American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-5, Issue-11, pp-278-284
www.ajer.org
Research Paper Open Access

Deployment of Data Center Network using B-Tree Methodology

Dr. Vibhakar Pathak[1], Suman Saurabh Sarkar[2]
[1]*Professor: Department of Computer Engineering Arya College of Engineering and IT Kukas, Jaipur, Rajasthan.*
[2]*M.Tech Scholar: Department of Computer Engineering Arya College of Engineering and IT Kukas, Jaipur, Rajasthan*

ABSTRACT: *Cloud computing is one of the major and fastest growing platform for making resources available to the users. Only required for the user is the account with any of the cloud service provider. For user only thing to remember is the user id and the password for the account and the knowledge for the operation of services he is seeking for. But for the developers of cloud it is much more. They need to decide the flow of resources that depends upon the background network. In terms of cloud computing this background network is known as Data Center Network. In this paper we have deployed the Data Center Network on the concept of B-Tree Topology. We have taken this topology in the account as the size of tree remains constant and it is very easy to estimate general behavior of the data flow. In this work we have also taken in account the overheads of network as the ultimate media of communication is the wired or wireless media and congestion and jitters are always expected in these types of networking media. With this we will explain the concept of traffic engineering. The major role to manage the data on the route is to put on the concepts of traffic Engineering.*

Keywords: *DCN (Data Center Network), Traffic Engineering, Virtualization FAT Tree, Ethernet, B-Tree, DSS, DVS*

I. INTRODUCTION

In the long tenure of technology development in the era of cloud computing it has been noticed that the major challenge is the storage. More than million of TB of data is to be stored to serve clients who all are moving forward to access data via internet to save expenses on permanent but not stable resources. With this a new problem is generated that how to serve clients with efficient speed. *Solution* has been observed as distributed storage for which the base model is Data Center Network. Another challenge that took place is how to make efficient distribution of cloud network. Lots of work has already been done. Various topologies as per the concept of data structure has been obtained that includes Minimum Spanning Tree, Fat Tree etc. We have tried to implement the same concept but with different topology i.e. B-Tree Structure. The approach what we have used is slightly different from the previous one as the later one remains fixed in size and always decided by the developer. In the approach to support our work we will explain the working and the result set of the previously done work on FAT Tree, Ethernet and then we will discuss the benefits of B-Tree over the previously developed methodologies. Before this we will explain DCN and traffic engineering that plays an important role in maintaining the flow of data among the Data Centre Networks. With this to make our platform more energy efficient we have also tried to make our deployment as a virtual environment so we will give a small understanding of virtual environment too.

Virtualization

When we say virtualization techniques then we must first understand what exactly virtualization is. As a description of requirement in simple words Virtualization is the path to future developments. And as in description of technical requirement we can explain virtualization as the most convenient and cheap source of computation. However to setup virtual environment is not very easy but after setup advantages gives us motivation to adopt it with good will.

II. CLOUD COMPUTING WITH VIRTUALIZATION

The combination of cloud computing with virtualization techniques is unique possibility for providing the resources and raw data to end user in cheapest way. Now further we will see some points in favor of this hybrid technology. Some points are as follows:

1) In cloud computing users are facilitate for using application (SaaS, i.e. Software as a Service) as well as the computing infrastructure (IaaS, i.e. Infrastructure as a Service).These facilities have to be provided by the Cloud Service Provider.
2) Advantage of this technology is that now user need not to own any kind of infrastructure and so no maintenance is required that utilizes lots of space, money, time, manpower. And with this most important mental stress. A user can order services of demand dynamically and will pay only for the "Available to Access" service.
3) With data availability cloud computing facilitates with operation on data and also updating value set and all this we are attempting on virtual environment.

III. VIRTUALIZATION OF CLOUD COMPUTING REFERS TO THREE MAJOR TASKS

To adopt virtualization in cloud computer infrastructure it is required to see the complete environment of cloud computing as a virtual environment. To accomplish this we have to perform some task and make the virtualized environment by following procedure:

1.) Virtualization of the Computing Resources.
2.) Virtualization of Storage
3.) Virtualization of Network with Security Services.

Basics of technologies used in implementation:

A) Layer 2 semantics in Data Center - Ethernet is a most dominant LAN technology which is oldest, popular, fast and most fundamental media of communication. It is the largest source of networking which is widely used and costs very low as compare to other networking technologies. In data center the high speed communication leverages proven, low cost hardware implementation of layer 2 devices. This is just a LAN where servers are interconnected to any port on Ethernet switch and make use of flat addressing.

Ethernet frame format- As mentioned above, Ethernet is CSMA/CD system, which means when a station has packet o send, it monitors the traffic or signal on a cable, voltage signifies other station is transmitting so it avoids sending the packet. Collision detection allows stations to see what other is doing (unicast/broadcast). This uses components of Ethernet such as coaxial cables, repeaters, hubs etc. The network is segmented so station on one segment might not see what other is doing until the packet is destined for itself. But broadcasts can be seen. Here, VLAN are used to separate broadcast domain as well, which restricts communication between stations on different virtual LAN's via router only i.e. at layer 3.
The Ethernet standard frame format is IEEE 802.3 which is usually of size 1518 bytes and minimum size is 64 bytes. The bits are transmitted in little endian format over the network which is different from FDDI or token ring that sends in MSB (most significant bit first).

B) Data center is a kind of core repository where data is stored, managed and distributed. This may have physical or virtual infrastructure. Different organizations uses data center for their purpose, business people use data center for better understanding their customers or scientist may use it for doing future predictions about weather forecasting say or gaining insight into data for retrieving vital information.
Large-scale data centers also enables ever-growing cloud computing and provide the core infrastructure to meet the computing and storage requirements for both enterprise information technology needs and cloud-based services. For supporting this servers are increasing in data centers exponentially which in turn leads to enormous challenges in designing an efficient and cost-effective data center network.

C) Traffic engineering involves adapting the routing of traffic to the network conditions, with the joint goals of good user performance and efficient use of network resources. Traffic engineering involves adapting the routing of traffic to the network conditions, with the joint goals of good user performance and efficient use of network resources.

Traffic engineering represents a methodology that improves the network performance by manipulating the flow of data in the network. This manipulation is done after analyzing the network and then making a forecast. Traffic engineering is now a requirement in different networking patterns like LAN, WAN, Ethernet Networks, PSTN and in cloud computing too. In cloud computing the major role id of the data centers that acts as a source of data among the information seekers. In this research we will follow following approaches for making efficient TE:

1.) Minimizing the maximum link usage in Ethernet Network
2.) Minimizing the maximum link usage in DCN
3.) The worst-case maximal link utilization in the DCNs with multiple spanning trees in case of link failures.
4.) Multi- Objective TE for DCN with Multiple Spanning Trees.

For above problems we have three parameters to take care of

1.) Maximum Link Utilization
2.) Total Network Load
3.) Number of links occupancy on which load has been distributed.

D) Concept of Virtualization in Switching Networks The advantages can be further understood by illustrations below:

1.) Switches are managed for modifying operations of switch such as enabling port mirroring or spanning tree protocol, creating virtual LAN's etc. VSS (virtual switching system) reduce this overhead of managing switches by 50% as there is single point of management i.e. is simple network. Hence improves efficiency of operations.
2.) Less overhead of configuring switches which are redundant because one configuration file has to be managed at node.
3.) Easy recovery in case of failure of virtual switch by using Ethernet and port aggregation protocol. This results in seamless communication for those applications that require network state information. Example applications that forwards information of routes, network address translators etc.
4.) VSS can be deployed at data center (server connectivity/distribution layer) that helps in scaling better by link load sharing.

E) Local Search TE problem's solution is always taken as a possible solution. One of the possible solutions is the Local Search Methodology. Local search facilitates us as powerful technique that can find high quality solutions for computationally complex optimization problems. This method starts with an initial possible solution and the quality of the solution improves with iterative movement from self's solution to neighbor's solution. The iteration stops on getting some termination command. LS Algorithm makes use of neighbor functions that defines set of neighbors for every solution.

Steps taken for making efficient Local Search:

1.) Local search algorithm has been reviewed and redeveloped for solving TE problem of minimization of maximum link utilization in Ethernet Network.
2.) Local search approach has been extended for network containing one spanning tree to cope up with the minimization of maximum link utilization in DCN of many VLAN.
3.) An algorithm is being proposed for minimizing the cases of that are worst in maximum link utilization and the service disruption in case of link failure.
4.) A Methodology is defined to search for problem in spanning tree space instead of link cost space. This approach helps in reducing the search time as the search space reduces significantly.
5.) Few assessment methods have been implemented for assessing the performance of these Multi objective TE methods for DCN.

IV. EXPERIMENTAL SETUP

To proceed for implementing the work we have taken following steps:

1.) Imported CloudSim in JVM.
2.) Developed a Program for Distribution of Data Center Network.
3.) Associated IP Address as Primary Key of the Distribution
4.) Associated Geographical Distance to each node represented by IP Address.
5.) Travelling time is measured in Microseconds from root node (Start of the DCN) to every Child Node (Storage of DCN).
6.) Compared the travelling time with the travelling time in DCN based on Ethernet Network.
7.) Proved that B-Tree based Distribution is much more efficient that the Ethernet based DCN Distribution.

V. RESULT GENERATED

Result has been generated in order to prove the significance of B-Tree architecture based Data Center Network over the Ethernet architecture of FAT Tree architecture based Data Center Network.
Following are the results:

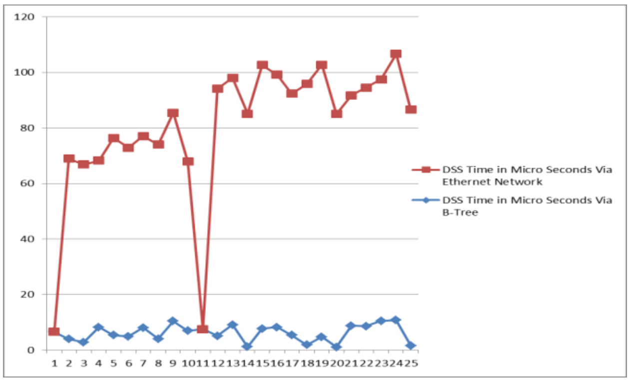

Fig 1. Comparision of Ethernet network and BTree routing delay in microseconds (as per TABLE I below)

In order to make this work more specific we have shown a tabular comparison of time taken by a 100 MB packet to travel from Aggregate switch to DSS node.
In the table we have mentioned the comparison between time in microseconds taken by B-Tree architecture and the time that was mentioned earlier taken by the Ethernet architecture.

IP Address	Geographical Distance	DSS Time in Micro Seconds Via B-Tree	DSS Time in Micro Seconds Via Ethernet Network
192.168.1.10	500	6.546	
192.168.1.20	600	3.9865	65
192.168.1.30	700	2.851	64
192.168.1.40	800	8.2376	60
192.168.1.50	900	5.386	71
192.168.1.60	1000	4.939	68
192.168.1.70	1100	8.098	69
192.168.1.80	1200	3.997	70
192.168.1.90	1300	10.452	75
192.168.1.100	1400	6.9407	61
192.168.1.110	1500	7.566	
192.168.1.120	1600	5.1271	89
192.168.1.130	1700	9.0839	89
192.168.1.140	1800	1.1448	84
192.168.1.150	1900	7.7254	95
192.168.1.160	2000	8.19746	91
192.168.1.170	2100	5.36	87
192.168.1.180	2200	1.8952	94
192.168.1.190	2300	4.7593	98
192.168.1.200	2400	1.066	84
192.168.1.210	2500	8.741	83
192.168.1.220	2600	8.5785	86
192.168.1.230	2700	10.508	87
192.168.1.240	2800	10.746	96
192.168.1.250	2900	1.6268	85

TABLE I.

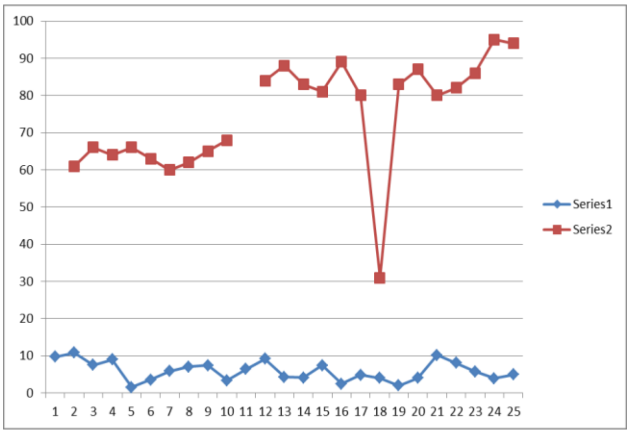

Fig.2 Comparision -- Ethernet network and BTree routing delay in micro seconds (as per TABLE II below.)

IP Address	Geographical Distance	DSS Time in Micro Seconds Via B-Tree	DSS Time in Micro Seconds Via Ethernet Network
192.168.1.10	584.005	9.716	
192.168.1.20	649.6242	10.841	61
192.168.1.30	1155.2815	7.479	66
192.168.1.40	1409.164	8.9629	64
192.168.1.50	1608.7232	1.57	66
192.168.1.60	1521.336	3.514	63
192.168.1.70	1770.442	5.87	60
192.168.1.80	1662.7.1	7.075	62
192.168.1.90	2035.771	7.445	65
192.168.1.100	1438.861	3.405	68
192.168.1.110	2307.0356	6.443	
192.168.1.120	2273.357	9.142	84
192.168.1.130	2639.234	4.255	88
192.168.1.140	2442.009	4.0966	83
192.168.1.150	2810.949	7.426	81
192.168.1.160	2064.908	2.4281	89
192.168.1.170	2609.0784	4.814	80
192.168.1.180	2327.623	4.033	31
192.168.1.190	3218.546	2.0146	83
192.168.1.200	3201.412	3.987	87
192.168.1.210	2741.926	10.142	80
192.168.1.220	3450.321	8.0489	82
192.168.1.230	3538.929	5.768	86
192.168.1.240	3773.794	3.909	95
192.168.1.250	3772.0432	4.9698	94

TABLE II

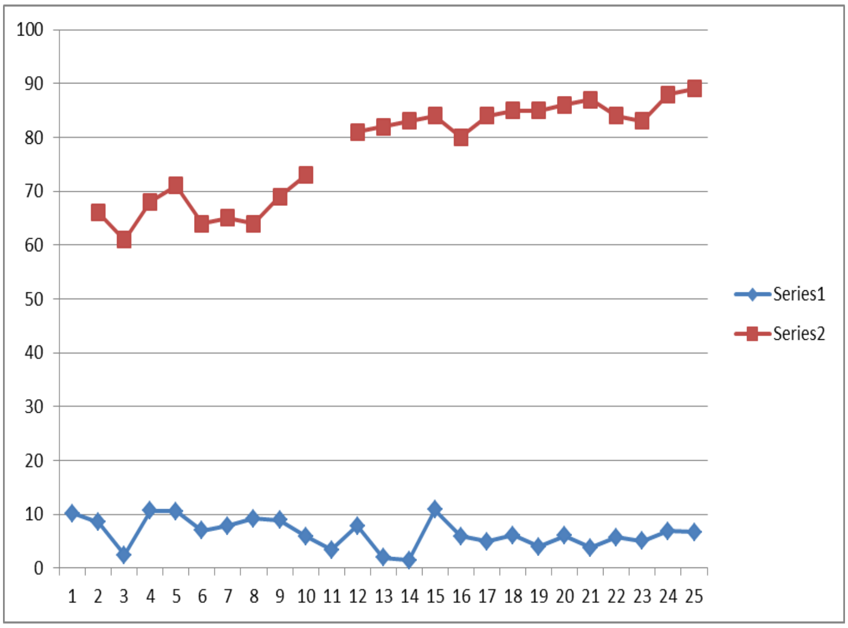

Fig 3. Comparision -- Ethernet network and BTree routing delay in micro seconds (as per TABLE III below.)

IP Address	Geographical Distance	DSS Time in Micro Seconds Via B-Tree	DSS Time in Micro Seconds Via Ethernet Network
192.168.1.10	891.56	10.145	
192.168.1.20	1228.037	8.525	66
192.168.1.30	906.578	2.46	61
192.168.1.40	1870.9301	10.701	68
192.168.1.50	1599.796	10.581	71
192.168.1.60	1861.574	6.98	64
192.168.1.70	2063.184	7.88	65
192.168.1.80	1892.01	9.254	64
192.168.1.90	1435.709	8.891	69
192.168.1.100	2199.106	5.918	73
192.168.1.110	2257.284	3.425	
192.168.1.120	2166.969	7.892	81
192.168.1.130	2572.459	2.0035	82
192.168.1.140	2696.583	1.458	83
192.168.1.150	2000.71	10.883	84
192.168.1.160	2270.94	5.882	80
192.168.1.170	2670.809	4.946	84
192.168.1.180	2965.228	6.116	85
192.168.1.190	3128.689	4.042	85
192.168.1.200	3004.045	6.049	86
192.168.1.210	3071.357	3.806	87
192.168.1.220	2851.718	5.747	84
192.168.1.230	3033.858	5.027	83
192.168.1.240	3236.428	6.929	88
192.168.1.250	3961.925	6.667	89

TABLE III

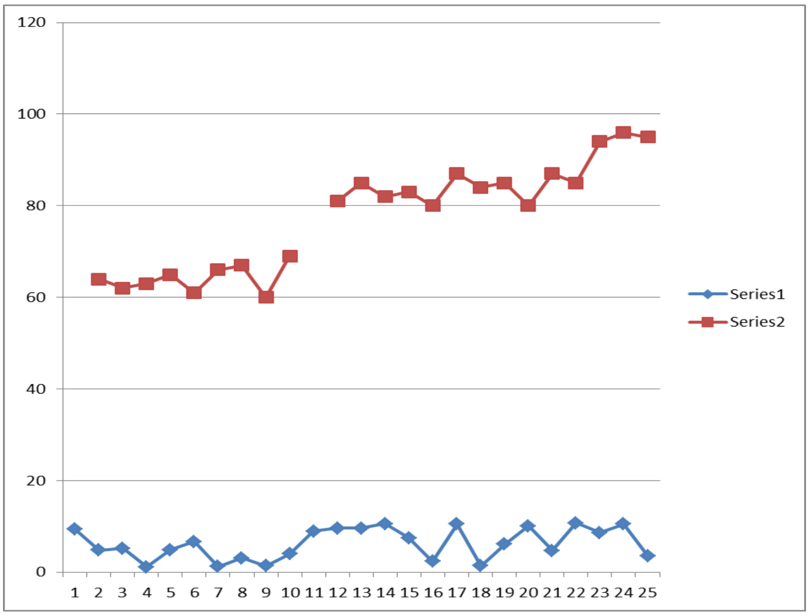

Fig. 4. Comparision -- Ethernet network and BTree routing delay in micro seconds (as per TABLE IV below.)

IP Address	Geographical Distance	DSS Time in Micro Seconds Via B-Tree	DSS Time in Micro Seconds Via Ethernet Network
192.168.1.10	1184.85	9.329	
192.168.1.20	989.465	4.81	64
192.168.1.30	1623.384	5.24	62
192.168.1.40	1164.442	1.15	63
192.168.1.50	1210.672	4.86	65
192.168.1.60	1657.987	6.69	61
192.168.1.70	1710.745	1.21	66
192.168.1.80	1302.1193	3.14	67
192.168.1.90	1446.731	1.37	60
192.168.1.100	1804.252	3.99	69
192.168.1.110	1830.343	9	
192.168.1.120	1901.546	9.644	81
192.168.1.130	2185.647	9.65	85
192.168.1.140	2390.969	10.59	82
192.168.1.150	2860.313	7.53	83
192.168.1.160	2411.59	2.35	80
192.168.1.170	2391.962	10.52	87
192.168.1.180	2270.624	1.36	84
192.168.1.190	2802.877	6.11	85
192.168.1.200	3051.053	10.14	80
192.168.1.210	2933.269	4.75	87
192.168.1.220	3216.196	10.75	85
192.168.1.230	3091.444	8.66	94
192.168.1.240	3435.365	10.52	96
192.168.1.250	3198.634	3.61	95

VI. CONCLUSION

The distribution of data center network based on B Tree methodology is found more efficient than distribution of DCN on Ethernet based architecture. This result has been judge because the size of B Tree is always fixed hence there is no scope of dynamic overheads. Still to make this result more accurate we have tried to include randomly generated overheads.

REFERENCES

[1]. searchservervirtualization.techtarget.com
[2]. Towards Virtual InfiniBand Clusters with Network and Performance Isolation
[3]. https://en.wikipedia.org/wiki/Virtualization
[4]. standards.ieee.org/events/ethernet/history.html
[5]. VL2: A Scalable and Flexible Data Center Network [A Microsoft Research]
[6]. http://www.rhyshaden.com/eth_intr.htm
[7]. https://tools.ietf.org/html/rfc2702
[8]. Traffic Engineering With Traditional IP Routing Protocols Bernard Fortz Jennifer Rexford Mikkel Thorup Institut d'Administration et de Gestion Internet and Networking Systems Universite Catholique de Louvain AT&T Labs Research
[9]. Traffic engineering techniques for data center Pr. Y. Deville (Promoteur) ICTEAM, UCL, Belgium and Pr. O. Bonaventure (Promoteur) ICTEAM, UCL, Belgium
[10]. http://www.cisco.com/c/en/us/products/collateral/switches/catalyst-6500-virtual-switching-system-1440/prod_qas0900aecd806ed74b.html
[11]. http://ieeexplore.ieee.org/abstract/document/6990724/?reload=true
[12]. accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.