MPLS

A Tutorial

Paresh Khatri

paresh.khatri@alcatel-lucent.com.au

Agenda

1. MPLS overview and LSP types
2. Label Distribution Protocol (LDP)
3. Questions

Introduction

Paresh Khatri (paresh.khatri@alcatel-lucent.com.au)

- Director – Advanced Consulting Engineering, APAC IP Competence Centre, ALU
- 12 years of design/architecture experience, including 6 years with Service Providers
- Key focus areas:
  - Large-scale IP/MPLS networks
  - L2/L3 VPNs
  - Carrier Ethernet
  - Next-generation mobile backhaul networks

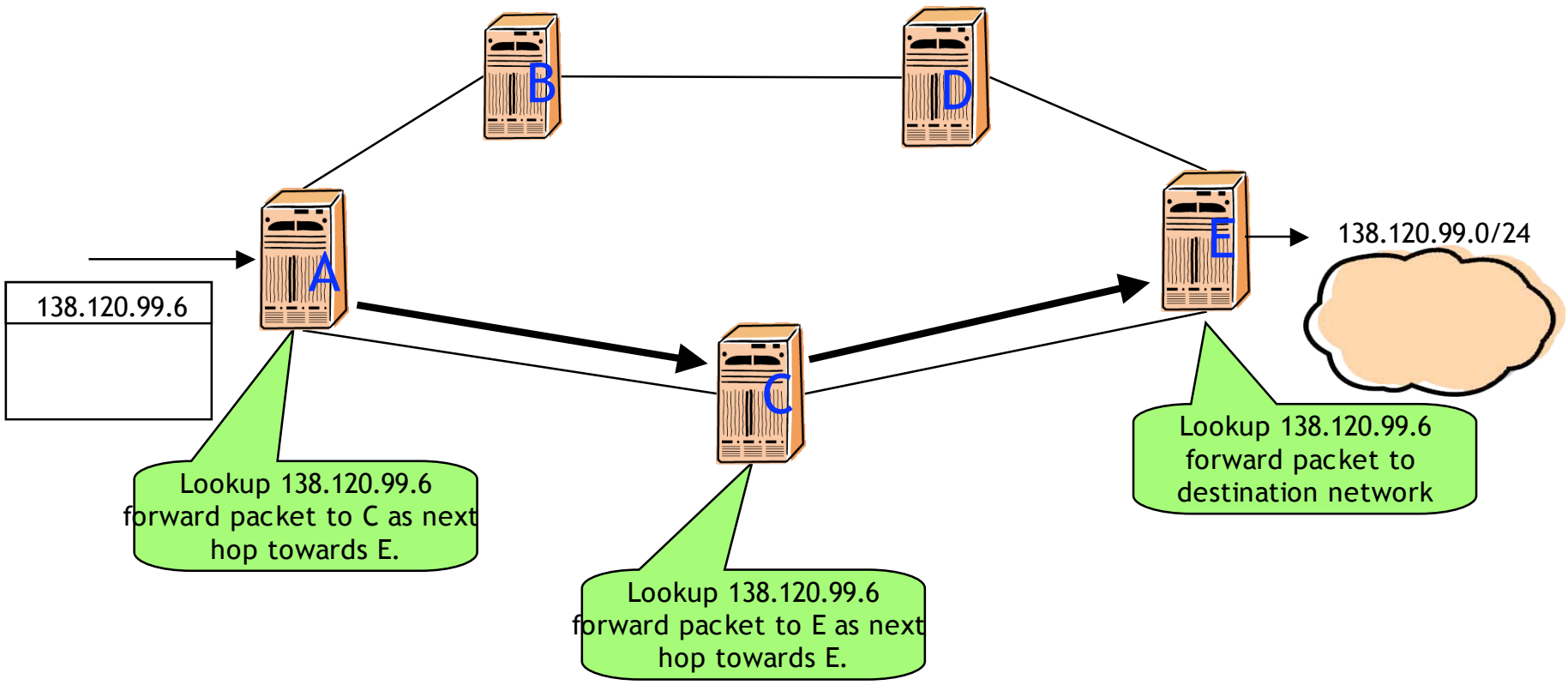
1 MPLS overview and LSP types

Agenda

1. MPLS overview and LSP types
    - 1.1 Motivation for MPLS
    - 1.2 Label Operations
    - 1.3 MPLS header
    - 1.4 LSP Types

Motivation: Classical Connectionless IP Networks

- Packets are forwarded, hop-by-hop, with no ordering or delivery guarantees.
- An independent forwarding decision is made at each hop.



Motivation: Classical Connectionless IP Network

- FEC represents *"a group of packets which are forwarded in the same manner (e.g., over the same path, with the same forwarding treatment)"*
- Each hop re-examines the packet's network layer header and assigns it to a FEC
- Choosing the next hop is a composition of two functions:
  - Partitioning the stream of packets into FECs
  - Assigning each FEC to a next hop
- Routing protocols are used to figure out:
  - what addresses are reachable in a network
  - the best path to use to reach an address
  - OSPF,ISIS,BGP are all well established and understood routing protocols used in IP networks

What is MPLS?

- MPLS stands for Multiprotocol Label Switching
- Assignment of a particular packet to a particular FEC is done just once, as the packet enters the network
- Packets are “labeled” before they are forwarded to the next hop
- All forwarding is driven by labels
- No further analysis of the packet’s network layer header at subsequent hops.

  Label is used as an index into a table which specifies the next hop and a new label. The old label is swapped with the new label and the packet is forwarded to its next hop

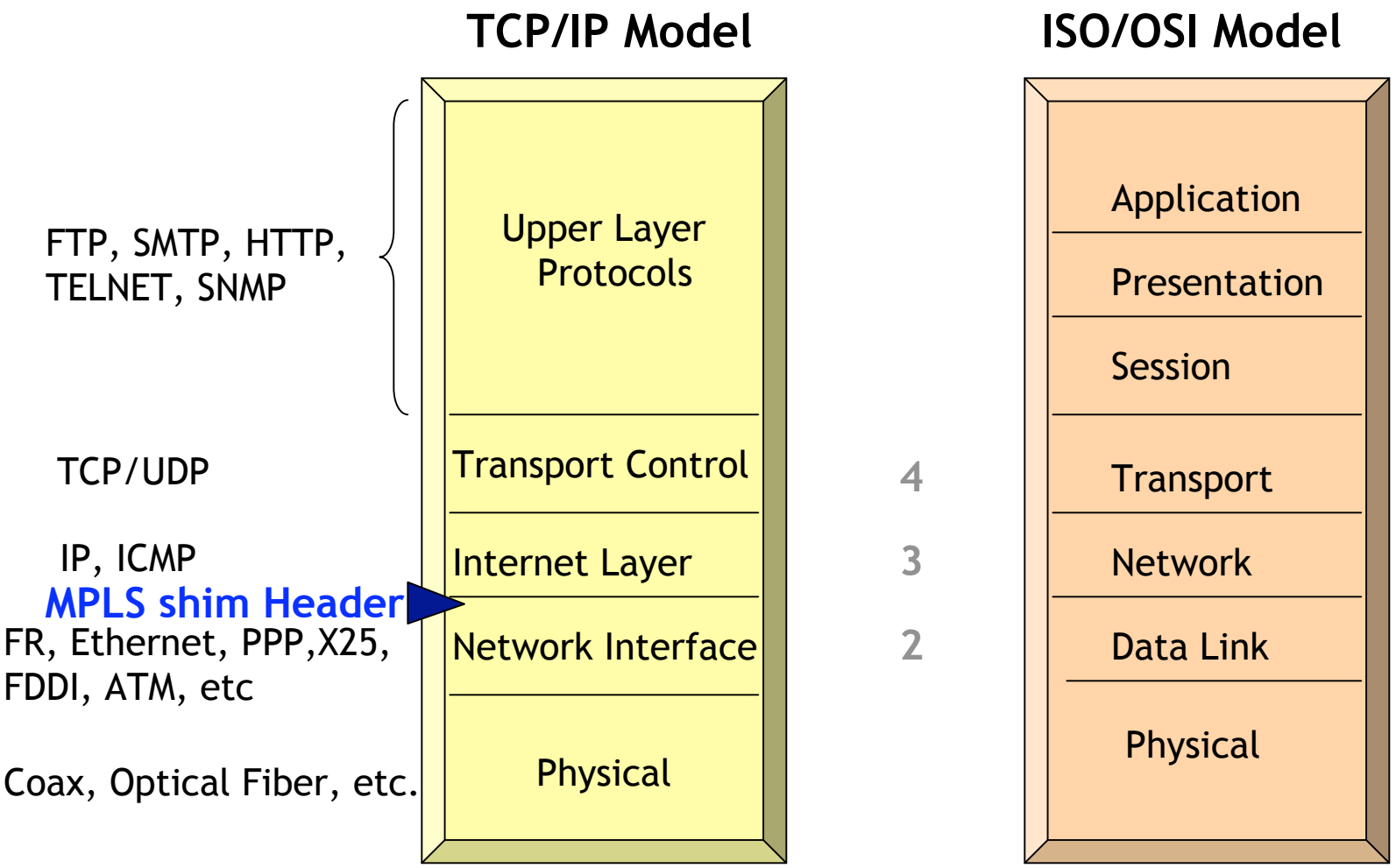
MPLS Advantages

MPLS brings the following benefits to IP networks:

- Integration of L2/L3 & Traffic engineering
  - The ability to set the path that traffic will take through the network
  - The ability to set performance characteristics for a class of traffic.
- Layer 2 transport

  New standards that allow service providers to carry Layer 2 services including Ethernet, Frame Relay, ATM and TDM over an IP/MPLS core.

Layered Model



MPLS Terminology

FEC - Forwarding Equivalence Class: "A subset of packets that are treated the same way by a router"

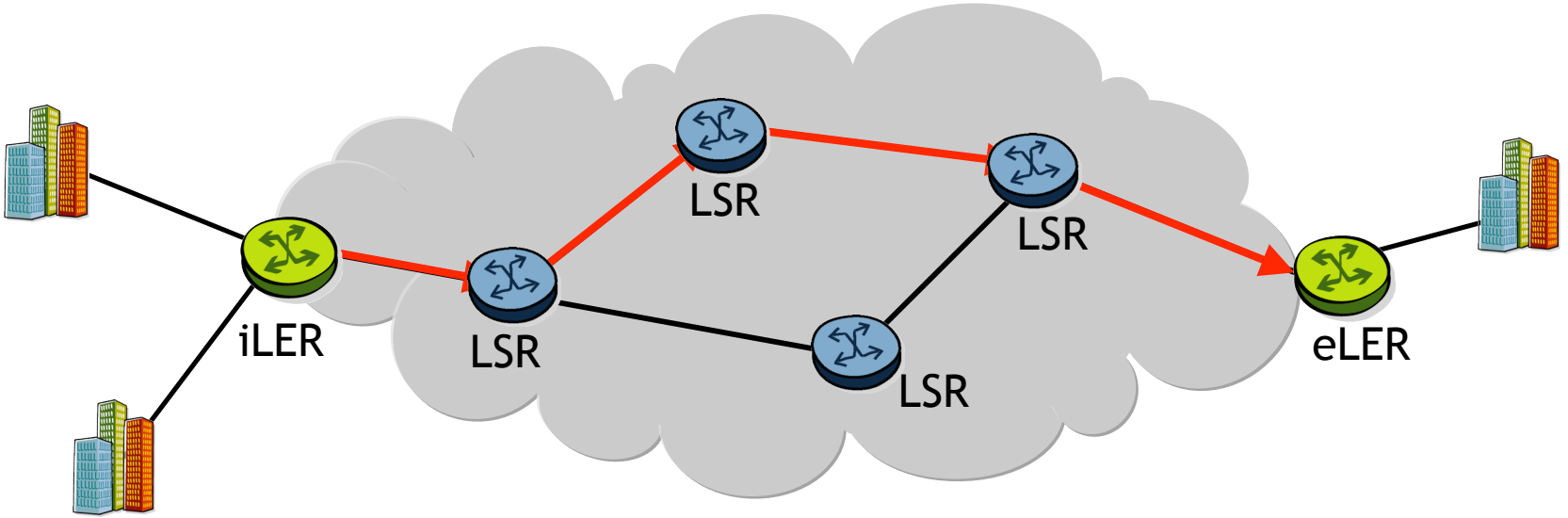
In IP routing, a packet is assigned to a FEC at each hop (i.e. L3 look-up).
In MPLS it is only done once at the network ingress.

iLER – ingress Label Edge Router that pushes MPLS label.

LSR - Label Switching Router, swaps MPLS label.

eLER – egress Label Edge Routes that pops MPLS Label.

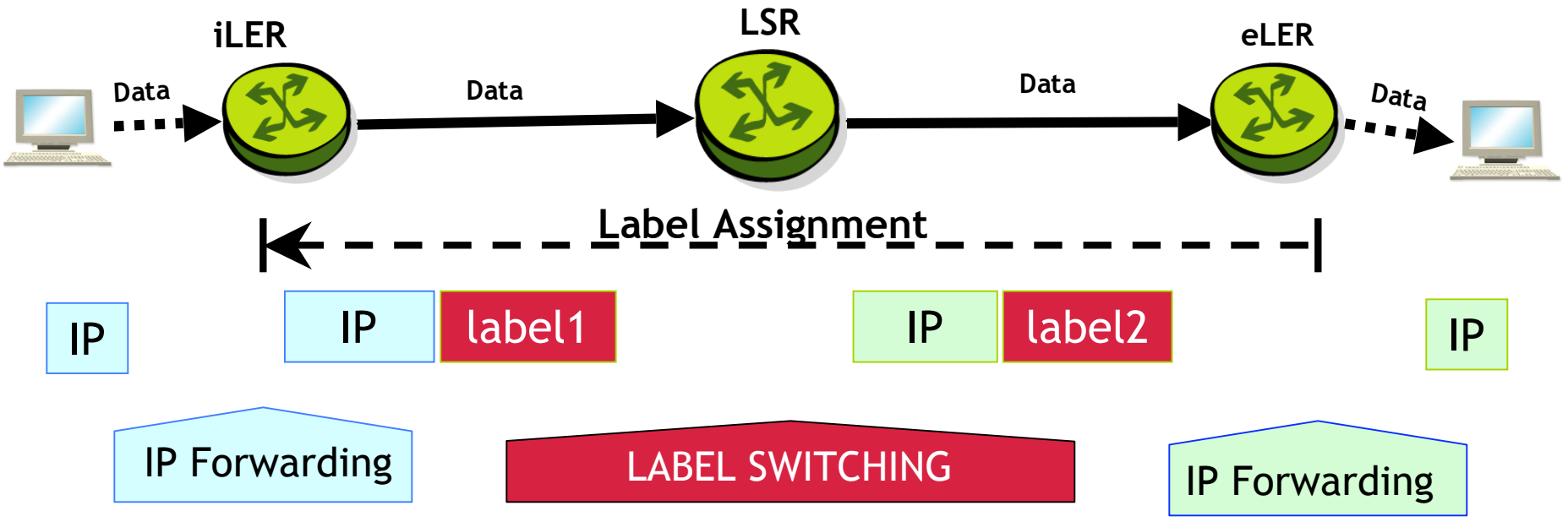
LSP - Label Switched Path. Path that MPLS Packets follow from iLER to eLER.



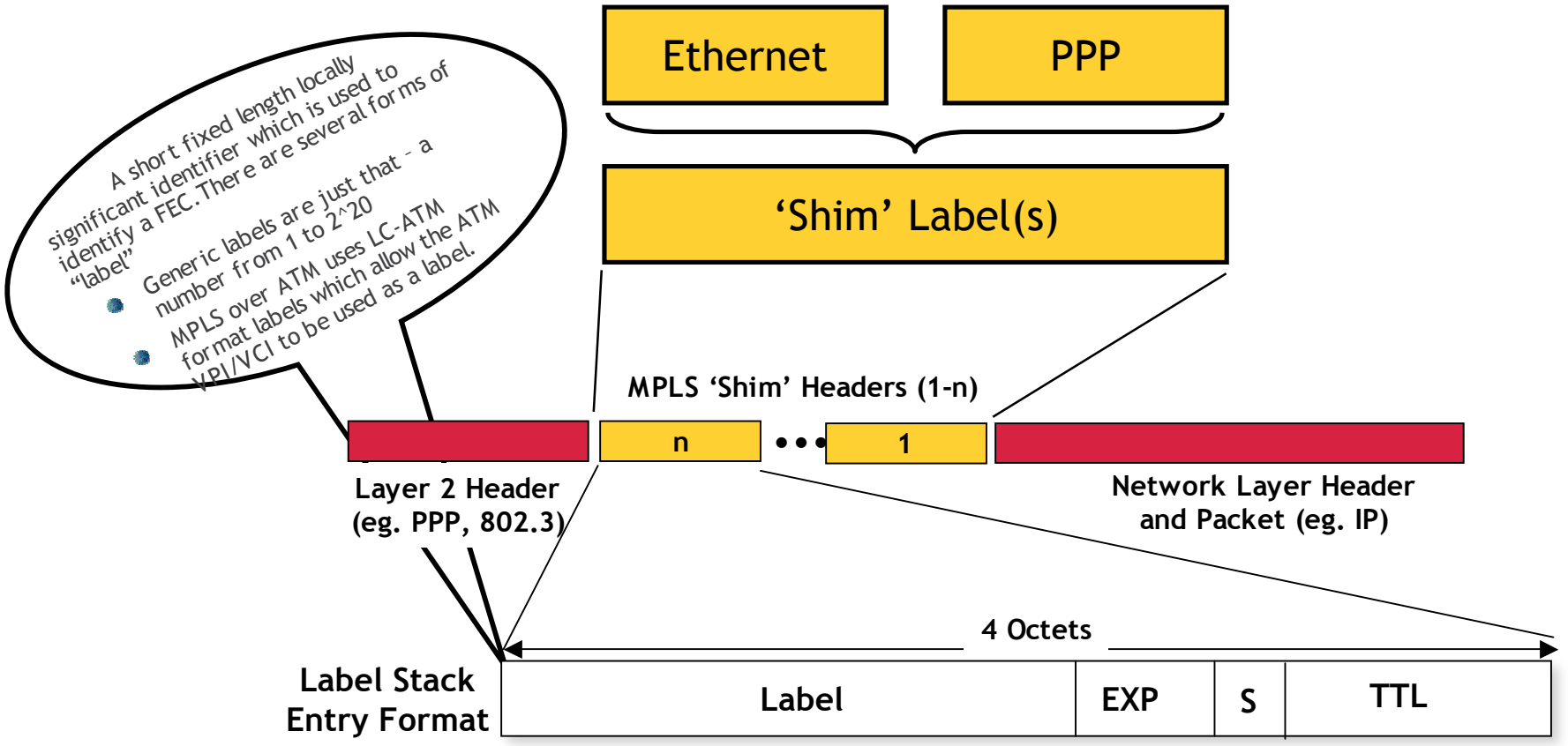
LSP Path that MPLS Packet follow

Label Operations

- LSP establishment either statically or dynamically using signalling
- Ingress LER pushed label1 to an IP packet
- Transit LSR swapped the label1 with label2
- Egress LER or Penultimate LSR popped off label2



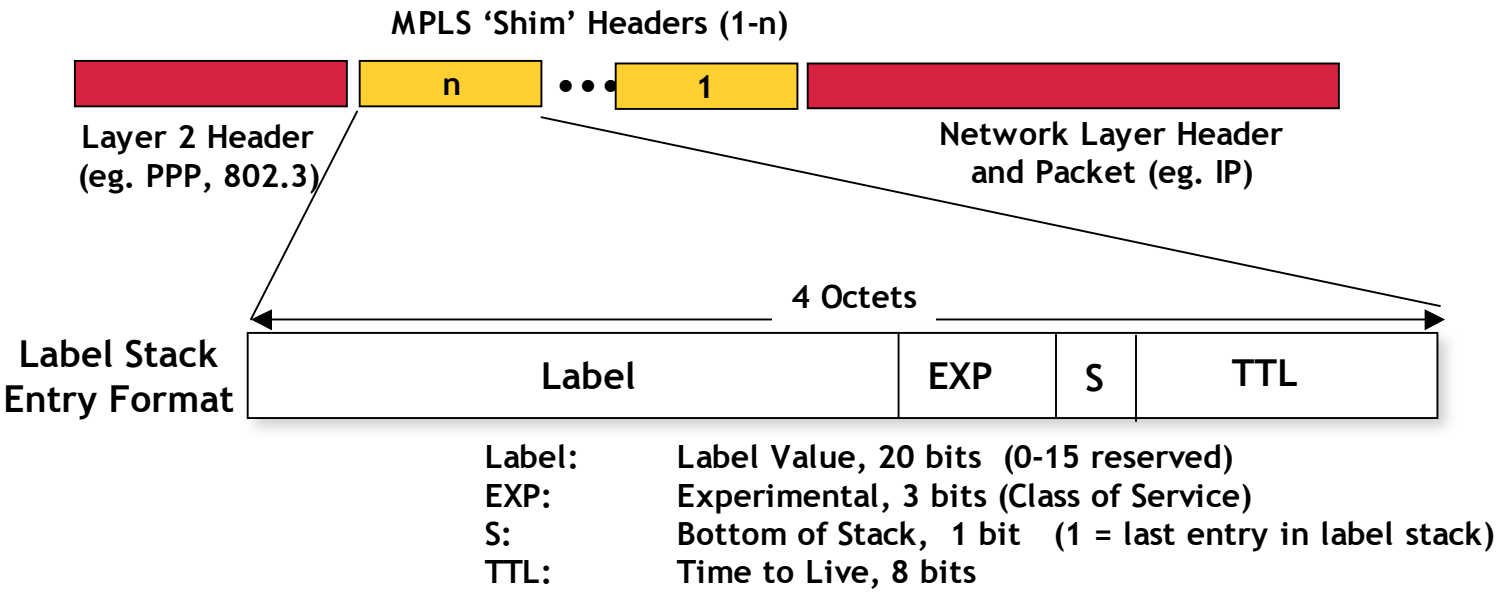
MPLS Packet Encapsulation & MPLS Header Fields



Packet-based encoding

Label: Label Value, 20 bits (0-15 reserved)
Exp.: Experimental, 3 bits (Class of Service)
S: Bottom of Stack, 1 bit (1 = last entry in label stack)
TTL: Time to Live, 8 bits

MPLS Header Field: EXP bits



- Various kinds of VLLs/traffic types are to be multiplexed over an LSP
- Exp bits handles the prioritization of mission critical traffic
- The 3 EXP bits designate the importance of a particular frame
- Classification/queuing at the P/PE routers usually takes place based on these EXP bits.

MPLS Header Fields: Label Allocation

Label values 0 – 15 reserved for MPLS use

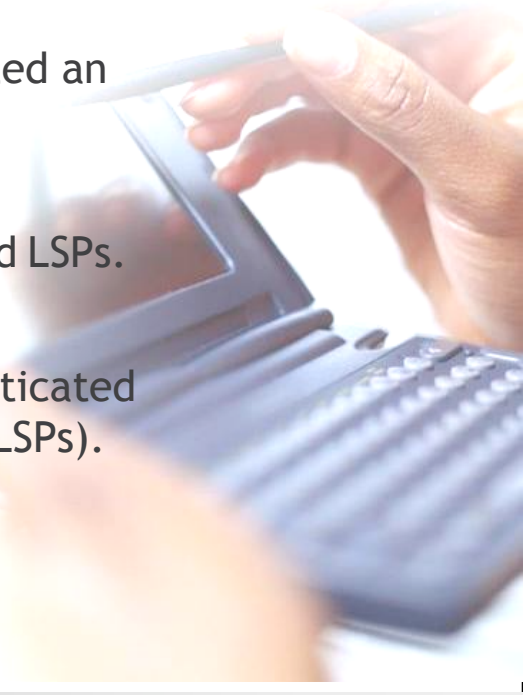
- **Value of 0 represents "IPv4 Explicit NULL label"**
  - **Value is only legal at the bottom of the stack**
  - **Implies that label must be popped & forwarding based on IP header**
- **Value of 1 represents Router Alert Label**
  - **Cannot be at the bottom of the stack**
- **Value of 2 represents "IPv6 Explicit NULL label"**
- **Value of 3 represents "Implicit NULL label"**
- **Values 4 – 15 reserved forMPLS future use**

```
B:TECSIM1>config>router>mpls# show router mpls label-range
===============================================================================
Label Ranges
===============================================================================
Label Type      Start Label    End Label      Aging        Total Available
-------------------------------------------------------------------------------
Static-lsp      32             1023           -            991
Static-svc      2048           18431          -            16384
Dynamic         32768          131071         0            98300
```

Used for services

Label Hierarchy

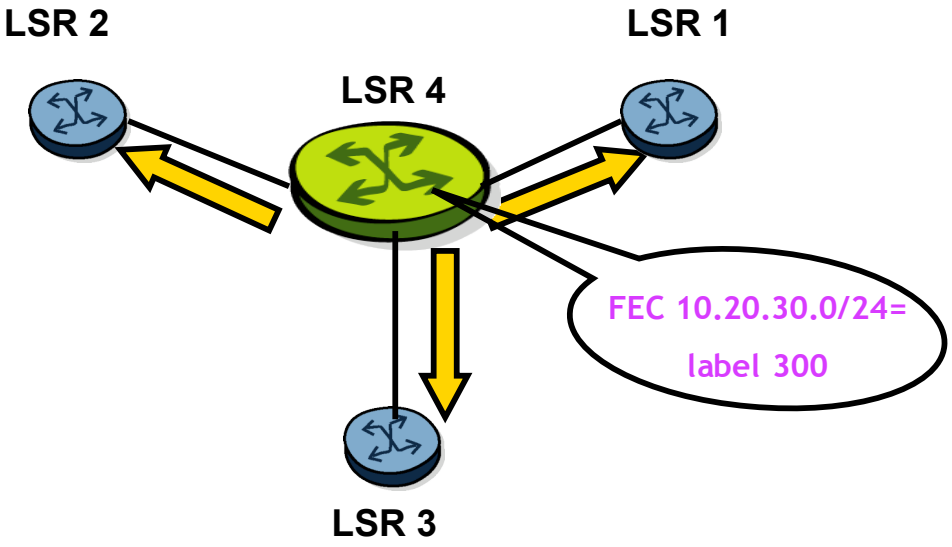
- Labels can be stacked on top of each other to form a kind of hierarchy of labels.
- When a destination LSR pops a label and finds another label, it must act on the new label, possibly swapping or maybe popping again.
- An LSP can traverse a network through another LSP, which is called an “LSP Tunnel” (e.g., pseudowires over MPLS tunnels).
- An LSP tunnel can and would be expected to carry many tunneled LSPs.
- LSP tunnels provide a mechanism to implement some very sophisticated path redundancy and reroute mechanism (not applicable to LDP LSPs).



MPLS Header Fields: Types of Label Space

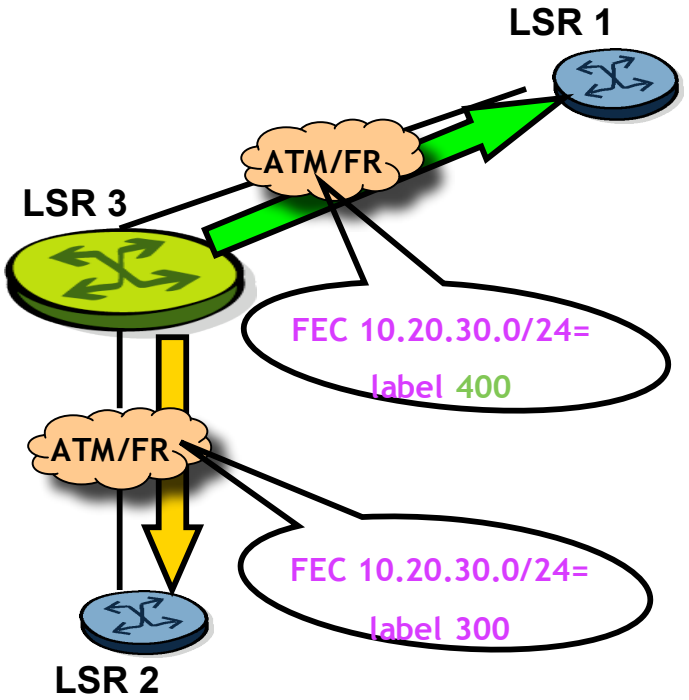
Per Platform label space

- One label per FEC per device.



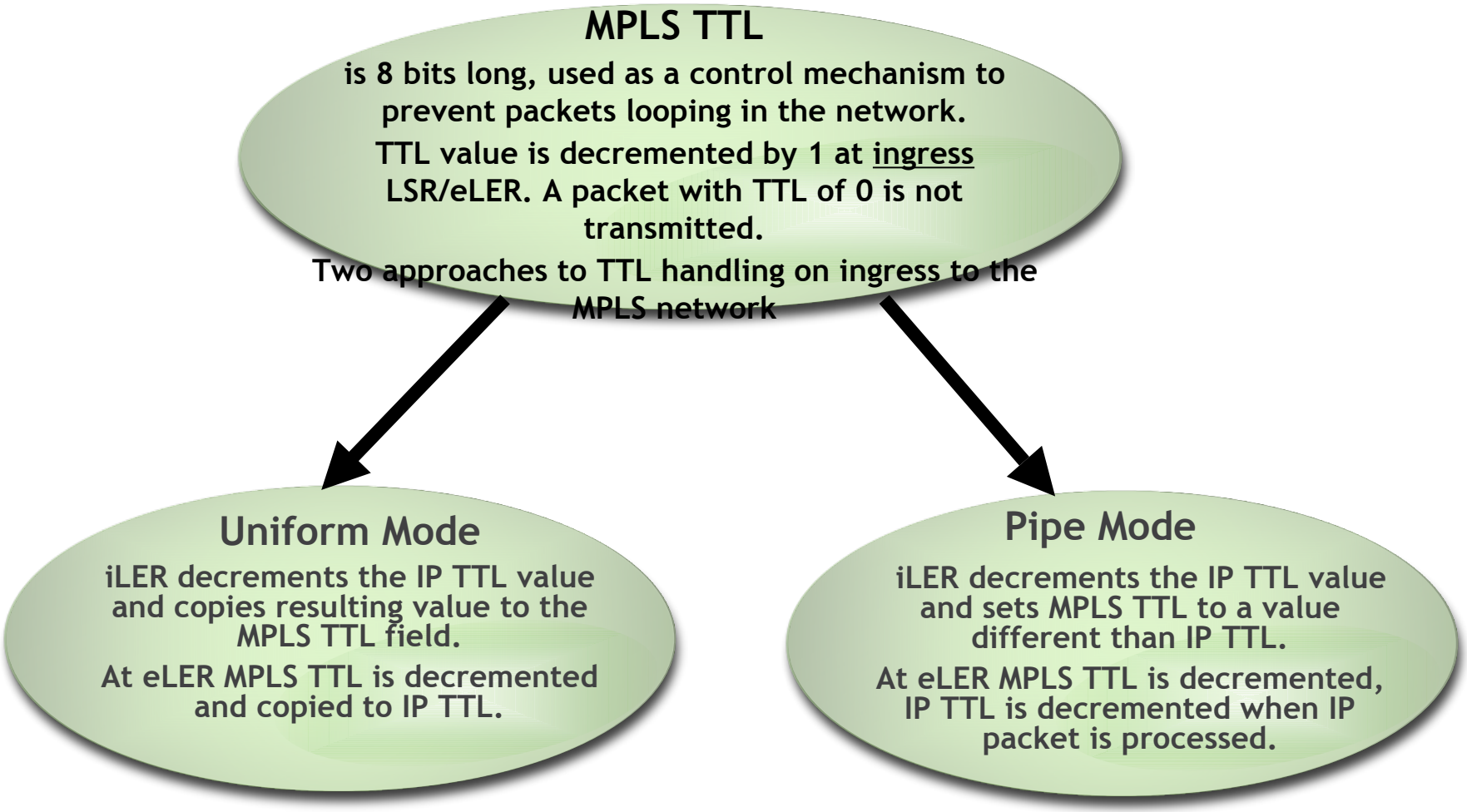
Per Interface label space

- One label per FEC per interface per device.



- Platform wide incoming labels are used for interfaces that can share the same labels. A single label is assigned to a FEC and used across all interfaces of the same router.
- Interface specific incoming labels are used for interfaces that can use interface resources for label. A separate label is used for each interface that the FEC is advertised on.

MPLS Header Fields: Time-to-Live (TTL)



Label/TTL Operations

Push (iLER)

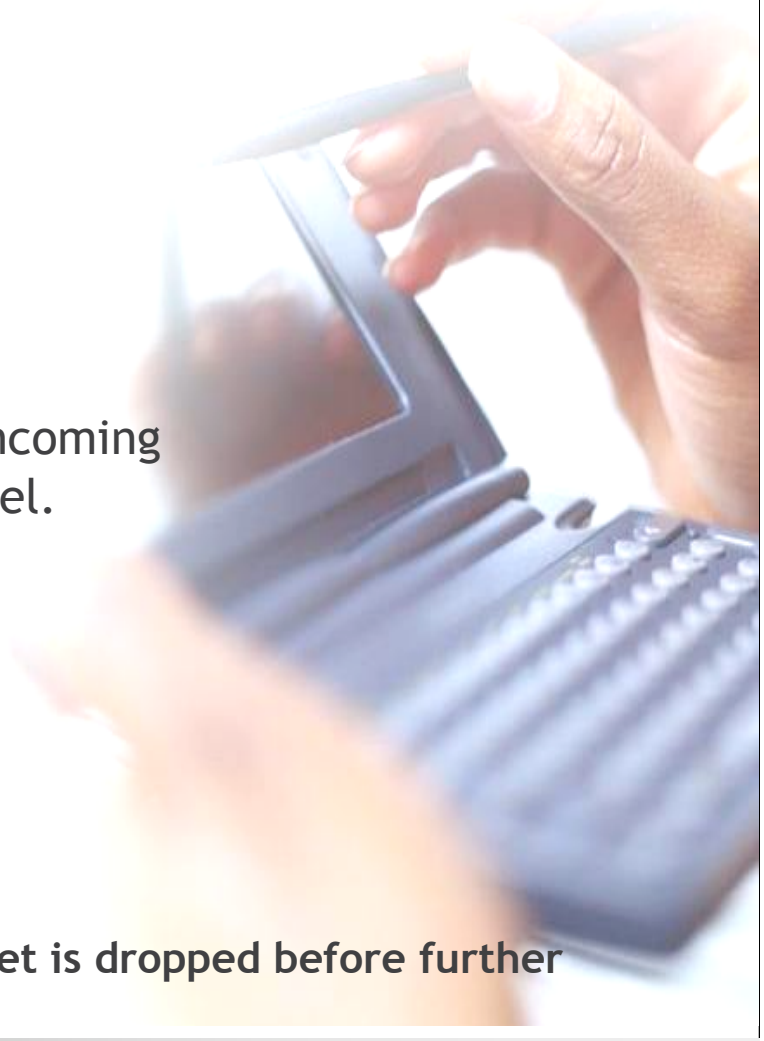
- Push the first label on the packet or
- Push a label on existing label stack
- Set the MPLS TTL value to 255

Swap (LSR)

- Combination of POP and PUSH operation
- At the ingress, decrement TTL value of incoming label and copy it to the new outgoing label.

Pop (eLER)

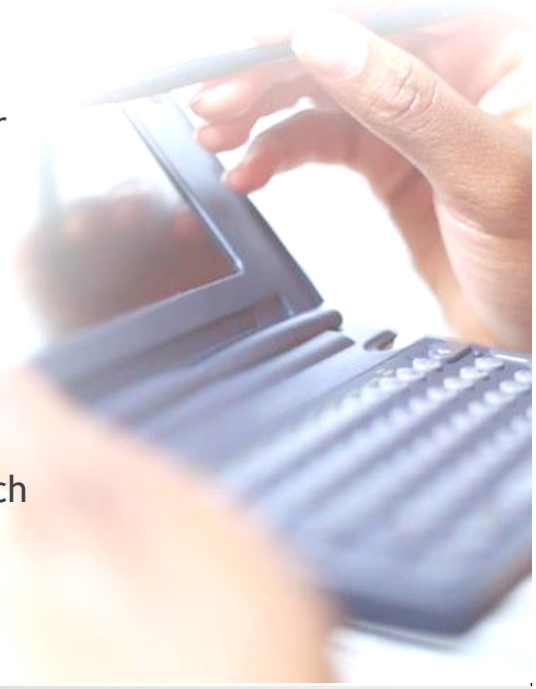
- At the ingress, decrement TTL value.
- Remove the top label from the packet
- Decrement the TTL value of the IP Packet



At any time TTL value hits 0, the MPLS or IP packet is dropped before further processing.

Label Switched Paths - LSPs

- MPLS directs a flow of IP traffic along a labeled switched path (LSP).
- An LSP is the sequence of LSRs that pushes or swaps labels on a packet as it traverses a network.
- An LSP has attributes such as traffic reservations, link color, explicit routes, hold/setup priorities and others depending on the label distribution protocol (LDP or RSVP-TE).
- An LSP has a state: up or down. When an LSP is up, the labels assigned for that LSP are used to label packets that will travel on the LSP.
- An LSP may reroute as a result of topological changes, it may reduce or increase its resource reservation, it may go down and come back up.
- A single LSP is UNIDIRECTIONAL. Data travels in one direction only along the path. For bidirectional traffic, two LSP's are required – one LSP in each direction.



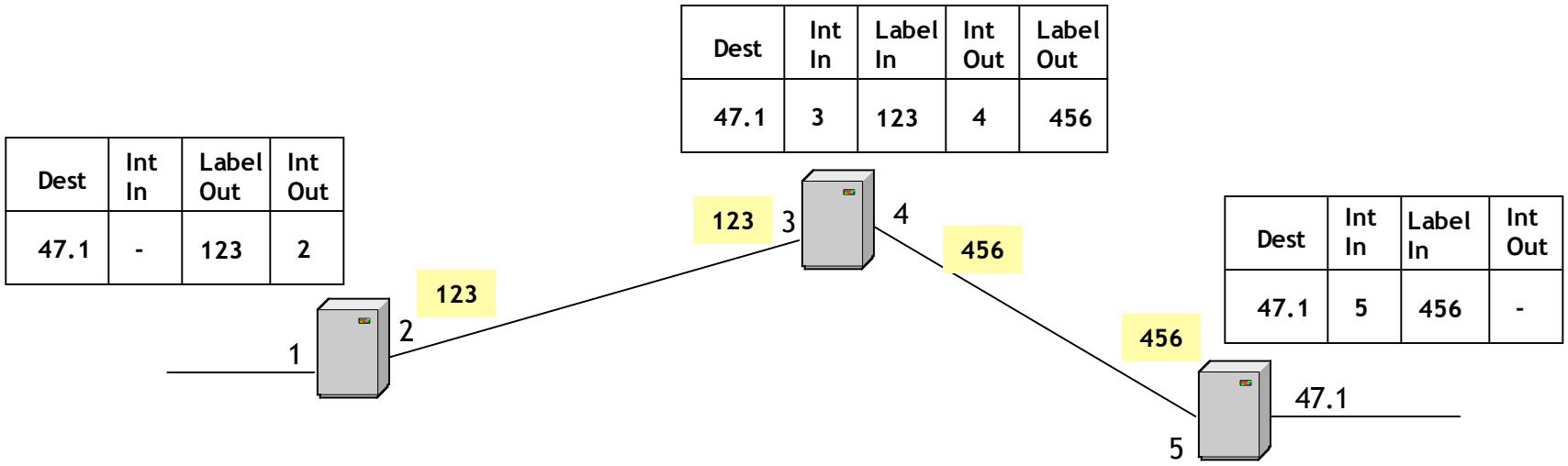
LSP Types

- **Static LSP:**
  - Specifies a static path i.e. no signalling RSVP/LDP is required.
  - Label are assigned manually by operator.
  - Configured manually at each router (ingress, transit, egress).
  - No dependence on IGP or local forwarding table.
  - No failure detection and no rerouting capabilities.

- **Signalled LSP:**
  - Are setup using RSVP-TE or LDP signalling protocols.
  - Signalling allows labels to be assigned automatically from ingress router to egress router.
  - Configuration is required only on the ingress/egress routers.
  - Dependant on IGP and local forwarding table.
  - Various protection techniques & FRR.

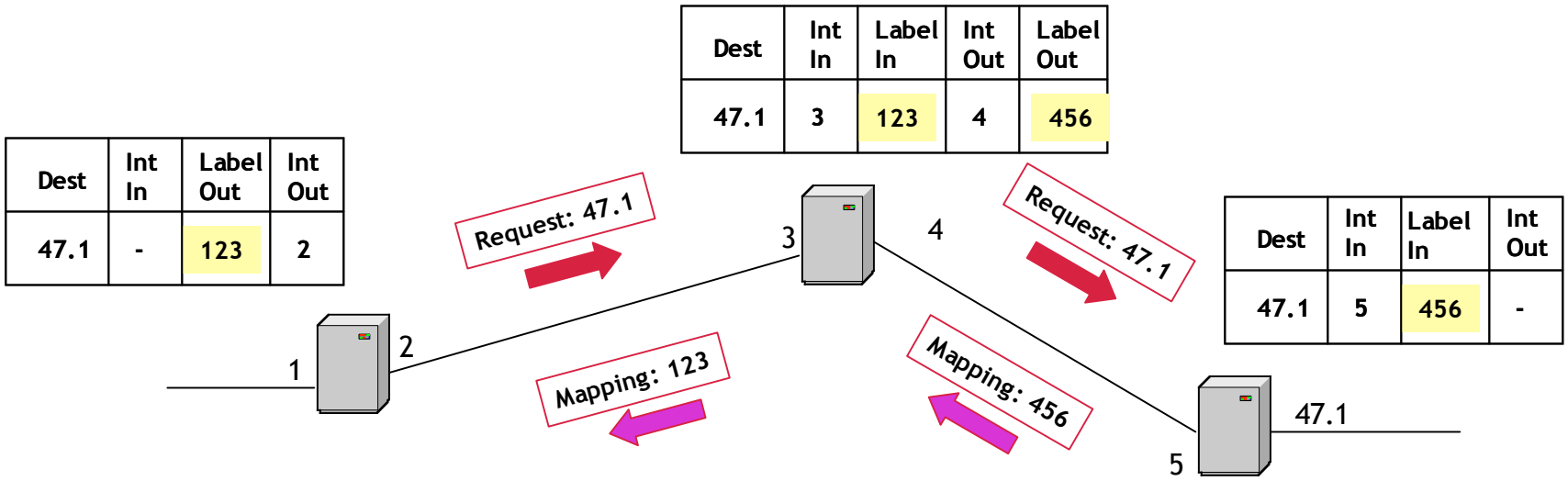
LSP Types: Static LSP

- All Routers are configured manually with labels.
- No signaling is required.



LSP Types : Signaled LSP

- LSPs are setup using a signaling protocol: LDP or RSVP-TE
- Signaling protocol facilitates:
  - Assignment of labels from egress router to the ingress router.
  - Signaling is triggered by the ingress/egress routers – no configuration is required on intermediate routers.
  - Path Selection



Motivation: MPLS Signalling and Label Distribution Protocols

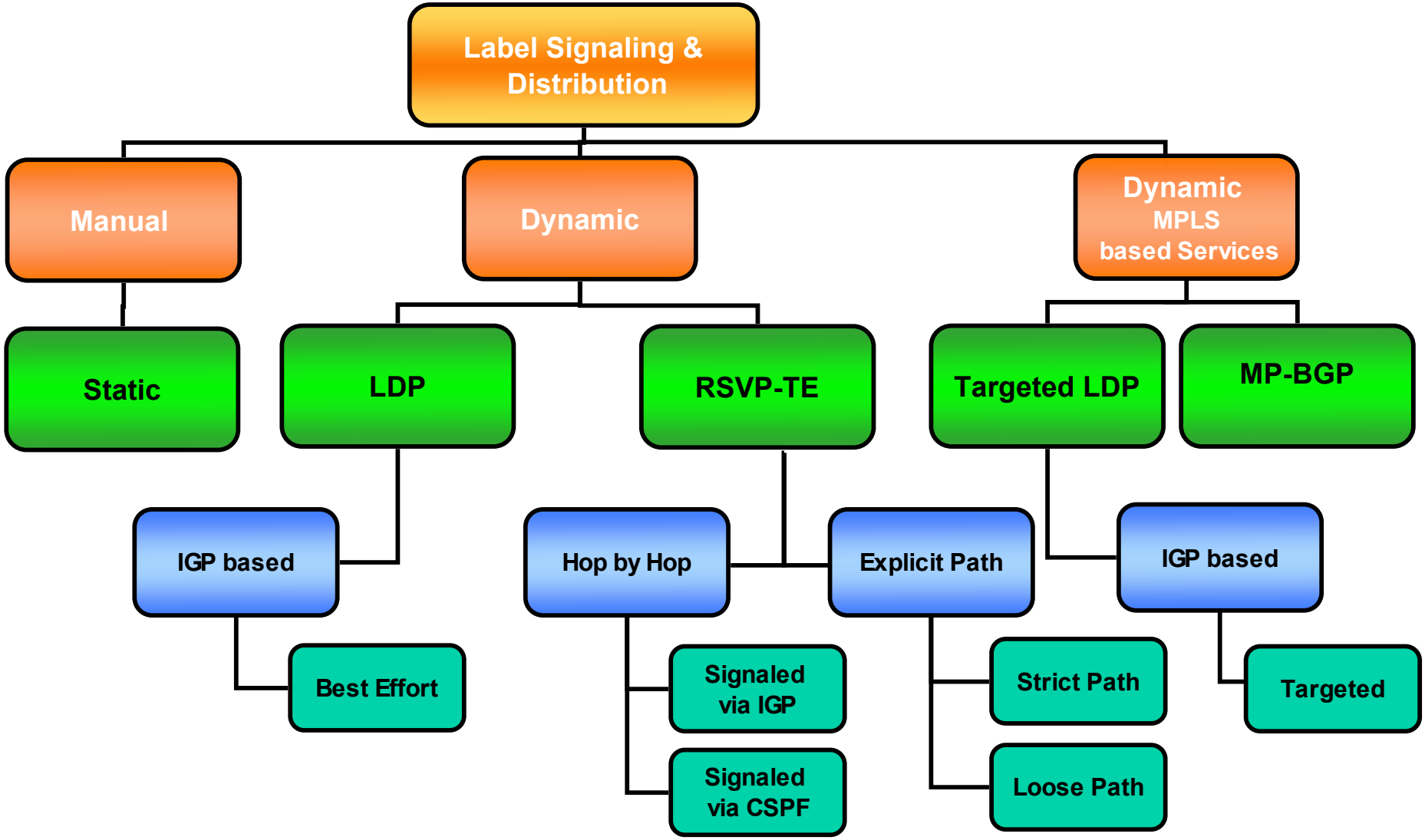
- A fundamental concept in MPLS is that two Label Switch Routers (LSR's) must agree on the labels used to forward traffic between and through them.

- This common understanding is achieved by using a set of procedures, generically called a label distribution protocol, by which one LSR informs another of label bindings it has made

MPLS Label Signaling & Label Distribution Protocols

- Signaling is a mechanism used to setup an LSP path through a network.
- Signaling involves the exchange of messages between LER/LSR routers.
- Exchanged messages include all the details required to setup and maintain LSP paths.
- Label distribution protocols define the procedures and messages by which MPLS LSR's inform each other of the label bindings it has made and their meaning.
- Example of MPLS signalling or label distribution protocols: LDP & RSVP.

Label Signaling and Distribution Summary



2 Label Distribution Protocol

Agenda

2. Label Distribution Protocol (LDP)

    2.1 LDP overview

    2.2 LDP operation

    2.3 Targeted LDP

LDP overview

LDP Overview

- LDP is a set of procedures and messages defined for distributing labels and establishing LSPs based on **RFC 5036**.
- Routers configured for the LDP protocol will establish an LDP session between them and become peers.
- The LDP sessions allow each LDP peer to exchange & learn the other's label/FEC binding (mapping).
- LDP message exchanges are carried in LDP PDUs over LDP session TCP connection.
- The LDP protocol is used for:
  1. Establishing Transport Tunnel LSP's
  2. Establishing Targeted LDP sessions between directly or non-directly connected routers.

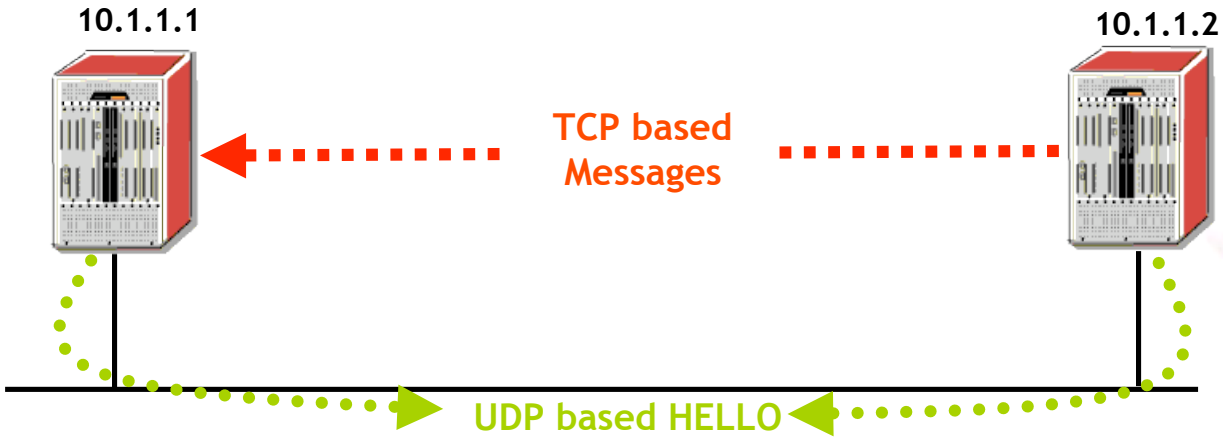
LDP Message Types

- **Discovery messages** - used to announce and maintain the presence of an LSR in a network (Hello)
- **Session messages** - used to establish, maintain, and terminate sessions between LDP peers (Init, KeepAlive)
- **Advertisement messages** - used to create, change, and delete label mappings for FECs (Label Request/Mapping/Withdraw/Release)
- **Notification messages** - used to signal a fatal error, provides advisory information such as message processing outcome or the state of LDP session.

Type	Name	Function
0x0001	Notification	Signal errors and other events
0x0100	Hello	Announces the presence of an LSR
0x0200	Initialization	Initiates the session establishment process
0x0201	KeepAlive	Monitors the integrity of the LDP session transport connection
0x0300	Address	Advertise the interface addresses to an LDP peer
0x0301	Address Withdraw	Withdraws a previously advertised interface address
0x0400	Label Mapping	Advertises a FEC-label binding to an LDP peer
0x0401	Label Request	Requests a FEC-label binding from an LDP peer
0x0402	Label Withdraw	Signals the peer that the previously advertised FEC-label mapping may no longer be used
0x0403	Label Release	Signals the peer that the LSR no longer needs specific FEC-label mappings previously requested of and/or advertised by the peer
0x0404	Label Abort Request	Aborts an outstanding Label Request message
0x3E00 - 0x3EFF	Vendor Private	Used to convey vendor-private information between LSR's
0x3F00 - 0x3FFF	Experimental	LDP Experimental Extensions

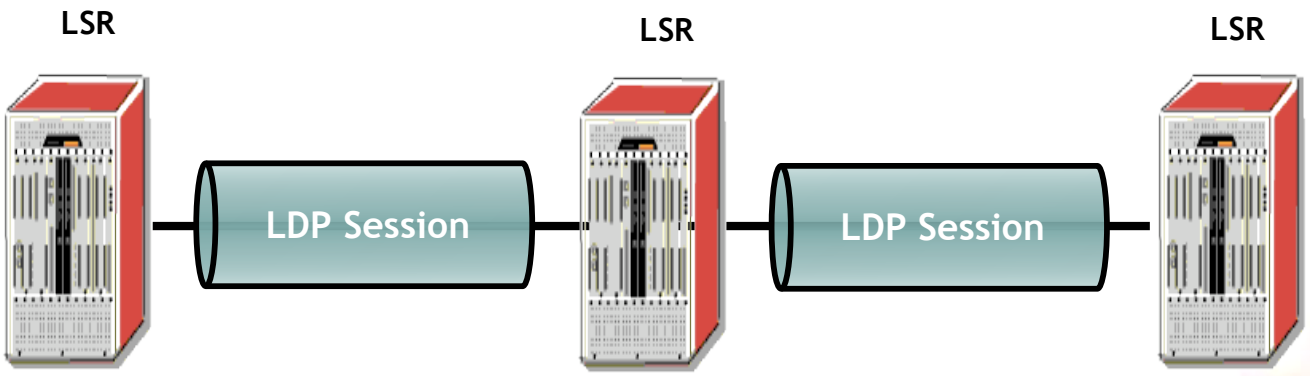
LDP Transport Protocols

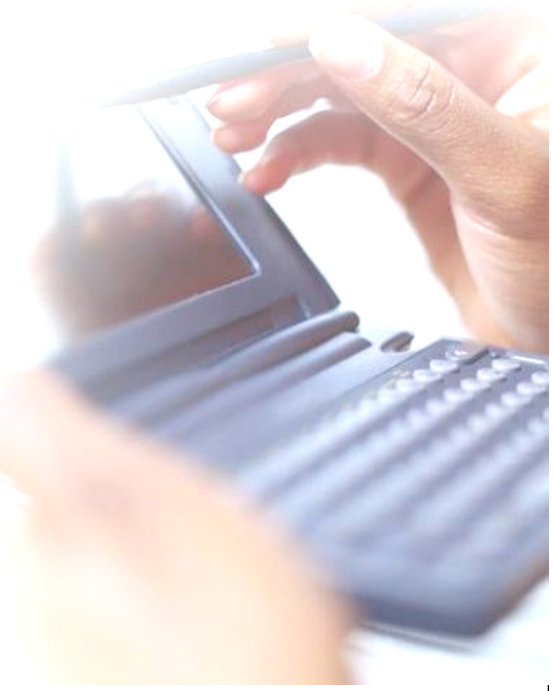
- LDP utilizes both UDP and TCP for transport services and uses port 646 for both.
- UDP is used as the transport protocol for the discovery mechanism
  - Discovery Hello messages periodically announce and maintain the presence of an LSR in a network
- TCP is used as the transport protocol for all messages except HELLO's. TCP based messages are:
  - Session messages to establish, maintain and terminate sessions between LDP peers
  - Advertisement messages to create, change and delete label mappings for FEC's
  - Notification messages to signal errors and other events



LDP Peers

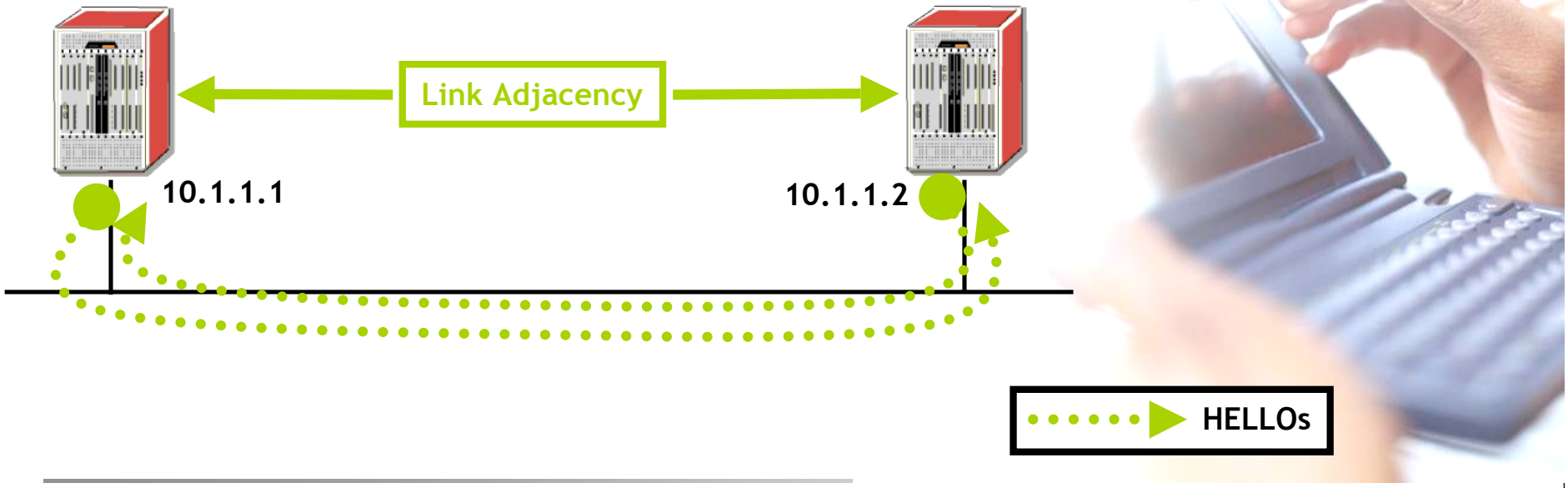
- Routers configured for the LDP protocol will establish an LDP session between them and become peers.
- LDP peers are directly connected





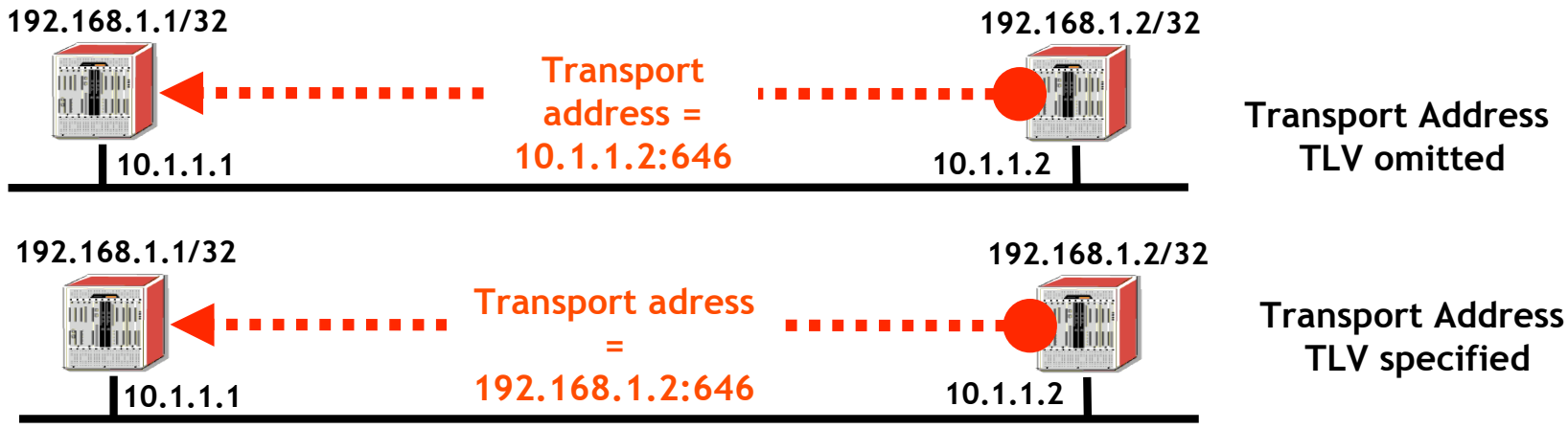
LDP Discovery

- LSRs periodically announce their presence on a network by sending HELLO messages out of each LDP enabled interface to the multicast address 224.0.0.2
- An LDP HELLO sent by an LSR carries the LDP Identifier for the label space the LSR intends to use for the interface.
- Receipt of a HELLO on an interface identifies a HELLO adjacency (Link Adjacency).



LDP Establishment: LDP Transport Address

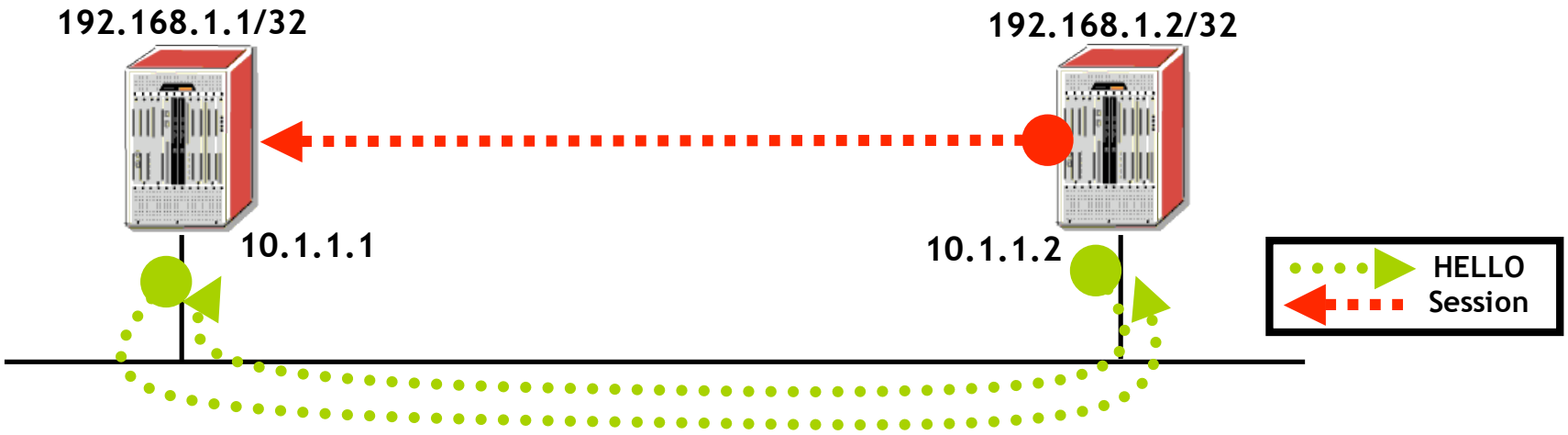
- Each LSR must select an LDP transport address: System or Interface



- Each LSR advertises the transport address for its end of the session using a HELLO message.
- The transport address may be different from the one used as the source of the HELLO
- An LSR MUST advertise the same transport address in all Hellos that advertise the same label space. This requirement ensures that two LSRs linked by multiple Hello adjacencies using the same label spaces play the same connection establishment role for each adjacency.

LDP Session Establishment

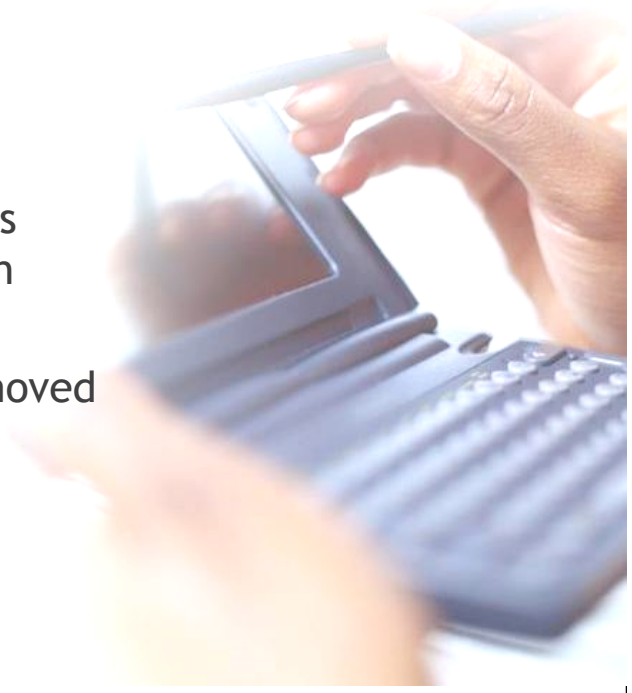
- Once the HELLO adjacency is established, the peer with the higher transport address (in Hello packet) becomes the active peer and initiates establishment of the LDP session.
- Session establishment is a two step process:
  - Transport connection (TCP) establishment
  - Session initialization
- Active peer establishes LDP TCP session by connecting to port 646 at the passive peer.
- Active node sends Init message to start the negotiation of session parameters.



LDP Session Maintenance

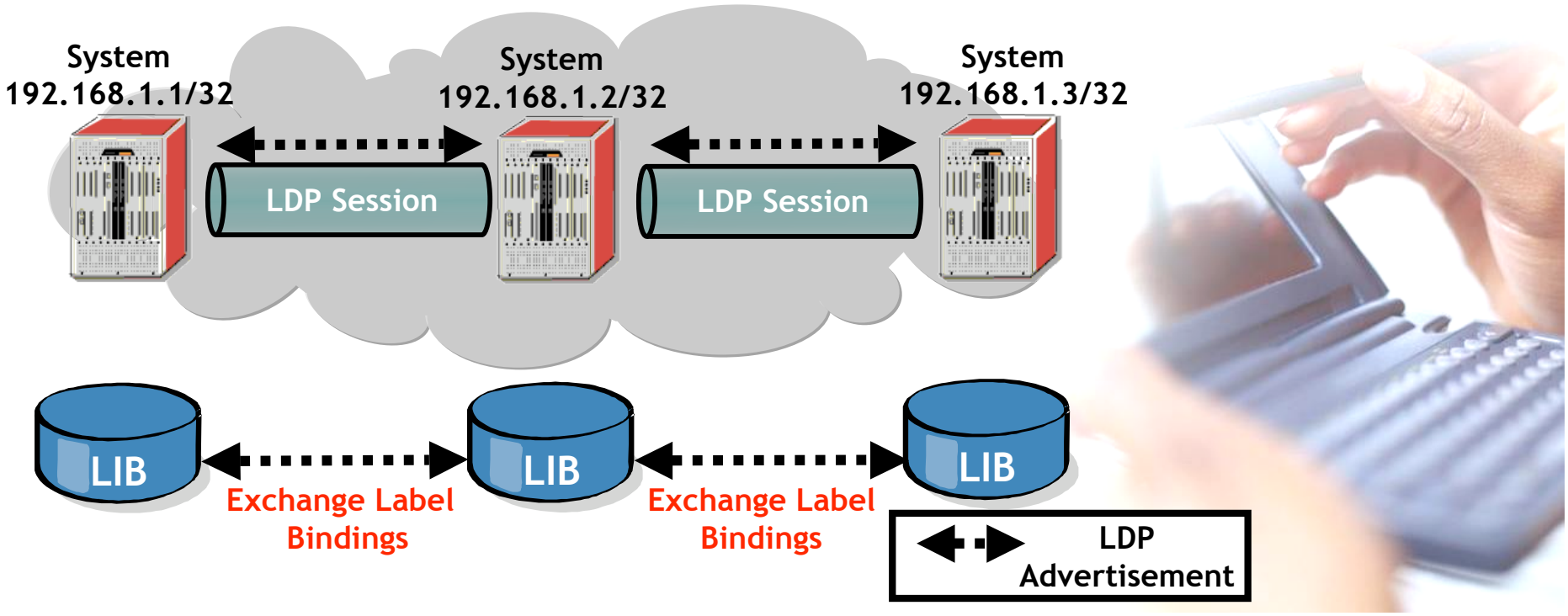
- **Hello Adjacency Maintenance**
  - Hello messages are sent periodically between LDP neighbors.
  - A “hold” timer associated with each Hello adjacency is refreshed when a Hello msg is received.
  - If the “hold” timer expires, the Hello adjacency is removed.
  - When the last Hello adjacency is removed, the TCP session is removed and LDP session is terminated

- **TCP and Session Maintenance**
  - Each LDP session maintains a “KeepAlive” timer which is refreshed upon the receipt of an LDP PDU on the session transport connection from the session peer.
  - If “KeepAlive” timer expires, the TCP connection is removed and LDP session is terminated.



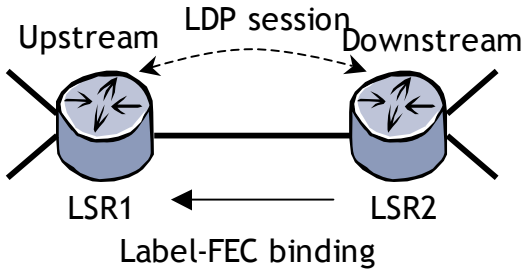
LDP Label Exchange

- There is a separate LDP session established per label space.
- An LDP session allows for the mutual exchange of FEC/label bindings using LDP label Mapping messages.
- By default, a label binding is generated and advertised for only the system address of the Alcatel-Lucent 7x50



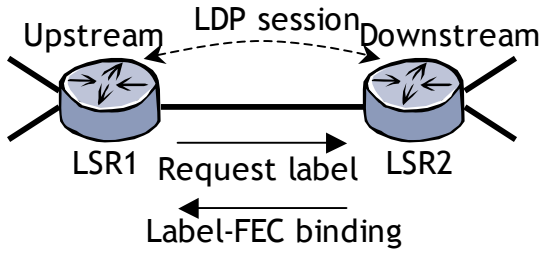
LDP - Label Advertisement Methods

Downstream Unsolicited



1. LSR2 discovers NH (IGP selected Next Hop) for a particular FEC
2. LSR2 sends label mapping message

Downstream-on-Demand



1. LSR1 recognizes LSR2 as NH for a FEC
2. LSR1 sends label request message
3. LSR2 responds with a label binding

LDP – Label Distribution Control

Independent Control

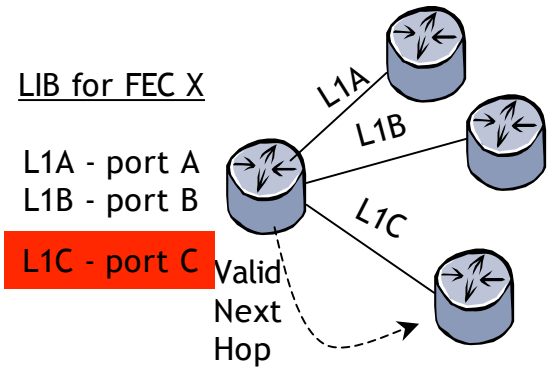
- Each LSR makes independent decision on when to advertise labels to upstream peers
- Label-FEC binding can be sent once NH has been recognized
- All LSRs may advertise labels (don't wait for downstream label binding)
- Faster convergence

Ordered Control

- Label-FEC binding is sent out if:
  - LSR is egress for FEC, or
  - Label binding has been received from downstream FEC-NH
- LSP is set up hop by hop from egress to ingress
- Only egress LSR can start LSP set up
- Setting up LSP takes longer

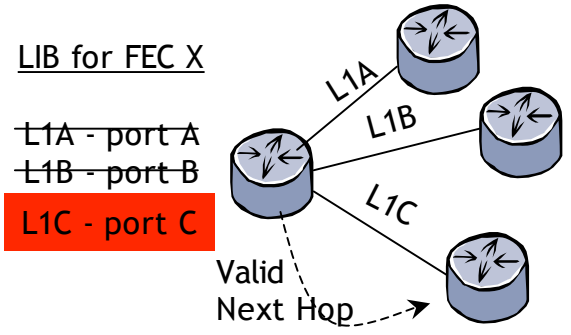
LDP - Label Retention Methods

Liberal Label Retention



- LSR retains bindings received from LSR's other than the valid NH LSR.
- If NH changes, it may begin using these labels immediately.
- Allows rapid adaptation to routing changes.
- Requires LSR to retain more labels.

Conservative Label Retention



- LSR will only retain label bindings received from FEC-NH.
- If NH changes, new binding must be requested from new next hop.
- Restricts adaptation to changes in routing.
- Fewer labels to be maintained in LSR.

Summary of LDP Characteristics

Features	LDP
UDP/TCP Based	Yes
Dependency on the IGP	Yes
Traffic Engineering or BW reservation	No
Signaled LSP's via the IGP	Yes
Signaled LSP's via CSPF	No
Explicit Path LSP's	No
Protection Mechanism	No
MPLS Convergence	Failure Detection + IGP /LDP Convergence
Scalable	Yes
Administrative Control	Medium
Configuration Complexity	Low

Contents of the LIB and LFIB

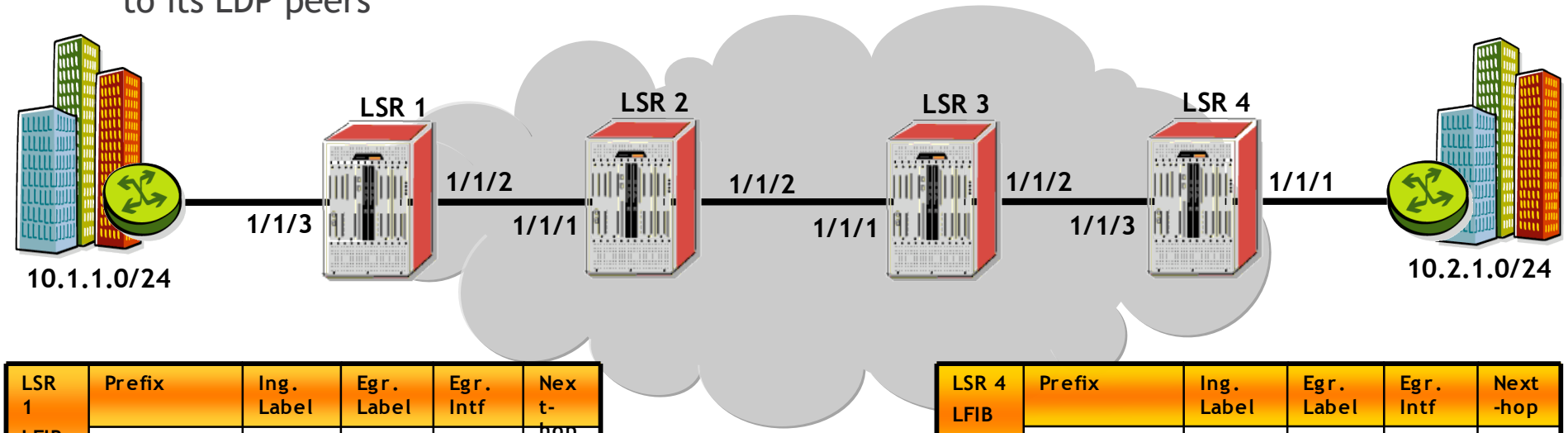
- The LIB is populated based on label exchange with neighbors.
- The LFIB is built from the LIB and the FIB

Table Name	Meaning	Contents	Populated By
RIB	Routing Information Base	Routing updates received	Routing Protocol Exchange - Each routing protocol has a separate RIB
FIB	Forwarding Information Base	Active routes	RTM selects the active routes from all protocol "Best" routes
LIB	Label Information Base	Locally generated and received MPLS labels	MPLS Label Exchange
LFIB	Label Forwarding Information Base	Labels used by the LSR	The labels assigned to the active routes (for each next-hop)

LDP operation

LDP Signaling

- Each LSR will originate a label for its system address by default
- Each LSR may originate a label for a FEC for which it has a next-hop that is external to the MPLS domain
- Each LSR will propagate labels for each FEC for which it is a possible next-hop to its LDP peers



LSR 1 LFIB	Prefix	Ing. Label	Egr. Label	Egr. Intf	Next-hop
	10.1.1.0/24	100	-	1/1/3	ext.

LSR 4 LFIB	Prefix	Ing. Label	Egr. Label	Egr. Intf	Next-hop
	10.2.1.0/24	100	-	1/1/1	ext.

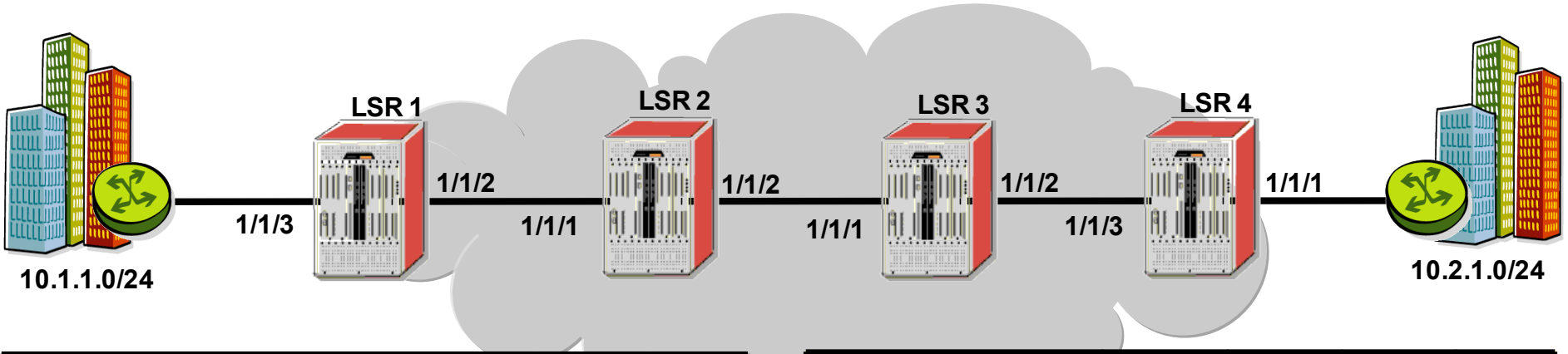
Control Plane

LSR 2 LFIB	Prefix	Ing. Label	Egr. Label	Egr. Intf	Next-hop
	10.1.1.0/24	101	100	1/1/1	LSR 1

LSR 3 LFIB	Prefix	Ing. Label	Egr. Label	Egr. Intf	Next-hop
	10.2.1.0/24	200	100	1/1/2	LSR4

LDP Signaling

- The labels will propagate bidirectionally to the edge of the MPLS domain

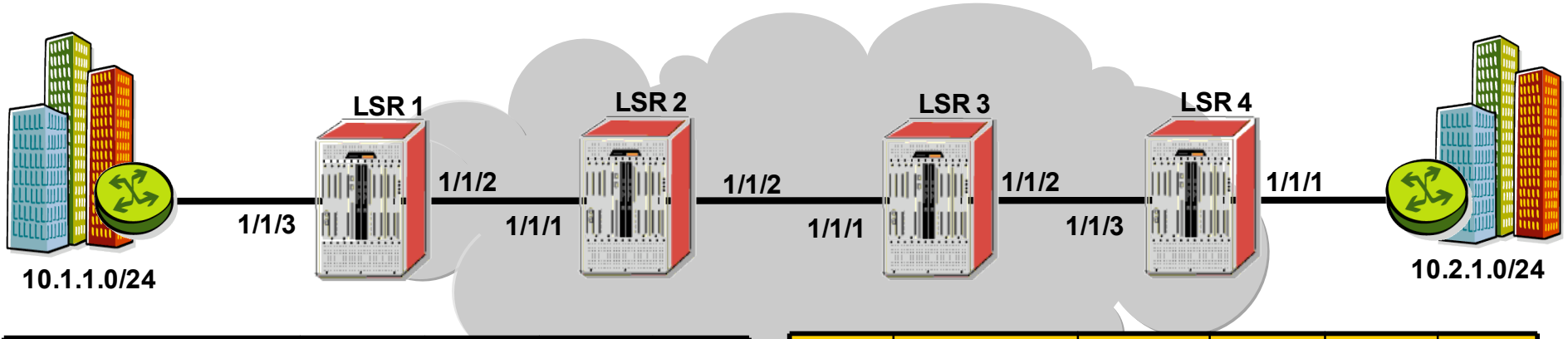


LSR 2 LFIB	Prefix	Ing. Label	Egr. Label	Egr. Intf	Next-hop
	10.1.1.0/24	101	100	1/1/1	LSR1
	10.2.1.0/24	300	200	1/1/2	LSR3

LSR 3 LFIB	Prefix	Ing. Label	Egr. Label	Egr. Intf	Next-hop
	10.1.1.0/24	300	101	1/1/1	LSR2
	10.2.1.0/24	200	100	1/1/2	LSR4

LDP Signaling

- When the labels have propagated across the MPLS domain
  - The ingress router has a blank ingress label
  - The egress router has a blank egress label

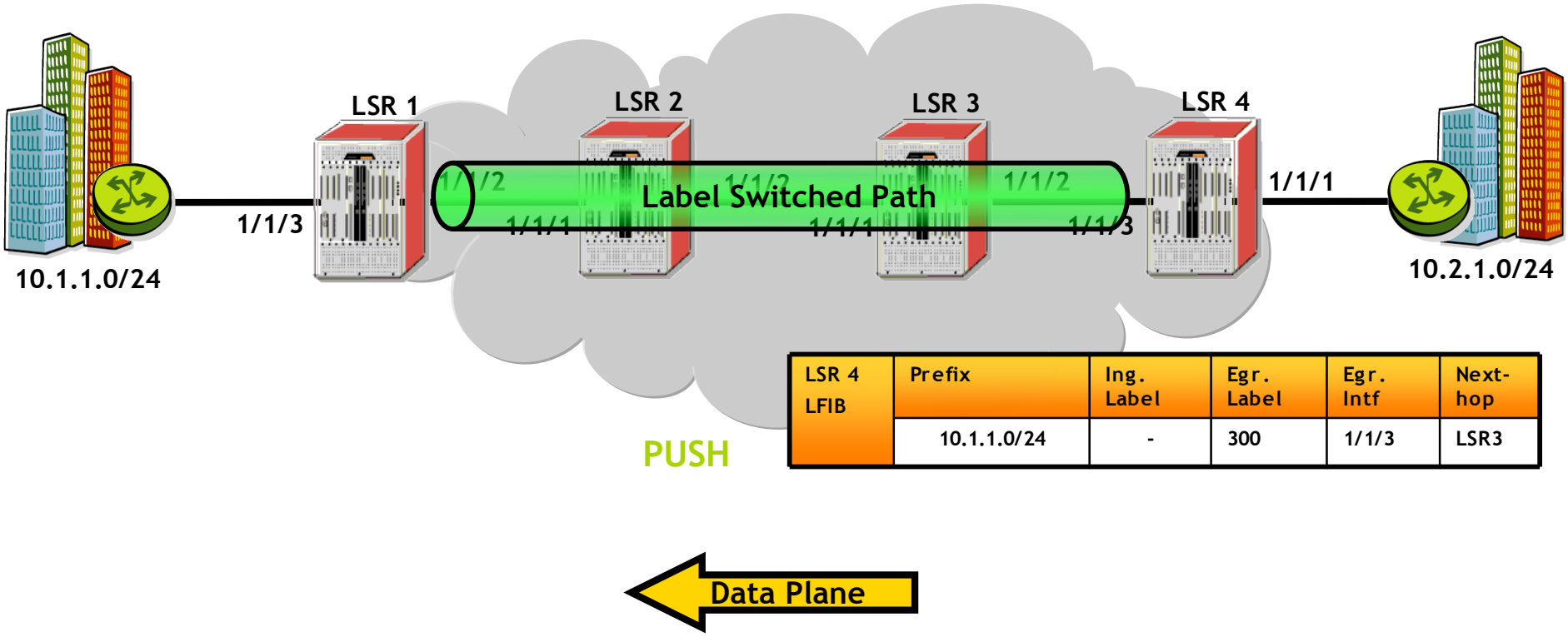


LSR 1 LFIB	Prefix	Ing. Label	Egr. Label	Egr. Intf	Next-hop
	10.1.1.0/24	100	-	1/1/3	ext.
	10.2.1.0/24	-	300	1/1/2	LSR2

LSR 4 LFIB	Prefix	Ing. Label	Egr. Label	Egr. Intf	Next-hop
	10.1.1.0/24	-	300	1/1/3	LSR3
	10.2.1.0/24	100	-	1/1/1	ext.

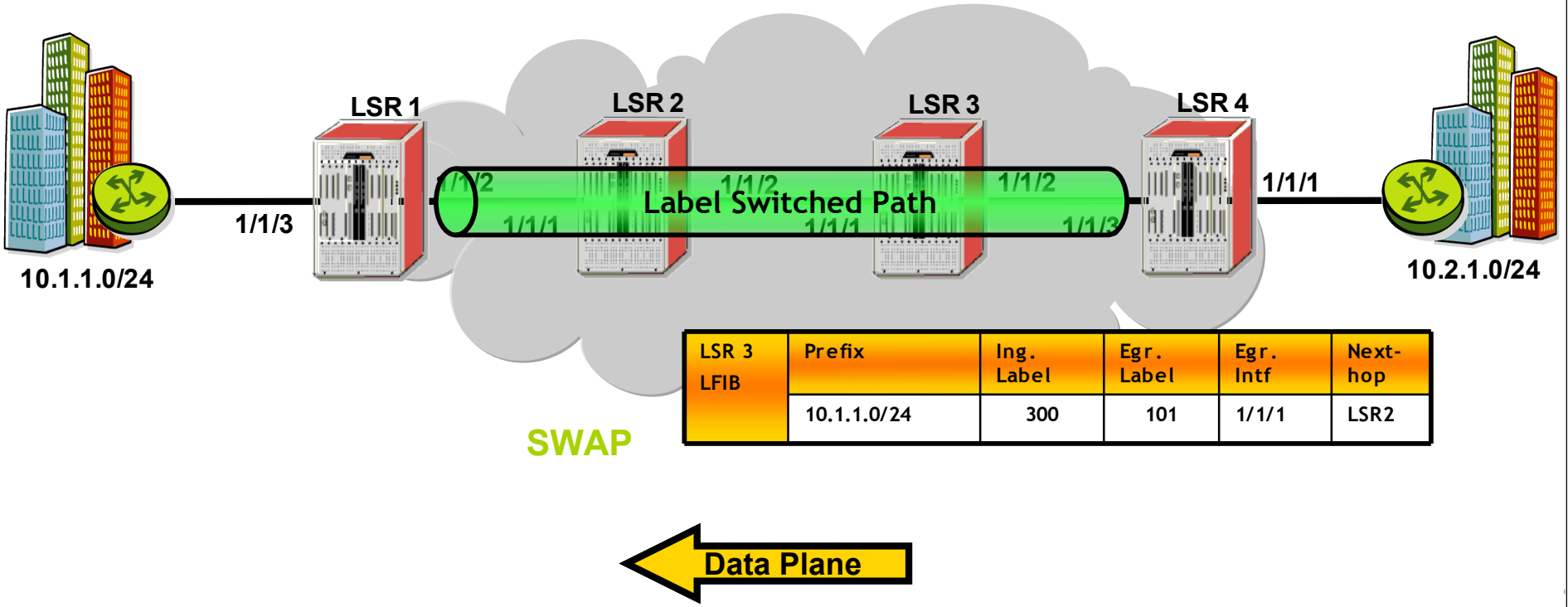
Data Plane Operation

- A packet arriving at LSR 4 for FEC 10.1.1.0/24 should arrive unlabeled and will have label 300 PUSHed onto it
  - It will then be forwarded to LSR 3 via interface 1/1/3



Data Plane Operation

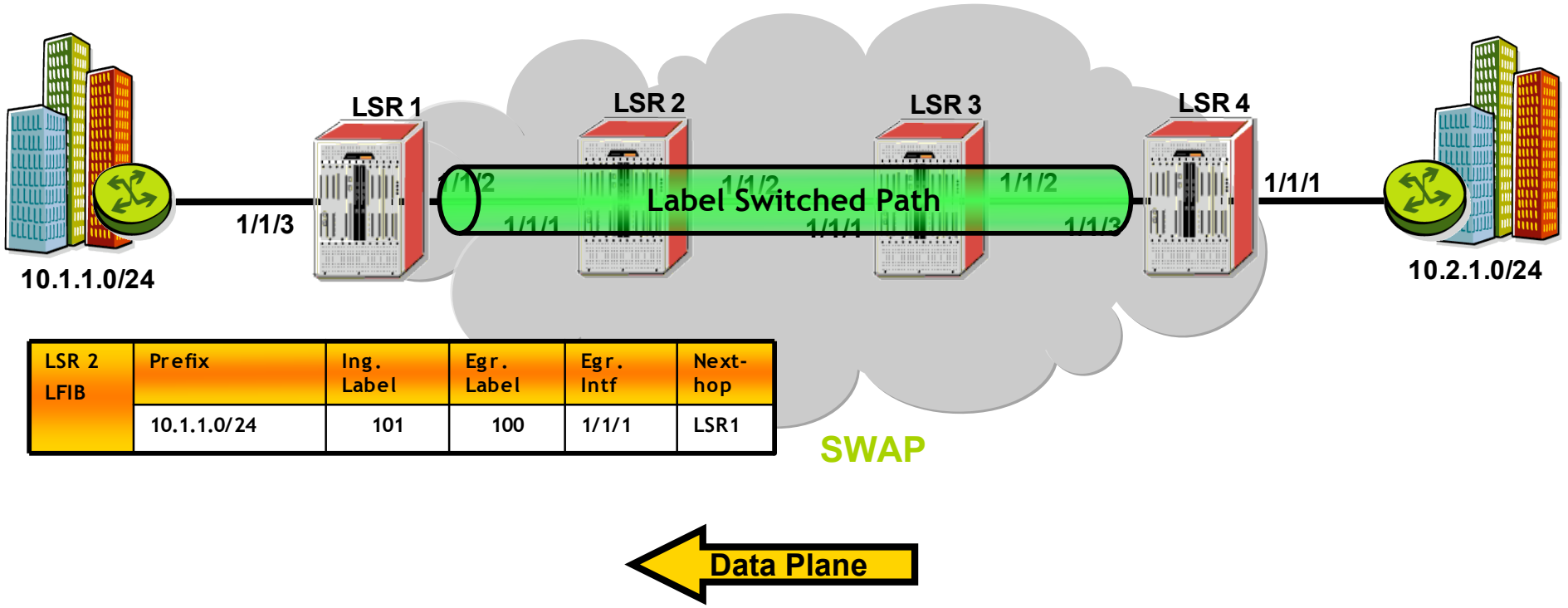
- The packet arriving at LSR 3 for FEC 10.1.1.0/24 should arrive with label 300
- Label 300 will be SWAPped for label 101
  - It will then be forwarded to LSR 2 via interface 1/1/1



LSR 3 LFIB	Prefix	Ing. Label	Egr. Label	Egr. Intf	Next-hop
	10.1.1.0/24	300	101	1/1/1	LSR2

Data Plane Operation

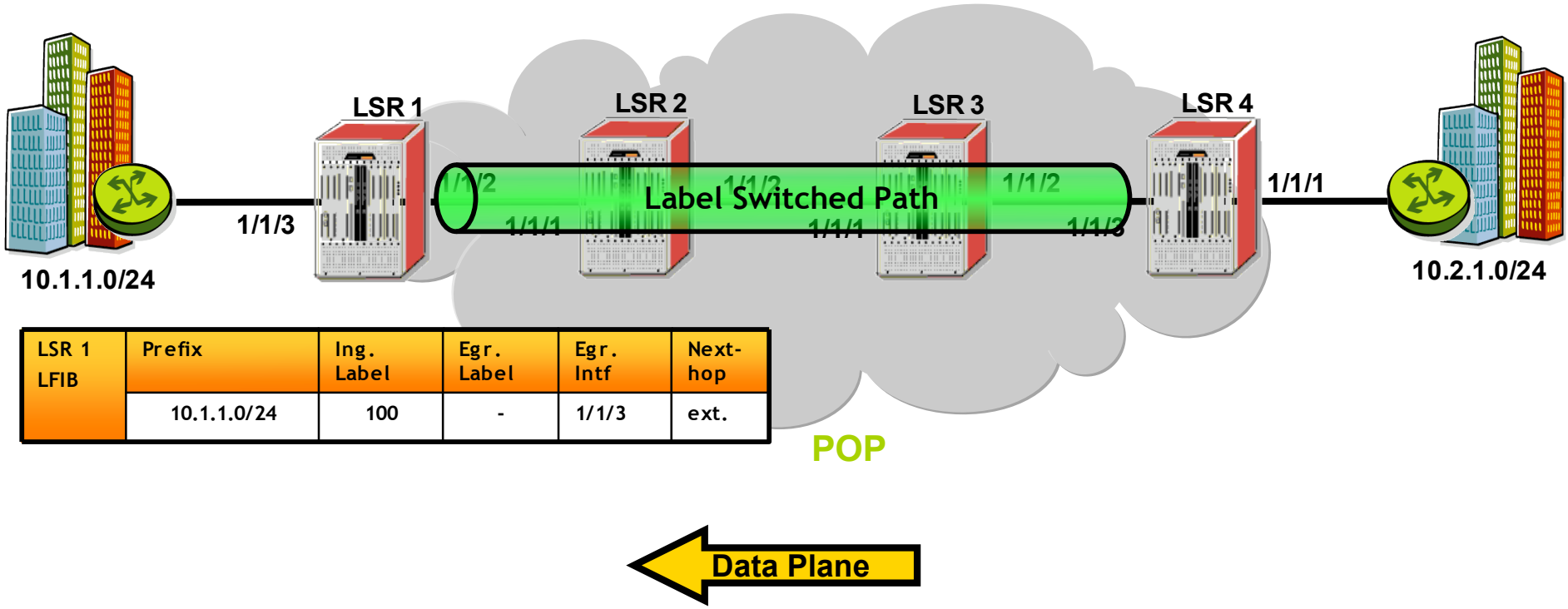
- The packet arriving at LSR 2 for FEC 10.1.1.0/24 should arrive with label 101
- Label 101 will be SWAPped for label 100
  - It will then be forwarded to LSR 1 via interface 1/1/1



LSR 2 LFIB	Prefix	Ing. Label	Egr. Label	Egr. Intf	Next-hop
	10.1.1.0/24	101	100	1/1/1	LSR1

Data Plane Operation

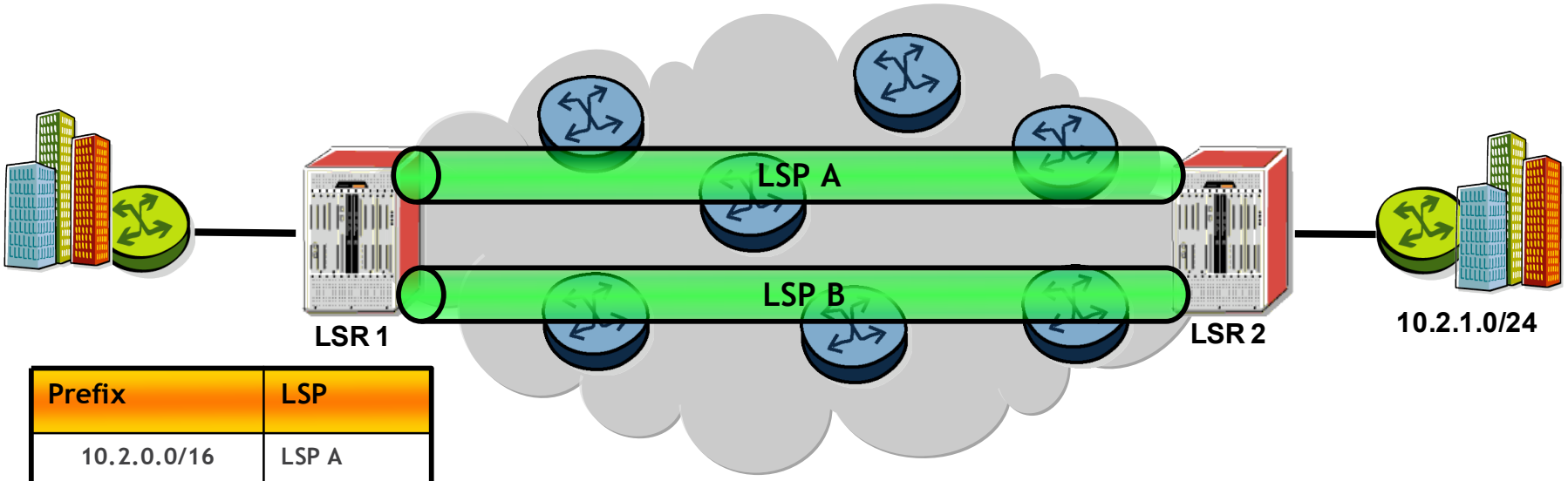
- The packet arriving at LSR 1 for FEC 10.1.1.0/24 should arrive with label 100
- Label 100 will be POPped since there is no outgoing label
  - It will then be forwarded via interface 1/1/3 outside the MPLS domain



LSR 1 LFIB	Prefix	Ing. Label	Egr. Label	Egr. Intf	Next-hop
	10.1.1.0/24	100	-	1/1/3	ext.

Label Switching Longest Match

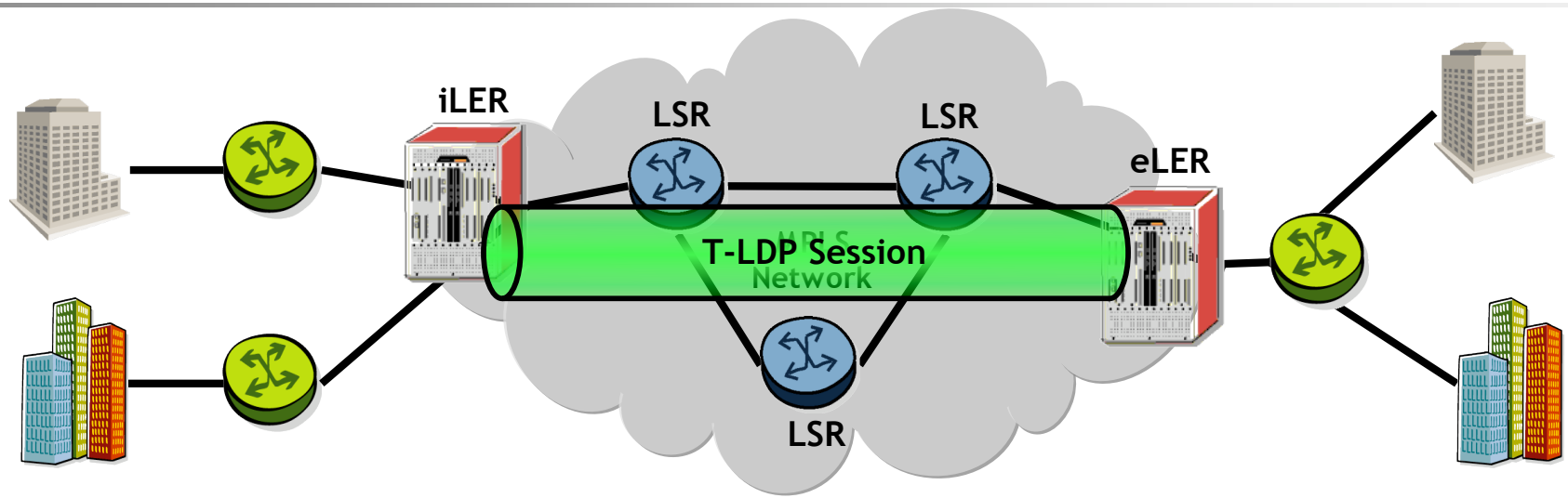
- A packet matches an LSP if and only if that LSP has a FEC which matches the packet's destination address
- A packet matches an LSP based on longest match to the prefix



Prefix	LSP
10.2.0.0/16	LSP A
10.2.2.0/24	LSP B

Targeted LDP (tLDP)

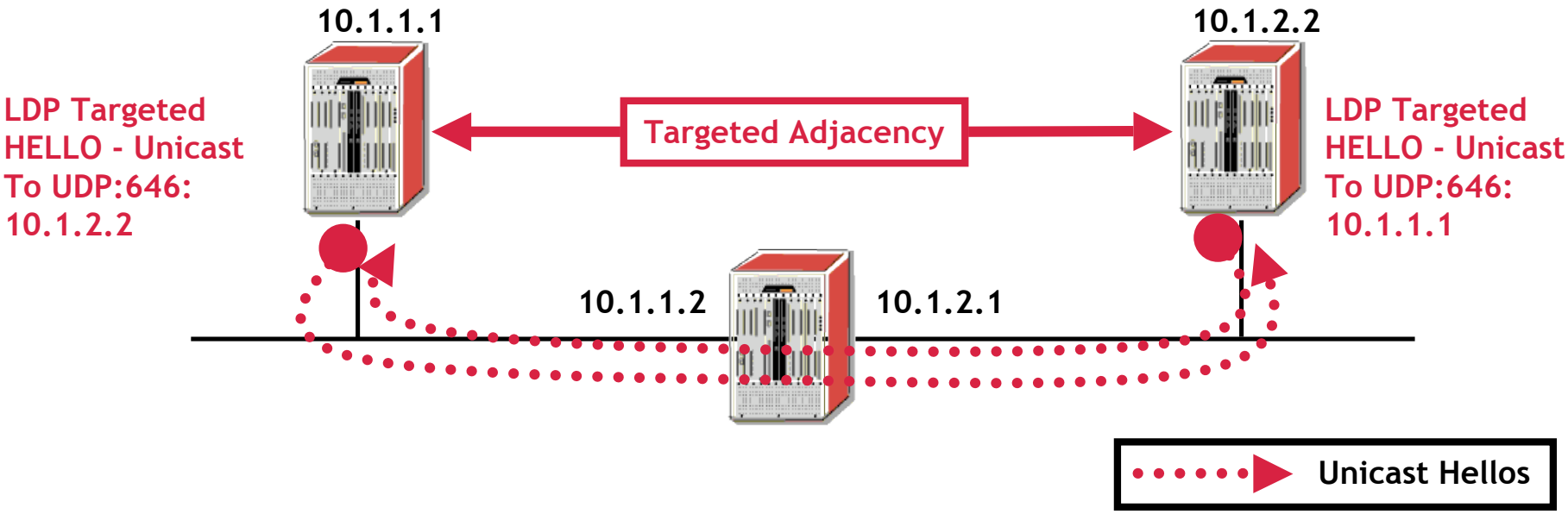
Targeted LDP Sessions



- Targeted LDP would be configured in the case of a service implementation such as PWE3 Services: aPipe, cPipe, etc..
- Targeted LDP sessions can be established between peers that are not directly connected
  - Provides a tunnel between ingress and egress LERs
- Link based sessions may still remain between the directly connected LSRs
  - Provides the hop by hop tunnel across the core

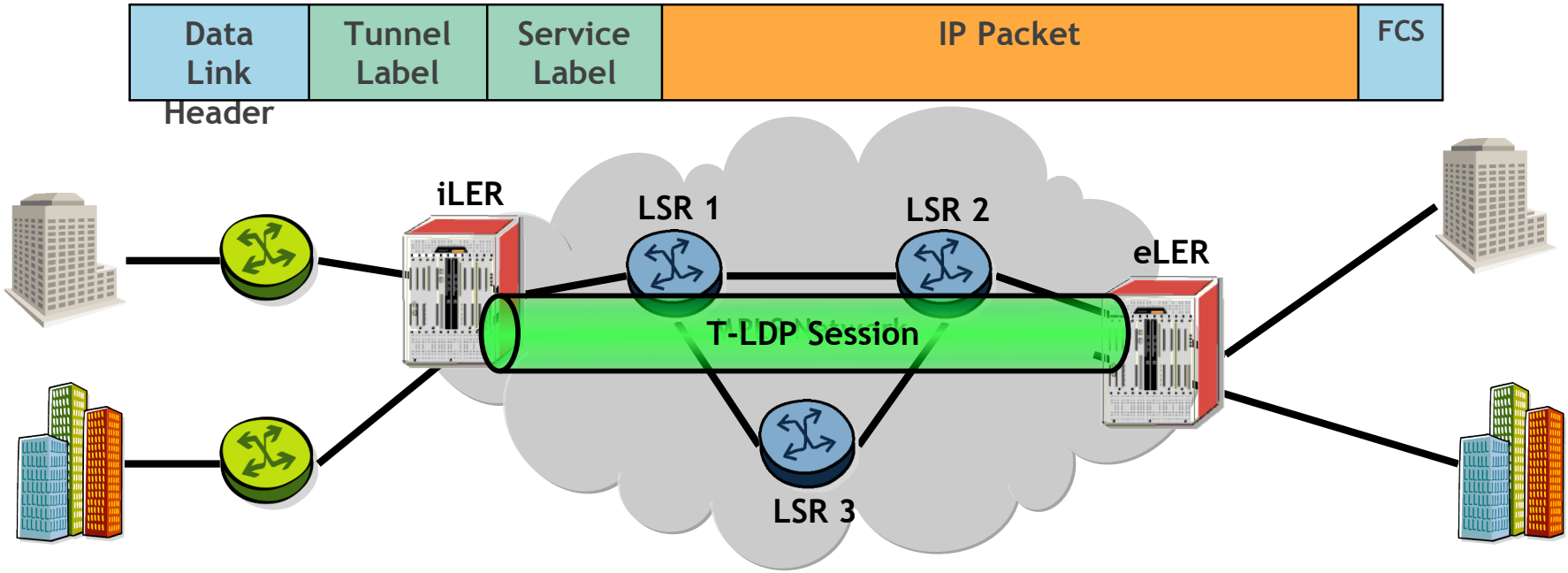
T-LDP Discovery - Non-Directly Connected LSRs

- The process is similar to regular neighbor establishment except the Targeted HELLOs are sent via unicast
- Receipt of an LDP Targeted HELLO identifies a "Hello adjacency“
  - Referred to as a Targeted Adjacency



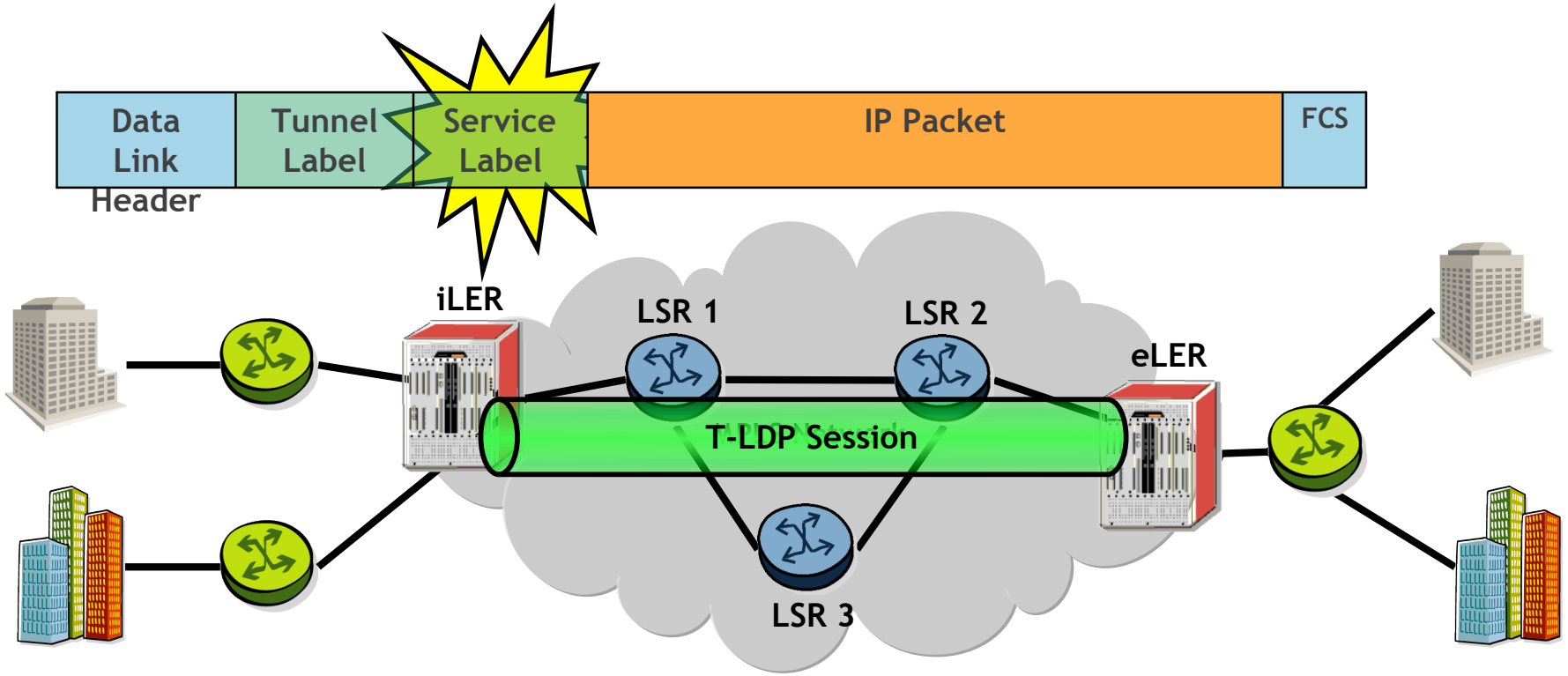
MPLS Label Stack

- An MPLS frame may have one or more labels applied to it
- The outer label is the Tunnel label and is used to switch the frame across the provider MPLS backbone
- The inner label is the Service label and is used by the egress LER to determine the egress interface or service



Service (Inner) Label

- The service labels, advertised by T-LDP, are used to identify to which service a packet belongs
  - It is PUSHed at the iLER and POPped at the eLER
  - Creates a per service tunnel that isolates traffic from other services.
- The Service labels are distributed via Targeted LDP (T-LDP)
- Labels are exchanged in Downstream Unsolicited (DU) mode



Pseudowire using LDP

- RFC4447 specifies the use of LDP to setup and maintain pseudowire
- Service packets are transmitted from one end of the pseudowire to the other end through an MPLS tunnel (LDP-DU LSP)
- LDP-DU is used to establish MPLS tunnels (*RSVP-TE tunnels may also be used)*
- T-LDP is used between LERs to exchange pseudowire labels

3

Questions ???

Thank You

paresh.khatri@alcatel-lucent.com

www.alcatel-lucent.com