a manner that minimizes the amount of reserved resources. The solution of this problem is by dynamically clustering the requested virtual data center and jointly optimizing virtual machine and virtual link allocation [9].

The comparative study between a set of Virtual Machine allocators for Cloud Data Centers to perform the joint allocation of computing and network resources. For each server, different allocators choose the network path to minimize electrical power consumption in the network [10].

The data storage management system on cluster based private cloud data center discusses how data is stored in the cluster based private data center [11].

Fault management in distributed system discusses how faults are managed in a distributed system [12].

Fault Tolerant System Architecture is also discussed in [13] to show how faults are handled in cloud network.

The method of Fault Detection in Cloud Computing systems is discussed in [14] and a survey of the Fault Tolerance techniques used in cloud storage is discussed in [15].

6. Proposed Work

We have proposed the following algorithm to choose a best fit data center using all the attributes of the data centers. Here, we have considered all the data centers as a set of nodes.

Step 1: A set of nodes { $N_1, N_2, N_3, ..., N_r$ } are present in the cloud network. work_load[], storage[], traffic[], network_speed[], latency[] are used to measure workload, storage, traffic growth and the network speed of each node.

```
Step 2: for i=1 to r
        Count[i]=0
// Initial Count Values are 0 for all the nodes

Step 3: min_work_load= work_load[0]
        For i=1 to r
        if work_load[i]< min_work_load
        then
        min_work_load= work_load[i]
// Choosing the node that has minimum workload

        for i=1 to r
        if workload[i] = min_work_load
        then
        Count[i]=Count[i]+1
// Incrementing the counter for the node that has minimum work_load

Step 4: Max_Storage= Storage [0]
        for i=1 to r
        if Storage[i]> Max_Storage
        then
        Max_Storage= Storage[i]
 // Choosing the node that has maximum storage

        for i=1 to r
        if Storage[i] = Max_Storage
        then
        Count[i]=Count[i]+1
 // Incrementing the counter for the node that has maximum storage

Step 5: min_traffic= traffic [0]
        for i=1 to r
        if traffic [i]<min_traffic
        then
        min_traffic = traffic[i]
// Choosing the node that has minimum traffic growth.
        for i=1 to r
        if traffic[i] = min_traffic
        then
        Count[i] =Count[i]+1
// Incrementing the counter for the node that has minimum traffic growth.

Step 6: max_network_speed= network_speed[0]
```

```
        for i=1 to r
        if network_speed[i]>max_network_speed
        then
        max_network_speed = network_speed[i]
// Choosing the node that has maximum network speed.

        for i=1 to r
        if network_speed[i] = max_network_speed
        then
        Count[i]=Count[i]+1
// Incrementing the counter for the node that has maximum network speed.

Step 7: min_latency= latency [0]
        for i=1 to r
        if latency [i]<min_latency
        then
        min_latency = latency[i]
// Choosing the node that has minimum latency.

        for i=1 to r
        if latency[i] = min_latency
        then
        Count[i] =Count[i]+1
// Incrementing the counter for the node that has minimum latency.

Step 8: Max_network_speed=Count [0]
        for i=1 to r
        if Count[i]>Max_Count
        Max_Count=Count[i]

// Choosing the node that has maximum count value.
```

We should find the median for each property of all the nodes. Now, properties of each node will be checked with other nodes whether they are same or closer to the median value or not, if so, then that node will be chosen as an ideal data center. If more than one node is same, then it will be required to choose any one of them as a data center.

Data Center Allocation is necessary when there is any fault in the data center that presently serves the request of the servers associated with it. When a fault is detected, it is necessary to move data from faulty data center to another data center. This is called Data Center Migration. Data Center has two different data backup system.

Local Backup: In local backup system data will be stored in the datacenter itself.

Centralized Backup: For centralized backup of data we need to form replica of those data centers so that during migration it is easier to fetch data from the replica servers.

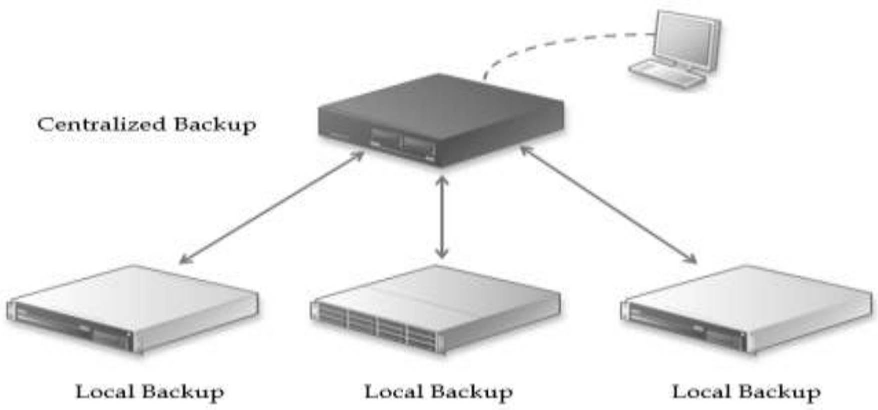

Figure 1. Local Backup and Centralized Backup

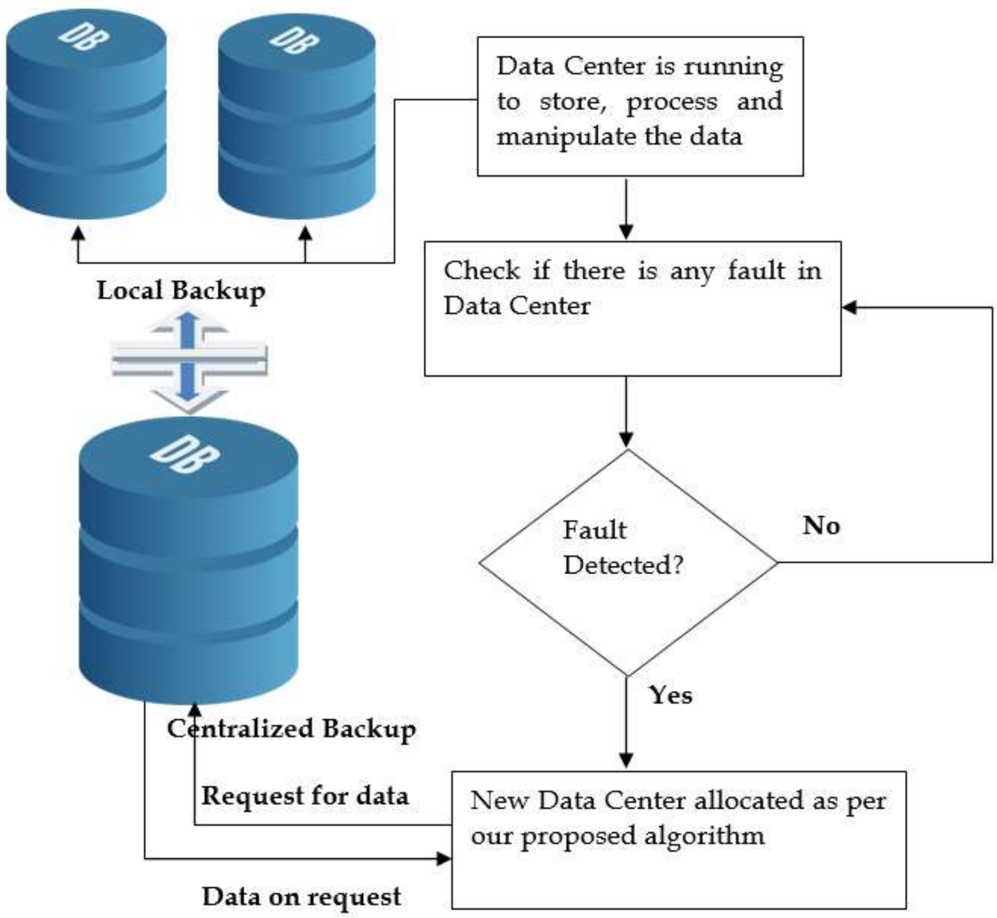

Figure 2. Flowchart of Data Center Allocation

Let us consider $D_1, D_2, D_{3...}, D_n$ be n number of data centers. We have observed some real-time data about different functional attributes. Data Center D_1 has Workload from the year 2015 to 2020 are 136 million, 189.8 million, 255.4 million, 322.0 million, 383.3 million and 440.0 million. Data Center D_1 has storage utilization from the year 2015 to 2020 are 171 exabyte, 251 exabyte, 370 exabyte, 513 exabyte, 689 exabyte and 915 exabyte. Data Center D_1 has Traffic Growth in Asia pacific region from 2015 to 2020 are 908 million, 1367 million, 1871 million, 2387 million, 2923 million and 3469 million. Data Center D_1 has Network Speed for downloading are 33.9 Mbps, 32.9 Mbps and 19.0 Mbps, 19.3 Mbps for uploading. Data Center has Latency 26 ms and 30 ms.

Data Center D_2 has Workload from the year 2015 to 2020 are 132.0 million, 191.8 million, 250.4 million, 300.0 million, 368.3 million and 445.0 million. Data Center D_2 has storage utilization from the year 2015 to 2020 are 168 exabyte, 245 exabyte, 350 exabyte, 510 exabyte, 680 exabyte and 910 exabyte. Data Center D_2 has Traffic Growth in Asia pacific region from 2015 to 2020 are 900 million, 1360 million, 1800 million, 2300 million, 2920 million and 3460 million. Data Center D_2 has Network Speed for downloading are 32.0 Mbps, 31.0 Mbps and 18.0 Mbps, 18.3 Mbps for uploading. Data Center has Latency 24 ms and 31 ms.

Data Center D_3 has Workload from the year 2015 to 2020 are 125.0 million, 180.8 million, 245.4 million, 310.0 million, 362.5 million and 420.0 million. Data Center D_3 has storage utilization from the year 2015 to 2020 are 160 exabyte, 240 exabyte, 330 exabyte, 520 exabyte, 700 exabyte and 905 exabyte. Data Center D_3 has Traffic Growth in Asia pacific region from 2015 to 2020 are 910 million, 1370 million, 1810 million, 2320 million, 2910 million and 3470 million. Data Center D_3 has Network Speed for downloading are 34.0 Mbps, 32.0 Mbps and 20.0 Mbps, 21.5 Mbps for uploading. Data Center has Latency 25 ms and 33 ms.

For a year, suppose 2017, we can compare each attribute to get the most appropriate data center among the available data centers in the cloud network.

Minimum Workload is 245.4 million for D_3

Minimum Storage 370 exa byte for D_1

Minimum Traffic 1800 million for D_2

Maximum Network Speed is 34 mbps for D_3

Minimum Latency is 24 ms for D_2

Now, we will calculate the median for the given data for the year 2017.

Table 1. Median Table

Year	work_load in million	storage in exabyte	traffic in million	network_speed in mbps	latency in ms
2017	250.4	350	1810	34.0	25

Either D_3 or D_2 can be selected as a data center as their advantageous attributes are nearer to the median.

7. Results

Table 2. Various Attributes of Data Center

Year (2017)	work_load in million	storage in exabyte	traffic in million	network_speed in mbps	latency in ms
Data Center	D2	D1	D2	D3	D2
Advantageous Value	245.4	370	1800	34	24
Median Value	250.4	350	1810	34	25

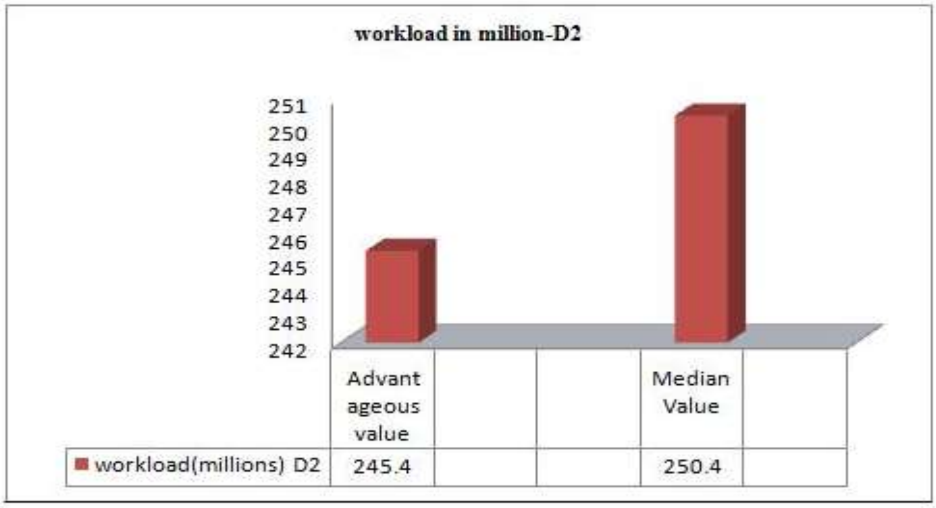

Figure 3. Work Load Comparison

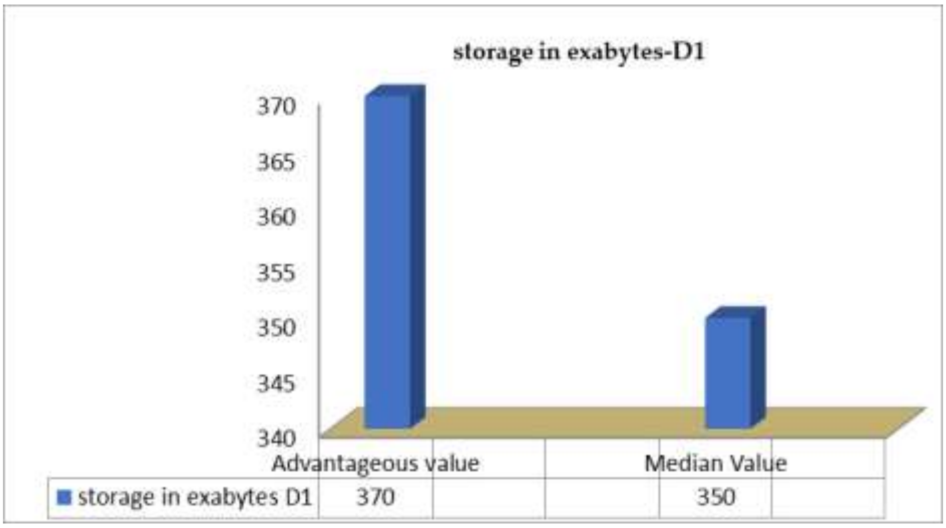

Figure 4. Storage Comparison

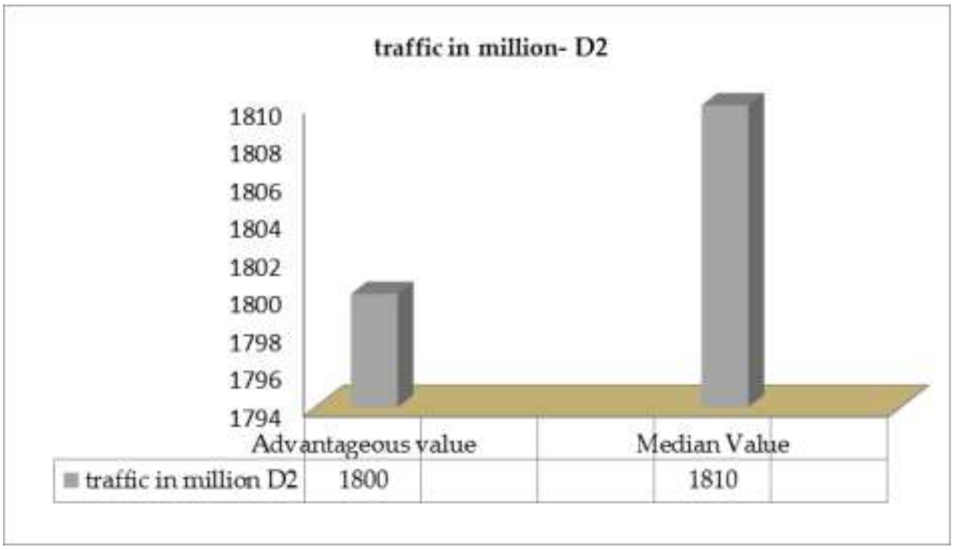

Figure 5. Traffic Comparison

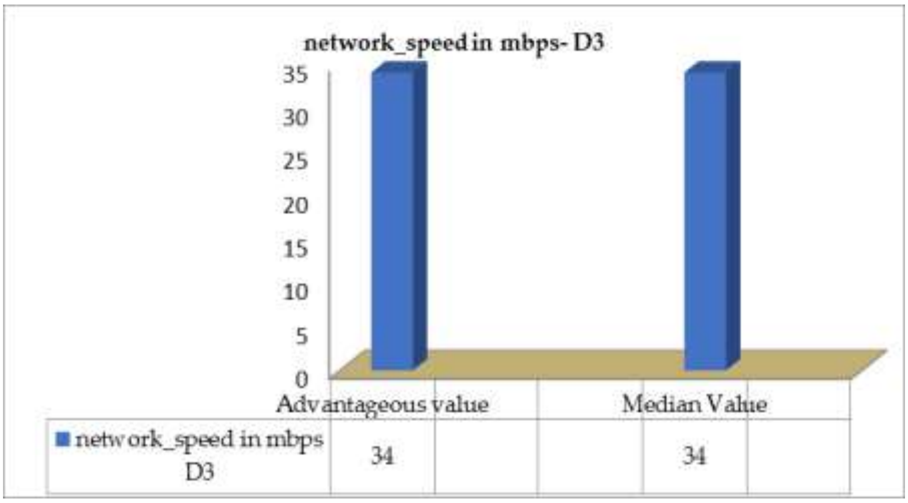

Figure 6. Network Speed Comparison

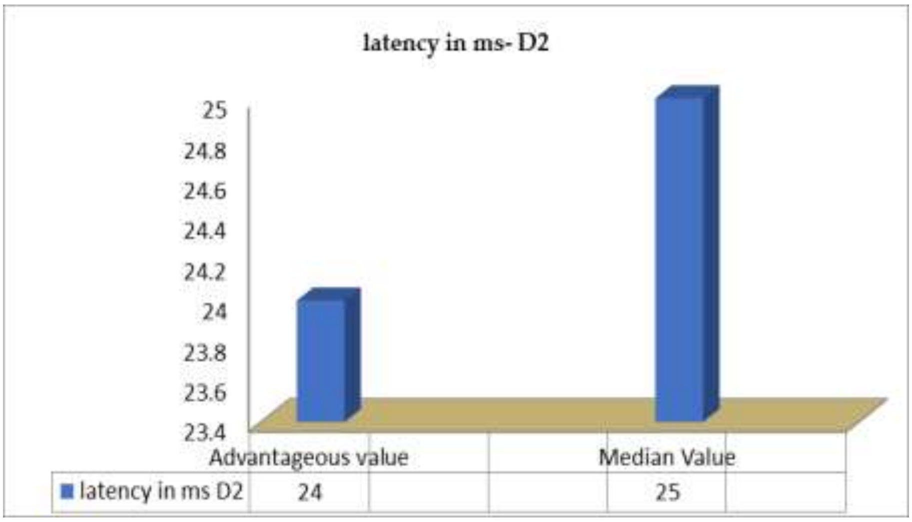

Figure 7. Latency Comparison

8. Conclusion

We have studied on the data centers and their desirable characteristics. Furthermore, we have vividly analyzed the services offered by the data centers globally. An exhaustive study has also been made towards fault detection and fault tolerance and we have collected plentiful realistic test data against all the attributes of the data centers from the various sources to work with our proposed algorithm which will suffice to bring a newer dimension in the open research problem – if a fault is detected then how the best fit data center will be chosen instead the faulty one.

Acknowledgments

We would like to convey our hearty respect and sincerest gratitude to our respected research advisor Asst. Prof. Dr. Debabrata Sarddar for sparing his valuable time and assimilating new flawless ideas at every stage of our work. Without his kind co-operation, motivation, and careful guidance, we would not have been able to carry out this research work successfully.

References

[1] Q. Guan, Z. Zhang and S. Fu, "A Failure Detection and Prediction Mechanism for Enhancing Dependability of Data Centers", International Journal of Computer Theory and Engineering, vol. 4, no. 5, **(2012)** October, pp. 726-730.

[2] Z. Amin, N. Sethi and H. Singh, "Review on Fault Tolerance Techniques in Cloud Computing", International Journal of Computer Applications, vol. 116, no. 18, **(2015)** April.

[3] T. Benson, A. Akella and D. A. Maltz, "Network Traffic Characteristics of Data Centers in the Wild", IMC'10 Proceeding of the 10th Proceeding of the 10th ACM SIGCOMM conference on Internet measurement, pp. 267-280.

[4] Cisco Global Cloud Index: Forecast and Methodology, White Paper, 2015–2020.

[5] J. Deng, S. C.-H. Huang, Y. S. Han and J. H. Deng, "Fault-Tolerant and Reliable Computation in Cloud Computing", Intelligent Automation, Inc., Rockville, MD, USA, pp. 2-5.

[6] A. Bala and I. Chana, "Fault Tolerance- Challenges, Techniques and Implementation in Cloud Computing", IJCSI International Journal of Computer Science Issues, vol. 9, no. 1, **(2012)** January, pp. 288-293.

[7] T. Zhu, Y. Xie, Y. Song, W. Zhang, K. Zhang and F. Gao, "IT Disaster Tolerance and Application Classification for Data Centers", International Congress of Information and Communication Technology (ICICT 2017), Elsevier ScienceDirect Procedia Computer Science, vol. 107, **(2017)**, pp. 341-346.

[8] A. Hammadi and L. Mhamdi, "A survey of fault tolerance architecture in cloud computing", Elsevier, Computer Communications, vol. 40, **(2014)** March, pp. 1-21.

[9] L. Shi, D. Katramatos and D. Yu, "Virtual data center allocation with dynamic clustering in clouds", Performance Computing and Communications Conference (IPCCC), 2014 IEEE International, **(2014)** December 5-7.

[10] G. Portaluri, D. Adami, A. Gabbrielli, S. Giordano and M. Pagano, "Power Consumption-Aware Virtual Machine Allocation in Cloud Data Center", Globecom Workshops (GC Wkshps), 2016 IEEE, **(2016)** December 4-8.

[11] C. Cho Khaing and T. Thu Naing, "The Efficient Data Storage Management System on Cluster-Based Private Cloud Data Center", Proceedings of IEEE CCIS2011.

[12] W. Zhou, "Fault Management in Distributed Systems", University of Pennsylvania ScholarlyCommons Technical Reports (CIS), **(2015)** May 1.

[13] K. Nørvag, "An Introduction to Fault-Tolerant Systems", IDI Technical Report 6/99, Revised July 2000, ISSN 0802-6394.

[14] Y. Jiang, J. Huang, J. Ding and Y. Liu, "Method of Fault Detection in Cloud Computing Systems", International Journal of Grid Distribution Computing, vol. 7, no. 3, **(2014)**, pp. 205-212.

[15] S. Mansur Hosseini and M. Ghobaei Arani, "Fault-Tolerance Techniques in Cloud Storage: A Survey", International Journal of Database Theory and Application, vol. 8, no. 4, **(2015)**, pp. 183-190.

Authors

Debabrata Sarddar, Assistant Professor in the Department of Computer Science and Engineering, University of Kalyani, Kalyani, Nadia, West Bengal, INDIA. He has done Ph.D. at Jadavpur University. He completed his M.Tech in Computer Science & Engineering from DAVV, Indore in 2006, and his B.E in Computer Science & Engineering from NIT, Durgapur in 2001. He has published around 200 research papers in different journals and conferences. His research interest includes wireless and mobile system, Cloud Computing and WSN.

Manas Kumar Sanyal is currently working as a professor in the Department in the Business Administration, University of Kalyani, Kalyani, Nadia, West Bengal, India. He is currently the Dean of faculty of Engineering, Technology and Management of the University of Kalyani. He has completed M.Tech and Ph.D. in Computer Science from Calcutta University. He is serving more than 18 years in teaching. He has presented papers in national and international Conferences and published papers in international journal. His research interest includes e-Governance, E-Commerce, ERP, Data warehouse, cloud computing, Semi Structured Data base.

Sougata Chakraborty is a Senior System Engineer at IBM India Private Limited in Kolkata. He completed M.Tech in Computer Science & Engineering from Jadavpur University in 2011. He had also completed his B.Tech in Information Technology from Murshidabad College of Engineering & Technology under West Bengal University of Technology in 2008. His research interests include Cloud Computing and Mobile Computing.

Priyajit Sen is presently working as an Assistant Professor in the Department of Computer Application, B.P. Poddar Institute of Management and Technology, Kolkata, India. He has passed M.Tech in Computer Science and Engineering at the Department of Computer Science and Engineering, University of Kalyani, Kalyani, Nadia, West Bengal, India in 2017. He has completed his MCA from Department of Computer Science and Engineering, University of Kalyani, Kalyani, Nadia, West Bengal, India in 2015. His research interest includes Mobile Computing, Wireless Sensor Network and Cloud Computing.