A REVIEW ON: NETWORKING IN CLOUD DATA CENTER

Ruchi Pandey[1], Parikshit Tiwari[2]
[1]Student, [2]Assistant Professor
Department Of Computer Science: Rewa Institute of Technology, Rewa (MP)

Abstract: ***This paper focuses on the systematic techniques for distribution of network traffic. The study of network traffic distribution is done on a cloud computing network. Cloud computing is an emerging paradigm of the computing world, which aims to share data over a single platform. Traffic engineering (TE) is known as an optimization method for improving the performance of a telecommunication network by analyzing, forecasting and efficiently adjusting the traffic (data) transmission in the network. Load balancing is an important aspect of cloud computing environment. Efficient load balancing scheme ensures efficient resource utilization by provisioning of resources to cloud user's on demand basis in pay-as-you-say manner. This paper represents a view of B Tree algorithm or techniques in cloud computing.***

Key words: Ethernet, Cloud sim, Traffic Engineering, B-tree.

I. INTRODUCTION

Traffic Engineering is an essential task in all kinds of telecomunication networks, including the Public Switched Telephone Networks (PSTN), Local Area Networks (LAN), Ethernet networks, Wide Area Networks (WAN), Cellular Telephone networks and the Internet networks. Cloud computing has generated a lot of interest and competition. This is due to service delivery model which provides that involves computing and storage for users in all market including financial, health care and government [1]. This advances in computing put cloud computing as one of the latest developments of computing models. Its development can be considered much advanced than that of distributed computing, parallel processing, grid computing and so on. With cloud computing, multi-level virtualization and abstraction can be achieved [2]. This can validated using an effective integration of variety of computing, storage, data, applications and other resources. Data centers with huge logical storage capacity are now base of many Internet and cloud computing services. Deploying by switched Ethernet networks, data centers are mostly dominated by different variants of Spanning Tree Protocol. TE problems in Data Center Networks (DCN) are very varied, both in problem formulations and evaluation criteria (objectives). Some typical problem formulations for TE in DCNs are off-line TE, online TE, Virtual LAN (VLAN) to spanning tree mapping, traffic flow to spanning tree assignment, and multiple spanning tree construction. The ultimate objective of a TE method is to avoid the congestion in the DCN. This paper has following sections including: literature survey, methodology used and conclusion.

II. PREVIOS WORK

Traffic engineering in Ethernet Network and Data centers is being now the focused research topic. TE in large DCN now has become a desire for many service providers. TE has been managed in different ways using different aspects. In the literature review we tried to focus only on some basic studies on data centers and different types of spanning tree implementation. Data centers have now made their position strong in the world of technology. Now a days daily many datacenters are being developed. Their numbers are increasing very quickly. In some report this has also been mentioned that their some data centers on earth that has more than 10000 servers to 100000 Servers. So why this is a very important technology. Answer is very simple. Fast , Quick and efficient compution is being produces provided a large set of data that could not be assessed with bare and traditional technology. Many Data centers support different types of application at the same instant of time. This means it gives the multi tasking environment on large scale with large number of users. Application runs on a set of virtual machines that are distributed on and via physical servers.

SPANNING TREE PROTOCOL:

While we are talking about Ethernet Network then we see that it contains multiple active links that ensures that there are backup links to support the situation in which any active link fails. But it is also very not fortunate that Ethernet networks does not support active cycles and in order to prevent these cycles Ethernet Switches are applied with some standard Spanning Tree Protocol. One of the most important Spanning Tree Protocol is IEEE 802.1d. This protocol help in reducing the size of topology of the switched network to a spanning tree network. Computation of the spanning tree is done by sending a specialized message called as configuration BDPU (Bridge Protocol Data Units). These messages are being sent by the Ethernet Switches. Each message has some information encapsulated as in the data packet.

Other Spanning Tree Protocols:

We have some more extended version of Spanning Tree Protocol:

- IEEE 802.1w – Rapid Spanning Tree Protocol – Designed to accelerate the convergence of the alternate spanning tree n case of link failures or if network topology changes. It has a major fault that it does not support high numbers of links in a single spanning tree.
- IEEE 802.1q – the spanning tree protocol for virtual LAN's enables large Ethernet Networks to be logically divided into Virtual LAN's. it could be used in a way that one DC user will get isolated

with another DC User. Servers of any VLAN can only be able to communicate with other servers that belong to the same VLAN.

- IEEE 802.1s – Multiple Spanning tree Protocol – Is a combination of IEEE 802.1w and IEEE 802.1q. It facilitates the service providers to benefit from the available links by spreading different spanning trees (one spanning tree per VLAN) over a single physical topology. It is to be remembered that MSTP Can only compute 16 different spanning trees and map each VLAN on to any one of these spanning trees. This lead to higher complexity as now each VLAN will require finding out its root and the cost of paths to have its own spanning tree.

TRAFFIC ENGINEERING IN DATCENTER NETWORK:
Traffic engineering for DCN can be done on the basis of many criteria. One of the criteria is traffic demand information. According to this criteria TE problems could be classified in to two main categories:

- Off-Line Traffic Engineering: This improves the performance of the switching protocols as per the given traffic demand. This demand is raised in form demand matrices. Demand matrices are generated by doing extensive study on DCN Traffic
- On-Line Traffic Engineering: In this approach routing is adapted dynamically according to the real-time changes in traffic. For this the network devices responsible for communication should work intelligently the local adjustment should be strong enough so that it could handle mold the present traffic situation.

Most of the researchers haven't talked about in local search algorithms for B tree, they talked about Fat Tree It is called Fat Tree because it is not a spanning tree like PortLand. All the ports on each switch are used to connect to the hosts or to the other switches. The Fat Tree topology is constructed as follows. All the switches in Fat Tree have k ports. There are k pods. Each pod contains k switches divided into 2 layers: Aggregation and Edge layer [2].

In the continuation of work we will focus on offline Traffic Engineering. The TE in large Ethernet Network such as DCN is defined by network's Quality of Service. The optimization of Quality of Service can be achieved on many evaluation metrics such as link utilization , link delay, number of links used, service disruption due to link failure, full network load, etc. Network operators can define TE problems with one of a set of metrics to be optimized as per the requirement. It is very impossible to take all the metrics in consideration for optimization as metrics are also interdependent of each other. This may lead to degradation of one metric if changed in another metric is reflected in order to make it more optimized.

2.2.1: Maximum Link Utilization
The main objective in any communication device based network is to make possibility of minimum congestion. This can be done by the evaluation of link utilization. This can b improved by minimizing the maximum utilization of link. Many TE Techniques are their to ensure the minimized maximum utilization of link in the Ethernet network with multiple spanning tree. The utilization of any link refer to the use of link against the available bandwidth. This can be understood by a ratio between load on link and bandwidth of link [3].
This can show a link as overloaded if load is greater than bandwidth as link load is the sum of total traffic on it so we can define it for DCN as sum of load of traffic from different VLAN's [4].

2.2.2) Total Load:
Total load on any network is considered to be the addition of loads of all the active links of that network. The simple formula for calculation of the total load having maximum links utilization is as follows:
Let L[i, j] denote the load of link (i, j), SumL denote the total load.
SumL = PL[i, j] (∀switchi, j)
We can interpret that the average link load on the network will be minimized if the total load of the network gets minimized. That is total load is always directed proportional to average link load. The TE with one spanning tree can be done by considering it as Optimum Communication Spanning Tree Problem (OCSTP or OCT).
OCT was first introduced by T.C. Hu in year 1974. It is a NP-hard combinational optimization problem that could be solved through many proposed solutions. Solution on small network are more preferable for this problem. By small network we mean those networks having dozen of nodes not more than that.
In large DCN's that supports many VLAN it is very tough to find a bound for this type of TE problem.

2.2.3) Number of Used Links
Our main objective is to anyhow reduce the number of used links. Recent studies in [5][6] tried to find the routing scheme that minimizes the number of active network elements including the number of used links in both IP and Ethernet networks.

In DCN's that are using Spanning tree this scheme of reducing used links could be useful if the no. of VLANs used in that network is more than 1. Reducing number of links also assist the network in fault tolerance. Faulty flow could be identified more easily in network having reduced number of links instead of full network comprising of high no. of used links for load balancing.

2.2.4) Service Disruption after link Failure:
TE Problem also focuses on minimization of Service Disruption. When change of topology is identified by MSTP such as if some switch or some in use link get disrupted due to any technical or non-technical reasons, then spanning tree re-computation is required in each VLAN affected by the disruption. In this minimization can be expressed in terms of number of links used to replace the failed links in each VLAN to maintain the spanning tree failure.

Local Search for Multiple Objective Optimizations:
Many researchers have been made to improve the quality of performance of local search while dealing with multiple-objective optimization. Some of the approaches are defined as follows:

- Multi-objective Steepest Descent Method (MSDM) It aims to find a Pareto descent direction for the search that maximizes the minimum degree of improvement of all objective functions. This can be done by solving a quadratic programming problem for finding the direction to move the current solution in each search step.
- Evolution Strategies (ES): This combines the Pareto archived evolution strategy Local Search with the use of population and recombination for solving a set of multiobjective knapsack problems. In each search step, ES randomly generates a new improving solution from the current solution according to a normal distribution.
- Combined Objectives Repeated Line-search (CORL) is a gradient-based LS method that computes the vectors for generating the convex cone of the Pareto descent directions. CORL is guaranteed to find the set of all non-dominated improving directions for any point in the parameter space of a multi-objective optimization problem.
- Pareto Descent Method (PDM) proposes a LS approach for moving solutions in the directions that simultaneously improve all objective functions. PDM finds the directions by solving linear programming problems.

The main drawback of these methods is that they are computationally expensive, restricted to some specific problems and inefficient for the multi-objective optimization problems with a huge solution space, such as TE in DCNs. In this thesis, we propose lightweight LS algorithms using efficient meta-heuristics and speeding up techniques that can give good solutions for dealing with the TE problems in DCN.

Proposed Methodology
For data generation techniques we use B tree which runs on cloudsim 3.0.3 as a simulator and to run the code we need Eclipse neon [7]. CloudSim provides concluded framework that enables modelling and simulation of app performance. By using CloudSim, users can focus on specific systems design issues that they are concerned about details related to cloud-based infrastructures and services. It is important because Cloud service providers offer elastic, on-demand, and measured infrastructure, platforms and software services. CloudSim, is -a toolkit for the modelling and simulation of Cloud computing environments- comes to the rescue. It provides system and behavioural modelling of the Cloud computing components. In this, we are using cloud simulator and needs java code that executes on Eclipse Neon as software's.

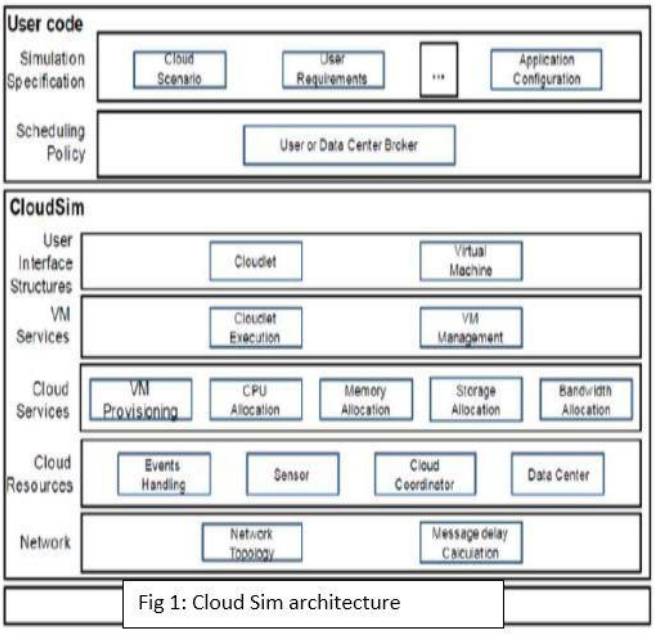

Fig 1: Cloud Sim architecture

III. CONCLUSION

This paper is based on networking cloud computing technology using data center which has a very vast potential and is still unexplored and there capabilitiesare interminable. One of the major issues of traffic engineering is load balancing because overloading of a system may lead to poor performance which can make the technology unsuccessful. So, there is eternally a requirement of efficient load balancing algorithm for efficient utilization of resources. This paper presents a proposed strategy to calculate the time by using CloudSim, which is concluded framework that enables modelling and simulation of app performance In future work I will discuss about B Tree algorithm which can help in meta-heuristics and speeding up techniques that can give good solutions for dealing with the Traffic Engineering problems in DCN.

REFERENCES

[1] Akhil Goyal,Bharti, "A Study of Load Balancing in Cloud Computing using Soft computing Techniques", International Journal of Computer Applications(0975 - 8887), Vol. 92, No. 9, April, 2014

[2] M. Al-Fares, A. Loukissas, and A. Vahdat. "A scalable, commodity data center network architecture." ACM SIGCOMM Computer Communication Review , vol. 38, no. 4, 63–74, 2008.

[3] Albert Greenberg , Parantap Lahiri , David A. Maltz , Parveen Patel , Sudipta Sengupta, Towards a next generation data center architecture: scalability and commoditization, Proceedings of the ACM workshop on Programmable routers for extensible services of tomorrow, August 22-22, 2008, Seattle, WA, USA

[4] Armbrust M, Fox A, Griffith R, Joseph A, Katz R, Konwinski A, Lee G, Patterson D, Rabkin A, Stoica

I, Zaharia M. A view of cloud computing. Communications of the ACM 2010

[5] Legrand A, Marchal L, Casanova H. Scheduling distributed applications: The SimGrid simulation framework. Proceedings of the Third IEEE/ACM International Symposium on Cluster Computing and the Grid, Tokyo, Japan, 2003

[6] Avetisyan AI, Campbel R, Gupta I, Heath MT, Ko SY, Ganger GR, Kozuch MA, O'Hallaron D, Kunze M, Kwan TT, Lai K, Lyons M, Milojicic DS, Lee HY, Soh YC, Ming NK, Luke J-Y, Namgoong H. Open cirrus: A global cloud computing testbed. IEEE Computer 2010

[7] Buyya R, Ranjan R, Calheiros RN. InterCloud: Utility-oriented federation of cloud computing environments for scaling of application services. Proceedings of the 10th International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP 2010), Busan, South Korea. Springer: Germany, 21–23 May 2010;

[8] Howell F, Mcnab R. SimJava: A discrete event simulation library for java. Proceedings of the First International Conference on Web-based Modeling and Simulation, San Diego, U.S.A.

