Design and Optimization of Metropolitan Area Networks using Multi-Protocol Label Switching

Alireza Emamjomeh¹, Hossein Emami², Ali Hashemi³

1- Department of Electrical Engineering, Majlesi Branch, Islamic Azad University, Isfahan, Iran Email: Alireza19838@yahoo.com

2- Department of Electrical Engineering, Majlesi Branch, Islamic Azad University, Isfahan, Iran Email: h.emami@iaumajlesi.ac.ir (Corresponding author)

3- Department of Electrical Engineering, Majlesi Branch, Islamic Azad University, Isfahan, Iran Email: a.hashemi@iaumajlesi.ac.ir

Received: March 2015

Revised: June 2015

Accepted: July 2015

ABSTRACT:

Looking in to the modern technology, we see new inventions and innovations which are unique. As an example, during the past years, new technologies have been created in computer networks like metropolitan area network (MAN) that has had a noticeable advance in urban communication. Therefore the aim of this article is to provide and compile information regarding MAN technology and MPLS. Utilizing Ethernet in urban networks needs a developing potential and strength that only exists in MPLS and IP levels. Two layer and three joint networks are the solutions that join simplicity and economy of Ethernet at IP/MPLS scale. Different transfer technologies have been used in MAN so metropolitan services should be given on a mixture of optic and data switches, MPLS have been developed for this work and using generalized-MPLS makes MPLS deployments possible even on none-IP-aware switches. At first, theories about using MAN technology to produce urban networks are explained. Then a set of problems regarding the management of resources and provisioning barriers are introduced following the solutions by means of IP/MPLS network which have been implemented and validated in eSNP simulator.

KEYWORDS: Ethernet, MPLS, MAC, IP, VLAN, VPN, CPU, Memory

1. INTRODUCTION

Over the years, the discussion about achieving knowledge and discovering unknowns have attracted a lot of people and occasionally a news or text or short understanding of a new science absorbs scholars and people motivated to learn. In the past decades ,Ethernet has been accounted of a built technology .Ethernet mediators enable you to have more bandwidths with lower cost .these Ethernet parameters (high bandwidth and low cost)and good performance ,simulates network Ethernet in using Ethernet as an access layer. Ethernet access has some good consequence for TDM network. It is clear that Ethernet is a known technology; simple but practical and with a wide range of usage. Now Ethernet is used as the superior technology widely in LAN and is practically a standard

for the connection between the service providers and customers. The success of Metro Ethernet Services caught the imagination of the world when the concept expanded to include worldwide services traversing national and global networks [16], since the owners of local networks are interested in communicating beyond the company's borders, Ethernet was noticed as a cheap and simple solution. The same change can be observed in residential environments that is of course with the development of the different technology of access to wide spread networks such as DSL and coaxial cable, Ethernet has entered local networks slowly and it's using in many cases, has provided high speed internet. In order to know MAN, it is necessary to define the approaches to build it on a massive range rather than a small office.

2. PREVIOUS RESEARCH

Ali, Chiruvolu, and Ge (2005) investigated the outlines novel algorithms for multipoint-TE in the metro Ethernet. They suggested a grouping scheme that extended the current label space in the provider domain and allowed for a large number of VLANs to be provisioned efficiently. They analyzed the issues of load balancing, multiple spanning trees, and interaction between grouping and bandwidth provisioning, and suggested solutions. They explained differentiated survivability in next-generation Ethernet and provided a novel scheme based on multiple spanning trees.

Padmaraj, Nair, Marchetti, Chiruvolu, Ali, and Ge (2005) investigated traffic engineering which is one of the integral components of QoS provisioning. They offered a scheme based on the generation and management of multiple spanning trees for near optimal traffic distribution.

Wang, Lynch,Ll, Klincewicz, Li, Doverspike, and Segal (2010) explained a methodology to enable the rapid introduction of metro Ethernet networks. They utilized a combination of numerous optimization algorithms and has integrated it into a pragmatic tool used by AT&T network planners. Case studies showed that the tool gave cost-effective solutions consistent with planner expectations and intuition.

Lianzhi (2010) examined reliability theory of computer communication network and simulation experiment revealed that utilizing the system redundancy, 2 link parallel redundancy, and interline hybrid series-parallel redundancy, could better solve reliability of Metro Ethernet.

Torki, Mirjalily, and Saadat (2011) proposed an efficient fast algorithm to find the best spanning tree by modeling and solving objective functions based on load balancing criterions. Utilizing this objective they could balance the traffic load on links and switches and showed the effectiveness of their approach.

3. HYPOTHESES

The existing network first was based on pure IP protocol that was just designed for a single service in which the end-users used that infrastructure, to connect to the internet. But due to the appearance of organizations acting as wholesalers and retailers, the demand for high speed interconnectivity whole over the city increased dramatically along with the primary services which made the demand for making a high speed common base connecting every part of the city to be able to serve both end-users and organizations in a massive range at the same time [12].

In this new model, networks are spanned city wide. Each organization has its own comprehensive network that ties with the common base as a carrier or even those organizations can act as carriers for the other ones while having their own customers. Apart from bandwidth requirements and the essence of quality of service (QOS) deployments in provider's network, the limitations in number of public IP addresses and high amount of CPU and memory usages at the core network are the main barriers to accommodate all customers' needs; at the same time, to avoid costly network expansions.

3.1. The statement of the hypotheses

3.1.1.CPU and memory limitations:

Almost in all cities around the world, TELCOs are the owners of the fiber connections underground. To act as a carrier network whole over the city, they have their own infrastructure containing access, edge and core routers, transmission multiplexers and central servers for monitoring, authentication, and authorization and accounting jobs. Besides, customers build up the same infrastructure to do the same jobs. However, those two mentioned networks above are completely properties of two different autonomous systems constructing separate management domains.



Fig.1 . Management domains belong to different autonomous systems

Customers use the leased lines over a shared bandwidth from TELCOs to interconnect their core routers and they serve end-users as ISPs or GSM/LTE mobile operators and etc. In fact, the provider satisfies the customer needs by configuring the QOS at the management domains boundaries such as traffic policing and congestion management and guarantees the consistent connectivity over the network with the least loss and delay rate [13].

But, to build up a full connectivity among end-users of different customer networks whole over the city, their traffic must be routed hierarchically over the nearest Vol. 9, No. 4, Decem, ber 2015

customer network and consequently throughout the provider network as a carrier and finally to the other customer network [10].



Fig. 2. Traffic pattern among different management domains

In terms of TCP/IP connection rules, a successful bidirectional communication among two end devices is accomplished by the virtue of their IP addresses' presence in the databases (routing tables) of all transient routers in a train, both on customers and provider networks [14].

That means, on provider network point of view, the growth of customers' networks becomes dependent to the growth of provider's core network. To put it more simply, as the number of customers and their end-users increases, the provider core routers must have their CPU and memory expanded to accommodate those routes entries in their routing tables. In other words, rapid expansion of services requires rapid expansion of the core routers. This is not cost effective and also increases delay in service provisioning even more than some months as the expansion is too time consuming as well. The following is the illustration of the above statement.



Fig. 3. Customers' growth impact on service provider's network

The above illustration implies that, as a rule of thumb, having three customers, if each customer expands its business twice during a month, the provider network must go for six times and even more for future expansions.

Now assume a provider which has initially p number of routes as its infrastructure apart from its customers' routes. i is a number assigned to each customer in order they sign a contract with provider. n shows the maximum numbers of customers that this provider has. K is the number of times that this provider's customers increased their number of routes independently and advertised them to the provider and r shows the current number of each customer's route at the change time. As a result, P(k) which is the current number of routes in the provider's database at each change time is as follow:

$$P(k) = p + \sum_{i=1}^{n} \sum_{j=1}^{k} (ri(j) - ri(j-1))$$
(1)

To show the customers' growth impact on the provider expansion needs, let's assume two customers with 0 numbers of routes initially as well as provider. Then illustrate their everyday growth 10 times by a set of random numbers and compare $\frac{\partial ri}{\partial k}$ with $\frac{\partial P}{\partial k}$ on the diagram visually.

Vol. 9, No. 4, December 2015

Table 1. Samples of customers' and provider's numbers

of routes			
Κ	r1(k)	r2(k)	P(k)
K=0	0	0	0
K=1	2	4	6
K=2	6	6	12
K=3	10	12	22
K=4	14	14	28
K=5	18	16	34
K=6	22	17	39
K=7	25	20	45
K=8	31	23	54
K=9	32	29	61
K=10	33	35	68





Referring to the above diagram, it is obvious that by the growth of customers' networks, the provider runs out of CPU and memory more rapidly.

However, in the new proposed solution, we aim to uncorrelated the customers' networks growth with provider's core routers. Instead of shifting this burden toward the edge routers. Accordingly, in this model P(K) on the core routers and just the edge ones follow up the previous formula.

3.1.2. IP addresses limitations:

The TCP/IP communication borrows the idea of post. Considering the conveyed information among two

network devices as a packet and each side of this conversation as a sender and receiver, each router's interface IP address is considered as a sender or receiver address (depending on the direction of the communication which must be bidirectional) and that interface's MAC address as a transient post office in the middle of sender and receiver. In a network that the corresponding devices communicate with each other, all the IP addresses in the whole network and MAC addresses in each segment (segment is a link which interconnects two adjacent routers with each other) must be unique [1]. This uniqueness puts barrier in front of provider to accommodate more customers while it has enough bandwidth and empty ports. In other words, as the number of customer increases, the provider runs out of IP addresses to be assigned to them. Consequently, it must obtain new sets of unused IP addresses to increase the chance of getting new customers. Not only but also, it must control the IP addresses conflicts among different customers using security solutions [7]. However, it is not cost effective and practical in very large scales.



Fig. 5. Filtering a new customer's IP addresses that are assigned to another customers or service provider's infrastructure

To describe the above issue mathematically, let's assume that T stands for the total number of IP addresses that this provider has and C represents the current number of customers. B is the total amount of bandwidth that is available on a shared link among

Vol. 9, No. 4, Decem, ber 2015

different customers and finally, *bi* is the amount of bandwidth randomly reserved by each customer on the link. Eventually, as the number of customers increases, the amount of available bandwidth and consequently the amount of available IP addresses reduce in parallel. So, the chance for each new customer to utilize the provider network and sign a contract will reduce by the following formula:

$$p1\% = \left(\frac{T-C}{T}\right) \times \left(\frac{B-\left(\sum_{i=1}^{C}bi\right)}{B}\right) \times 100$$
(2)

The first part of the above formula represents the impact of IP addresses limits and the second part shows the impact of bandwidth limit imposed on new customers.

If the provider can afford to get so many IP addresses in stock to make the first part of the formula approximately 1, then the chance reduces just by the impact of bandwidth consumptions by the other customers as the following:

$$p2\% = \left(\frac{B - \left(\sum_{i=1}^{C} bi\right)}{B}\right) \times 100$$
(3)

Assume that a provider has 30 IP addresses and wants to accommodate 30 customers over a link policed at 1430Mbps to be shared among them. If we randomly divide this amount of bandwidth among 30 customers, then here is the calculated chance for each new customer to get the required service in comparison to when the provider has unlimited number of IP addresses:

 Table 2. Samples of each new customer's chance to

 utilize provider's network as the numbers of customers

 increase w/o IP address constraints

increase w/o ii address constraints			
С	bi	<i>p1%</i>	<i>p2%</i>
0	0	100	100
1	59	92.67832168	95.87412587
2	6	89.09090909	95.45454545
3	23	84.46153846	93.84615385
4	73	76.90909091	88.74125874
5	87	68.88111888	82.65734266
6	10	65.56643357	81.95804196

7	17	61.92307692	80.76923077
8	50	56.66666667	77.27272727
9	72	50.56643357	72.23776224
10	91	43.91608392	65.87412587
11	55	39.28438228	62.02797203
12	61	34.65734266	57.76223776
13	23	31.82051282	56.15384615
14	56	27.86013986	52.23776224
15	42	24.65034965	49.3006993
16	43	21.6037296	46.29370629
17	40	18.84848485	43.4965035
18	52	15.94405594	39.86013986
19	15	14.23076923	38.81118881
20	13	12.63403263	37.9020979
21	86	9.566433566	31.88811189
22	37	7.813519814	29.3006993
23	6	6.738927739	28.88111888
24	56	4.993006993	24.96503497
25	59	3.473193473	20.83916084
26	32	2.48018648	18.6013986
27	77	1.321678322	13.21678322
28	16	0.806526807	12.0979021
29	80	0.216783217	6.503496503
30	93	0	0



Fig. 6. Comparison of a new customer chance to utilize provider's network w/o IP address constraints

The above diagram shows an apparent drop for new customers' chances of getting services when there are IP address limits in provider network. However, in this proposal we aim to introduce a solution to re-use the IP

Vol. 9, No. 4, December 2015

addresses assigned to different customers and isolate their databases from provider database. In this new solution, the new customers' chances are increased just by competing to overcome provider bandwidth limits.

3.1.3 Proposed network as a solution:3.1.3.1 Organized network topology:

As mentioned before, when it comes to manageability and scalability, a service provider network should be divided into several parts and then interconnected hierarchically over a common base. These parts include customer services' edges, server farms, BRAS farms and NOC (network operation centers). With reference to ITIL and TMN recommendations, there must be redundancies at network layers by duplicating each part at least twice [9].



Fig. 7. provider's network redundancy

In terms of hierarchical design, the common base routers should be divided into three parts: UPE, NPE and core routers. Then, the UPE devices should be aggregated by middle-speed links to NPE devices and consequently all NPE devices should be concentrated by high speed links to core routers. This method of design predetermines the traffic pattern of each type as well as providing redundancy at link levels [8]. Moreover, the hierarchical cost design in this model makes load balancing at bottom-tier (between UPE and NPE) and up-tier (between NPE and core).



Fig. 8. provider's network hierarchical design 1



Fig. 9. Provider's network hierarchical design 2

3.1.3.2. Virtual private networks:

Telecom networks, to act as a shared media to transport different kinds of protocols build up virtual private networks in most cases.

Initially, the frame-relay was widely used to build up overlay VPNs to make dedicated connections among different customers' sites over telecommunication networks. That means, static virtual circuits (SVC) would be defined among two edge devices of a provider consequently, simulating a layer 2 connection in OSI model among customer sites which isolates that customer's network from provider and the other customers as well [2].

However, due to the lack of speed in such networks they became out of date soon. As an alternative, ATM and SONET/SDH solutions were introduced [5].

But they were replaced by Ethernet connections since the introduction of carrier Ethernet over optical networks at higher speeds with lower costs per port as well as the evolution of WAN Ethernet defined in RFC#3637 which has made it flexible to be integrated Vol. 9, No. 4, Decem, ber 2015

into the SONET/SDH networks [4].

Moreover, the ATM and SONET/SDH networks still have connection-oriented nature and transport packets by circuit switching rather than packet switching which has lower flexibility in transportation of bulk data. That is, SVCs must still be established manually to make a dedicated point-to-point connection. Thus, it brings extra loads of management efforts such as building up redundant paths, bandwidth management and troubleshooting. Nowadays, ATM networks are widely used in broadband communications (ADSL) and SONET/SDH networks are good for transporting voice packets due to having high MTU. However, they are not good choice for transportation of bulk data mixed with VOICE and IPTV packets.

Since the introduction of IPSEC VPNs in RFC#2401, it has become possible to build up secure connections among different sites and/or clients over an unsecured shared media like internet by the virtue of packets' authentication and encryption. However, this type of VPN doesn't solve the problem of CPU and memory resource management in telecommunication networks in high scale rather it puts extra burden. It is not also a scalable solution in massive ranges due to its point-topoint connectivity nature.

as the definition of MPLS protocol in RFC#3031, the benefits brought by IP and Ethernet protocols such as providing higher speeds with lower costs, automatic bandwidth management plus provisioning and connectionless communication were mixed with the benefits brought by ATM such as building up end-toend layer 2 overlay VPNs over a shared network [6], [15]. Later on, in RFC#2547, the BGP protocol was enhanced to work on top of MPLS protocol to build up layer 3 VPNs as a scalable and effective solution. MPLS protocol works based on TCP/IP stack. By the evolution of GMPLS, it has become possible to establish MPLS LSPs (Label switching paths) without the use of TCP/IP stack and deployment of IGP protocols [3], [11].

3.1.3.3. Reducing CPU and memory usage in provider networks using MPLS VPNs:

As mentioned before, the growth of customers' networks is tied with the growth of provider network and having n number of customers, the speed of growth in provider network is n times more than each customer. The main reason is that the customers' databases are correlated with the provider's database. Naturally, by taking look at the organized network structure we introduced in 5.1.3.1, this question comes to mind that how it is possible to isolate database of each part from the other parts?

To solve this problem, firstly we must make a separate database for each part of the network at UPEs which is called VRF (virtual routing forwarding) table. One part's database just contains the routing table of that part only such as common base, customers' services, server farms, BRAS farms and NOC. The VRF contents are advertised by the corresponding UPEs to the remote UPEs introducing themselves as the next hops to reach them. UPEs of each part just need to know the database of that part along with the common base's database while the NPEs and CORE routers just need to have common base's database.

Secondly, each UPE's location must be determined by a numerical label inside the common base called MPLS label. To reach a destination of a part, the source UPE searches for the corresponding IP address inside the relevant database and finds out which other UPE inside the common base accommodates that destination, then labels that packet with the correct destination UPE's label and sends the packet toward that UPE. The NPEs and CORE routers just look at the attached label and understand that the packet must be conveyed to which UPE without of knowing the location of the destination IP address. When the packet reaches at the destination UPE, it detaches the label and searches the corresponding part's database and routes the packet toward the destination.



Fig. 10. Reducing the core network's routing table by just limiting their databases to know the common base routes

3.1.3.4. Overcoming the IP addresses limitations in provider networks using MPLS VPNs:

As mentioned before, while the number of customers increases, there must be enough IP addresses in stock to accommodate their needs. Referring to the previous formula in 5.1.2; even the provider must have much more IP addresses than the number of customers to increase the customers' chances to get new contracts which is not cost effective.

To solve this problem, we must make a solution to reuse IP addresses assigned to each customer. This can be done by adding extra information to each customer's routes. That is, each customer's routes must be labeled numerically to distinguish the corresponding routes from other customers. This label is called route distinguisher. When this label is attached to a customer's routes while being advertised to other UPEs, makes that route unique even if the IP addresses are the same. Hence, each customer's database is assigned a unique route distinguisher. The same approach can be taken if we decide to re-use the IP addresses assigned to the customers in other parts of the provider network such as NOC, BRAS and server farms.



Fig. 11. IP address reuse

4. THE RESULT OF HYPOTHESES TESTING

A provider network which was built up in an old fashion was taken. This network was previously built up totally based on SDH and IP networks. It had been several months that it had faced losses due to lack of IP addresses which had prevented it to get new customers as well as having so many problems because of high memory and CPU utilizations in its core network as a result of having too many routes in their routing tables which had caused losing in SLA contracts to the extent that some of their customers decided to not continue using their network as a carrier. By applying several changes and cutovers they upgraded their network into MPLS network. As a result the number of CPU and memory utilizations in their core network reduced remarkably. Here is the record of the number of routes in routing tables of their core routers during some months before and after migration into the MPLS network following its graphical illustration.

Table 3	Samples of provider's network database
	volume before and after MPLS

provider DB before	provider DB after
MPLS	MPLS
8126	8126
8230	8126
8310	8127
8424	8127
8910	8127
9230	8128
9313	8128

Vol. 9, No. 4, Decem, ber 2015

9512	8128
10073	8128
10225	8128
10546	8131
10910	8131



Fig. 12. Comparison of provider's network database volume before and after MPLS

At the same time, the network growth of one customer has been monitored and recorded along with the growth of provider network before and after migrating into the MPLS network. To calculate the dependence of these two databases, we utilized Pearson's correlation coefficient formula which is the most familiar measure of dependence between two quantities as the following:

$$r(x,y) = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sqrt{\sum (x-\bar{x})^2 \sum (y-\bar{y})^2}}$$
(4)

In the above formula, as much as the calculated coefficient reaches 1 it shows more dependency among x and y data sets. In contrast, when it reaches 0, it implies that the two data sets are more uncorrelated.

Here is the record of a customer routing table growth along with the provider growth before and after migrating into the MPLS network following the calculated correlation coefficient in both cases.

 Table 4. Samples of provider's network database

 volume before and after MPLS beside the customer's

 growing database volume

growing database volume			
customer	provider DB	provider DB	
DB	before MPLS	after MPLS	
400	8126	8126	
410	8230	8126	
415	8310	8127	
425	8424	8127	
470	8910	8127	
510	9230	8128	
518	9313	8128	
536	9512	8128	
587	10073	8128	
601	10225	8128	
625	10546	8131	
632	10910	8131	

Before MPLS:

r(r, y) = 0.994549	131936801 (5)	
T(x, y) = 0.994549	121320001 (.	3)	

After MPLS:

 $r(x, y) = 0.872523898683138 \tag{6}$

After migrating to MPLS network, the correlation coefficient is reduced by 1/10.

Regarding the IP address resources, this provider could utilize the assigned IP addresses again and again. Consequently, as the number of customers increases, there is no need to get new sets of IP addresses to accommodate them. In other words, when re-using the previously assigned IP addresses becomes possible, T - C = T in calculation of p1%so p1% = p2%. Thus, by increasing the number of customers, if the provider decides to keep up the new customers' entrance chances to 50% constantly over a 1430Mbps shared link, it just costs to upgrade that link capacity approximately every 725Mbps of its occupation.

5. CONCLUSION

MPLS (multi-protocol label switching) technology is a new method to get a shared quick communication infrastructure to provide a variety of services based on the clients' needs.By implementing the MPLS VPNs in provider networks, it becomes possible to accommodate all clients in a massive range without running out of resources such as IP addresses, CPU and memory. Moreover, it makes a sustainable common base whole over a city without the need for upgrading it as the number of city inhabitants grows up. By utilizing this network and having high-speed connections (10Gbps, 40Gbps, and 100Gbps) in MPLS core, it is possible to meet all clients' needs on a shared infrastructure in long term.

REFERENCE

- [1] A., Frank John, M. Ruediger Arnd, "Method of configuring a valid IP address and detecting duplicate IP addresses in a local area network", U.S. Patent, No. 5, 724,510., 3 Mar. 1998.
- [2] A., Loa, Tove Madsen, "Provider provisioned virtual private network (VPN) terminology", *RFC 4026*, March, 2005
- [3] E. Mannie, "Generalized multi-protocol label switching (GMPLS) architecture", *Interface*, 501, 19, 2004.
- [4] N. Fabio, J. Solé Pareta Xavier Masip Bruin, and S. Sánchez Lopez, "Optical Network Design and Modeling", 2007, 1-1
- [5] E. Nestle, F. Robert Schunneman, "TIME DIVISION MULTIPLEXING", U.S. Patent No. 3,599,160. 10 Aug. 1971.
- [6] Sh. Laxman, Sh. Ambe, "Fastpath implementation for transparent local area network (LAN) services over multiprotocol label switching (MPLS)", U.S. Patent No. 20,030,174,706., 18 Sep. 2003
- [7] P. Ferguson, D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", *RFC 2827*, May 2000.

- [8] R. R. Boorstyn, H. Frank, "Large-scale network topological optimization", *IEEE Transactions Communications*, 1977.
- [9] J. Brian Weatherred, et al., "Link aggregation path selection method", U.S. Patent No. 6,535,504., 18 Mar. 2003.
- [10] R. Govindan, A. Reddy, "An Analysis of Internet Inter-Domain Topology and Route Stability, Proceedings of the INFOCOM '97", Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Driving the Information Revolution, pp. 850, April 09-11, 1997
- [11] E. Mannie, "Generalized multi-protocol label switching (GMPLS) architecture", *Interface*, 501, 19, 2004.

Vol. 9, No. 4, Decem, ber 2015

- [12] R. Santitoro, "Metro Ethernet Services-A Technical Overview", http://www.metroethernetforum.org, Apr. 2003, pp. 1-19, v.2.5, Metro Ethernet Forum, U.S.A.
- [13] Y. Bernet et al., "A Framework for Integrated Services Operation over Diffserv Networks", *RFC 2998*, pp. 1-29, Nov. 2000.
- [14] J. Moy, "Link-state routing. In Martha E. Steenstrup", editor, *Routing in Communications Networks*, pp. 135-157. Prentice Hall, 1995.
- [15] E. Rosen, A. Viswanathan, and R. Callon. "Multiprotocol Label Switching Architecture", *Internet RFC 3031*, January 2001.
- [16] Metro Ethernet Forum, MEF. Retrieved 30 May 2011.