Traffic Engineering (TE) & Multiprotocol Label Switching (MPLS)

Redes de Comunicações II

Licenciatura em Engenharia de Computadores e Informática

Prof. Amaro de Sousa (asou@ua.pt)

DETI-UA, 2024/2025

Traffic supported by a Network

- A **traffic flow** from a sender host A to a receiver host B is the sequence of IP packets sent by host A to host B during time
  - The **throughput** is the rate, in bits/second, of the traffic flow

---

- Example: host A downloads a file of 5 GBytes file from host B through TCP at an average download rate of 5000 IP packets/second.
  - Average download throughput of the file download:

    Assuming that each IP download packet has 20 Bytes (IP header) + 20 Bytes (TCP header) + 1460 Bytes (Data) = 1500 × 8 bits, then, the download throughput is 1500 × 8 × 5000 = 60000000 bps = 60 Mbps (Megabits/sec)
  - Average upload throughput of the file download:

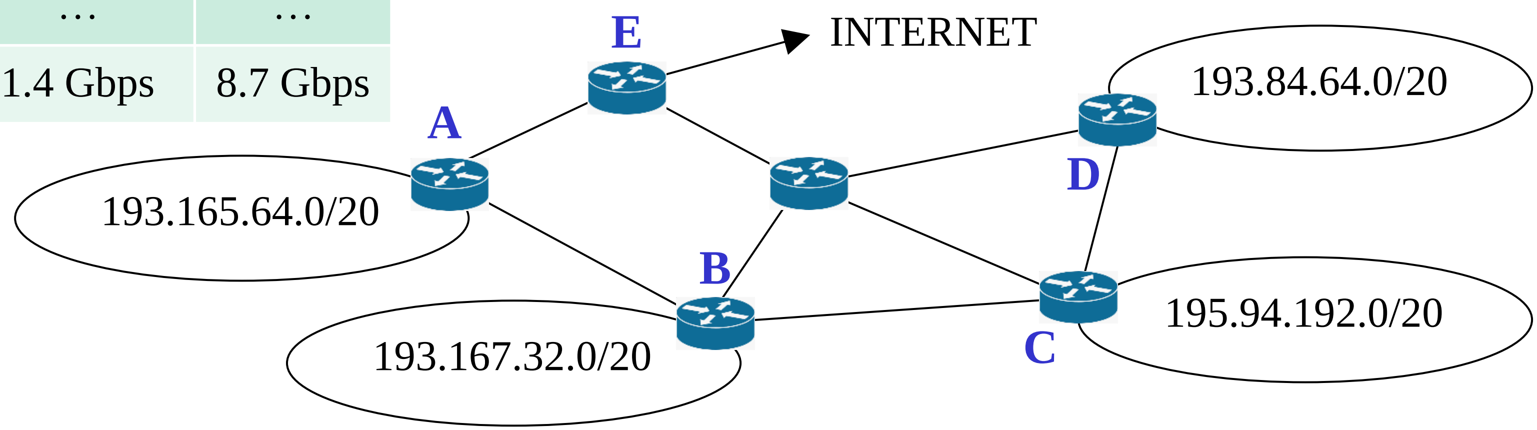
    Assuming that each 2 download TCP segments generate one upload TCP segment (to acknowledge the received data) in an IP packet with 20 Bytes (IP header) + 20 Bytes (TCP header) = 40 × 8 bits, then, the upload throughput is 40 × 8 × 2500 = 800000 bps = 800 Kbps (Kilobits/sec)
  - File download time:

    The file size 5 GBytes is segmented in = $5 \times 1024^3 / 1460 \approx 3677$ TCP segments. So, the download time is around 3677 / 5000 = 0.74 seconds

Traffic supported by a Network

- An **aggregated flow** from one network element A to another element B is the set of traffic flows from A to B.
  - The **throughput** of an aggregated flow is the total rate, in bits/second, of all its traffic flows
- The **traffic throughput of a service** supported by a network is defined (in its simplest version) by the average throughput over time of the aggregated flows between all node pairs running the service.

N1	N2	N1→N2	N2→N1
A	B	2.5 Gbps	3.1 Gbps
A	C	1.3 Gbps	1.7 Gbps
…	…	…	…
D	E	1.4 Gbps	8.7 Gbps



Network Dimensioning vs. Traffic Engineering

- **Network Dimensioning** is the process of determining the required capacity on each element of the network to be able to carry the traffic throughput of all supported services
  - Network elements are links, routers, switches, etc…
  - Traffic throughputs are hard to predict!
- **Traffic Engineering** is the process selecting the routing paths to best fit the traffic of all supported services in the capacity resources of the network
  - Ensure the most desirable/appropriate routing paths for each aggregated flow of each service.
  - The aim is to use efficiently the capacity resources of the network.
    - For example, to prevent over-utilized (congested) links whilst other links are under-utilized.

Life Cycle of a Network

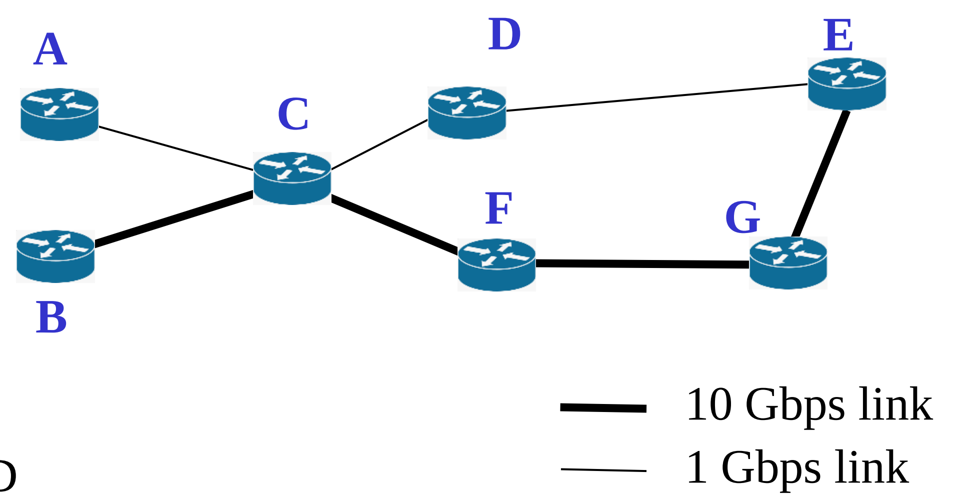
- A network life cycle is composed by two management phases:
  - A network dimensioning phase
    - The operator estimates the traffic throughput of the services the network must support in a given future time (for example, in the next year).
    - Then, it determines the required capacity of the network elements to carry the estimated traffic throughputs.
      - The aim is to minimize the investment cost.
    - Finally, the operator upgrades the network.
  - A traffic engineering phase
    - The operator changes the routing of the network in accordance with the traffic changes of the supported services.
    - This phase continues while network resources are enough to support all services
      - An efficient traffic engineering allows this phase to last for a longer time (minimizing the need for a new investment cost).
    - A new network dimensioning phase is required when the network capacity resources are no longer enough to support all services due to:
      - (i) throughput grow of the existing aggregated flows, (ii) new aggregated flows of the initial services and/or (iii) the introduction of new services.

Traffic Engineering with IGP Routing Protocols

- Traffic engineering is not ideal when using standard IGP protocols (as RIP or OSPF)

- With RIP, all flows are routed in C to next-hop D
  - There is a huge overflow (i.e., a drop) of 10 Gbps in link C-D
- With OSPF, and assigning appropriate OSPF port costs:
  - All flows can be routed in C through next-hop F
    - overflow of 1 Gbps in link C-F
  - All flows cab be routed in C through next-hops D and C (with ECMP)
    - overflow of 4.5 Gbps in link C-D

N1	N2	N1→N2	N2→N1
A	E	1.0 Gbps	-
B	E	5.0 Gbps	-
C	E	5.0 Gbps	-



Traffic Engineering with IGP Routing Protocols

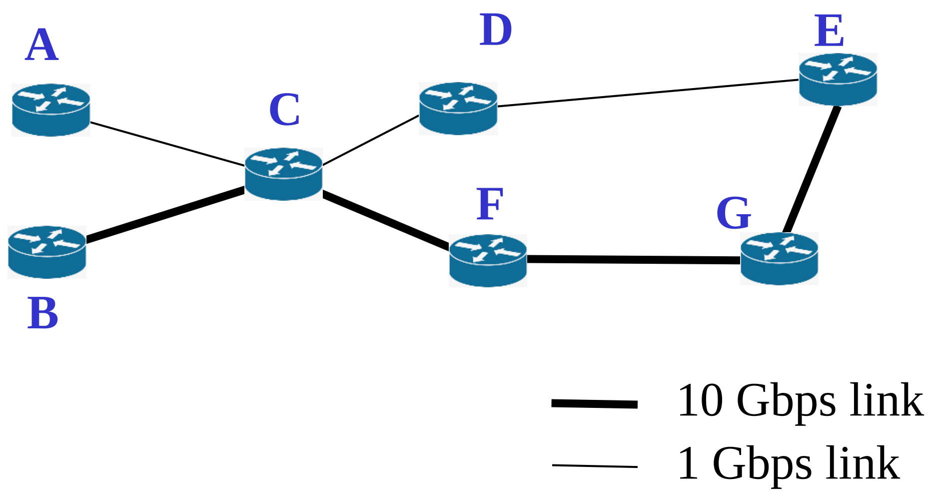
- In order to fit the traffic flow throughputs in the capacity resources of the network, we need to route each traffic flow through:

A – E: A → C → D → E

B – E: B → C → F → G → E

C – E: C → F → G → E

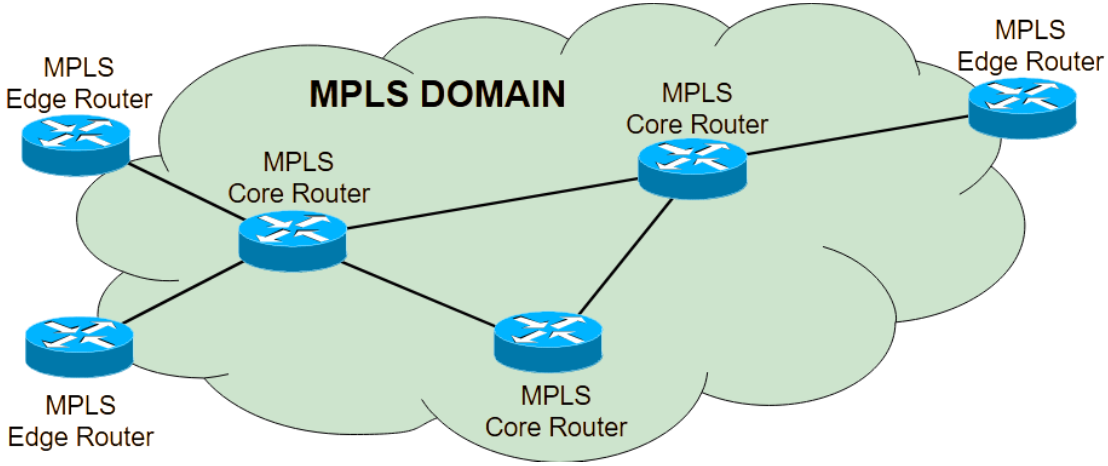
N1	N2	N1→N2	N2→N1
A	E	1.0 Gbps	-
B	E	5.0 Gbps	-
C	E	5.0 Gbps	-



- Note that:
  - this solution requires C to route flows based not only on the destination but also on the source
  - IGP protocols (as RIP or OSPF) are destination-based routing protocols

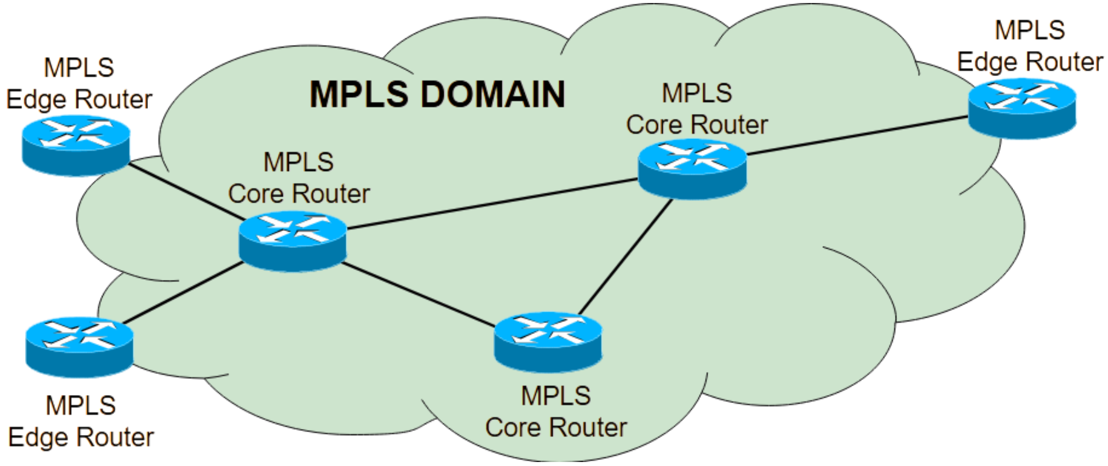
Multiprotocol Label Switching (MPLS)

- MPLS is a routing technique that forwards packets from one router to the next router based on labels rather than on network addresses.
  - While network addresses identify endpoints, labels identify established routing paths between endpoints.
  - MPLS can encapsulate packets of different protocols (IPv4 and IPv6), hence the 'multiprotocol' component of the name.
- In an **MPLS domain**, there are **Edge Routers** and **Core Routers**.

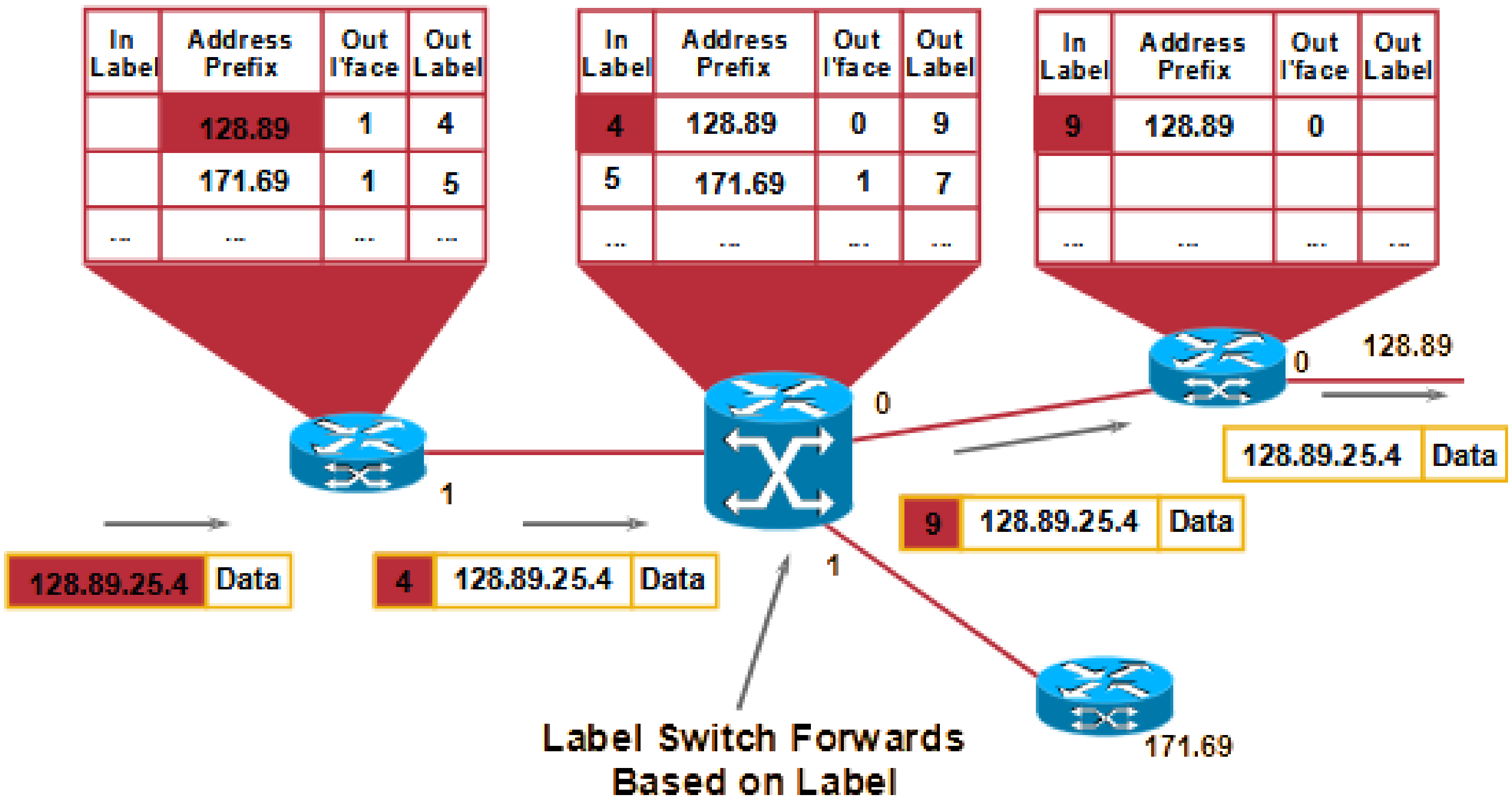


Multiprotocol Label Switching (MPLS)

- Packets are labelled at the entry Edge Router with a label of the first hop.
- Each Core Router forwards the packet based on its incoming label and swaps the label to the one of the next hop.
- The exit Edge Router removes the label.
- Advantages:
  - Simplification of the packet routing process on routers.
  - Full traffic engineering capability.
  - Simplification of the network management (a single protocol layer).

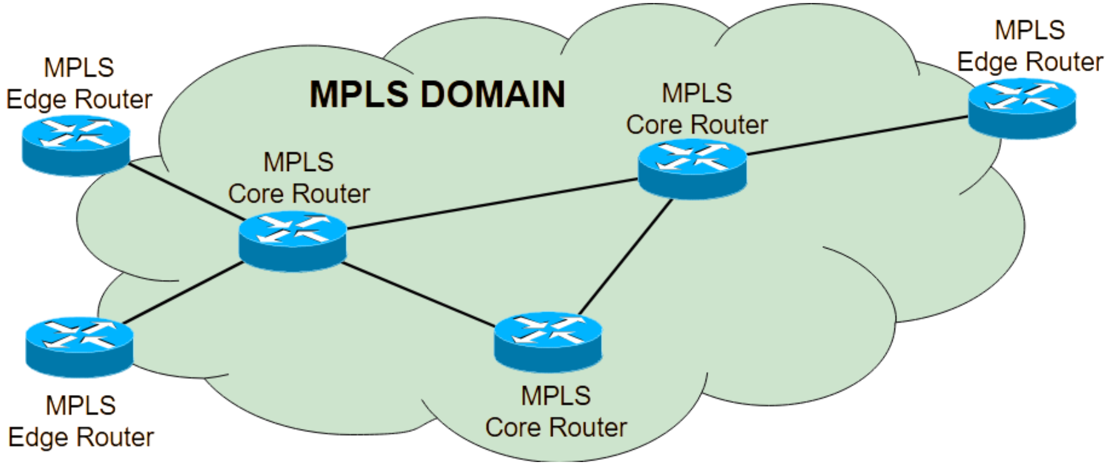


MPLS Switching Illustration



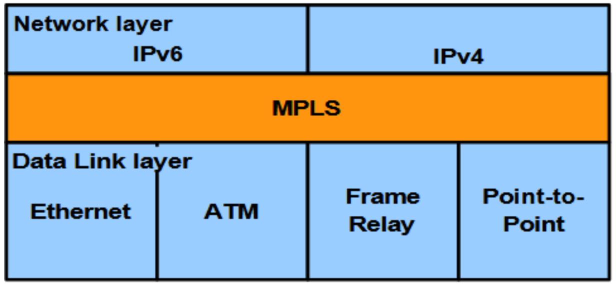
MPLS Fundamentals

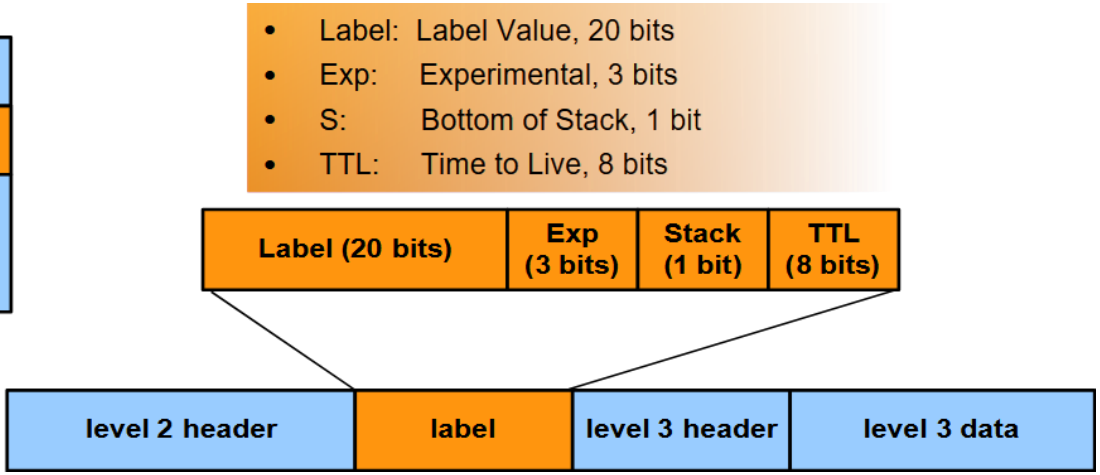
- As a packet enters an MPLS domain, it is assigned a label based on its **Forwarding Equivalence Class (FEC)** defined at the entry **Edge Router**.
- FECs are groups of packet flows forwarded through the same **Label Switched Path (LSP)** by **Label Switching Routers (LSR)**.
  - Need a mechanism to assign and distribute labels to establish LSP paths.
- In **Core Routers**, routing is based on label-swapping and forwarding.
- MPLS works in two planes:
  - Control Plane - responsible for setup label swapping tables among LSRs.
  - Forwarding Plane - uses the label carried by each incoming packet and the label swapping table of the LSR to forward the packet.



MPLS Labels

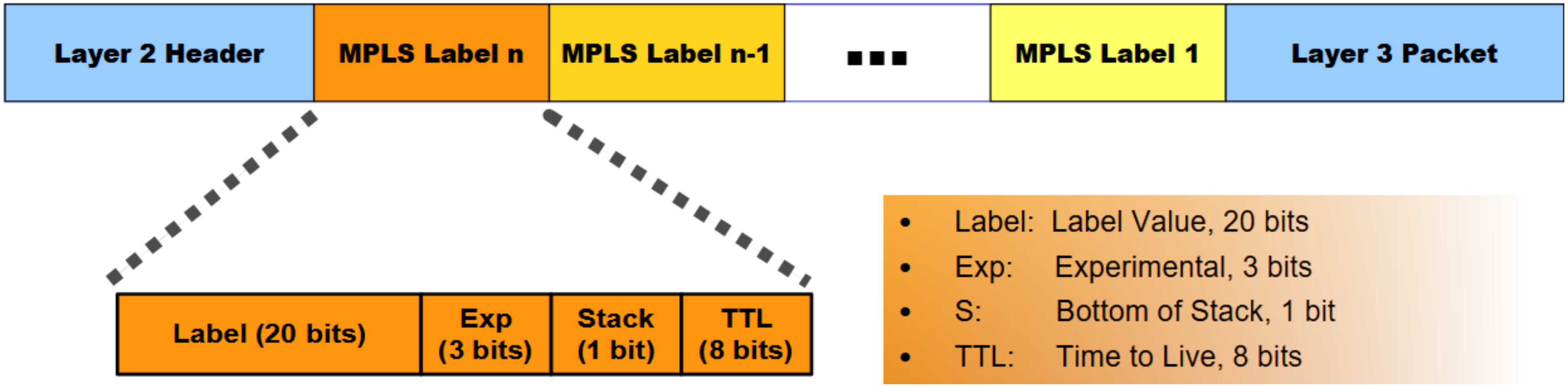
- Label is a 20-bit field that carries the value of the Label.
  – TTL field is IP independent – Similar purpose.
- On some Data Link (Layer 2) old technologies, label is given by the appropriate fields of their header.
  – ATM : VPI (Virtual Path ID) and VCI (Virtual Channel ID) fields.
  – Frame Relay: DLCI (Data Link Connection Identifier) field.
- On other Data Link technologies (Point-to-Point, Ethernet), the label is inserted between Layer 2 and Layer 3 headers.





MPLS Label Stacking

- Labels can be used in stacks to support more efficiently multiple services:
  - Outer (last) label (first received) is used to switch the packets between Edge Routers in the MPLS core.
  - Inner labels are used to designate services, FECs, etc.
- S (Bottom of Stack) bit is set to one for the last entry in the label stack (i.e., for the bottom of the stack), and zero for all other labels.

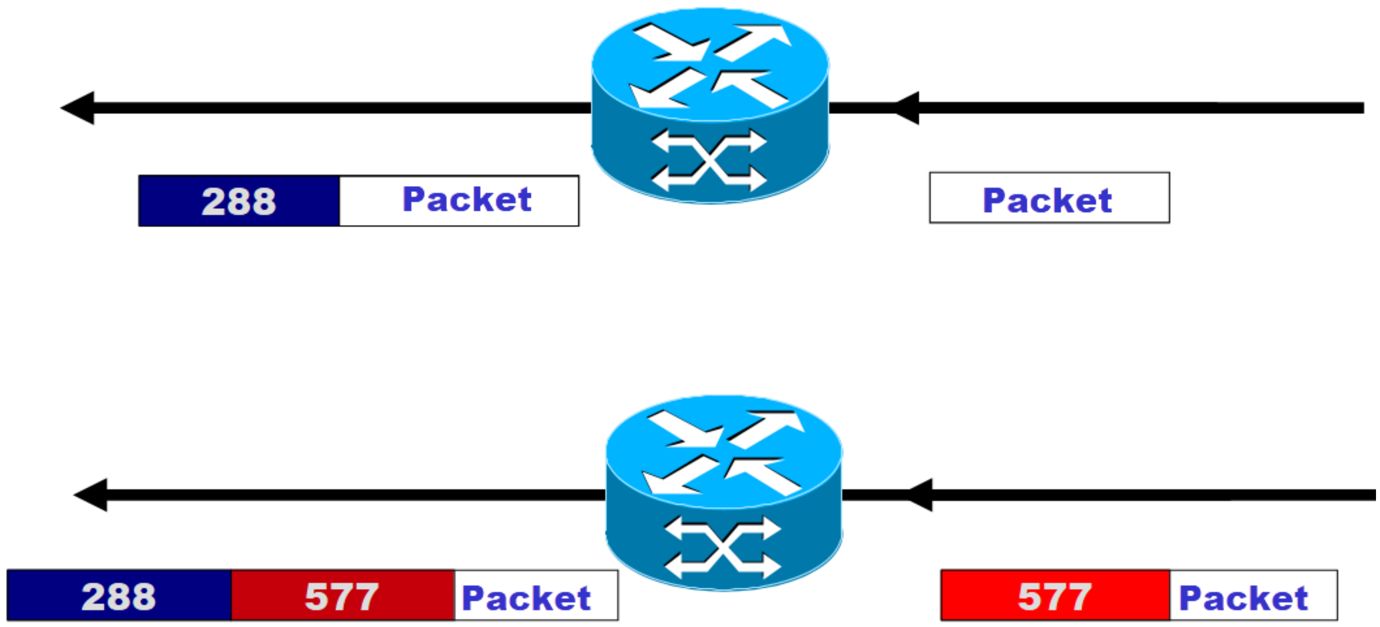


- Label: Label Value, 20 bits
- Exp: Experimental, 3 bits
- S: Bottom of Stack, 1 bit
- TTL: Time to Live, 8 bits

MPLS Forwarding via Label Swapping

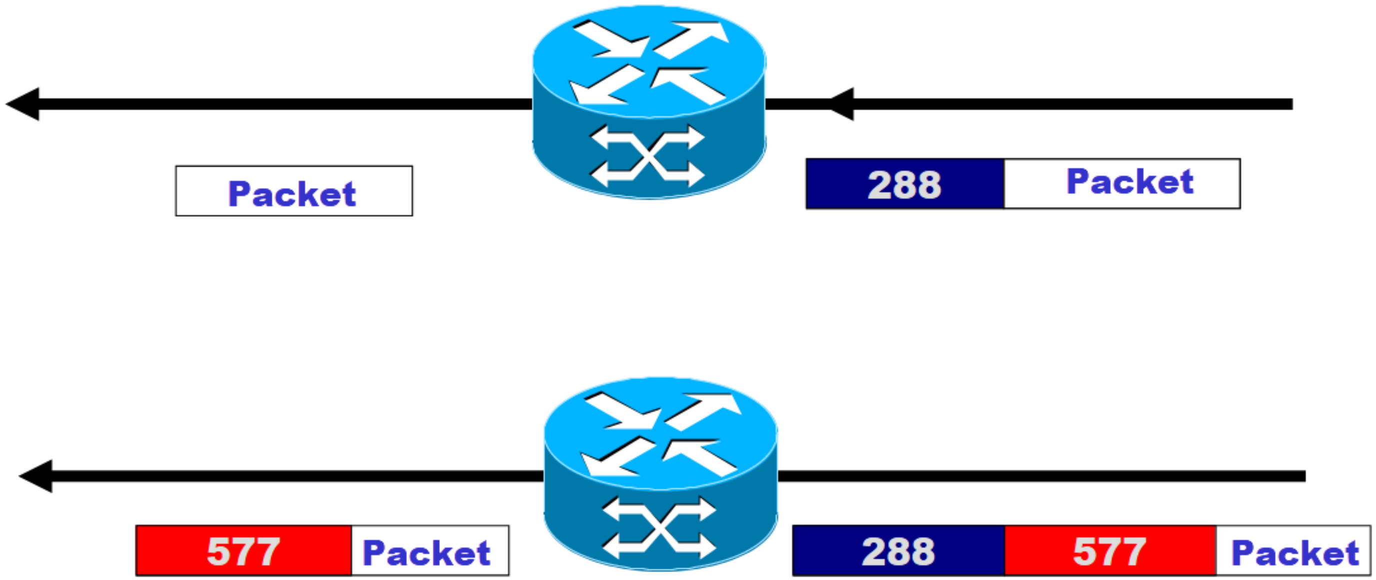
- Swapping is the process of changing the label of an incoming packet in the outgoing packet.
  - Labels are short and based on fixed-length fields.
  - Forwarding based on labels is much more efficient than forwarding based on IP routing tables.
  - Throughput capacity of MPLS routers is much higher than throughput capacity of IP routers.

MPLS Pushing Labels



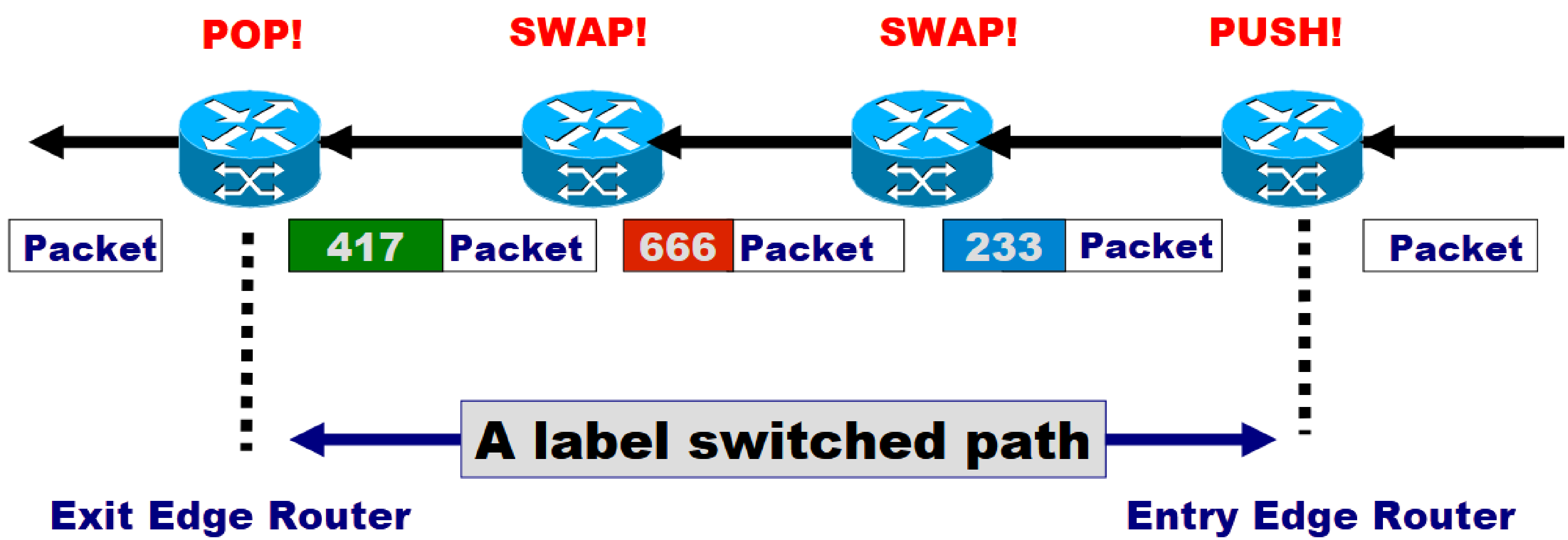
- Pushing is the process of adding a label to an incoming packet in the outgoing interface.

MPLS Popping Labels



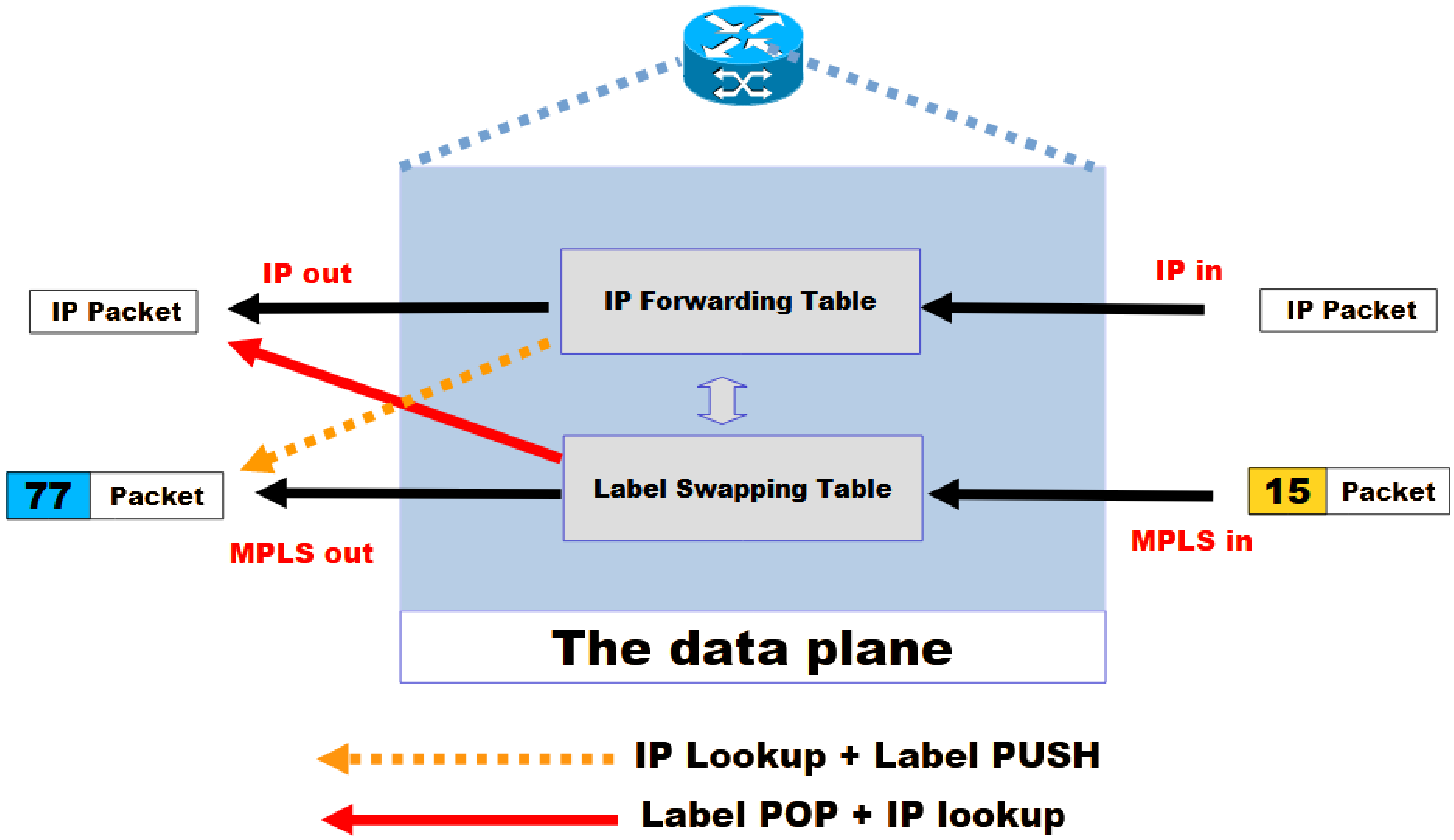
- Popping is the process of removing a label from an incoming packet in the outgoing interface.

A Label Switched Path (LSP)

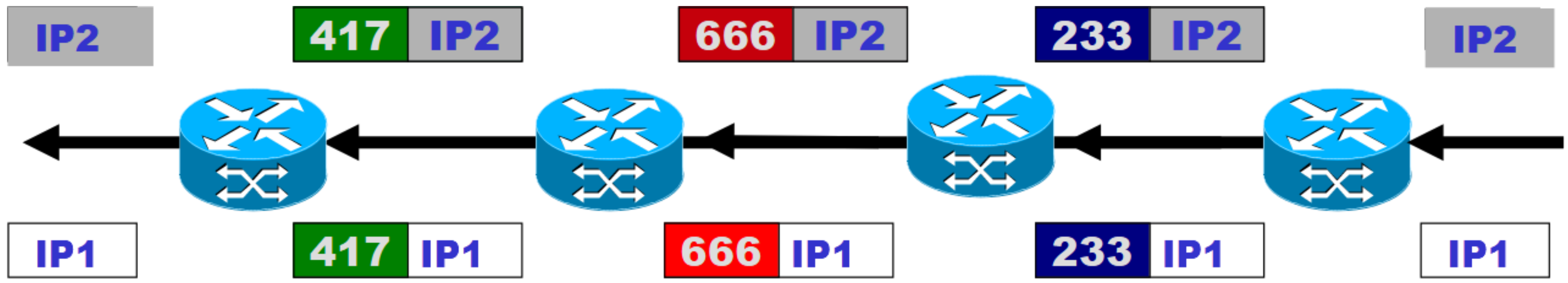


- An LSP is often referred to as a MPLS tunnel, as payload headers are not inspected inside of an LSP

Label Switched Router (LSR)



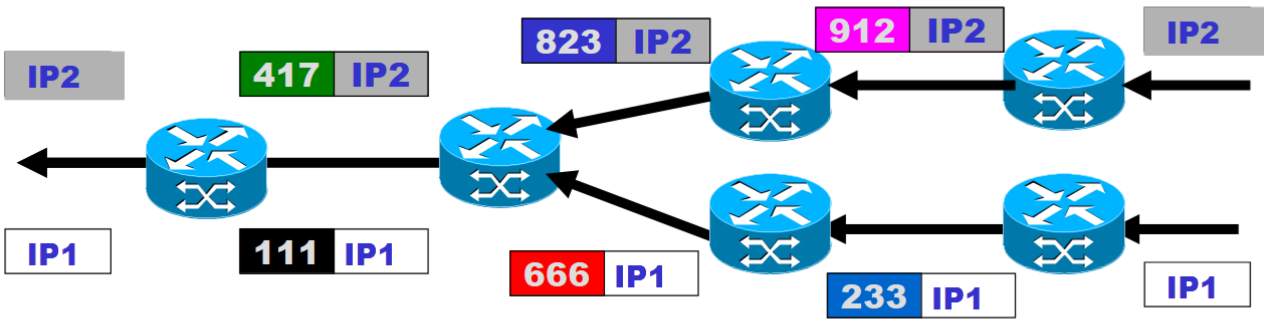
Forwarding Equivalence Class (FEC)



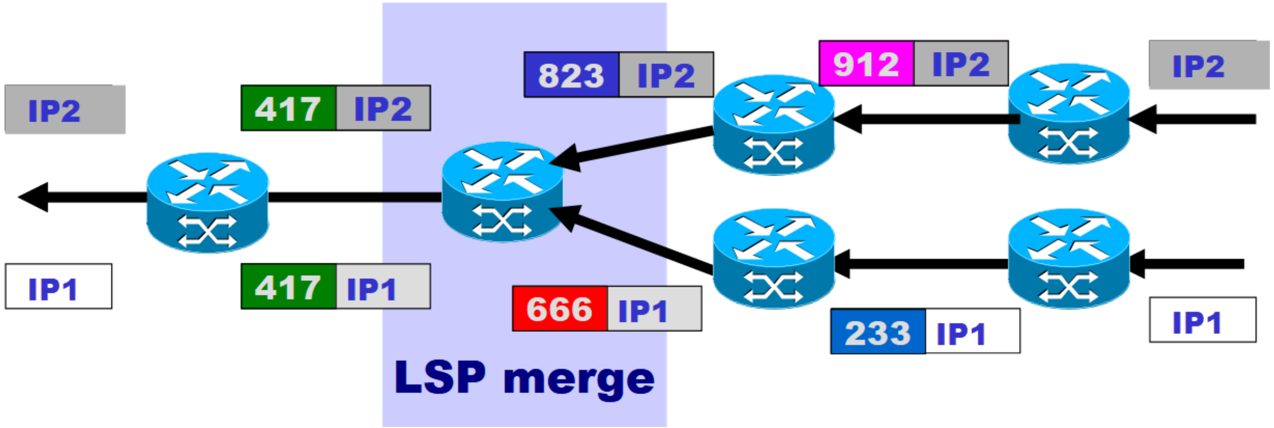
- Packets IP1 and IP2 are in the same FEC: they are forwarded in the same way.
  - IP layer headers are not inspected inside an MPLS LSP.
  - This means that inside of the tunnel, the LSRs do not need full IP forwarding tables.

Merging LSPs

- Two FECs to the same exit Edge Router with a common part in the two routing paths:

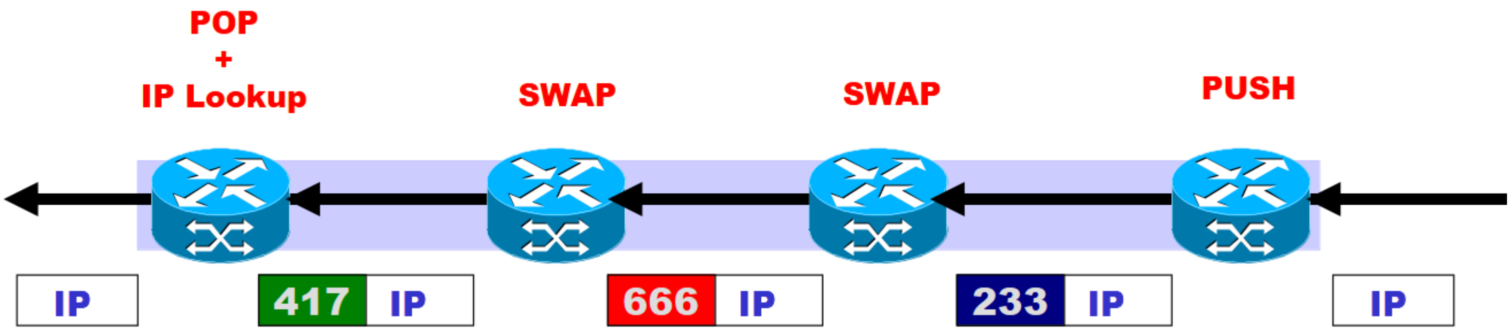


- The common part can be merged to reduce the size of the Label Swapping tables:

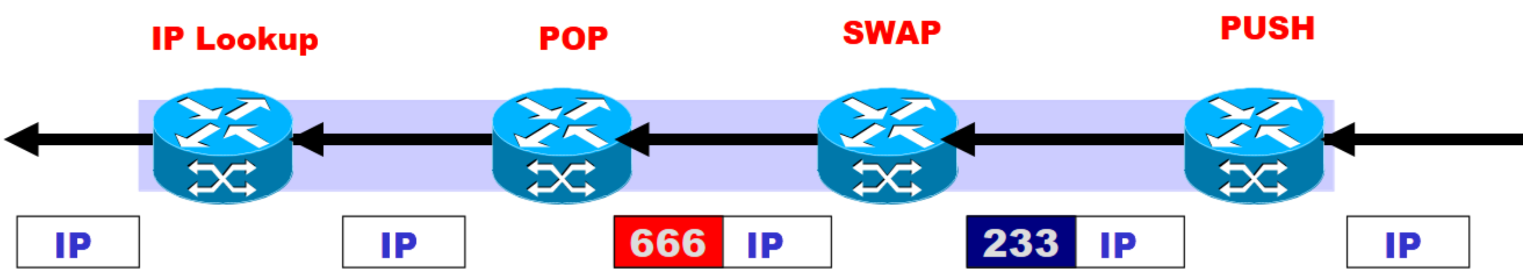


Penultimate Hop Popping (PHP)

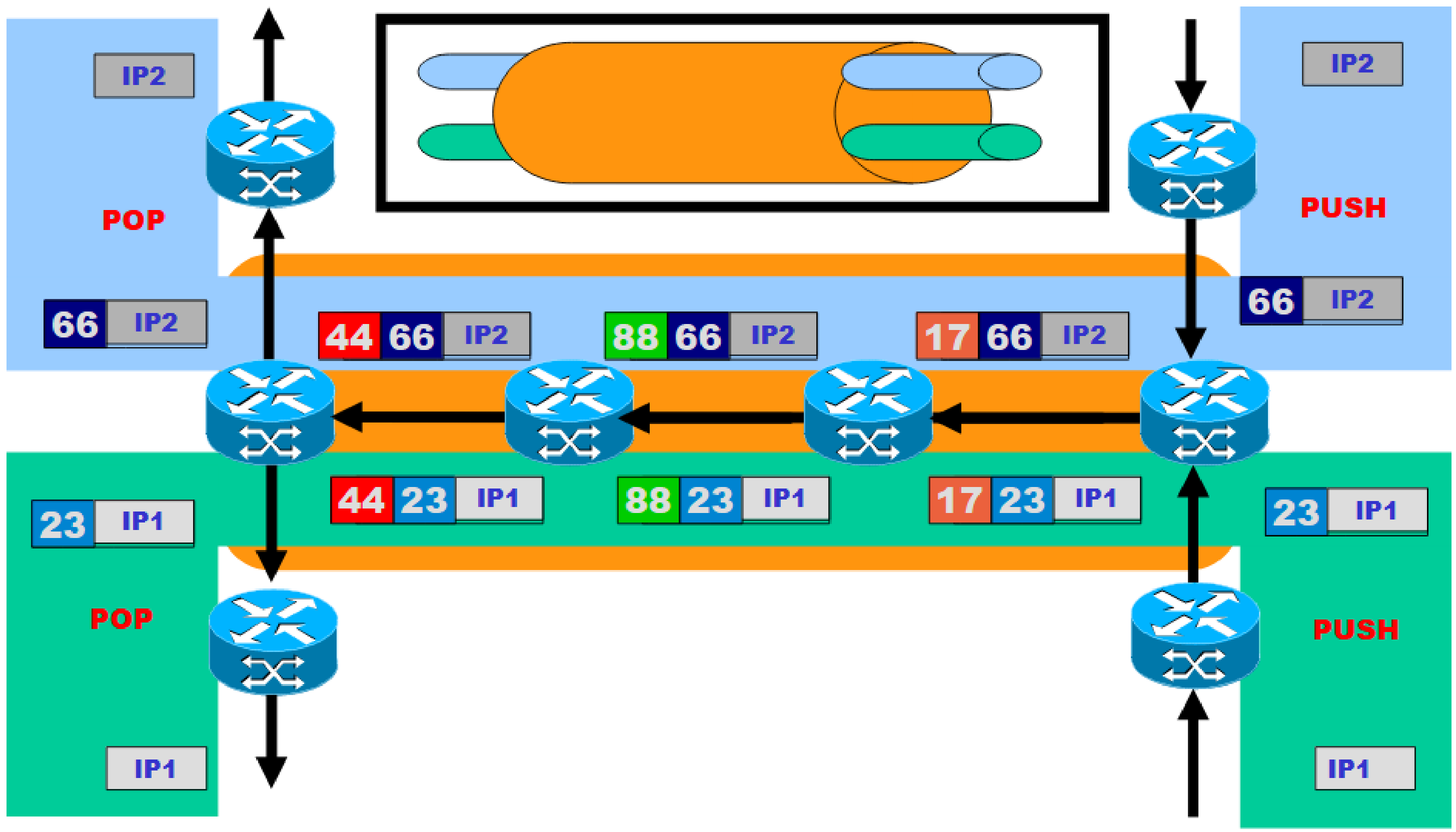
- An LSP from an entry Edge Router to an exit Edge Router:



- PHP is the process of label popping in the last Core Router before the packet reaches its exit Edge Router
  - it reduces the Edge Router load (usually, Edge Routers have more processing tasks than Core Routers)



LSP Hierarchy via Label Stacking



- The blue and the green LSPs are supported over the orange LSP in the central part of the MPLS network