Rev. Roum. Sci. Techn.– Électrotechn. et Énerg.
Vol. **63**, *2*, pp. 202–209, Bucarest, 2018

Dedicated to the memory of Academician Andrei Țugulea

PROPOSED SCENARIOS OF THE LOAD BALANCING MECHANISM IN MULTIPROTOCOL LABEL SWITCHING NETWORKS

EMA-MARIA GALEȘ[1], VICTOR CROITORU[1], MARIUS IORDACHE[2]

Key words: Algorithm, Multiprotocol label switching (MPLS), Load balancing, Tunnel, Virtual private network (VPN), Multiprotocol-border gateway protocol (MP-BGP), Label.

The present paper emphasizes the crucial benefits of Internet protocol/multiprotocol label switching (IP/MPLS) networks. This protocol implies broad development of the information transmission between users and successful implementation of algorithms and methods that might generate a balance in the data load passing through the links between devices. MPLS proposes imperative solutions to network architectures that are affected by rapid changes of the properties of traffic, such as traffic peaks that must be controlled. Additionally, this document underlines the efficiencies of MPLS networks such as lower costs, quality of service attributes, scalability and vigorous traffic routing by creating two scenarios and analyzing their characteristics.

1. INTRODUCTION

As the networking domain has rapidly evolved during the last years, the need for dynamism in a communication system is crucial. Time, costs, efficiency and reliability are valuable aspects that are to be taken into consideration when delivering various services to the customers, from the point of view of a service provider. Consequently, there is a small number of possibilities regarding the manner in which the information of any type is transmitted through the networks [1].

The solution is represented by multiprotocol label switching (MPLS), which is increasingly used in the recent networks, because it greatly differs from the common IP routing by many particular features. In the Internet Protocol (IP) world, many aspects of MPLS are seen as foreign, because of the fact that they can be considered extremely complicated or they can be mistakenly presented. Even if there are imple-mentations that utilize only basic characteristics of MPLS, it is still a very powerful tool that greatly influences any network [1].

2. IP/MPLS NETWORKS

First of all, it does not use IP addresses to route the traffic, but some labels that have local meaning (between two routers). Secondly, it enables the mapping of real-time information (voice, video) to links that have a low latency, which is difficult in the case of IP routing. Thirdly, the circuit-based forwarding is easily replaced by packet-based structures. Figure 1 presents the placement of MPLS in the open systems interconnection (OSI) stack.

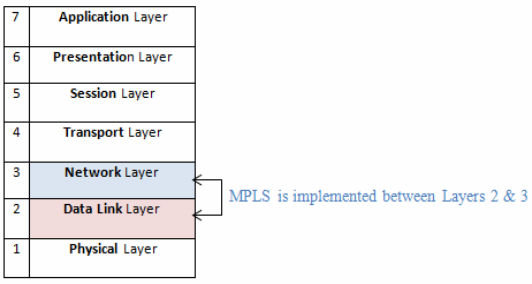

Fig. 1 – The position of MPLS in the OSI model.

MPLS is increasingly used in the recent networks, because it greatly differs from the common IP routing by many particular features. In the IP world, many aspects of MPLS are seen as strange, because of the fact that they can be considered extremely complicated or they can be mistakenly presented. Even if there are implementations that utilize only basic characteristics of MPLS, it is still a very powerful tool that greatly influences any network [1].

2.1. EVOLUTION

MPLS is found somewhere between the data link layer and the network layer and it supports end-to-end circuits over any type of transport medium, using any network layer protocol. It combines the performance and simplicity of Layer 2 with the flexibility and scalability of Layer 3 [2].

The main advantage is the transmission based on labeling the ingress packets based on their destination address or a preconfigured criteria and the switching of the traffic over a common infrastructure. MPLS gives the possibility of transporting IP version 4 (IPv4), IPv6, Ethernet, Point to Point Protocol (PPP) and other Layer 2 technologies over it.

2.2. MPLS ARCHITECTURE

The MPLS networks are built in order to connect different local area networks (LANs), by use of some equipment that have specific roles. In order to understand how MPLS works, one has to study the MPLS labels and their role when attached to the packet in the traffic forwarding. The MPLS header has the structure of 32 bits (4 bytes), divided into 4 sections. The label stack may have one or more labels in its structure, depending on the MPLS applications (MPLS virtual private network (VPN) and Any Transport over MPLS (AToM) need two labels in the stack). The first label is the top label and the last one is the bottom label. The number of labels between these two can be countless. The bottom of stack (BoS) for all the labels except the bottom label is set to 0. The MPLS label stack is positioned before the Layer 3 packet, precisely the header of the transported protocol and after the Layer 2 header [1].

Figure 2 presents a possible architecture of an MPLS network, including the most important constitutive elements.

[1] "Politehnica" University of Bucharest, Faculty of Electronics, Telecommunications and Technology of Information, Romania, E-mail: gales.ema@gmail.com, croitoru@adcomm.pub.ro
[2] Orange Romania, E-mail: marius.iordache@orange.com

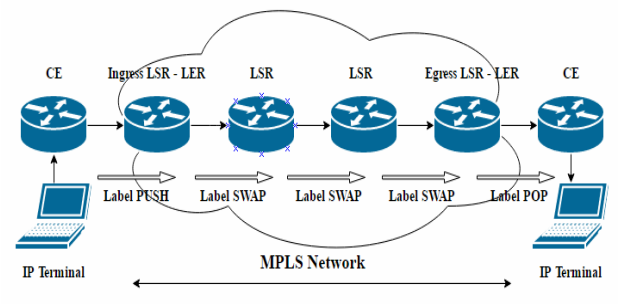

Fig.2 – MPLS network architecture.

2.3. MPLS OPERATION

The label switch routers (LSRs) are the networking equipment found inside the network and they can also be called transit LSRs and they split into:

a) Ingress LSRs, which receive packets that are not labeled and insert a label/a label stack in front of them, sending afterwards the packets on a data link.

b) Egress LSRs, which receive labeled packets and remove the label(s), sending them afterwards on a data link. Ingress and egress LSRs are also named edge LSRs, because they are situated at the edge of an MPLS network.

c) Intermediate LSRs, which receive a labeled packet and afterwards they make an operation on it, they switch it and send the packet on the corresponding data link.

A LSR is capable of performing 3 operations: pop, push or swap. The label switch path (LSP) is a sequence of LSRs that switch a labeled packet through an MPLS network or just through a part of it, meaning that it is actually the path that packets take when forwarded through an MPLS network. A LSP begins with ingress LSR and ends with an egress LSR, both connected further to customer edge (CE) routers [2].

Figure 3 is an example of the operation of MPLS in a network, with labels specific to each pair of neighboring routers and the connections between them.

Label distribution protocol (LDP) is based on the following mechanism: every interior gateway protocol (IGP) IP prefix in the routing table has its specific local binding, meaning that the IPv4 prefix has a label attached to it. These bindings become remote bindings when they are split among the LSRs which become LDP neighbors. They store these remote and local bindings in a table called label information base (LIB). Usually, the remote bindings are more than one, because the LSR has more than one adjacent LSR [2].

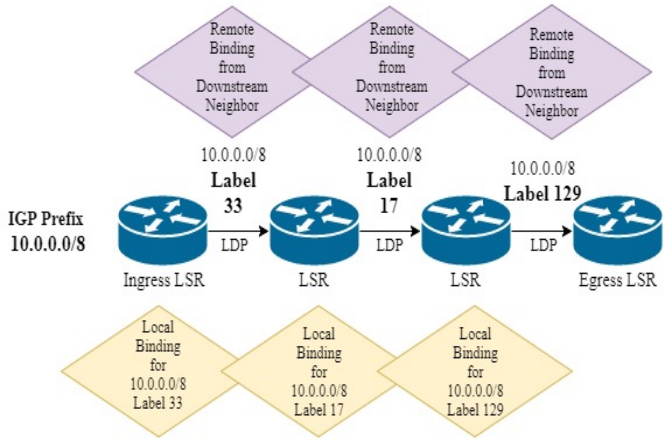

Fig. 3 – IPv4-over-MPLS network running LDP.

3. THE LOAD BALANCING MECHANISM

Traffic engineering (TE) is defined by all the techniques applied to a network in order for it to function properly and be optimized when necessary. The optimization is realized by redirecting the traffic to those lightly loaded paths such that the load among the paths to be balanced as per the diverse metrics calculated. The purpose of doing this is to avoid the congestion across the network and this fact leads to a competition between the Internet service providers (ISPs) in order to provide quality of service (QoS) [3].

There are two principal methods of balancing the traffic: per-destination and per-packet. By default, the load-sharing of type per-destination is configured on a device and it actually implies the hashing algorithm of the source and destination IP addresses. Also, it can be related to "per-flow" load-balancing, because it sends packets of the same destination on the same link, this way assuring an ordered stream at the final destination [4].

In order to do the load balancing for the labeled packets, the paths that are unlabeled (IP) are not considered, if there are also labeled paths for the same prefix. It happens because the possibility might exist to lose the traffic flow going over the unlabeled path. When considering MPLS implemented on an IPv4 network, the packets succeed in reaching the destination even without labels attached to them. At the links where MPLS is not enabled, the packets become unlabeled and reach their state of being labeled again at the next link where MPLS is implemented. When becoming unlabeled, there is an IP lookup in those packets and because there is IPv4 running everywhere, the delivery of the unlabeled packet to the destination without problems should successfully happen. In contrast, in MPLS VPN or AToM, the packet that is unlabeled at a certain moment does not reach the final destination [5].

3.1. LAYER 2 LOAD BALANCING

Load-balancing at Layer 2 uses EtherChannel, meaning that the traffic is balanced across all the links in the EtherChannel. It transforms the bits from the addresses in the frame into a digit used to choose one of the links in the EtherChannel. It is usually used to interconnect LAN switches, routers, servers and customers. EtherChannel includes fast EtherChannel, Gigabit EtherChannel, port channel and port group [6].

The distribution of the data load in the channel is done based on a hash. Port aggregation control protocol (PAgP) automatically builds the links in the EtherChannel, by sending specific packets that negotiate the characteristics of the channel. The constraints imposed by PAgP refer, generally, to the fact that all the configured ports must be in the same virtual LAN (VLAN) or defined as trunk ports, they must also run at the same speed and there are only the auto-desirable, desirable-desirable and on-on combinations of PAgP that permit the construction of a bundle [6].

Link aggregation control protocol (LACP) is an alternative to PAgP. The LACP packets are sent only between active-active or active-passive ports. LACP does not interoperate with PAgP, all ports that are implied in the process of building a channel being obliged to run one of them [6].

At layer 2, load-balancing is done on the basis of the source and destination media access control (MAC) addresses.

Basically, the EtherChannel load-balancing algorithm is based on the exclusive (ORXOR) operation between the source and destination addresses. The hashing process cannot be enabled to load balance the traffic among the ports in an EtherChannel. The hashing mechanism calculates a number in the interval 0–7 and based on this value, the specific port in the EtherChannel is selected. The most suitable manner to have a good load-balancing is to set an even number of links in EtherChannel, simultaneously controlling the throughput of the link, because if the physical ports are in a number equal to a power of 2 (2, 4, 8), the traffic is equally balanced. In the case of having a port-channel with 3 interfaces that are bundled, the load-balancing algorithm needs two bits to use from the hash. Additionally, these two bits result in four combinations (00, 01, 10, and 11) connected to the physical port. In consequence, the data flow is split unequally, not being in percentage of 33 % on each channel, but on 50 % on one and 25 % on the other two links [7].

3.2. LAYER 3 LOAD BALANCING

A router automatically learns its parallel paths to a certain destination using a standard routing protocol and balances the traffic over these routes according to the routing table. The load-balancing implies the distribution of the traffic across different links in a channel if considering layer 3 (L3) routing details.

3.2.1. Layer 3 ECMP

Equal-cost multi-path (ECMP) load-balancing represents the possibility to have links with their own IP address of the interface with the interior gateway protocol (IGP) configuration, this way obtaining the next-hop routes.

This type of load balancing can be made by having either L3 information or layer 4 (L4) information (with source and destination ports). In this manner, the hashing algorithm is processed and all the data from the same flow go on the same route. If there is the case of a flow which needs more resources, like bandwidth, than another flow, the way of making use of the path can be different and consequently, the data not be equally spread [4].

3.2.2. Link Bundling

Link bundling refers to the possibility of putting together a variety of links and transforming them into a single logical link. The objective is to improve the bidirectional bandwidth, redundancy and to acquire load balancing between two devices (for instance, routers). The bundle implies a virtual interface. Moreover, the components inside it can be dynamically added or deleted from it. This interface can have an IP address and other abilities, resulting in the fact that the entire data load sent to the entire bundle is further sent to one of the links inside of it [6]. EtherChannel is used to form bundles of Ethernet interfaces, having no mechanism of checking whether the links in the bundle are compatible or not. Load balancing can be enabled on all the members of the bundle. It can be done per-destination or per-packet.

Figure 4 presents the fact that the load balancing is based on the decision made at the line card (LC) ingress level. When sending the data flow to a certain LC, path or member, it is surely the one forwarding the traffic further. In the above figure, there are two paths: the first one is on LC2 and the second one is through the bundle, which has two members on two LCs [7].

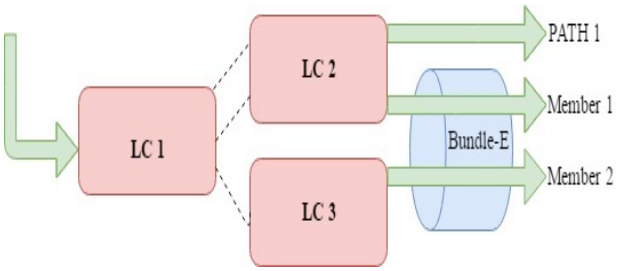

Fig.4 – Architecture of L3 load balancing.

The traffic is forwarded based on the decision of the ingress network processing unit (NPU) on the LC1, taking into consideration the hash computation. If it indicates PATH1 to transmit the data flow, then LC1 only sends it to LC2. If it shows the second path (the bundle-ether) to forward the traffic, then the link aggregation (LAG) selects just that specific member of the bundle and sends it to the NPU of the LC of that certain member that is going to transmit the traffic [7].

In addition, the multiple ports inside the link bundle behave as one link and present the great advantage that they might extent to various line cards to build only one interface, the damage of a link having no influence on the rest of them. A crucial benefit is constituted of the fact that the traffic is split over all available components of the link bundle. The members of the link bundle are essentially of the same type and have the same speed.

The load balancing of type “per-destination” allocates the packets on one of the members in the link bundle with the purpose of having the load balanced, taking into consideration the hashing algorithm, precisely the hash previously calculated using the source and the destination addresses together with user routing information. Consequently, all the traffic that has to go to a certain destination from a specific source goes on the same link [7].

3.2.3. Bundle in layer 2 and layer 3 scenarios

We can take into consideration the distinctive structures of the link bundle, depending on the layers at which it is implemented. Therefore, the hash is computed in a different manner. In the situation of having an IP address configured on the interface of the link bundle, ECMP load sharing can be applied. In the situation of having an attachment circuit (AC), it is necessary to perform L2 (Layer 2) load balancing, being based on the source and destination MAC (Media Access Control) addresses of the devices and on the router-IDs (identification). In the case of having two routers on each edge of the AC, the MAC addresses are not changeable, resulting in the fact that L3 load balancing might be implemented on the L2-based VPN structure [8].

Figure 5 underlines three use case scenarios. In the first case, there is a link bundle AC, having configured the transport of L2 and the VPN at L2. The second situation emphasizes having a pseudo-wire (PW) over the interface of the bundle-ether, case where the AC has no significant function, because it includes MPLS load sharing together with L2 VPN and IP addresses. The third case underlines

the presence of a bundle-ether, where the routing process is made by the LAG. Here the header of the bundle-ether is accompanied by the header of MPLS and both lead further by the IP header, resulting in the chance of implementing load sharing of L3 [8].

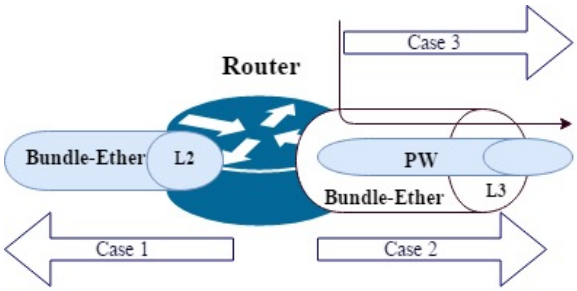

Fig.5 – L2/L3 load balancing use case scenarios [12].

When implementing L3 load-balancing, there are some special configurations and restrictions that must be taken into account. Otherwise, problems like the automatic disabling of the EtherChannel interfaces may appear, in order to avoid loops in the network. So, in the case of EtherChannel configuration, the following steps must be taken into account: there can be a maximum number of eight physical ports on a module and they have to support EtherChannels; all LAN ports must use the same protocol (PAgP, LACP), because there is no possibility to have two EtherChannel protocols in one EtherChannel; all the ports have to be of the same speed and in the same duplex mode; no LAN port must be shut down in an EtherChannel, in order to have no link failures and avoid the necessity to transfer the data flow to the other available port in the EtherChannel; no EtherChannel would be built if having one of the ports as a destination port; there must be L3 addresses assigned to the port channel logic interface and not to the physical ports [8].

3.2.4. MPLS VPN load balancing

The MPLS VPN is an implementation of MPLS that has recently gained a lot of popularity because of its important characteristics such as scalability and the ability to separate the network into smaller ones, a benefic aspect of a large enterprise network. MPLS VPN is used for transmission at L2 or L3 and the network structure is extremely important in the traffic load balancing [9].

The load balancing algorithms implemented at the VPN level might be diverse, depending on the way the links to the tunnels are organized. When considering L2 VPNs, the Ethernet frames are transmitted in a manner very similar to the transmission between two physical devices such as switches in different buildings [9]. In the case of L3 VPNs, the two ends of the constructed tunnel over VPN are situated on two different subnets, having different IP addresses and transporting IP packets across the VPN.

Furthermore, to load share the data, LACP can be used in order to add various links together or both ECMP and IGP configuration in order to have the routes associated to different next-hops and not to a single one. There is the possibility to take into consideration the bandwidth of the routes and reserve it for distinct purposes of the traffic flows, but also to not pay attention to the bandwidth [9].

4. IMPLEMENTATION OF A LAYER 3 IP/MPLS NETWORK

MPLS comes with an important benefit, namely the fact that it respects the label switching based on exact matching of the lookups, which is cheap, easy to use and noticeably decreases the load on the core routers. Other crucial aspects of MPLS are represented by the possibility of controlling the data flow through the network and of prioritizing a diversity of services while preventing congestion, by implementing TE [10].

4.1. MPLS VPNv4

We bring in multiprotocol-border gateway protocol (MP-BGP) and peer-to-peer VPN. These protocols imply the existence of a route reflector (RR) as well as VRF (VPN routing/forwarding). We created a scenario which outlines the likelihood of a L3 MPLS-based VPN to access the Internet. The practical usage is to support the VPN connectivity between corporate sites. VPNs serve as a method to share bandwidth between customers using an Internet service provider (ISP) network as a backbone. A VPN might also be seen as an association of sites that have a common routing table. The customer routers are connected by one or more interfaces to the service provider (SP) routers, that link, in turn, every interface to a VPN. The VRF is basically the routing table of a VPN [11].

More than one VPN can be enabled on a service provider router and the device can effectively make the difference between their links. Also, the router is the point where the bonding between the various VRFs and the specific L3 interfaces is made, because a L3 interface must correspond to a single VRF. The locally known routes are advertised by the CE routers to the provider edge (PE) routers and in turn they find out the remote routes that take part in the VPN process from the PEs. The VPN paths for the SP are not all included into the configuration of the PEs, those that are directly linked to it being an exception. internal border gateway protocol (iBGP) announces between PEs the VPNs that are advertised by CEs to PEs, taking into account that PEs have each VRF corresponding to each connected site. In fact, we use multiprotocol-border gateway protocol (MP-BGP), which enables IPv4 unicast addresses and transmits the VPN labels into the MPLS VPNs [12].

VPNs of version 4 can be considered, taking into account that the CE routers are announcing their routes to PE routers using dynamic routing protocols, such as OSPF (open shortest path first). The PEs in turn connect to CEs by two VPNs (customer1 and customer2), enabling the address-family IPv4 unicast VRFs. Consequently, every VPN existing between them has its own VRF that additionally has its paths inside of it. The two PEs have a VPNv4 connection between them compared to the provider (P) routers that are placed in the central network and run MPLS. So PEs announce to each other the links included in the VRFs, paths that are actually VPNv4 routes with route targets (RT), RD (route distinguisher) and a VPN label. This way the MPLS L3 VPNs are built. IPv4 routing represents the base of the two customers' communication, whereas inside the core network, iBGP is used to advertise the collected paths from the VPNs between PEs [12].

Concisely, the IPv4 address-family is utilized to reveal the simple IPv4 addresses and the VPNv4 address-family puts the 64 bits of the RD into them, in this manner forming exclusive VRFs. The RT describes the structure of the network of the VPNv4 and the label of the VPN, which makes the PEs able to know each VPN marked by the data load flowing into them [12].

4.2. ROUTE REFLECTOR

As long as routing loops in a network are concerned, BGP is a powerful tool to warn them, based on the principle that the routes learned from an iBGP peer are not announced to the other iBGP neighbors. This method requires a logical full mesh topology in order to permit the transmission of the traffic in the entire network, but this is not a solution for the scalability of it, particularly in extensive structures. Consequently, a convenient idea is that of the use of a route reflector in the process of BGP [13].

A BGP route reflector is simultaneously the point of concentration (central point/server) in the network and the router that advertises the paths received from an MP-BGP peer to other MP-BGP peers, respecting certain constraints. This way, it ensures the extensibility of the network, making it possible for the BGP routers in the topology to connect only to the RR and not necessarily to all the other routers in a full mesh model. The routers peered with the route reflector are called route reflector clients and they individually need to be configured as such. Moreover, the RR represents a single point of failure, meaning that in case of damage, the whole circuit of routers is interrupted from working. So, in order to provide reliability as well as redundancy in the network, there is recommended to implement a second RR [13].

Another significant role of the RR is underlined by the fact that it avoids the necessity to implement specific supplementary commands on all the existent PE routers when adding new PEs in the service provider system. However, this addition of a PE demands a neighbor announcement on it pointing to the RR and a neighbor announcement referring to the new PE on the RR router [13].

A route reflector can have two kinds of peers: route reflector clients and non-clients. The routes from a client are transmitted to clients, non-clients and external BGP (eBGP) associates, whereas the routes from a non-client are revealed only to clients and eBGP peers. The RR and its clients define a cluster, which is characterized by the cluster-id inserted to each path made known by the RR to the clients and non-clients [13].

4.3. LAYER 3 VPNV4 TRAFFIC BALANCING MECHANISM-PROPOSAL

We created a L3 VPNv4 MPLS structure with the purpose of indicating the crucial benefits of MPLS as well as of VPNs of version 4. This first scenario is formed from eight routers with various functions in the network. We utilized UNetLab (Unified Networking Lab), a platform that allows the simulation, construction, examination and correction of both real and potentially built networks [12].

We enabled two VPNs on the two interfaces between R5 (Router 5) and R6 (Router 6), correspondingly router 7 (R7) and router 8 (R8). This way, we deal with one VRF assigned to customer1 on Ethernet0/0 and one VRF associated to customer2 on Ethernet0/1. The RD is the object making the difference between the two VRFs, which consists of the autonomous system (AS) and a special digit for each of them. The two PEs from the topology are found in the same AS (BGP depends on the AS).

As long as OSPF is concerned, it upholds multiple VRFs, meaning that it can block the possible creation of loops in the network, because it reorganizes the paths between BGP and OSPF. Having this characteristic, the router can be seen as multiple diverse virtual routers that have their own routing tables together with VRF tables for the VPNs [14].

The two implemented customers are found in two areas of OSPF, because this aspect gives the possibility to the different customers to have their routes completely separated. In this manner, the paths can be seen as VPNs, which is a also a convenient condition for MPLS [14].

Figure 6 is identified with the logical diagram of the first scenario (created in UNetLab). As is can be seen, it is formed from a core area, with OSPF area 0 set up on the routers and intermediate system-to-intermediate system (IS-IS), which helps with the examination of the dynamic routing protocols. In addition, MPLS is used to highlight the paths of customer 1 and customer 2 (by the transmission of labeled packets) and the extremely important role of the RR, which uses MP-BGP to deliver the traffic flowing through these two VPNs to the specific clients [12].

As long as the core area is concerned, all the paths between the central routers are dynamically routed by OSPF and IS-IS, which have the administrative distance (AD) different. This AD makes the difference between the two protocols, because as this number is smaller, the erefore, OSPF represents the one elected in our case, because its AD is equal to 110, whereas IS-IS has the AD exactly 115 [12].

We enabled further the fundamental characteristics of MPLS, this way making an analysis of the transmission of the data load through the routes in the network. The transportation of the traffic is strongly dependent on the labels that the router found at the starting point of the network insert on the header of the packets.

Figure 7 shows the design of the eight routers used to examine the load balance of the traffic and they have the following functions:

– R1, R2, R3, R4 - P (provider) routers
– R5, R7 - PE (provider edge) routers
– R6, R8 - CE (customer edge) routers.

R1, R2, R3, R4, R5 and R7 (the routers in the middle) have MPLS capacities enabled, while R6 and R8 operate by using only IP addresses. CEs link straight to PEs and they do not know about the VPNs, because only the PEs deal with the forwarding of the MPLS packets. P routers serve as the backbone routers in the SP structure [12].

The PEs are fed with the routing information by R6 and R8 that make use of OSPF, which supports multi-instance. Between the ingress LSR and the egress one, meaning R5 and R7, the routing data is learned through MP-BGP and the use of MPLS labels as shown in Fig. 7.

R6 is recognized as CE1 (Site 1) and we might acknowledge that if it sends an IPv4 packet with the destination address 62.0.0.8 (loopback0 address of R8), which is recognized as CE2 (site 2), R6 performs an IP lookup in the routing table and sends further the traffic to R5 (PE1) in the suitable way. R5, which is the ingress LSR in the topology, inserts the MPLS label into the packet and

makes the decision based on which the next-hop BGP router would be (PE2) [12].

After all, R5 sends the MPLS packets to the first P in the topology (R3 or R4) and the packets then follow the calculated route to the final target point.

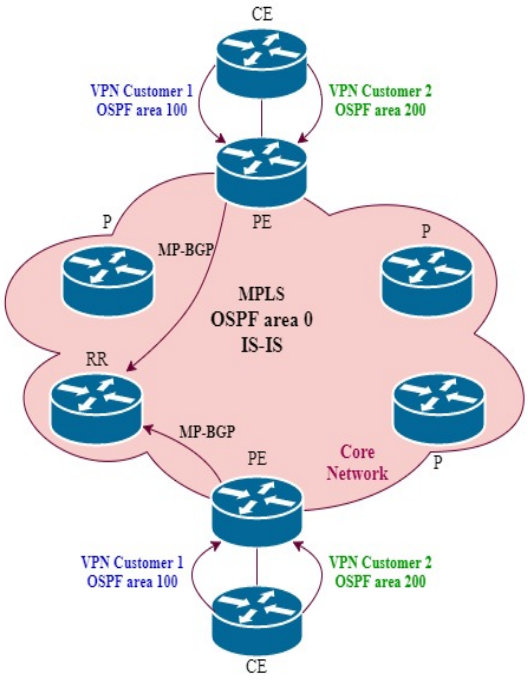

Fig. 6 – The logical scheme of a L3 MPLS VPNv4-based topology [12].

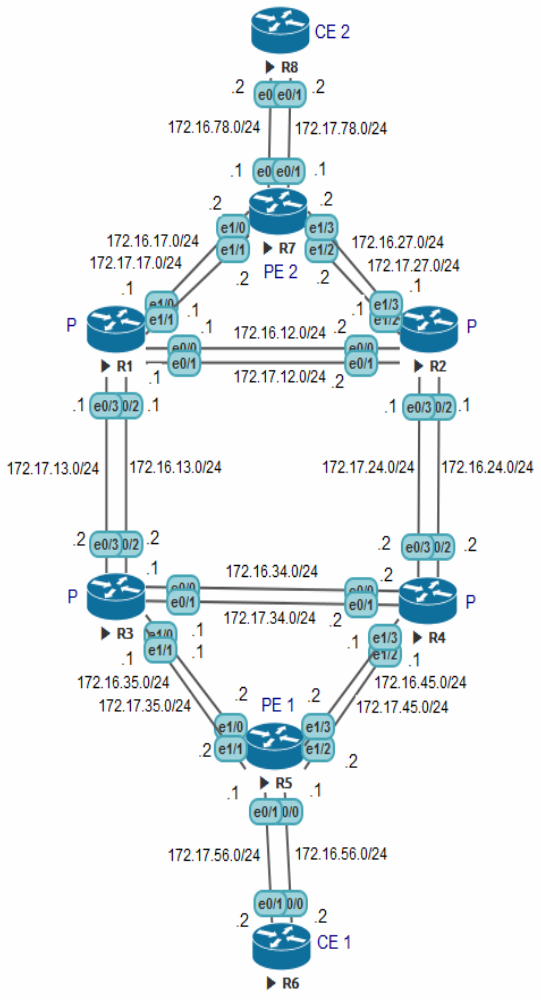

Fig. 7 – L3 MPLS VPNv4-based practical topology [12].

R5 makes the load sharing of the data on these two links (the two routes going to PE2). Each P router on a certain LSP sends the traffic to the penultimate-hop router, R1 or R2, making use of the labels. The function of R1 and R2 is to pick up the interface that leads to PE2 and to pop the MPLS label, process that is called penultimate hop popping (PHP). This process leads further to the transmission of the packets to the final point in the network. CE2 receives the IPv4 packets from PE2 that makes an IP lookup in the routing table [15].

Taking into consideration the fact that the two enabled VPNs need two areas of OPSF, R6 and R8 have two loopback addresses (loopback0 and loopback1), whereas all the other routing devices in the entire network have only loopback0 enabled. We set up OSPF and IS-IS in the network and on the PEs, as previously mentioned, we implemented BGP and MPLS, including LDP. BGP makes use of the loopback addresses on R5 and R7. These addresses are important, because of the fact that they are 'alive' up to the moment they are actually disabled.

When taking a comprehensive look on the whole network, we notice the fact that the PEs (R5, R7) have their roles fulfilled only by MP-BGP, being in contrast with the functions of all the P routers (R1, R2, R3, R4), that accomplish them by use of all the other protocols, except for MP-BGP. CE1 and CE2 have the two VPNs enabled on each of them, involving those specific VRFs on PE1 and PE2.

The command "show mpls forwarding-table 'IP address' detail" has the output on R1 presented in Fig. 8. It shows the label assigned by R1, the label attached by the next-hop router, the final IP address that the packets with this label are sent to and the interface used to forward these packets. In addition, the IP address of the next-hop router is shown, together with the "MAC/Encaps" field, which explains the number of bytes of the L2 header (Ethernet II)/ the number of bytes of the encapsulated packet: L2 header and the label header.

Figure 9 underlines the fact that the L2 Header contains 14 bytes, while the MPLS Header is made of 4 bytes, which explains the 18 bytes of the encapsulation of the packet. The "MRU" represents the maximum received unit of the labeled packet and the "label stack" underlines the number of labels in the stack of the forwarded packet. The "per-destination load-sharing" field shows that, by default, the packets are forwarded taking into account the destination address.

R1

```
R1#sh mpls forwarding-table 62.0.0.8 detail
Local      Outgoing   Prefix           Bytes Label   Outgoing    Next Hop
Label      Label      or Tunnel Id     Switched      interface
36         36         62.0.0.8/32      0             Et1/0       172.16.17.2
       MAC/Encaps=14/18, MRU=1500, Label Stack{36}
       AABBCC000701AABBCC0001018847 00024000
       No output feature configured
   Per-destination load-sharing, slots: 0
           36         62.0.0.8/32      590           Et1/1       172.17.17.2
       MAC/Encaps=14/18, MRU=1500, Label Stack{36}
       AABBCC000711AABBCC0001118847 00024000
       No output feature configured
   Per-destination load-sharing, slots: 1
```

Fig. 8 – Traffic analysis-MPLS forwarding-table in detail command.

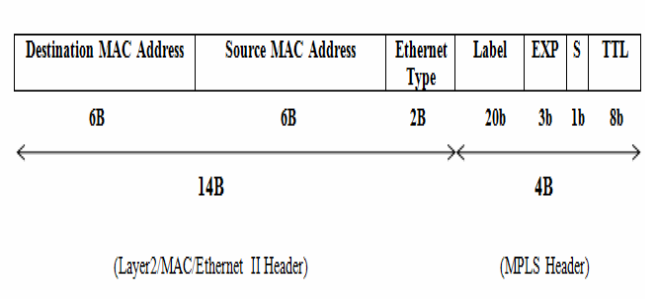

Fig. 9 – MAC/Encapsulation header.

Furthermore, it implies that, in the output from the above, where we have two routes to the same router, all the packets to the first destination, use the first path and all the packets directed to the second destination, choose the second route. Even if, by this method, there is kept the order of the packets, the links might be underutilized, because if one path is occupied with the majority of the traffic, the other ones are left with available unused bandwidth. Anyway, for a more equally balanced data flow, there should be more destination addresses.

VPNs are enabled in order to provide separate routes for specific applications. This way, if, for instance, R6 receives traffic from a VPN, it executes an IP lookup conditioned by the input interface and, in consequence, it sends the packets to R5. This router then looks into the exact VRF, discovers a route and depending on it, R5 imposes an MPLS label in the header of the packets. After that, it sends it to one of the P routers. No matter the path, when arriving to R7, the packets have already lost their MPLS label. In turn, R7 uses the label to find the VPN routing table and also makes an IP lookup to find the route to R8, which, when being contacted, makes an IP lookup further to obtain the final destination, the VPN [12].

When considering that here we have two MP-BGP peered PEs, the number of the needed peering is calculated by the formula $n(n–1)/2$ and it equals $2(2–1)/2 = 1$, respecting the full mesh status of the network. Moreover, we made R3 a RR, which is directly connected to R5 and R7, by their loopback interfaces. It sends control and data plane addressing only for the VPN customers.

Figure 10 helps us evaluate the data load passing through the network and we can select the next path: from the CE2 router, when seeking the path to PE1 or R6 CE1, we notice that the load-balancing is established depending on the destination address. It represents a model of tracing the route into Customer 2 VRF, noticing the expired labels at each hop router, the label inserted in order to distinguish the VRF and the PHP process that takes part at the level of R4. Consequently, a VPN uses two labels: one for the identification of the distinctive VRF and one (on top of the stack) for the identification of the MPLS network (vanishes at each next-hop P router in the SP domain) [12].

Finally, in this situation, when having two VPNs on different interfaces, we can observe the fact that the traffic is divided on both branches of the network. In this manner, the load-balancing is dependent on the per-destination hashing algorithm: for each and only final target, there is a certain path that is selected, based on the inserted labels, on the IP addresses as well as L4 port numbers [12].

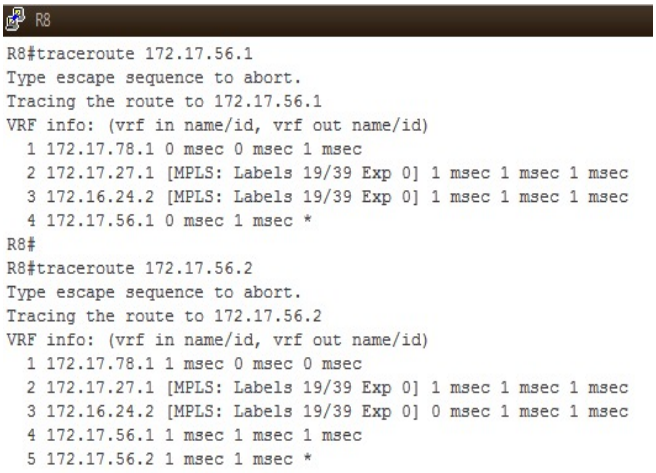

Fig.10 – Traffic analysis-traceroute command [12].

4.4. LAYER 3 ETHERCHANNEL IMPLEMENTATION- PROPOSAL

Another algorithm of balancing the traffic is to create a virtual group of physical ports, symbolized by a single logical interface, with the evident purpose of having link redundancy when some damages appear (the bundle automatically redistributes the load on the existent accessible members) and also having advantageous sharing of the traffic on the combined links. Another beneficial advantage is the increase of the bandwidth available in the network, when the data flow is split all over the paths in the bundle [16, 17].

The EtherChannel is constructed with independent Ethernet links and is represented by one MAC address, one IP address and one set of network layer configurations. However, the configurations regarding the physical and link layers are done at the level of each member of the bundle. The most valuable aspect is the load balancing of the traffic within the members of the created bundle, because it is realized taking into account the entire flow and not the individual packets. The information is allocated to every path depending on their bandwidth related to the whole EtherChannel [16, 17].

We created, in another network simulation program – Packet Tracer, a network with four switches: two L3 switches and two L2 switches, connected to some end terminals, as it can be seen in Fig. 11. The innovative thing is that we created between the two L3 switches, an EtherChannel, putting the two physical ports into one logical interface, having an IP address and some other characteristics, with the purpose of increasing the bandwidth and having the possibility to redistribute the traffic in the case of failure of one channel member [16, 17].

The EtherChannel using the two links between SW1 (switch1) and SW2 (switch2) was created by putting the two interfaces in the same channel group (channel-group 1). After that, we enabled the load balancing of the traffic on the port-channel interface and there are many options, but we chose to do this using the source and destination IP addresses and eventually the source and destination MAC addresses for the data flow of non-IP type. The more the components used in the hashing process of splitting the traffic on the members of the EtherChannel, the better the distribution of the packets across the network [16, 17].

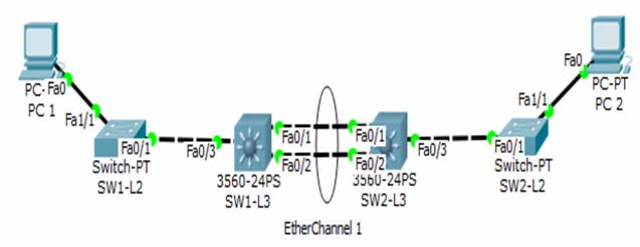

Fig. 11 – Layer 3 EtherChannel-based practical topology.

The two interfaces used are FastEthernet0/1-0/2. We wanted to use these two links effectively in the process of balancing the traffic that goes from one LAN to another. It is in fact the hashing algorithm done in order to do the load balancing on the links.

The details regarding the EtherChannel are seen in Fig. 12 when enabling the "show etherchannel summary" command. It analyzes the status of the ports in the channel, these being of type "P" and meaning "in port-channel". Also, the protocol used for the connection between the two pairs of interfaces is LACP, the port-channel being in the "RU" state: enabled at L3 and being in use.

Another useful command is "show etherchannel load-balance", shown in Fig. 13. It underlines the fact that the manually configured method of distributing the traffic across the links in the port-channel is based on the source and destination IP address for the packets of L3 (IPv4 or IPv6) and on the source and destination MAC addresses for the frames encapsulated at L2.

SW1-L3

Physical | Config | CLI

```
SW1-L3#
SW1-L3#sh etherchannel summary
Flags:  D - down        P - in port-channel
        I - stand-alone s - suspended
        H - Hot-standby (LACP only)
        R - Layer3      S - Layer2
        U - in use      f - failed to allocate aggregator
        u - unsuitable for bundling
        w - waiting to be aggregated
        d - default port


Number of channel-groups in use: 1
Number of aggregators:           1

Group  Port-channel  Protocol    Ports
------+-------------+-----------+-----------------------

1      Po1(RU)           LACP   Fa0/1(P) Fa0/2(P)
```

Fig.12 – EtherChannel analysis-Etherchannel summary command.

SW1-L3

Physical | Config | CLI

```
SW1-L3#
SW1-L3#sh etherchannel load-balance
EtherChannel Load-Balancing Configuration:
        src-dst-ip

EtherChannel Load-Balancing Addresses Used Per-Protocol:
Non-IP: Source XOR Destination MAC address
  IPv4: Source XOR Destination IP address
  IPv6: Source XOR Destination IP address
```

Fig. 13 – Etherchannel analysis-EtherChannel load balance command.

5. CONCLUSIONS

MPLS could be considered as the most appropriate protocol used in a sophisticated backbone network. We implemented a scenario based on MPLS and on L3 VPNs, therefore making an analysis of the load-balancing of the traffic flows across specific routes. The survey also highlighted OSPF and IS-IS (dynamic routing) as well as MP-BGP. The second case was the one configuring L3 switches having between them port-channels represented by a logical interface, a MAC address and specific MPLS labels, EtherChannels that ensured higher speed between the devices and simultaneous redundancy of the routes.

One of the greatest problems nowadays is concerning the increase of the data load on the links. In this situation, the implementation of EtherChannels/bundles of physical ports was beneficial in improving the bandwidth allocation and the speed of information transportation. However, taking into account the multitude of the links that have to do load-sharing of the increased traffic based on the calculated hash, it is not an ideal load balance. The hashing algorithm was not excellent and it would be necessary to include more parameters when performing the hashing algorithm for each transported service/application.

As a consequence, there would be necessary to have more entropies/resolution criteria for the calculus of the hash value when splitting the traffic flows, with the purpose of obtaining an optimal hashing method. In a future study, there will be some other scenarios emphasizing the advantages of implementing traffic engineering for the end-to-end load-balancing in an IP/MPLS network, aiming to ensure a much better sharing of the traffic between applications.

Received on March 4, 2018

REFERENCES

1. Luc De Ghein, *MPLS Fundamentals*, CiscoPress, 2006.
2. Juniper, *Evolving Backbone Networks with an MPLS Supercore*, 2015.
3. Cisco, *MPLS Traffic Engineering*, CiscoPress, 1999.
4. Ravindra Kumar Singh, *Load Balancing in IP/MPLS Networks: A Survey*, 2012.
5. Azhar Sayeed, *MPLS and Next-Generation Networks: Foundations for NGN and Enterprise Virtualization*, CiscoPress, 2006.
6. Cisco, *Catalyst 6500 Series Switch Cisco IOS Software Configuration Guide*, Release 12.1E, CiscoPress, 2015.
7. Cisco, *Understanding EtherChannel Load Balancing and Redundancy on Catalyst Switches*, CiscoPress, 2007.
8. Cisco, *MPLS Label Distribution Protocol (LDP)*, CiscoPress, 2002.
9. Cisco, *Multiprotocol BGP Extensions for IP Multicast Commands*, CiscoPress, 2013.
10. R. A. Steenbergen, *MPLS for Dummies*, nLayer Communications.
11. Cisco, *Internet Access from an MPLS VPN Using a Global Routing Table*, CiscoPress, 2005.
12. Ema-Maria Galeş, V. Croitoru, M. Iordache, *Layer 3 VPNv4 Traffic Balancing Mechanism Scenario,* COMM 2018, in reviewing.
13. Juniper, *Understanding BGP Route Reflectors*, 2017.
14. Cisco, *Understanding Redistribution of OSPF Routes into BGP,* CiscoPress, 2012.
15. Cisco, *Cisco IOS XR MPLS Command Reference for the Cisco CRS Router, Release 4.2.x*, CiscoPress, 2016.
16. Cisco, *Cisco ASR 9000 Series Aggregation Services Router L2VPN and Ethernet Services Configuration Guide,* Release 4.3.x, CiscoPress, 2016.
17. Ema-Maria Galeş, *Algorithms and Methods of Balancing the Traffic in IP/MPLS Networks*, PhD thesis, „Politehnica" University of Bucharest, 2017.