International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

COPY RIGHT

IJIEMR Transactions, online available on 30th Nov 2019. Link :http://www.ijiemr.org/downloads.php?vol=Volume-08&issue=ISSUE-11

Title **A NOVEL OPTIMAL DISTRIBUTED HADOOP NETWORK VIRTUALIZATIONFOR CLOUD BASED DATA CENTRE**

Volume 08, Issue 11, Pages: 406-418.

Paper Authors

G.JOEL SUNNY DEOL, DR.O.NAGARAJU

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

A NOVEL OPTIMAL DISTRIBUTED HADOOP NETWORK VIRTUALIZATIONFOR CLOUD BASED DATA CENTRE

G.JOEL SUNNY DEOL[#], DR.O.NAGARAJU[*]
[#]Research Scholar, Dept of CSE , Acharya Nagarjuna University
[*]Asst.Professor, Department of computer science, Government Degree College,Macherla-522426
[1]sunnydeolgosu@gmail.com, [2]profonr@gmail.com

Abstract— With the continuous increasing migration of application to the cloud-based data centers is also increasing the demand for optimization for data centers. The optimization for cloud based data centers focuses on enhancement of performance of virtualized computing and storage components. Multiple research efforts have been made to achieve the optimal solution for compute and storage solutions. Though in order to achieve the complete efficiency by the data centers, it is also important to realize the optimization of load balancing and data center network. The optimization of data center load balancing is already been addressed in many recent research outcomes. However, the data center network optimization is rarely been addressed by the recent research efforts.Considering the present growth in cloud based data center application and the growth in generated data by those applications, it is also increasing the challenge in high availability and network security. Thus the researchers and infrastructure developers have started considering the data center network with higher priority and distinguished focus with the light of network as a service. Hence in order to propose any further improvements, it is the demand of the study to explore the generic data center network and realize the challenges. Hence driven by the challenges in this work we propose a novel optimal distributed Hadoop data centre network considering the optimization factors. This work additionally explores and recommends the impact of the parameters influencing the data center networks. Another outcome of the work is to consider the performance of the existing network protocols with the proposed novel performance evaluation matrix and perform a comparative study of the recent research outcomes with the proposed novel optimal distributed Hadoop data centre networkto reach to the conclusion. The work demonstrates overall 73% improvement of network performance index value over the existing data center network protocols.

Keywords— NaaS, Data Center, PI Factors, Basic Tree, FAT - Tree, Valiant Load Balancing, Hadoop Framework, Optimal Topology

I. INTRODUCTION

The highest momentous improvement in the field of application hosting is the cloud computing with the enhancement and significant migration of the applications into cloud based data centers.Driven by the advantages like virtualization, pay per use cost effectiveness, elastic scalability of the resources and most importantly accessibility of the resources over sustainable internet protocols the cloud based data center are the most acceptable among small, medium and enterprise scale information technology industry and researches. The core advantage of the cloud computing is realized with

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

the technological foundation of computing, storage and network component virtualization within the cloud based data centers [1]. The most significant component of the cloud service is the internet as the services are hosted and accessed over the internet in majority of the instances. In case of the cloud based public or dedicated data centers, the customers access the services only over internet. Also from service provider aspect, the cloud based services are not only limited to computing and storage, the new dimension of cloud computing is also to provide the network as a service to the customers [2].

Also this is to be understood that the network connecting the data center to the outer world is not only the point of consideration, the network connecting different parts of the distributed data center is also the point of focus of the optimization oriented research.

With the understanding of the recent outcomes from parallel researches and statistics, it is evidently understood that the network of the data centers have significant impact on the performance of cloud based services. Hence it is mandatory of the networks to match the guidelines generated in order to employ quality of service for high performance computing demands.

In order to understand the significant of the network performance for providing the cloud based service, here we elaborate the situation of any application deployed over the cloud based data centers for accessing the social media data for analytics. The application will access and storage the data collected from the social media using the appropriate plug-in and will storage the data into the scalable storage solutions. In order to achieve the higher computational capabilities, the application might use the scalable properties of the computing resources. However in order to achieve the quality of service the application must justify the service delay requirements with the network resources for timely response [4].

Moreover the communication based service applications hosted on the data centers implements functionalities which need to be enabled with greater support from the underlying data center network.

Hence this keeps the demand of continuous research on cloud based data center networks alive.

The rest of the work is organized such as in Section II we discuss and apprehend the network as service offering from the cloud computing, in Section III we understand and define the impacts of data center generic architecture, in Section IV we propose the factors influencing the data center network performances, in Section V we realise the pros and cons of the outcomes from the parallel researches, in Section VI we propose the Optimal Distributed Hadoop Network for Cloud Based Data Centers, in Section VII we propose the performance evaluation matrix, in Section VIII we realise the results from implementation of the proposed Optimal Distributed Hadoop Network for Cloud Based Data Center in the light of the performance evaluation matrix and in Section IX we conclude this work.

II. NETWORK AS A SERVICE

The main objective of a Network As A Service or NAAS [3] is to enable the users of the data center to use the data center network efficiently. In this part of the research we understand the functionalities, architecture and finally the implementation requirements. The architecture of the NAAS is the combination of the virtual servers, NAAS boxes, Network Switches and finally the data processing units during the transmissions [Figure – 1].

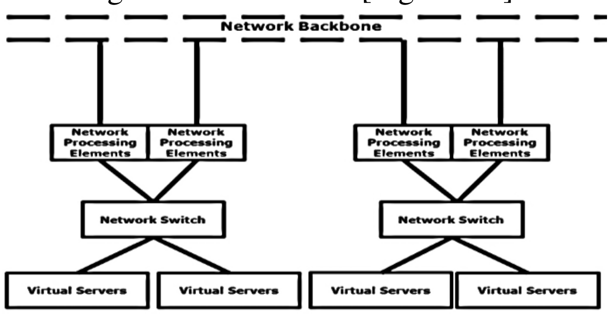

Fig. 1 NAAS Architecture

A. *NAAS Functionalities*

The functionalities of the NAAS is tightly coupled and cannot be considered separately, however for the sole purpose of understanding here we present the list of functions separately below:

- **Visibility of the Network**: There are a set of applications which are build and function on the intersections of the data center networks like the telecommunication networks. The locality of the network also plays a major role in the performance for the highly loaded and occupied data centers. In the other hand, the black box approach is considered where the generic visibility of the network must be high.

- **Request Forwarding Scheme:**Most of the data center servers are configured with only one Network Interface Card and generally configured as open tree protocol for interconnection. Hence the second most important parameter is to forward the incoming connections from one data center server to another. The majority of the servers are configured to forward the packets based on some pre-defined schemes like content aware, location aware or load aware.

- **Packet Processing**: The next core functionality expected from the NAAS is the processing of the packets in the network during transmissions. The processing of the data packets during the network is to perform tasks like Map-Reduce or aggregation of the data packets. The processing during transmission can significantly reduce the overhead on the data center processing and storage.

Hence understanding the functionalities in detail will certainly help to improve the functionalities while not losing the core functional benefits in this work.

B. *NAAS Architecture*

In this part of the work, we understand the core architecture of the Network As A Service. The NAAS is the generic connection source between the data center servers of the virtual servers configured as virtual machines. The major components in NAAS are the virtual servers, Network Switches and most important the Network Processing Elopements.

- **Network Switch**: The Network switches are responsible for managing the interconnection between the servers and other components of the data centers. The network switches are generally connected with a dedicated high bandwidth network wired links to ensure the proper connectivity.

- **Network Processing Elements**: The major advantage of the NAAS is to employ processing of the data packets during the transmission. The elements for processing the data are realized by deploying the special devices with dedicated software programs for processing of data.

Hence understanding the core components of the NAAS architecture will certainly help in realizing the components to be taken under considerations for improvements in the present architectures in the light of outcomes from the parallel researches.

C. *Implementation Requirements*

In this part of the work, we understand the basic requirements in order to implement NAAS for data centers and meet the desired functionalities.

- **Integration Capabilities**:The NAAS brings advantages by reducing the load on the storage and servers residing on the data centers. However the processing components must comply with the existing data center hardware for interfaces.

- **Programming Interfacing Capabilities**: The NAAS processing components are generally the software programs to control the processing of data packets and load balancing operations. Hence further programming implementation is required to match the exact requirements of the data center customers. This is to be understood that the native programming APIs must comply with the fundamental requirements of the programming languages. Henceforth the NAAS programing implementation APIs must be compatible with popular programming languages.
  Creation of a NAAS implementation starts with creating the core of the service [Table – 1].

TABLE I
CREATION OF NAAS SERVICE

```
{
 "naasService": {
  "name": "Domain_1",
    "naasDomain": {
      "instanceId": "8220ccfb-e836-4a6d-
be32-
        205fdae095b0",
      "name": "Multitier_Network"
     }
   }
}
```

The second step is to create the end point resources by employing a POST request [Table – 2].

TABLE II
CREATION OF NAAS END POINT RESOURCE

```
{
   "endpoint": {
     "endpointType": "SWITCH_PORT",
         "name": "10.94.45.171:ge-0/0/21"
   }
}
```

The final step is to invoke the layers into the implementation [Table – 3].

TABLE III
CREATION OF NAAS LAYERED SERVICES

```
{
   "l2Domain": {
     "prefix": "10.20.22.1/24",
     "isolationId": 20,
     "name": "WebTierSubnet",
     "members": {
       "link": [
         {
           "instanceId": "cac60025-c69e-
459c-a9a2-712771465c8e",
           "name": "10.94.45.171:ge-
0/0/21"
         }
       ]
     },
     "isolationType": "VLAN"
   }
}
```

Thus this APIs are compatible with popular programming languages and do not create any additional encumbrance for the developers of NAAS.

- **Isolation Capabilities**: The NAAS data processing components are the software implementations and most of the data center NAAS is configured for a wide range of customer application. Henceforth it is

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

important to realise a complete isolation between the NAAS processing components for different customers.

Hence in order to implement any new model of data center network using NAAS must obey the above mentioned properties to meet the acceptance guidelines from data center standards.

III. GENERIC DATA CENTER NETWORK

In this section of work, we identify the core components of the data center network in order to propose the scope for further enhancements in the outcomes of the parallel researches [Figure – 2].

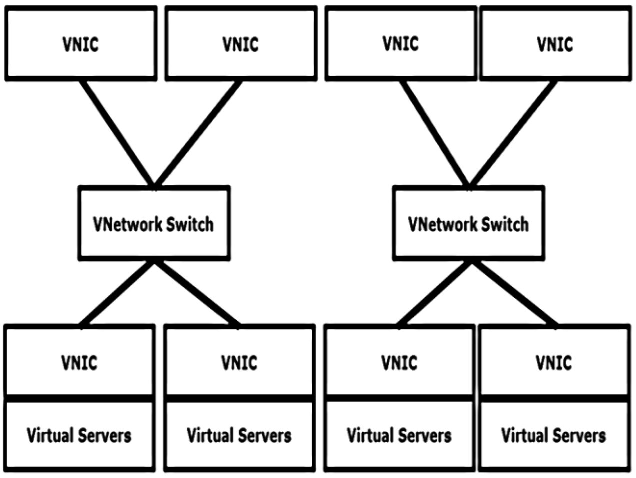

Fig. 2 Data Center Generic Network

Virtualization of Data Center components like compute, storage and network is the major advantage of cloud based data centers. Technological enhancements with software based virtual network components are been widely adopted in all cloud based data centers [3].

Virtualization of network provides higher capabilities for enterprise level networks in order to provide the best in time services. The virtual network components are configured between the local data center components and also deployed between geodetically separated virtual servers from multiple data centers [4].

Through virtual networks, the virtual network interface cards are associated with virtual switches and the virtual switches provide the connectivity between virtual machine network interface cards. The interconnection is purely relying on the network topology used for the data center network [5] [6].

The detail understanding of the data center network components provides the scope to understand that the study of network topologies implemented can provide the scope of further enhancement. The study of the existing network topologies are considered in section V and subsequently the proposed Hadoop topology is addressed in section VI of this work.

IV. PERFORMANCE INFLUENCING FACTORS

The cloud based data centers host a number of applications which are diversified in nature. The applications are different in nature based on the network communication requirements and network based resource accessibilities. Hence it is to be understood that the network performance to be measured based on diversified parameters which truly influences the network performance [5].

Here in this section of the work, we understand the parameters influencing the performance of the network and the weightage of the mentioned parameters to calculate the overall performance [Table – 4].

TABLE IV
PARAMETERS INFLUENCING THE NETWORK PERFORMANCE

Type	Details of the Parameters	
	Name	**Definition**
Buffer Parameters	Q_BIT (t)	The number of bytes stored in the send buffer at a given time t
	Q_MAX_BIT(t)	The maximum

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

		limit of the send buffer at any given time instance t
Application Parameters	$TRANS_N$ (t)	The total number of successful transmissions in a given time window t
	TIMEOUT (t)	The total number of time outs in a given time frame t
	R_TIMEOUT	The time windows for receiving packets for an application deployed on the network
Responsive Parameters	RTT (t)	The round-trip time for the network topology with standard application load at any given time instance
	TOT_RTT	The total round trip time for the network topology with standard application load

The number of bytes stored in the send buffer at a given time t for a specific network adaptor N is calculated as following:

$$Q\_BIT_N(T)=\sum_{i=1}^{n}\Delta Q\_BIT_i(t) \quad \text{...Eq 1}$$

Defined as the cumulative bytes per sub node.

The maximum limit of the send buffer at any given time instance t for a specific network adaptor N is calculated as following:

$$Q\_MAX\_BIT_N(T)=\lceil Q\_BIT_N(T)\rceil \quad \text{... Eq 2}$$

The total number of successful transmissions in a given time window t for a specific network adaptor N is calculated as following:

$$TRANS_N(T)=\sum_{t=1}^{n}TRANS_{n_0}(t)+TRANS_{n_1}(t)+...+TRANS_{n_n}(t) \quad \text{... Eq 3}$$

The round-trip time for the network topology with standard application load at any given time instance for a specific network adaptor N is calculated as following:

$$RTT_N(T)=\|T_1-T_2\| \quad \text{... Eq 4}$$

Where, T1 and T2 denotes the time instance to dispatch the packet and time instance to receive the acknowledgement respectively.

The total round trip time for the network topology with standard application load for a specific network adaptor N is calculated as following:

$$TOT\_RTT_N(T)=\sum_{t=1}^{n}RTT_N(t) \quad \text{... Eq 5}$$

V. Outcomes from the Parallel Researches

The core of the performance for any cloud based data centers are depending on the data center networks as the services are to be accessed through the network connecting the data center with the customers. Moreover the internal components are also connected via the data center internal networks. In this work, during the study of data center architecture and network, we have realized that the performance will solely depend on the network topologies as the hardware components used for managing the network also comply with the generic characteristics.

A numerous number of efforts are been made to improve the topology for data center networks and the parallel research outcomes demonstrates the core logic behind the topologies. Here in this work, we consider few popular topologies.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

A. *Basic Tree*

The two or three tier network topology commonly known as Basic Tree topology [Figure – 3] is consisting of collection of server with the collection of network switches as Edge layer, collection of domain name servers with load balancing components as aggregation layer and finally the core layer [7].

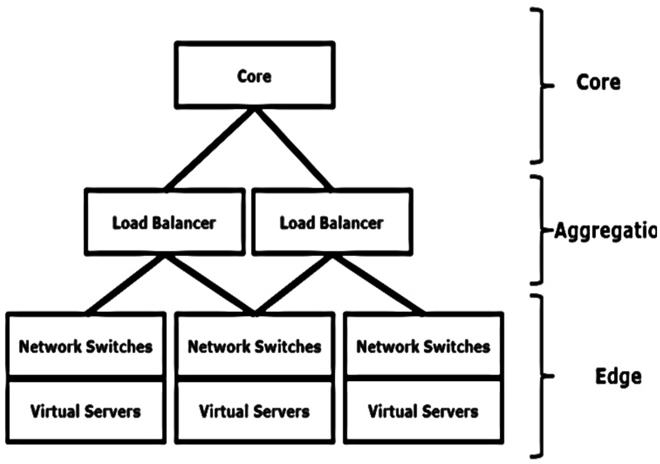

Fig. 3 Basic Tree Topology

However, it is been observed that the Basic Tree topology can increase the congestions and reduce the effective bandwidth of the network.

B. *FAT- Tree*

In order to overcome the bandwidth limitations of the Basic Tree topology, another three layer topology called Fat – Tree topology [Figure – 4] is introduced. The Fat Tree topology uses multiple alternate network channels to distribute the loads [8].

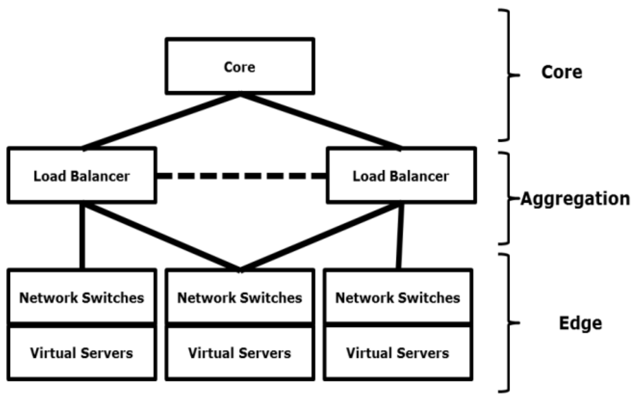

Fig. 4FAT Tree Topology

However, it is been observed that, realizing the connection for any physical data center might be very complex task.

C. *Valiant Load Balancing*

In order to overcome the complexities of realizing the FAT- Tree topology the other researches have demonstrated the use of Valiant Load Balancing Topology [Figure – 5]. In order to overcome the wiring complexities, the VLB introduces the agile networking between the load balancing components into the network to distribute the loads over the complete network [9].

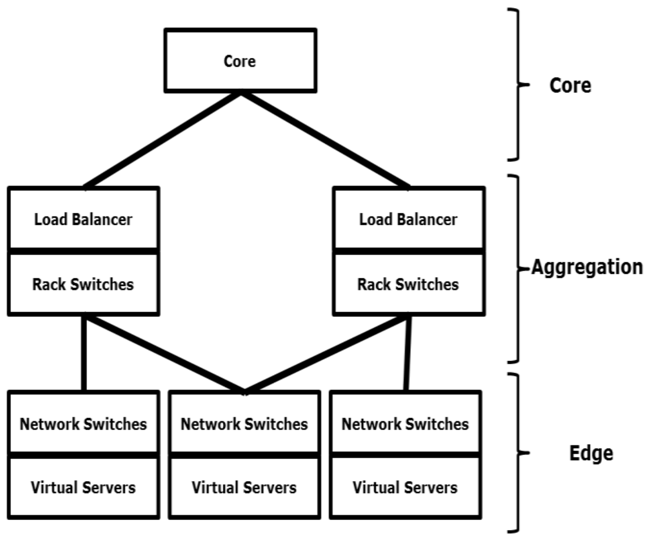

Fig. 5VLB Topology

However, it is been observed that the during the pick hours of high load on the data centers, the VLB topologies do not provide the outputs to the expectations.

After detail understanding the short comings of the existing systems, in this work we propose the Novel Optimal Distributed Hadoop Network for Cloud Based Data Center in section VI.

VI. PROPOSED NOVEL HADOOP FRAMEWORK

With the detail understanding of the outcomes from the parallel research efforts, it is easy to understand the following shortcomings:

- The data center networks are limited in scope to effectively use the bandwidth.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

- Also in order to increase the effective use of the bandwidth in the networks, the architectural complexities are been increased.
- Furthermore, the increasing loads on the load balancers have accelerated the demand for Network As A Service implementations in the data center network.

Hence in this work, we propose a Novel Optimal Distributed Hadoop Network for Cloud Based Data Center [Figure – 6] in order to overcome the above mentioned problems in the following ways.

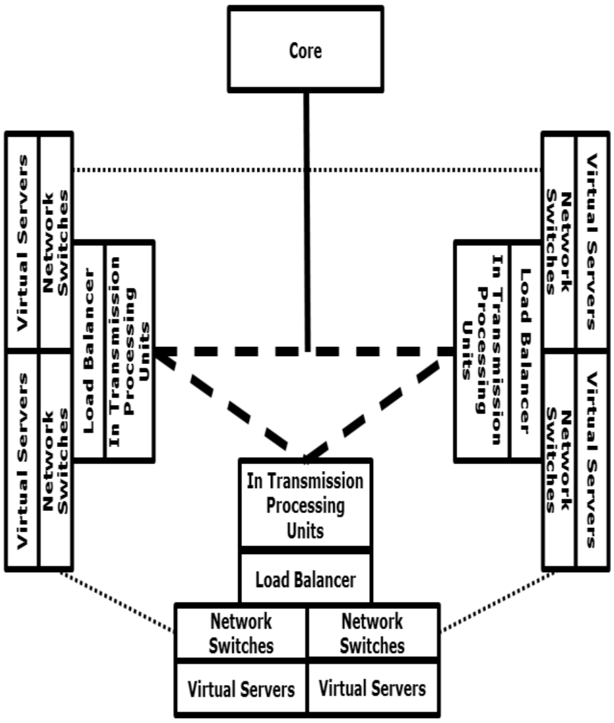

Fig. 6Proposed Hadoop Topology

- The dual agile connection modes between the load balancers, In Transmission Processing Units and the network switches have made the highest level of effectiveness of the bandwidth utilizations.
- The simple architecture reduces the complexity of realization of this network topology and results in high response times.
- The implementation of the In Transmission Processing Units along with the load balancers reduces the load on the load balancing components.

Hence with the implementation proposal, in section VIII we simulate all the topologies and compare the performances for conclusion.

VII. PERFORMANCE EVALUATION MATRIX

The proposed weightage for the parameters defined based on the core definition of the parameters available for network performance measurements [10]. The total weightage of the governing formula is assumed as K, Hence the individual weightage for the parameters are calculated in terms of K in this work [Table – 5].

TABLE V
PARAMETERS INFLUENCING THE NETWORK PERFORMANCE WITH WEIGHTAGE

Type	Details of the Parameters	
	Name	**Weightage**
Buffer Parameters	Q_BIT (t)	K/10
	Q_MAX_BIT(t)	K/5
Application Parameters	$TRANS_N$ (t)	K/30
	TIMEOUT (t)	K/10
	R_TIMEOUT	K/5
Responsive Parameters	RTT (t)	K/30
	TOT_RTT	K/10

Hence forth the general formulation as an outcome of this work will help in comparing the performance of the network topologies proposed by the outcomes of the parallel research and the Hadoop topology proposed in this work.

Hence the network performance index of any network data center is represented as:

$$NPI = Q\_BIT_N(T)*\frac{K}{10} + Q\_MAX\_BIT_N(T)*\frac{K}{5} + TRANS_N(T)*\frac{K}{30} + TIMEOUT(t)*\frac{K}{10} + R\_TIMEOUT*\frac{K}{5} + RTT_N(T)*\frac{K}{30} + TOT\_RTT*\frac{K}{10}$$

… Eq 6

Finally, we achieved the formulation for the Network Performance Index for calculating the performance of the data center networks.

VIII. Results

The results of this work is been simulated using a java based applications with 10 nodes in the edge level and the simulation results are been listed in this section.

The simulation results for Basic Tree topology have demonstrated the short comings mentioned in the previous section of this work [Table – 6].

TABLE VI
SIMULATION RESULTS OF BASIC TREE TOPOLOGY

Topology : Basic Tree								
(10 Nodes, 17 Edges)								
Nodes (10)								
0	711	141	5	5	0 AS_NONE			
1	579	903	5	5	1 AS_NONE			
2	767	449	5	5	2 AS_NONE			
3	440	953	4	4	3 AS_NONE			
4	20	565	3	3	4 AS_NONE			
5	521	270	2	2	5 AS_NONE			
6	87	558	3	3	6 AS_NONE			
7	351	23	3	3	7 AS_NONE			
8	979	42	2	2	8 AS_NONE			
9	467	983	2	2	9 AS_NONE			
Edges (17)								
0	0	1	773.3486	-1.0	1025.6561	0	1	E_AS_NONE U
1	0	2	313.04953	-1.0	663.5874	0	2	E_AS_NONE U
2	1	2	491.3858	-1.0	243.49432	1	2	E_AS_NONE U
3	3	0	856.0286	-1.0	27.468811	3	0	E_AS_NONE U
4	3	1	147.71933	-1.0	926.00024	3	1	E_AS_NONE U
5	4	3	571.79016	-1.0	470.75586	4	3	E_AS_NONE U
6	4	0	810.71387	-1.0	1018.53217	4	0	E_AS_NONE U
7	5	3	687.7863	-1.0	871.8953	5	3	E_AS_NONE U
8	5	2	304.2318	-1.0	485.35498	5	2	E_AS_NONE U
9	6	4	67.36468	-1.0	94.22162	6	4	E_AS_NONE U
10	6	1	600.9068	-1.0	387.891	6	1	E_AS_NONE U
11	7	0	378.8456	-1.0	650.3594	7	0	E_AS_NONE U
12	7	2	595.4259	-1.0	150.36511	7	2	E_AS_NONE U
13	8	6	1030.495	-1.0	155.38013	8	6	E_AS_NONE U
14	8	2	458.90414	-1.0	719.7897	8	2	E_AS_NONE U
15	9	7	966.9829	-1.0	957.3366	9	7	E_AS_NONE U
16	9	1	137.6372	-1.0	620.627	9	1	E_AS_NONE U

The simulation results for FAT Tree topology have demonstrated the short comings mentioned in the previous section of this work [Table – 7].

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

TABLE VII
SIMULATION RESULTS OF FAT TREE TOPOLOGY

Topology : FAT Tree
(10 Nodes, 17 Edges)
Nodes (10)
0 442 54 7 7 -1 RT_NONE
1 526 129 7 7 -1RT_NONE
2 440 118 5 5 -1RT_NONE
3 132 362 3 3 -1 RT_NONE
4 725 497 2 2 -1 RT_NONE
5 430 444 2 2 -1 RT_NONE
6 654 619 2 2 -1 RT_NONE
7 547 485 2 2 -1 RT_NONE
8 764 38 2 2 -1 RT_NONE
9 612 380 2 2 -1 RT_NONE
Edges (17)
0 0 1 112.60995 -1.0 884.68835 -1 -1 E_AS_NONE U
1 0 2 64.03124 -1.0 768.7857 -1 -1 E_AS_NONE U
2 1 2 86.70064 -1.0 293.87988 -1 -1 E_AS_NONE U
3 3 1 457.739 0.0 61.036804 -1 -1 E_RT_BACKBONE U
4 3 2 392.93765 0.0 495.1988 -1 -1 E_RT_BACKBONE U
5 4 1 418.3599 0.0 279.36945 -1 -1 E_RT_BACKBONE U
6 4 2 474.20038 0.0 568.5163 -1 -1 E_RT_BACKBONE U
7 5 0 390.18457 0.0 620.9853 -1 -1 E_RT_BACKBONE U
8 5 3 309.07605 0.0 132.79205 -1 -1 E_RT_BACKBONE U
9 6 0 603.4642 0.0 170.8064 -1 -1 E_RT_BACKBONE U
10 6 1 506.4425 0.0 1012.12756 -1 -1 E_RT_BACKBONE U
11 7 0 443.60568 0.0 511.10663 -1 -1 E_RT_BACKBONE U
12 7 1 356.61884 0.0 577.20013 -1 -1 E_RT_BACKBONE U
13 8 1 254.80385 0.0 260.2948 -1 -1 E_RT_BACKBONE U
14 8 0 322.39728 0.0 569.78564 -1 -1 E_RT_BACKBONE U
15 9 2 313.41345 0.0 93.033936 -1 -1 E_RT_BACKBONE U
16 9 0 367.6629 0.0 445.3393 -1 -1 E_RT_BACKBONE U

The simulation results for VLB topology have demonstrated the short comings mentioned in the previous section of this work [Table – 8].

TABLE VIII
SIMULATION RESULTS OF VLB TOPOLOGY

Topology : VLB
(10 Nodes, 17 Edges)
Nodes (10)
10 562 767 5 5 0 T_BORDER
11 925 609 4 4 0RT_NONE
12 708 246 2 2 0RT_NONE
13 288 953 2 2 0RT_NONE
14 562 88 2 2 0RT_NONE
15 401 236 3 3 0T_BORDER
16 163 608 10 10 0T_BORDER
17 858 449 5 5 0RT_NONE

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

18	112	745	4	4	0 R_BORDER
19	662	241	2	2	0RT_NONE
Edges (17)					
17 16 17 712.9558 -1.0 -1.3368013 0 0 E_AS_NONE U					
18 16 18 146.18481 -1.0 -1.9500041 0 0 E_AS_NONE U					
19 17 18 802.57837 -1.0 -1.9850034 0 0 E_AS_NONE U					
20 19 16 619.4272 0.0 -1.9805245 0 0 E_RT_BACKBONE U					
21 19 18 746.0 0.0 -1.2450812 0 0 E_RT_BACKBONE U					
22 10 16 429.51367 0.0 -1.6992847 0 0 E_RT_BACKBONE U					
23 10 17 434.44217 0.0 -1.5928284 0 0 E_RT_BACKBONE U					
24 11 17 173.4618 0.0 -1.3496339 0 0 E_RT_BACKBONE U					
25 11 16 762.0007 0.0 -1.1095054 0 0 E_RT_BACKBONE U					
26 12 16 654.26984 0.0 -1.8308802 0 0 E_RT_BACKBONE U					
27 12 17 252.40642 0.0 -1.3751826 0 0 E_RT_BACKBONE U					
28 13 16 366.94687 0.0 -1.8686869 0 0 E_RT_BACKBONE U					
29 13 10 331.16763 0.0 -1.1958687 0 0 E_RT_BACKBONE U					
30 14 10 679.0 0.0 -1.3516251 0 0 E_RT_BACKBONE U					
31 14 11 634.98816 0.0 -1.3850412 0 0 E_RT_BACKBONE U					
32 15 16 441.61975 0.0 -1.4059677 0 0 E_RT_BACKBONE U					
33 15 11 643.19904 0.0 -1.0018892 0 0 E_RT_BACKBONE U					

The simulation results for Novel Optimal Distributed Hadoop Network for Cloud Based Data Center topology have demonstrated the advantages mentioned in the previous section of this work [Table – 9].

TABLE IX
SIMULATION RESULTS OF NOVEL OPTIMAL DISTRIBUTED HADOOP NETWORK TOPOLOGY

Topology :Novel Optimal Distributed Hadoop Network					
(10 Nodes, 17 Edges)					
Nodes (10)					
0	539	477	7	7	-1RT_NONE
1	770	490	2	2	-1RT_NONE
2	66	729	7	7	-1RT_NONE
3	237	549	6	6	-1RT_NONE
4	549	47	2	2	-1RT_NONE
5	652	647	2	2	-1RT_NONE
6	906	847	2	2	-1RT_NONE
7	50	452	2	2	-1RT_NONE
8	307	832	2	2	-1RT_NONE
9	360	735	2	2	-1RT_NONE
Edges (17)					
0 0 1 231.36551 -1.0 158.6145 -1 -1 E_AS_NONE U					
1 0 2 535.9412 -1.0 247.8468 -1 -1 E_AS_NONE U					
2 1 2 743.4628 -1.0 748.70337 -1 -1 E_AS_NONE					

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

U
3 3 2 248.27606 0.0
613.4642 -1 -1
E_RT_BACKBONE U
4 3 0 310.46417 0.0
725.3574 -1 -1
E_RT_BACKBONE U
5 4 2 835.71106 0.0
474.49402 -1 -1
E_RT_BACKBONE U
6 4 0 430.11627 0.0
581.73645 -1 -1
E_RT_BACKBONE U
7 5 0 204.12987 0.0
495.11273 -1 -1
E_RT_BACKBONE U
8 5 3 426.41412 0.0
315.96613 -1 -1
E_RT_BACKBONE U
9 6 3 732.36945 0.0
970.14246 -1 -1
E_RT_BACKBONE U
10 6 0 521.142 0.0
818.1544 -1 -1
E_RT_BACKBONE U
11 7 2 277.4617 0.0
327.232 -1 -1
E_RT_BACKBONE U
12 7 3 210.66086 0.0
131.88849 -1 -1
E_RT_BACKBONE U
13 8 2 262.08777 0.0
185.98303 -1 -1
E_RT_BACKBONE U
14 8 0 424.0861 0.0
307.3092 -1 -1
E_RT_BACKBONE U
15 9 2 294.06122 0.0
349.24957 -1 -1
E_RT_BACKBONE U
16 9 3 222.99103 0.0
520.798 -1 -1
E_RT_BACKBONE U

Alsowe compare the Network Performance Index for the topologies and realise the improvements [Table – 10] based on the simulation results obtained from [Table – 6 to 9] and the Network Performance Index [Eq. 6].

TABLE X
NETWORK PERFORMANCE INDEX

Topology Type	NPI
Basic Tree	100 (Value Normalized)
FAT Tree	112 (Value Normalized)
VLP	117 (Value Normalized)
Optimal Distributed Hadoop Network for Cloud Based Data Center topology (Proposed)	160 (Value Normalized)

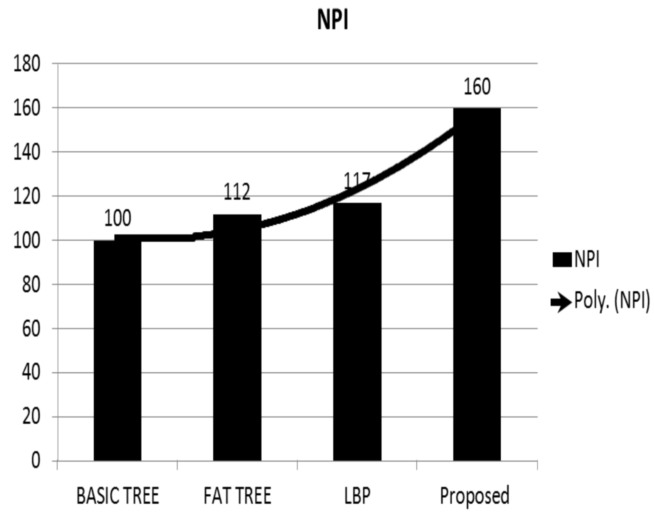

Fig. 7 NPI Graph

Hence forth it is a significant improvement in the network performance for cloud based data center network topology.

IX. CONCLUSION

In this work, we have understood the architectural components of Network as a service with the core basic functionalities, which helped to improve the functionalities while not losing the core functional benefits in this work. Also the architectural understanding helped in realizing the components to be taken under considerations for improvements in the present architectures in the light of outcomes from the parallel researches. The work also focuses on the implementation requirements for implementing the proposed model of data center networking NAAS to meet the acceptance guidelines from data center standards. This work carried out a detailed study of the Generic Data Center Network to understand the scope of improvement in topologies as the hardware components already comply with the standards. During the study of the cloud based data center networks performance influencing factors which helped in defining the performance evaluation matrix and based on that the final Network performance index for evaluating network topologies for cloud based networks. The final outcome of this work is the novel optimal distributed Hadoop network topology. The performance of the topologies are been compared and we have observed an improvement of the performance [Figure – 7]. Hence the final outcome of the work demonstrates overall 73% improvement of network performance index value over the existing data center network protocols.

REFERENCES

[1] C.-Y. Hong, M. Caesar and P. B. Godfrey , "Finishing flows quickly with pre-emptive scheduling”, ACM SIGCOMM Comput. Commun. Rev. , vol. 42 , no. 4 , pp.127 -138 , 2012

[2] D. Li , J. Yu and J. Yu , "Exploring efficient and scalable multicast routing in future data center networks" , Proc. IEEE Conf. Comput. Commun. (INFOCOM¿¿¿11) , pp.1368 -1376 , 2011

[3] D. Li , C. Guo and H. Wu , "Scalable and cost-effective interconnection of data center servers using dual server ports" , IEEE/ACM Trans. Netw. , vol. 19 , no. 1 , pp.102 -114 , 2011

[4] C. Guo , G. Lu and Y. Xiong , Datacast: A scalable and efficient group data delivery service for data centers , 2011

[5] Y. Vigfusson , H. Abu-Libdeh and M. Balakrishnan , "Dr. multicast: Rx for data center communication scalability" , Proc. ACM Eur. Conf. Comput. Syst. (EuroSys¿¿¿10) , pp.349 -362 , 2010

[6] M. Al-Fares, A. Loukissas, and A. Vahdat, ``A scalable, commodity data center network architecture," in Proc. ACM SIGCOMM, 2008, pp. 6374.

[7] A. Greenberg et al., ``VL2: A scalable and exible data center network," ACM SIGCOMM Comput. Commun. Rev., vol. 39, no. 4, pp. 5162, Oct. 2009.

[8] D. Lin, Y. Liu, M. Hamdi, and J. Muppala, ``FlatNet: Towards aatter data center network," in Proc. IEEE GLOBECOM, Dec. 2012, pp. 24992504.

[9] B. Heller et al., ``ElasticTree: Saving energy in data center networks," in Proc. NSDI, vol. 3. 2010, pp. 1921.

[10] R. N. Mysore et al., ``PortLand: A scalable fault-tolerant layer 2 data center network fabric," ACM SIGCOMM Comput. Commun. Rev., vol. 39, no. 4, pp. 3950, 2009.