PERFORMANCE EVALUATION OF A REENGINEERED DATA CENTER NETWORK USING A LINK STATE PROTOCOL IMPLEMENTATION

Christiana. C. Okezie
Electronic and Computer Engineering Department, Nnamdi Azikiwe University, NIGERIA.
christianaobioma@yahoo.com

Udeze Chidiebele. C
R & D Department, Electronics Development Institute, Awka, NIGERIA.
udezechidi@yahoo.com

Okafor Kennedy .C
R & D Department, Electronics Development Institute, Awka, NIGERIA.
arissyncline@yahoo.com

ABSTRACT

The operational challenges of the contemporary data center network (DCN) has given rise to the need to change its IT service delivery to a modern data center network model that is more efficient, service-oriented, and responsive to business needs. A new dynamic infrastructure for such a modern data center network was modeled, which allows organizations to be better positioned to adopt integrated new technologies, such as consolidation, virtualization and cloud computing. This will help them to deliver dynamic access to IT services and resources thereby enabling IT departments to spend less time fixing IT problems and more time solving real business challenges. Open Shortest Path First (OSPF) as a link state protocol in addressing network convergence and cost in the proposed data center network was evaluated in this work. OPNET simulator was used for modeling and parameter characterization.

Keywords: Traditional DCN Infrastructure, Virtualization, Cloud Computing, OSPF, OPNET

INTRODUCTION

A Data Center is the consolidation point for provisioning multiple services that drive an enterprise business process (Emersion, 2004). It is also known as the server farm or the computer room. The data center is where the majority of enterprise servers and storage systems are located, operated and managed like the Enterprise Resource Planning solutions, (ERPs), application servers, security systems (IDS). It generally includes redundant or backup power supplies, redundant data communications connections, environmental controls (e.g. Air conditioning, fire suppression) and security devices. Ethernet switching technology is the foundation, upon which many of these services are built. (Emersion, 2007)

The four primary components of a traditional data center include [Oracle 2009]:

White Space

This typically refers to the usable raised floor environment measured in square feet (anywhere from a few hundred to a hundred thousand square feet).

Support Infrastructure

This refers to the additional space and equipment required to support data center operations — including power transformers, your uninterruptible power source (UPS), generators, computer room air conditioners (CRACs), remote transmission units (RTUs), chillers, air distribution systems, etc. In a high-density, tier 3 class data center (i.e. a concurrently maintainable facility), this support infrastructure can consume 4-6 times more space than the white space and must be accounted for in data center planning.

IT Equipment

This includes the racks, cabling, servers, storage, management systems and network gear required to deliver computing services to the organization.

Operations

The operations staff assures that the systems (both IT and infrastructure) are properly operated, maintained, upgraded and repaired when necessary. In most companies, there is a division of responsibility between the technical operations group in IT and the staff responsible for the facilities support systems.

The functional layers designed to serve the needs of the Data Center are typically built around three layers: The Core Layer, the Aggregation Layer and the Access Layer.

The Core Layer

This is central to the Data Center network and provides interconnection between the Aggregation Layers and the Access layer. Typically, the Core Layer utilizes high performance low latency switches providing high densities of 10GE. The use of 10GE to link up to the Aggregation Layer Switches is highly recommended. Switches at this layer operate exclusively as Layer 3 devices. The target Cisco device recommended to serve the needs of the Core Layer is the Catalyst 6500 switch.

The Aggregation Layer

It acts as a Services Layer for the Data Center. Services such as Load Balancing, SSL Optimization, and Firewalling with NAC servers, etc are typically found at this layer. Multiple Access Layer switches will also use the Aggregation Layer as an interconnection point. The use of 10GE links to uplink into the Core Layer is a common practice. More of an emerging trend is the use of 10GE links to downlink into the Access Layer providing higher bandwidth and future proofing the network.

The Access Layer

It provides connectivity for the many servers that deliver application and web services to the business as well as the interconnections for a server cluster design. Access Layer switches can be configured for both Layer 2 and Layer 3 deployments. An access Layer switches may be required to support both single and dual homed servers. Architecture of a traditional data center network is shown in Figure1.

CONTEXT AND REVIEW OF LITERATURE

Data centres run the applications that deliver business processes and services. These applications provide critical information and rich, differentiated content for users. Users now demand an agile, responsive infrastructure that provides exactly the access that they need [5].

This can be 24x7x365 for services that must be “always on and accessible from anywhere,” or a series of scheduled updates set to meet user needs for time-based information (hourly, daily, weekly, monthly, or quarterly).

In [6], a technical white paper, HP delivers the foundation for the data center of the future, today, by providing a unified, virtualization-optimized infrastructure. HP networking solutions deliver the following services:

a. Flatter and more efficient data center networks with fewer layers, less equipment and cabling, and greater port densities.

b. High performance, low latency intra-data center connectivity for VM migration and bandwidth-intensive server-to-server communications.

c. Virtualization-aware security to protect intra-server communications flows and virtual resources.

d. Unified administration to remove costly, time-consuming and error-prone change management processes and improve business agility.

e. Multi-site, multi-vendor management to connect and control thousands of physical and virtual resources from a single pane of glass.

ISSN-L: 2223-9553, ISSN: 2223-9944
Vol. 3, No. 1, July 2012

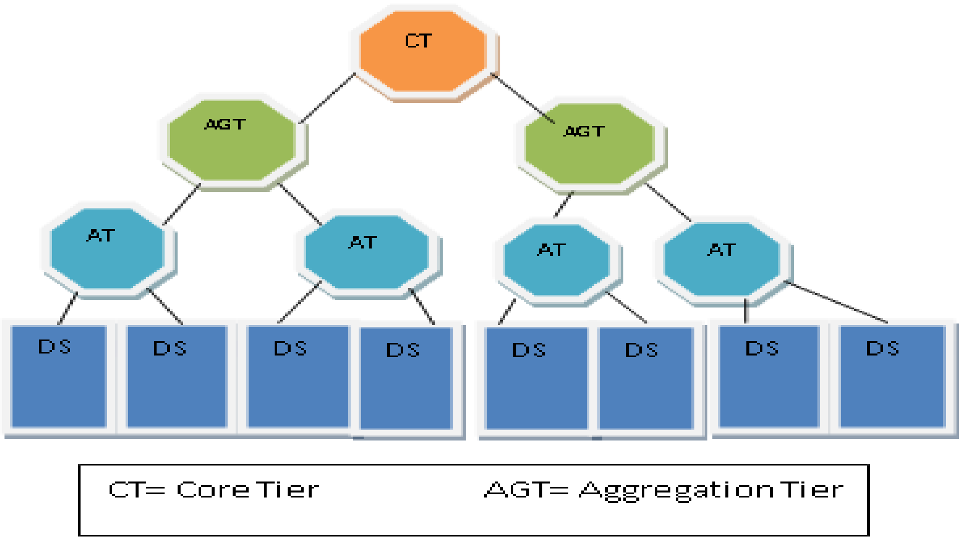

Figure1: Traditional Data Center Networks

Deloitte's Center for the Edge carried out extensive research and industry insights to explore the topics of cloud computing and next-generation web services. The authors (Thomas and Brown, 2009) presented their work in three sets of papers from different perspectives: business drivers, architectural models, and transformation strategies. They outlined a vision for the role of the cloud, describing how it can evolve and extend today's service-oriented architectures into business platforms for the next-generation economy but fails to explain scalable network convergence with link state protocol like OSPF.

IBM Corporations in [8] presented a technical overview of the new enterprise data center, including a description of its key characteristics, a description of the functions and capabilities of the underlying architecture and a description of an evolutionary approach to implementation through stages of adoption. The new enterprise data center strategy allows companies to focus on the services provided by the infrastructure, rather than on the underlying technology that enables these services. The paper in [8] explains that through a service-oriented approach to IT delivery, the new enterprise data center provides enterprises with greater flexibility and agility, helping them to move toward dynamic infrastructures. As such, stipulates that cloud computing is the natural outcome of next generation data centers.

Consequently, the work by IBM Corporation then provides a good foundation for re-engineering of the legacy data center networks to a modern data center model that is more efficient, service-oriented, and responsive to business needs using OSPF in our context.

THE TRADITIONAL DCN AND ITS LIMITATIONS

Most traditional data center networks are based on three-tier architectures designed to support conventional "north-south" client-server traffic flows in and out of the data center [9]. A typical three-tier data center network is comprised of an access tier, an aggregation tier and a core tier. The access tier is made up of cost-effective Ethernet switches connecting rack servers and IP-based storage devices (typically 100Mbps or 1GbE connections). The access switches are connected via Ethernet to a set of aggregation switches (typically 10GbE connections) which in turn are connected to a layer of core switches or routers that forward traffic to an intranet, the Internet and between aggregation switches. Layer 2 VLANs are typically implemented across the access and aggregation tiers and Layer 3 routing is implemented in the core. Bandwidth is typically over-provisioned in the access tier, and to a lesser extent in the aggregation tier. The server infrastructure and the networking infrastructure are typically administered independently, by separate teams using distinct toolsets. Each server is typically dedicated to a specific function (i.e. web server, application server, database server) and can be reasonably well protected using conventional security solutions such as intrusion prevention system.

From the research we carried out, the following are the limitations of the traditional data center networks which are the drivers for the re-engineering:

Network Convergence and Downtime

IT operations are a crucial aspect of most organizational operations. One of the main concerns is business continuity; companies rely on their information systems to run their operations. If a system becomes unavailable due to downtime, company operations may be impaired or stopped completely. It is necessary to provide a reliable infrastructure for IT operations, in order to minimize any chance of disruption. Traditional data center networks have problems of network convergence and downtime thereby disrupting business operations in the enterprise market segments like ISPs, financial and educational Institutions, oil and gas sector. The issue of network convergence and downtime was handled in our proposed reengineered DCN by the link state routing protocol chosen for the routers and key technologies such as virtualization, consolidation and cloud computing.

High Infrastructure Economy

The cost of deployment as well as maintaining the IT infrastructure in data center networks is very immense and hence calls for a better approach to cost reduction as well as service availability. Internet and business applications are increasingly being moved to large data centers that hold massive server and storage clusters. Current data centers can contain tens of thousands of servers, and plans are already being made for data centers holding over a million servers.

The massive amounts of computational power required to drive these systems results in many challenging and interesting distributed systems, as well as IT cost resource management problems. In this work, server virtualization is proposed as a remedy to infrastructure demands in the traditional DCNs.

Latency and Delay

This is the amount of time that it takes for a packet to be transmitted from one point in a network to another point in the network. There are a number of factors that contribute to the amount of delay experienced by a packet as it traverses the network. Delays include; forwarding delay, queuing delay, propagation delay, and serialization delay. The end-to-end delay can be calculated as the sum of the individual forwarding, queuing, propagation, and serialization delays occurring at each node and link in the network. Moreover in relation to the traditional DCN, the overall architecture generates over 30 percent network latency in switching and traffic delay thereby affecting responsiveness to business demands and services.

Throughput

Throughput is a term used to describe the capacity of a system to transfer data. There are different ways to define and measure throughput, this includes: the packet rate across the network; the packet rate of a specific application flow; the packet rate of host-to-host aggregated flows; the packet rate of network-to-network aggregated flows. Since the demand for data exchange in DCNs is extremely large compared with other networks, the first design goals is to maximize the throughput. Actually, here the throughput should be good put, i.e., retransmissions are harmful for DCNs. If there are no retransmissions, generally, maximization of the throughput is equivalent to maximizing the link utilization. The amount of bandwidth allocated to different types of packets affect throughput. The bulky architecture of the traditional data center networks which results to about 30 percent network latency in switching, the overall throughput is negatively affected. Server virtualization and consolidation made for high throughput in our proposed reengineered DCN

Congestion

Data centers deploy a variety of middle boxes (e.g., Firewalls, load balancers and SSL Off loaders, Web caches and intrusion prevention boxes) to protect, manage and improve the performance of applications and services they run. Since existing (traditional) data center networks provide limited support for middle boxes, administrators typically overload path selection mechanisms to coerce traffic through the desired sequences of middle boxes placed on the network path. These ad-hoc practices result in a data center network that is hard to configure and maintain, and which creates unnecessary bottlenecks and congestion. Network congestion is affected by the following factors : the amount of traffic generated and placed on the network by its users, the network architecture in terms of its link capacities, their connectivity, and the characteristics of each link layers viz: Core, aggregate and Access layers vis-à-vis data centers, operating parameters of sources, routers and destinations - such as available buffers, available processing power and system design, flow control mechanisms used in the Transport Layers of sources of data flow and their respective destinations, congestion avoidance and congestion recovery schemes in DCN devices (switches, routers) and client sources, packet admission mechanisms in the Network and Link layers of sources and destinations, packet discarding and packet loss mechanisms in routers and destinations. Virtualization which is one of the key technologies in our proposed reengineered DCN, handles most of these factors thereby reducing the issue of DCN congestion to the barest minimum.

Security

Information security is a very serious concern in all DCNs, and for this reason they have to offer a secure environment which minimizes the chances of a security breach. A data center must keep high standards for assuring the integrity and functionality of its hosted computer environment. This is accomplished through redundancy of both fibre optic cables and power, which includes emergency backup power generation. In recent times DCN became more vulnerable to security attacks due to the proliferation of the web-based technologies. And any security attack on the data center can destroy the whole organization’s network and data.

From detailed study carried out in the literature review, a conceptual model of the proposed re-engineered DCN architecture is presented in this context and is shown in Figure 2a while Figure 2b shows the process model architecture. Figure 3 shows the implementation model with OPNET guru. The system comprises of the internet cloud domain, load balancer for load segmentation, a multi-protocol speed redundancy layer, and virtualisation/consolidation domain. From the architecture, the speed redundancy layer was modelled to completely replace the core and distribution layer of the traditional DCN.

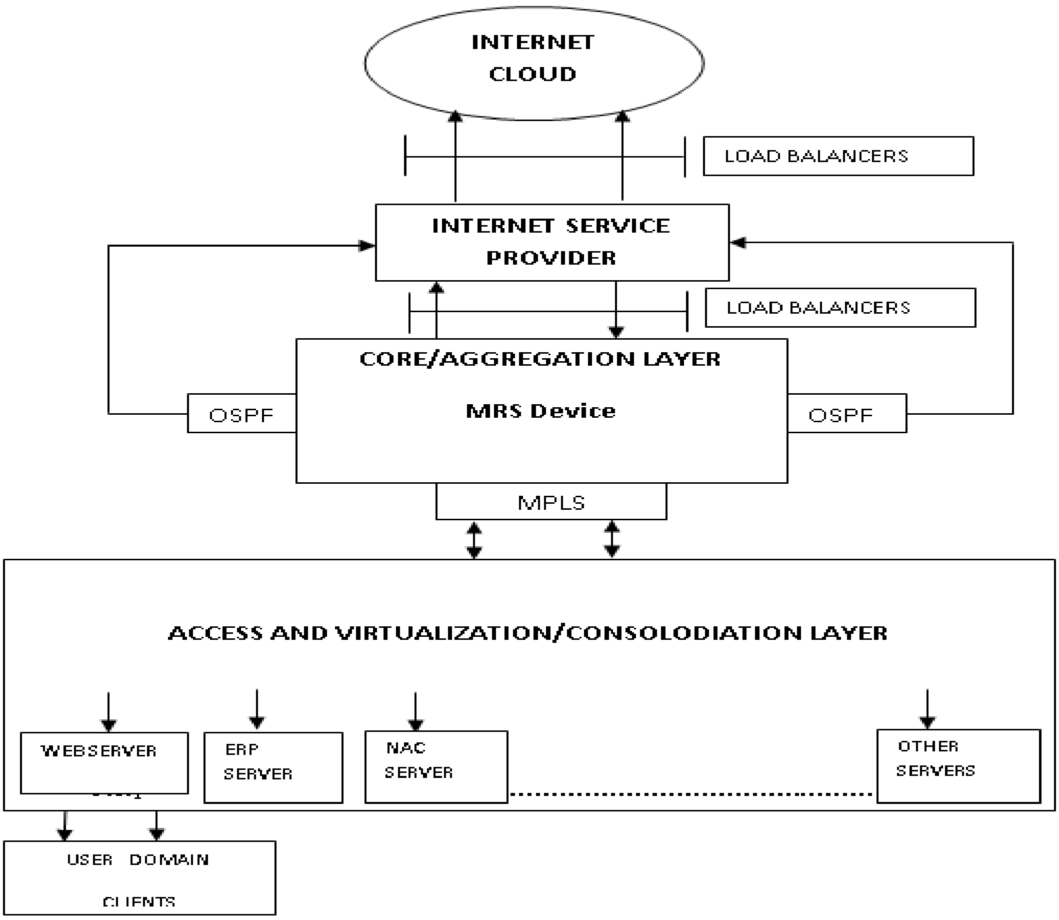

Figure 2a: A Conceptual Model of the Re-engineered DCN Architecture

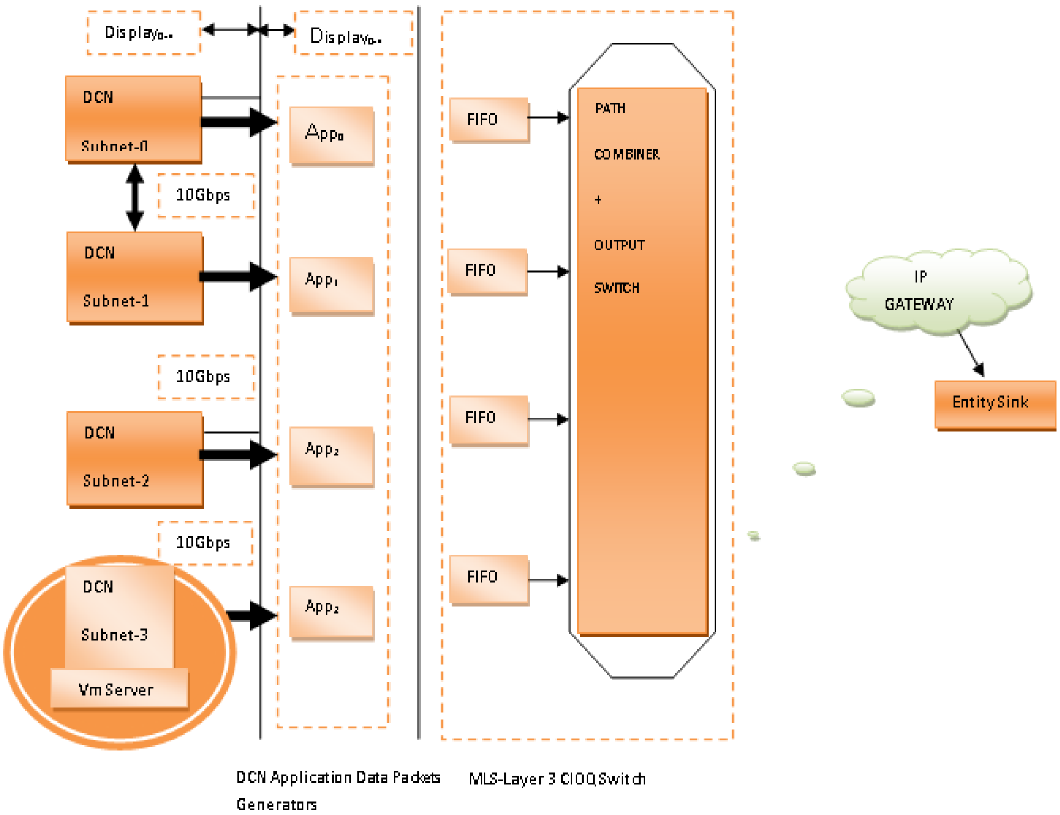

Figure 2b: The Process Model Architecture

Multi-protocol Routing Switch (MRS) was adapted to carry out routing and switching function running two redistribution protocol OSPF (open shortest path first) and MPLS (multi-protocol label switch). This can be extended to accommodate chains of routine and distribution models within the DCN. A link interface FCOE (Fibre Channel over Ethernet) or OC (Optic Carrier) establishes connection to the virtualisation/consolidation layer. Primarily, this layer runs on an enterprises server architecture having two layers of caching services, one for a virtual device service or virtual machines(VM) outer loop and the other for the servers attached to the VMs and its related applications. This layer is designed to have resilience, scalability, robustness owing to its hardware configurations. Also, beside the caching services, bandwidth optimisation is realised in this layer. Finally, the user domain is the termination point of the proposed model. Client machines (workstations, PDAs, PCs, Laptops) can gain access to the DCN through the User Domain. Owing to VM logic instantiation in the enterprises server, upon authorisation users can make connection and access resources in the network. Beside the security configurations done at the speed redundancy layer, high level security is implemented at the virtualisation layer since the layer supports extensive platform security. Different VMs are configured on the virtualisation/consolidation layer (VM_1 ...VM_{128}) using a VM ware tool to enable connections to different kinds of server such as web application servers, ERP servers, NAC servers etc. This work only focuses on the effect of OSPF as a link state protocol on our re-engineered DCN.

METHODOLOGY

Routing Protocol for Reengineered DCN Routers (OSPF)

A routing protocol is a set of processes, algorithms, and messages that are used to exchange routing information and populate the routing table with the routing protocol's choice of best paths. Routing protocols are needed for the following reasons:

a. They discover remote networks
b. They maintain up-to-date routing information
c. They choose the best path to destination networks

d. They have the ability to find a new best path if the current path is no longer available

Routing protocols operate in the following ways:

i. The router sends and receives routing messages on its interfaces.
ii. The router shares routing messages and routing information with other routers that are using the same routing protocol.
iii. Routers exchange routing information to learn about remote networks.
iv. When a router detects a topology change, the routing protocol can advertise this change to other routers.

The most commonly used routing protocols are as follows:

I. Routing Information Protocol (RIP): A distance vector interior routing protocol
II. Interior Gateway Routing Protocol (IGRP): The distance vector interior routing protocol developed by Cisco
III. Open Shortest Path First (OSPF): A link-state interior routing protocol
IV. Intermediary System- Intermediary System (IS-IS): A link-state interior routing protocol
V. Enhanced Interior Gateway Routing Protocol (EIGRP): The advanced distance vector interior routing protocol developed by Cisco
VI. Border Gateway Protocol (BGP): A path vector exterior routing protocol

Routing protocols can be classified as distance vector or link-state. Table 1 shows a comparison between a distance vector routing protocols and link state routing protocol with their examples. In this work, link state protocol implementation (OSPF) was chosen for the performance evaluation of the re-engineered DCN. This is due to the numerous advantages of link state routing protocols over distance vector routing protocols as shown in Table 1. Also, from the comparative study carried out for all the routing protocols, OSPF was found to be the best routing protocol that will suit large, scalable, low latency and cost effective data center network as in the case of our reengineered DCN.

Table 1: Distance Vector Routing versus Link State Routing Protocols

Distance Vector Routing Protocols	Link State Routing Protocols
Distance vector protocols periodically send complete routing tables to all connected neighbors.	After the network has *converged*, a link-state update is only sent when there is a change in the topology.
Distance vector protocols typically use the Bellman-Ford algorithm for the best-path route determination.	Link state routing protocols typically use the Dijkstra algorithm for the best-path route determination.
In large networks such as data center networks, the periodic routing updates can become enormous, causing significant traffic (congestion) on the links, hence, it is not suitable for such network	Since, a link-state update is only sent when there is a change in the topology the traffic (congestion) that would have resulted from periodic updates is avoided; hence, it is suitable for large network such as DCN.
Distance vector protocols use routers as signposts along the path to the final destination. The only information a router knows about a remote network is the distance or metric to reach that network and which path or interface to use to get there. They do not have an actual map of the network topology.	A router configured with a *link state* routing protocol can create a "complete view," or topology, of the network by gathering information from all the other routers. A *link-state router* uses the link-state information to create a topology map and to select the best path to all destination networks in the topology.
Distance vector protocols work best in situations where The network is simple and flat and does not require a hierarchical design. Worst-case convergence times in a network are not a concern.	link-state routing protocols work best in situations where The network design is hierarchical, usually occurring in large networks. Fast convergence of the network is crucial.
Examples of distance vector routing protocols include: RIP, IGRP, and EIGRP.	Examples of link-state routing protocols include: OSPF, IS-IS etc.

ISSN-L: 2223-9553, ISSN: 2223-9944
Vol. 3, No. 1, July 2012

SIMULATION RESULTS

In this section, we provide simulation results that support our proposed model. We used OPNET Modeler [14] for our modeling and simulations. OSPF was configured for broadcast multi-access for Ethernet network. All simulations presented in this paper use the following parameters and configurations in the OSPF parameter table comprising process attribute, interface information attribute and loopback interface attribute. Link propagation delays are 0.5µ secs with link speeds of 10Gbps.Using reference bandwidth of 1000Mbps, for client gateway 3-0, Cost = 5, for client gateway 2-3, Cost = 10, Administrative weight = 110, SPF calculation parameter = Periodic, OSPF Area Identifier = 0.0.0.1, 0.0.0.2, 0.0.0.3, IP Addressing = Auto Assigned, Links = Bidirectional PPP_DS3, Hello interval = 10sec, Router Dead interval = 40sec, Interface transmission delay = 0.5µ secs, Retransmission interval = 5secs.

DISCUSSION (STABILITY CONVERGENCE)

From our model in Figure 2, calculating the client gateway routes is based on a well-known algorithm from graph theory-Dijkstra’s shortest-path algorithm. OSPF introduces another layer of hierarchy into routing by allowing a domain to be partitioned into areas. This means that a router within a domain does not necessarily need to know how to reach every network within that domain—it may be sufficient for it to know how to get to the right area. Thus, there is a reduction in the amount of information that must be transmitted to and stored in each node. In addition, OSPF allows multiple routes to the same destination to be assigned the same cost and will cause traffic to be distributed evenly over those routers. This work presents stability convergence as a metric needed in our re-engineered DCN, enabling the entire network to respond swiftly to changes in topological database with low latency. Figure 4 and 5 show the plots for OSPF c onvergence behavior.

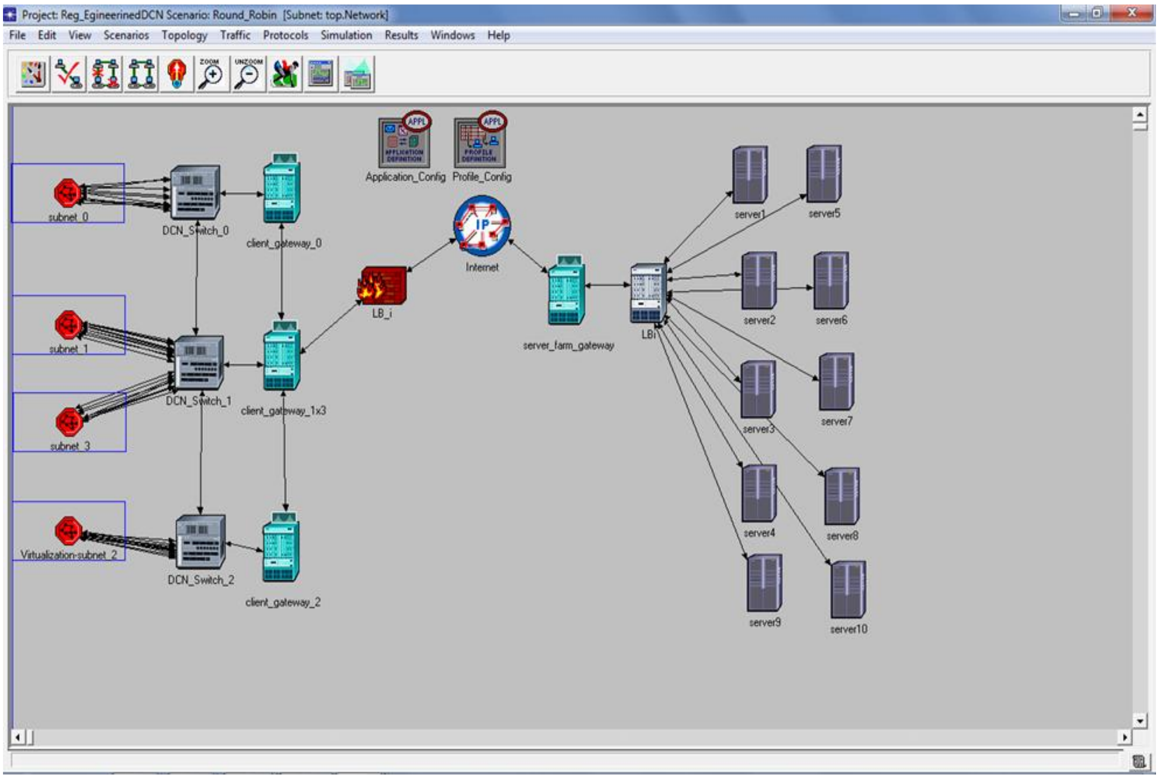

Figure 3: An Implementation Model of the Re-engineered DCN Architecture with OPNET Modeler

From Figure 4, OSPF periodic hellos are suppressed and the periodic refreshes of link state advertisement (LSAs) are not flooded over the demand circuit. These packets bring up the link only when they are exchanged for the first time, or when a change occurs in the information they contain. This operation allows the underlying data link layer to be closed when the network topology is stable. After initial topology synchronization, with the set parameters in above, a stable convergence was observed giving rise to an efficient packet transmission in our DCN model.

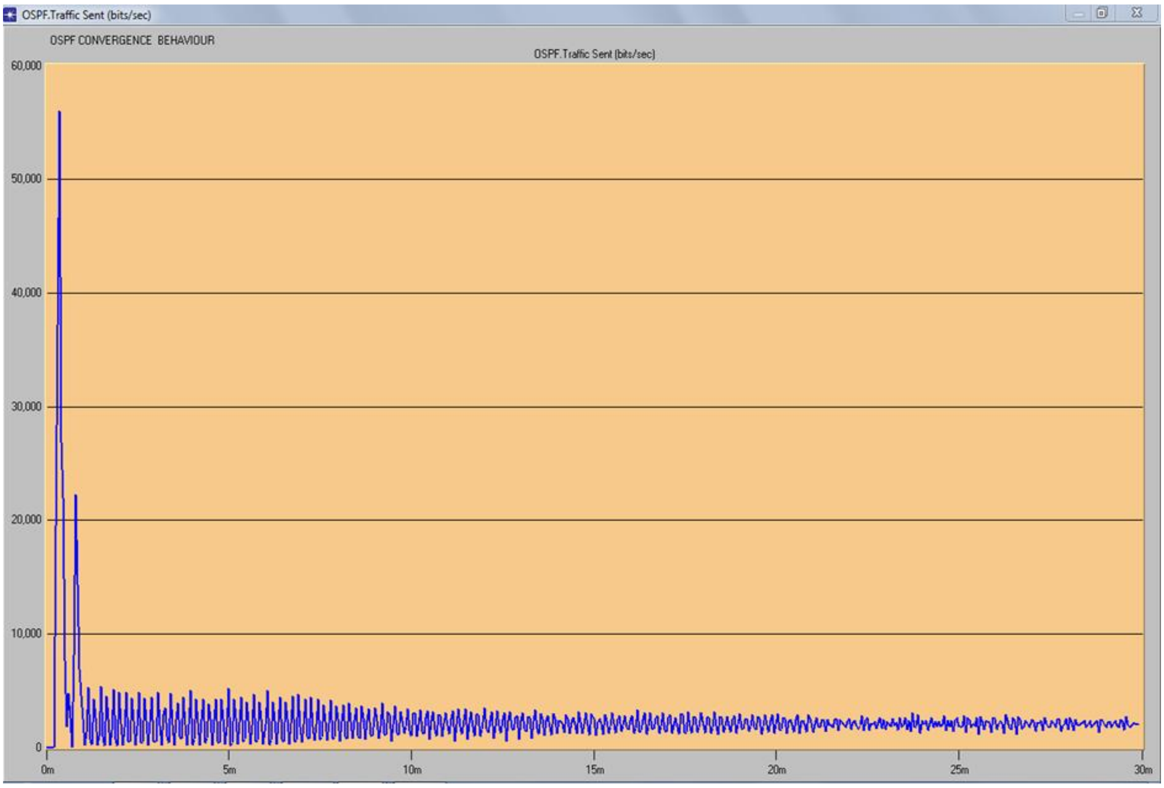

Figure 4: OSPF on demand circuits Convergence

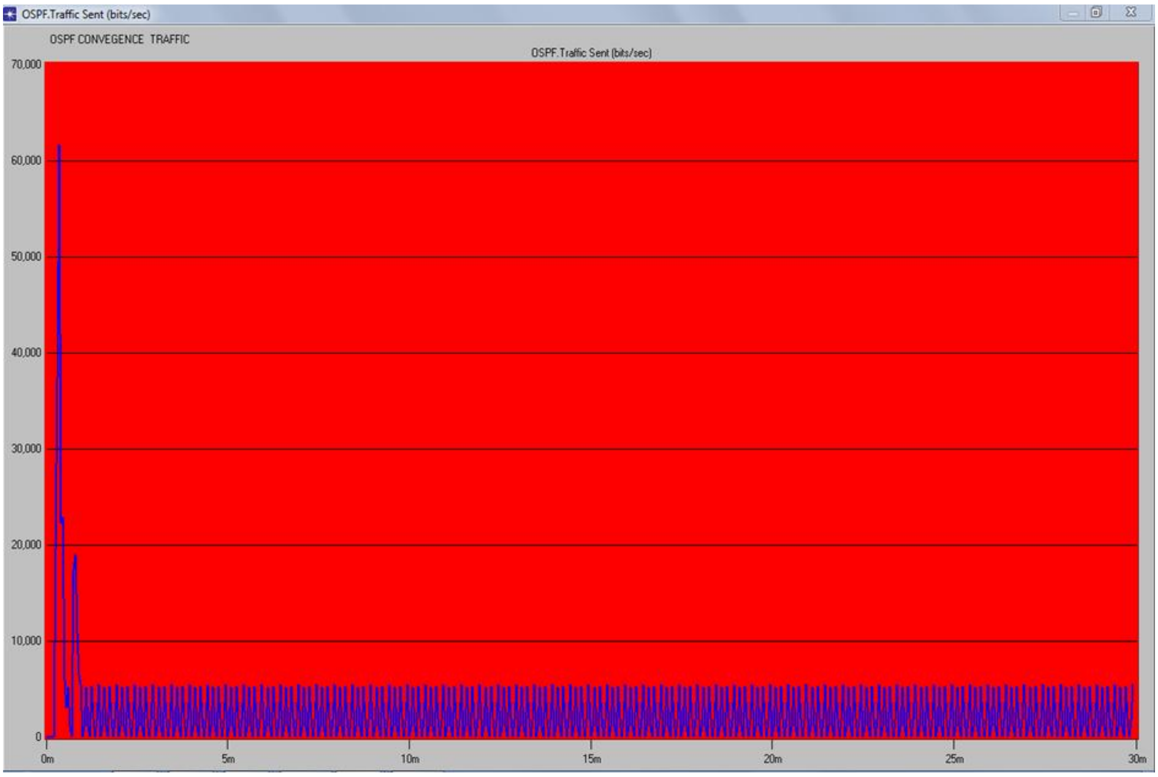

Figure 5: Absolute OSPF traffic Convergence

At a central site with the layered switches, telecommuters or branch offices can connect to an OSPF backbone. In this case, OSPF for on demand circuits allows the benefits of OSPF over the entire domain, without excess connection costs. Periodic refreshes of hello updates, LSA updates, and other protocol overhead are prevented from enabling the on demand circuit when there is no "real" data to transmit.

Also from Figure 5, Overhead protocols such as hellos and LSAs are transferred over the on demand circuit only upon initial setup and when they reflect a change in the topology. This means that critical changes to the topology that require new SPF calculations are transmitted in order to maintain network topology integrity. Periodic refreshes that do not include changes, however, are not transmitted across the link. Convergence in this context is completely stabilized, hence higher bit rate throughput for the segmented user groups in the reengineered DCN.

CONCLUSION

This work has presented OSPF as a scalable link state protocol for our proposed DCN which was modelled with OPNET Modeller. Most IT infrastructures were not built to support the explosive growth in computing capacity and information that we see today. Besides, many data centres have become highly distributed and somewhat fragmented such that they are limited in their ability to change quickly (absolute convergence) and support the integration of new types of technologies or to easily scale to power the business as needed. Thus this work re-engineers legacy data center networks for a more scalability with optimal cost and good service delivery using the link state OSPF protocol. The implementation of the re-engineered data center network architecture presented in this paper provides the best possible solution for enterprise network convergence.

REFERENCES

Emerson Network Power white paper, Balancing Scalability and Reliability in the Critical Power System: When does 'N+1' Become 'Too Many +1'? http://www.EmersonNetworkPower.com, 2004.

Emerson Network Power white paper, *Power Management Strategies for High-Density IT Facilities and Systems*, http://www.EmersonNetworkPower.com, 2007.

Emerson Network Power white paper, *Managing Extreme Heat: Cooling Strategies for High Density Systems*, http://www.EmersonNetworkPower.com, 2007.

Oracle white paper, *Architectural Strategies for Cloud Computing*, http://www.Oracle.com, 2009.

www.juniper.net: Technical White Paper; Cloud-Ready Data Center Reference Architecture

www.hp.com: Technical White Paper; Building Virtualization-Optimized Data Center Networks. Febuary 2011, page 2.

Thomas B Winans, John Seely Brown. Cloud Computing: A Collection of Working Papers. Deloitte Development LLC, 2009.

http://ibm.com/datacenter. Technial White Paper. The New Enterprise Data Center. IBM Corporation, May, 2008.

http://www.datacenterknowledge.com/archives/2008/11/07/amazon-building-large-data-center-in-oregon/

www.juniper.net: Technical White Paper; Cloud-Ready Data Center Reference Architecture

www.wikipedia.com

H. C. Inyiama, Okafor Kennedy. C, Udeze Chidiebele. C: Virtualization: A Cost Effective Approach to Infrastructure Management in Data Center Networks. (A Case for Elastic Sky X Server) (Unpublished).

R. J. Adair, R. U. Bayles, L. W. Comeau, R. J. Creasy, "A Virtual Machine System for the 360/40", IBM Corporation, Cambridge Scientific Center Report No. 320-2007, 1966.

OPNET (2006). *OPNET network simulator.* Retrieved February 21, 2006, from http://www.opnet.com

P. H. Gum. System/370 extended architecture: facilities for virtual machines. IBM Journal of Research and Development, 27(6):530–544, Nov. 1983.

L. Seawright and R. MacKinnon. VM/370 – a study of multiplicity and usefulness. IBM Systems Journal, pages 4–17, 1979.

S. Devine, E. Bugnion, and M. Rosenblum. Virtualization system including a virtual machine monitor for a computer with a segmented architecture. US Patent, 6397242, Oct. 1998.

Connectix. Product Overview: Connectix Virtual Server, 2003. http://www.connectix.com/products/vs.html

A. Whitaker, M. Shaw, and S. D. Gribble. Scale and performance in the Denali isolation kernel. In Proceedings of the 5th Symposium on Operating Systems Design and Implementation (OSDI 2002), ACM Operating Systems Review, Winter 2002 Special Issue, pages 195–210, Boston, MA, USA, Dec. 2002.

A. Awadallah and M. Rosenblum. The vMatrix: A network of virtual machine monitors for dynamic content distribution. In Proceedings of the 7th International Workshop on Web Content Caching and Distribution (WCW 2002), Aug. 2002.

L. Peterson, D. Culler, T. Anderson, and T. Roscoe. A blueprint for introducing disruptive technology into the internet. In Proceedings of the 1st Workshop on Hot Topics in Networks (HotNets-I), Princeton, NJ, USA, Oct. 2002.

J. Gelinas. Virtual Private Servers and Security Contexts, 2003. http://www.solucorp.qc.ca/miscprj/s_context.hc

A. Bavier, T. Voigt, M. Wawrzoniak, L. Peterson, and P. Gunningberg. SILK: Scout paths in the Linux kernel. Technical Report 2002-009, Uppsala University, Department of Information Technology, Feb. 2002.

Leena Jain and Sushil Bhardwaj, “Enterprise Cloud Computing: Key Considerations for Adoption” International Journal of Engineering and Information Technology Vol 2, (2010). IJEIT 2010, 2(2), 113-117 ISSN 0976-0253 (Online).

D. Nurmi et al., “The Eucalyptus Open-source Cloud-computing System”

Rajkumar Buyya et al, “Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility, Future Generation Computer Systems, 25 (2009) 599-616, Journal homepage: http://www.elsevier.com/locate/fgcs.