MPLS Introduction

The slides are based on a set developed by MPLS Forum;

MPLS Technology and Applications, B. Davie and Y. Rekhter, Morgan Kaufman, 2001**.**
Traffic Engineering with MPLS by E. Osborne and A. Simha, Cisco Press 2003; and
IP Switching and Routing Essentials, S. Thomas, Wiley, 2002

Section 2: Agenda

- How Does Traditional Routing Work?
  - ✓ Brief overview
  - ✓ The hyperaggregation problem
- MPLS Architecture
  - ✓ Data Plane and Control Plane
  - ✓ MPLS Terminology
  - ✓ How Does It Work?
  - ✓ Label Distribution Protocol (LDP)
  - ✓ Penultimate Hop Popping, Aggregation, TTL
  - ✓ ATM Issues

MPLS – How It All Started

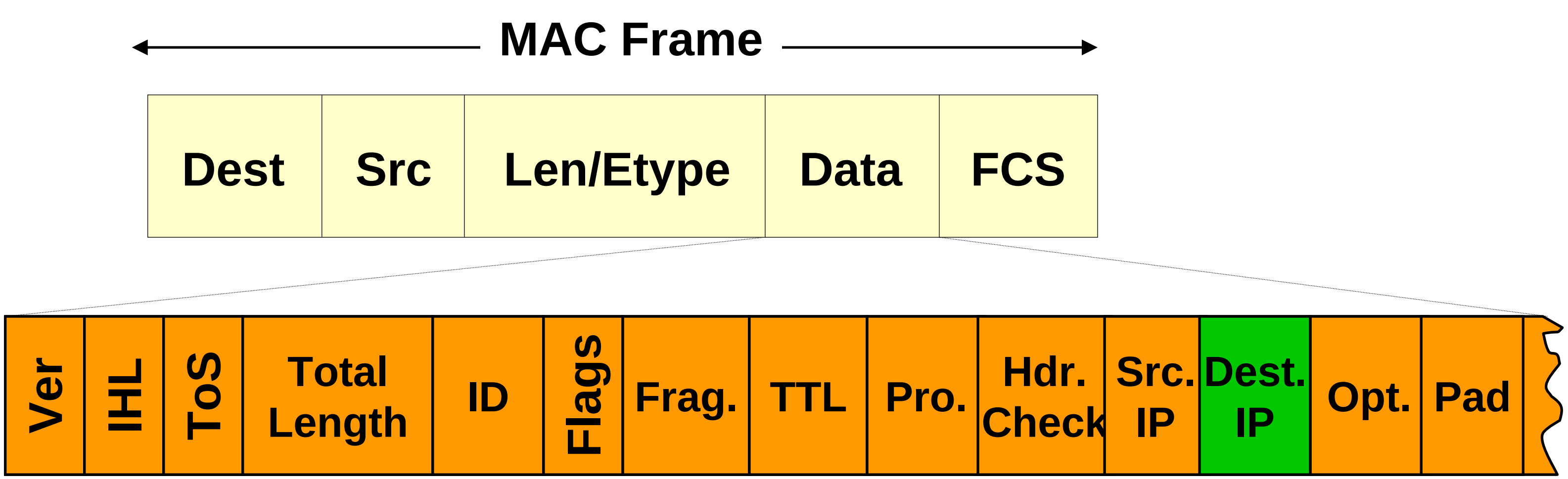
- **Early Multi-Layer Switching Initiatives**
  - ✓ IP Switching (Ipsilon/Nokia)
  - ✓ Tag Switching (Cisco)
  - ✓ IP Navigator (Cascade/Ascend/Lucent)
  - ✓ ARIS (IBM)
- **IETF Working Group chartered in spring 1997**
- **IETF Solution should address the following problems:**
  - ✓ Enhance **performance** and **scalability** of IP routing
  - ✓ Facilitate **explicit routing** and traffic engineering
  - ✓ Separate control (routing) from the forwarding mechanism so each can be modified independently
  - ✓ Develop a single forwarding algorithm to support a wide range of routing functionality

How Did We Get Here?

- Growth and Evolution of the Internet
  - ✓ Both number of users and bandwidth requirements
    - ➢ ISPs need higher **performance** switching/routing products
  - ✓ Increased number of nodes, more routes in routing tables, more flows passing thro a point, …
    - ➢ Scalability: the ability to grow the network
  - ✓ Evolution of routing functionality of the Internet
    - ➢ Classless Interdomain Routing (CIDR): prefixes can be any length. Requires changes to in the forwarding algorithm of all IP routers and affect performance.
- Label Switching
  - ✓ Forwarding algorithm is independent of control paradigms and can be put in hardware or tune software once
  - ✓ Shorten the time to develop & deploy new routing functionality

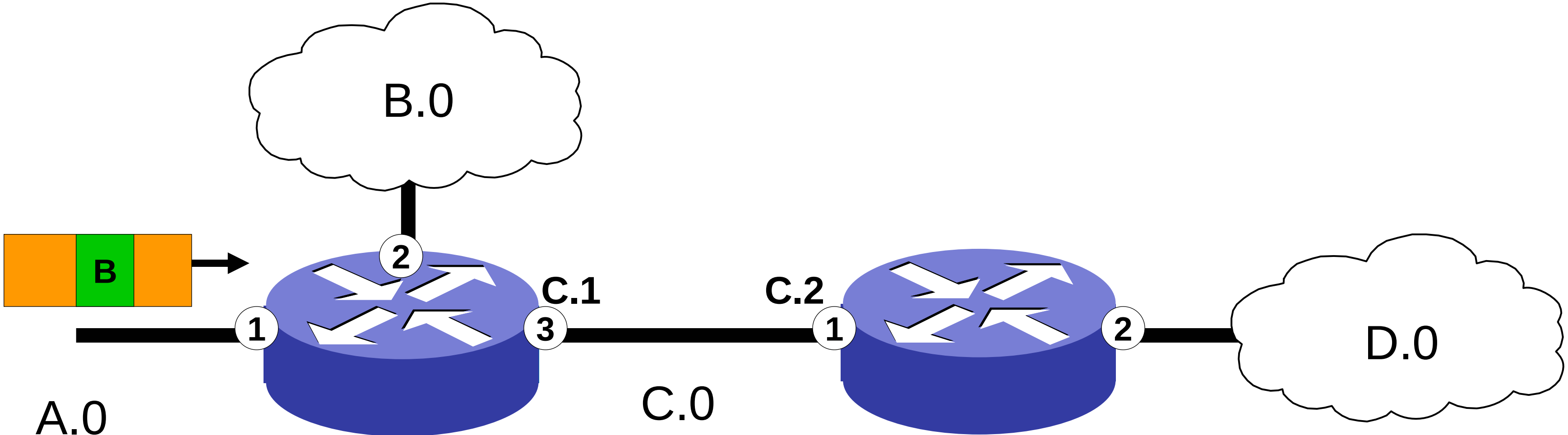
Traditional IP Routing

- Examines the *destination IP address* for each packet received*



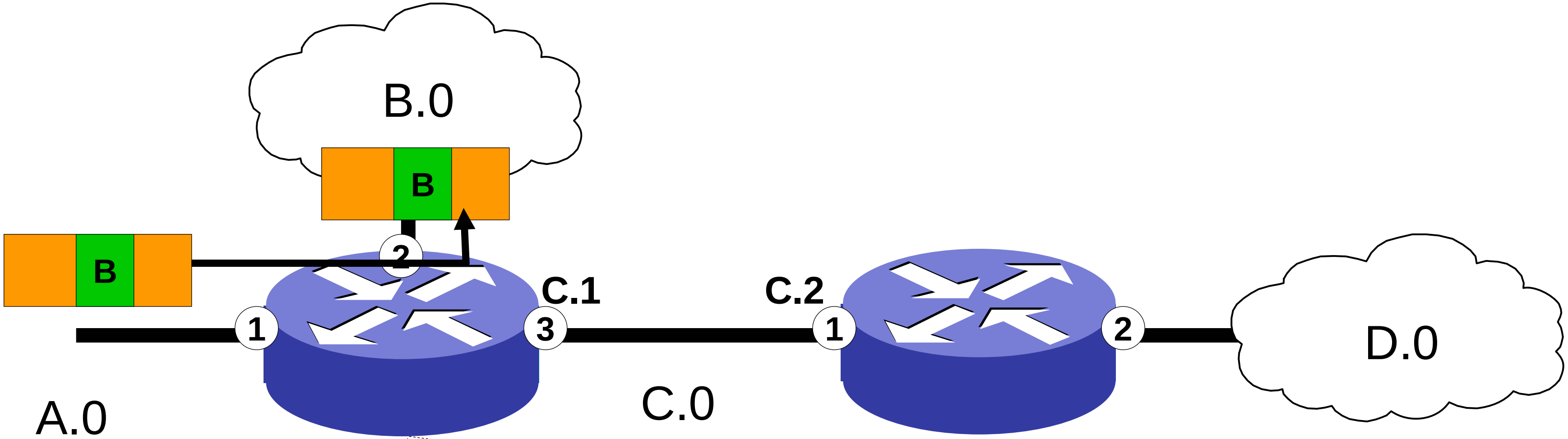
**This is a strict interpretation of connectionless routing*

Case 1: A Direct Route



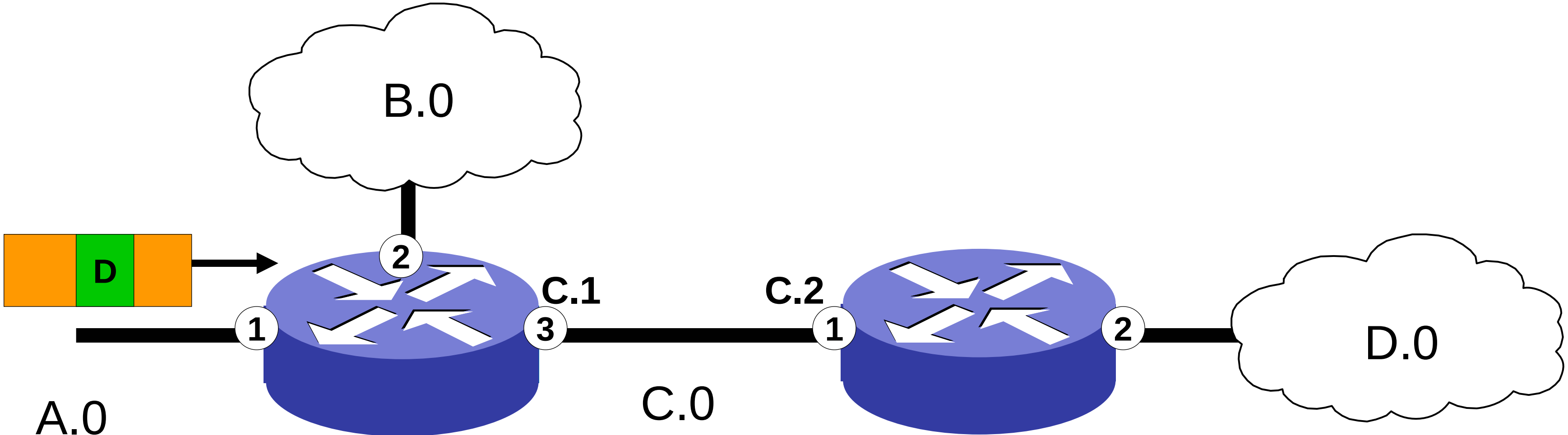
Dest.	Next Hop	Cost	Port
A.0	direct	0	1
B.0	direct	0	2
C.0	direct	0	3
D.0	C.2	1	3

Case 1: A Direct Route



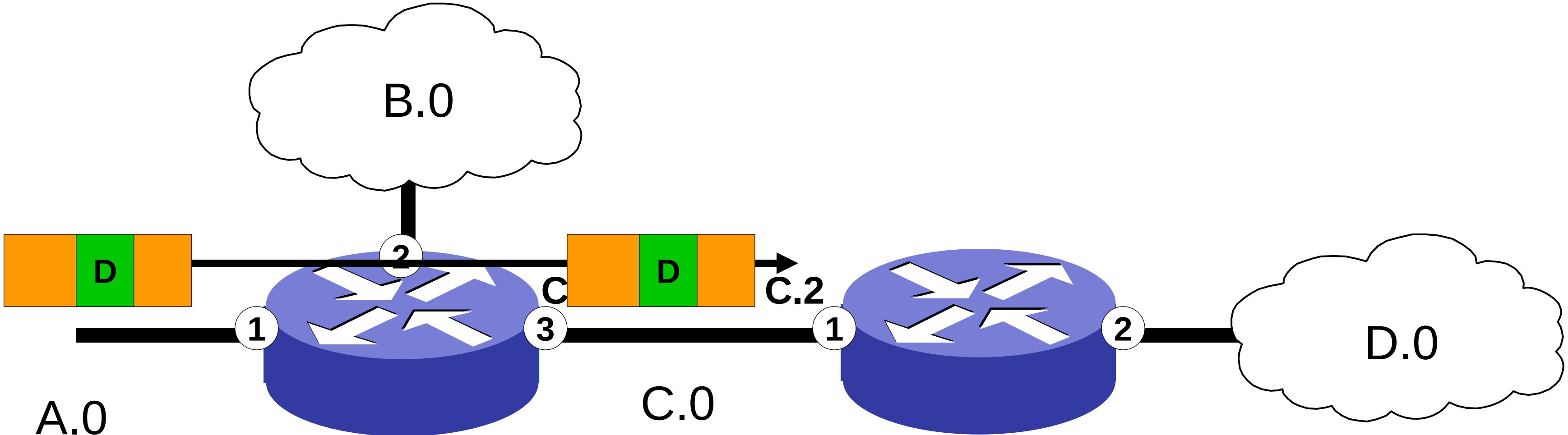
Dest.	Next Hop	Cost	Port
A.0	direct	0	1
B.0	direct	0	2
C.0	direct	0	3
D.0	C.2	1	3

Case 2: An Indirect Route



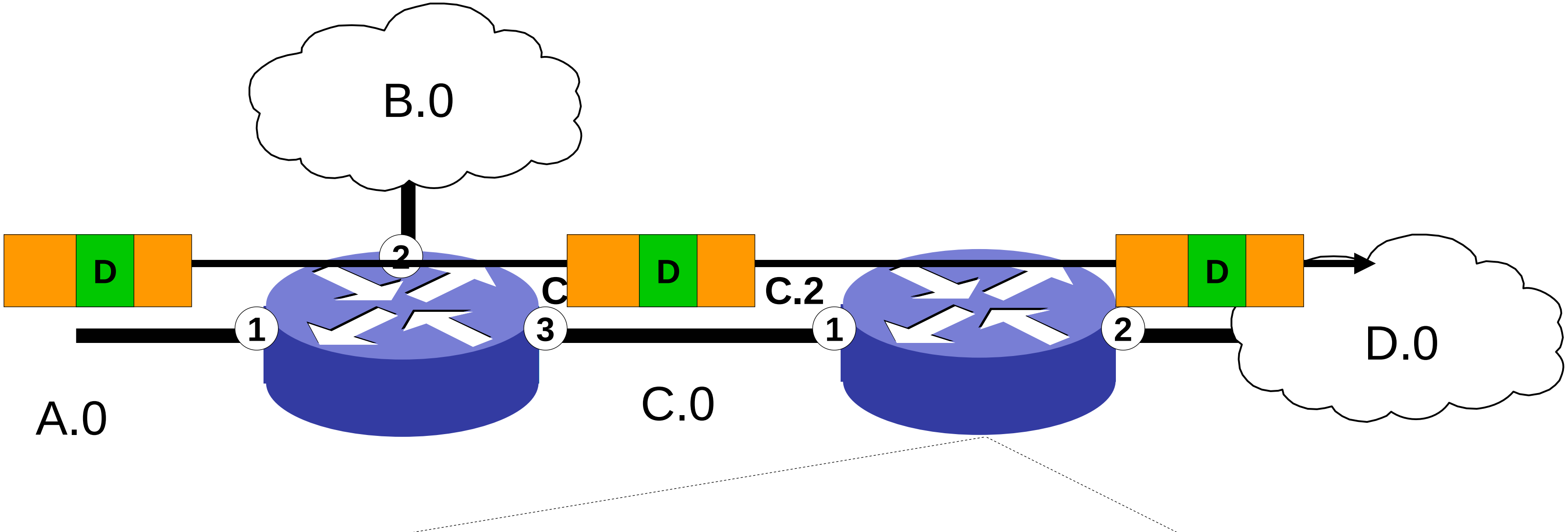
Dest.	Next Hop	Cost	Port
A.0	direct	0	1
B.0	direct	0	2
C.0	direct	0	3
D.0	C.2	1	3

Case 2: An Indirect Route



Dest.	Next Hop	Cost	Port
A.0	C.1	1	1
B.0	C.1	1	1
C.0	direct	0	1
D.0	direct	0	2

Case 2: An Indirect Route



Dest.	Next Hop	Cost	Port
A.0	C.1	1	1
B.0	C.1	1	1
C.0	direct	0	1
D.0	direct	0	2

Per-Hop Routing:
Three Important Questions

- Q: What field on the packet do we use to make the forwarding decision?
  - ✓
- Q: When we use this field as an index into the Routing Table…what do we look up?
  - ✓
- Q: What other vital piece of information does the Routing Table contain?
  - ✓

Per-Hop Routing:
Three Important Questions

- Q: What field on the packet do we use to make the forwarding decision?
  - ✓ A: The destination IP address
- Q: When we use this field as an index into the Routing Table…what do we look up?
  - ✓
- Q: What other vital piece of information does the Routing Table contain?
  - ✓

Per-Hop Routing:
Three Important Questions

- Q: What field on the packet do we use to make the forwarding decision?
  - ✓ A: The destination IP address
- Q: When we use this field as an index into the Routing Table…what do we look up?
  - ✓ A: The next hop IP address
- Q: What other vital piece of information does the Routing Table contain?
  - ✓

Per-Hop Routing:
Three Important Questions

- Q: What field on the packet do we use to make the forwarding decision?
  - ✓ A: The destination IP address
- Q: When we use this field as an index into the Routing Table…what do we look up?
  - ✓ A: The next hop IP address
- Q: What other vital piece of information does the Routing Table contain?
  - ✓ A: An internal reference to the output I/F

How Are These Routing Tables Populated?

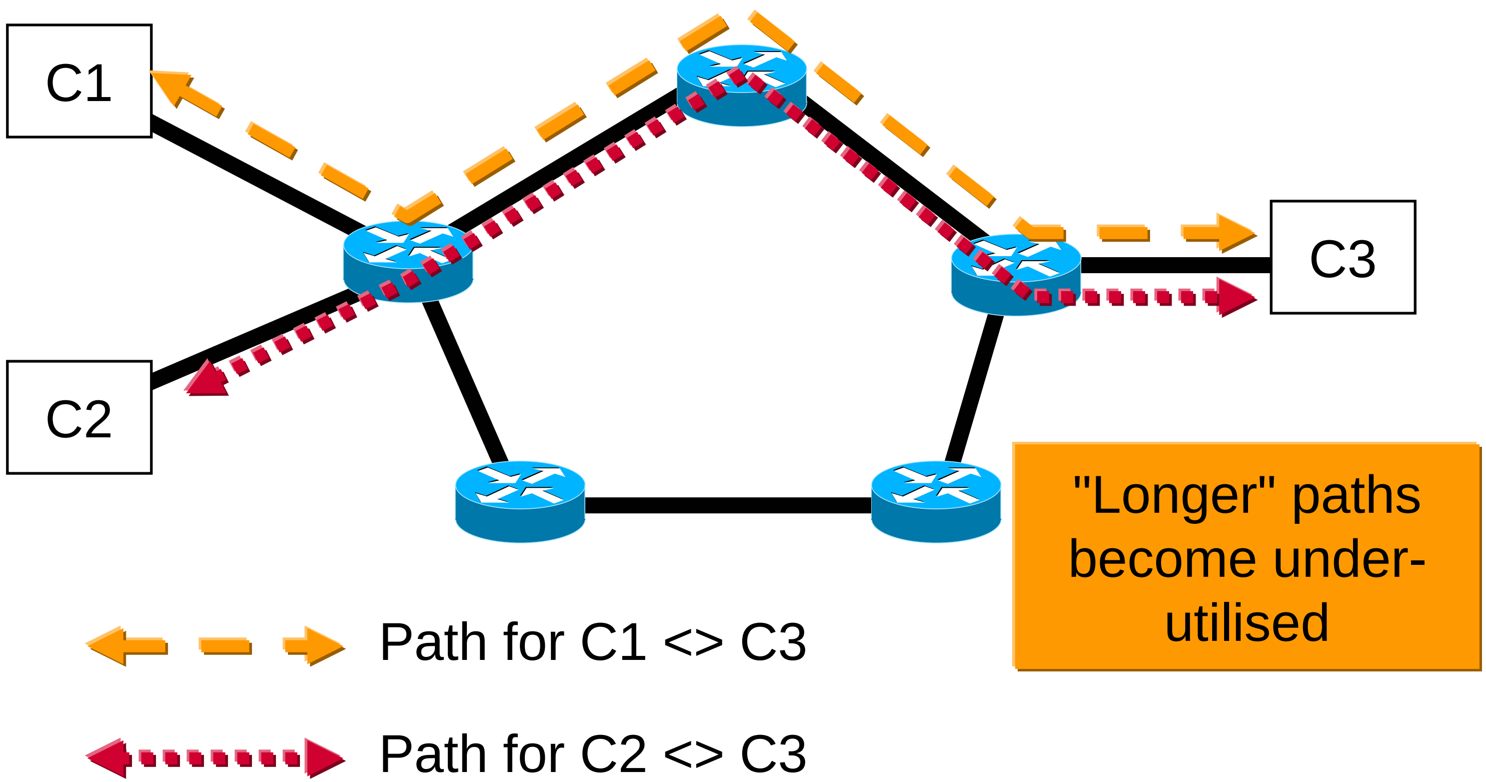
- Option 1: Direct Routes
  - Router is directly connected to this network
  - Router is told the address of its ports
- Option 2: Manually
  - a.k.a. Static Routes
- Option 3: Automatically
  - a.k.a. Routing Protocols
    - IGPs: RIP, OSPF, ISIS
    - EGPs: BGP
- Option 4: Default Route

Drawbacks of Conventional Routing

- Performance
  - ✓ *In the past*, routing was perceived as processor-limited
  - ✓ Each forwarding decision might require ~1000 machine instructions
  - ✓ Longest prefix match was difficult to transfer to silicon
  - ✓ *Today*, it is possible to build wire-speed routing in silicon
- Connectionless IP does not support Traffic Engineering
  - ✓ The "hyperaggregation problem"
- Difficulty of implementing QoS architectures

The Hyperaggregation Problem

- Routing Protocols Create A Single "Shortest Path"



Some Terminology...

- Network Engineering
  - "Put the *bandwidth* where the *traffic* is"
    - Physical cable deployment
    - Virtual connection provisioning
- Traffic Engineering
  - "Put the *traffic* where the *bandwidth* is"
    - On-line or off-line optimisation of routes
    - Implies the ability to diversify routes

Section 2: Agenda

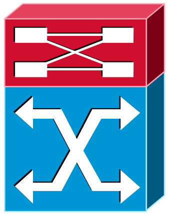
- How Does Traditional Routing Work?
  - ✓ Brief overview
  - ✓ The hyperaggregation problem
- MPLS Architecture
  - ✓ Data Plane and Control Plane
  - ✓ MPLS Terminology
  - ✓ How Does It Work?
  - ✓ Label Distribution Protocol (LDP)
  - ✓ Penultimate Hop Popping, Aggregation, TTL
  - ✓ ATM Issues

Label Switching Router

- Device that forwards *labelled packets* using *label only*
  - ✓ Has "slow" path for unlabelled traffic

- Router-based LSR

- ATM LSR



Forwarding Fundamentals

- Network Layer Routing Functional Components: Control and Forwarding
  - Control: construction and maintenance of the forwarding table. Each router has:
    - One or more routing protocols (e.g. OSPF, BGP) to exchange routing information
    - Procedures to convert routing information to a forwarding table.
  - Forwarding: actual forwarding of packets from input to output across a switch or a router.
    - Needs two sources of information: forwarding table and the information carried in the packet.
    - Consists of procedures to make a forwarding decision on a packet. The procedures:
      - Define the info from the packet to find an entry in the table
      - Specify how to find the entry.

Datagram Forwarding

- Every datagram contains a host destination’s address
- From the host destination address *“find out”* associated subnetwork
  - If destination host directly connected, then forward to host
  - If destination host not directly connected, then forward to another router based on best route
  - Forwarding table maps subnetwork number into next hop router through an interface
  - Each host has a default router sitting on the LAN (sometimes called Gateway)
  - Routing table : subnet number + mask ➔ I/F, next hop
  - Each router maintains (by learning / manual provisioning) a forwarding/routing table

Forwarding Fundamentals

How to find the an entry in the table? Examples:

- ✓ forwarding of unicast packets
  - Information from a packet used to find an entry: network layer destination address
  - Procedure: longest match algorithm
- ✓ forwarding of unicast packets with Types of Services
  - Information used to find an entry: network layer destination address and the Type of Service value
  - Procedure: longest match algorithm on the dest address and the exact match algorithm on the Type of Service value
- ✓ forwarding of multicast packets
  - Information used to find an entry: network layer source and destination addresses, and the incoming interface
  - Procedure: both the longest match and the exact match algorithms

Forwarding Equivalence Class

- Partition all possible packets into disjoint subsets.
- Packets within the same subset are treated the same (e.g., they are all set the next hop, incoming I/F, IP precedence values in the header, the packets’ destination port #, or other scheme), even if they differ from each other in the network layer header.
- Such subsets are referred to as FECs.
- Mapping of info in packets to the entries in FT any-to-one.
- All packets in the same FEC are treated the same. Examples: forwarding the packets down a certain path, providing the packet some preferential treatment within the core, or even dropping the packet.

Forwarding Equivalence Class

- **Forwarding Equivalence Class (FEC)**
  - ✓**Stream/flow of IP packets:**
    - ➢**Forwarded over the same path**
    - ➢**Treated in the same manner**
    - ➢**Mapped to the same label**
    - ➢**Multiple FEC's may be mapped to the same FEC**
    - ➢**For QoS use the Exp bits for mapping**
  - ✓**FEC/label binding mechanism**
    - ➢**Binding is done once at the ingress**
    - ➢**Currently based on destination IP address prefix**
    - ➢**Future mappings based on SP-defined policy**

Providing Consistent Routing

- Correct routing requires consistent forwarding across multiple routers. We need:
  - ✓ Control component:
    - ➢ Consistent distribution of routing information used by the routers for building forwarding tables
    - ➢ Consistent procedures to construct forwarding tables
    - ➢ Ex: OSPF procedures distribute link-state information among routers and they use the same SPF algorithm, resulting in consistent set of FECs and their next hops.
  - ✓ Forwarding component:
    - ➢ Consistent procedures for extracting the info from the packets
    - ➢ Consistent method to find an appropriate entry in the table, resulting consistent mapping of packets into FECs across multiple routers.
    - ➢ Ex: Use only the dest addr and the longest match algorithm
  - ✓ Challenge is in the distributed environment.
    - ➢ Consistent results for a distributed environment
    - ➢ Convergence time, especially for large networks

Label Switching: The Forwarding Component

- Two sources of information: forwarding table and a label carried in the packet
- **What is a label?**
  - ✓ Short, fixed-length packet identifier
  - ✓ Unstructured
  - ✓ Link local significance
  - ✓ Decouple forwarding of packets from IP headers

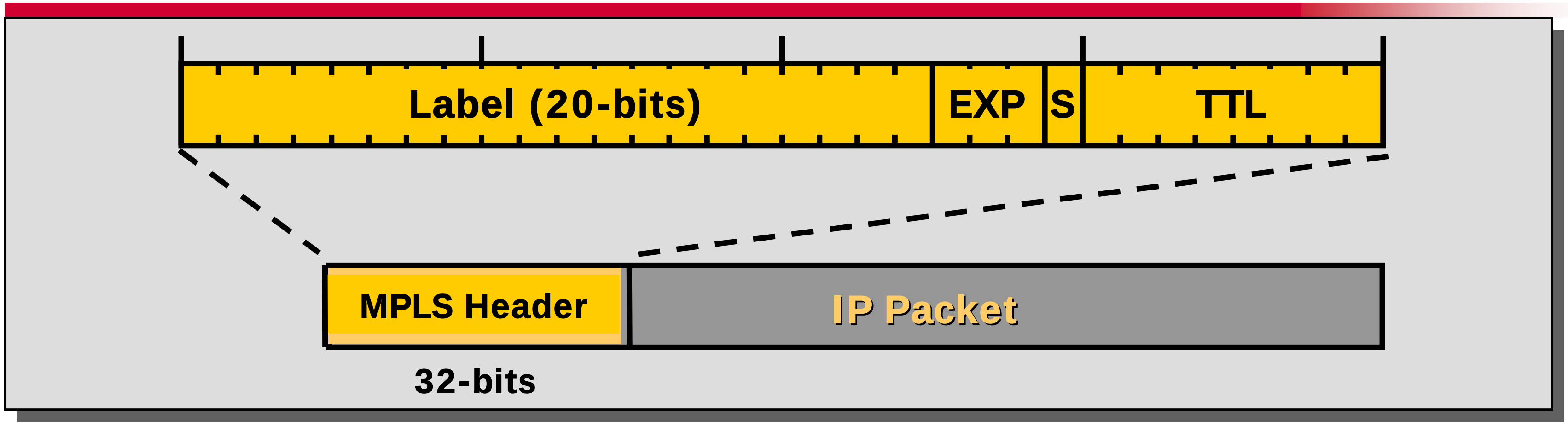
Label Switching Forwarding Tables

- A typical entry in a label switching table consists of a sequence of entries, each consisting of:
  - ✓ An incoming label and interface
  - ✓ One or more subentries (for multicast forwarding), each has:
    - ➢ An outgoing label
    - ➢ An outgoing interface
    - ➢ Next hop address
  - ✓ May include information related to what resources the packet may use, e.g., an outgoing queue.
  - ✓ The table is indexed by the value contained in the incoming label.
- An LSR may maintain a single forwarding table or a forwarding table per each of its interfaces.

Carrying a Label in a Packet

- Where to put the label?
- Need to consider Layer 2, data link layer, technologies, such as ATM and Frame Relay.
- ATM and Frame Relay can carry a label. But using this approach limit the usefulness of label switching, because some media such as Ethernet, Token Ring, or point-to-point links do not support it.
- Can’t use data link layer header. Instead, we use a small “shim” label header that is inserted between the link layer and the network layer headers.

MPLS Label

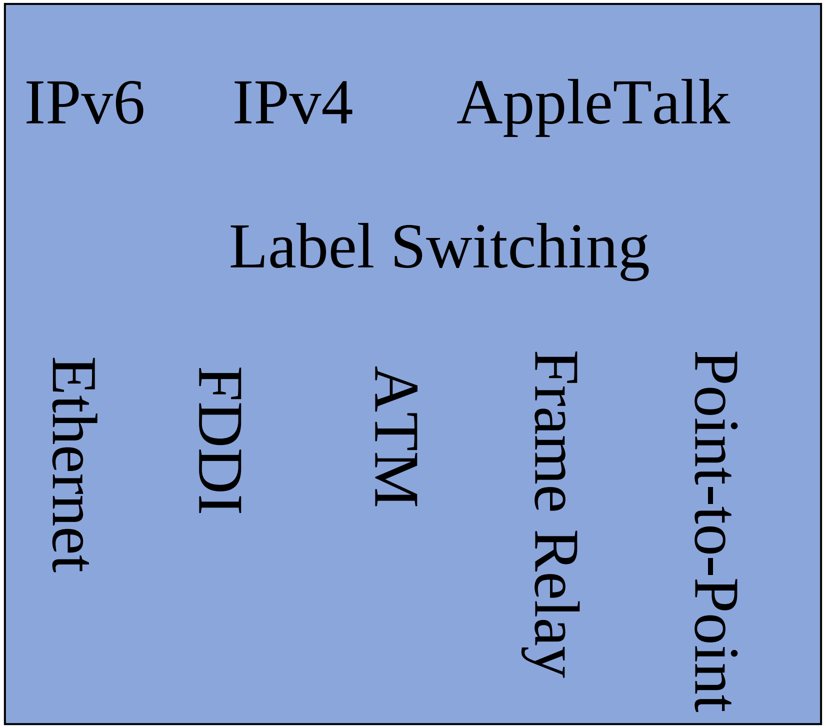


- **Fields**
  - ✓ **Label**
  - ✓ **Experimental (can be used for CoS mapping)**
  - ✓ **Stacking bit**
  - ✓ **Time to live**

Label Switching Forwarding Algorithm

- The forwarding algorithm used by the forwarding component of label switching is based on label swapping.
- There is only one **single forwarding algorithm**.
- In the conventional approach, different functionality provided by the control component (e.g., unicast, unicast with ToS, multicast) requires multiple algorithms in the forwarding component.
- With label switching, only one is needed: label swapping.

Multiprotocol: Both Above and Below



Network Layer Protocols

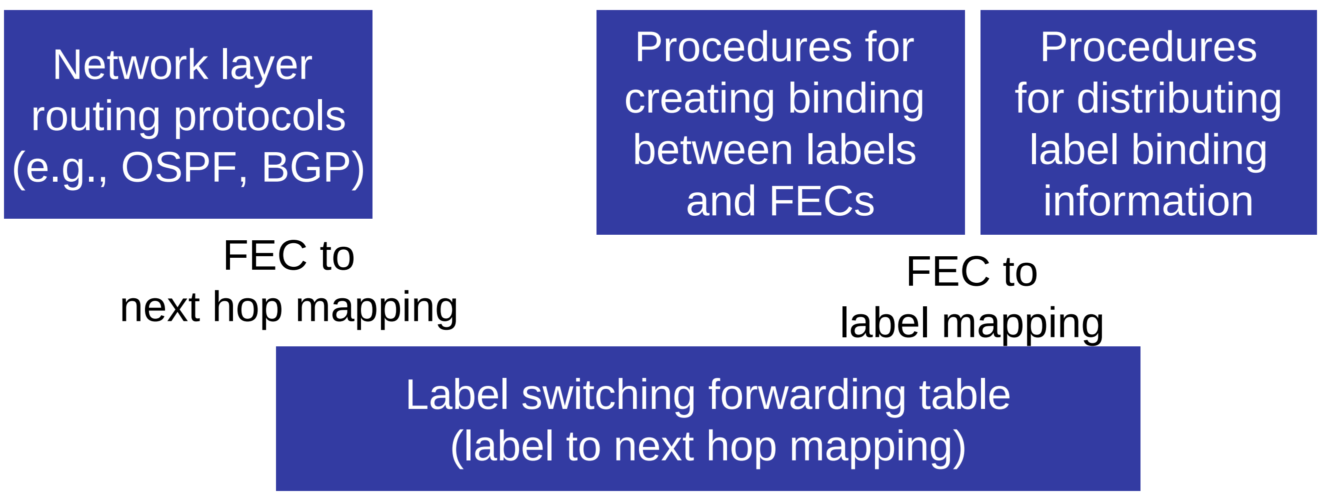
Link Layer Protocols

Label Switching: The Control Component

- Responsible for:
  - ✓ Distributing routing information among LSRs
  - ✓ The procedures for LSRs to convert the routing information into a forwarding table that is used by the forwarding component
- Similarly, it must provide consistent distribution of routing information and consistent procedures for building the forwarding table.
- Includes all the routing protocols (e.g., OSPF, BGP, …) for the conventional routing.
- But it needs to contain **mappings between labels and next hops**
  - ✓ Local
  - ✓ Neighbors

Label Switching: The Control Component

- So, additional procedures are needed for LSRs:
  - ✓ Create binding between labels and FECs
  - ✓ Inform other LSRs of the bindings it creates
  - ✓ Construct and maintain the forwarding table



Local vs Remote Binding

- The label switching control component provides two types of label bindings for incoming & outgoing labels:
  - ✓ Local binding: router creates the binding with a label chosen and assigned locally
  - ✓ Remote binding: router receives from other LSR label binding info

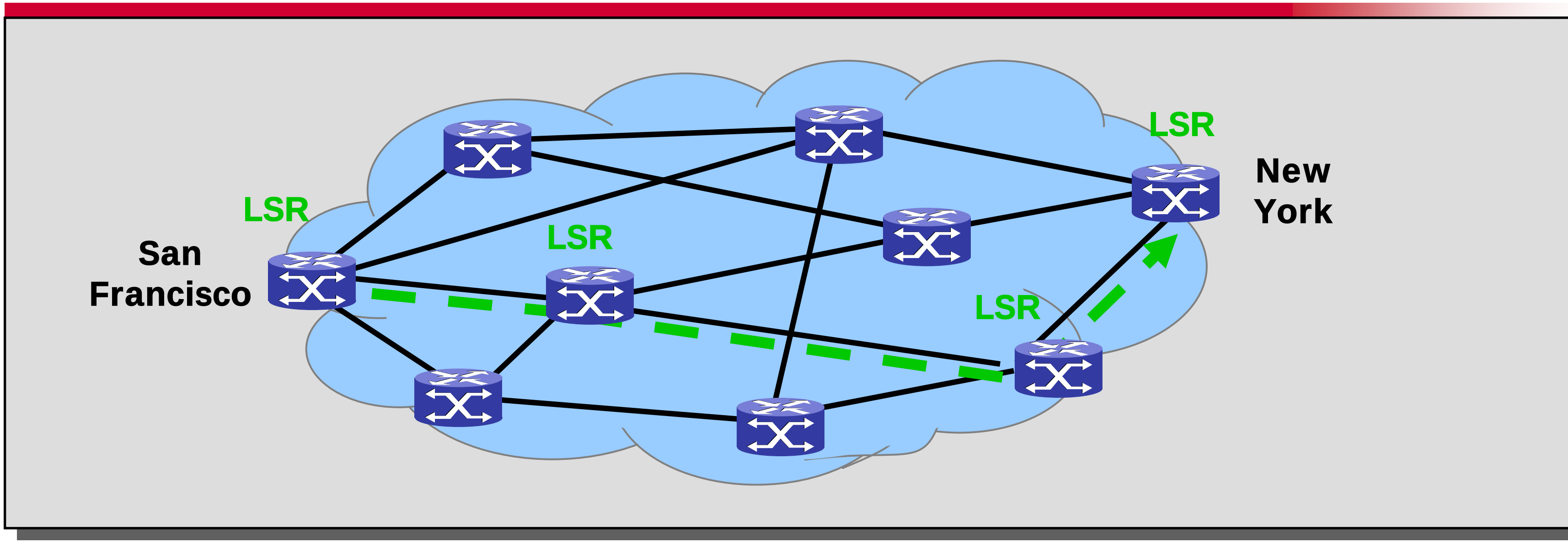
Upstream vs Downstream Binding

- Local and remote bindings are usually used together.
  - ✓ Labels from the local binding are used as incoming labels and labels from the remote binding are used as outgoing labels
  - ✓ The opposite
- Downstream label binding: binding between a label and a FEC for the packet is created by a downstream LSR.
- Upstream label binding: the opposite

Control-Driven vs Data-Driven

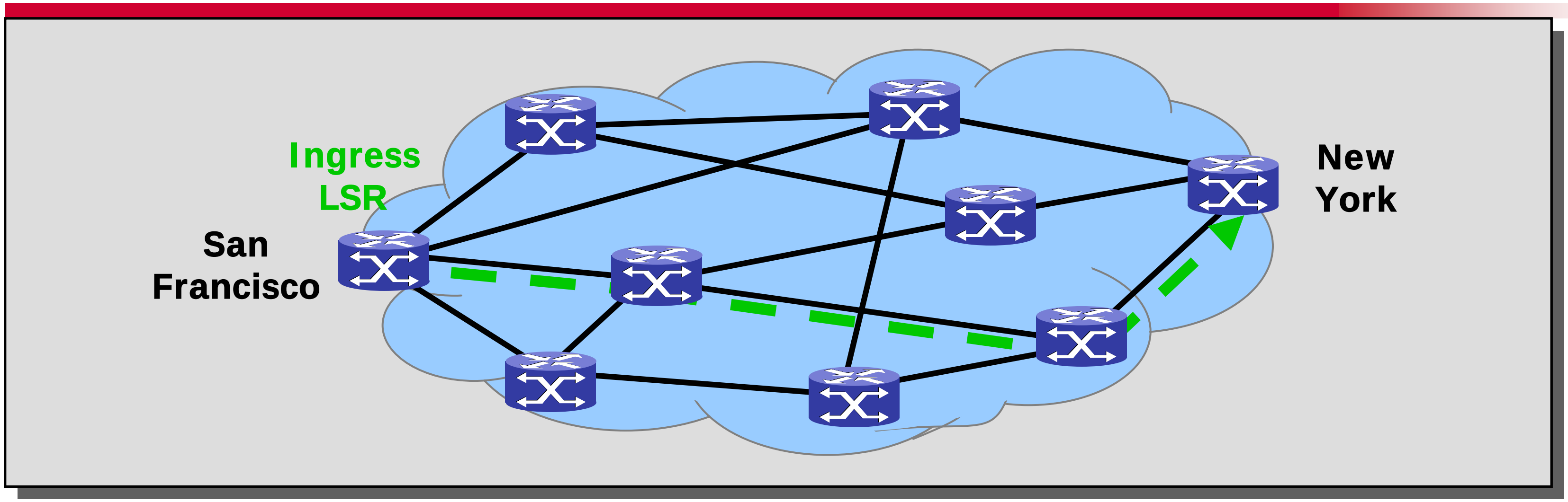
- LSRs can create or destroy a binding between a label and an FEC triggered either by
  - ✓ Data packets that have to be forwarded by the LSR, or
  - ✓ By control (routing) information (e.g., OSPF routing updates, RSVP PATH/RESV messages) that has to be processed by the LSR.
- The former is data-driven; the latter, control-driven
- Have impact on performance, scalability, robustness, depending on the traffic scenarios
- Generally, control-driven is simpler and more robust.

Label Switching Router (LSR)



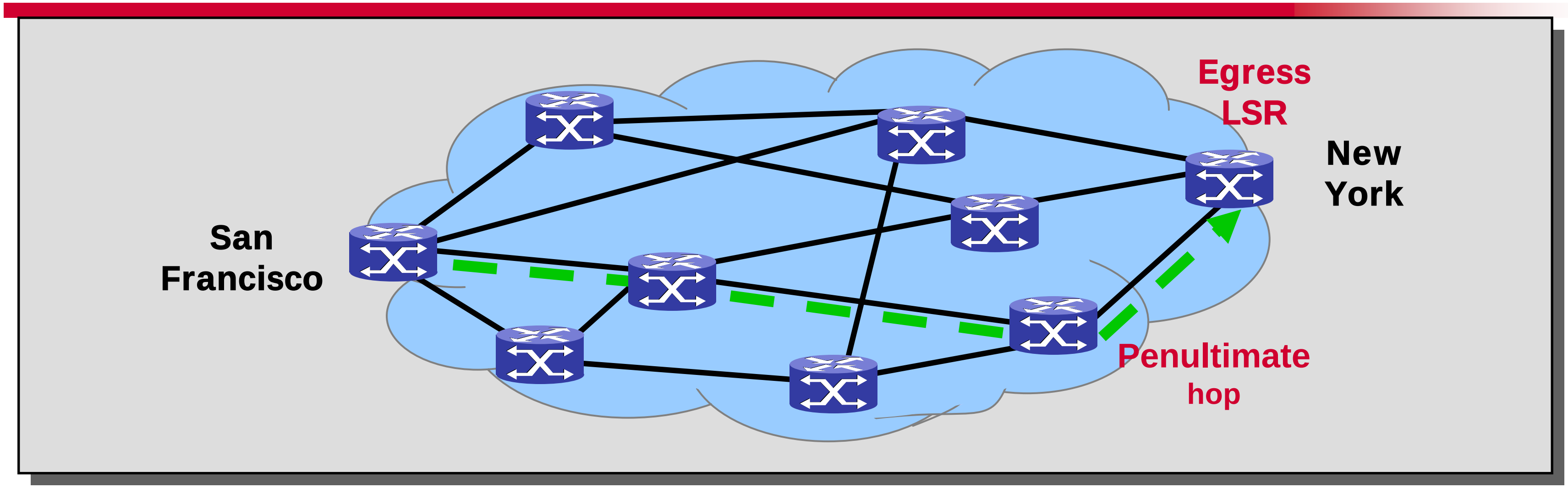
- **Label-Switching Router (LSR)**
  - ✓ **Forwards MPLS packets using label-switching**
  - ✓ **Capable of forwarding native IP packets**
  - ✓ **Executes one or more IP routing protocols**
  - ✓ **Participates in MPLS control protocols**

Ingress Router
Label Edge Router (LER)



- **Ingress LSR**
  - ✓ **Examines inbound IP packets**
  - ✓ **Classifies packet to an FEC**
  - ✓ **Generates MPLS header and assigns (binds) initial label**
  - ✓ **Upstream from all other LSRs in the LSP**
  - ✓ **All other routers inside the MPLS domain look at the labels only, not at the IP address**

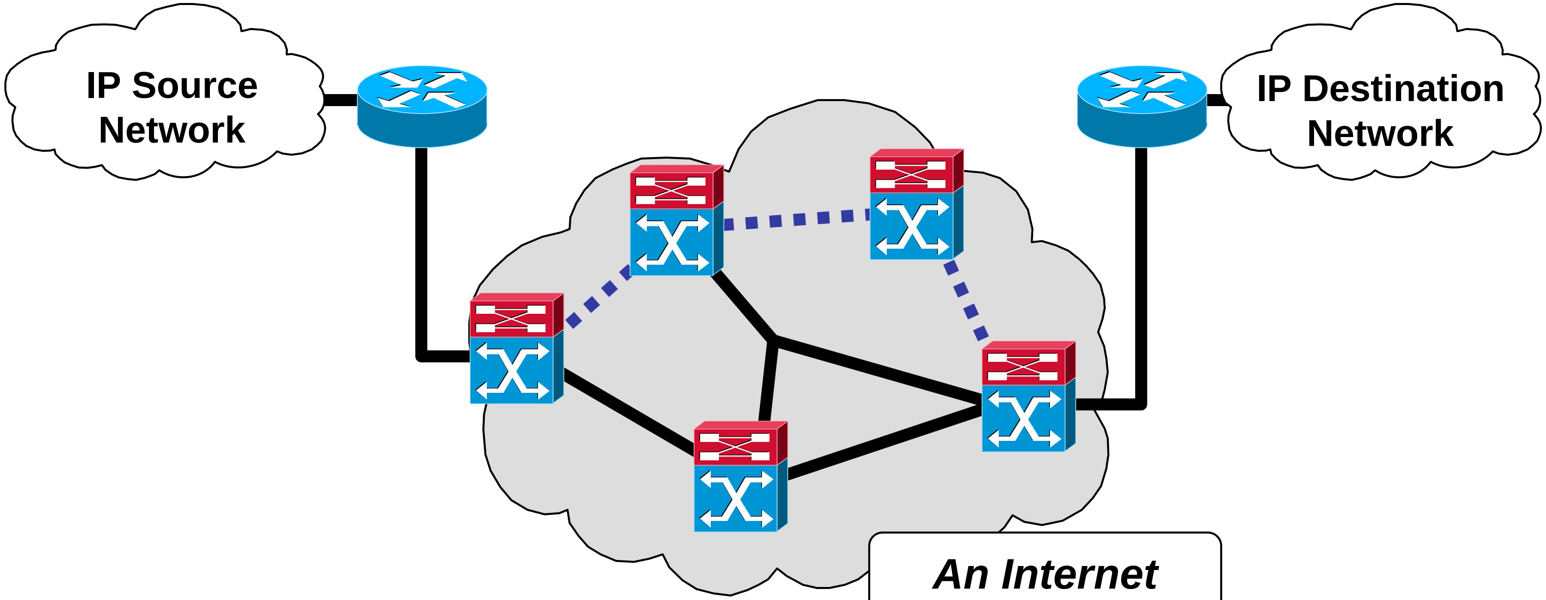
Egress Router
Label Edge Router (LER)



- **Egress LSR**
  - ✓ **Processes traffic as it leaves the MPLS domain – based on IP packet destination address**
  - ✓ **Removes the MPLS header – unless the "Penultimate hop" router already had removed it.**
  - ✓ **Downstream from all other LSRs in the LSP**

Label Switched Path

- The path followed by labelled packets that are assigned to the same FEC



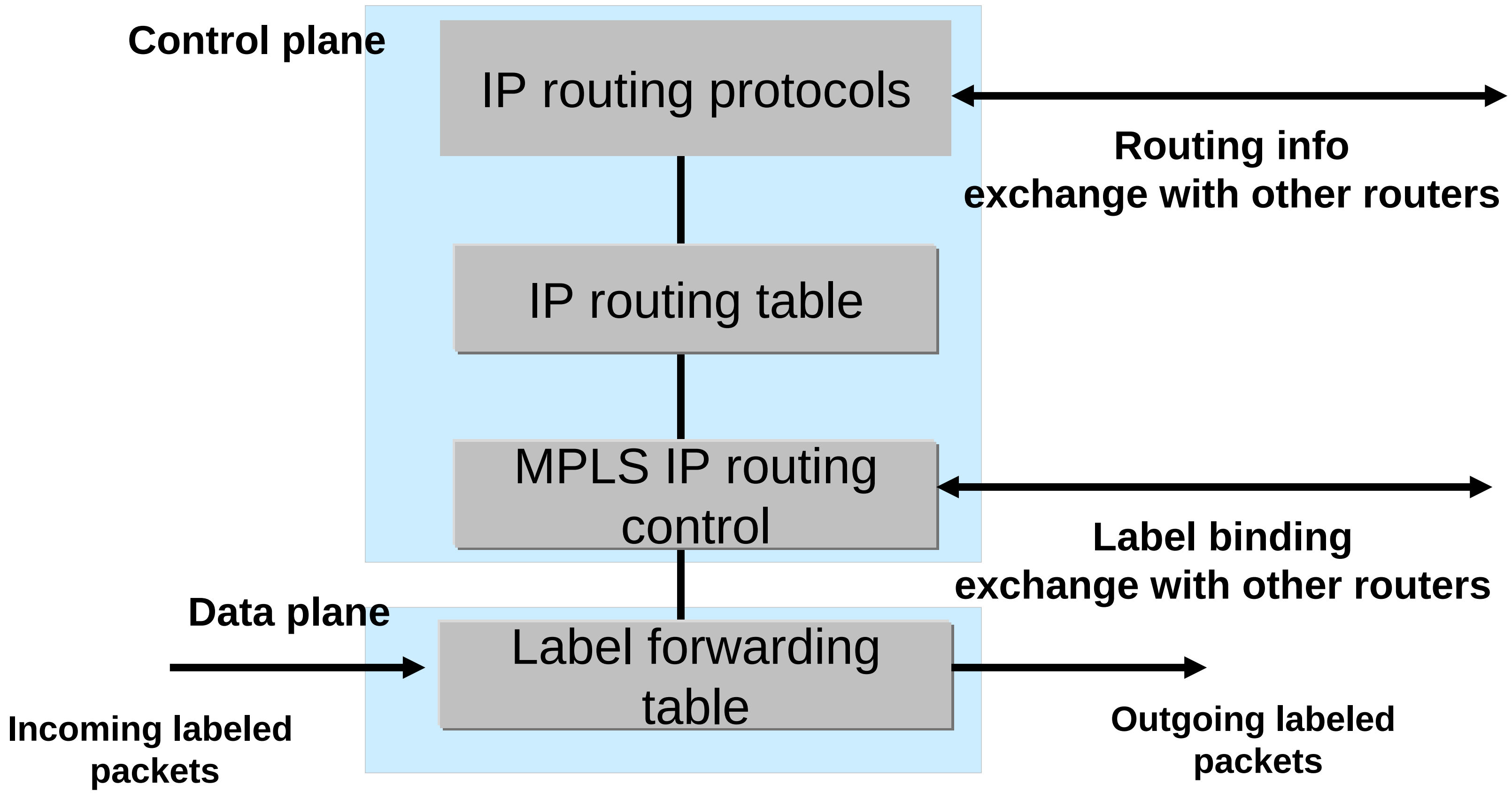
Data Plane vs Control Plane (1)

- For Routers…
  - ✓ The Routing Table
  - ✓ Routing Protocols
- For ATM Switches
  - ✓ PNNI
- For MPLS LSRs
  - ✓ Topology, Label Distribution and Explicit Routing Protocols

The data plane actually carries the information while the control plane sets up pathways through the data plane

MPLS LSRs and MPΛS OXCs (later GMPLS) solve performance scalability problem by decoupling control and data planes

Control Plane vs Data Plane (2)



Section 2: Agenda

- How Does Traditional Routing Work?
  - ✓ Brief overview
  - ✓ The hyperaggregation problem
  - ✓ Data Plane and Control Plane
- MPLS Architecture
  - ✓ MPLS Terminology
  - ➡ How Does It Work?
  - ✓ Label Distribution Protocol (LDP)
  - ✓ Penultimate Hop Popping, Aggregation, TTL
  - ✓ ATM Issues

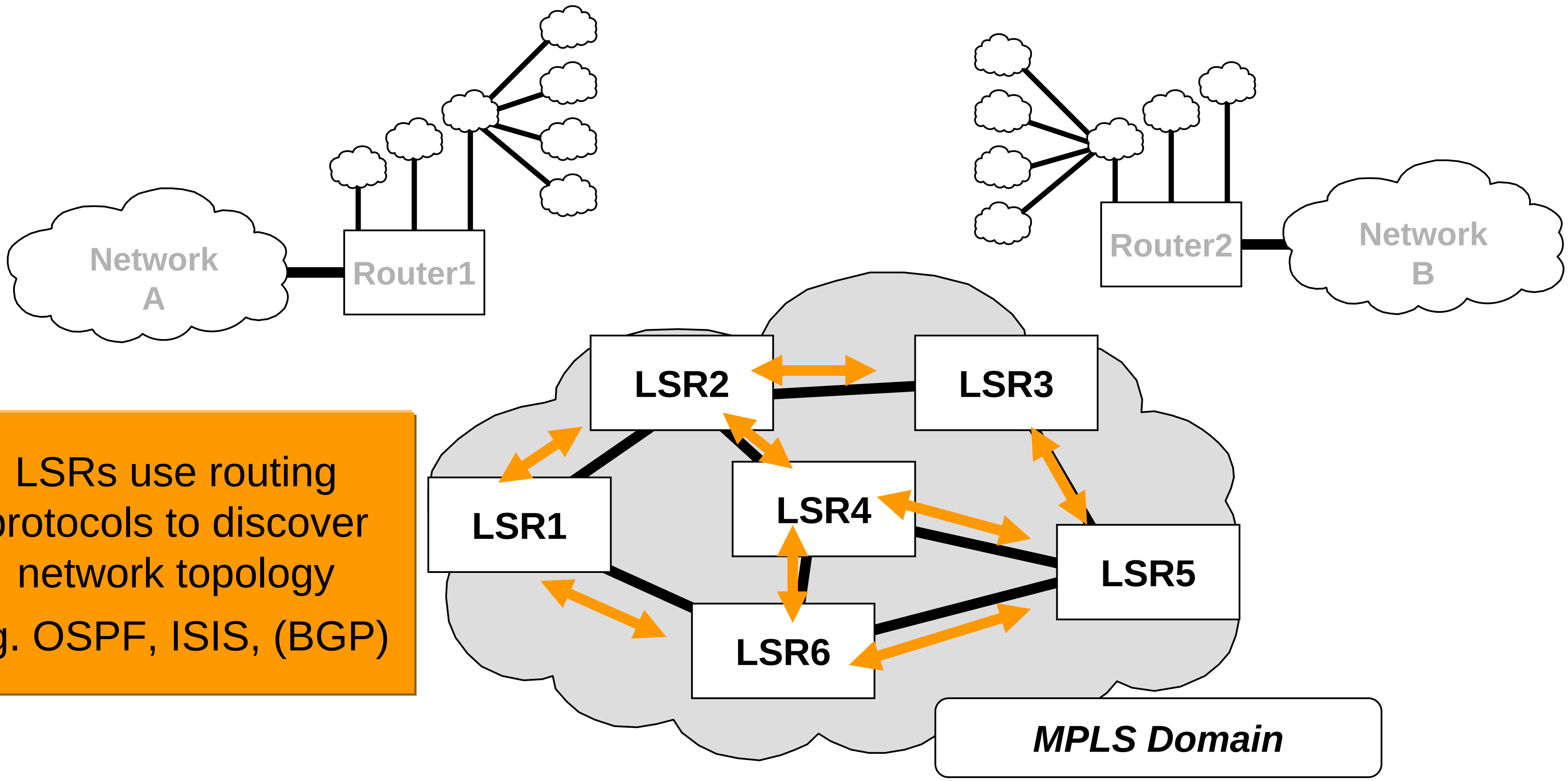
Steps in the process

- Topology determination
- Path creation
- Data forwarding

Steps in the process

- Topology determination
- Path creation
- Data forwarding

Toplogy Determination
What happens when we switch on?



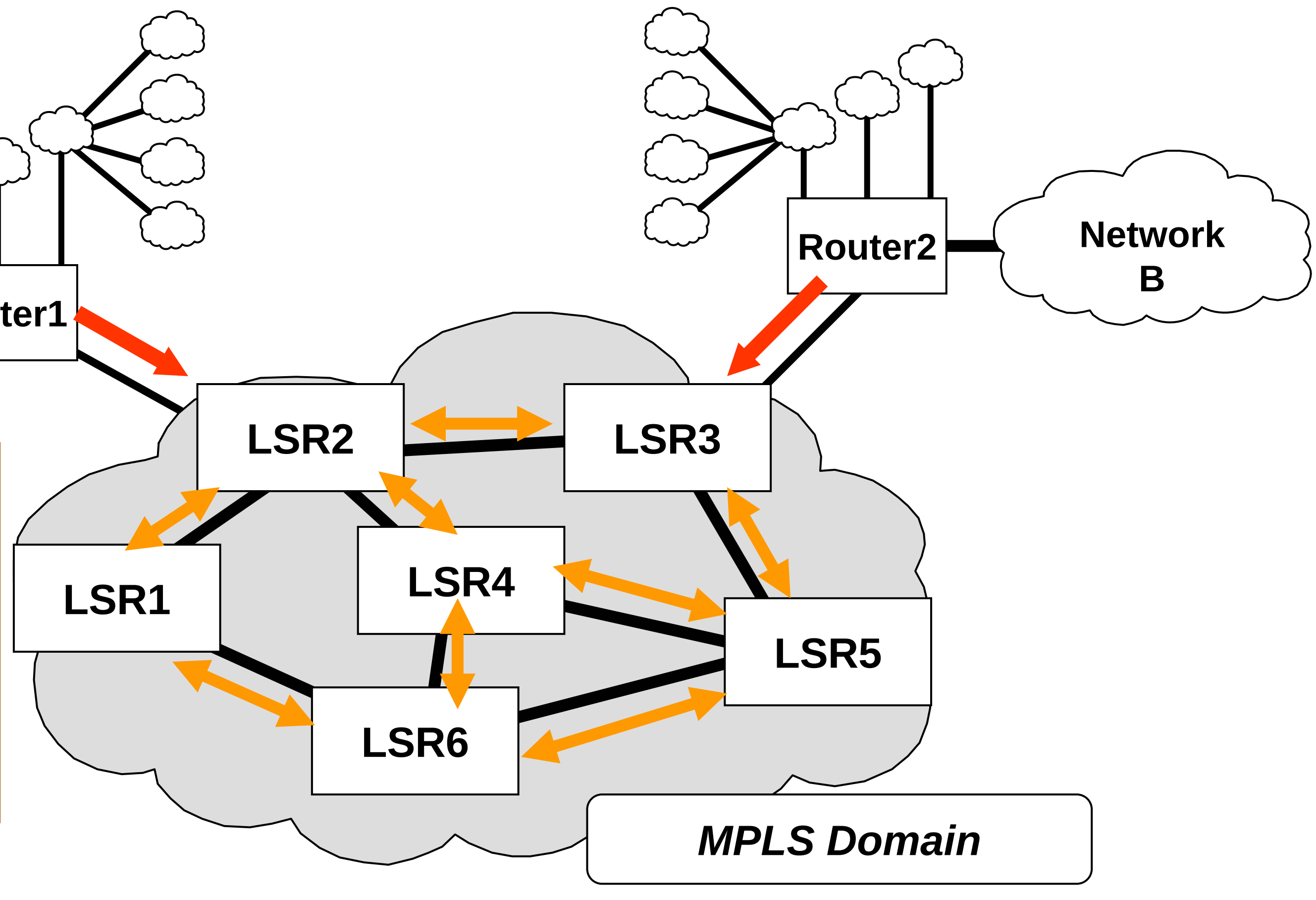
LSRs use routing protocols to discover network topology

eg. OSPF, ISIS, (BGP)

Why is there more than one protocol?

- OSPF, ISIS, BGP-4
  - ✓ OSPF and ISIS are IGPs, BGP is an EGP
  - ✓ Service providers prefer ISIS to OSPF (generalisation)

Topology Determination
What happens when we add IP networks?



IP networks advertise their addresses using routing protocols into the MPLS cloud

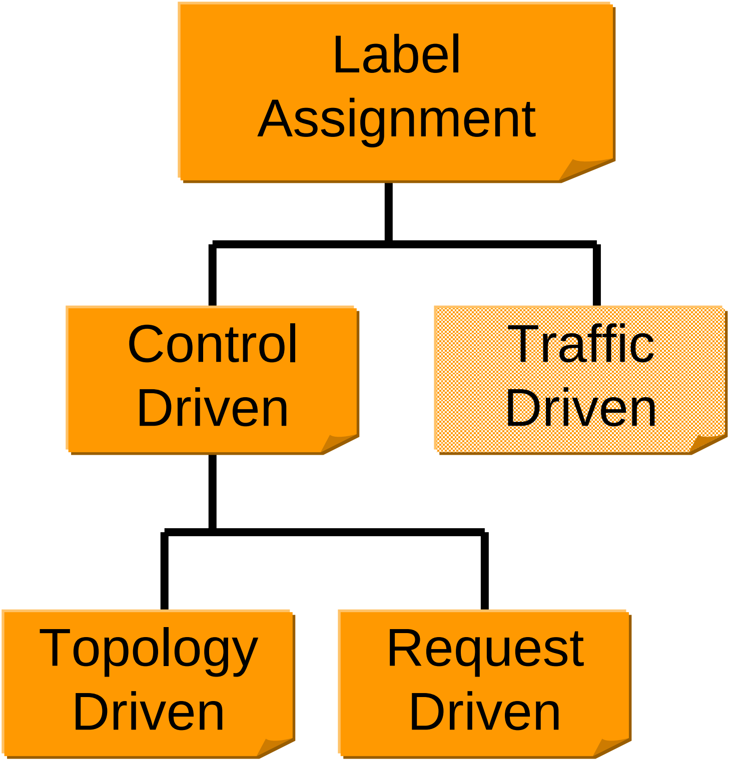
Steps in the process

- Topology determination
- Path creation
- Data forwarding

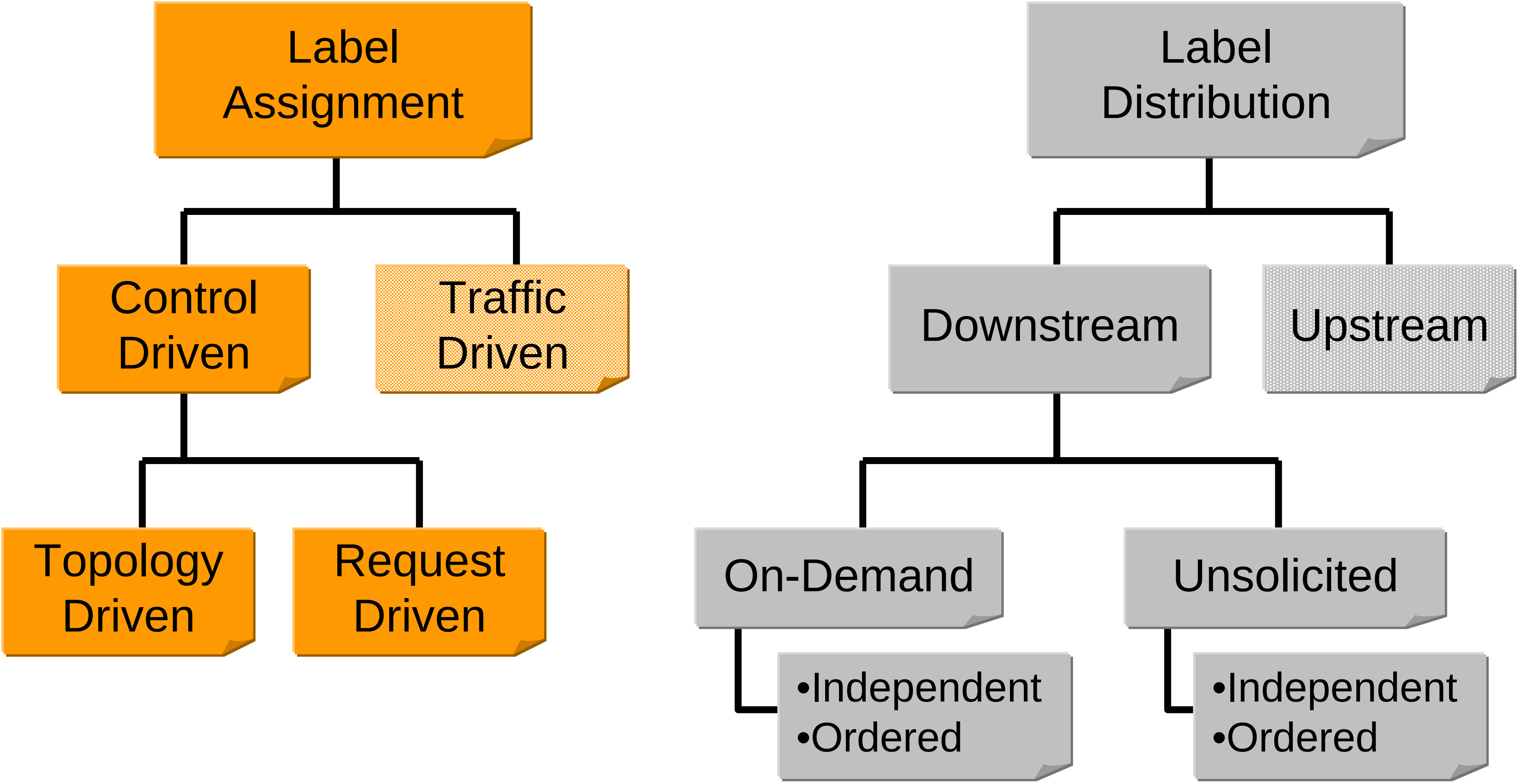
Label Distribution Concepts

- Actually distributed include a label, an IP prefix, and a mask length, (and QoS information if needed,) but called label distribution
- Ordered vs Independent Control
  - ✓ Ordered: The LSR waits to receive bindings from up/down-stream nodes before sending the labels it generated to its down/up-stream neighbors: RSVP.
  - ✓ Independent: LSRs are free to distribute label bindings to their neighbors: LDP.
- Unsolicited vs On-Demand
  - ✓ Unsolicited: An LSR advertises labels for all prefixes in its IGP to all neighbors
  - ✓ On-Demand: LSRs do not hand out labels for prefixes unless they are asked.
- Liberal vs Conservative Retention: if LSRs receive labels bindings that are not routing next hops, they may choose to:
  - ✓ Liberal: Keep the bindings for future use in case the LSRs that sent these bindings become next hops
  - ✓ Conservative: Discard label bindings that are not currently useful.

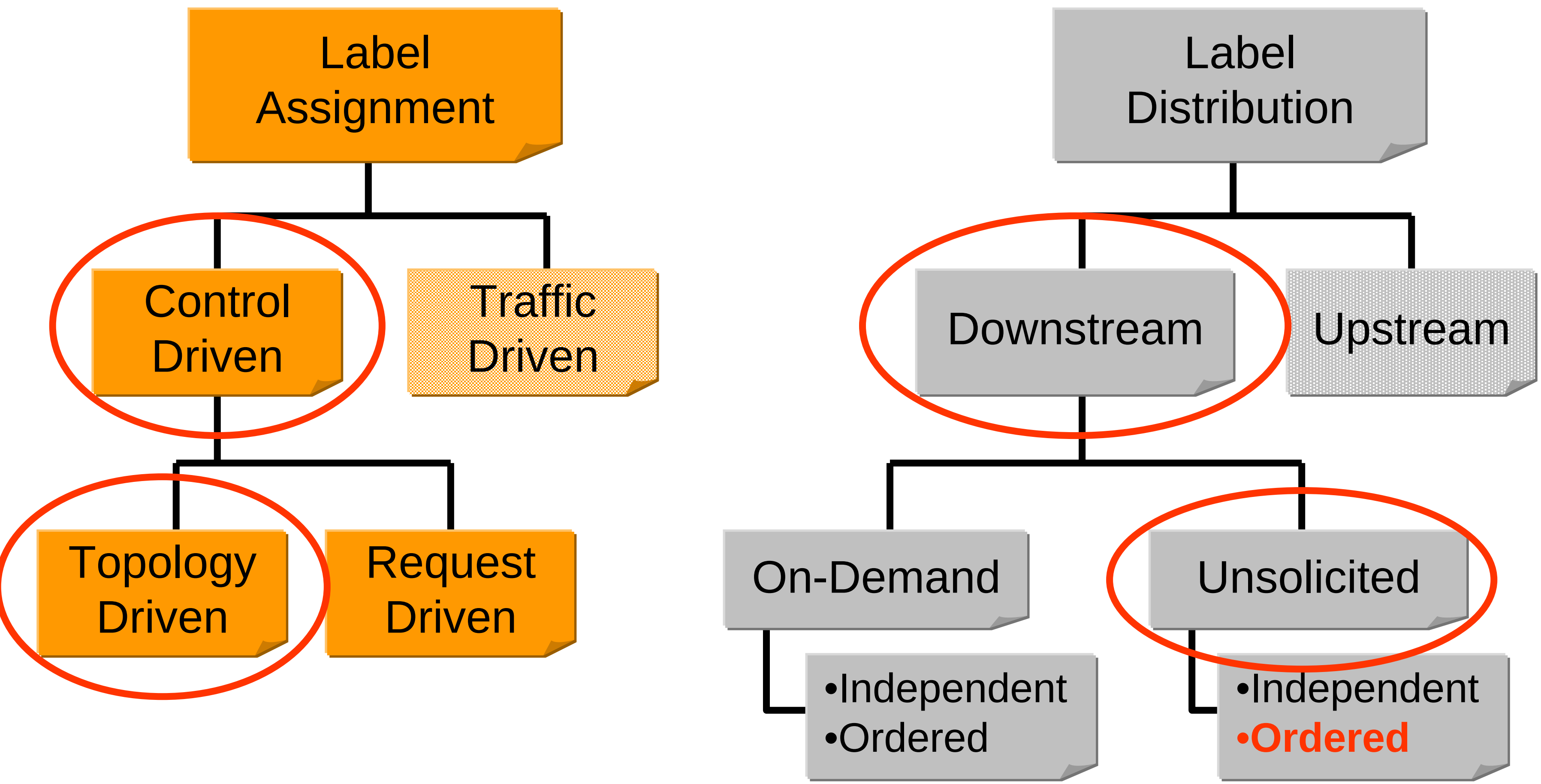
Label Assignment & Distribution



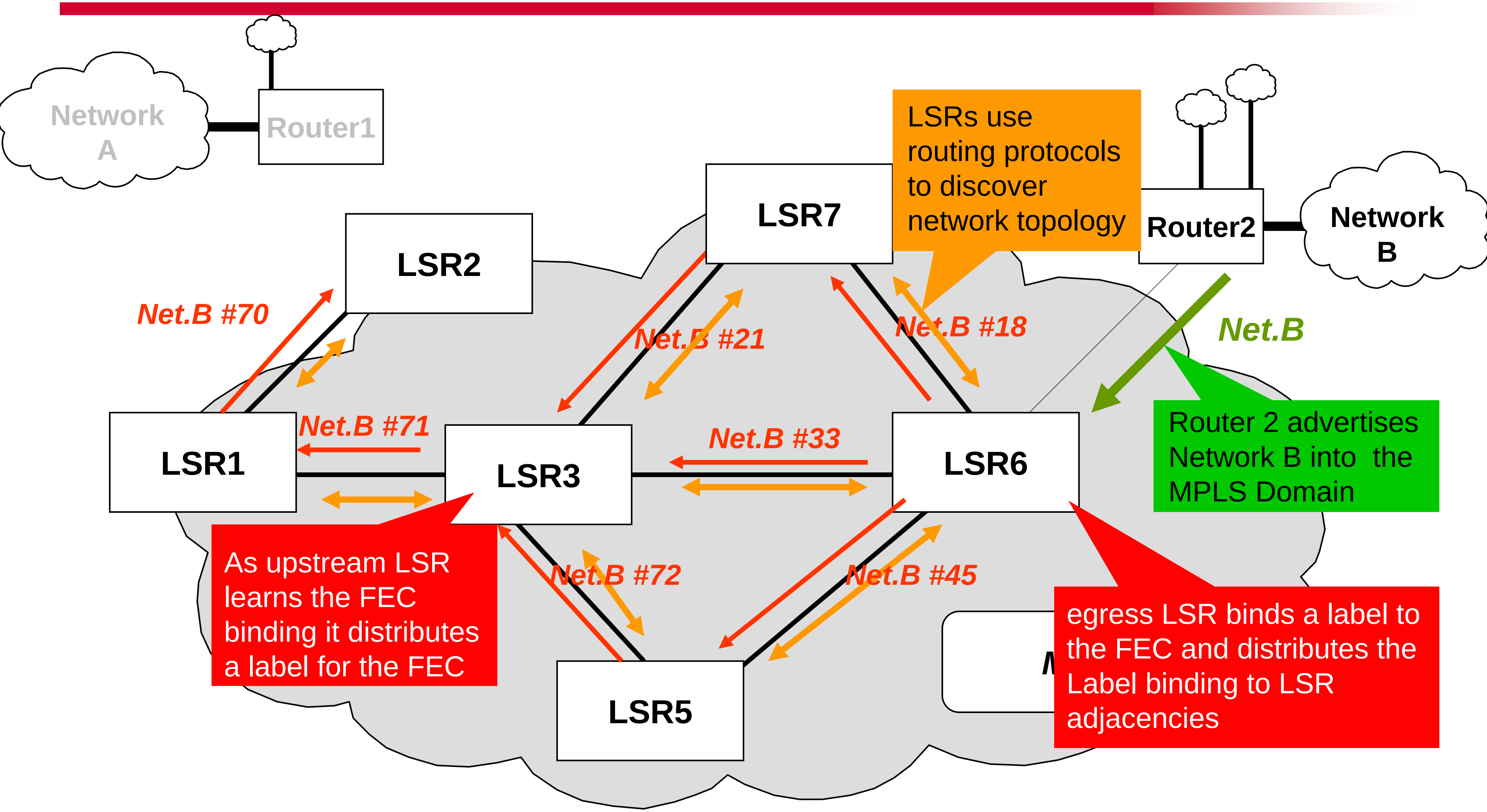
Label Assignment & Distribution



Label Assignment & Distribution



Label Distribution
Downstream Unsolicited - *Ordered*



Downstream Unsolicited - Ordered

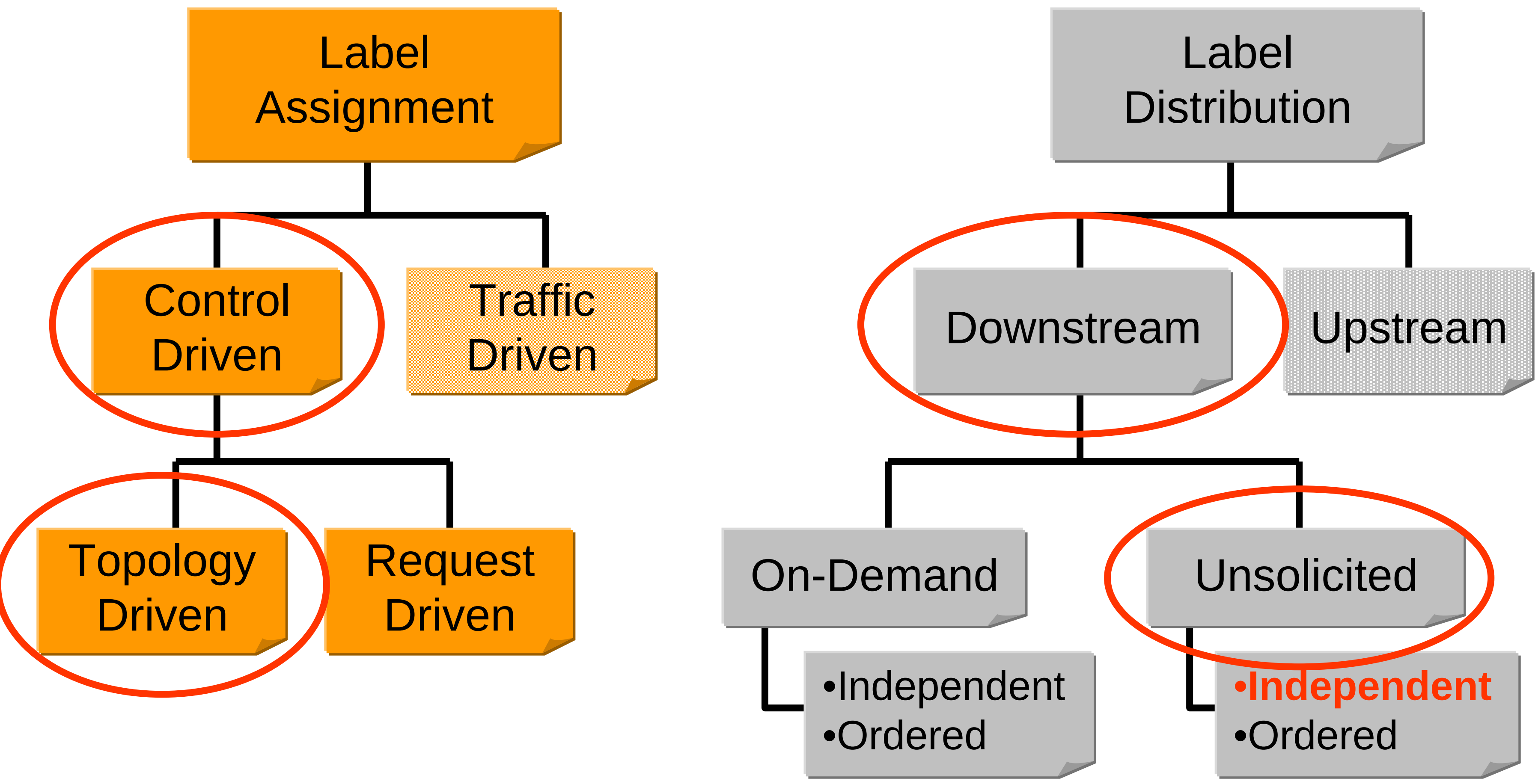
Downstream Unsolicited

- LSRX and LSRY are said to have an “LDP adjacency” (LSRX being the downstream of LSRY)
- LSRX discovers a ‘next hop’ for a particular FEC
- LSRX generates a label for the FEC and communicates the binding to LSRY
- LSRY inserts the binding into its forwarding tables
- If LSRX is the next hop for the FEC, LSRY can use that label knowing that its meaning is understood

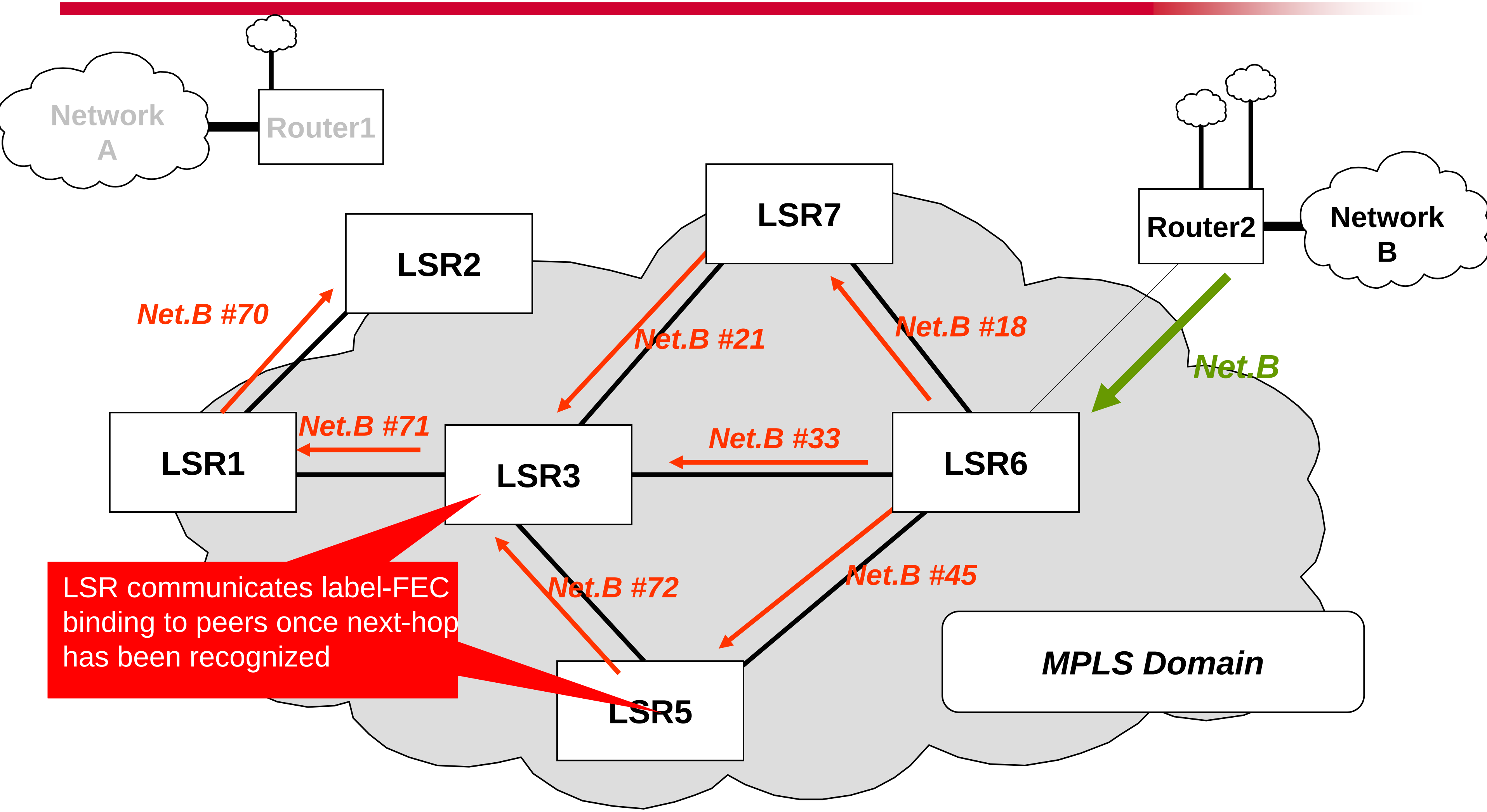
Ordered Control

- Label-FEC binding is communicated to peers if:
  - LSR is the ‘egress’ LSR to particular FEC
  - label binding has been received from downstream LSR
- LSP formation ‘flows’ from egress to ingress

Label Assignment & Distribution



Label Distribution
Downstream Unsolicited - *Independent*



Unsolicited - Independent

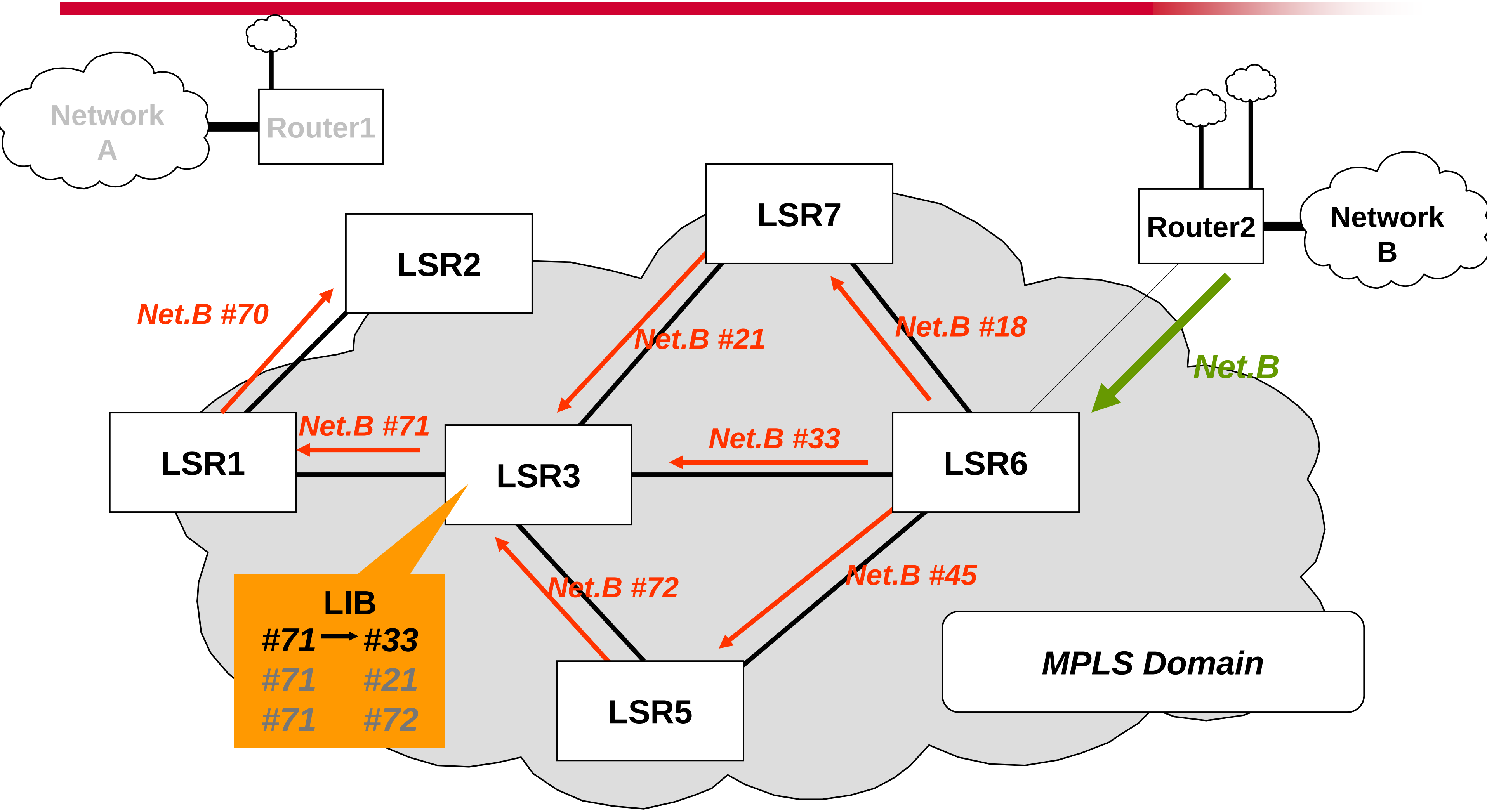
Downstream Unsolicited

- LSRX and LSRY are said to have an "LDP adjacency" (LSRX being the downstream of LSRY)
- LSRX discovers a 'next hop' for a particular FEC
- LSRX generates a label for the FEC and communicates the binding to LSRY
- LSRY inserts the binding into its forwarding tables
- If LSRX is the next hop for the FEC, LSRY can use that label knowing that its meaning is understood

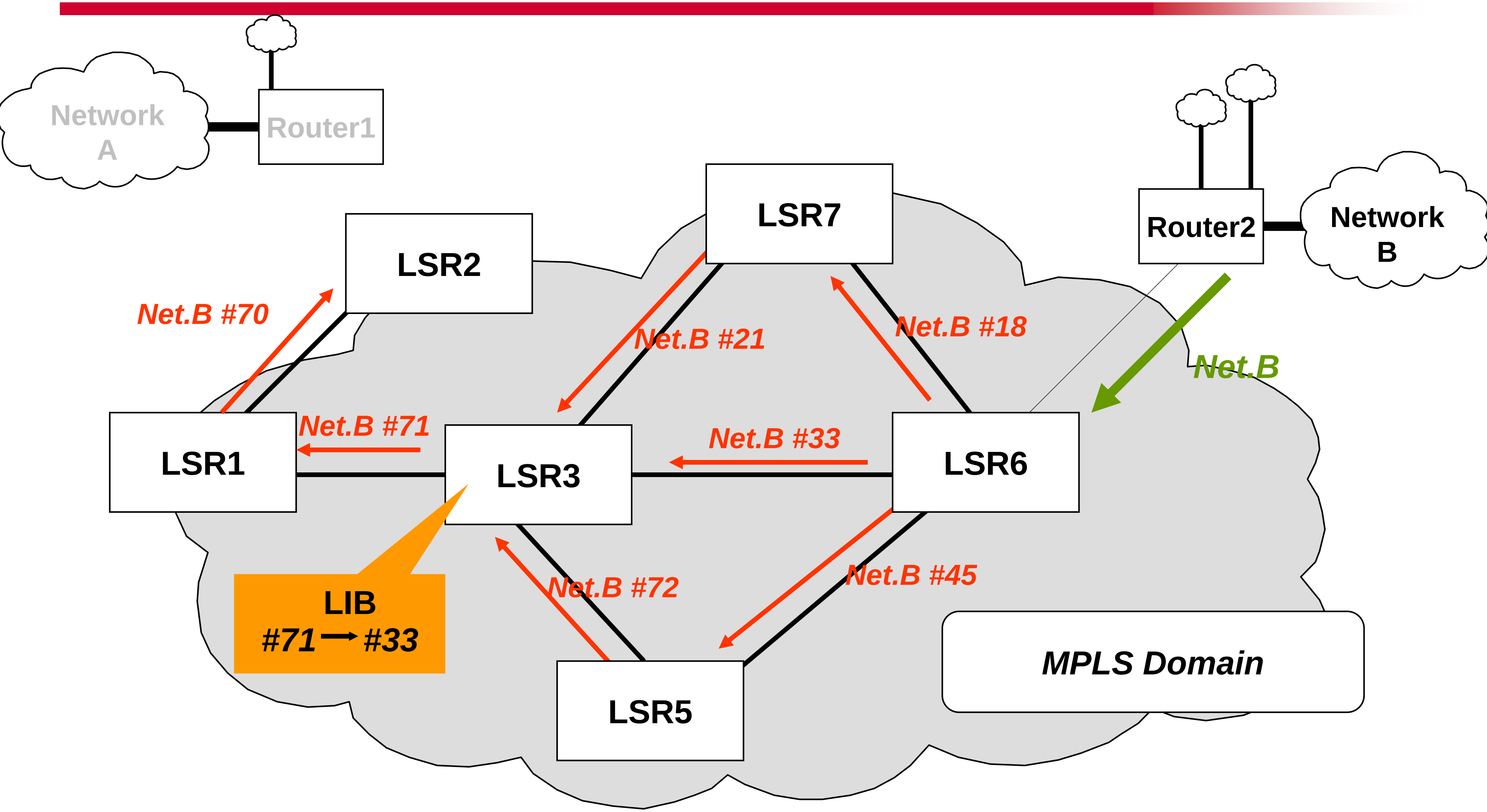
Independent Control

- Each LSR makes independent decision on when to generate labels and communicate them to upstream peers
- Communicate label-FEC binding to peers once next-hop has been recognized
- LSP is formed as incoming and outgoing labels are spliced together

Label Distribution
Liberal Label Retention



Label Distribution
Conservative Label Retention



Liberal vs Conservative Label Retention

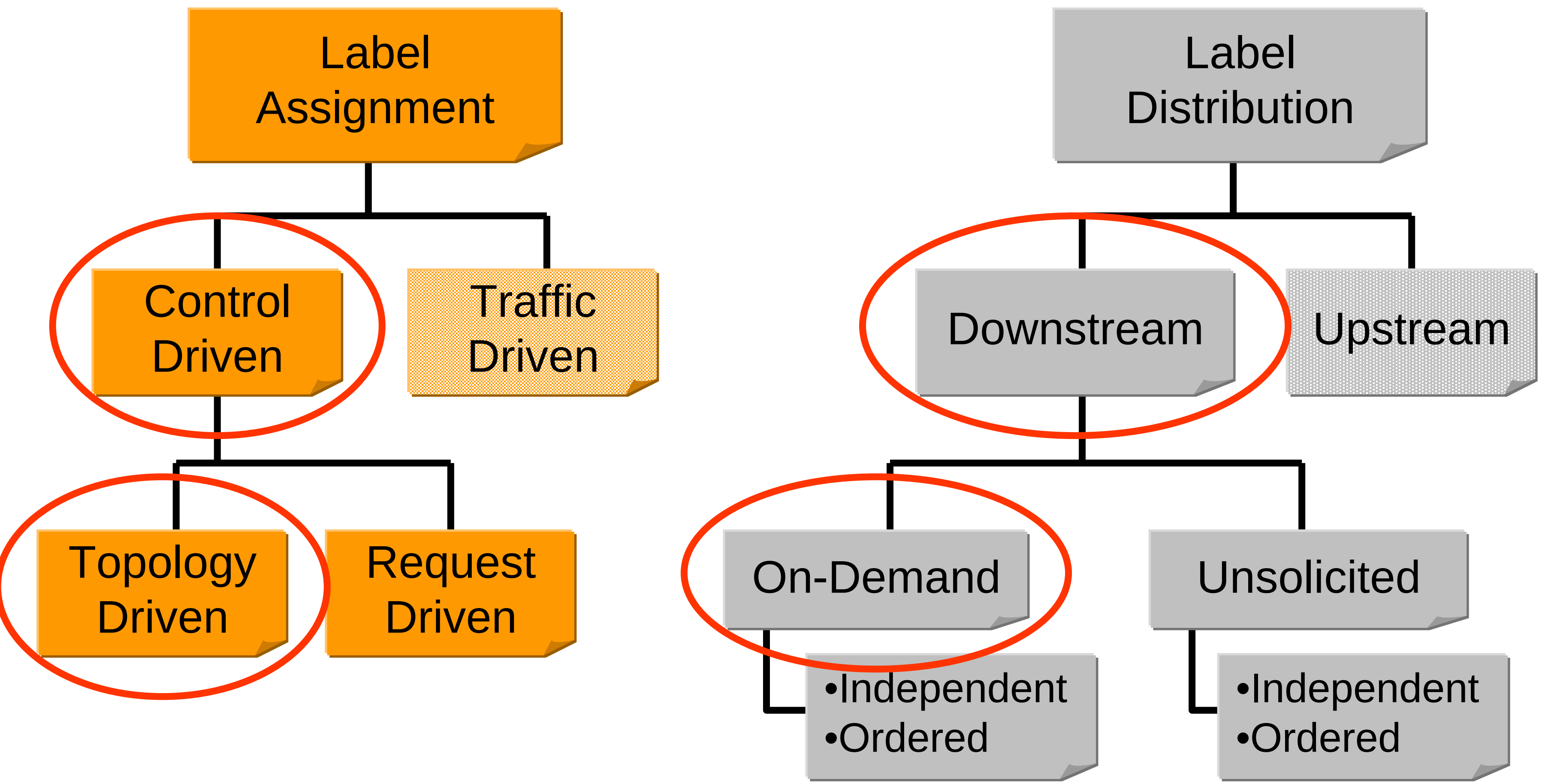
Liberal

- LSR maintains bindings received from LSRs other than the valid next hop
- If the next-hop changes, it may begin using other bindings immediately
- May allow more rapid adaptation to routing changes
- Requires an LSR to maintain many more labels

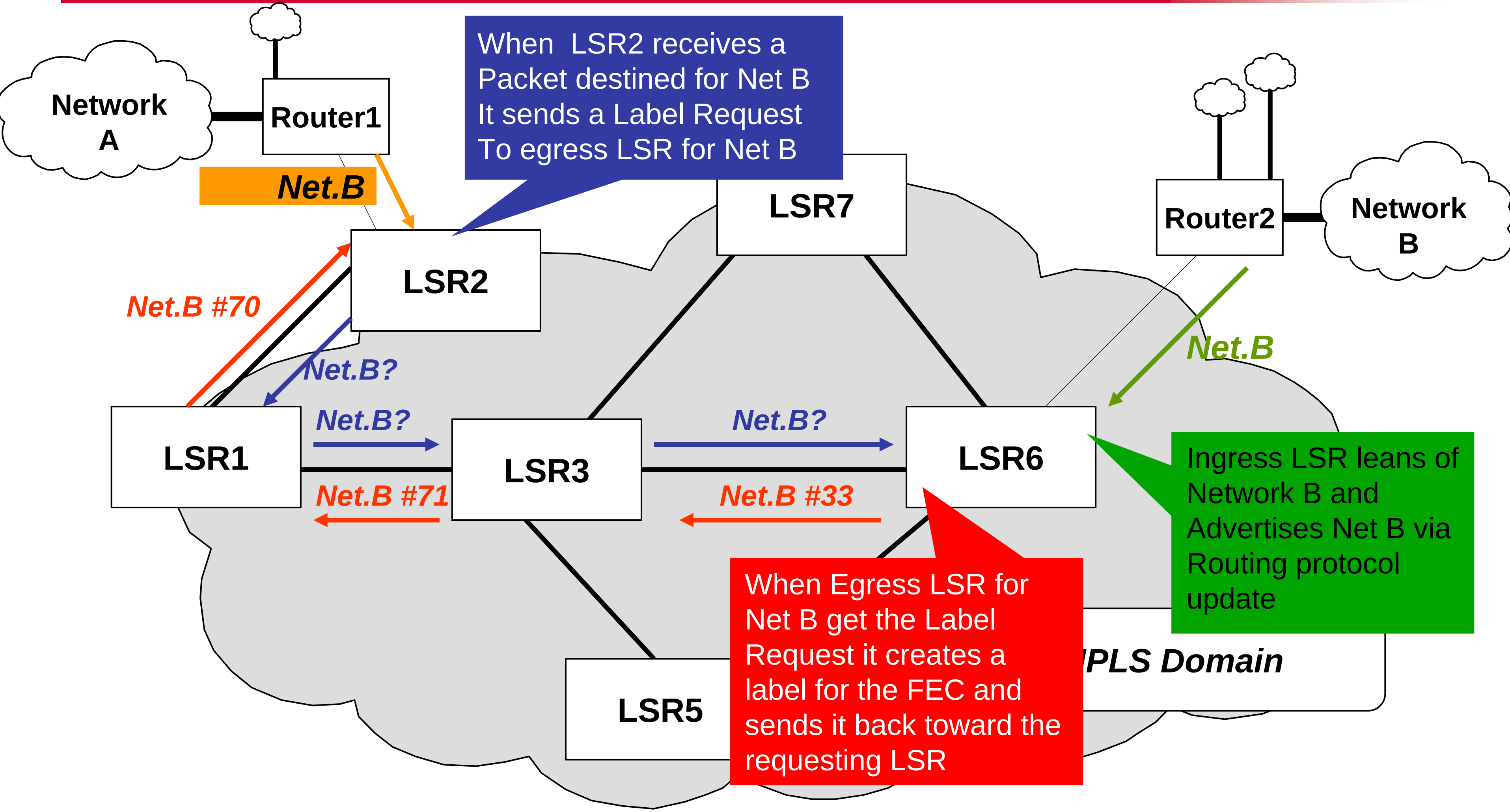
Conservative

- LSR only maintains bindings received from valid next hop
- If the next-hop changes, binding must be requested from new next hop
- Restricts adaptation to changes in routing
- Fewer labels must be maintained

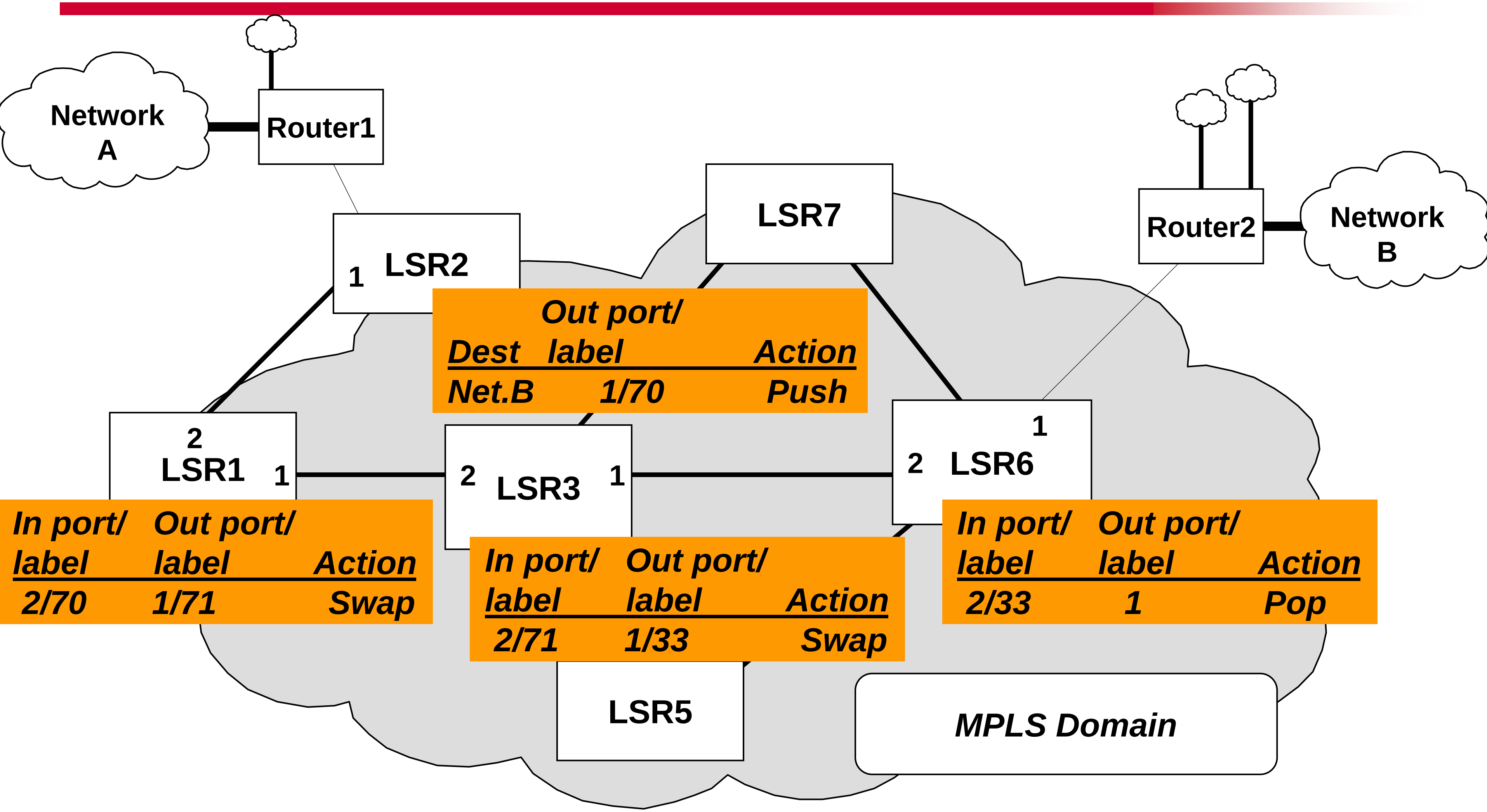
Label Assignment & Distribution



Label Distribution Downstream-On-Demand



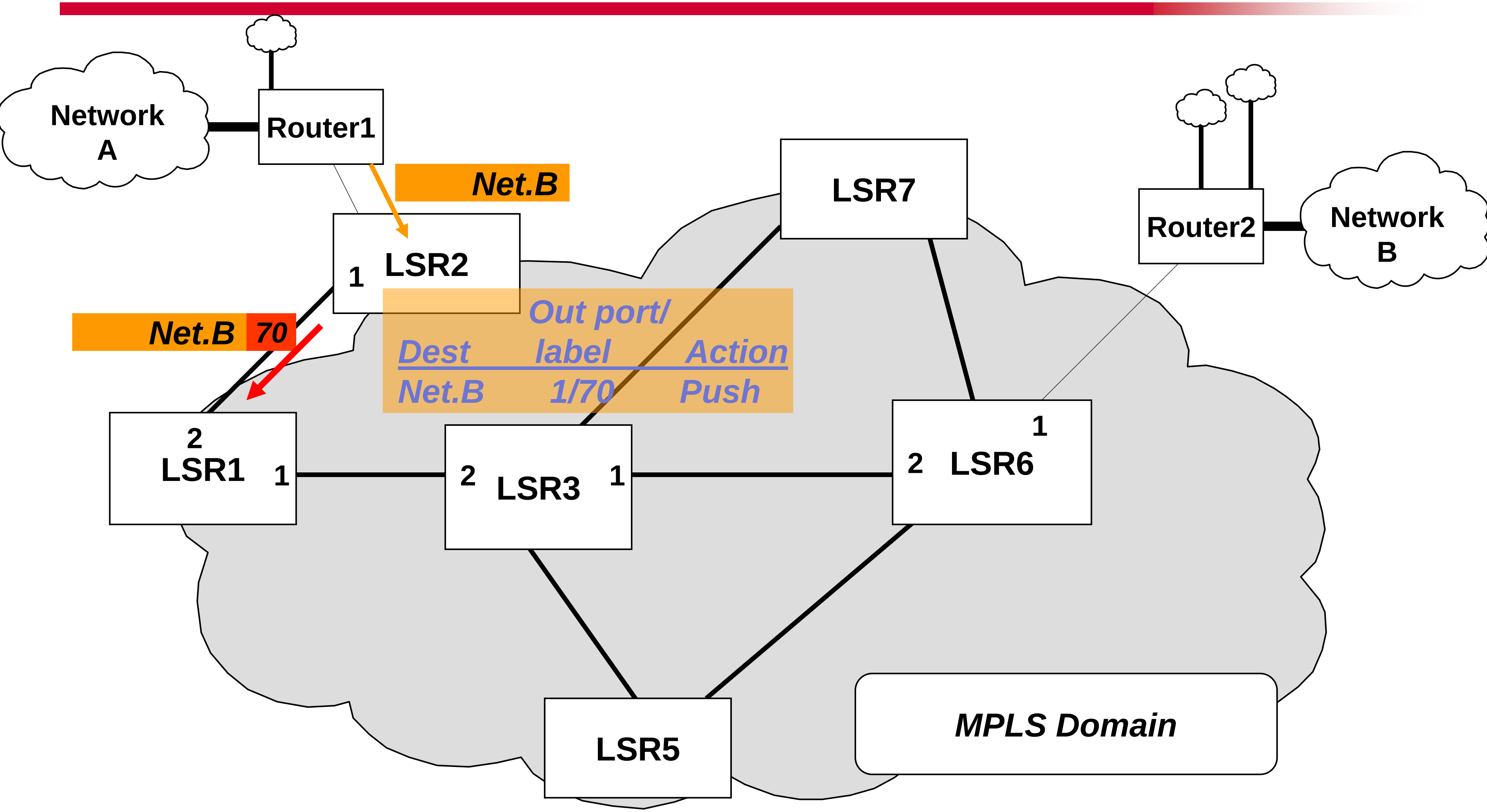
Label Switched Path – *Created & Forwarding Table Updated*



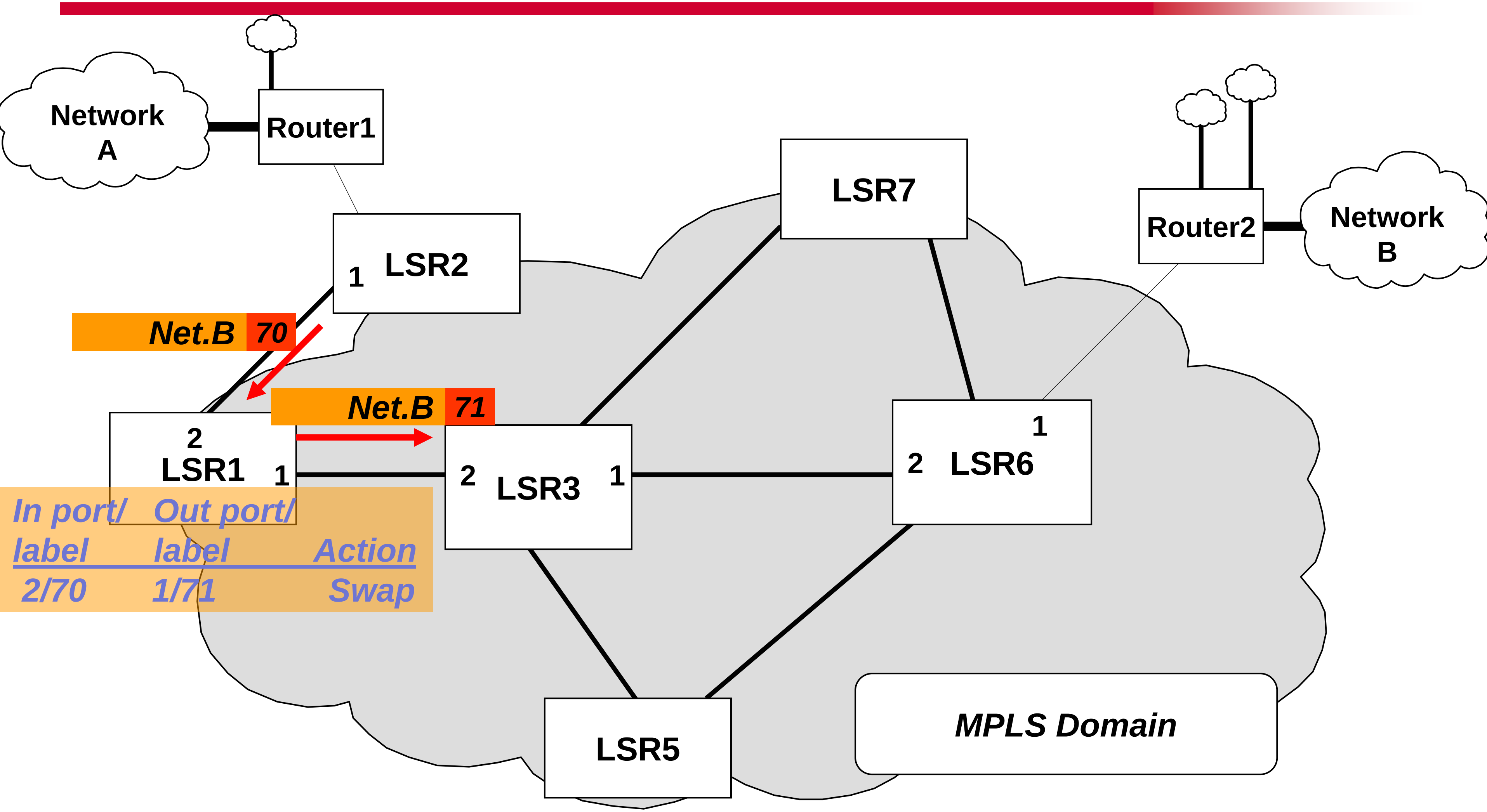
Steps in the process

- Topology determination
- Best path determination
- Data forwarding

Data Forwarding – *Unlabelled packet*

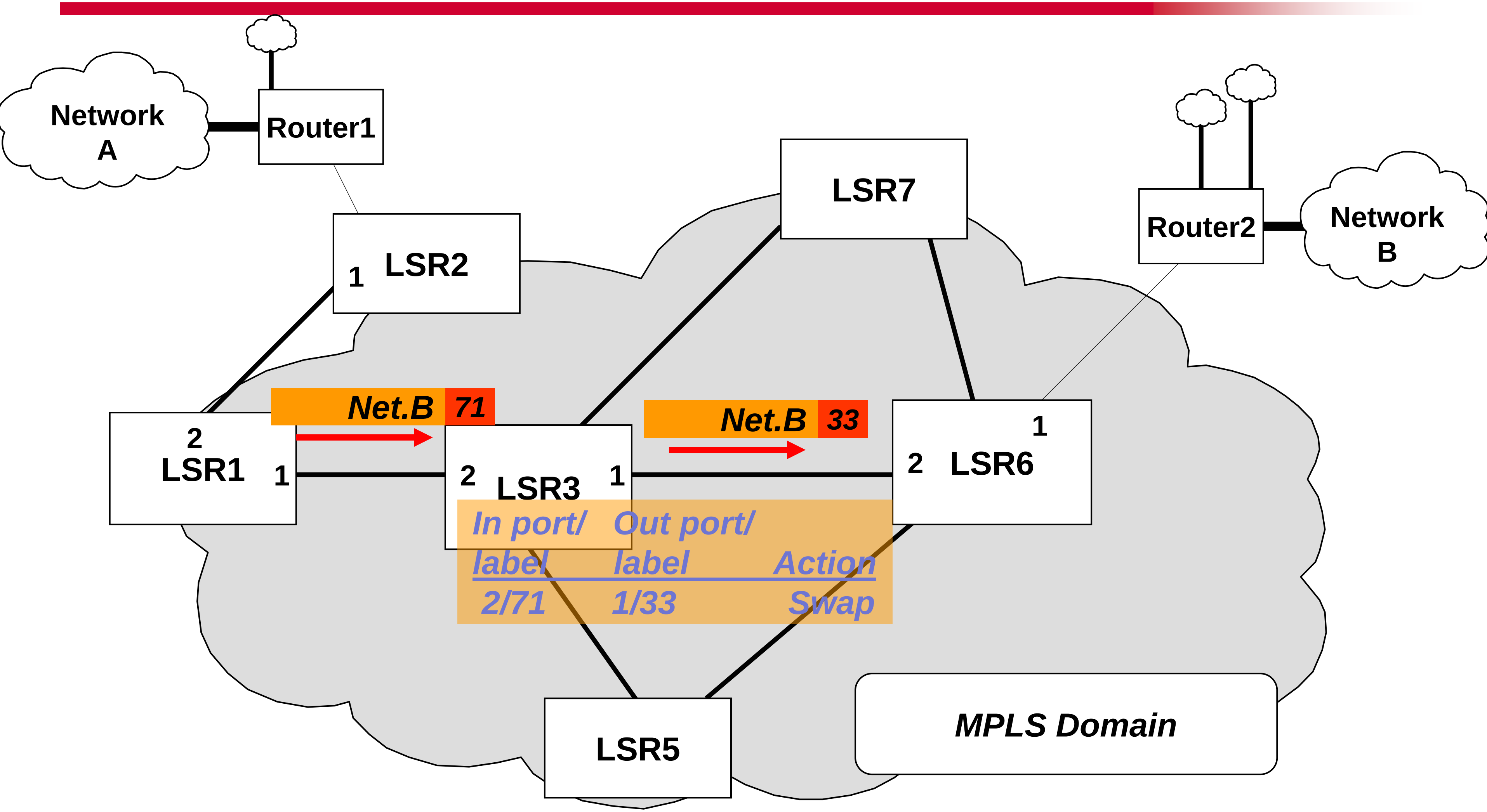


Data Forwarding – *LSR1 – LSR3*

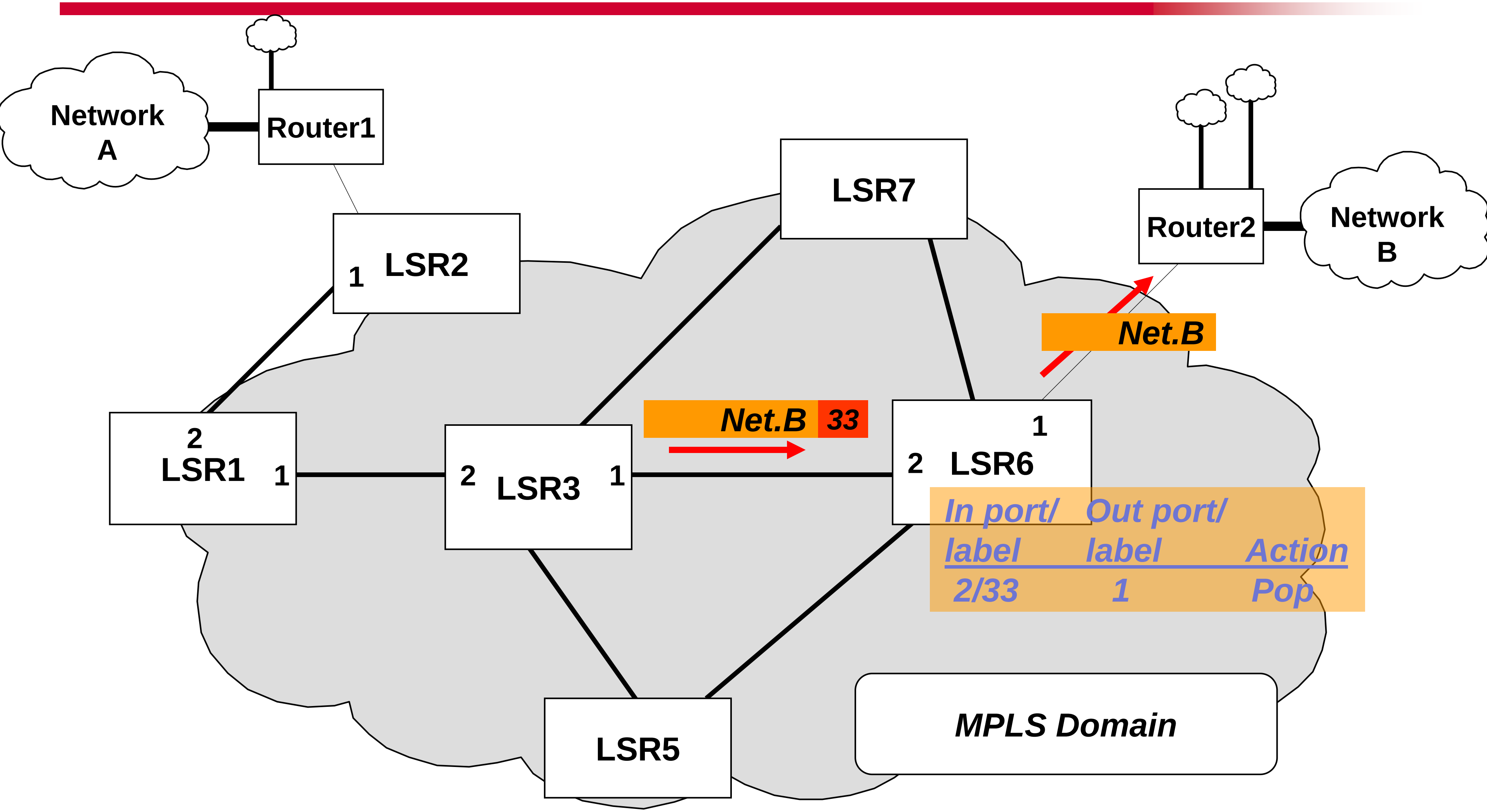


In port/ label	Out port/ label	Action
2/70	1/71	Swap

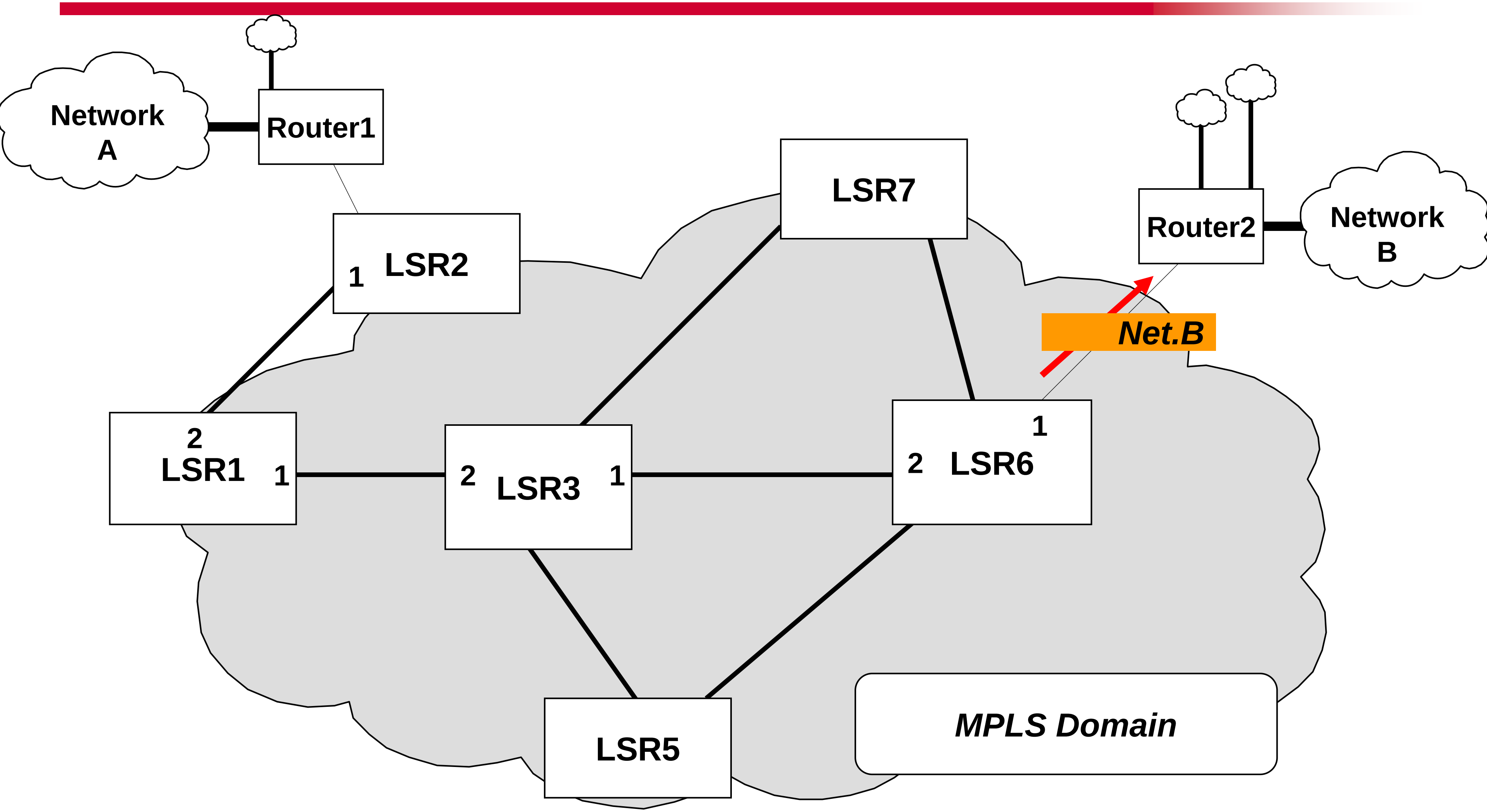
Data Forwarding – *LSR3 – LSR6*



Data Forwarding – *LSR6 – Router*



Data Forwarding – *Unlabelled packet delivered*



Label Switching: *Three Important Questions*

- Q: What field on the *labelled* packet do we use to make the forwarding decision?
  - ✓
- Q: When we use this field as an index into the LIB…what do we look up?
  - ✓
- Q: What other vital piece of information does the LIB contain?
  - ✓

Label Switching:
Three Important Questions

- Q: What field on the *labelled* packet do we use to make the forwarding decision?
  - ✓ A: The *outermost* label
- Q: When we use this field as an index into the LIB…what do we look up?
  - ✓
- Q: What other vital piece of information does the LIB contain?
  - ✓

Label Switching:
Three Important Questions

- Q: What field on the <u>*labelled*</u> packet do we use to make the forwarding decision?
  - ✓ A: The <u>*outermost*</u> label
- Q: When we use this field as an index into the LIB…what do we look up?
  - ✓ A: The output I/F (or queue) reference
- Q: What other vital piece of information does the LIB contain?
  - ✓

Label Switching: *Three Important Questions*

- Q: What field on the ***labelled*** packet do we use to make the forwarding decision?
  - ✓ A: The ***outermost*** label
- Q: When we use this field as an index into the LIB…what do we look up?
  - ✓ A: The output I/F (or queue) reference
- Q: What other vital piece of information does the LIB contain?
  - ✓ A: The outbound label value

Section 2: Agenda

- How Does Traditional Routing Work?
  - ✓ Brief overview
  - ✓ The hyperaggregation problem
  - ✓ Data Plane and Control Plane
- MPLS Architecture
  - ✓ MPLS Terminology
  - ✓ How Does It Work?
  - ✓ Label Distribution Protocol (LDP)
  - ✓ Penultimate Hop Popping, Aggregation, TTL
  - ✓ ATM Issues

We need a protocol to create these LSPs

- Requires a signaling protocol to:
  - ✓ Coordinate label distribution
  - ✓ Loop prevention
  - ✓ *Explicitly route the LSP*
  - ✓ *Bandwidth reservation*
  - ✓ *Class of Service (DiffServ style)*
  - ✓ *Pre-emption of existing LSPs*

MPLS signaling protocols

- Label Distribution Protocol - LDP
- RSVP-TE
- CR-LDP
- BGP4-TE

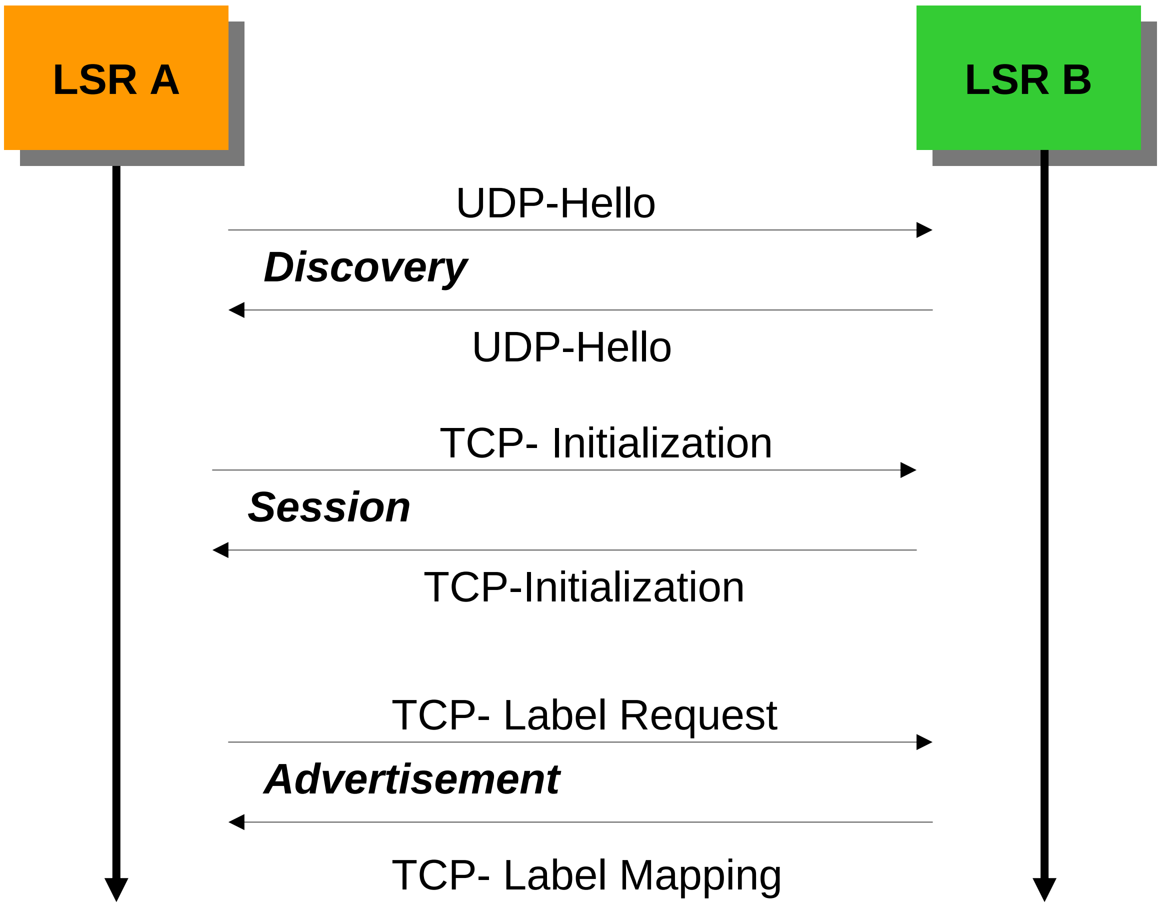
MPLS signaling protocol - LDP

- Label Distribution Protocol - LDP
- RSVP-TE
- CR-LDP
- BGP4-TE

LDP Message Types

- Discovery
  - ✓ Announce and maintain the presence of LSR
  - ✓ Send a Hello message periodically via UDP
- Session
  - ✓ Establish, maintain, and delete sessions over TCP
- Advertisement
  - ✓ Create, change, and delete label mappings for FECs over TCP
- Notification
  - ✓ Provide status, diagnostic, and error information over TCP

LDP Message Exchange



LDP - *Discovery*

- Basic Discovery
  - ✓ LDP *Hellos* sent periodically on interface
  - ✓ LDP *Hellos* sent as UDP packets addressed to well-known LDP *discovery port*
  - ✓ Hello includes a Hello *hold time*
  - ✓ Hello carries the LSR Identifier and Label Space LSR intends to use on the interface
  - ✓ Receipt of Hello establishes *Hello Adjacency*

LDP – *Session Initialization*

- Active LSP (higher LSR Identifier) sends Session Initialization (TCP) message to passive LSR
- Session Initialization contains:
  - ✓ LDP Protocol version
  - ✓ Label distribution and control method
  - ✓ Timer values
  - ✓ VPI/VCI or DLCI label ranges
- If passive LSR accepts parameters it responds with a *KeepAlive;* else a Notification *reject*

LDP - *Advertisement*

- Label Binding –
  - LSR distributes a label mapping for a FEC to an LDP peer
    - LDP defines two FEC Types:
      - IP Address Prefix
      - Host Address
- Label Withdrawal
  - LSR informs an LDP peer that it may not continue to use specific bindings the LSR had previously advertised
- Label Release
  - LSR informs a peer that it no longer needs a binding previously received

LDP – Advertisement *Label Request*

- LSR may transmit a Label Request when:
  - ✓ LSR recognizes a FEC via the forwarding table, and the next hop is an LDP peer, and there is no mapping for the FEC
  - ✓ LSR recognizes that the next hop to a FEC changes, and LSR doesn’t already have a mapping from that next hop for the FEC
  - ✓ LSR receives a Label Request for a FEC from an upstream LDP peer, the next hop is an LDP peer, and the LSR doesn’t already have a mapping from that next hop.

LDP – Advertisement
Label Request Message

0	Label Request (0x401)	Message Length
Message ID		
FEC TLV		
Optional Parameters		

Optional Parameters (used for loop detection):

Hop Count TLV – Specifies the running total of LSRs along the LSP being setup

Path Vector TLV – Specifies the LSRs along the LSR being setup

LDP – Advertisement
Label Mapping Message

0	Label Mapping (0x400)	Message Length
Message ID		
FEC TLV		
Label TLV		
Optional Parameters		

Optional Parameters:
- Hop Count TLV
- Path Vector TLV

LDP – Advertisement
Label Mapping - Independent Control

- If LSR configured for independent control a label mapping message is sent when:
  - ✓ LSR recognizes a new FEC via forwarding table and LSR configured for Downstream Unsolicited mode
  - ✓ LSR receives a label request from an upstream peer for a FEC in the LSR forwarding table
  - ✓ The next hop for a FEC changes to another LDP peer
  - ✓ Attributes of a mapping change
  - ✓ Receipt of a mapping from downstream next hop AND
    - ➢ no upstream mapping has been created, OR
    - ➢ attribute of mapping changed

LDP – Advertisement
Label Mapping - Ordered Control

- If LSR configured for ordered control a label mapping message is sent when:
  - ✓ LSR recognizes a new FEC via forwarding table and LSR is the egress for the FEC
  - ✓ LSR receives a label request from an upstream peer for a FEC in the LSR forwarding table, and the LSR is the egress for the FEC or has a downstream mapping for the FEC
  - ✓ The next hop for a FEC changes to another LDP peer
  - ✓ Attributes of a mapping change
  - ✓ Receipt of a mapping from downstream next hop AND
    - ➢ no upstream mapping has been created, OR
    - ➢ attribute of mapping changed

LDP – Advertisement
Label Withdraw Message

- Revokes a previously assigned binding
- Reason may include:
  - ✓ FEC removed from in the forwarding table due to routing change
  - ✓ LSR configuration change

0	Address Withdrawal (0x301)	Message Length
Message ID		
Address List TLV		

LDP – Advertisement *Label Release Message*

- LSR sends a Label Release to inform LSP peer that FEC-label mapping is no longer needed
- Reason may include:
  - ✓ LSR that sent mapping is no longer next hop for FEC
  - ✓ LSR receives a label mapping from an LSR which is not the next hop and the LSR in conservative mode
  - ✓ LSR receives a Label Withdraw message

LDP - *Notification*

- Sent by LSR to inform a LDP peer of a significant event:
  - Received LDP with unsupported protocol version
  - LDP message type unsupported
  - TLV not supported
  - KeepAlive timer expired
  - Session Initialization failure
    - parameters unacceptable

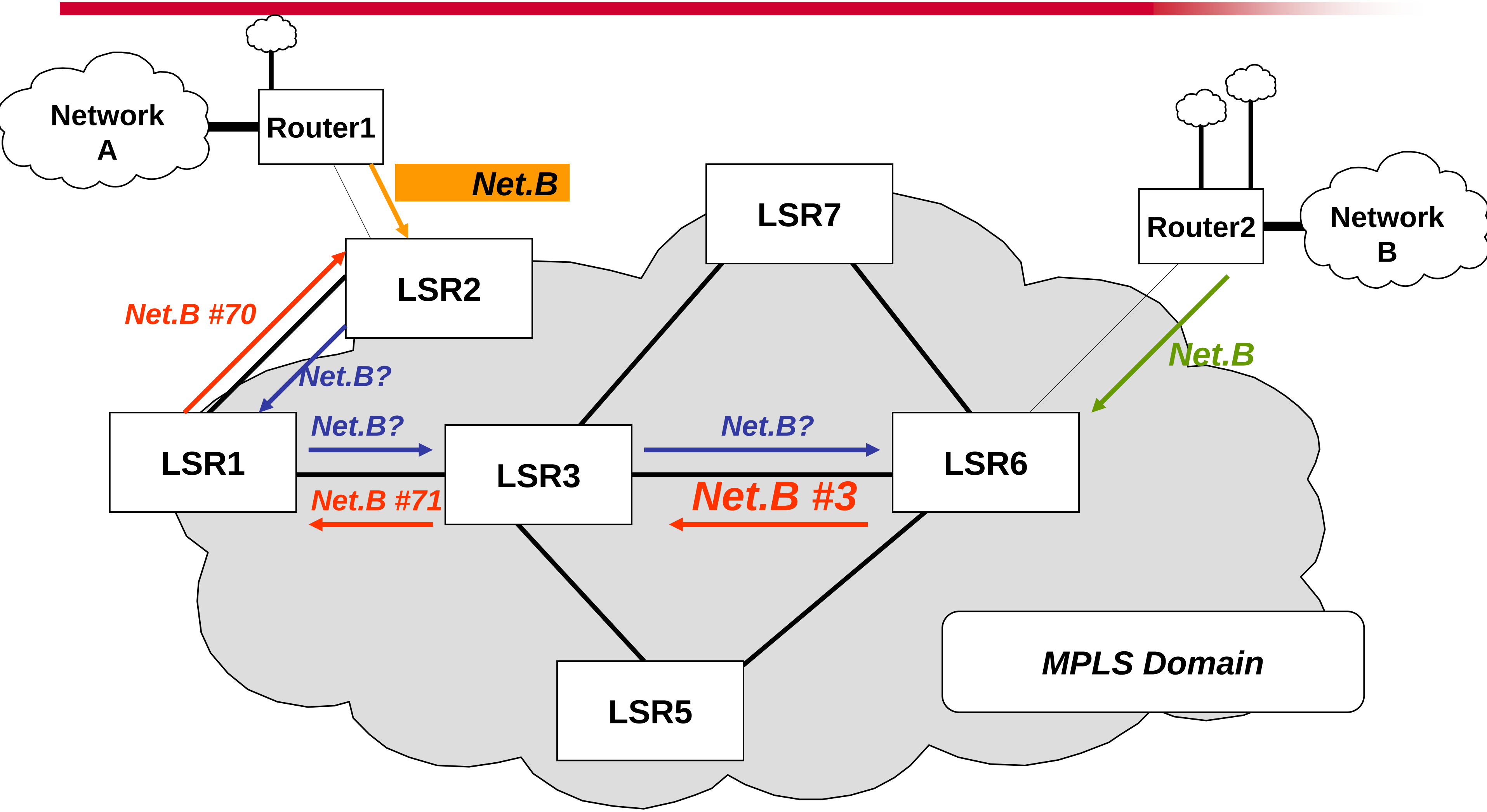
Section 2: Agenda

- How Does Traditional Routing Work?
  - ✓ Brief overview
  - ✓ The hyperaggregation problem
  - ✓ Data Plane and Control Plane
- MPLS Architecture
  - ✓ MPLS Terminology
  - ✓ How Does It Work?
  - ✓ Label Distribution Protocol (LDP)
  - ✓ Penultimate Hop Popping, Aggregation, TTL
  - ✓ ATM Issues

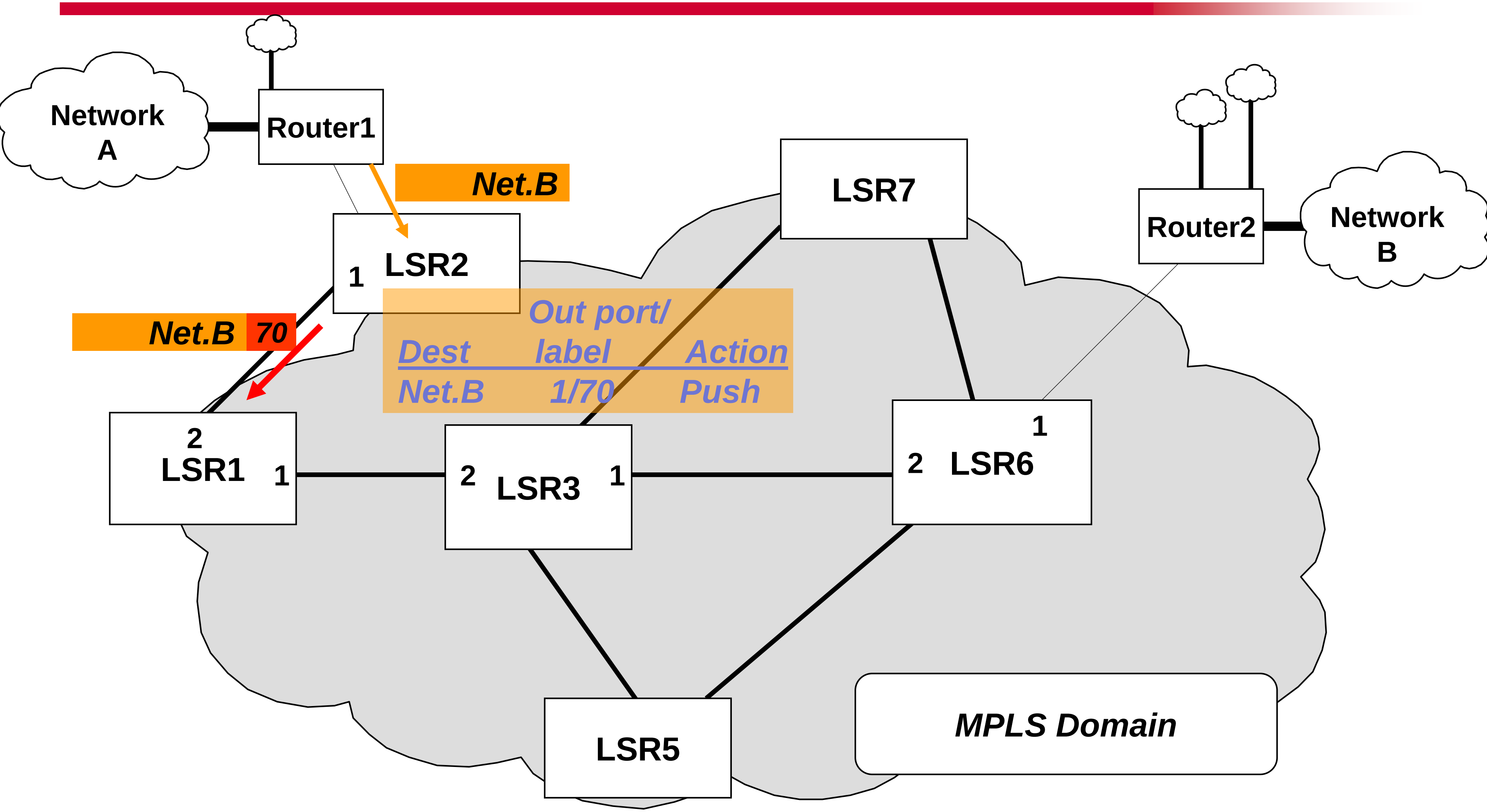
Penultimate Hop Popping

- Requested by egress LSR to neighbor LSR by using a special label value
- Penultimate LSR forwards packet without an MPLS label
- Used to prevent an egress LSR from having to perform two lookups on a packet
  - ✓ The packet has no Label so only the Routing FIB needs to be looked at to determine where to forward the packet

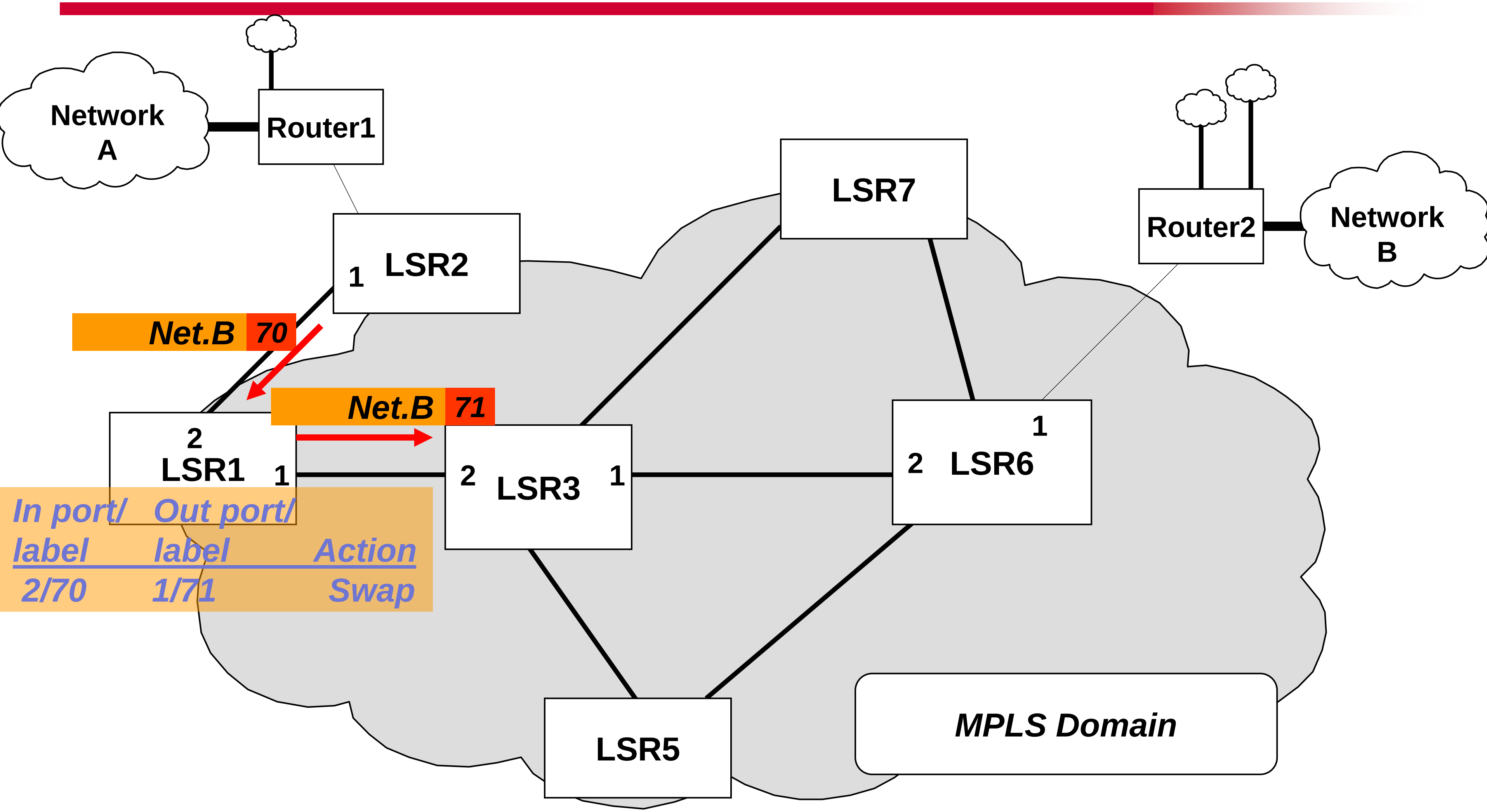
PHP signaled by Egress LSR



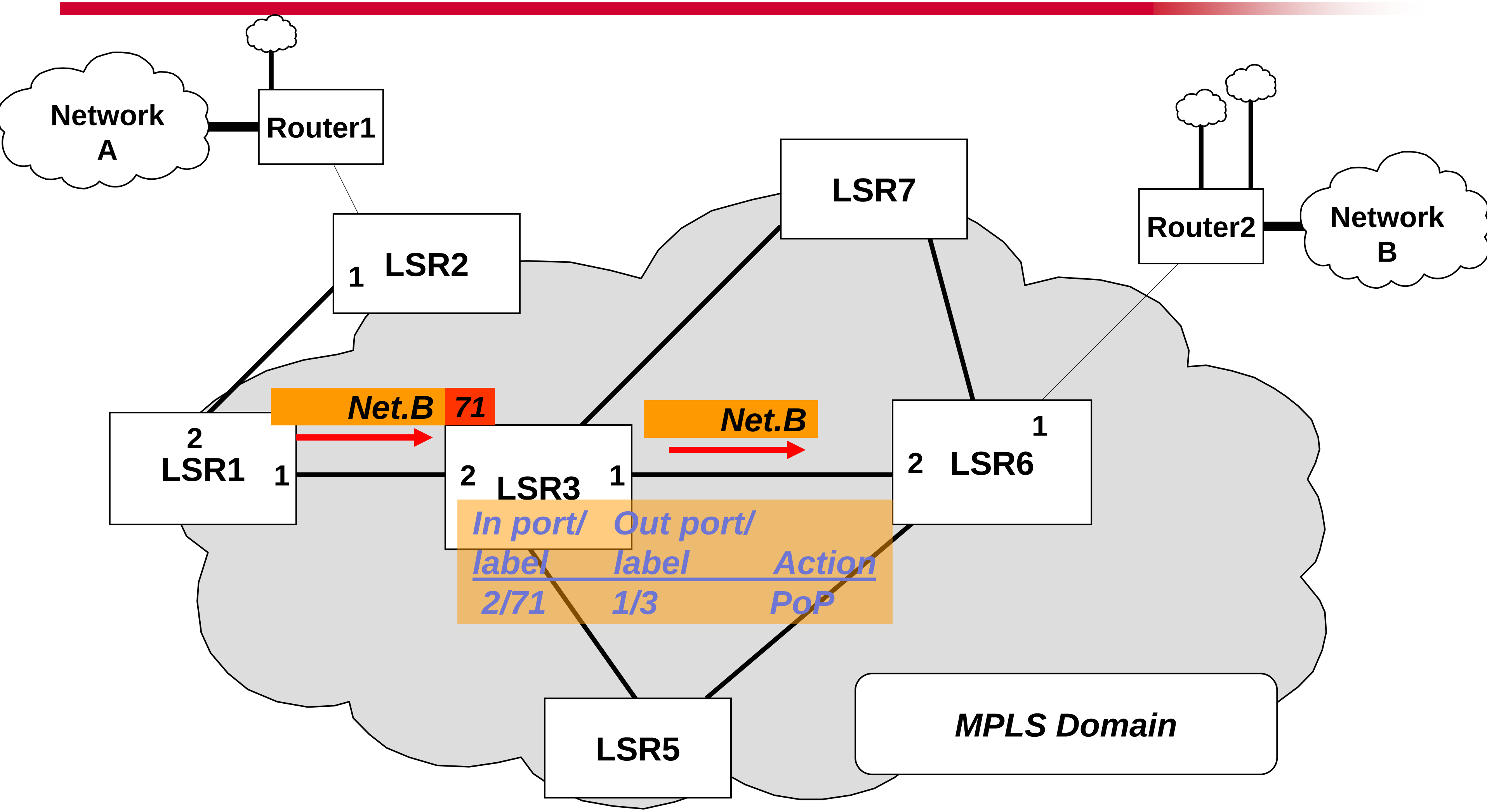
Data Forwarding with PHP



Data Forwarding with PHP



Data Forwarding with PHP

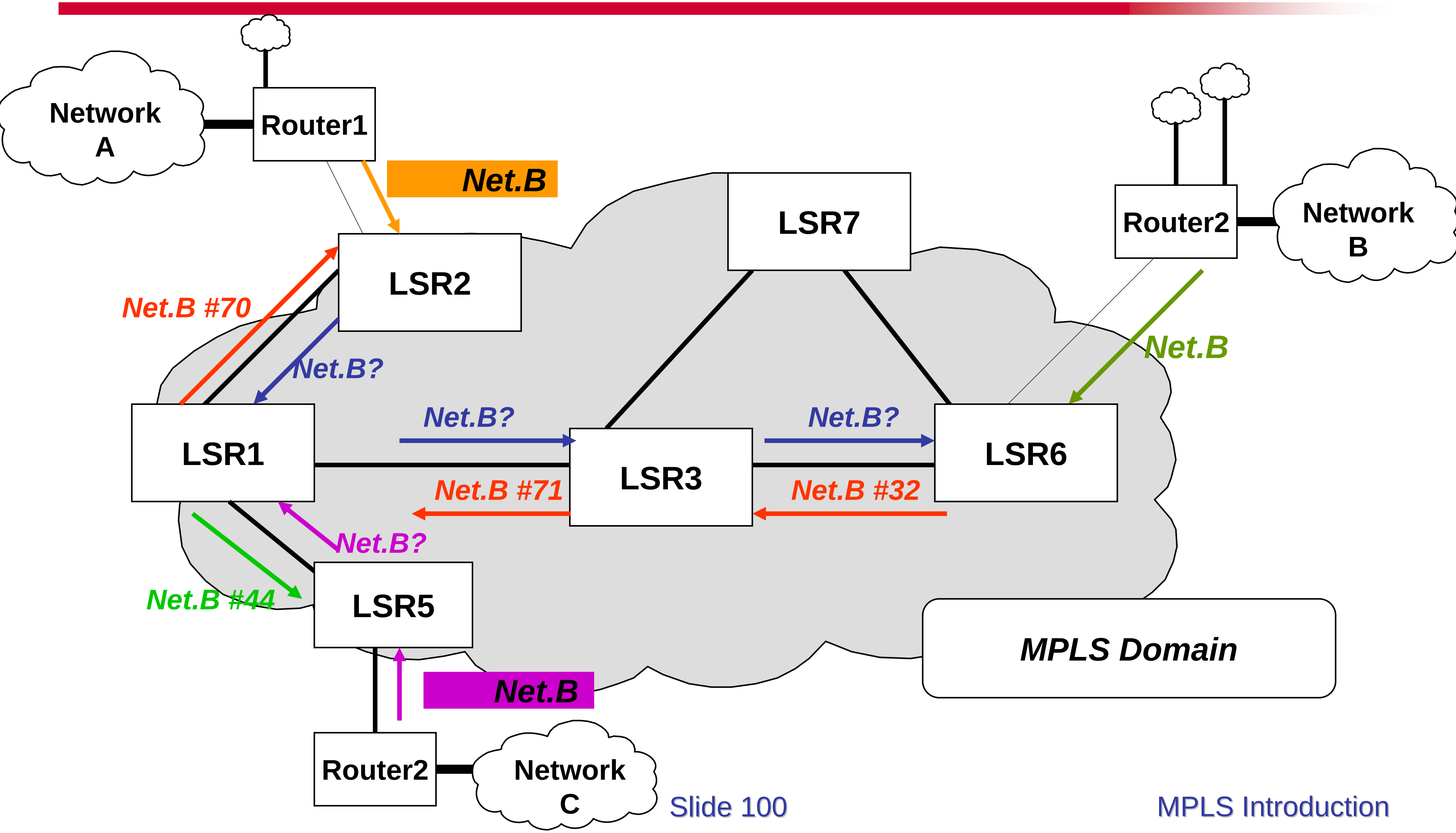


In port/ label	Out port/ label	Action
2/71	1/3	PoP

Route Aggregation

- Produces a single label to a union of all traffic with a common FEC
- Reduces the number of labels needed to handle a particular set of packets
- Reduces the amount of label distribution control traffic

Route Aggregation



Time To Live

- There are two approaches:
  - At ingress IP TTL is copied into the MPLS Header
  - MPLS TTL is decremented at each LSR
  - At egress MPLS TTL is copied into IP TTL
    - Prevents loops
- Or
  - At ingress IP TTL copied into MPLS Header
  - At egress MPLS TTL is decremented and copied into IP TTL
    - Hides structure of MPLS domain

Section 2: Agenda

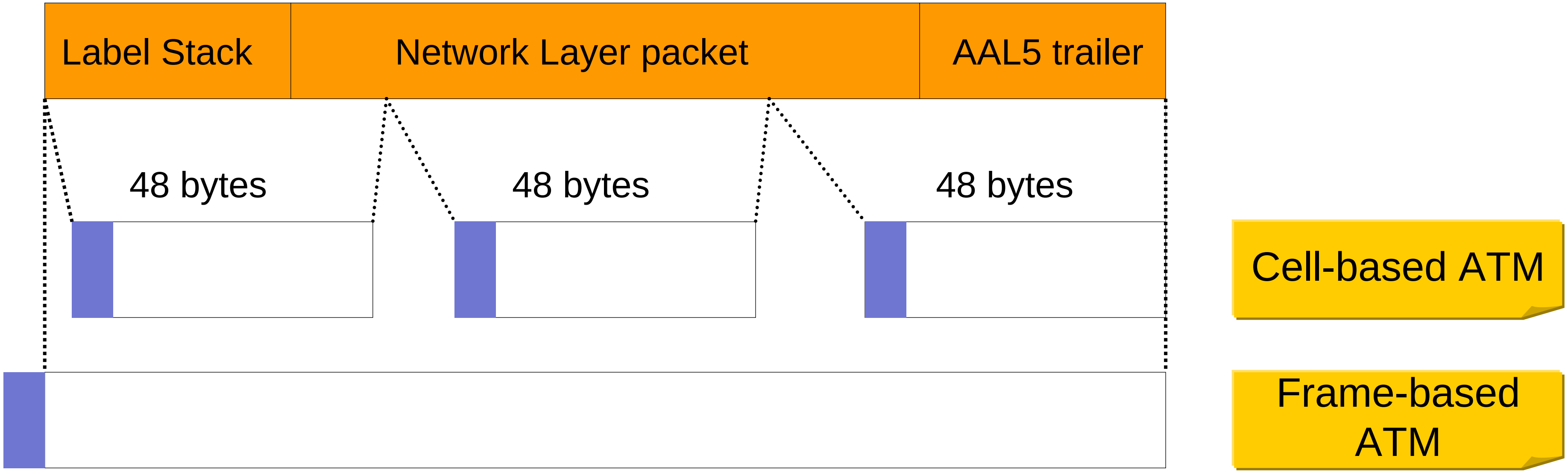
- How Does Traditional Routing Work?
  - ✓ Brief overview
  - ✓ The hyperaggregation problem
  - ✓ Data Plane and Control Plane
- MPLS Architecture
  - ✓ MPLS Terminology
  - ✓ How Does It Work?
  - ✓ Label Distribution Protocol (LDP)
  - ✓ Penultimate Hop Popping, Aggregation, TTL
  - ➡ ATM Issues

ATM Issues

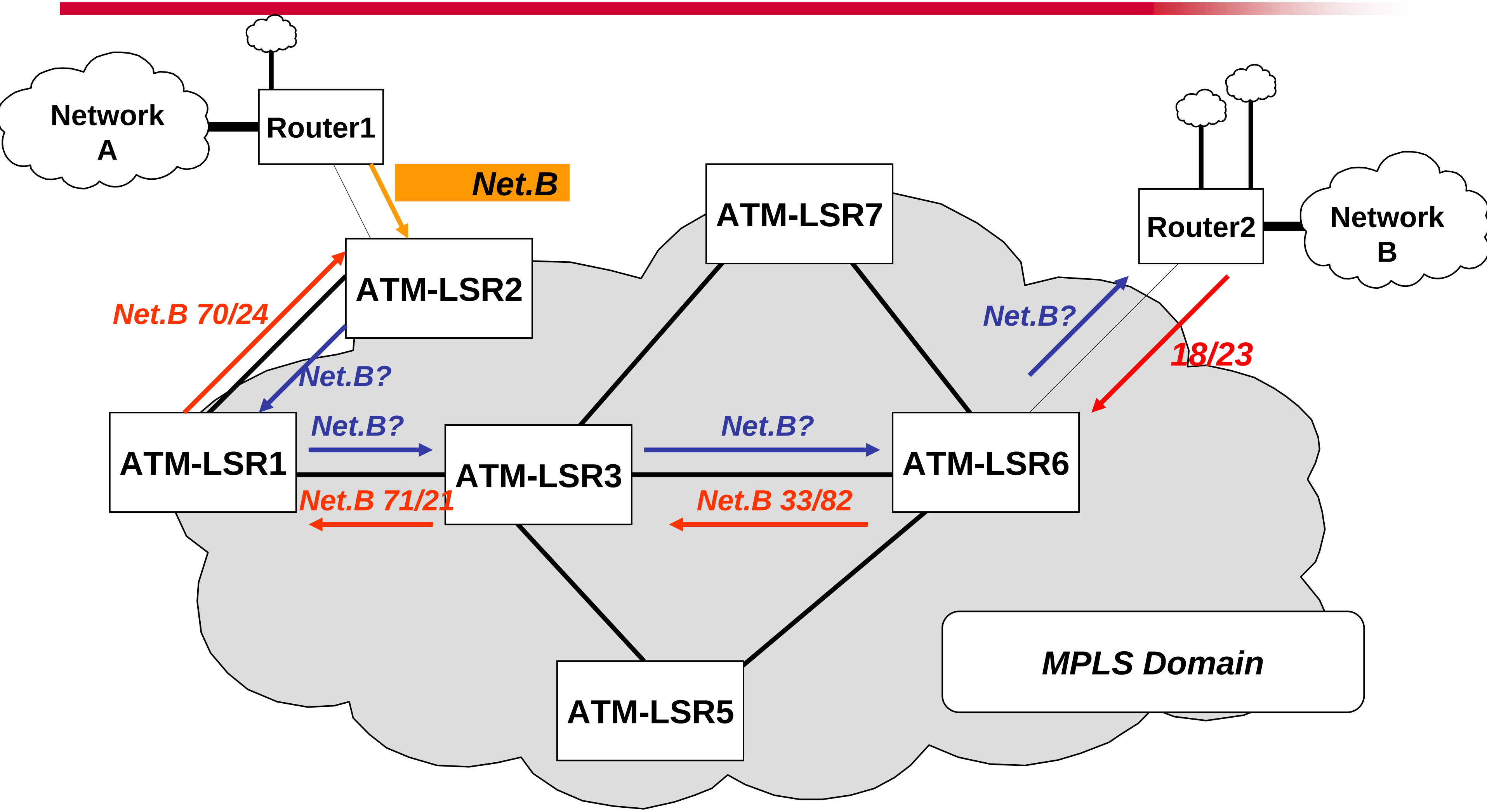
- Encapsulation of labeled packets on ATM links
- Looping and TTL adjustments
- Cell interleave and VC-merge

ATM Encapsulation

- MPLS label stack is inserted into a packet before segmenting the packet into cells
- The VPI/VCI is associated with the outmost label



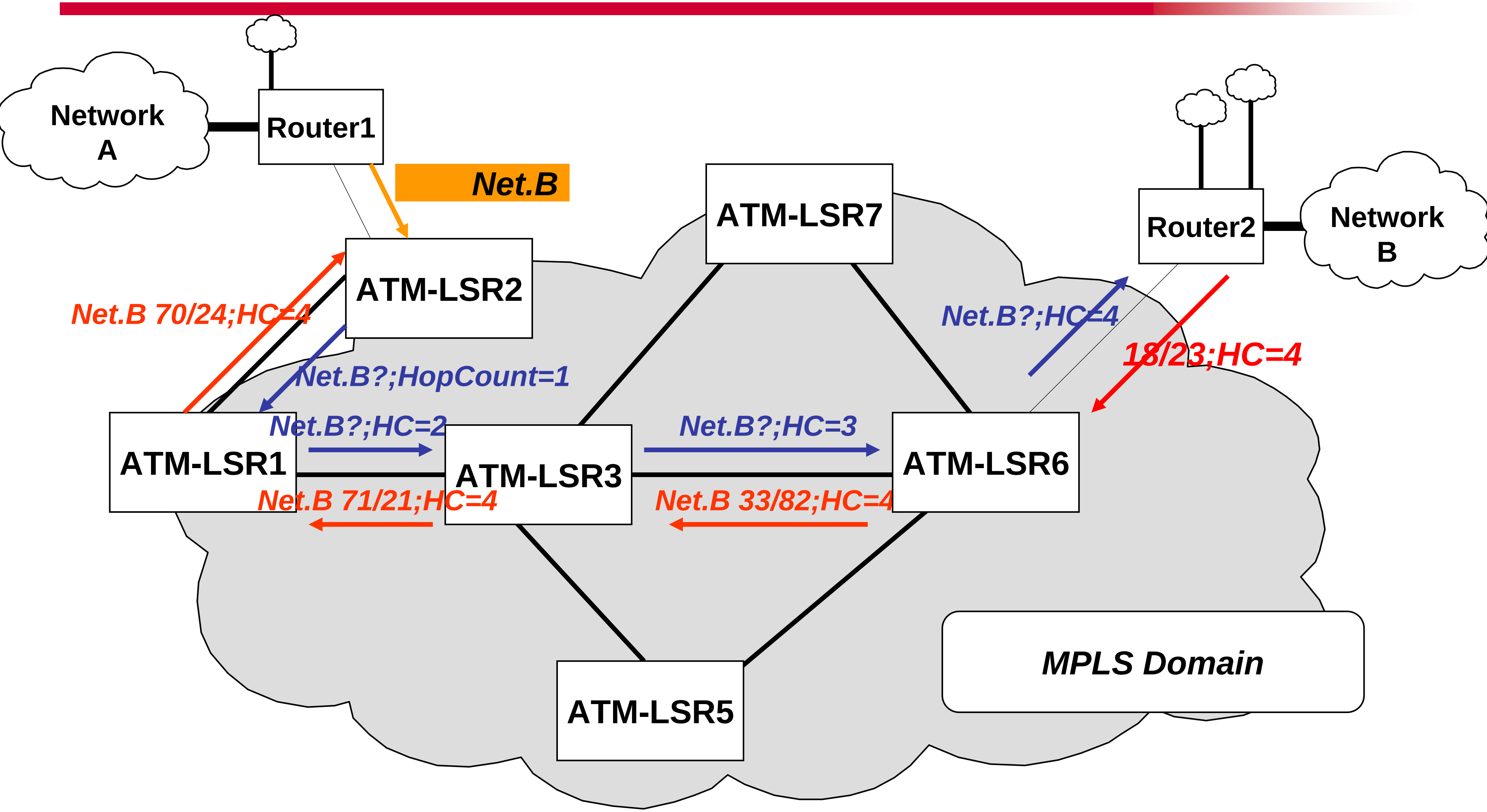
ATM Label Distribution Downstream-On-Demand



ATM TTL Adjustments

- ATM-LSR have no capability to decrement TTL
- Edge LSR relies on Hop Count of downstream-on demand with ordered control LDP to determine number of LSR hops across MPLS domain
- The edge LSR adjusts the IP TTL according to Hop Count before transmitting packet
- Loop avoidance is via Path Vector

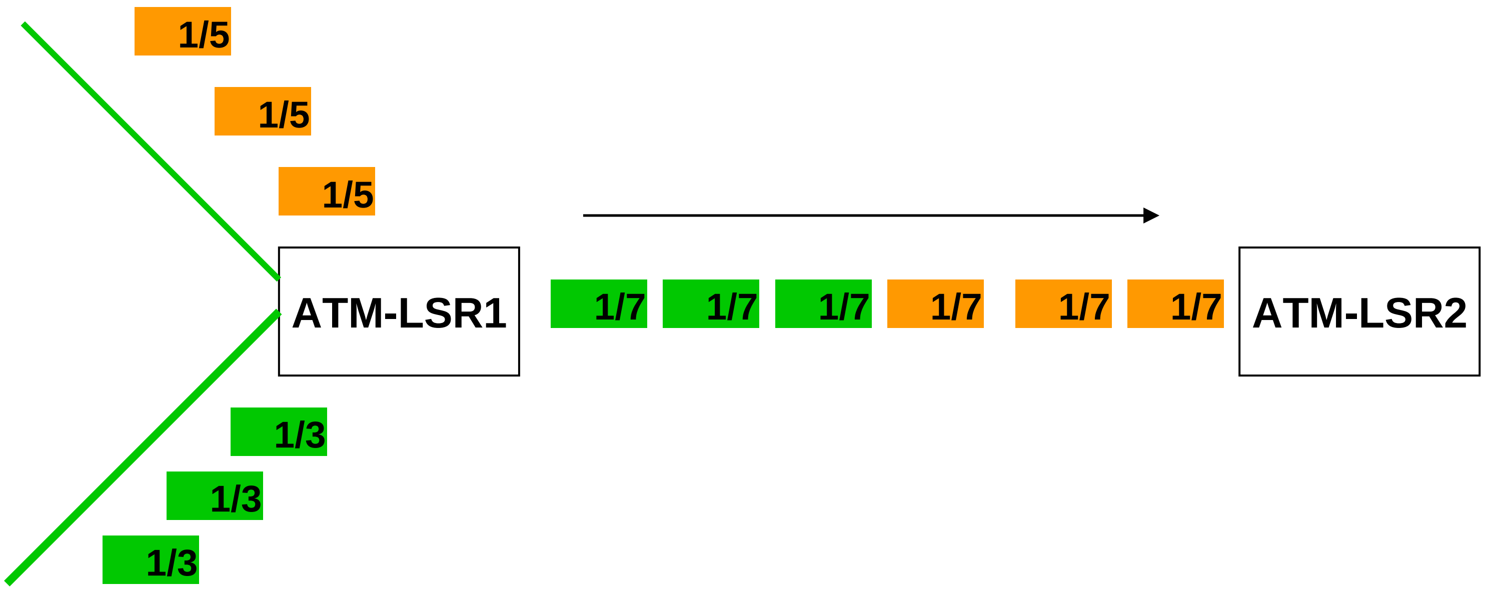
ATM Label Distribution
TTL Determination - Hop Count



Cell Interleaving

- VC-Merge – Label is VPI/VCI and ATM-LSRs are required to buffer cells from one packet until the entire packet is received
  - ✓ Requires multiple labels per FEC
  - ✓ Introduces delay
- VP-Merge – Label is VPI and the VCI is used as an identifier to distinguish different frames sent with the same VPI
  - ✓ Limits number of labels to the VPI space

VC-Merge



VP-Merge

More on TE - TE before MPLS

- IP TE:
  - ✓ Popular, used successfully by ISPs
  - ✓ Major control mechanism: change the cost of a particular link or links
  - ✓ Coarse, manual, limited
- ATM TE:
  - ✓ Place PVCs across the network from src to dest
  - ✓ Fine-grained control over the traffic flow
  - ✓ Full mesh PVCs between a set of routers
  - ✓ Periodically resize and reposition those PVCs based on observed traffic from routers
  - ✓ Scalability issues: $O(n^2)$ or $O(n^3)$
    - ➢ Link failure
    - ➢ Node failure
  - ✓ Cell tax

Example: the Fish Problem

- IP shortest path – all IP traffic routed this way. Congestion may happen, while other links may be under-utilized.
- ATM – build two PVCs and set their costs to the same for load sharing
- ATM TE is more flexible than traditional IP TE
  - ✓ No other devices connected to the networks are affected by any metric change.
  - ✓ ATM TE is more powerful, but has scalability and cost problem.
- MPLS TE: ATM TE capabilities + IP TE simplicity – ATM TE scalability

Benefits of MPLS

- Decoupling routing and forwarding
  - ✓ Enables applications like TE
- Better integration of IP and ATM
  - ✓ IP over ATM (carrying IP over ATM VCs), overly model:
    - ➢ scalability limits
    - ➢ Mapping issue for QoS
  - ✓ MPLS bridges the gap between IP and ATM: VPI/VCI values are used as labels for cells, called Label-Controlled ATM (LC-ATM) or IP+ATM
- Basis of building new applications/services, e.g., TE, VPNs

MPLS TE

- Like ATM VCs, MPLS TE LSPs (or TE tunnels) let the headend of a TE tunnel control the path its traffic takes to a dest.
- Unlike ATM VCs, no need full mesh of routing neighbors
- Like ATM, MPLS TE can reserve bandwidth when it builds LSPs.
- Unlike ATM, there is no forwarding-plane enforcement of a reservation. A reservation is made in the control plane only.
- The Fish problem with MPLS TE

Using MPLS TE in Real Life

- Three typical real-life applications:
  - ✓ Optimizing network utilization
    - ➢ Strategic or full-mesh or partial-mesh approach: build LSPs that meet bandwidth demands
    - ➢ Get as much as you can from the infrastructure and delay upgrading, which translates directly into cost by not having to buy bandwidth
  - ✓ Handle unexpected congestion as needed
    - ➢ Tactical approach
  - ✓ Quick recovery from link and node failures
    - ➢ FRR allows you to drastically reduce packet loss