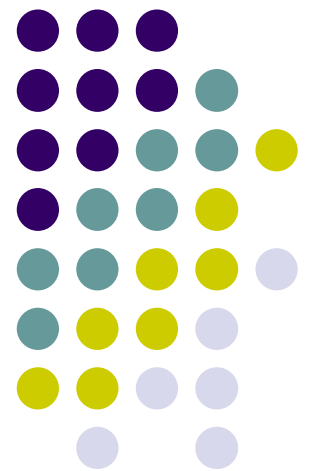# MPLS - MultiProtocol Label Switching

Curs in colaborare cu
Ramona Marfievici
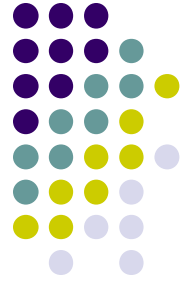
Agenda

- De ce MPLS
- Conceptul si terminologia MPLS
- Operatii MPLS
- Concluzii

Agenda

- De ce MPLS?
- Conceptul si terminologia MPLS
- Operatii MPLS
- Concluzii

De ce MPLS

- Cerintele utilizatorilor
  - calitatea serviciilor
  - securitate, robustete
  - clase de servicii, servicii ieftine
  - trafic any-to-any
  - optiuni pentru servicii
    - ATM, FR, IP, Ethernet
    - optiuni multiple pentru VPN
  - any service, any time, anywhere

Sistem autonom (Autonomous System - AS) este un grup de reţele şi routere care aparţin unei autorităţii cu o singură administrare

Internetul – colectie de furnizori servicii Internet (ISPs), conectati (accesati) prin PoPs (point of Presence) si Network Access Points
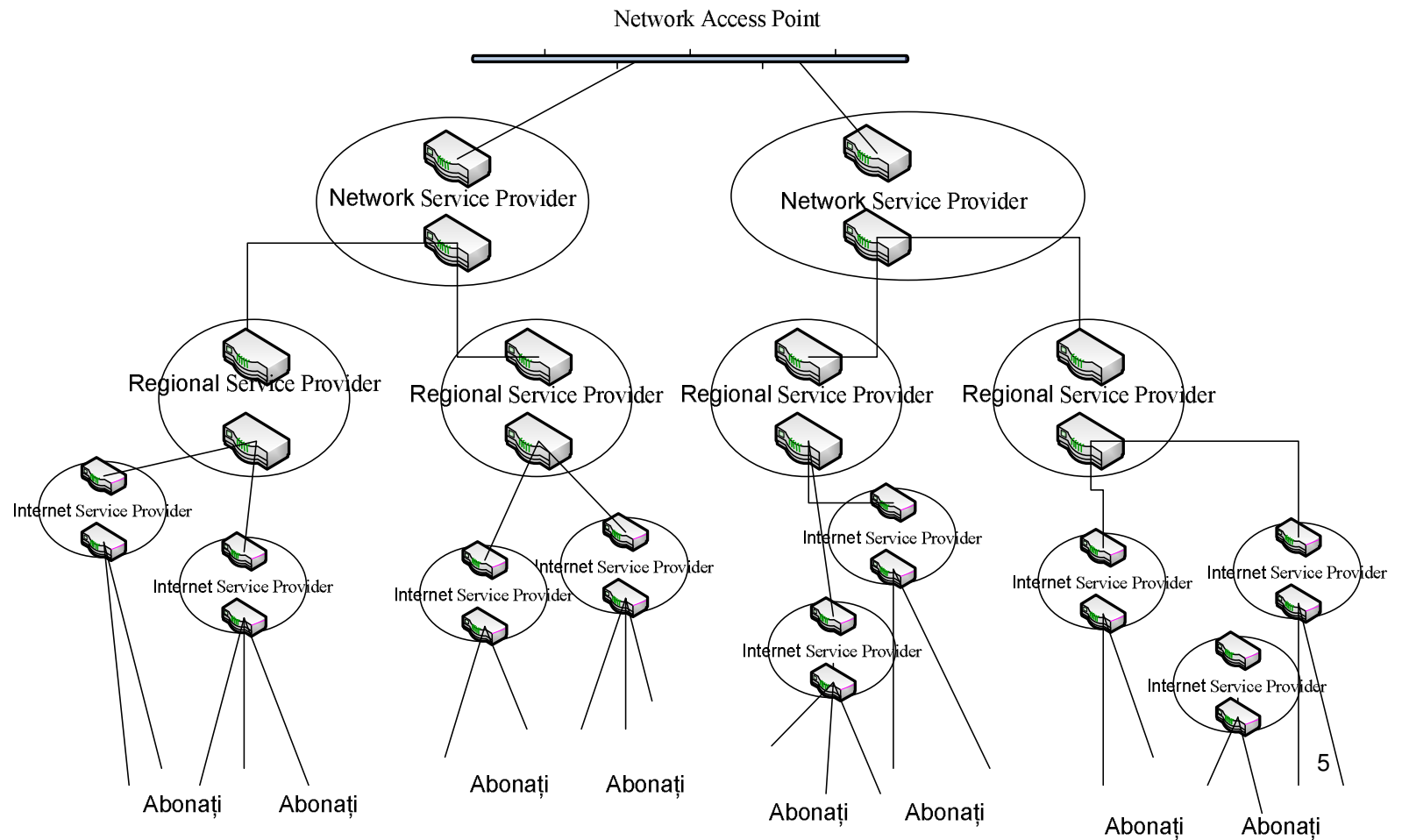
De ce MPLS (cont.)

- Solutii:
  - IP
  - ATM, Frame Relay
  - IP over ATM
  - LANE (LAN Emulation)
 - IP switching (Tag switching, ARIS)

IP, ATM, Frame Relay

- IP
  - flexibilitate, scalabilitate
  - conectivitate infinita
  - best-effort
  - serviciu neorientat pe conexiune
- ATM, Frame Relay
  - QoS, management de trafic
  - nu au conectivitate any-to-any
- diferite dar bune

IP-over-ATM

- IP peste circuite virtuale

- Traficul IP transformat in trafic ATM

- Topologie, trafic : ATM

- Topologie logica complicata (mesh de circuite virtuale)

- Lipsa topologie de nivel 3

  - toate ruterele sint vecine

- Problema fundamentala: nu se stie de existenta IP

Topologia fizica

Topologia logica

8
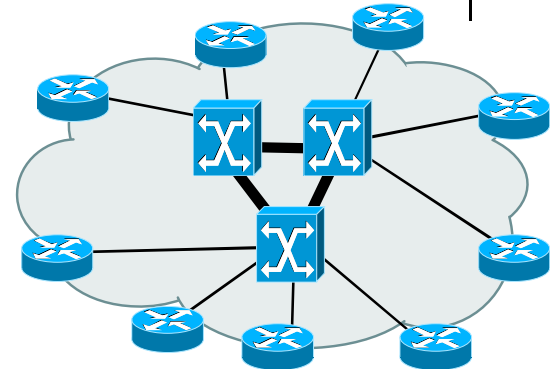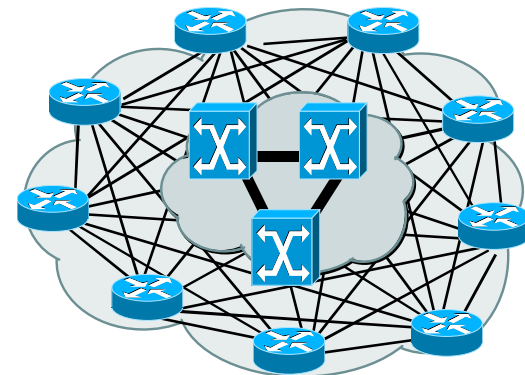
IP + ATM, MPLS: o noua paradigma

- Framework pentru implementarea QoS
- Combina ce e mai bun din cele doua lumi:
    - QoS din ATM, Frame Relay
    - flexibilitate si scalabilitate din IP

Avantaje MPLS

- **Simplified forwarding**: Packet forwarding is based on exact match for a short label, rather than a longest match applied to a lon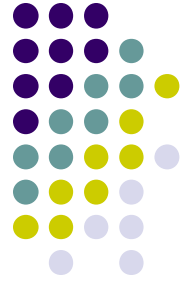ger address as required by datagram forwarding. Also, the label headers are simpler, resulting in a simpler forwarding paradigm.

- **Efficient explicit Routing**: Overhead of Source Routing in the case of pure datagram routing is prohibitive, since the entire explicit route (ER) is carried in each packet. In MPLS, the explicit route is carried only once, when the label switched path is being set up. Thus ER is more practical in MPLS (see next slides).

- **Traffic Engineering** is the process of selecting the paths chosen by data traffic in order to balance the traffic load on the various links, routers and switches in the network.

Avantaje MPLS

- MPLS allows data streams from any particular ingress to any particular egress to be individually identified, thereby providing a straight-forward mechanism to measure the traffic associated with each ingress-egress node pair. In addition, efficient ER ensures that any particular stream of data takes the preferred path.

- Since MPLS allows efficient ER, it follows that MPLS also facilitates **QoS routing**.

- **Complex Mappings from IP packet to FEC (**Forward Equivalence Class) at the ingress node of an MPLS domain offers an efficient method to support provisioned QoS for data traffic. ISPs can offer differentiated services to preferred customers, by providing filtering, based on src and dest address, incoming interface, etc. and then labeling the packet in some way in the MPLS domain.

Avantaje MPLS

- **Partitioning of Functionality**: it is possible to hierarchically partition the processing functionality to the different network elements. More heavy processing takes place on the edges of the network, near the customers, and on the core network, the processing is as simple as possible, eg. pure label based forwarding.

- Common Operation over Packet and Cell media

Alegerea caii (rutei) MPLS

- alegerea unei cai pentru un flux de date

OBS: in cale, pachetele sunt comutate pe baza de etichete

- **hop-by-hop sau rutare explicita**
- hop-by-hop
  - fiecare LSRouter alege independent urmatorul hop
  - protocoalele existente sunt bazate pe prefixul adresei destinatiei

- rutare explicita
  - toate sau o parte din LSR dintr-o cale sunt specificate
  - pre-configuare sau dinamic
  - ideal, dinamic => traffic enginnering, dar trebuie cunoscuta topologia domeniului si parametri QoS ai domeniului
  - se pot defini cai de backup => scade timpul de convergenta
- caile LSPath sunt unidirectionale

Explicitly Routed LSP

Overload !!

LER 1

Overload !!

LER 4

Forward to
LSR 2
LSR 3
LSR 4
LSR X

LSR 2

LSR 3
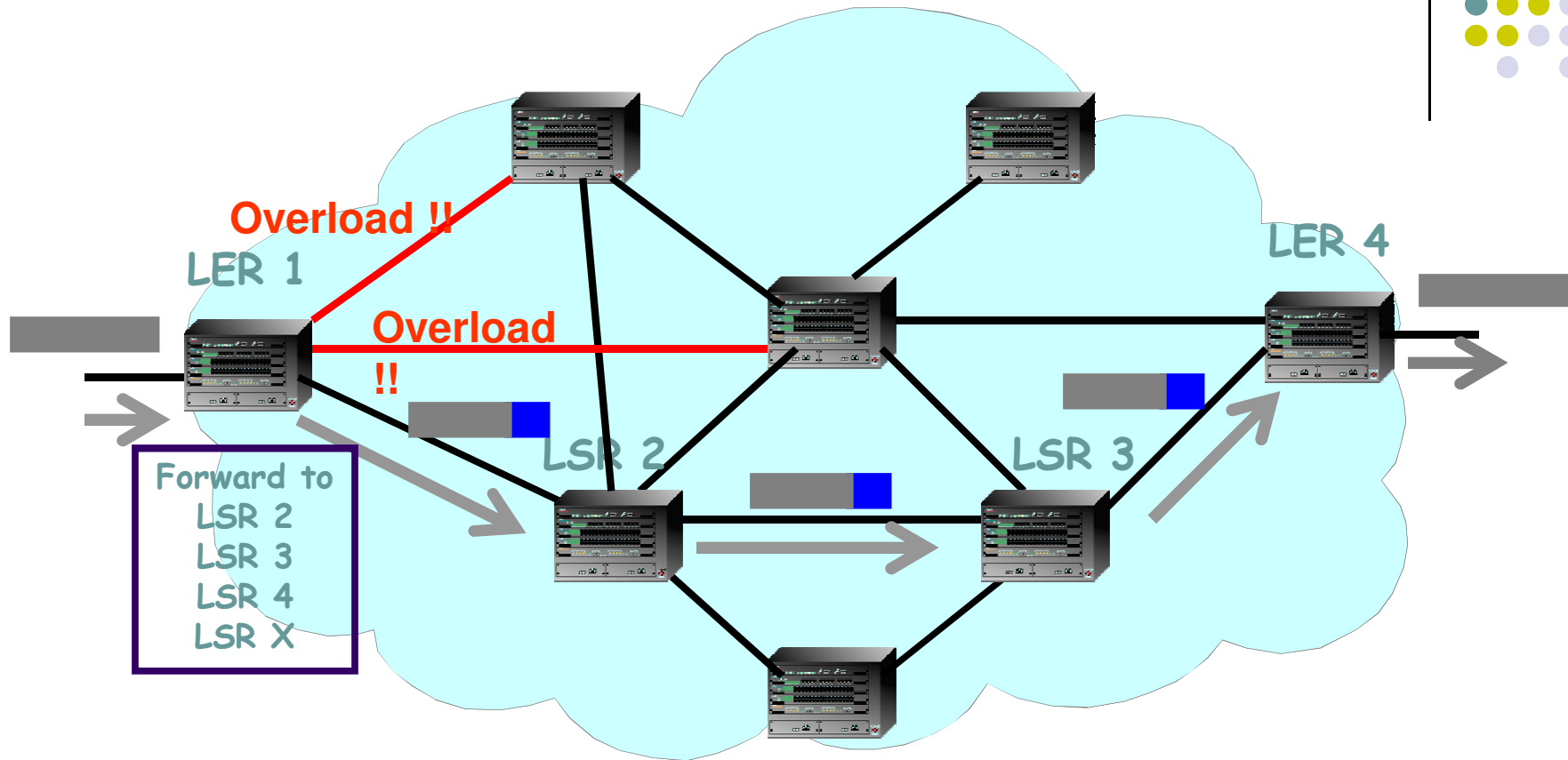
- **End-to-End forwarding decision determined by ingress node.**
- **Enables Traffic Engineering**

14

MPLS si ISO model
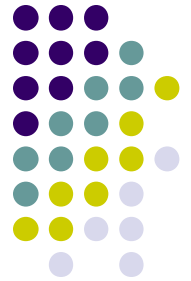
IETF proiecteaza un nou protocol, fara a fi afectate cele anterioare (backward compatible)

MPLS – ofera servicii la nivel 2+

	7 to 5
Applications	
TCP UDP	4
IP	3
MPLS	
PPP FR ATM (*)	2
Physical (Optical - Electrical)	1
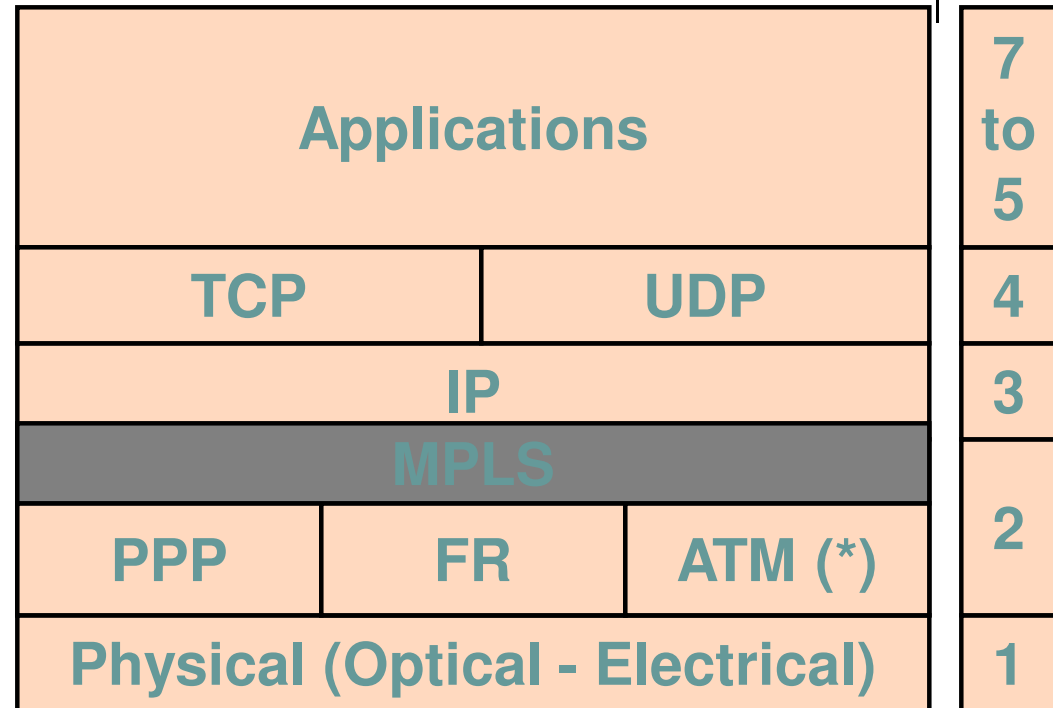
(*) ATM overlay model (fara adresare si P-NNI) – considerat ca un ISO layer 2 protocol.

15

Agenda

- De ce MPLS
- Conceptul si terminologia MPLS
- Operatii MPLS
- Concluzii

Conceptul MPLS

- Switching de nivel 2+
- Datagrama "intilneste" circuitul
- Decupleaza rutarea de forwardare
- Dezvoltat in cadrul IETF - fast-forwarding

http://www.ietf.org/html.charters/mpls-charter.html

- Independent de tehnologie
- Forward pe baza etichetei

Conceptul MPLS: Arhitectura MPLS

	Ingress Node	Core Node	Egress Node
LSP	Classification Label assignment	Label swapping	Label removal
Routing protocol	OSPF	OSPF	OSPF
FEC table	Local table	Local table	Local table
Attributes	Precedence		
Label table	Local table	Local table	Local table
Label Switch	Layer 2 / Layer 1	Layer 2 / Layer 1	Layer 2 / Layer 1

Ingress Node   Core Node   Egress Node
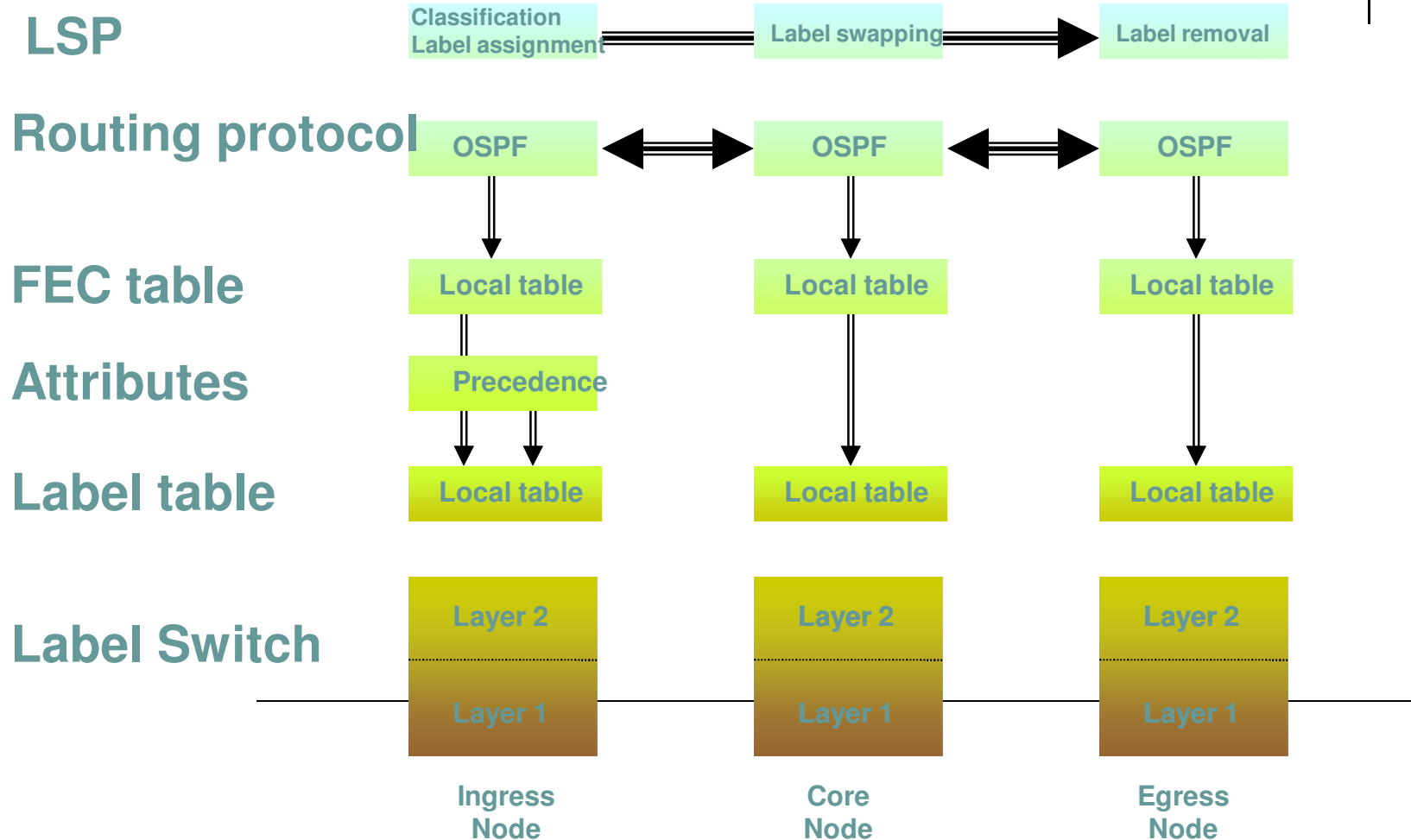
Conceptul MPLS (cont.)

La intrare:
- clasificarea pachetelor
- etichetare
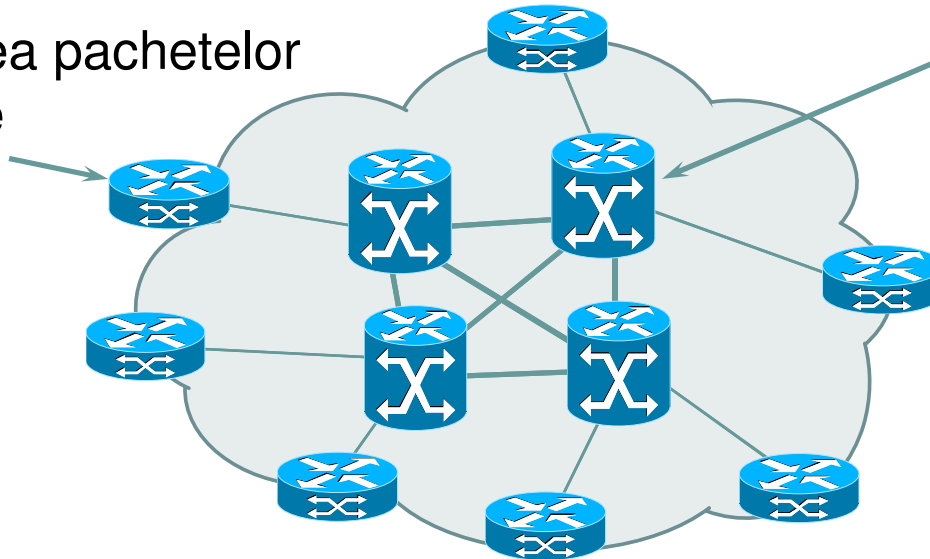
In retea:
- comutare cu etichete
- eticheta indica serviciul si destinatia

- comutarea cu etichete
- protocoale de nivel retea: IP, IPX, AppleTalk
- eticheta: unde si cum sa transmit pachetul

Conceptul MPLS (cont.)

IGP domain with a label distribution protocol
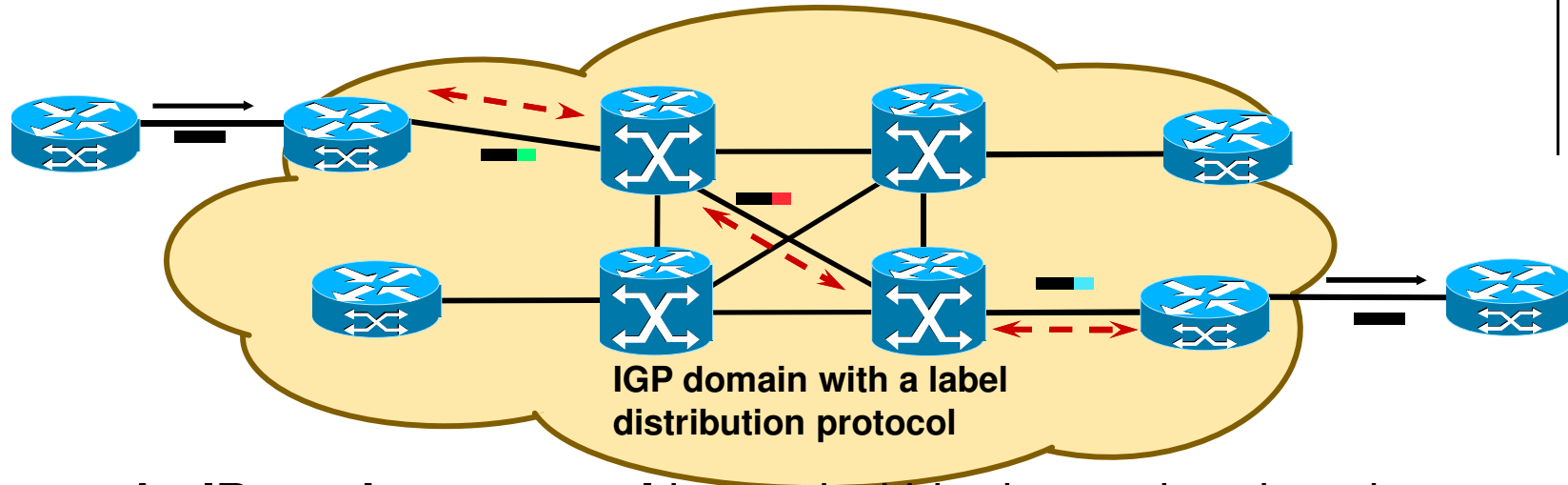
- An **IP routing protocol** is used within the routing domain (e.g.:OSPF, IS-IS)

- A **label distribution protocol** is used to distribute address/label mappings between adjacent neighbours

- The ingress LSR receives IP packets, performs packet classification, assign a label, and forward the labelled packet into the MPLS network

- Core LSRs switch packets/cells based on the label value

- The egress LSR removes the label before forwarding the IP packet outside the MPLS network

MPLS Cloud

LER

L3 Routing

LER

L3 Routing

LER

L3 Routing

LSR
Label Swapping

LSR
Label Swapping

LER

L3 Routing

LER
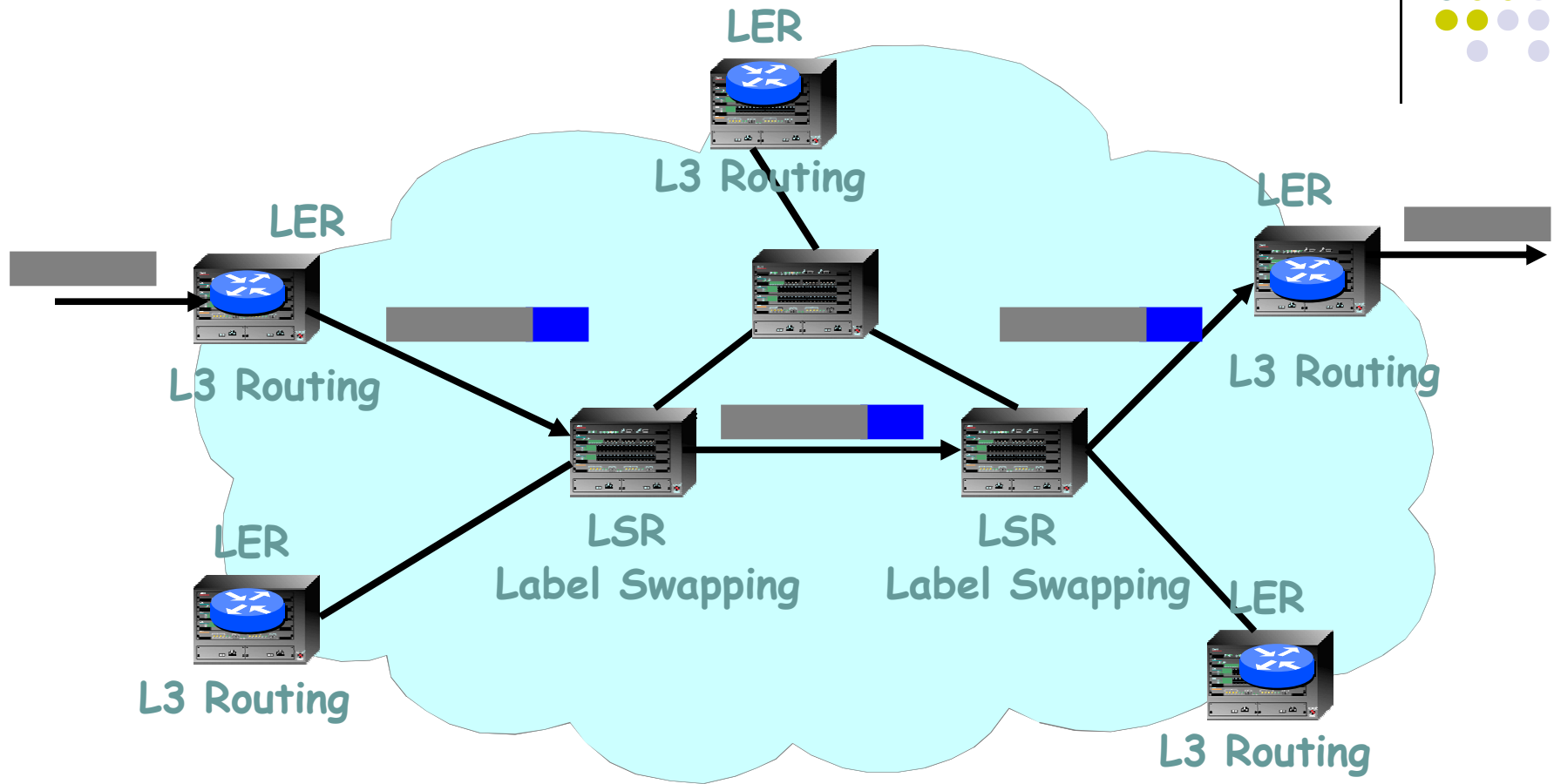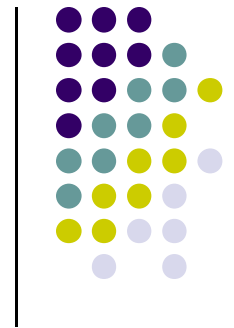
L3 Routing

IP Packet

IP Packet w/ Label

Edge LSR Features

– **Routing protocols**

– **FEC Classification**

– **Initiates LSP setup for Downstream On Demand method**

– **Adaptation of non-MPLS data to MPLS data**

– **Layer 2 translation for MPLS data**

– **Terminated MPLS-VPN**

– **At least one LDP protocol**

– **Edge LSR is counted into the TTL count as a regular router**
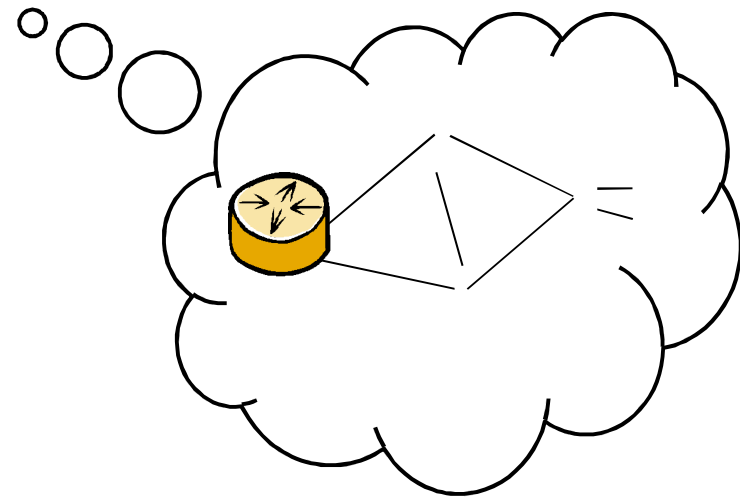
Core LSR Features

- – **Routing protocols**
- – **Propagates Downstream On Demand method (request and mapping)**
- – **Layer 2 translation**
- – **High speed label forwarding/switching**

- – **At least one LDP protocol**

23
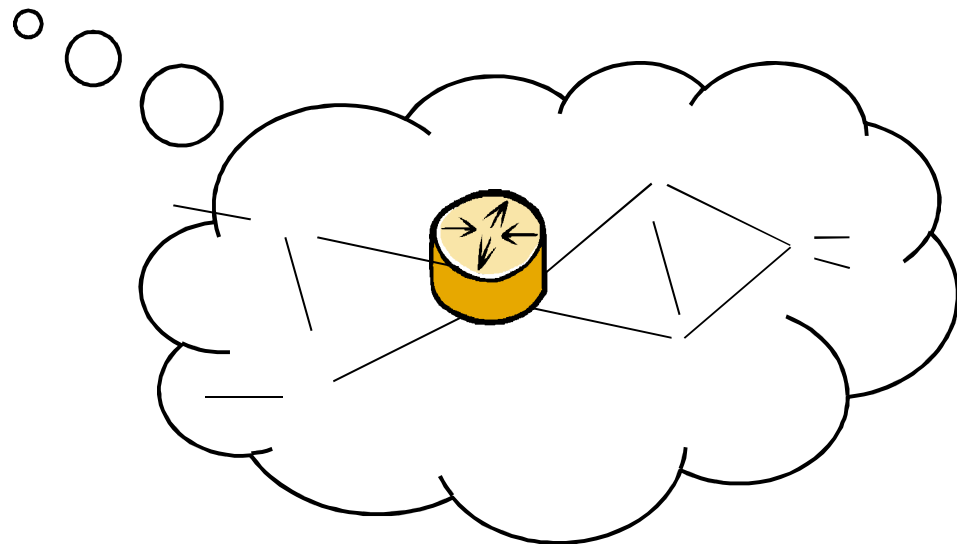
Terminologie MPLS

- etichete MPLS
- clase de echivalenta
- rutere MPLS
- cai comutate
- penultimate/ultimate hop popping
- protocoale de semnalizare

Some MPLS Terms...

- **LER** - Label Edge Router
- **LSR** - Label Switch Router
- **FEC** - Forward Equivalence Class
- **Label** - Associates a packet to a FEC
- **Label Stack** - Multiple labels containing information on how a packet is forwarded.
- **Shim** - Header containing a Label Stack
- **Label Switch Path** – unidirectional path that a packet follows for a specific FEC; may differ from routing protocol's shortest path
- **LDP** - Label Distribution Protocol, used to distribute Label information between MPLS-aware network devices
- **Label Swapping** - manipulation of labels to forward packets towards the destination.

Etichete MPLS

- identificator de dimensiune fixa; identifica o clasa de echivalenta la forwarding (FEC)

- semnificatie locala (ruter); semnificatia unui identificator de layer 2

Label (20 bits)	Exp (3 bits)	S (1 bit)	TTL (8bits)

- eticheta: 20 biti (0…1048575)

- Exp: biti experimentali (3biti)

- Stackbit (1bit) – stiva last-in first-out (setat indica bottom of stack)

- Time To Live (8biti)

- etichetele 0-15 rezervate IETF

- Numita si **MPLS shim header ,** daca antetul DLink nu poate transporta eticheta, se incapsuleaza la nivel DL un antet de eticheta

- cei 32 biti formeaza **Label Stack Entry**

Eticheta MPLS - incapsulare

PPP Header(Packet over SONET/SDH) | PPP Header | Eticheta | Layer 3 Header

Ethernet | Ethernet Hdr | Eticheta | Layer 3 Header

Frame Relay | FR Hdr | Eticheta | Layer 3 Header

ATM Cell Header | GFC | VPI | VCI | PTI | CLP | HEC | DATA

Eticheta

Subsequent cells | GFC | VPI | VCI | PTI | CLP | HEC | DATA

Eticheta

27
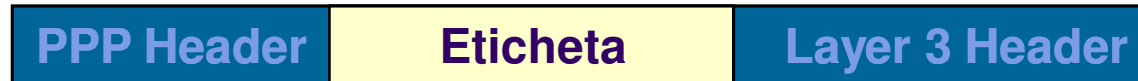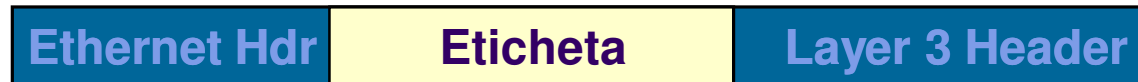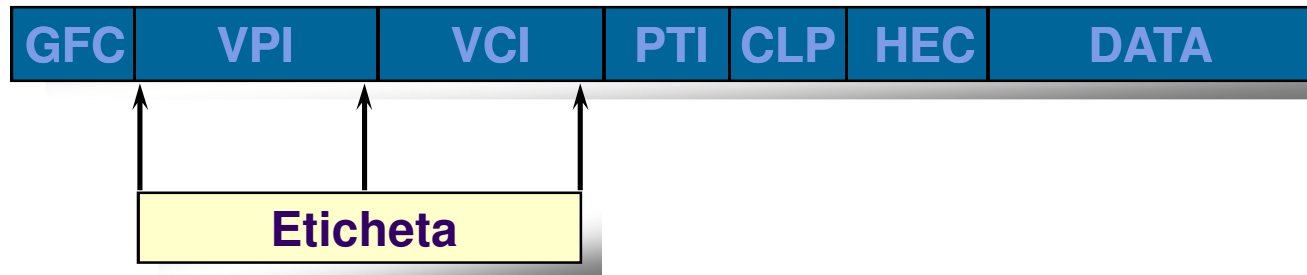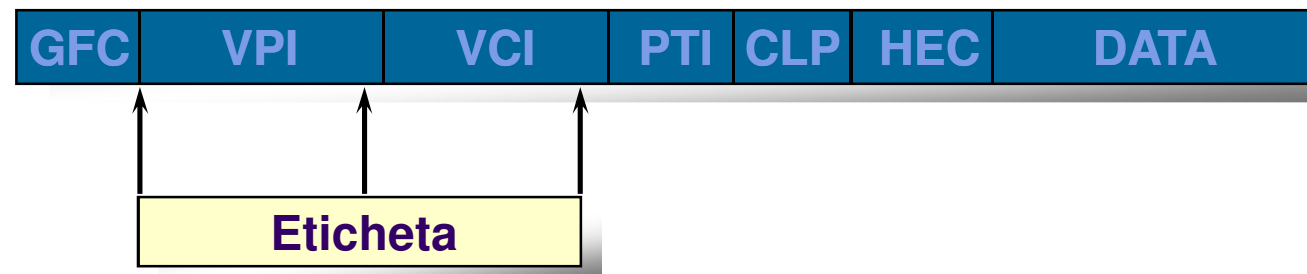
Clasele de echivalenta (FEC)

- subset de pachete comutate in aceeasi maniera (interfata, nexthop, eticheta)
- Un pachet poate fi mapat la o clasa de echivalenta (FEC particular) bazat pe criterii de:

  - **destination IP address,**
  - **source IP address,**
  - **TCP/UDP port,**
  - **in case of inter AS-MPLS: Source-AS and Dest-AS,**
  - **class of service,**
  - **application used,**
  - **…**
  - **any combination of the previous criteria.**

Clase de echivalenta

•tabela de rutare extinsa la ruterul de intrare (ingress)
•un FEC este asociat cu cel putin o eticheta
•nod de intrare (ingress): mapare FEC
  •prefix adresa IP destinatie
  •identificator ruter
  •flux (SA/DA)
  •QoS

Ingress Label	FEC	Egress Label
6	138.120.6/24 - xxxx	9

Ingress Label	FEC	Attribute	Egress Label
6	138.120.6/24 - xxxx	A	9
6	138.120.6/24 - xxxx	B	12

Rutere MPLS

Ruter tranzit
- ruter
- switch ATM

Cale comutata (LSP)
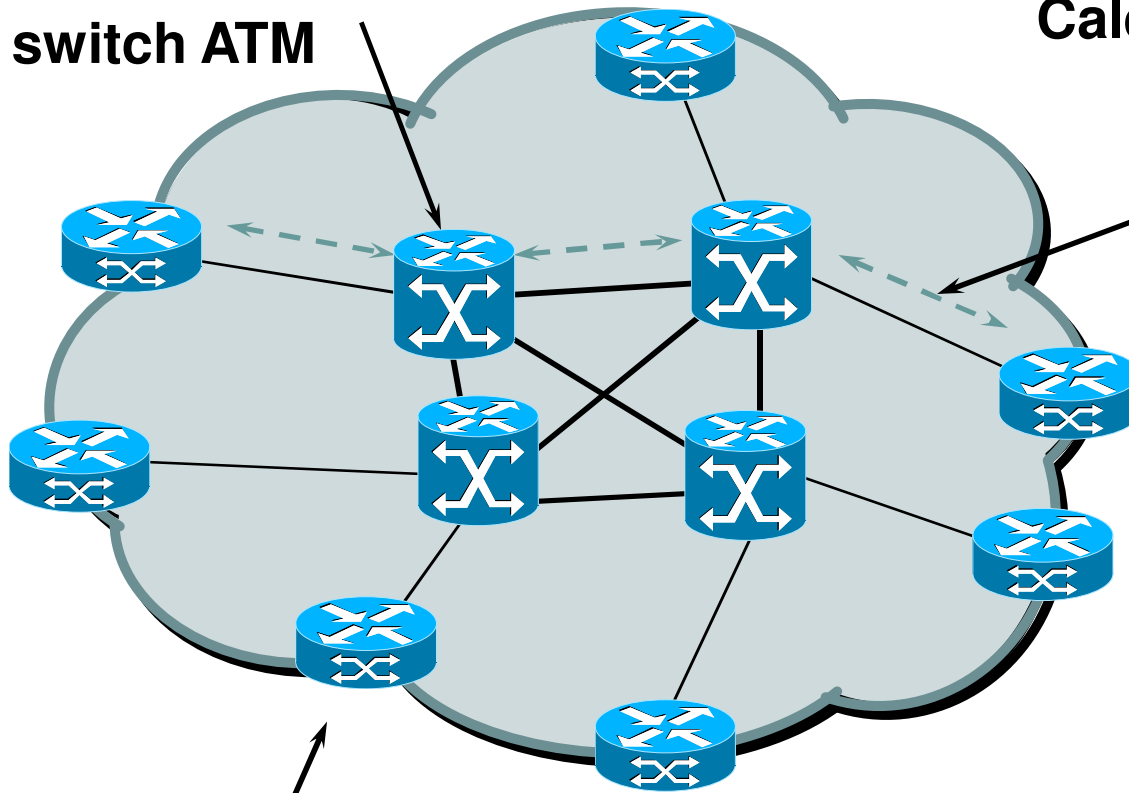
Ruter ingress - de intrare
- switch ATM sau ruter

Ruter egress – de iesire

Asignarea etichetelor

- Etichetele - semnificatie locala

- LSR asigneaza etichete pentru FEC

- Etichete asignate local si transmise intre vecini

- LIB (label information base): interface in, label in, interface out, label out

- LSR isi cunosc vecinii, prin protocoalele de rutare

Exemplu: Rtr-C este vecin in aval (downstream) pentru Rtr-B
pentru adresa 171.68.10/24

171.68.40/24   **171.68.10/24**

Rtr-A   **Rtr-B**   **Rtr-C**

Scheme de asignare etichete

- ## Topology Driven
  - Label assignment in response to routing protocols (OSPF and BGP) updates

- ## Control Driven
  - Label assignment in response to RSVP, CR-LDP requests

- ## Traffic Driven
  - Label assignment in response to flow detection & triggering

Label Switched Path

Ingress Interface	Ingress Label	FEC	Egress Interface	Egress Label
1	5	138.120	3	12

Ingress Interface	Ingress Label	FEC	Egress Interface	Egress Label
1	12	138.120	4	x

MPLS switch

MPLS switch

127.20 $_1$

4 138.120

3

$_2$ **MPLS switch** 192.168

$_2$

$_3$

$_2$

$_1$

$_2$

MPLS switch

Ingress Interface	Ingress Label	FEC	Egress Interface	Egress Label
1	x	138.120	3	5

33
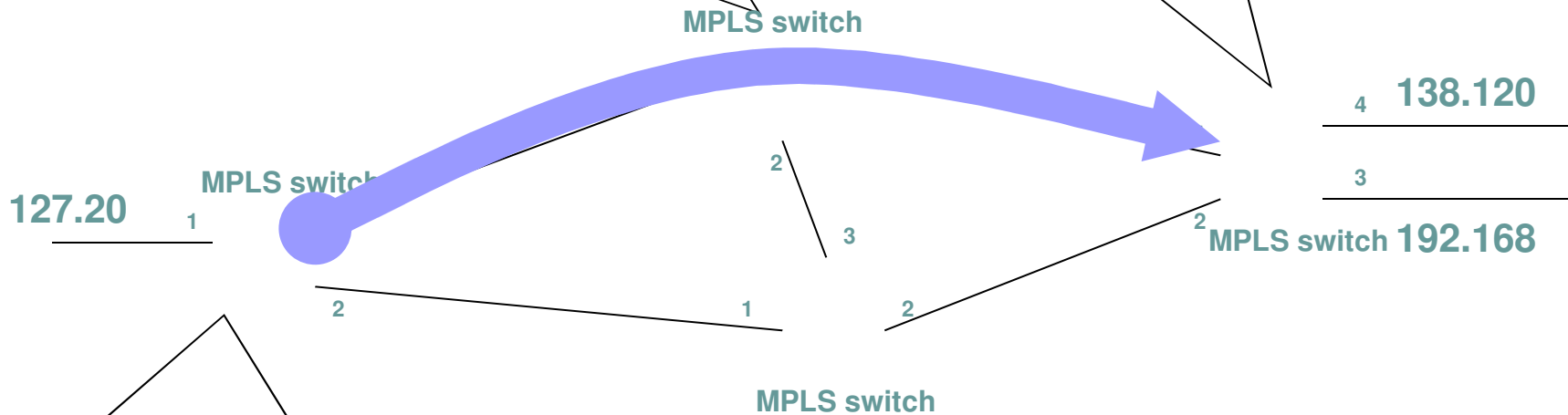
Hop by Hop IP forwarding

Ingress Interface	Ingress Label	FEC	Egress Interface	Egress Label
1	Default	None	3	Default

Ingress Interface	Ingress Label	FEC	Egress Interface	Egress Label
1	Default	None	4	x

??

MPLS switch

138.120.6.12

138.120.6.12

??

138.120

138.120.6.12

138.120.6.12

MPLS switch

138.120.6.12

127.20

138.120.6.12

??

2

3

1

2

3

2

MPLS switch 192.168

2

MPLS switch

Ingress Interface	Ingress Label	FEC	Egress Interface	Egress Label
1	x	None	3	Default
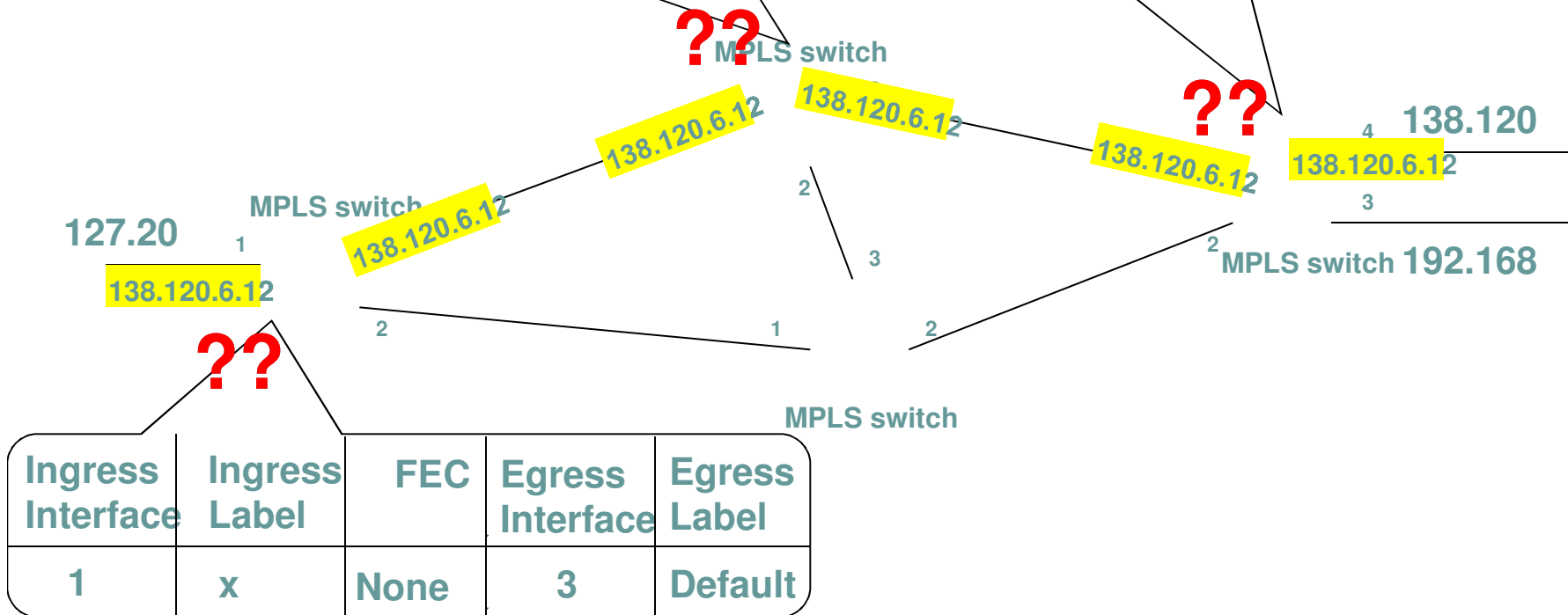
IP forwarding using LSP

Ingress Interface	Ingress Label	FEC	Egress Interface	Egress Label
1	5	138.120	3	12

Ingress Interface	Ingress Label	FEC	Egress Interface	Egress Label
1	12	138.120	4	x

MPLS switch

138.120.6.12

138.120.6.12

138.120.6.12

138.120.6.12

138.120

138.120.6.12

MPLS switch

127.20

138.120.6.12

MPLS switch 192.168

MPLS switch

Ingress Interface	Ingress Label	FEC	Egress Interface	Egress Label
1	x	138.120	3	5
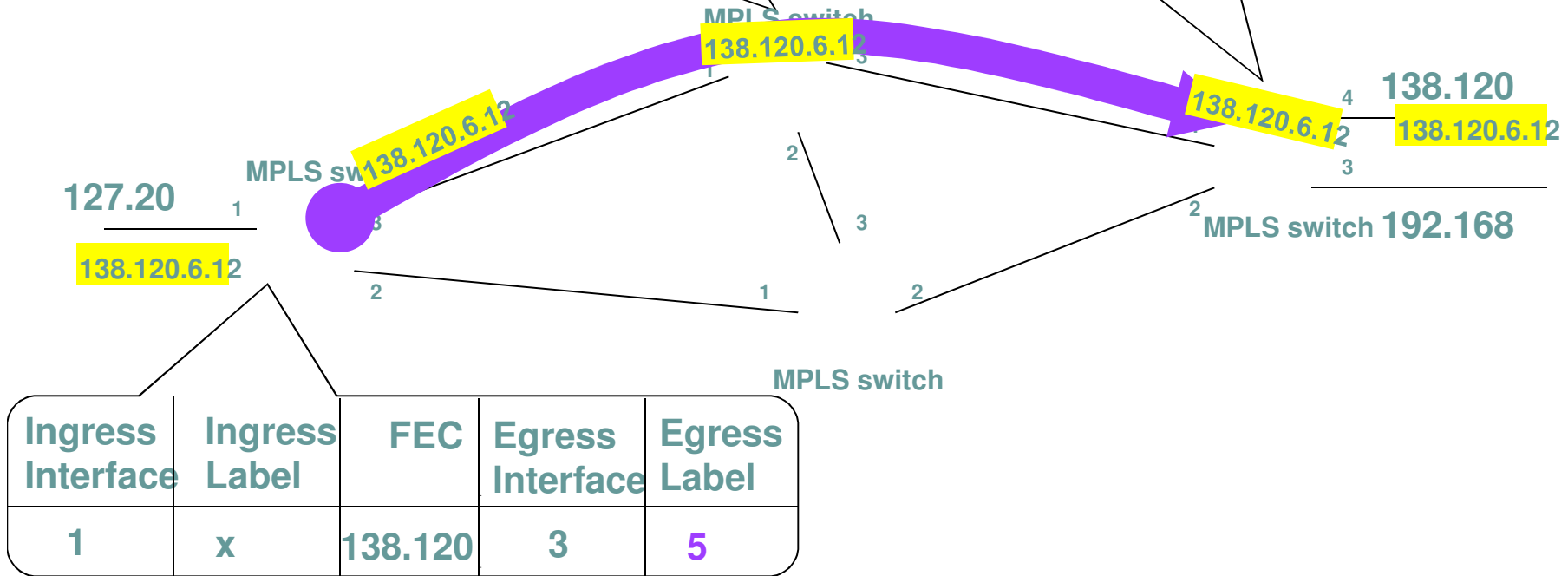
Distributia etichetelor

Distributie nesolicitata

Utilizeaza eticheta **30** pentru 171.68.10/24

Utilizeaza eticheta **40** pentru 171.68.10/24

171.68.40/24   **Rtr-A**   **Rtr-B**   **Rtr-C**   171.68.10/24

In I/F	In Lab	Address Prefix	Out I/F	Out Lab
0	-	171.68.10	1	**30**
...	...	...	...	...

In I/F	In Lab	Address Prefix	Out I/F	Out Lab
0	30	171.68.10	1	**40**
...	...	...	...	...

In I/F	In Lab	Address Prefix	Out I/F	Out Lab
0	40	171.68.10	1	-
...	...	...	...	...

Rute IGP
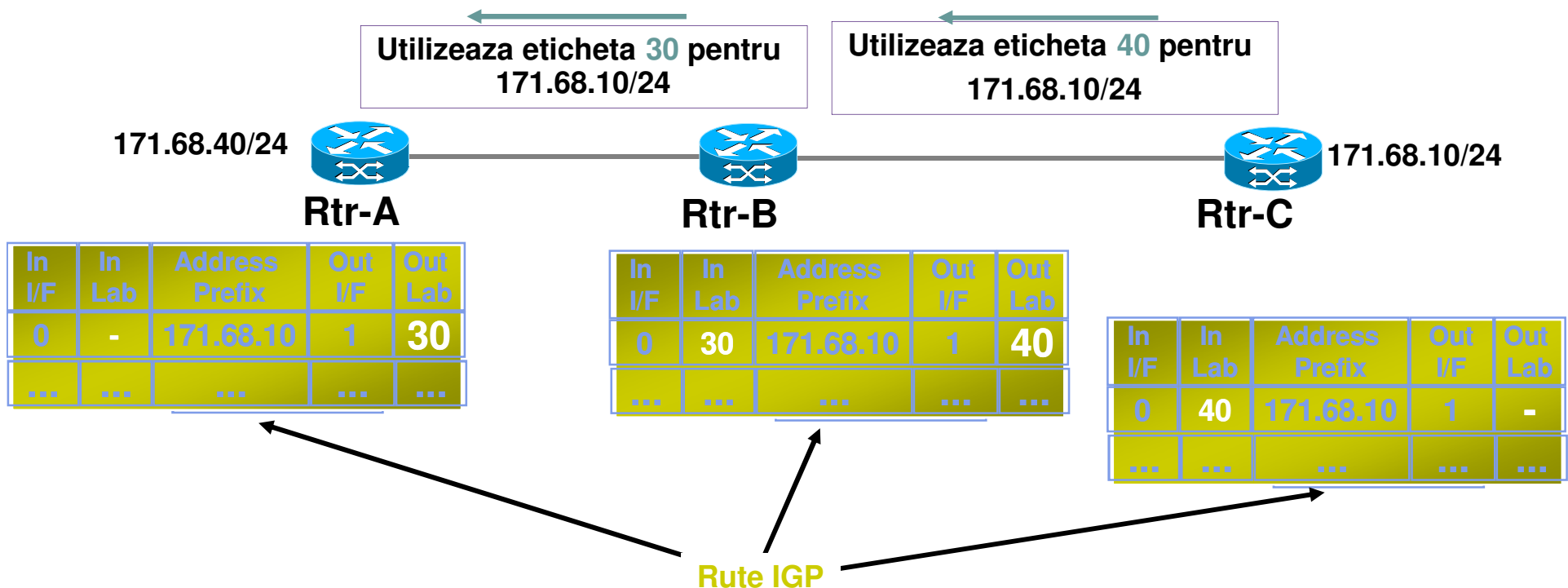
- LSR distribuie etichete vecınilor (upstream)

Unsolicited Downstream

Ingress Interface	Ingress Label	FEC	Egress Interface	Egress Label
1	5	138.120	3	12

Ingress Interface	Ingress Label	FEC	Egress Interface	Egress Label
1	12	138.120	4	x

MPLS switch

1   3

MPLS switch

2   4   138.120

127.20   1

3   1

Mapping 5   2   3

Mapping 12

2   **MPLS switch 192.168**

2   1   2
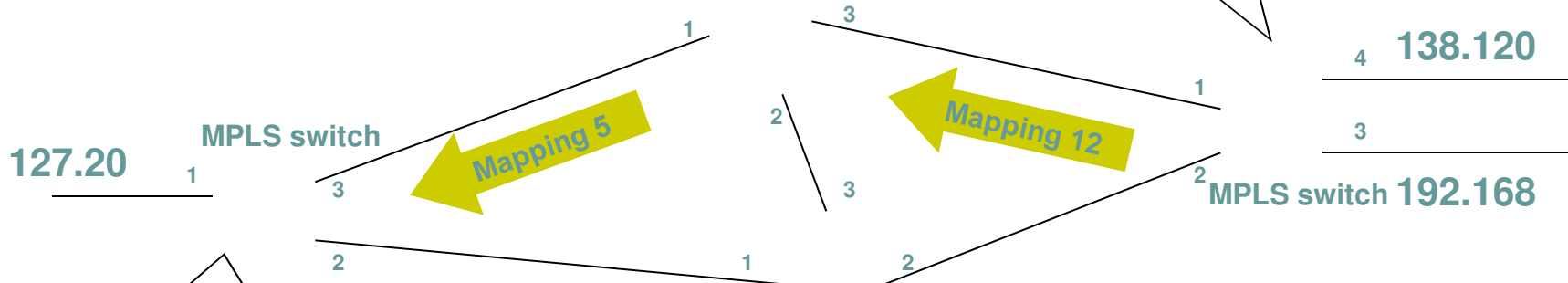
MPLS switch

Ingress Interface	Ingress Label	FEC	Egress Interface	Egress Label
1	x	138.120	3	5

The downstream node defines the label and advertises it to the upstream node.

Distributia etichetelor (cont.)

Distributie la cerere

Utilizeaza eticheta **40** pentru 171.68.10/24

Utilizeaza eticheta **30** pentru 171.68.10/24

171.68.10/24

171.68.40/24 **Rtr-A**   **Rtr-B**   **Rtr-C**

Cerere eticheta pentru destinatia 171.68.10/24

Cerere eticheta pentru destinatia 171.68.10/24
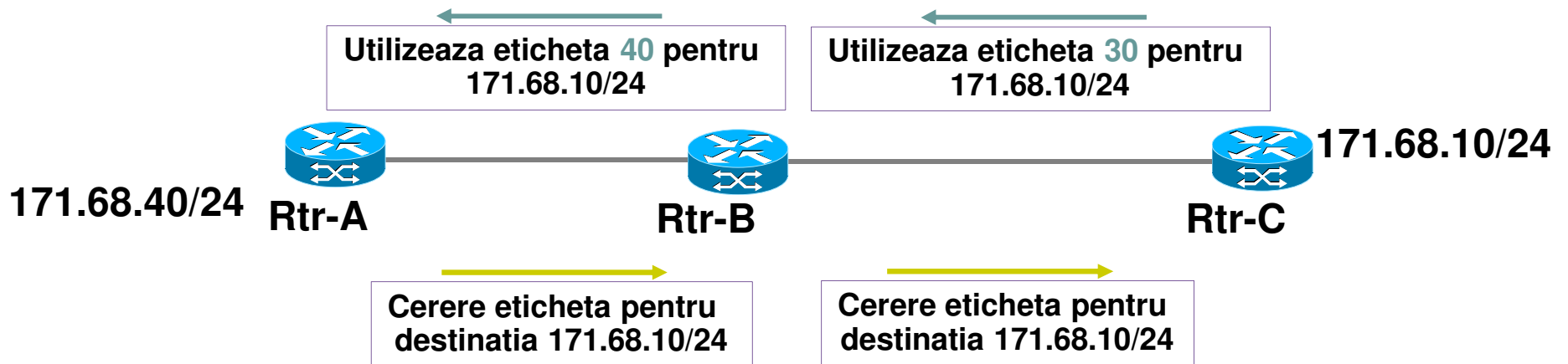
- **LSR (upstream) cer etichete vecinilor (downstream)**
- **LSR (downstream) distribuie etichete la cerere**

Downstream stream on demand

Ingress Interface	Ingress Label	FEC	Egress Interface	Egress Label
1	5	138.120	3	12

Ingress Interface	Ingress Label	FEC	Egress Interface	Egress Label
1	12	138.120	4	x

MPLS switch

Request 138.120

Request 138.120

Mapping 5

Mapping 12

127.20

MPLS switch

138.120

MPLS switch 192.168

MPLS switch
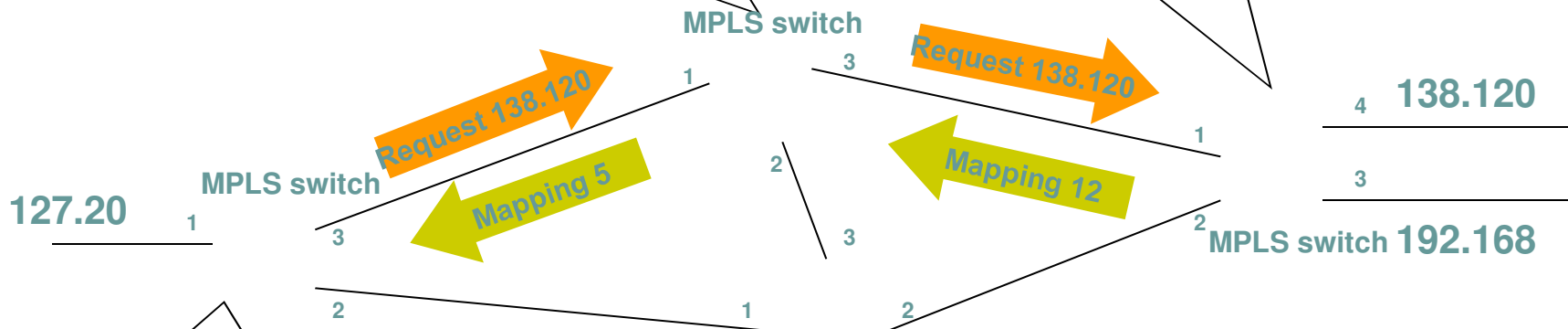
Ingress Interface	Ingress Label	FEC	Egress Interface	Egress Label
1	x	138.120	3	5

The label is requested by the upstream node and the downstream node defines the label used.

39

Distributia etichetelor (cont.)

- Independent
    - LSR asigneaza eticheta-FEC independent
    - transmite eticheta vecinilor
- Ordered
    - LSR asigneaza eticheta-FEC si transmite vecinilor daca este ruter egress pentru FEC sau a primit o asignare de la un next-hop

Stocarea etichetelor

- Liberal
  - etichete de la toti vecinii
  - timp de convergenta imbunatatit
  - necesita memorie (label space)
- Conservator
  - etichete de la vecini next-hop
  - se sterg toate etichetele pentru FECs fara nexthop

Protocoale de semnalizare LSP/ distributie etichete

- Resource Reservation Protocol (RSVP)
- Label Distribution Protocol (LDP)
- Constrained Routing LDP (CR-LDP)
- BGP

Label Distribution Protocol

- protocol de distributie a etichetelor
- construirea cailor comutate (Label Switched Path)
- etichetele asignate FECs pentru prefix de adrese destinatie de tip unicast
- LSP sint unidirectionale
- LDP opereaza asupra **perechi rutere** adiacente sau nu
- Mecanisme:
    - "descoperirea" ruterelor
    - mesaje: Discovery, Adjacency, Label Advertisment, Notification
    - nivel transport: TCP, UDP (Discovery)

LDP Messages

- **Discovery messages**

  - Used to discover and maintain the presence of new peers

  - Hello packets (tip UDP) sent to all-routers multicast address

  - Once neighbor is discovered, the LDP session is established over TCP

LDP Messages (cntd)

- **Session messages**
  - **Establish, maintain and terminate LDP sessions**

- **Advertisement messages**
  - **Create, modify, delete label mappings**

- **Notification messages**
  - **Error signalling**

Sesiuni LDP

- doua rutere in proxima vecinatate, cu capabilitati LDP, pot crea o sesiune LDP

  - Se foloseste UDP la nivel transport

  - periodic, mesaje Hello

  - Adr. nivel transport/sursa – identificator de sesiune

  - conexiune TCP (o singura conexiune)

  - LDP session Id

Sesiuni LDP (cont)

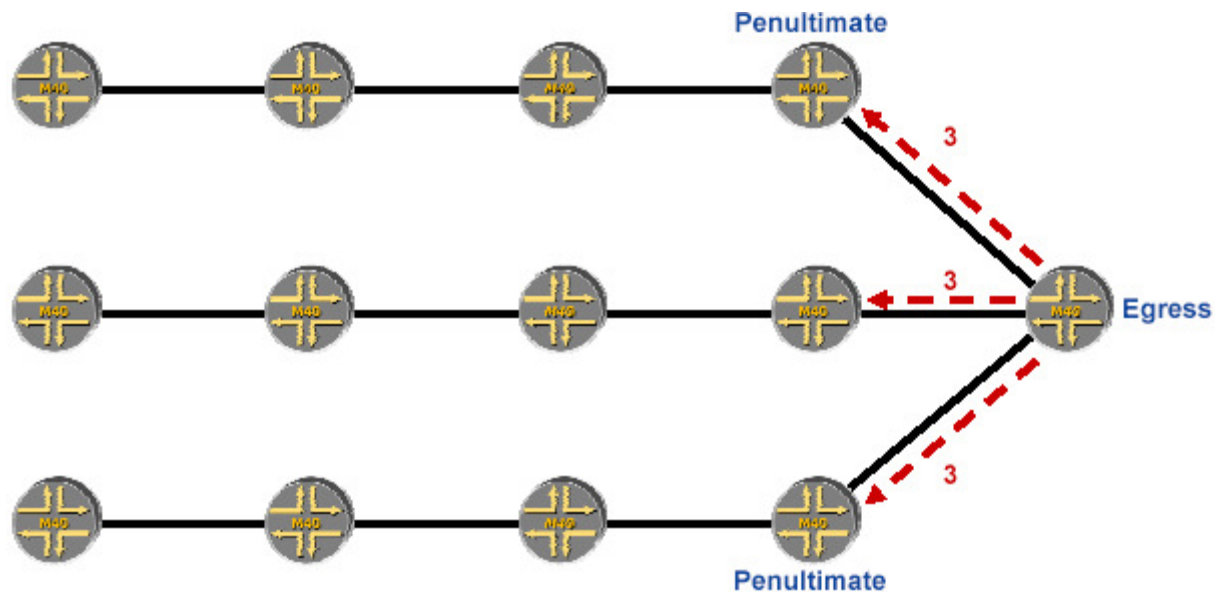- doua rutere la distanta pot forma, de asemenea, o sesiune LDP
  - UDP la nivel transport
  - msg Hello multicast
  - vecin descoperit => sesiune TCP
- mapare FEC-eticheta
- full-mesh al informatiilor: input-output mapping (eticheta - port)
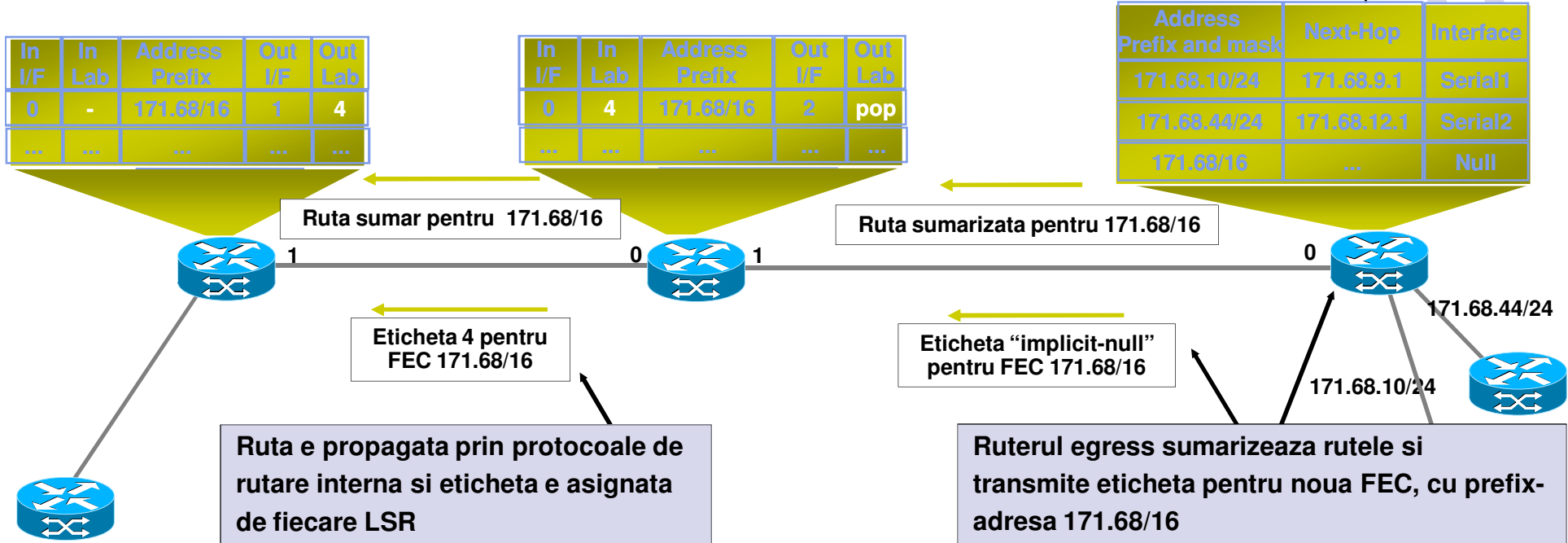
Penultimate hop popping

- eticheta stearsa in penultimul ruter
- ruterul egress semnalizeaza eticheta 3 (Implicit NULL Label)
- decongestie ruter egress

Label Switch Path (LSP) Penultimate Hop Popping

In I/F	In Lab	Address Prefix	Out I/F	Out Lab
0	-	171.68/16	1	4
...	...	...	...	...

In I/F	In Lab	Address Prefix	Out I/F	Out Lab
0	4	171.68/16	2	pop
...	...	...	...	...

Address Prefix and mask	Next-Hop	Interface
171.68.10/24	171.68.9.1	Serial1
171.68.44/24	171.68.12.1	Serial2
171.68/16	...	Null

Ruta sumar pentru 171.68/16

Ruta sumarizata pentru 171.68/16

Eticheta 4 pentru FEC 171.68/16

Eticheta "implicit-null" pentru FEC 171.68/16

171.68.44/24

171.68.10/24

Ruta e propagata prin protocoale de rutare interna si eticheta e asignata de fiecare LSR

Ruterul egress sumarizeaza rutele si transmite eticheta pentru noua FEC, cu prefix-adresa 171.68/16

Ruterele egress executa 'IP routing table lookup' pentru a gasi rutele specifice (171.68.44/24 de ex.)

Ruterele egress pot sa nu primeasca pachete cu etichete, pentru a salva un lookup
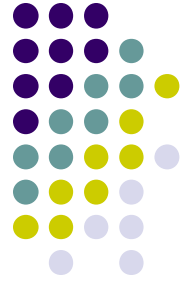
49

Ultimate hop popping

- eticheta stearsa in ultimul ruter
- ruterul egress semnalizeaza eticheta 0 (Explicit NULL)

Agenda

- De ce MPLS?
- Conceptul si terminologia MPLS
- Operatii MPLS
- Concluzii

Operatii MPLS

1a. Protocoale de rutare (OSPF, EIGRP) stabilesc topologia retelei

1b. Protocoale de semnalizare pentru asignarea si distributia etichetelor

4. Ruterele de iesire sterg eticheta si ruteaza pachetul spre destinatie
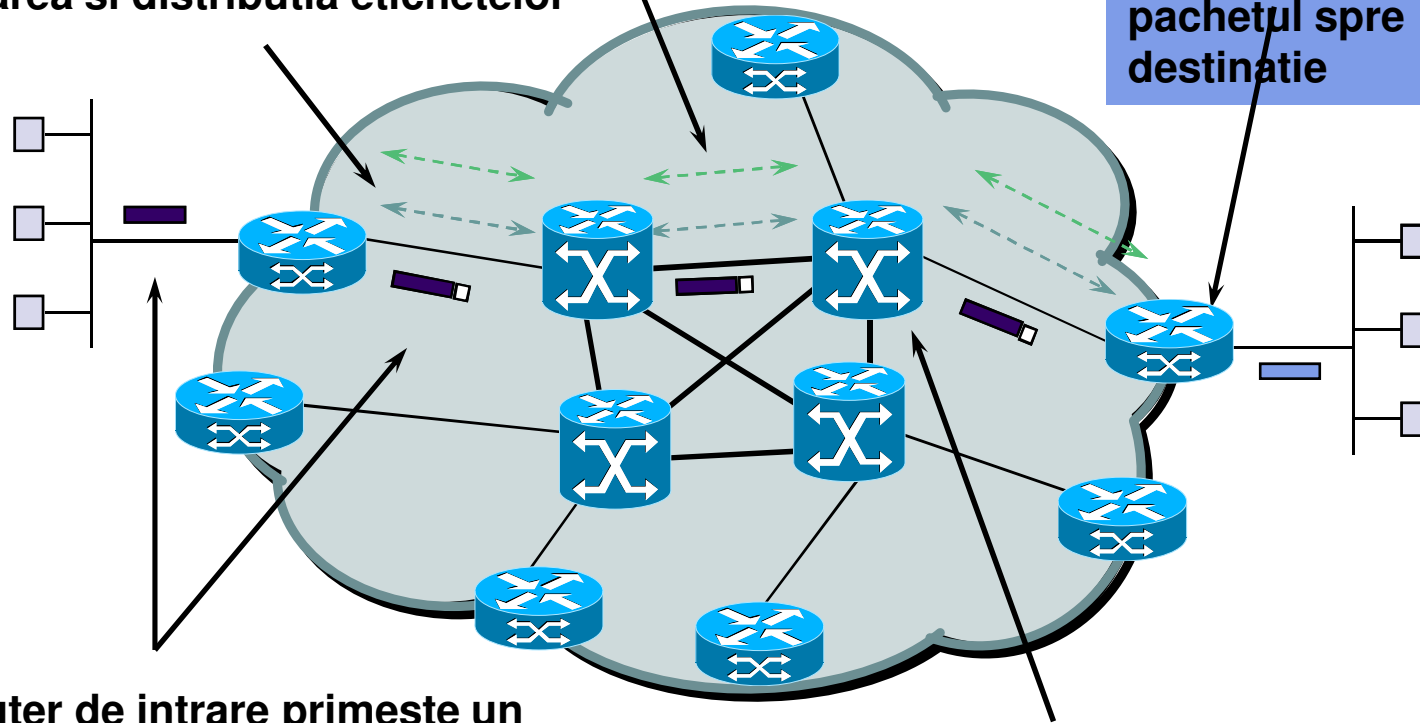
2. Un ruter de intrare primeste un pachet, evalueaza serviciile de care are nevoie, asigneaza FEC, eticheteaza pachetul

3. Ruterele tranzit comuta pachetele pe baza etichetelor

52

Exemplu Informatii de rutare

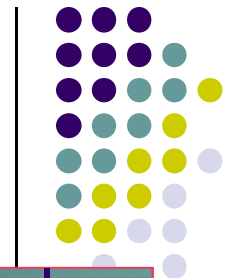| In Label	Address Prefix	Out I'face	Out Label
	128.89	1	
	171.69	1	
	...	...	

In Label	Address Prefix	Out I'face	Out Label
	128.89	0	
	171.69	1	
	...	...	

In Label	Address Prefix	Out I'face	Out Label
	128.89	0	
	...	...	

1

0

0   **128.89**

poti ajunge la 128.89 prin mine

poti ajunge la 128.89 si 171.69 prin mine

1

actualizarea rutelor (OSPF, EIGRP, IS-IS)

poti ajunge la 171.69 prin mine
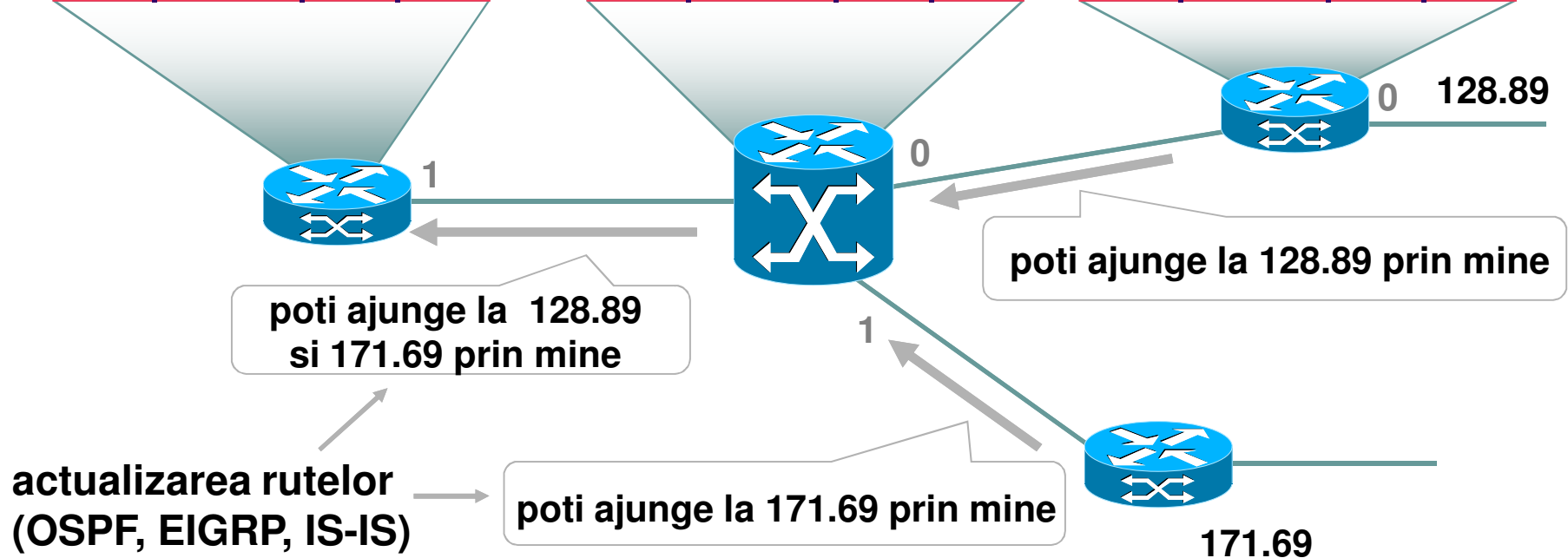
171.69

Exemplu:
Asignarea etichetelor

In Label	Address Prefix	Out I'face	Out Label
-	128.89	1	4
-	171.69	1	5
...	...	...	...

In Label	Address Prefix	Out I'face	Out Label
4	128.89	0	9
5	171.69	1	7
...	...	...	...

In Label	Address Prefix	Out I'face	Out Label
9	128.89	0	-
...	...	...	...

1

0

0   **128.89**

utilizeaza eticheta 9 pt 128.89

utilizeaza eticheta 4 pt 128.89
si eticheta 5 pt 171.69

1

**distributie etichete
(RSVP, LDP)**

utilizeaza eticheta 7 pt 171.69

171.69

54

Exemplu: Comutarea pachetelor

In Label	Address Prefix	Out I'face	Out Label
-	128.89	1	4
-	171.69	1	5
...	...	...	...

In Label	Address Prefix	Out I'face	Out Label
4	128.89	0	9
5	171.69	1	7
...	...	...	...

In Label	Address Prefix	Out I'face	Out Label
9	128.89	0	-
...	...	...	...

128.89

128.89.25.4	Data

9	128.89.25.4	Data

0

1

128.89.25.4	Data

4	128.89.25.4	Data

1

LSR comuta pachetele pe baza etichetei

171.69

MPLS Example: More Details

In Label	Address Prefix	Out I'face	Out Label
7	128.89	1	4
2	171.69	1	5
7	117.59	1	4

In Label	Address Prefix	Out I'face	Out Label
4	128.89	0	X
5	171.69	1	7
4	117.59	0	X

In Label	Address Prefix	Out I'face	Out Label
X	128.89.25	0	-
X	117.59	1	-
...	...	...	...

128.89.25

128.89.25.4	Data

128.89.25.4	Data

7	128.89.25.4	Data

4	128.89.25.4	Data

117.59
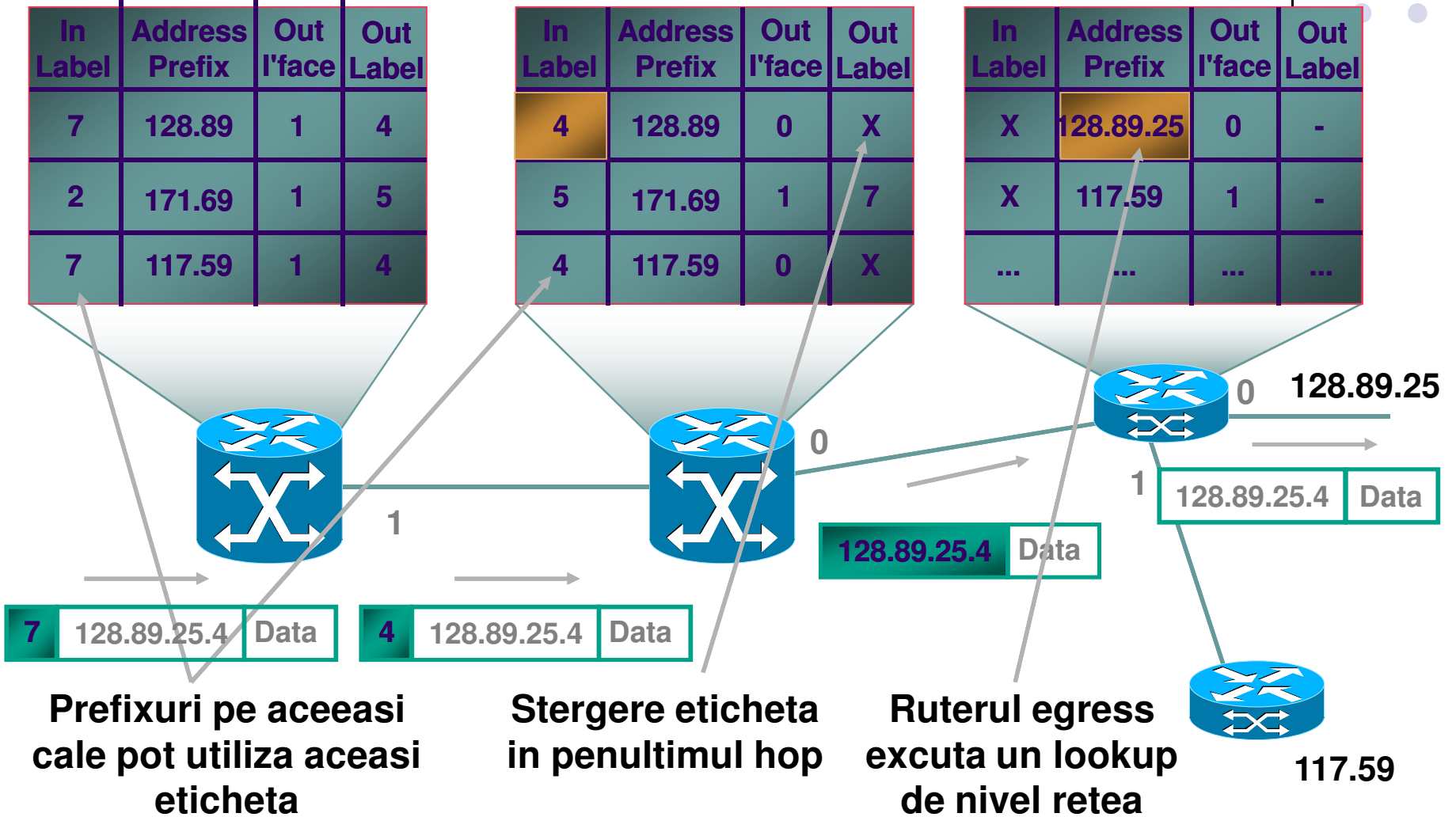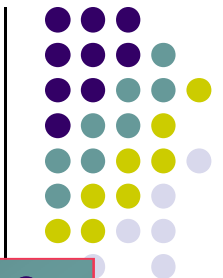
Prefixuri pe aceeasi cale pot utiliza aceasi eticheta

Stergere eticheta in penultimul hop

Ruterul egress excuta un lookup de nivel retea

Buclarea in retele MPLS

- In IP networks Time-To-Live (TTL) is used to prevent packets to travel indefinitely in the network

- MPLS _may_ use same mechanism as IP, but not on all encapsulations

  - TTL is present in the label header for PPP and LAN headers (shim headers)

  - ATM cell header does not have TTL

- Some suggested options:

  - hop-count object in LDP

  - Path Vector object in LDP

Loops and TTL
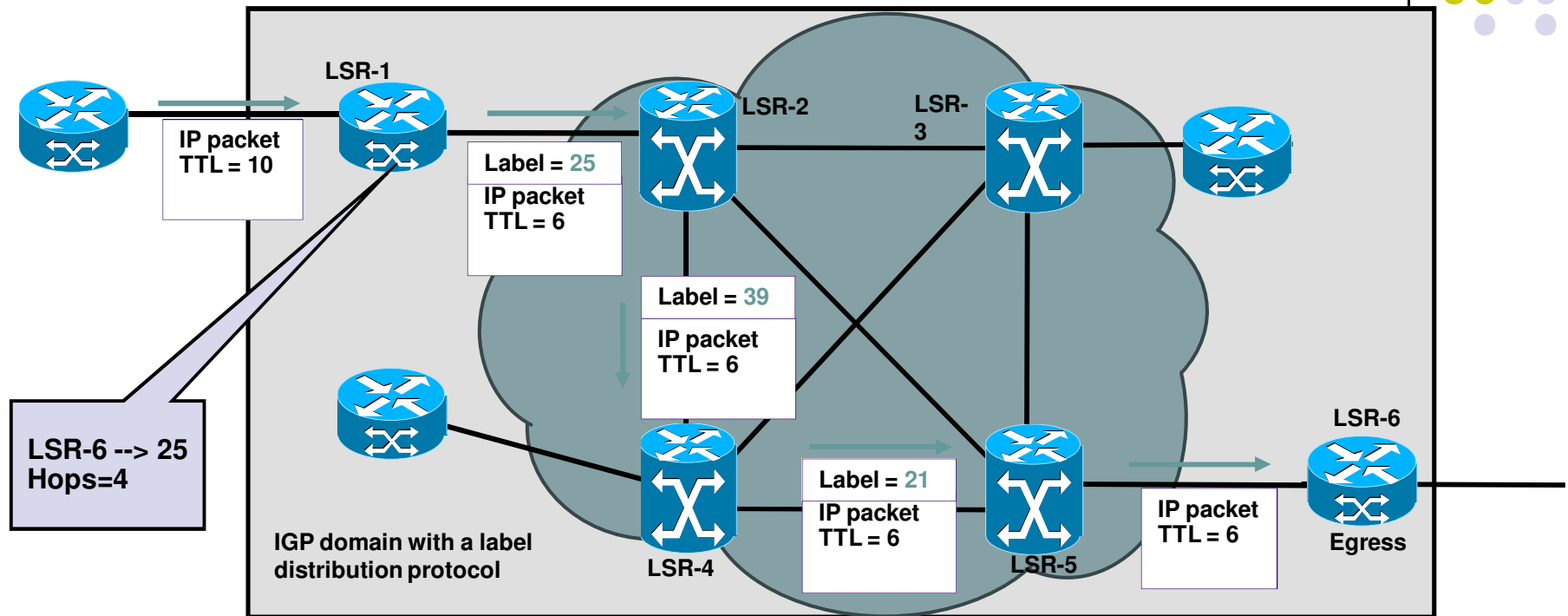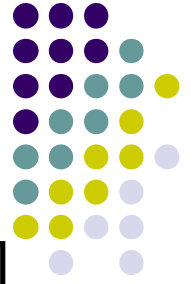
- **TTL is decremented prior to enter the non-TTL capable Label Switching Path (LSP)**

  If TTL becomes 0, the packet is discarded at the ingress point

- **TTL is examined at the LSP exit**

Concluzii

- LSR – forward pachete pe baza etichetei
- Eticheta: protocoale de rutare IP, traffic engineering, QoS, VPN
- Decupleaza rutarea de forwardare
- Clasificarea pachetelor
- Protocoale de distributie etichete: RSVP, LDP, CR-LDP pot coexista
- Etichetele au semnificatie locala