NORTEL NETWORKS™

How the world shares ideas.

NANOG

MPLS Tutorial and Operational Experiences

Peter Ashwood-Smith,
Bilel Jamoussi,
October, 1999

Tutorial Outline

- **Overview**
- Label Encapsulations
- Label Distribution Protocols
- MPLS & ATM
- Constraint Based Routing with CR-LDP
- Operational Experiences with Similar Protocols
- Summary

“Label Substitution” what is it?

One of the many ways of getting from A to B:

- ***BROADCAST***: Go everywhere, stop when you get to B, never ask for directions.
- ***HOP BY HOP ROUTING***: Continually ask who’s closer to B go there, repeat … stop when you get to B.

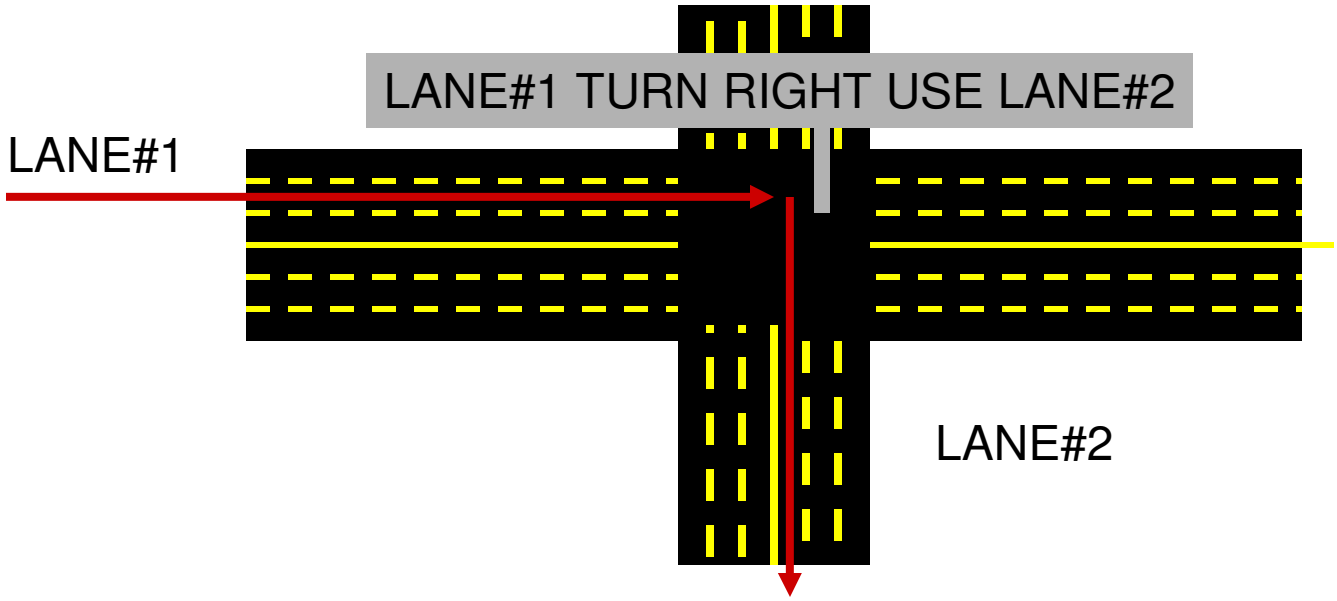
  “Going to B? You’d better go to X, its on the way”.
- ***SOURCE ROUTING***: Ask for a list (that you carry with you) of places to go that eventually lead you to B.

  “Going to B? Go straight 5 blocks, take the next left, 6 more blocks and take a right at the lights”.

Label Substitution

NANOG

Have a friend go to B ahead of you using one of the previous two techniques. At every road they reserve a lane just for you. At ever intersection they post a big sign that says for a given lane which way to turn and what new lane to take.



A label by any other name ...

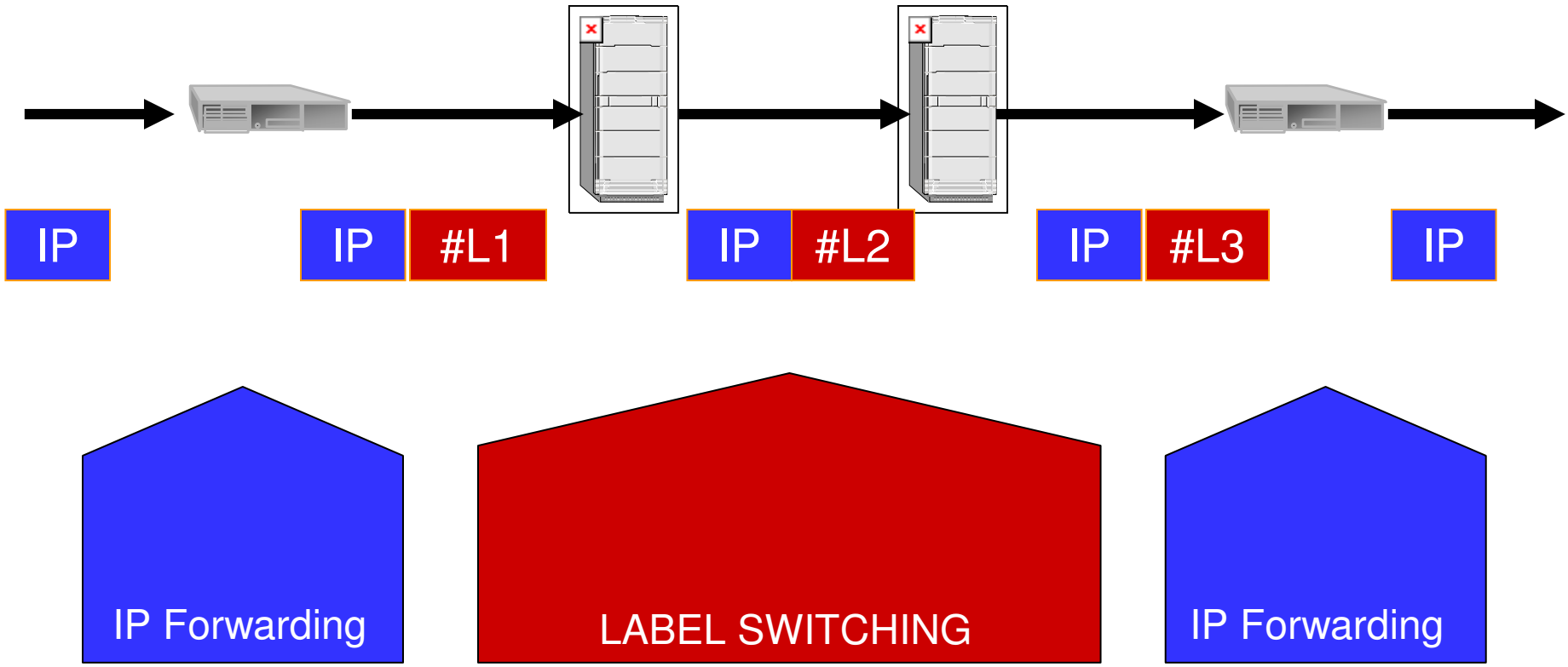
There are many examples of label substitution **protocols already in existence.**

- **ATM - label is called VPI/VCI and travels with cell.**
- **Frame Relay - label is called a DLCI and travels with frame.**
- **TDM - label is called a timeslot its implied, like a lane.**
- **X25 - a label is an LCN**
- **Proprietary PORS, TAG etc..**
- **One day perhaps Frequency substitution where label is a light frequency?**

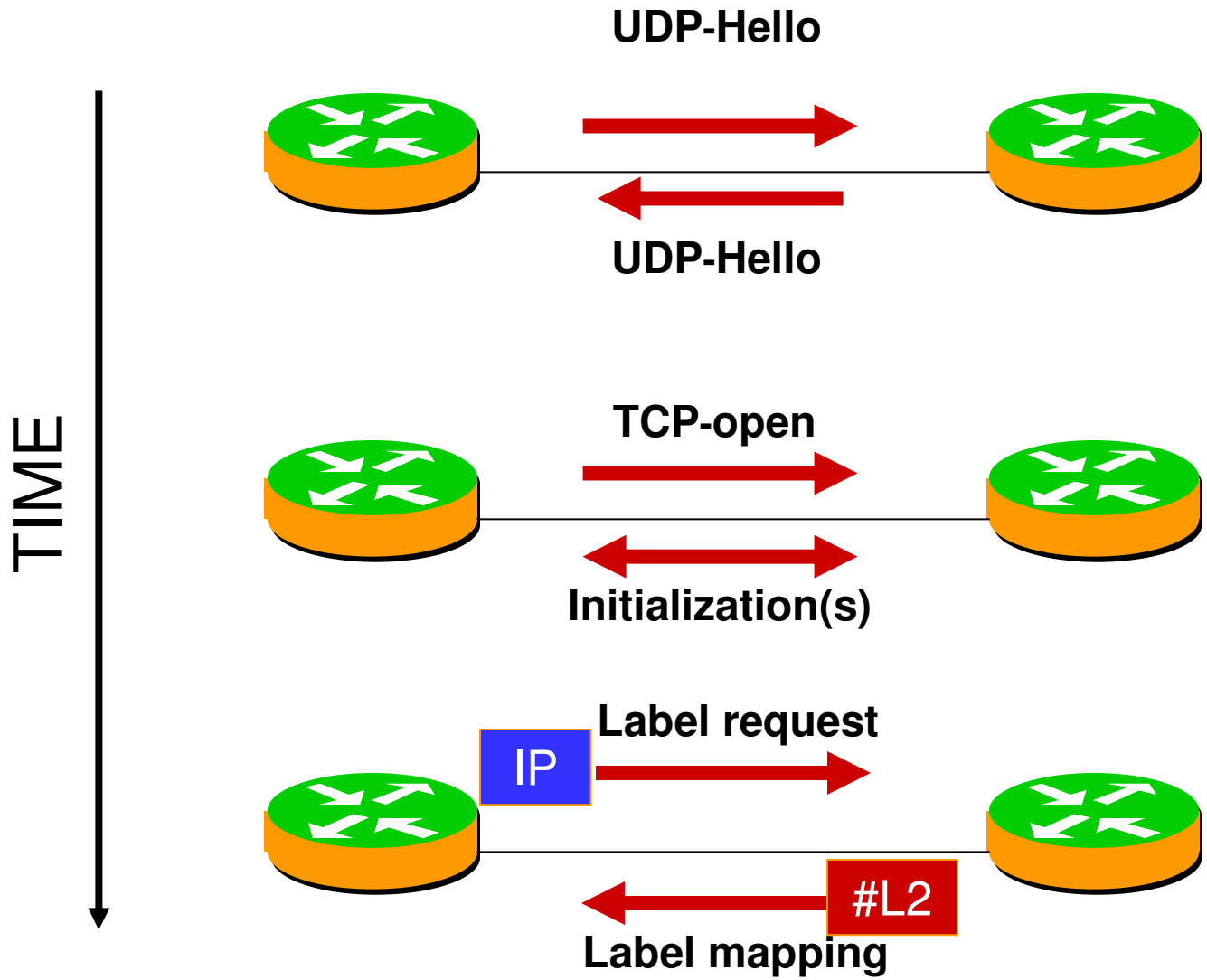
SO WHAT IS MPLS ?

- Hop-by-hop or source routing to establish labels
- Uses label native to the media
- Multi level label substitution transport

ROUTE AT EDGE, SWITCH IN CORE



MPLS: HOW DOES IT WORK ?

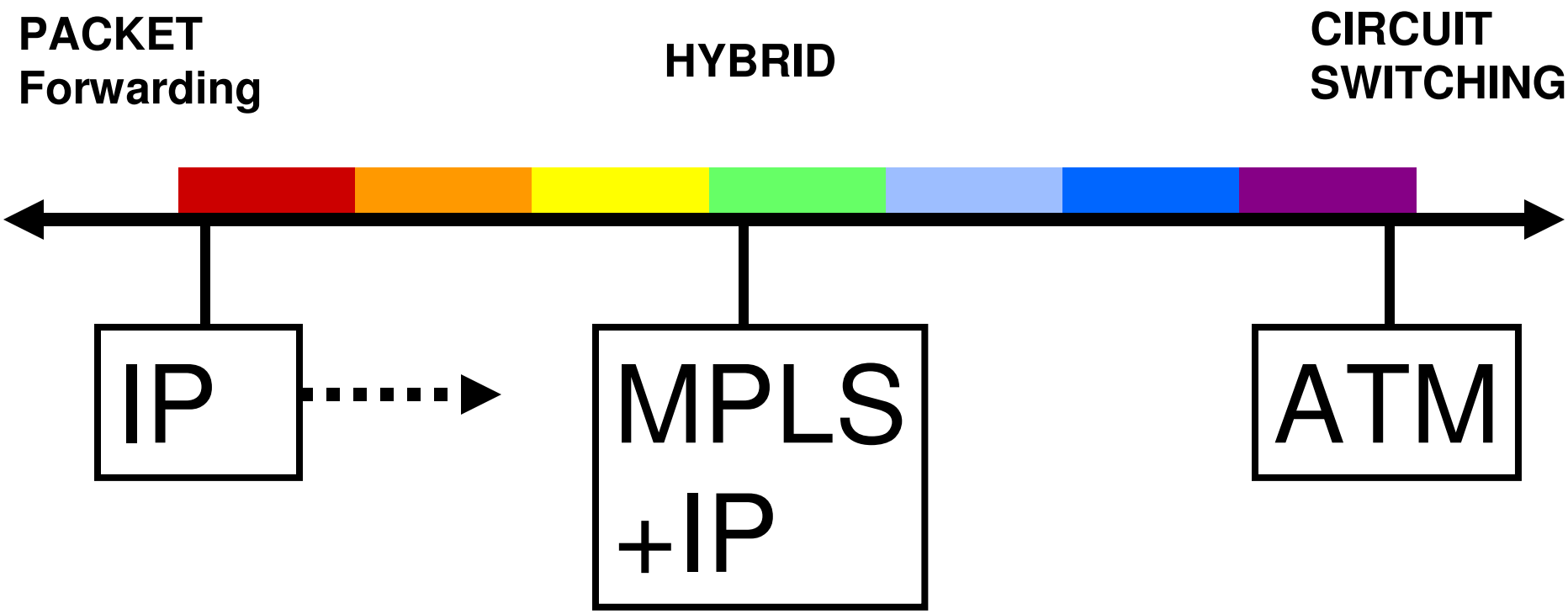


WHY MPLS ?

NANOG

- **Leverage existing ATM hardware**
- **Ultra fast forwarding**
- **IP Traffic Engineering**
  - Constraint-based Routing
- **Virtual Private Networks**
  - Controllable tunneling mechanism
- **Voice/Video on IP**
  - Delay variation + QoS constraints

BEST OF BOTH WORLDS

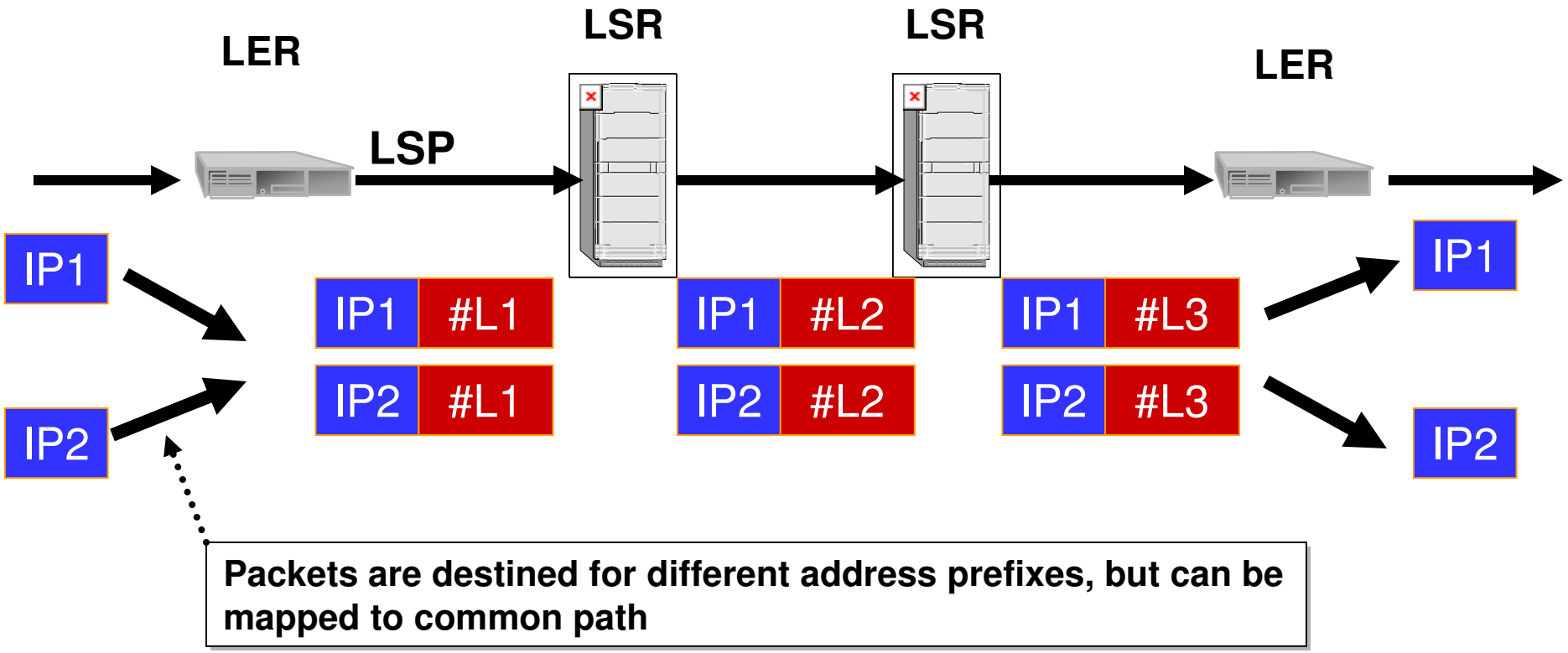


- MPLS + IP form a middle ground that combines the best of IP and the best of circuit switching technologies.
- ATM and Frame Relay cannot easily come to the middle so IP has!!

MPLS Terminology

- **LDP: Label Distribution Protocol**
- **LSP: Label Switched Path**
- **FEC: Forwarding Equivalence Class**
- **LSR: Label Switching Router**
- **LER: Label Edge Router**

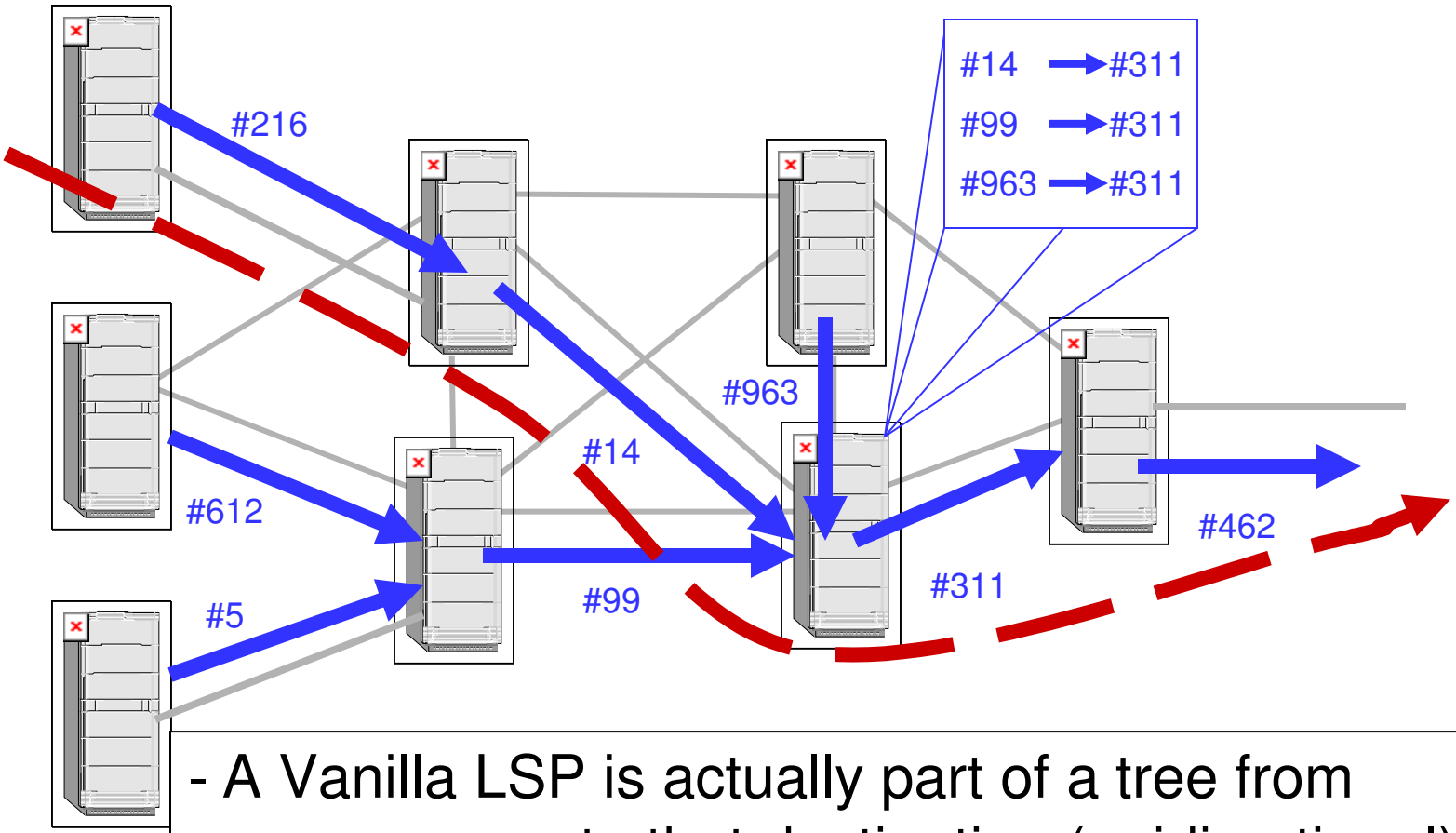
Forwarding Equivalence Classes



- FEC = “A subset of packets that are all treated the same way by a router”
- The concept of FECs provides for a great deal of flexibility and scalability
- In conventional 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.

LABEL SWITCHED PATH (vanilla)

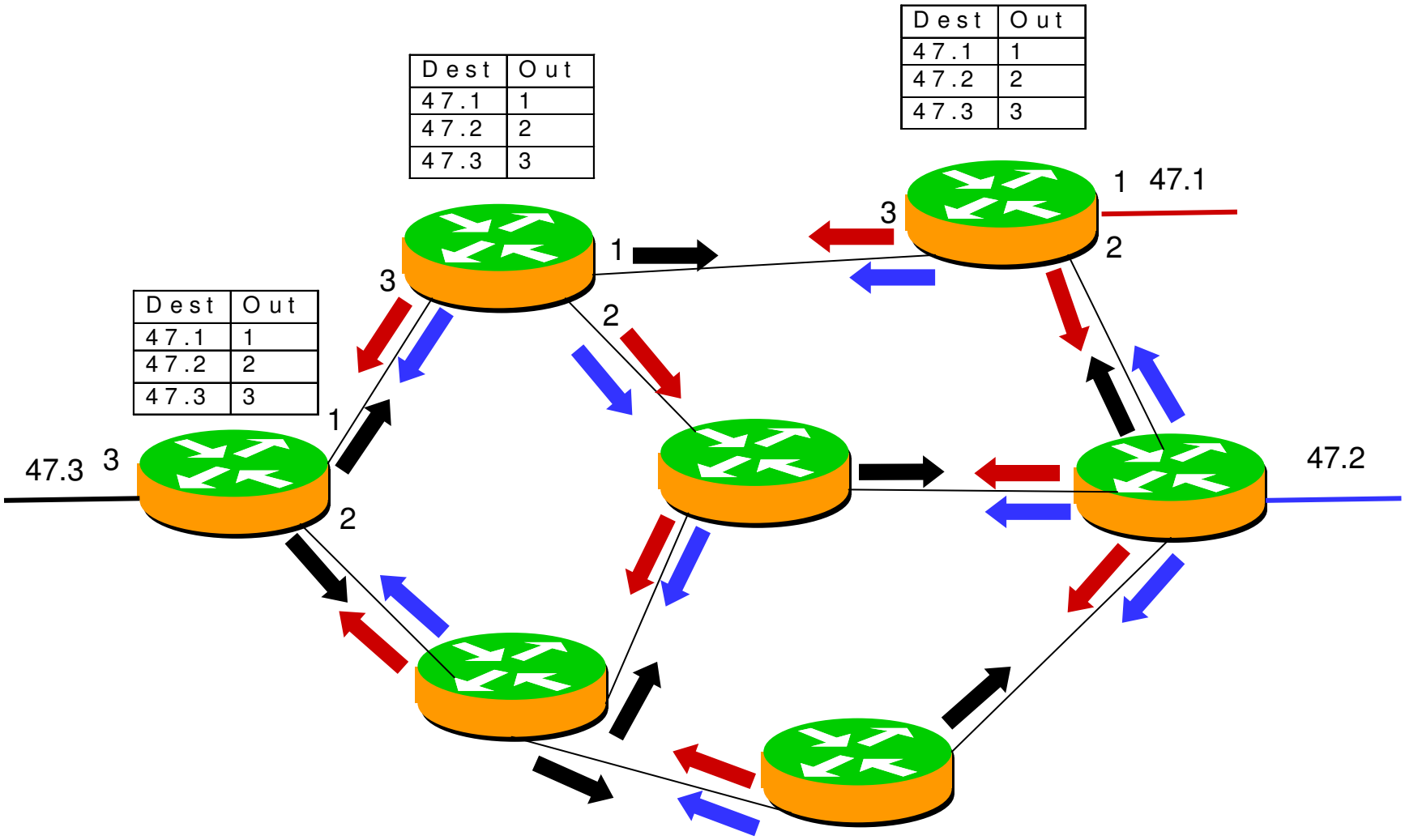
NANOG



- A Vanilla LSP is actually part of a tree from every source to that destination (unidirectional).
- Vanilla LDP builds that tree using existing IP forwarding tables to route the control messages.

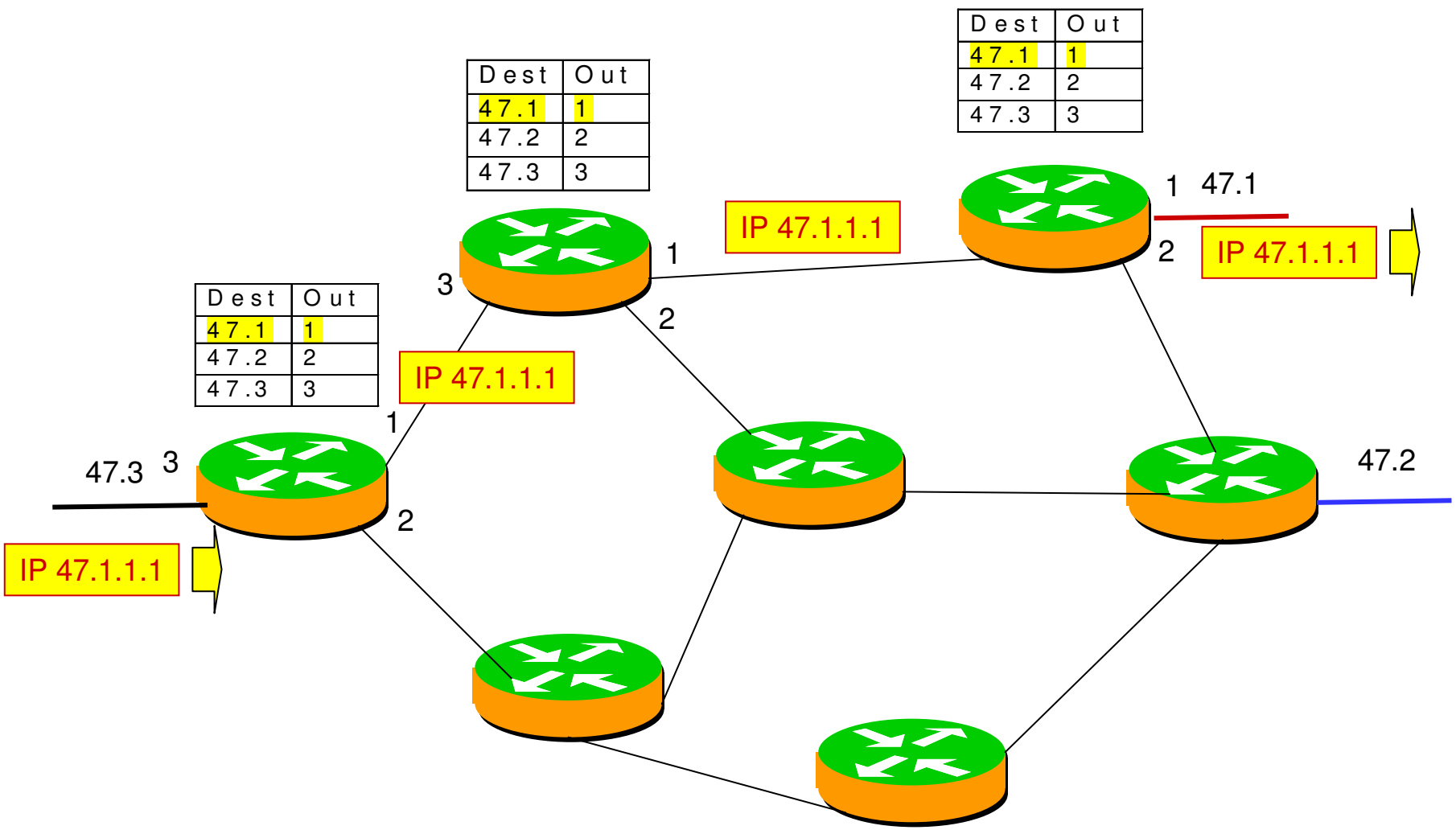
MPLS BUILT ON STANDARD IP

NANOG

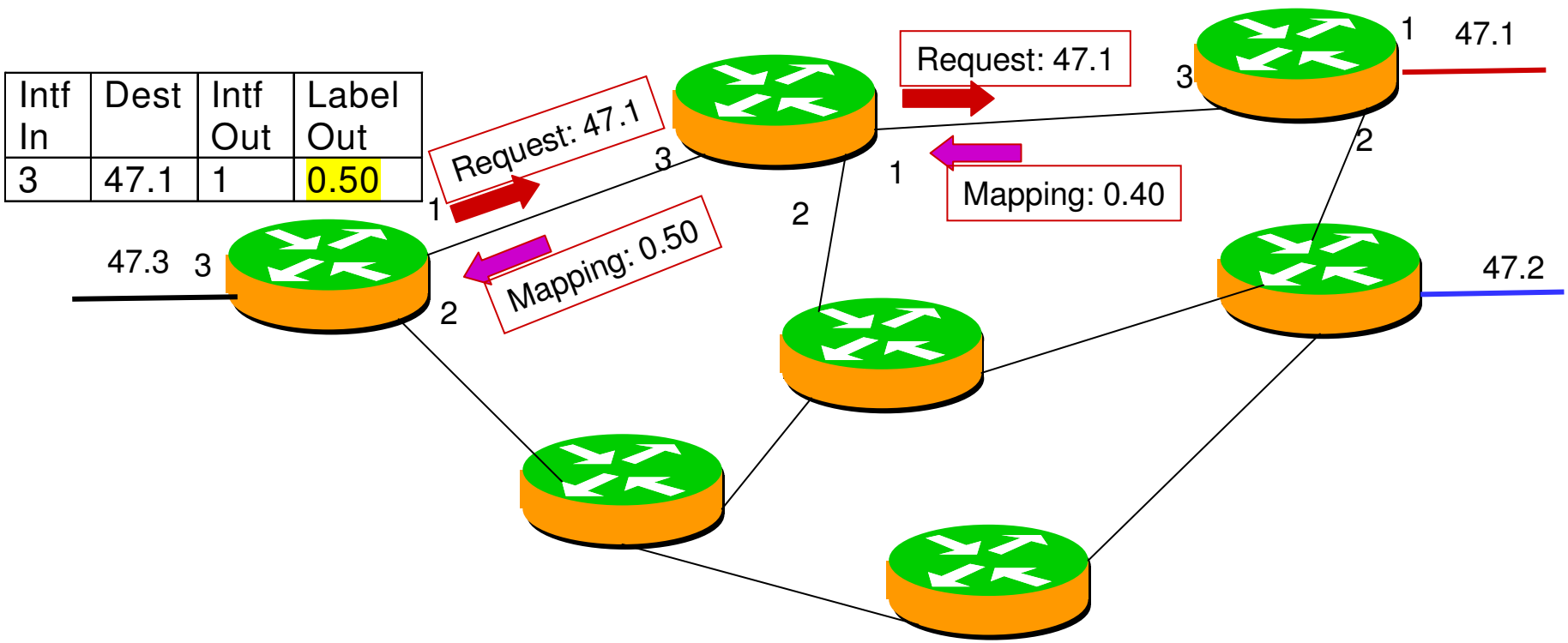


- **Destination based forwarding tables as built by OSPF, IS-IS, RIP, etc.**

IP FORWARDING USED BY HOP-BY-HOP CONTROL



MPLS Label Distribution



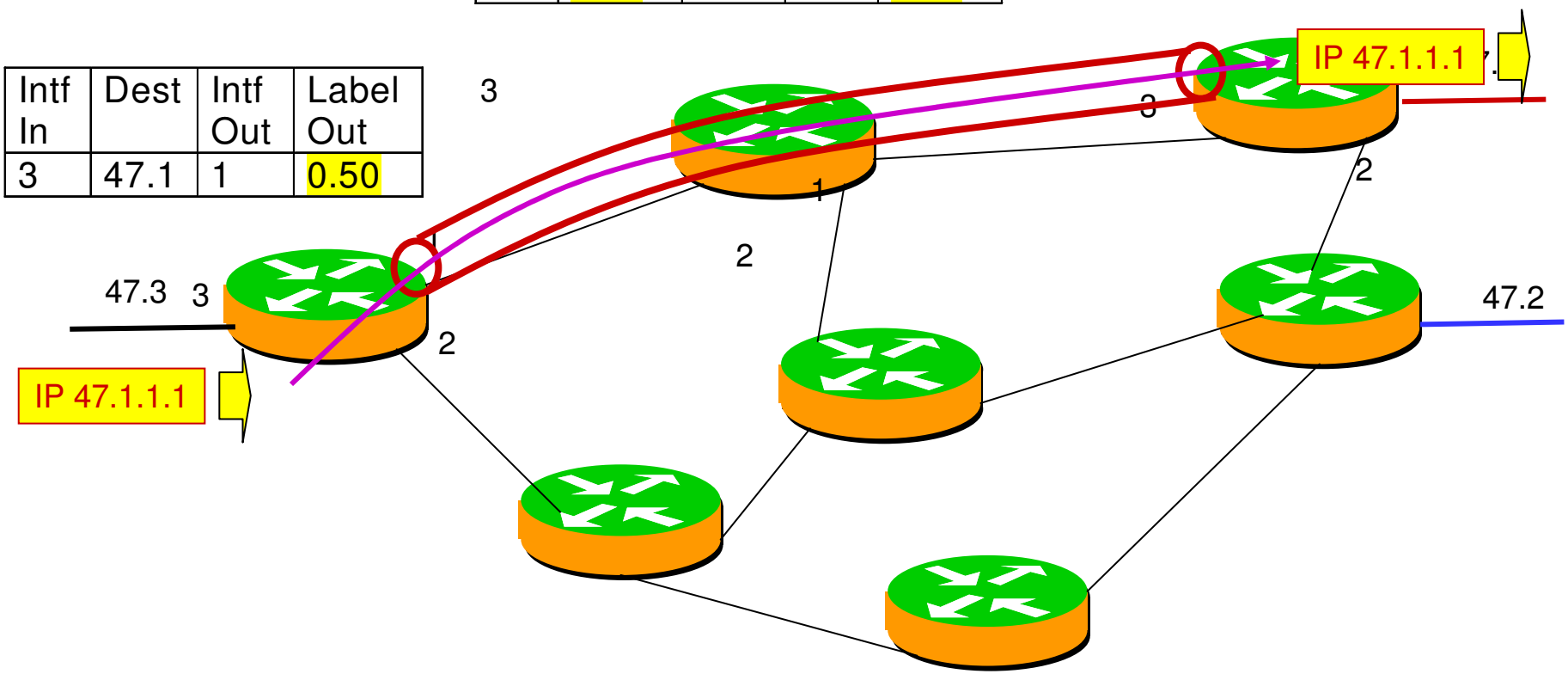
Label Switched Path (LSP)

NANOG

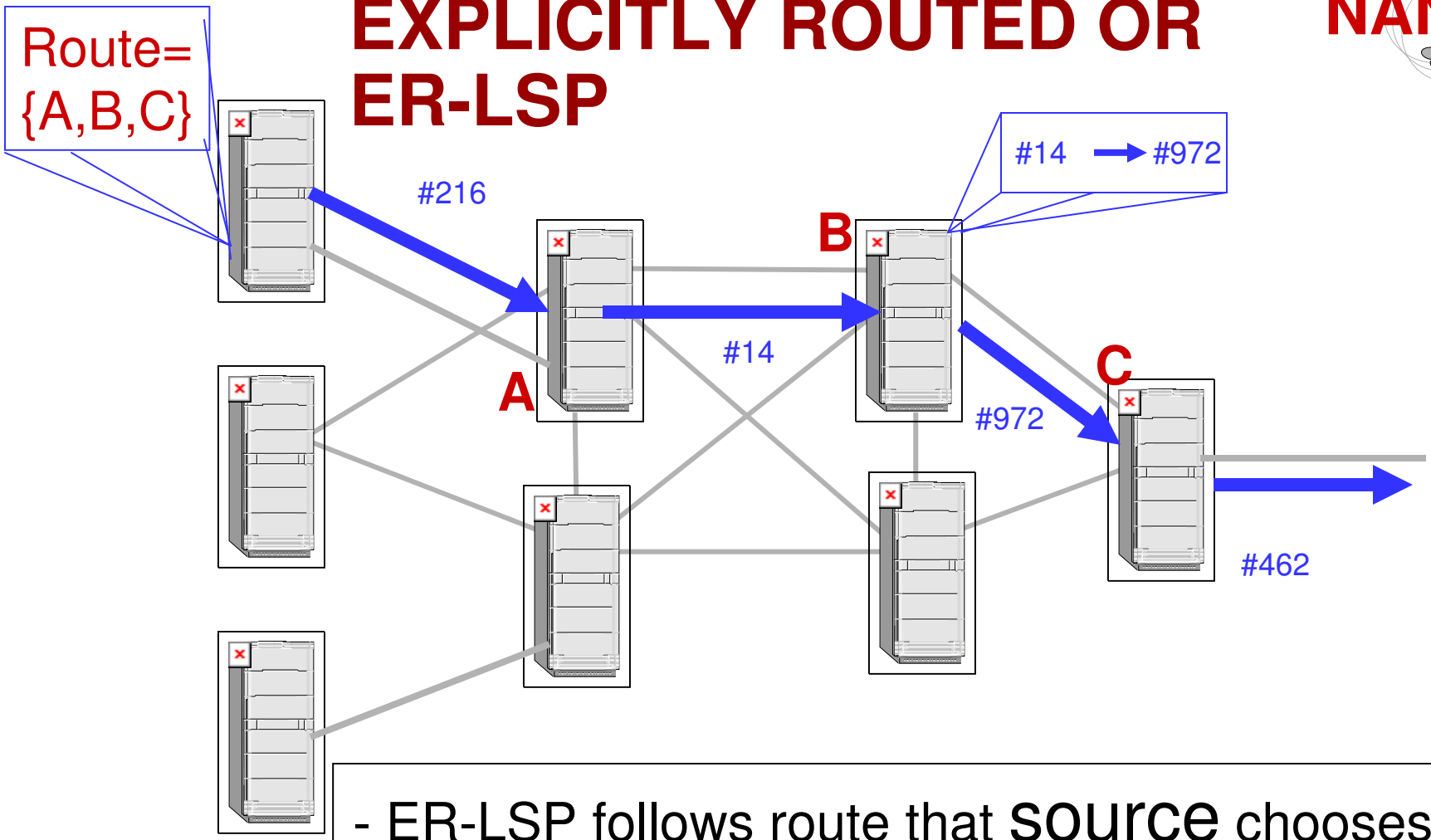
Intf In	Label In	Dest	Intf Out	Label Out
3	0.50	47.1	1	0.40

Intf In	Label In	Dest	Intf Out
3	0.40	47.1	1

Intf In	Dest	Intf Out	Label Out
3	47.1	1	0.50



EXPLICITLY ROUTED OR ER-LSP



- ER-LSP follows route that source chooses. In other words, the control message to establish the LSP (label request) is ***source routed***.

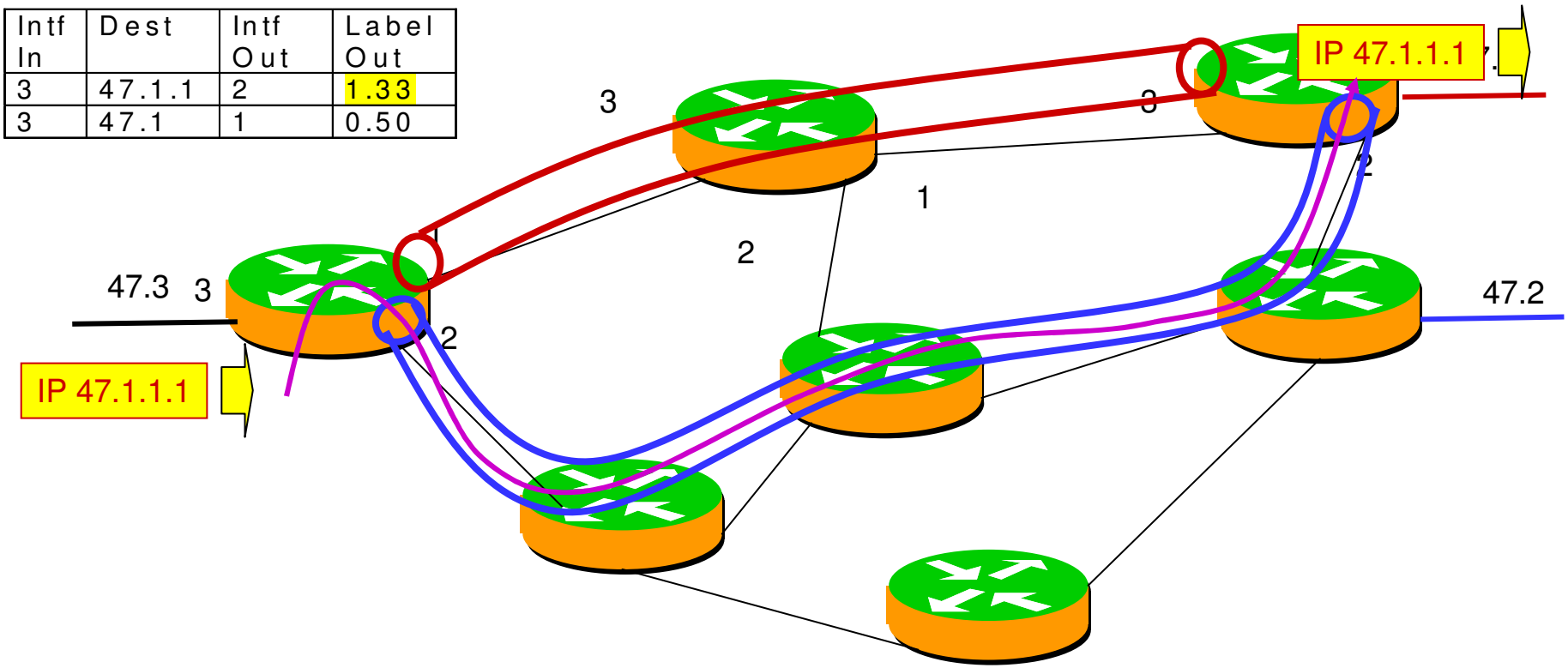
EXPLICITLY ROUTED LSP
ER-LSP

NANOG

Intf In	Label In	Dest	Intf Out	Label Out
3	0.50	47.1	1	0.40

Intf In	Label In	Dest	Intf Out
3	0.40	47.1	1

Intf In	Dest	Intf Out	Label Out
3	47.1.1	2	1.33
3	47.1	1	0.50



ER LSP - advantages

NANOG

- Operator has routing flexibility (policy-based, QoS-based)
- Can use routes other than shortest path
- Can compute routes based on constraints in exactly the same manner as ATM based on distributed topology database. (traffic engineering)

ER LSP - discord!

NANOG

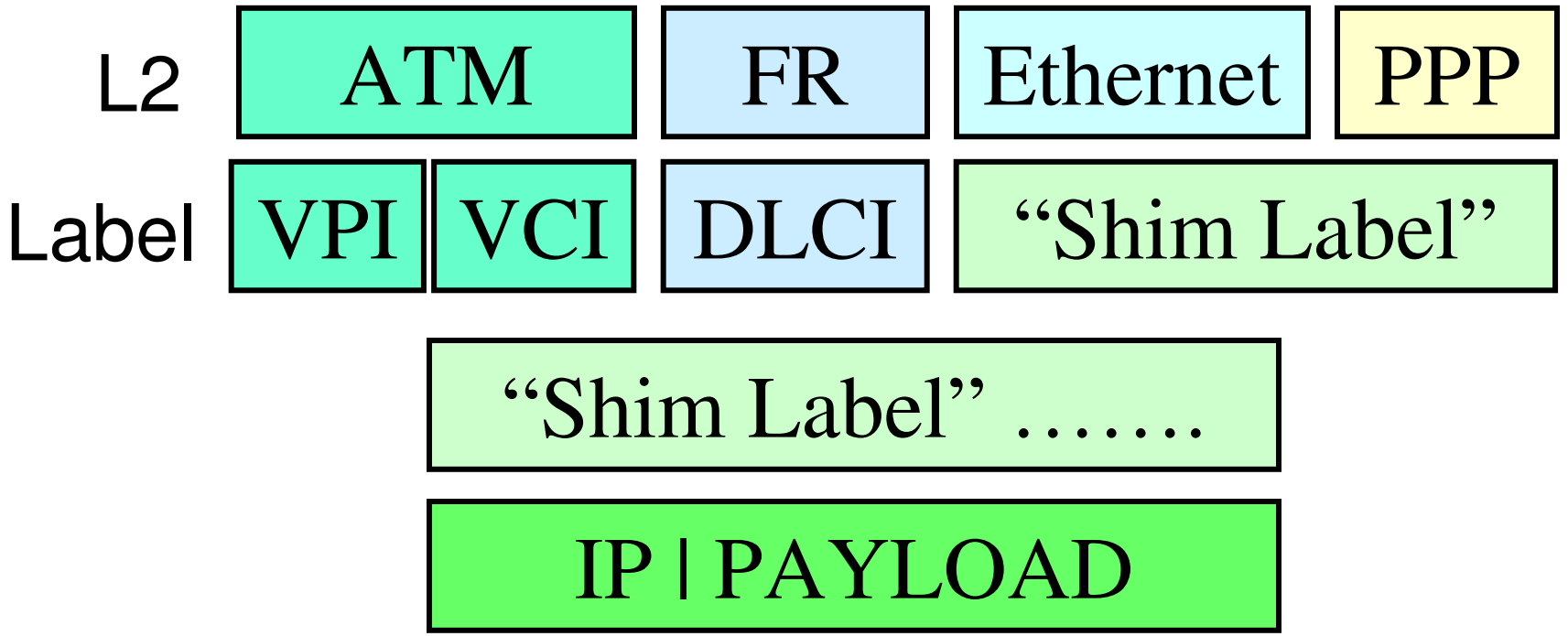
- Two signaling options proposed in the standards: CR-LDP, RSVP extensions:
  - CR-LDP = LDP + Explicit Route
  - RSVP ext = Traditional RSVP + Explicit Route + Scalability Extension
- ITU has decided on LDP/CR-LDP for public networks.
- Survival of the fittest not such a bad thing although RSVP has lots of work in scalability to do.

Tutorial Outline

NANOG

- Overview
- **Label Encapsulations**
- Label Distribution Protocols
- MPLS & ATM
- Constraint Based Routing with CR-LDP
- Operational Experiences with Similar Protocols
- Summary

Label Encapsulation



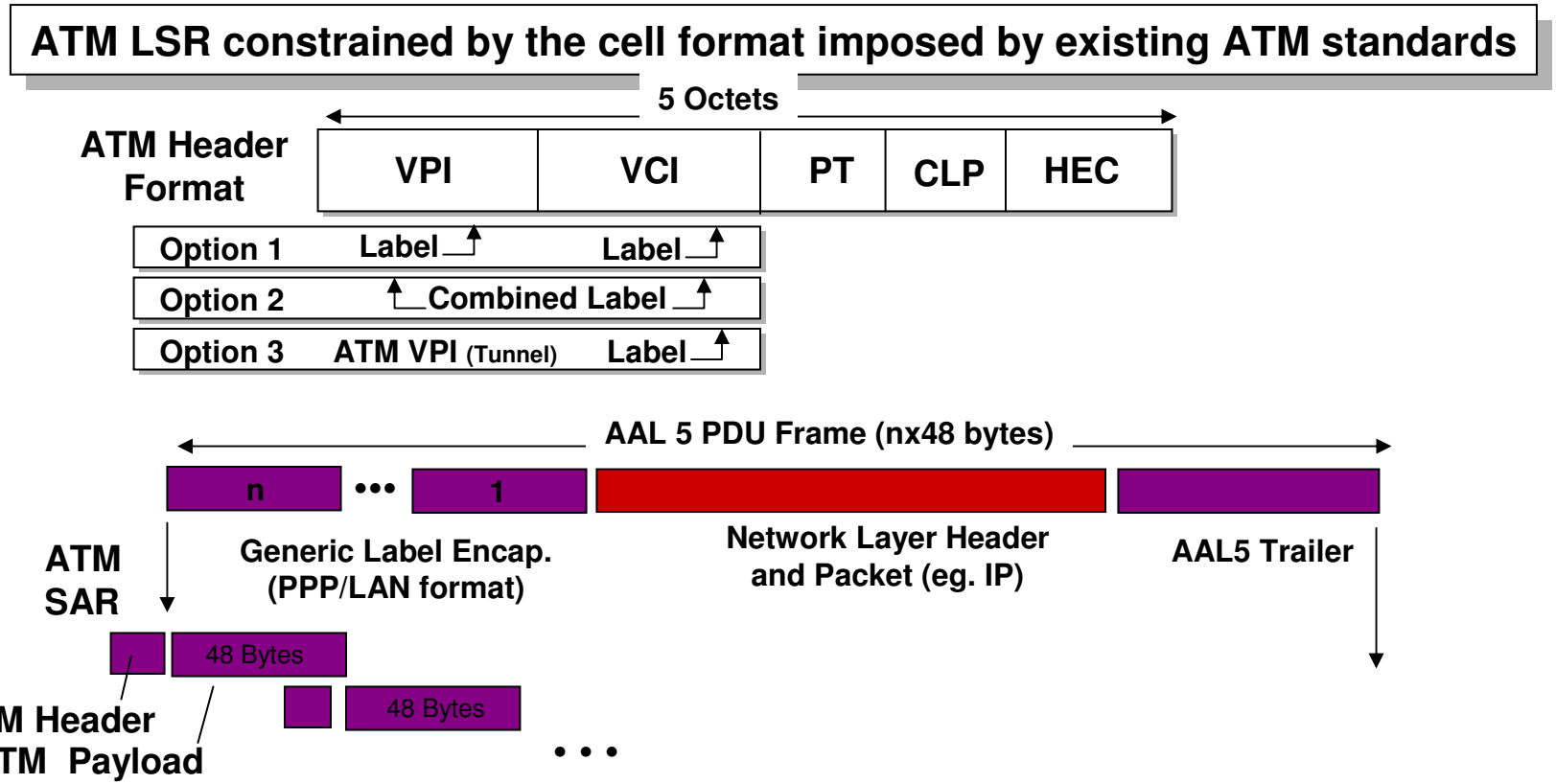
MPLS Encapsulation is specified over various media types. Top labels may use existing format, lower label(s) use a new "shim" label format.

MPLS Link Layers

- MPLS is intended to run over multiple link layers
- Specifications for the following link layers currently exist:
  - ATM: label contained in VCI/VPI field of ATM header
  - Frame Relay: label contained in DLCI field in FR header
  - PPP/LAN: uses 'shim' header inserted between L2 and L3 headers
- Translation between link layers types must be supported

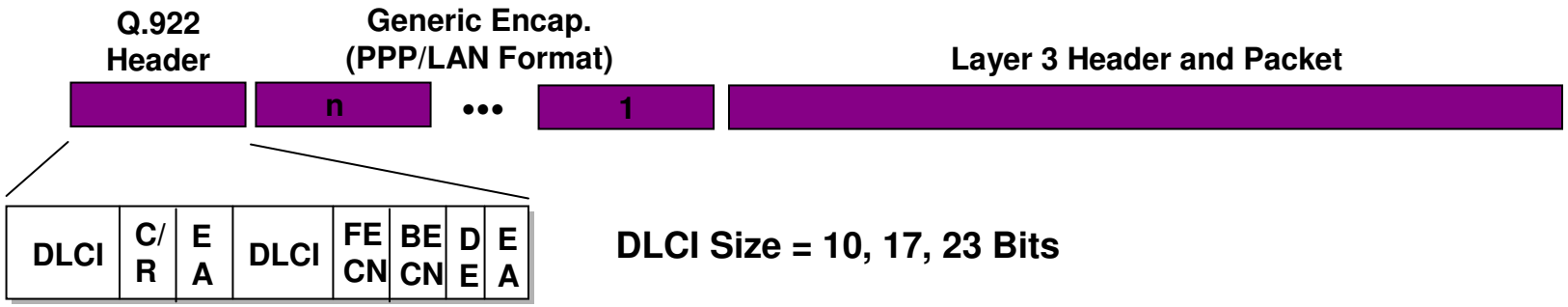
MPLS intended to be "multi-protocol" below as well as above.

MPLS Encapsulation - ATM



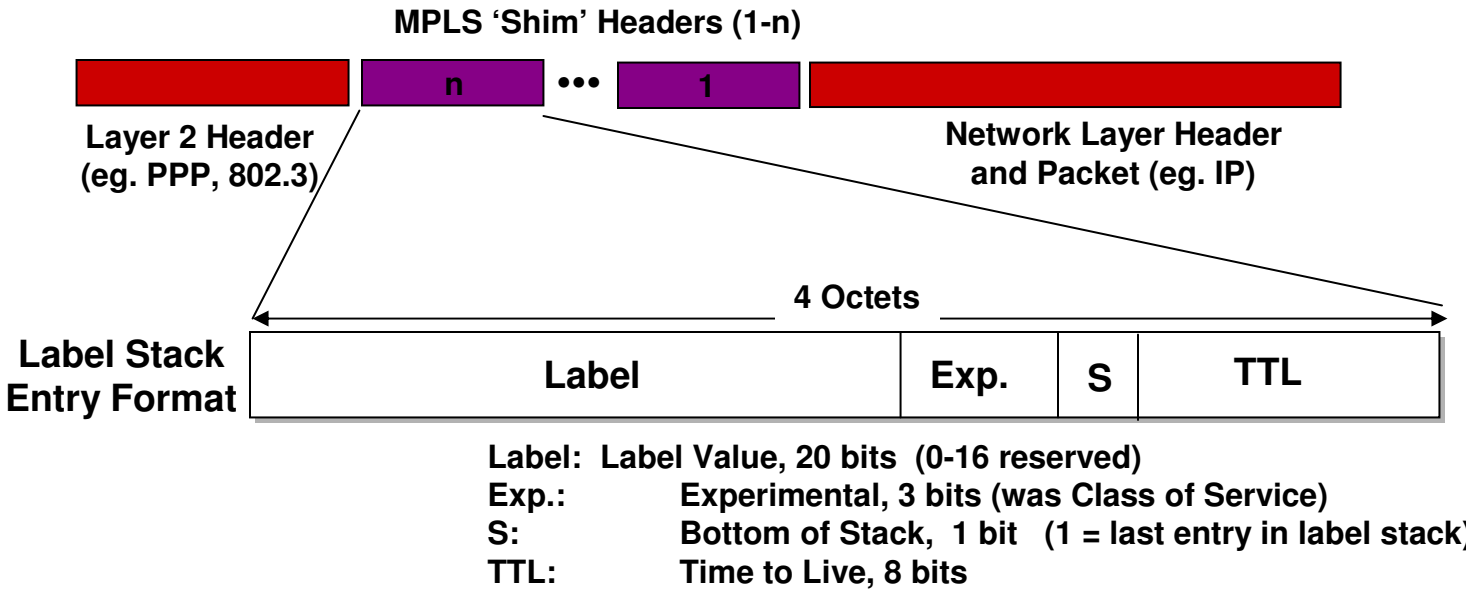
- Top 1 or 2 labels are contained in the VPI/VCI fields of ATM header
  - one in each or single label in combined field, negotiated by LDP
- Further fields in stack are encoded with 'shim' header in PPP/LAN format
  - must be at least one, with bottom label distinguished with 'explicit NULL'
- TTL is carried in top label in stack, as a proxy for ATM header (that lacks TTL)

MPLS Encapsulation - Frame Relay



- Current label value carried in DLCI field of Frame Relay header
- Can use either 2 or 4 octet Q.922 Address (10, 17, 23 bytes)
- Generic encapsulation contains n labels for stack of depth n
  - top label contains TTL (which FR header lacks), 'explicit NULL' label value

MPLS Encapsulation - PPP & LAN Data Links



Label: Label Value, 20 bits (0-16 reserved)
Exp.: Experimental, 3 bits (was Class of Service)
S: Bottom of Stack, 1 bit (1 = last entry in label stack)
TTL: Time to Live, 8 bits

- Network layer must be inferable from value of bottom label of the stack
- TTL must be set to the value of the IP TTL field when packet is first labelled
- When last label is popped off stack, MPLS TTL to be copied to IP TTL field
- Pushing multiple labels may cause length of frame to exceed layer-2 MTU
  - LSR must support "Max. IP Datagram Size for Labelling" parameter
  - any unlabelled datagram greater in size than this parameter is to be fragmented

MPLS on PPP links and LANs uses 'Shim' Header Inserted Between Layer 2 and Layer 3 Headers

Tutorial Outline

- **Overview**
- Label Encapsulations
- **Label Distribution Protocols**
- **MPLS & ATM**
- **Constraint Based Routing with CR-LDP**
- **Operational Experiences with Similar Protocols**
- **Summary**

Label Distribution Protocols

- **Overview of Hop-by-hop & Explicit**
- **Label Distribution Protocol (LDP)**
- **Constraint-based Routing LDP (CR-LDP)**
- **Extensions to RSVP**

Hop-by-Hop vs. Explicit Routing

Hop-by-Hop Routing

- Distributes routing of control traffic
- Builds a set of trees either fragment by fragment like a random fill, or backwards, or forwards in organized manner.
- Reroute on failure impacted by convergence time of routing protocol
- Existing routing protocols are destination prefix based
- Difficult to perform traffic engineering, QoS-based routing

Explicit Routing

- Source routing of control traffic
- Builds a path from source to dest
- Requires manual provisioning, or automated creation mechanisms.
- LSPs can be ranked so some reroute very quickly and/or backup paths may be pre-provisioned for rapid restoration
- Operator has routing flexibility (policy-based, QoS-based,
- Adapts well to traffic engineering

Explicit routing shows great promise for traffic engineering

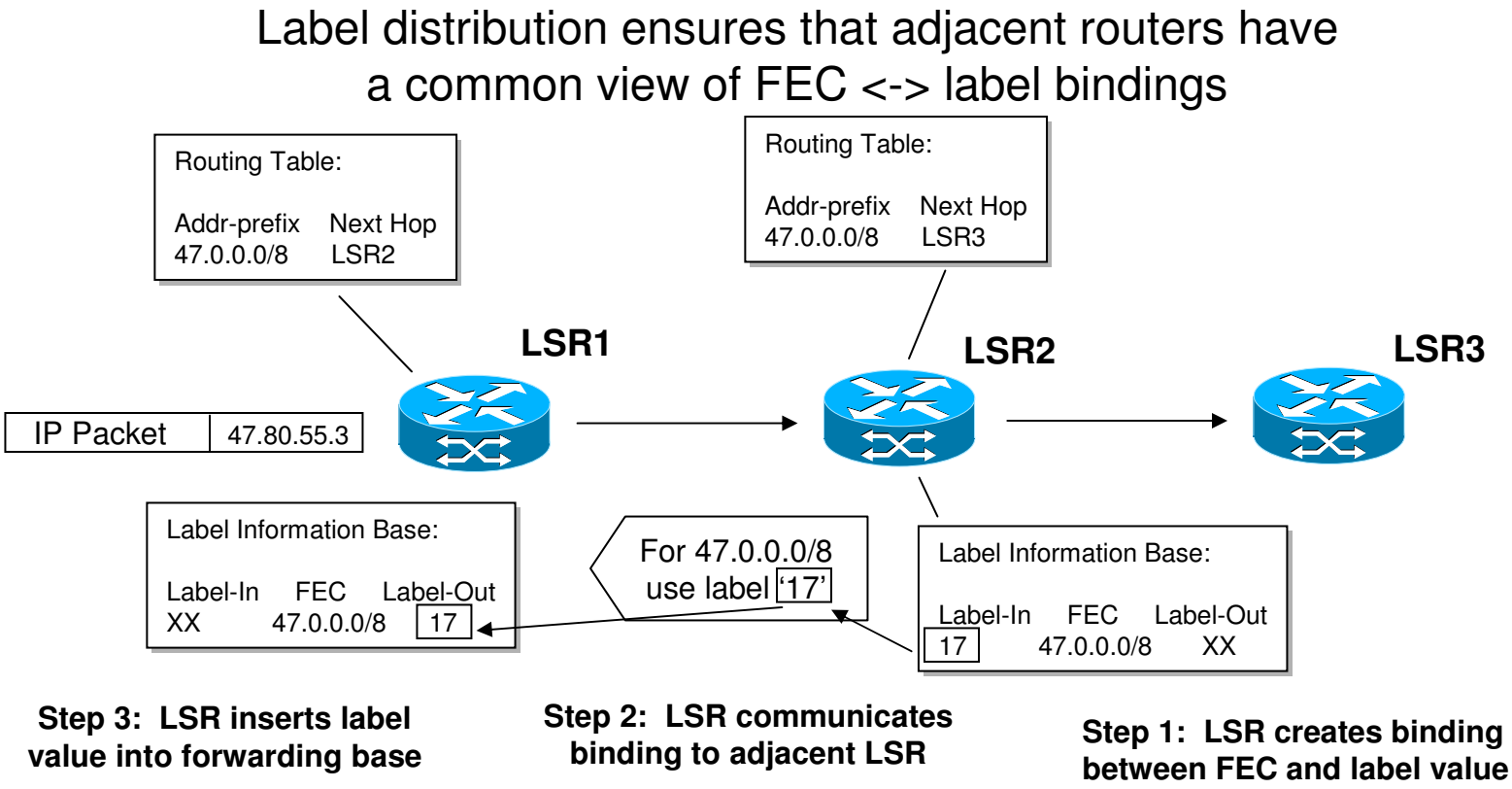
Explicit Routing - MPLS vs. IP Source Routing

- Connectionless nature of IP implies that routing is based on information in each packet header.
- Source routing is possible, but path must be contained in each IP header.
- Lengthy paths increase size of IP header, make it variable size, increase overhead.
- Some gigabit routers require ‘slow path’ option-based routing of IP packets.
- Source routing has not been widely adopted in IP and is seen as impractical.
- Some network operators may filter source routed packets for security reasons.
- MPLS enables the use of source routing by its connection-oriented capabilities.
    - paths can be explicitly set up through the network
    - the ‘label’ can now represent the explicitly routed path
- Loose and strict source routing can be supported.

Label Distribution Protocols

- Overview of Hop-by-hop & Explicit
- **Label Distribution Protocol (LDP)**
- Constraint-based Routing LDP (CR-LDP)
- Extensions to RSVP
- Extensions to BGP

Label Distribution Protocol (LDP) - Purpose



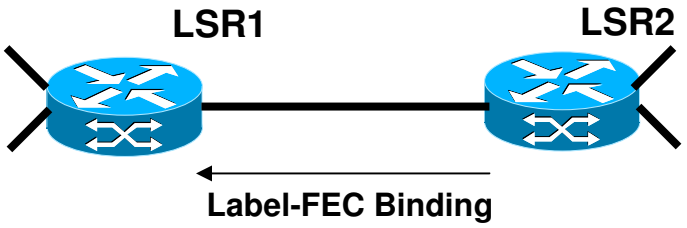
Common understanding of which FEC the label is referring to!

Label distribution can either piggyback on top of an existing routing protocol, or a dedicated label distribution protocol (LDP) can be created.

Label Distribution - Methods

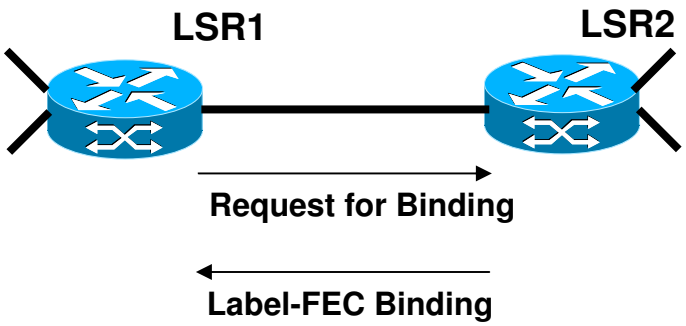
Label Distribution can take place using one of two possible methods

Downstream Unsolicited Label Distribution



- LSR2 and LSR1 are said to have an "LDP adjacency" (LSR2 being the downstream LSR)
- LSR2 discovers a 'next hop' for a particular FEC
- LSR2 generates a label for the FEC and communicates the binding to LSR1
- LSR1 inserts the binding into its forwarding tables
- If LSR2 is the next hop for the FEC, LSR1 can use that label knowing that its meaning is understood

Downstream-on-Demand Label Distribution

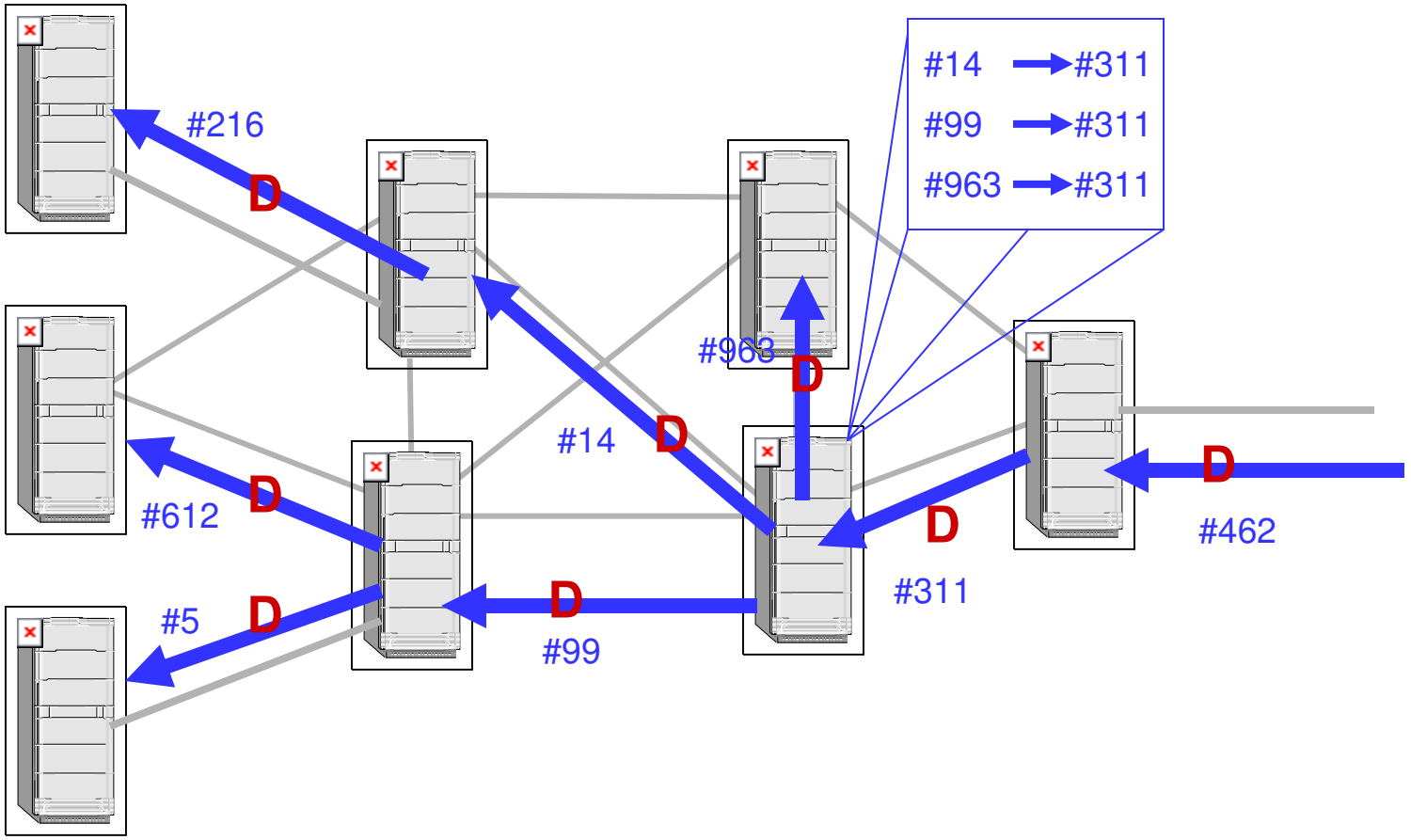


- LSR1 recognizes LSR2 as its next-hop for an FEC
- A request is made to LSR2 for a binding between the FEC and a label
- If LSR2 recognizes the FEC and has a next hop for it, it creates a binding and replies to LSR1
- Both LSRs then have a common understanding

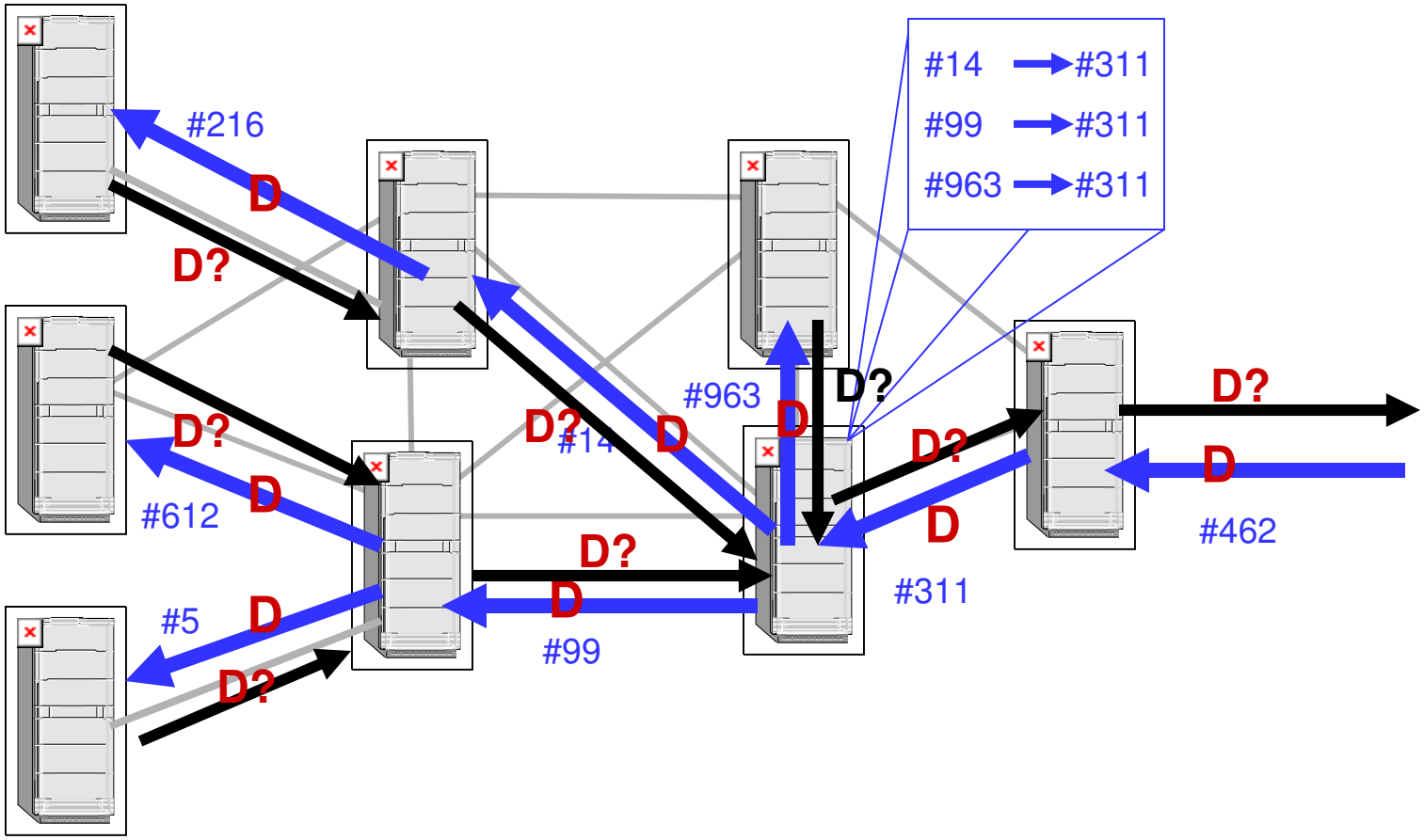
Both methods are supported, even in the same network at the same time

For any single adjacency, LDP negotiation must agree on a common method

Downstream Mode Making SPF Tree Copy In H/W

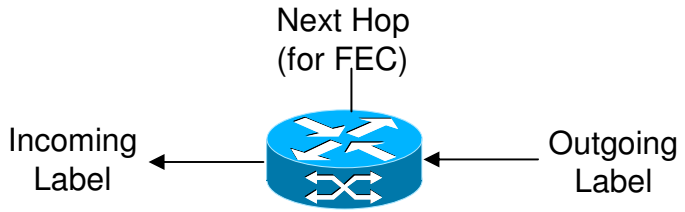


Downstream On Demand Making SPF Tree Copy In H/W



Distribution Control: Ordered v. Independent

MPLS path forms as associations are made between FEC next-hops and incoming and outgoing labels

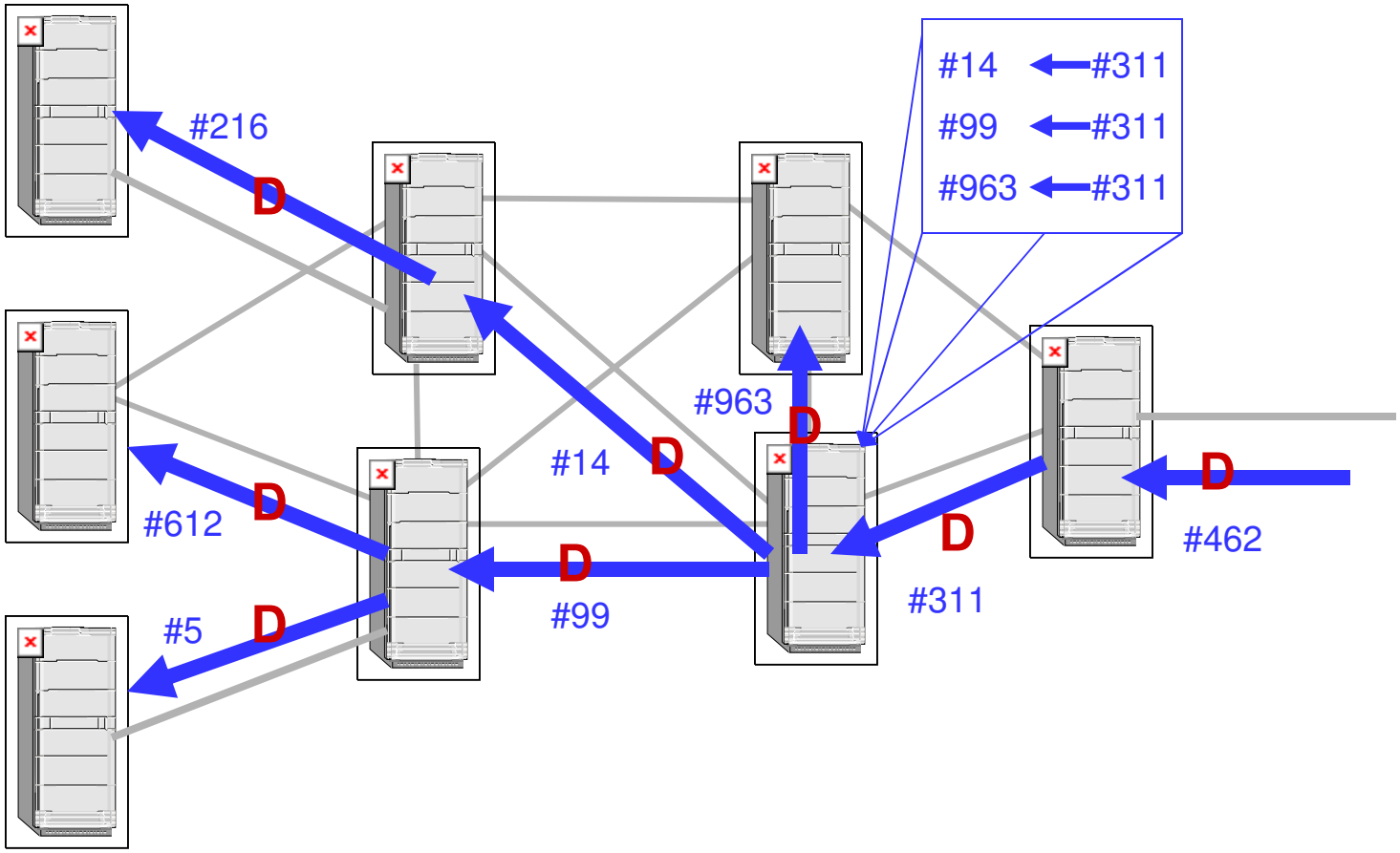


	Independent LSP Control	Ordered LSP Control
Definition	• Each LSR makes independent decision on when to generate labels and communicate them to upstream peers • Communicate label-FEC binding to peers once next-hop has been recognized • LSP is formed as incoming and outgoing labels are spliced together	• Label-FEC binding is communicated to peers if: - LSR is the 'egress' LSR to particular FEC - label binding has been received from upstream LSR • LSP formation 'flows' from egress to ingress
Comparison	• Labels can be exchanged with less delay • Does not depend on availability of egress node • Granularity may not be consistent across the nodes at the start • May require separate loop detection/mitigation method	• Requires more delay before packets can be forwarded along the LSP • Depends on availability of egress node • Mechanism for consistent granularity and freedom from loops • Used for explicit routing and multicast

Both methods are supported in the standard and can be fully interoperable

INDEPENDENT MODE

NANOG

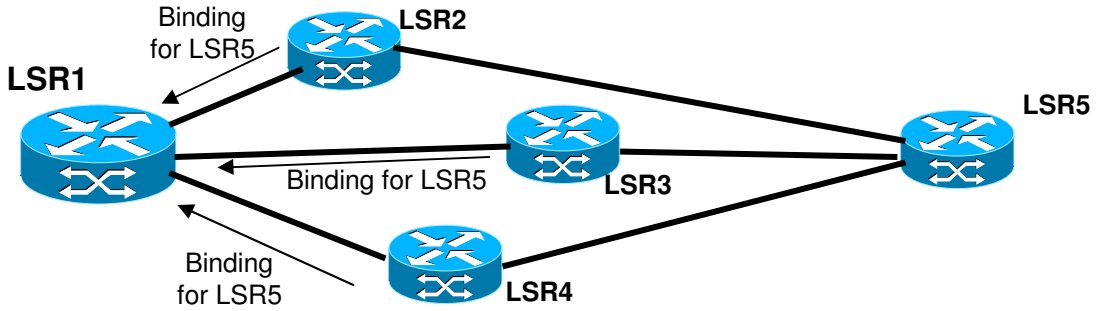


NORTEL NETWORKS™
How the world shares ideas.

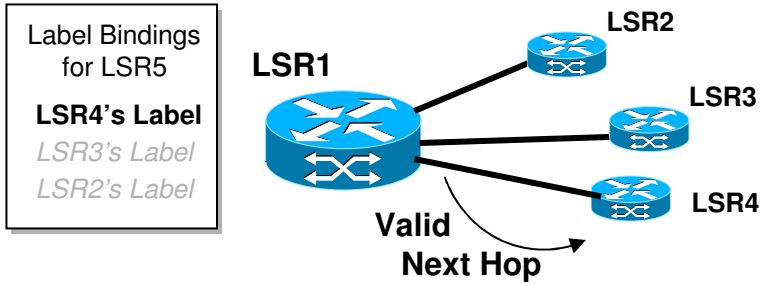
Label Retention Methods

An LSR may receive label bindings from multiple LSRs

Some bindings may come from LSRs that are not the valid next-hop for that FEC

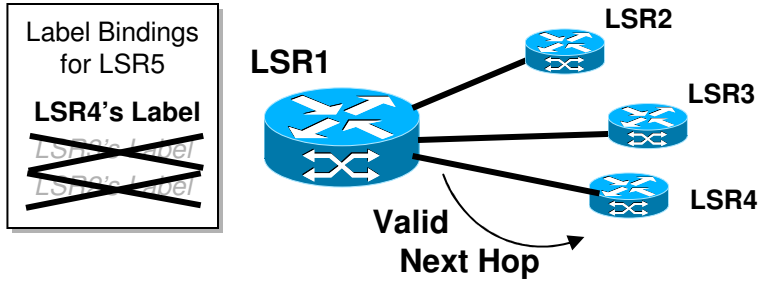


Liberal Label Retention



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

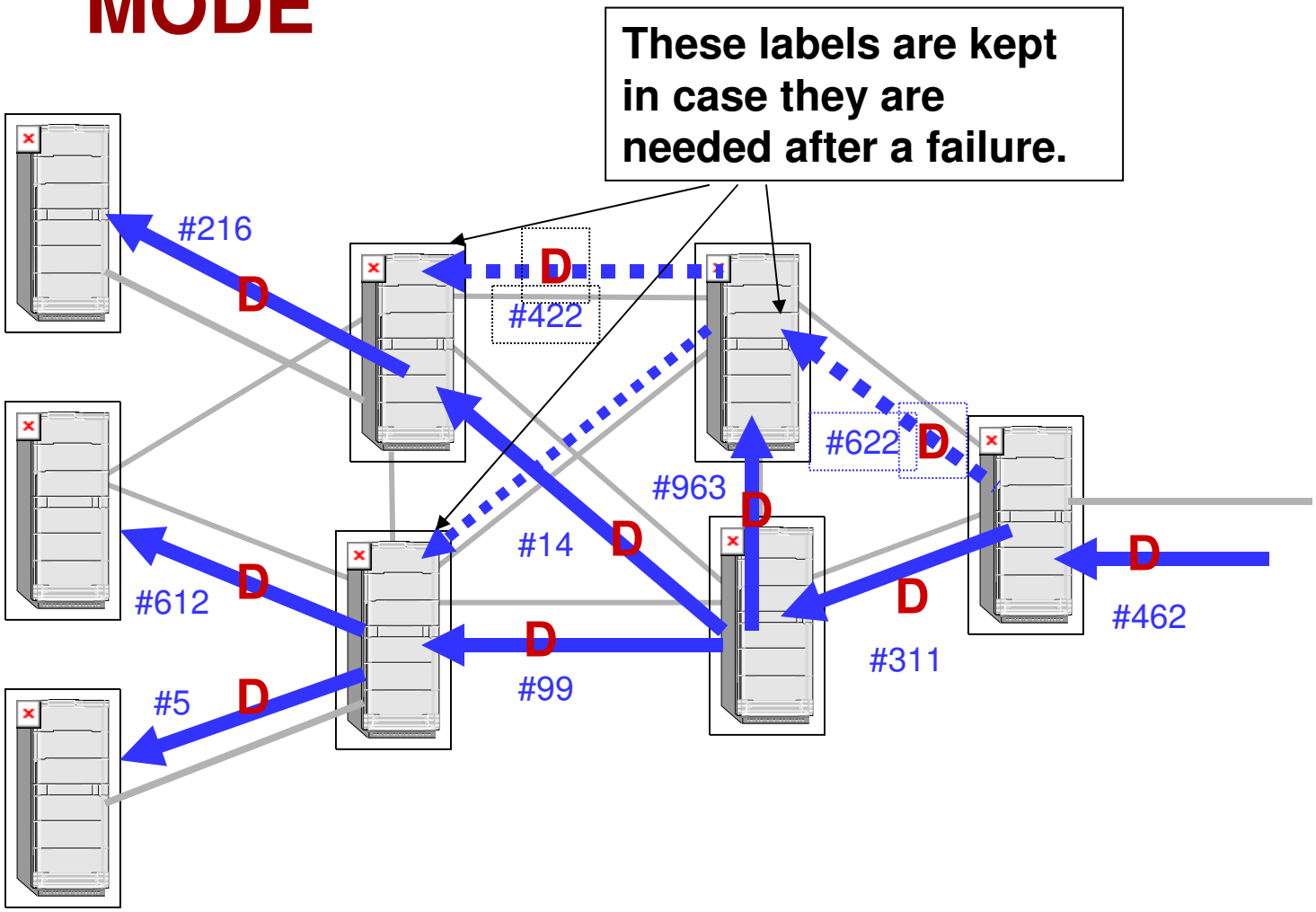
Conservative Label Retention



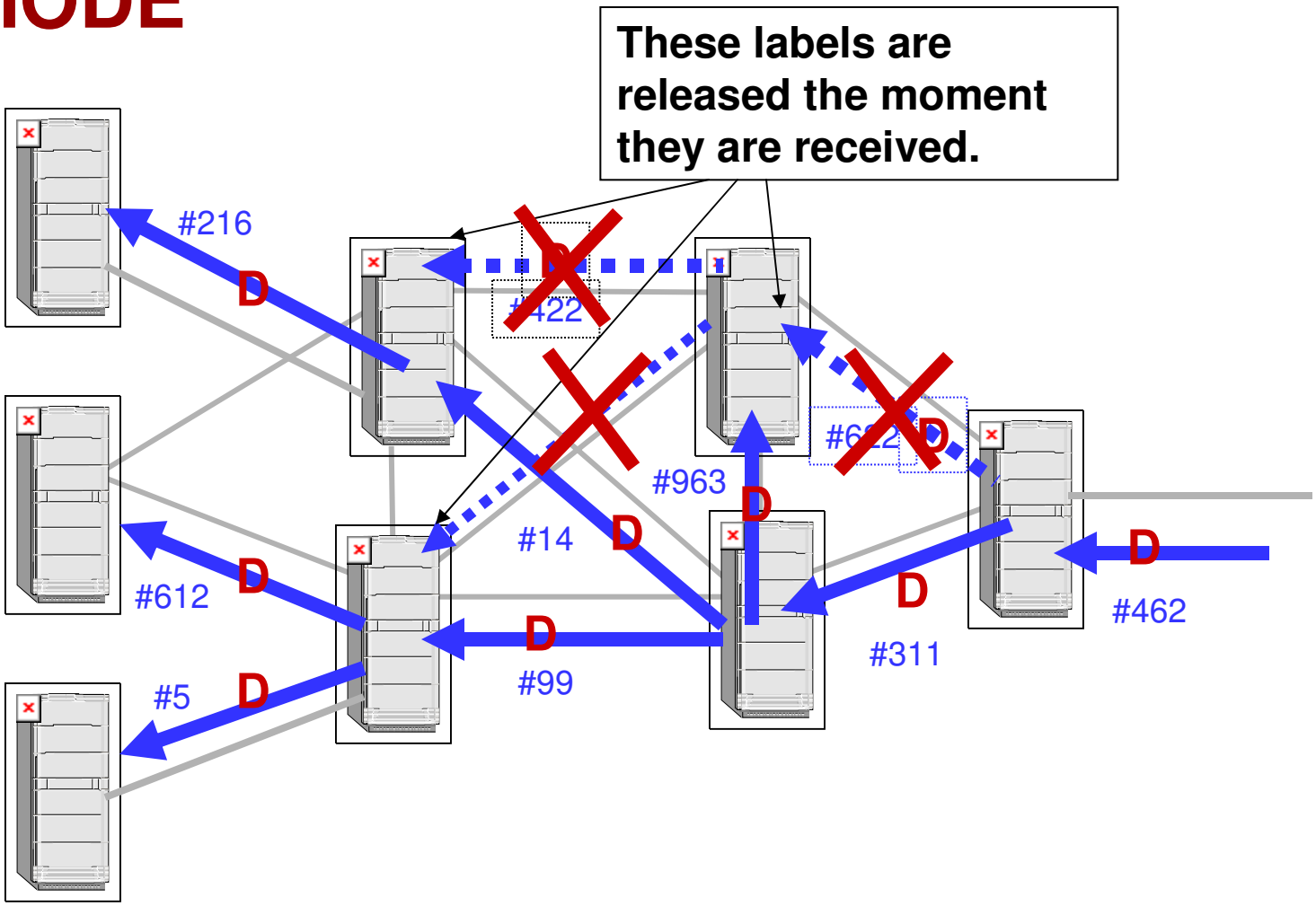
- LSR only maintains bindings received from valid next hop
- If the next-hop changes, binding must be requested from new next hop
- Restricts adaptation to changes in routing
- Fewer labels must be maintained by LSR

Label Retention method trades off between label capacity and speed of adaptation to routing changes

LIBERAL RETENTION MODE



CONSERVATIVE RETENTION MODE



LDP - STATUS

- Last Call Ended going to IESG for RFC also ITU SG13 has adopted for IP on ATM.
- Multi Vendor interoperability demonstrated for Downstream on demand mode on OC-3/ATM by (Nortel Networks, Ericson, Cisco, H&J, Ficom … 7 vendors) at Atlanta Interop/99
- Source code for these PDUs publicly available: www.NortelNetworks.com/mpls
- LINUX implementation using above code publicly available.

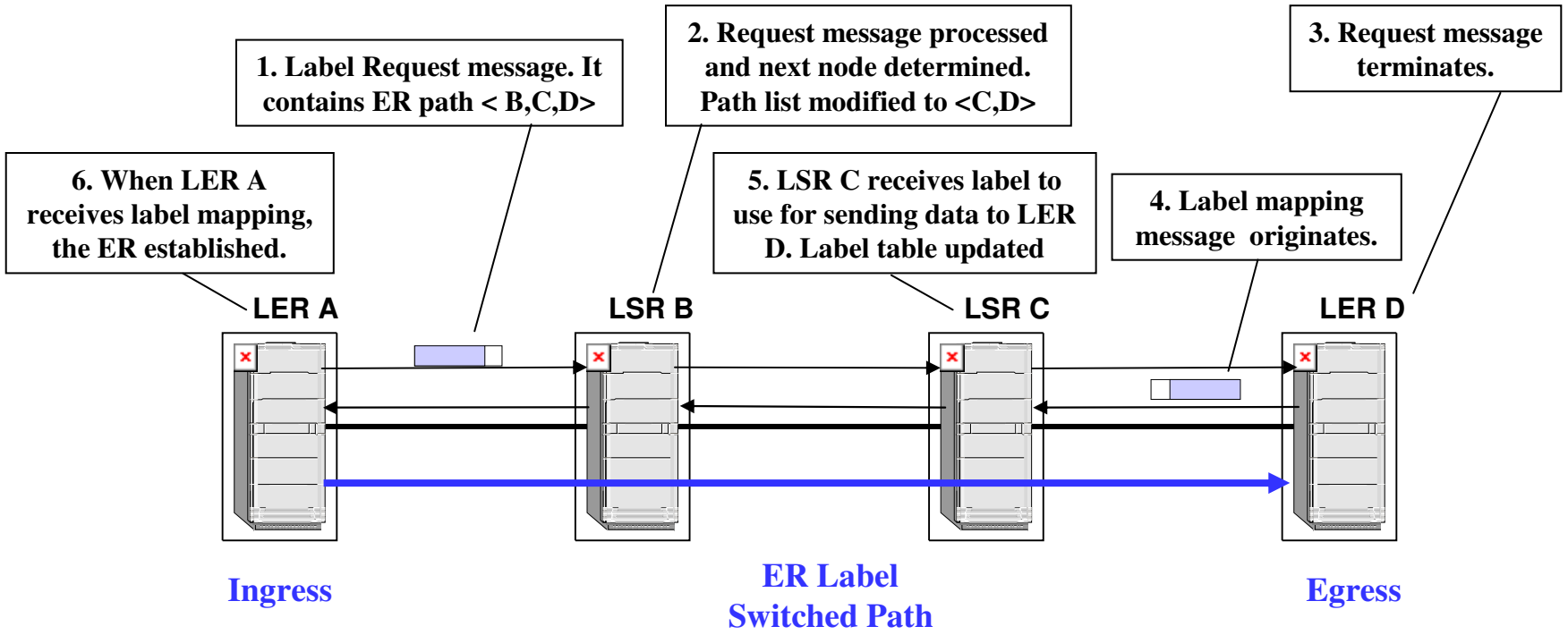
Label Distribution Protocols

- Overview of Hop-by-hop & Explicit
- Label Distribution Protocol (LDP)
- **Constraint-based Routing LDP (CR-LDP)**
- Extensions to RSVP

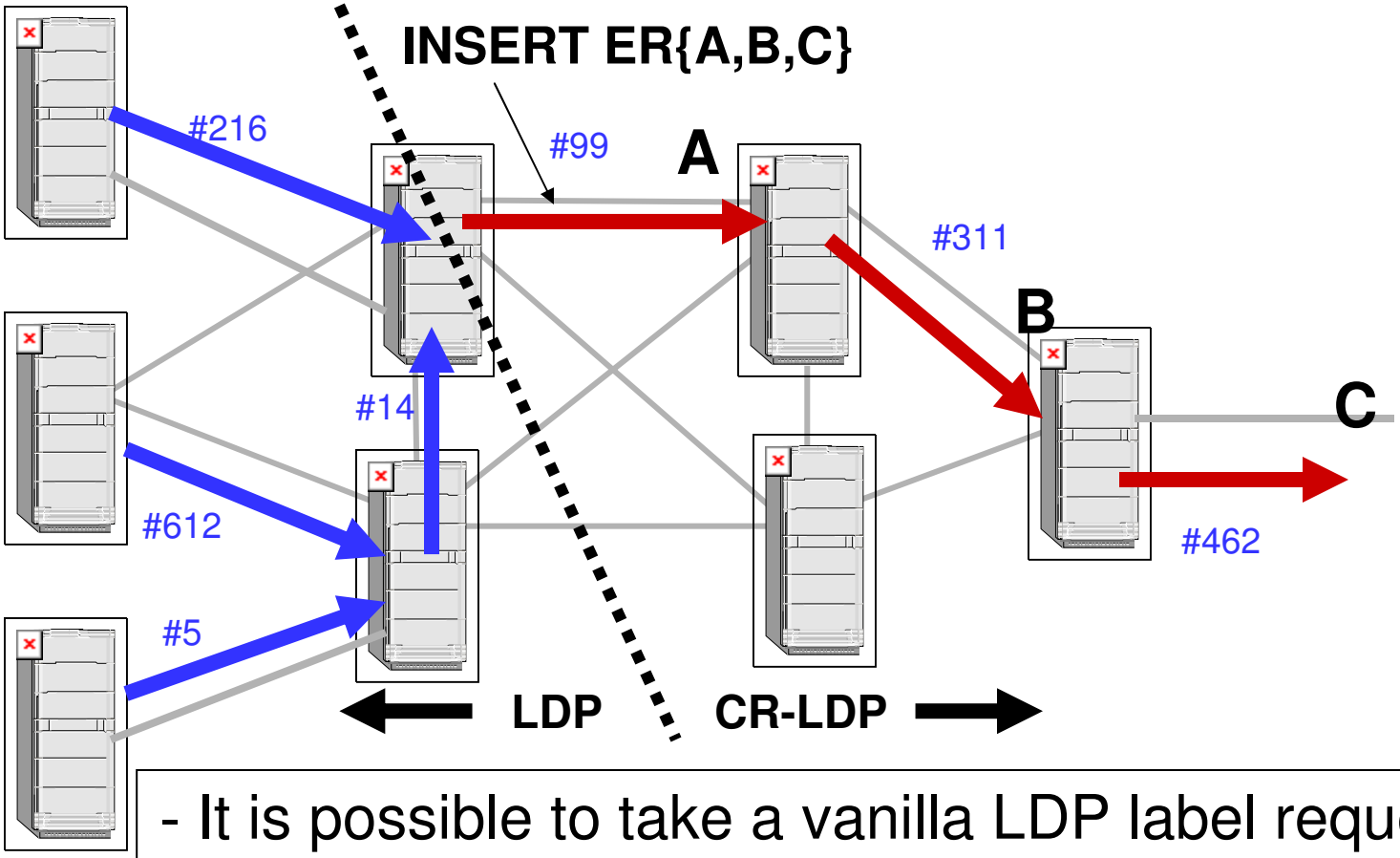
Constraint-based LSP Setup using LDP

- Uses LDP Messages (request, map, notify)
- Shares TCP/IP connection with LDP
- Can coexist with vanilla LDP and inter-work with it, or can exist as an entity on its own
- Introduces additional data to the vanilla LDP messages to signal ER, and other “Constraints”

ER-LSP Setup using CR-LDP



LDP/CR-LDP INTERWORKING



- It is possible to take a vanilla LDP label request let it flow vanilla to the edge of the core, insert an ER hop list at the core boundary at which point it is CR-LDP to the far side of the core.

Basic LDP Message additions

- LSPID: A unique tunnel identifier within an MPLS network.
- ER: An explicit route, normally a list of IPV4 addresses to follow (source route) the label request message.
- Resource Class (Color): to constrain the route to only links of this Color. Basically a 32 bit mask used for constraint based computations.
- Traffic Parameters: similar to ATM call setup, which specify treatment and reserve resources.

CR-LDP Traffic Parameters

U	F	Traf. Param. TLV	Length
Flags		Frequency	Reserved	Weight
Peak Data Rate (PDR)			
Peak Burst Size (PBS)			
Committed Data Rate (CDR)			
Committed Burst Size (CBS)			
Excess Burst Size (EBS)			

32 bit fields are short IEEE floating point numbers

Any parameter may be used or not used by selecting appropriate values

Flags control "negotiability" of parameters

Frequency constrains the variable delay that may be introduced

Weight of the CRLSP in the "relative share"

Peak rate (PDR+PBS) maximum rate at which traffic should be sent to the CRLSP

Committed rate (CDR+CBS) the rate that the MPLS domain commits to be available to the CRLSP

Excess Burst Size (EBS) to measure the extent by which the traffic sent on a CRLSP exceeds the committed rate

CRLSP characteristics not edge functions

- **The approach is like diff-serv's separation of PHB from Edge**
- **The parameters describe the "path behavior" of the CRLSP, i.e. the CRLSP's characteristics**
- **Dropping behavior is not signaled**
  - Dropping may be controlled by DS packet markings
- **CRLSP characteristics may be combined with edge functions (which are undefined in CRLDP) to create services**
  - Edge functions can perform packet marking
  - Example services are in an appendix

Peak rate

- **The maximum rate at which traffic should be sent to the CRLSP**
- **Defined by a token bucket with parameters**
  - Peak data rate (PDR)
  - Peak burst size (PBS)
- **Useful for resource allocation**
- **If a network uses the peak rate for resource allocation then its edge function should regulate the peak rate**
- **May be unused by setting PDR or PBS or both to positive infinity**

Committed rate

- **The rate that the MPLS domain commits to be available to the CRLSP**
- **Defined by a token bucket with parameters**
  - Committed data rate (CDR)
  - Committed burst size (CBS)
- **Committed rate is the bandwidth that should be reserved for the CRLSP**
- **CDR = 0 makes sense; CDR = +∞ less so**
- **CBS describes the burstiness with which traffic may be sent to the CRLSP**

Excess burst size

- Measure the extent by which the traffic sent on a CRLSP exceeds the committed rate
- Defined as an additional limit on the committed rate's token bucket
- Can be useful for resource reservation
- If a network uses the excess burst size for resource allocation then its edge function should regulate the parameter and perhaps mark or drop packets
- EBS = 0 and EBS = +∞ both make sense

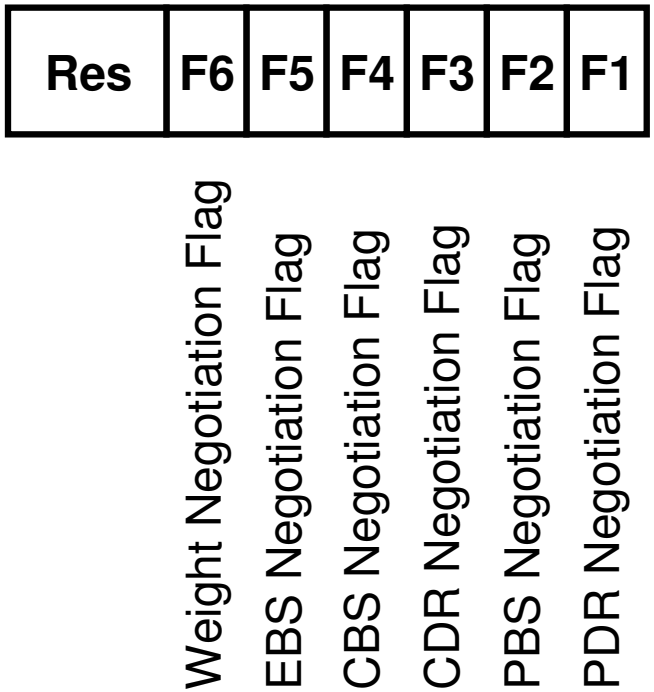
Frequency

- **Specifies how frequently the committed rate should be given to CRLSP**
- **Defined in terms of "granularity" of allocation of rate**
- **Constrains the variable delay that the network may introduce**
- **Constrains the amount of buffering that a LSR may use**
- **Values:**
  - Very frequently: no more than one packet may be buffered
  - Frequently: only a few packets may be buffered
  - Unspecified: any amount of buffering is acceptable

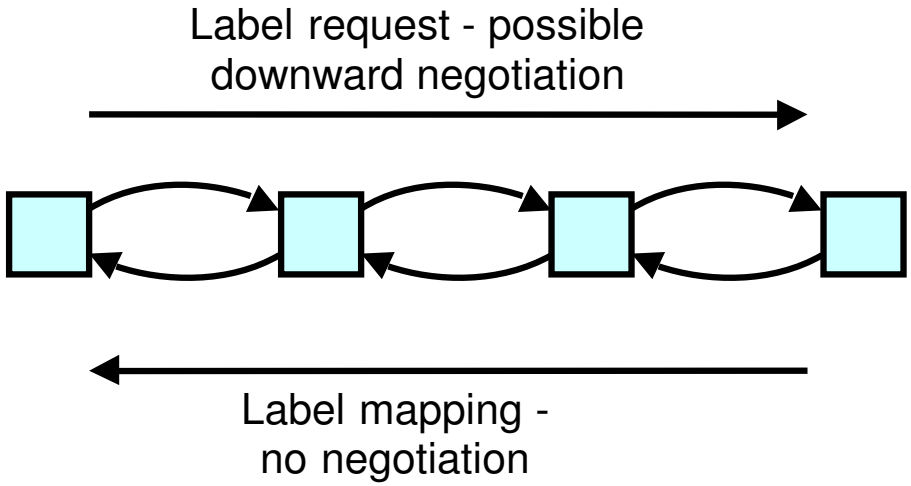
Weight

- **Specifies the CRLSP's weight in the "relative share algorithm"**
- **Implied but not stated:**
  - CRLSPs with a larger weight get a bigger relative share of the "excess bandwidth"
- **Values:**
  - 0 — the weight is not specified
  - 1-255 — weights; larger numbers are larger weights
- **The definition of "relative share" is network specific**

Negotiation flags



If a parameter is flagged as negotiable then LSRs may replace the parameter value with a smaller value in the label request message. LSRs discover the negotiated values in the label mapping message.



CR-LDP PREEMPTION

A CR-LSP carries an LSP priority. This priority can be used to allow new LSPs to *bump* existing LSPs of lower priority in order to steal their resources.

This is especially useful during times of failure and allows you to rank the LSPs such that the most important obtain resources before less important LSPs.

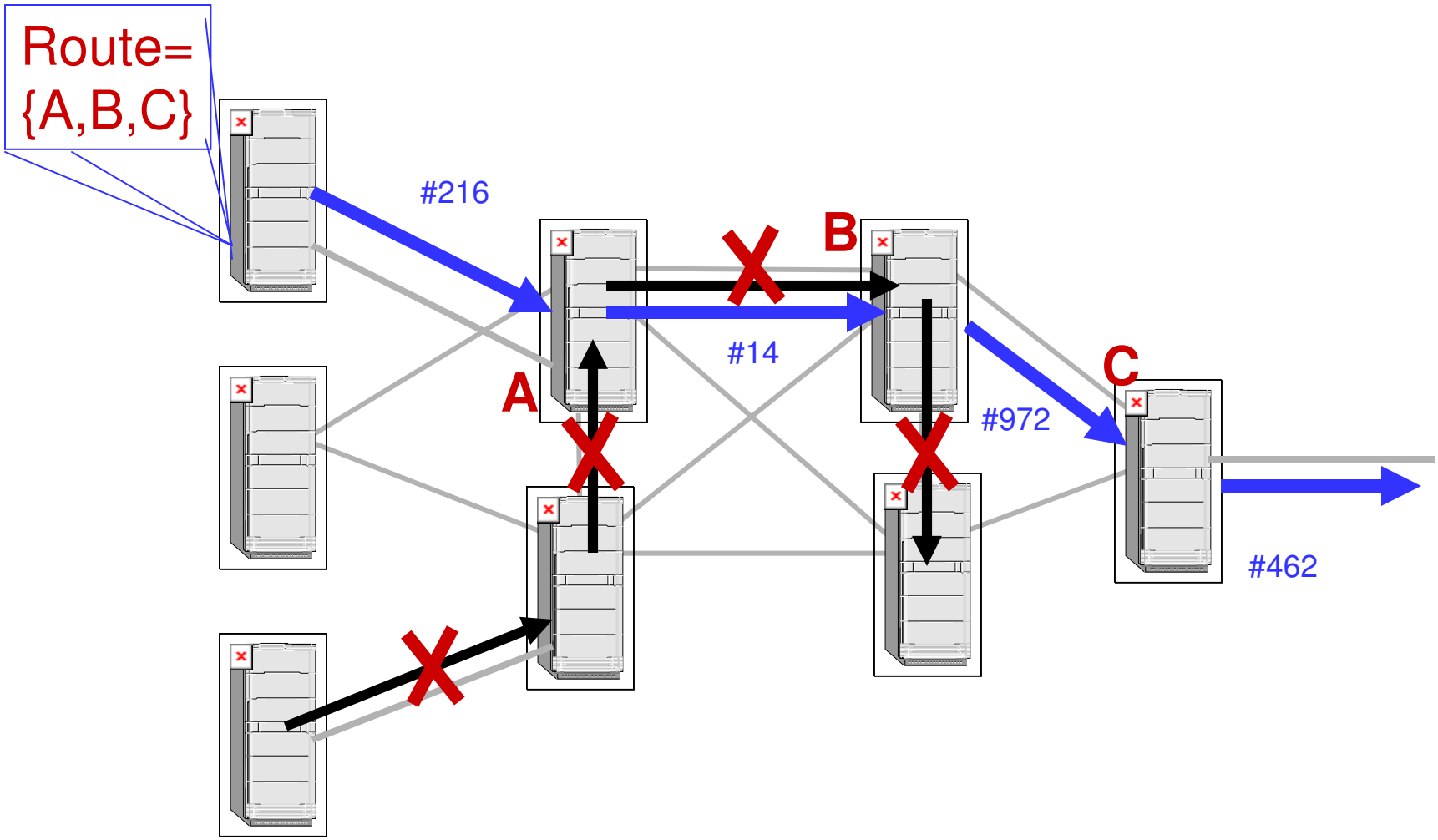
These are called the setupPriority and a holdingPriority and 8 levels are provided.

CR-LDP PREEMPTION

When an LSP is established its setupPriority is compared with the holdingPriority of existing LSPs, any with lower holdingPriority may be bumped to obtain their resources.

This process may continue in a domino fashion until the lowest holdingPriority LSPs either clear or are on the worst routes.

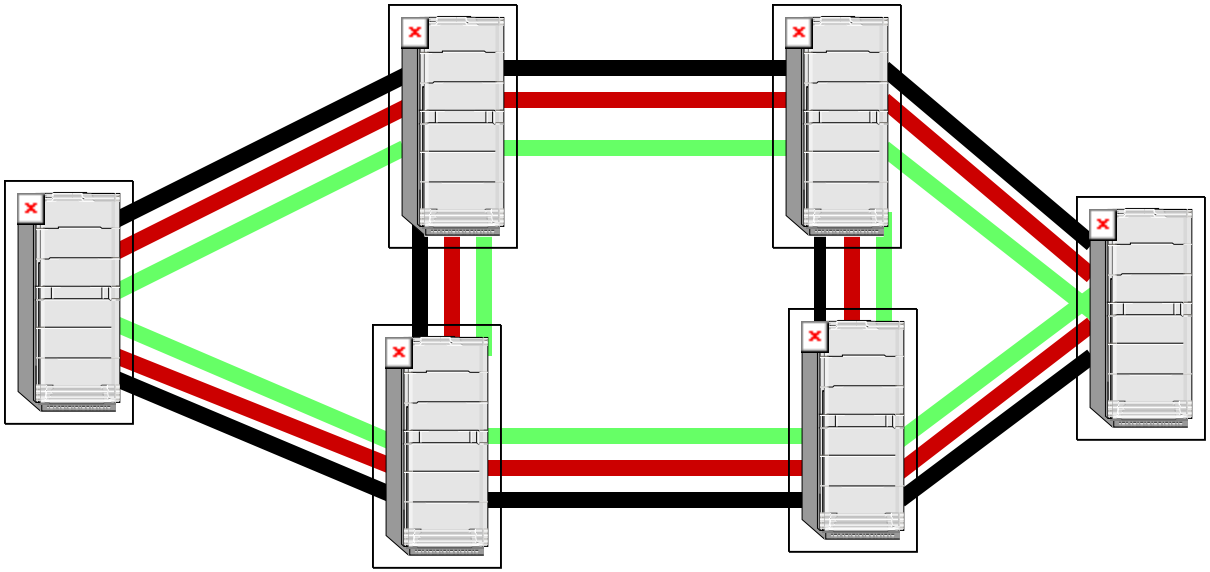
PREEMPTION A.K.A. BUMPING



TOPOLOGY DB FOR BUMPING

LOW PRI

HIGH PRI



Topology Database sees 8 levels of bandwidth, depending on the setup priority of the LSP, a subset of that bandwidth is seen as available.

The highest priority sees all bandwidth used and free at levels lower that it, etc. to the lowest priority which only sees unused bandwidth.

CR-LDP Status

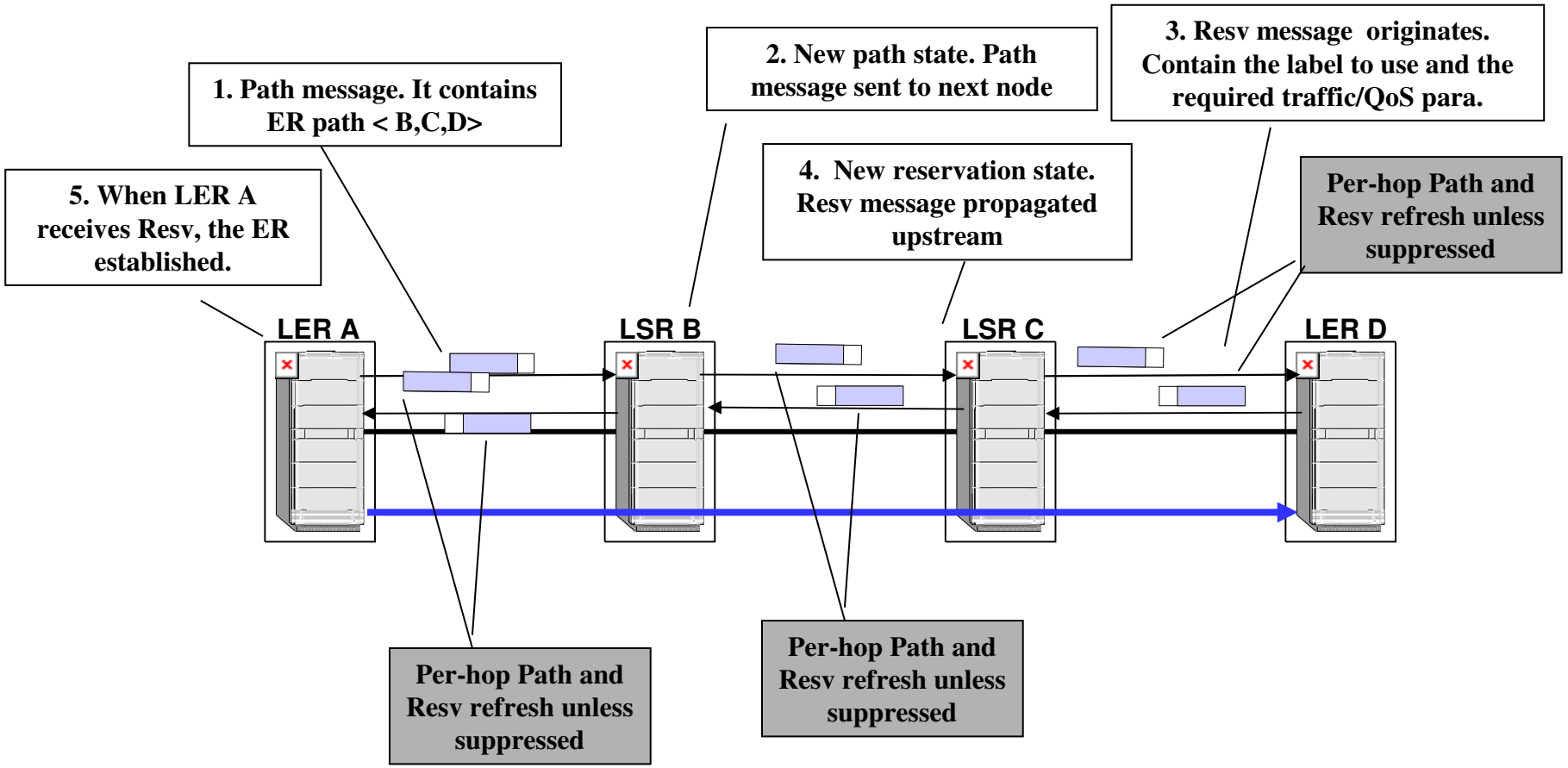
NANOG

- Through last call, going to IESG for RFC.
- Demonstrated Interoperability Nov/98, Sept/99 Nortel, Ericson, Ficom, H&J … 7 vendors.
- Source code for these PDUs publicly available: www.NortelNetworks.com/mpls
- LINUX implementation available publicly U of W.
- The ITU (SG13) has recommended CR-LDP for Traffic engineered IP on ATM in public networks by unanimous vote in Geneva Sept/99.

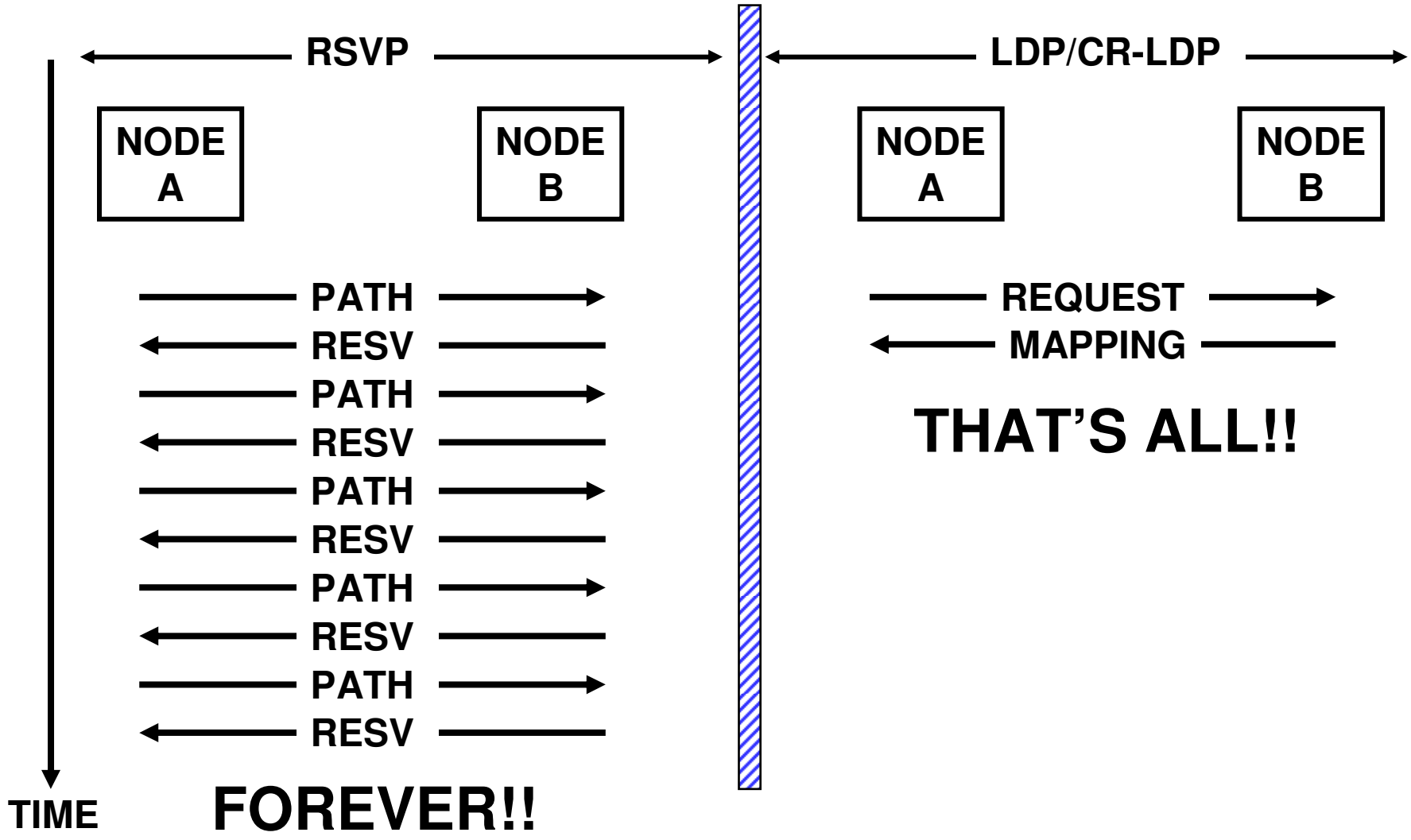
Label Distribution Protocols

- Overview of Hop-by-hop & Explicit
- Label Distribution Protocol (LDP)
- Constraint-based Routing LDP (CR-LDP)
- **Extensions to RSVP**

ER-LSP setup using RSVP



THE BASIC DIFFERENCE: RSVP REFRESHES CONTINUALLY!!



Tutorial Outline

NANOG

- Overview
- Label Encapsulations
- Label Distribution Protocols
- **MPLS & ATM**
- Constraint Based Routing with CR-LDP
- Operational Experiences with Similar Protocols
- Summary

MPLS & ATM

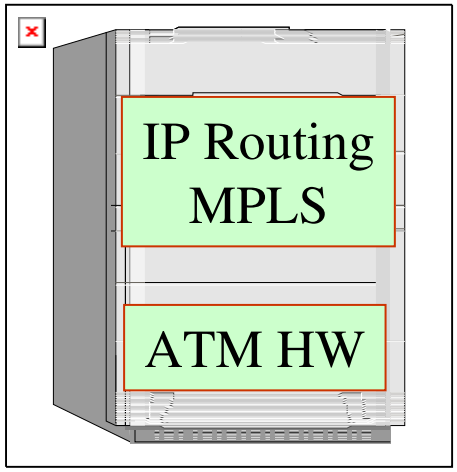
- **Various Modes of Operation**
  - Label-Controlled ATM
  - Tunneling Through ATM
  - Ships in the night with ATM
- **ATM Merge**
  - VC Merge
  - VP Merge

MPLS & ATM

NANOG

Several Models for running MPLS on ATM:

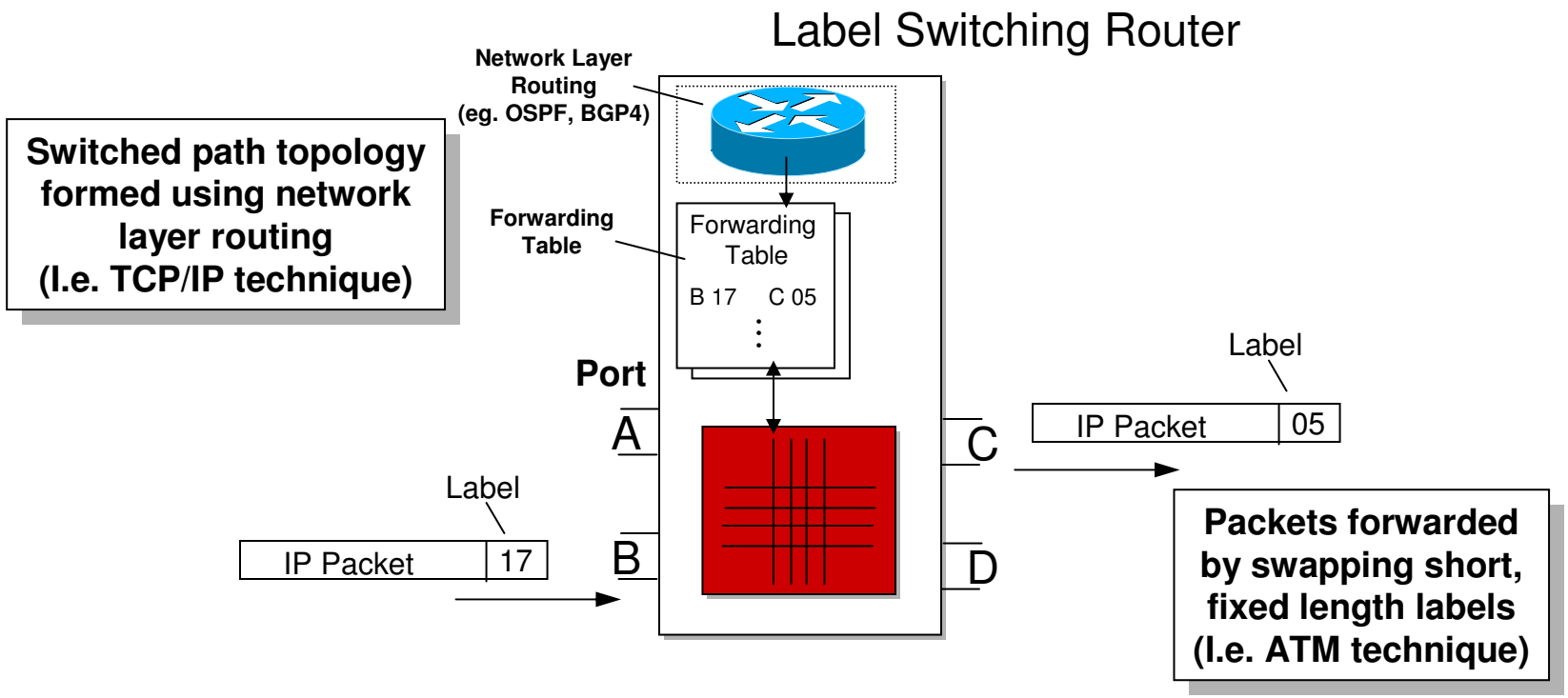
1. Label-Controlled ATM:
   - Use ATM hardware for label switching
   - Replace ATM Forum SW by IP/MPLS



NORTEL NETWORKS™
How the world shares ideas.

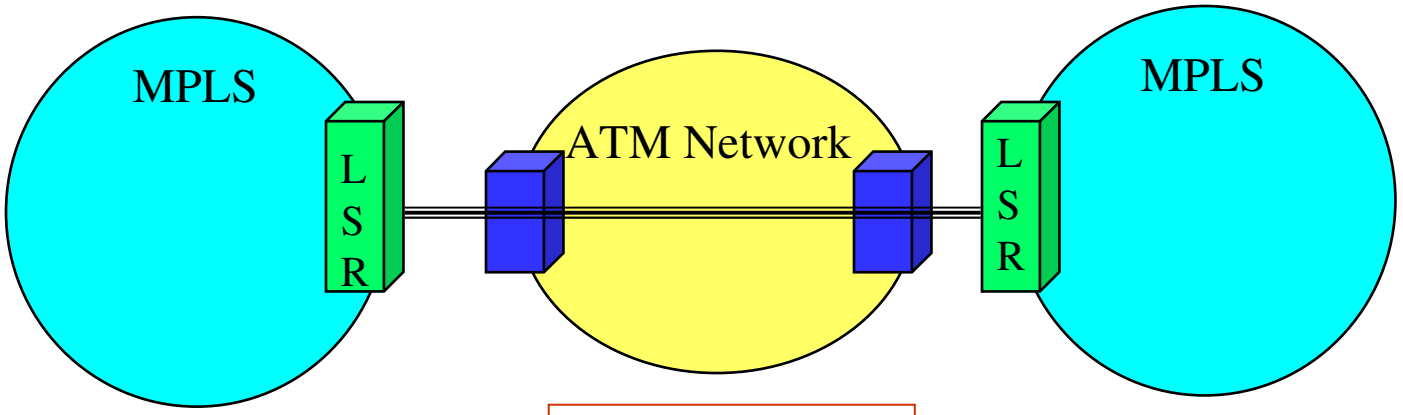
Label-Controlled ATM

- Label switching is used to forward network-layer packets
- It combines the fast, simple forwarding technique of ATM with network layer routing and control of the TCP/IP protocol suite

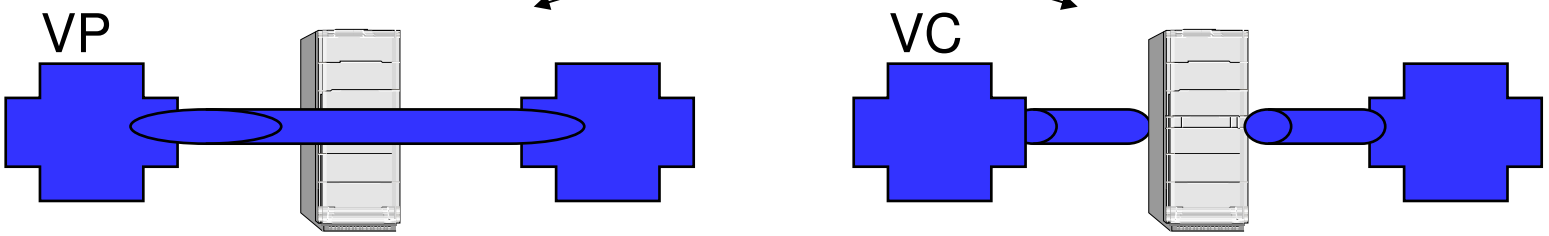


ATM Label Switching is the combination of L3 routing and L2 ATM switching

2. MPLS Over ATM



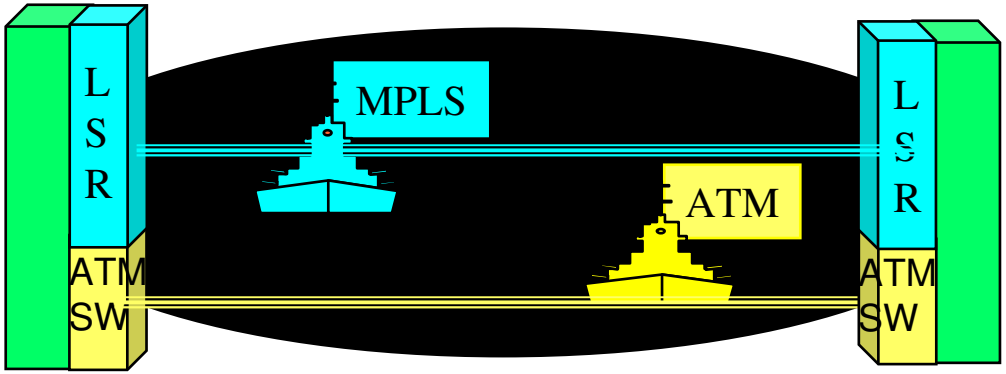
Two Models



Internet Draft:

↑*VCID notification over ATM Link*

3. Ships in the Night

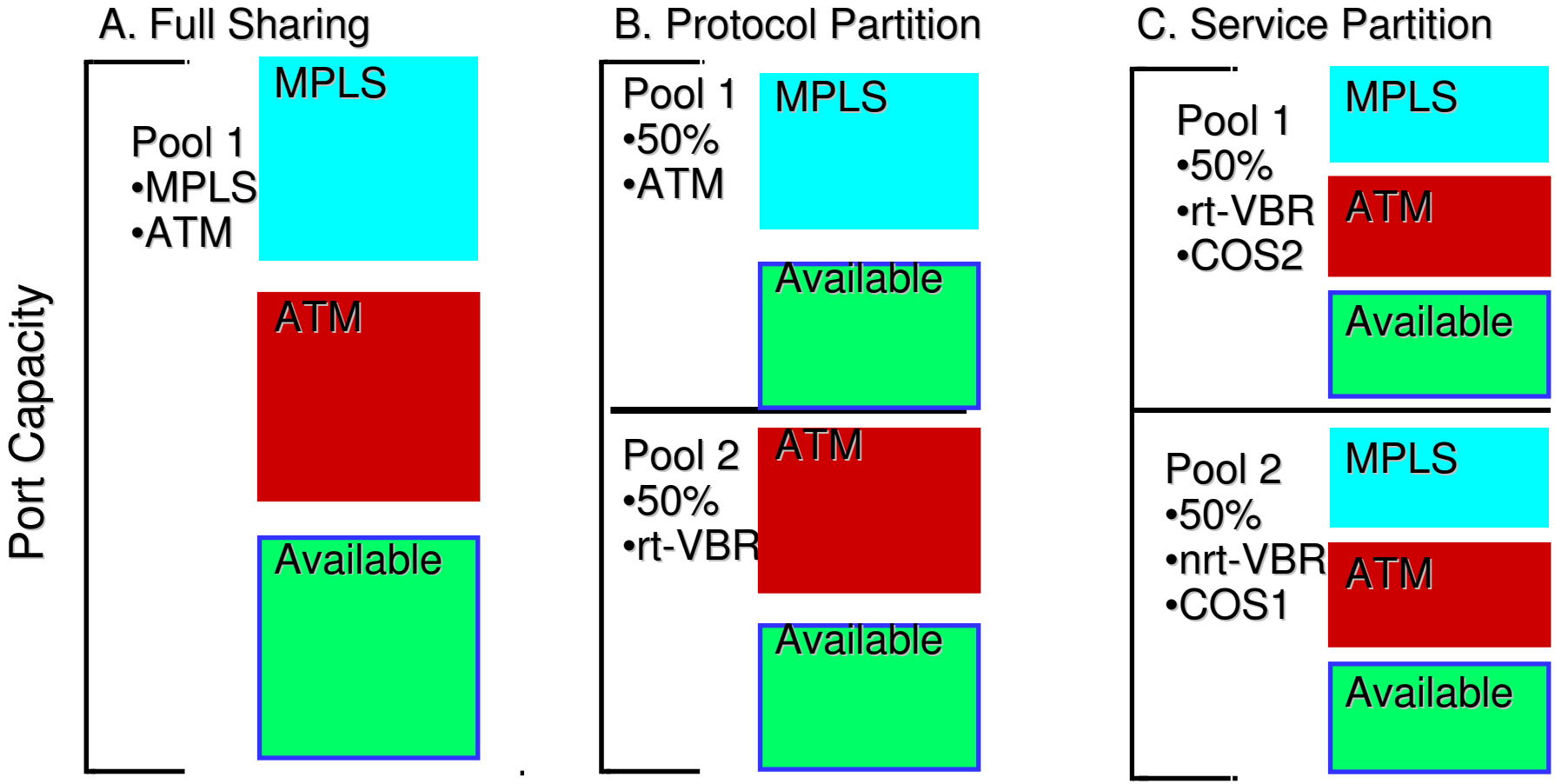


- ATM Forum and MPLS control planes both run on the same hardware but are isolated from each other, i.e. they do not interact.
- This allows a single device to simultaneously operate as both an MPLS LSR and an ATM switch.
- Important for migrating MPLS into an ATM network

Ships in the night Requirements

- **Resource Management**
  - —VPI.VCI Space Partitioning
  - —Traffic management
    - –Bandwidth Reservation
    - –Admission Control
    - –Queuing & Scheduling
    - –Shaping/Policing
  - —Processing Capacity

Bandwidth Management



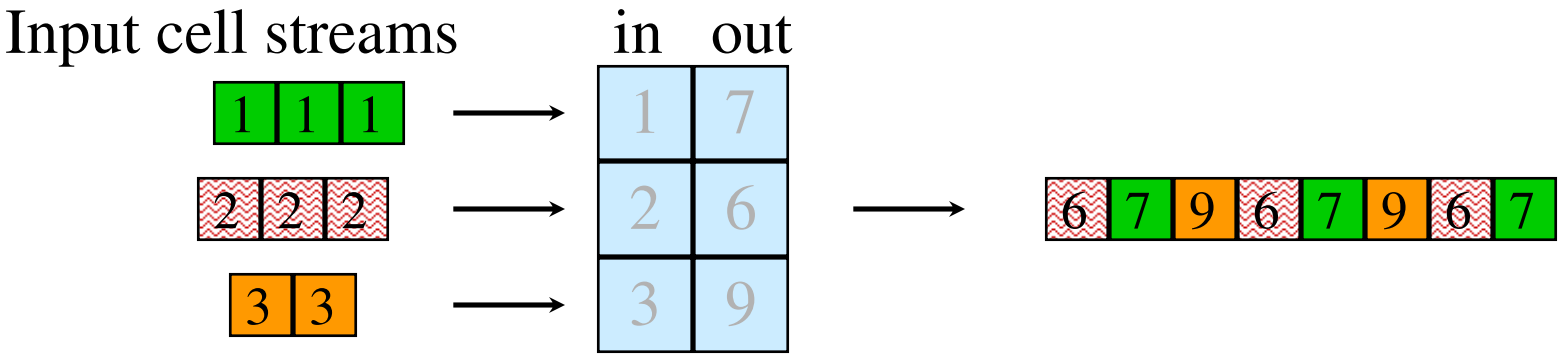
- Bandwidth Guarantees
- Flexibility

ATM Merge

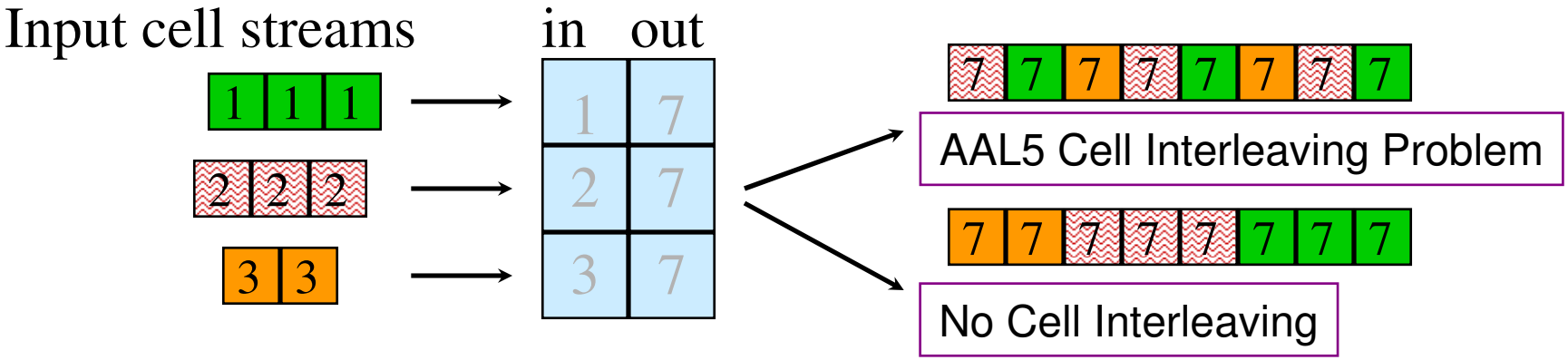
NANOG

- Multipoint-to-point capability
- Motivation
  - —Stream Merge to achieve scalability in MPLS:
    - O(n) VCs with Merge as opposed to $O(n^2)$ for full mesh
    - less labels required
  - —Reduce number of receive VCs on terminals
- Alternatives
  - —Frame-based VC Merge
  - —Cell-based VP Merge

Stream Merge

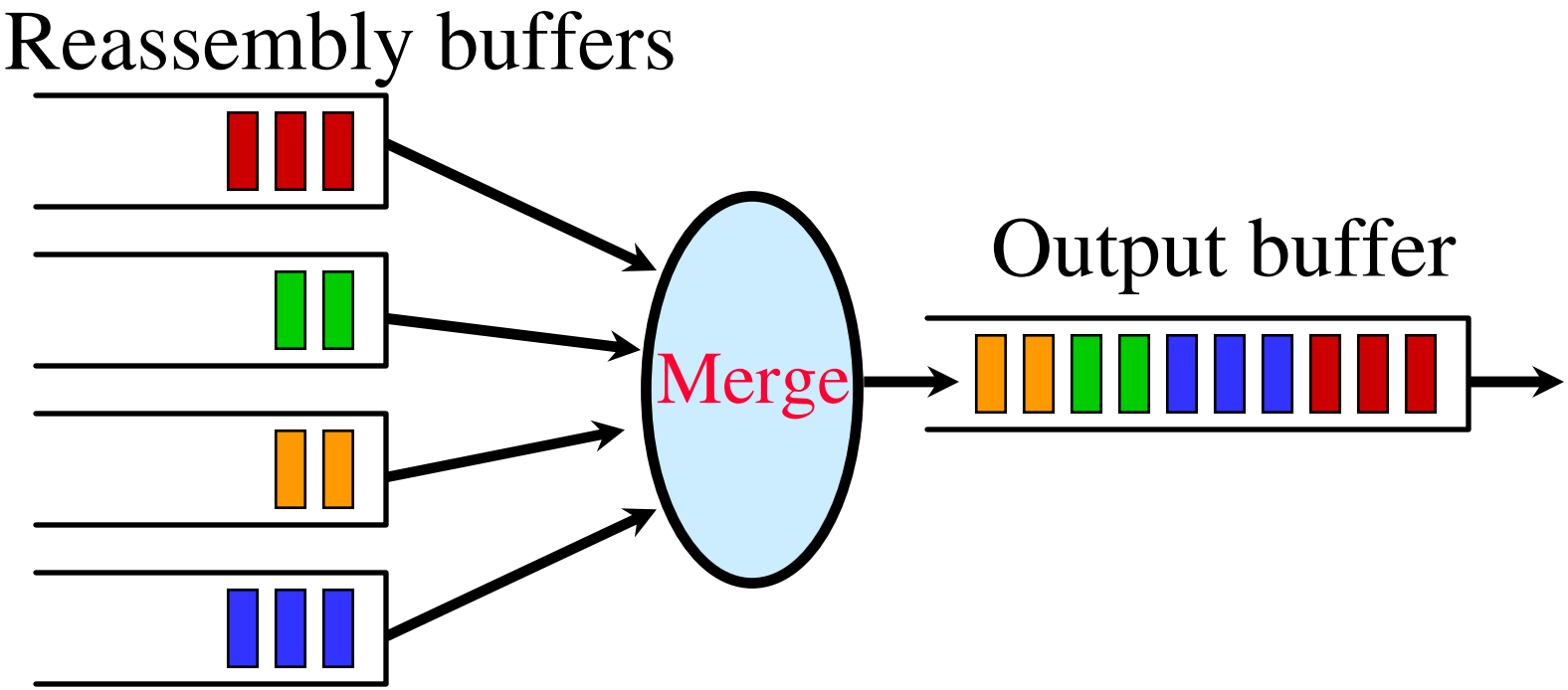


Non-VC merging (*Nin--Nout*)

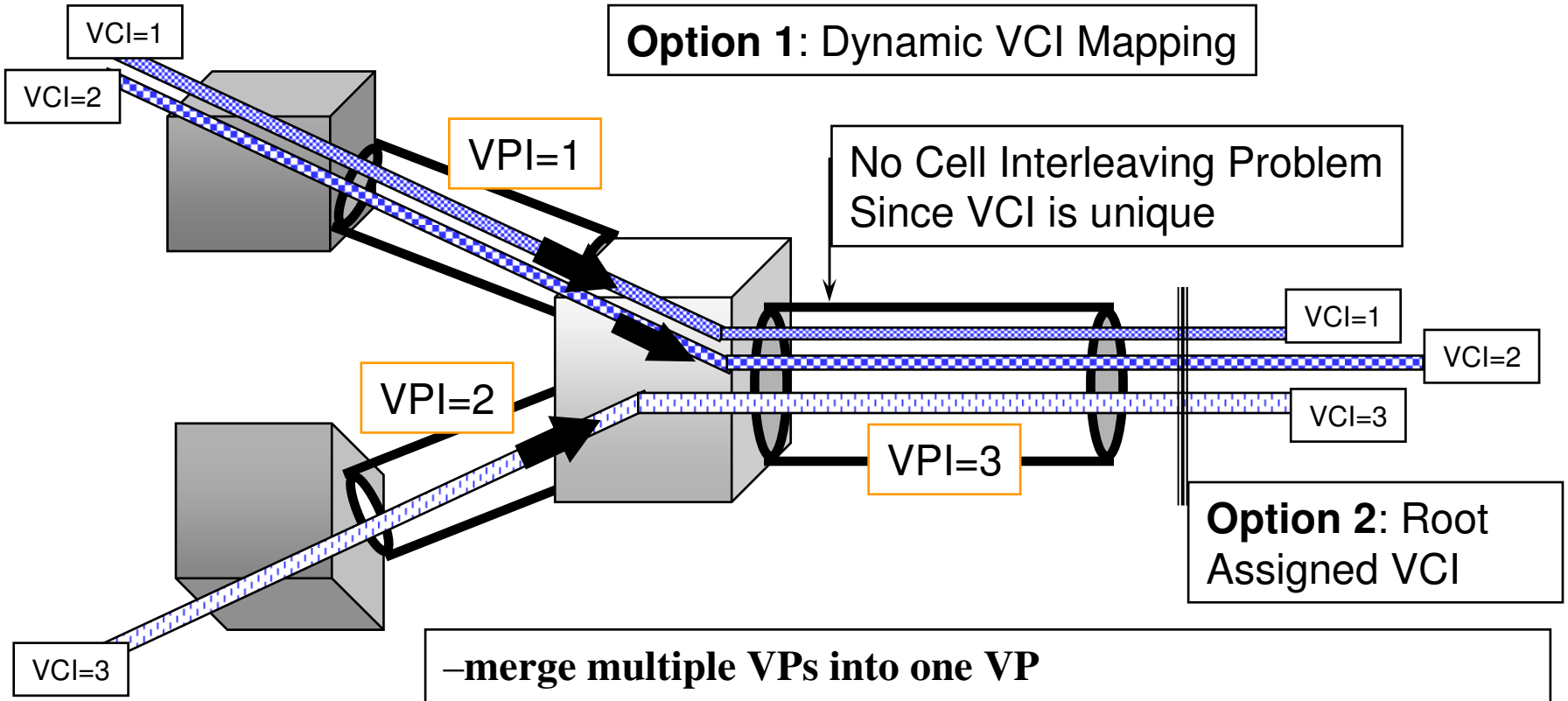


VC merging (Nin-1out)

VC-Merge: Output Module



VP-Merge



–merge multiple VPs into one VP

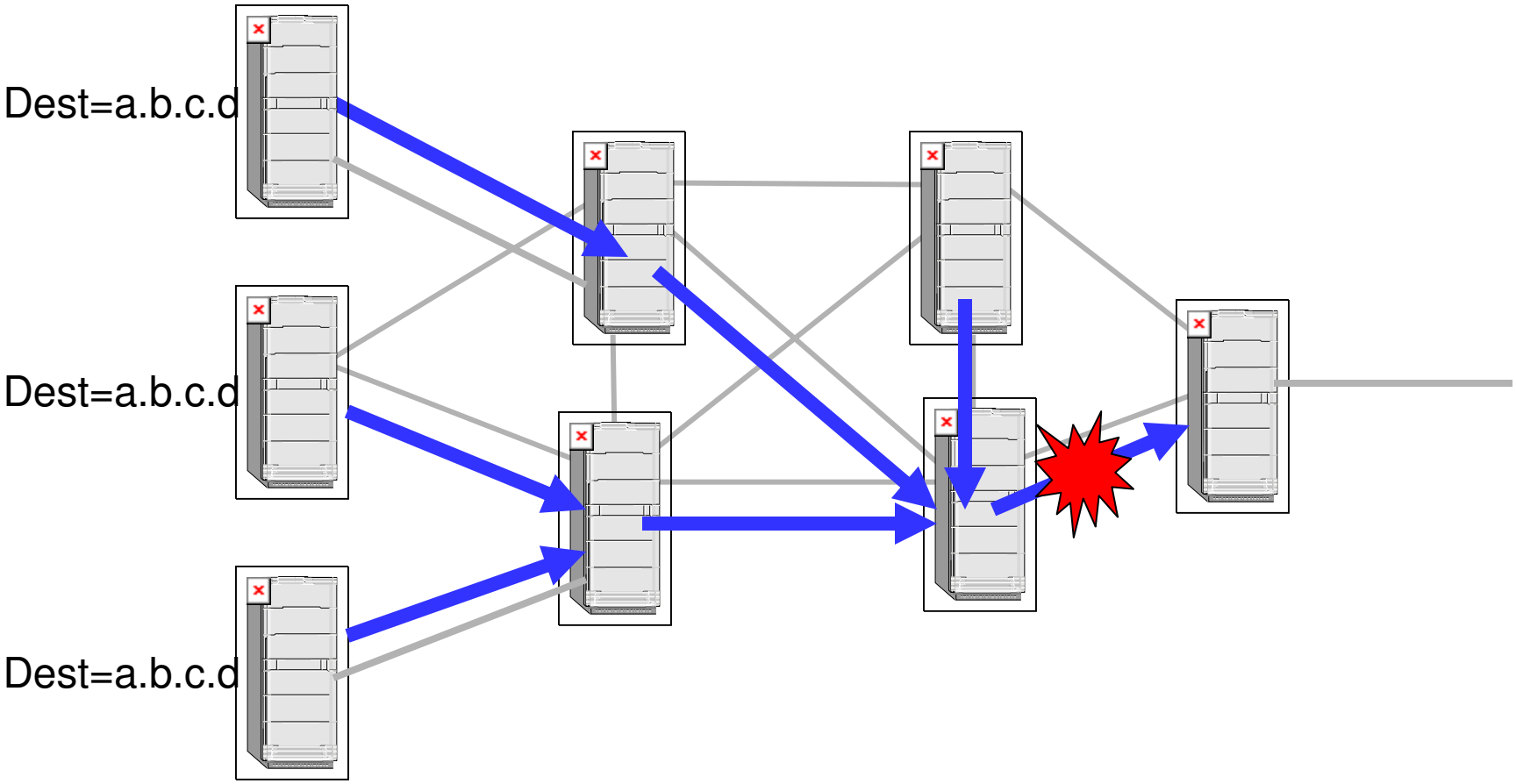
–use separate VCIs within VPs to distinguish frames

–less efficient use of VPI/VCI space, needs support of SVP

Tutorial Outline

- Overview
- Label Encapsulations
- Label Distribution Protocols
- MPLS & ATM
- **Constraint Based Routing with CR-LDP**
- Operational Experiences with Similar Protocols
- Summary

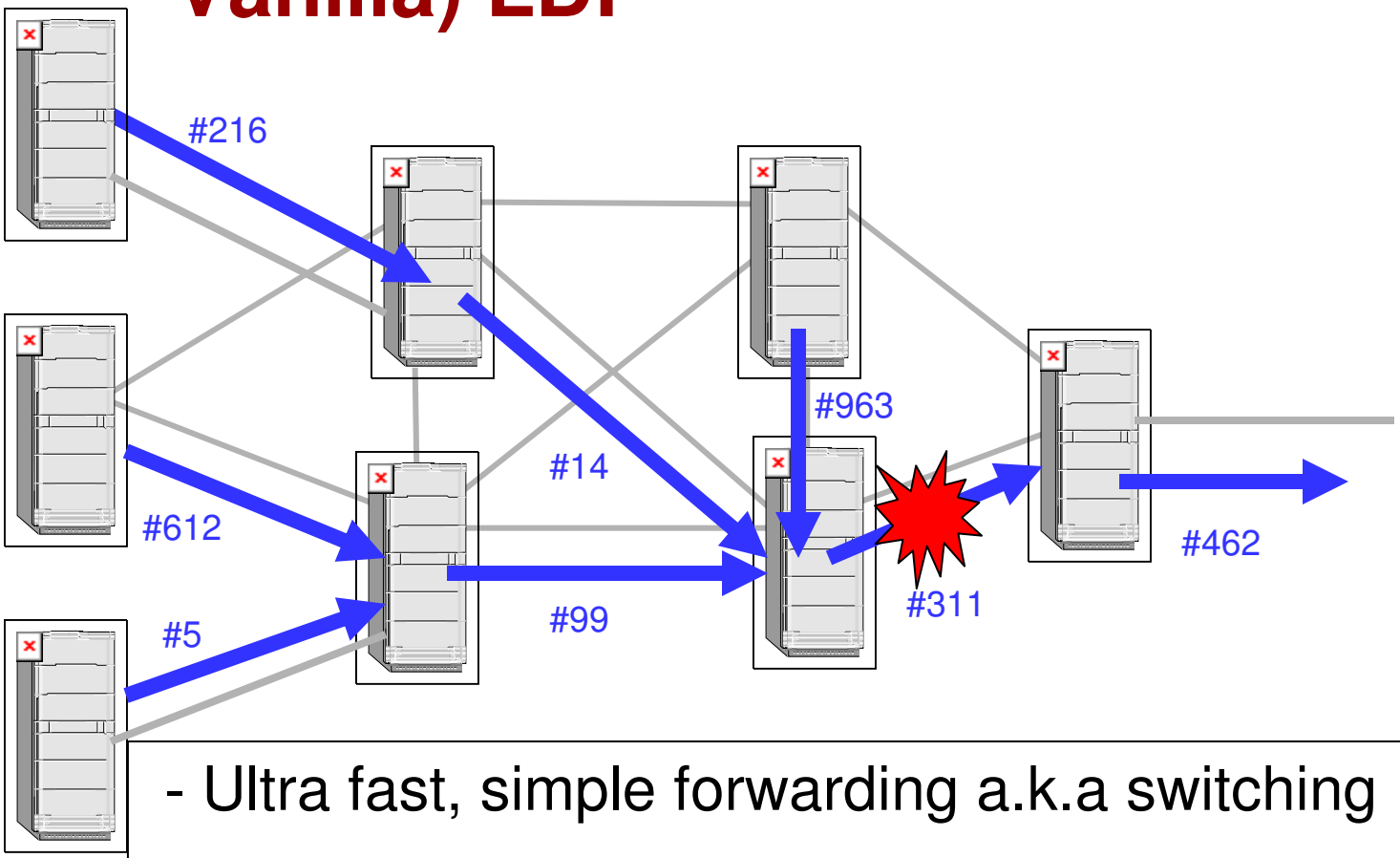
IP FOLLOWS A TREE TO DESTINATION



- IP will over-utilize best paths and under-utilize less good paths.

HOP-BY-HOP(A.K.A Vanilla) LDP

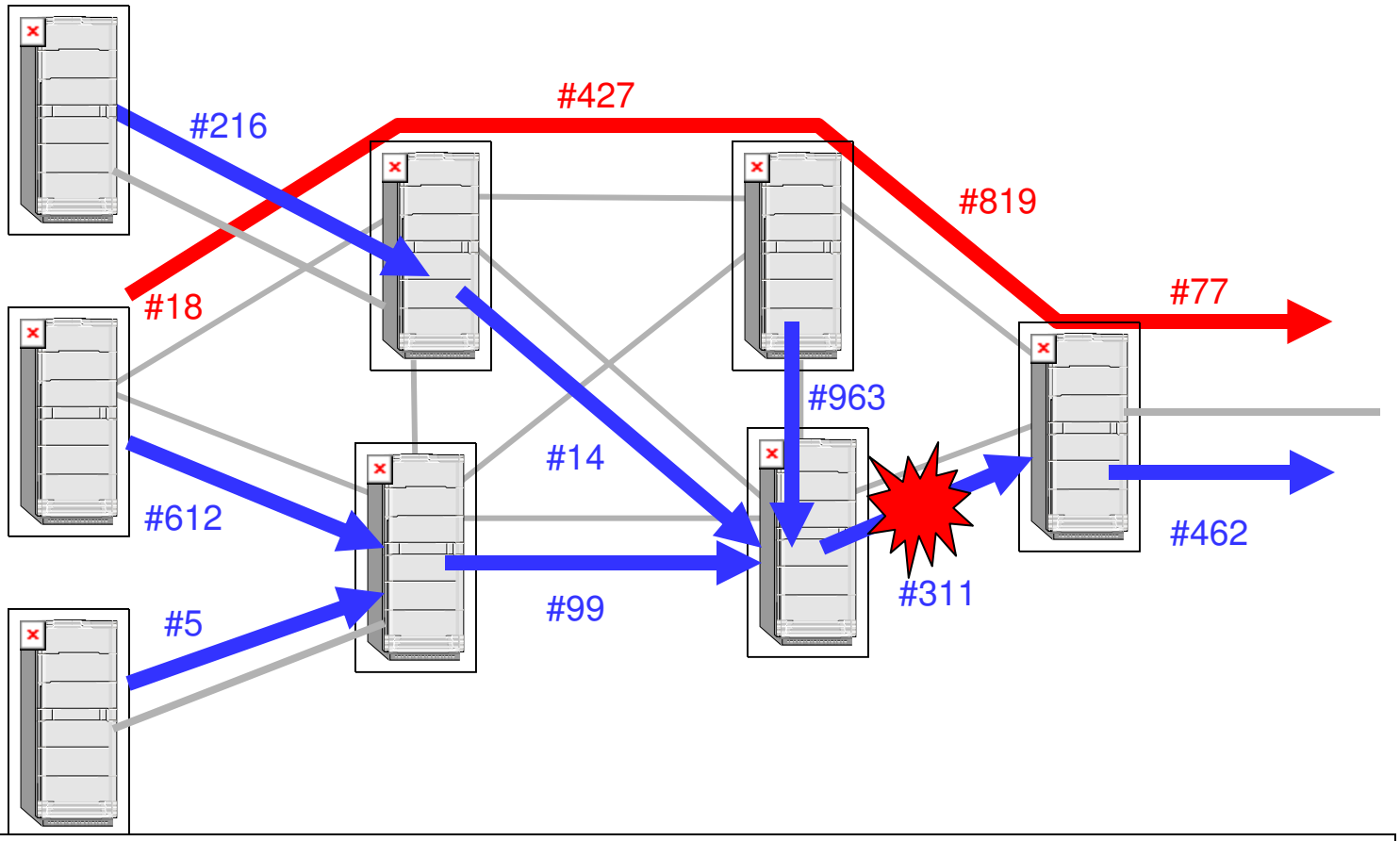
NANOG



- Ultra fast, simple forwarding a.k.a switching
- Follows same route as normal IP datapath
- So like IP, LDP will over-utilize best paths and under-utilize less good paths.

Label Switched Path (Two Types)

NANOG



Two types of Label Switched Paths:

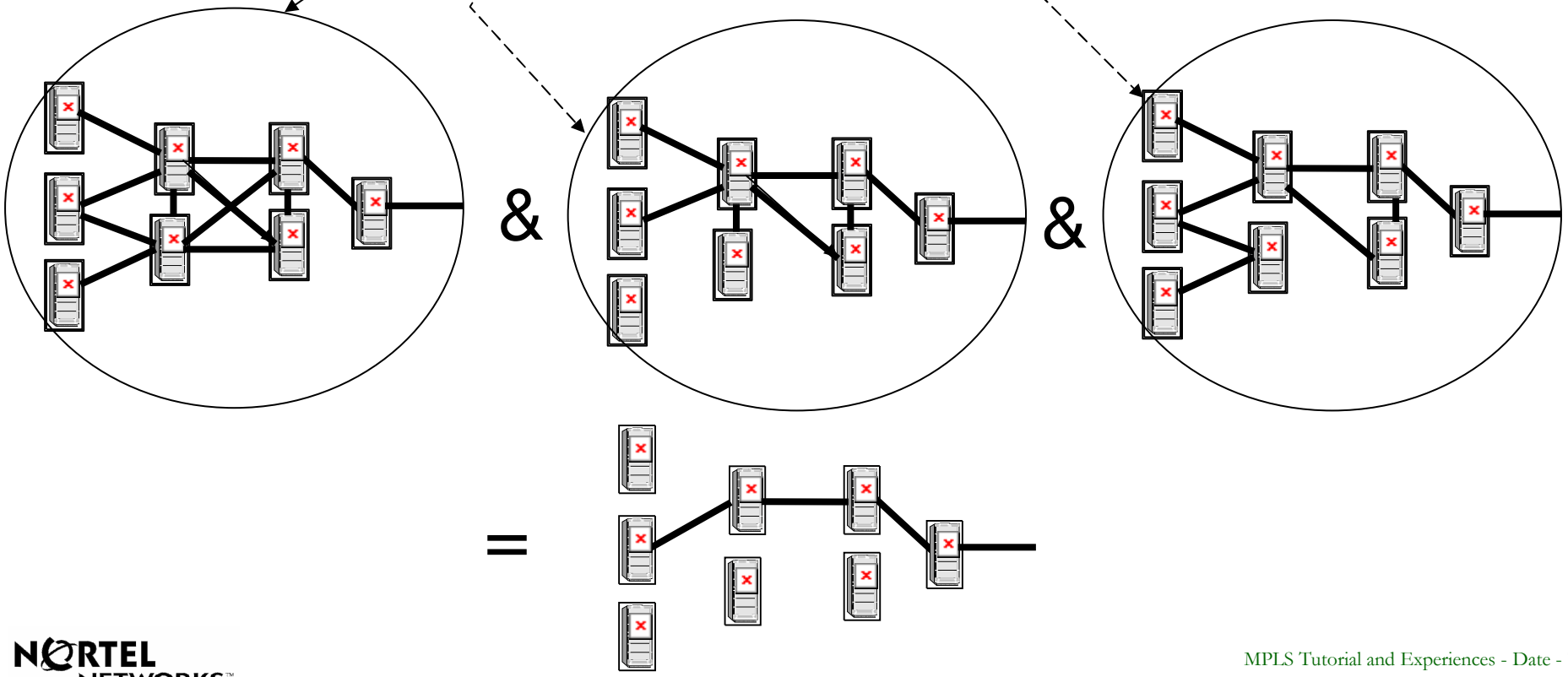
- Hop by hop ("Vanilla" LDP)
- Explicit Routing (LDP+"ER")

NORTEL NETWORKS™
How the world shares ideas.

CR-LDP

NANOG

- CR = “Constraint” based “Routing”
- eg: USE: (links with sufficient resources AND
  (links of type “someColor”) AND
  (links that have delay less than 200 ms)

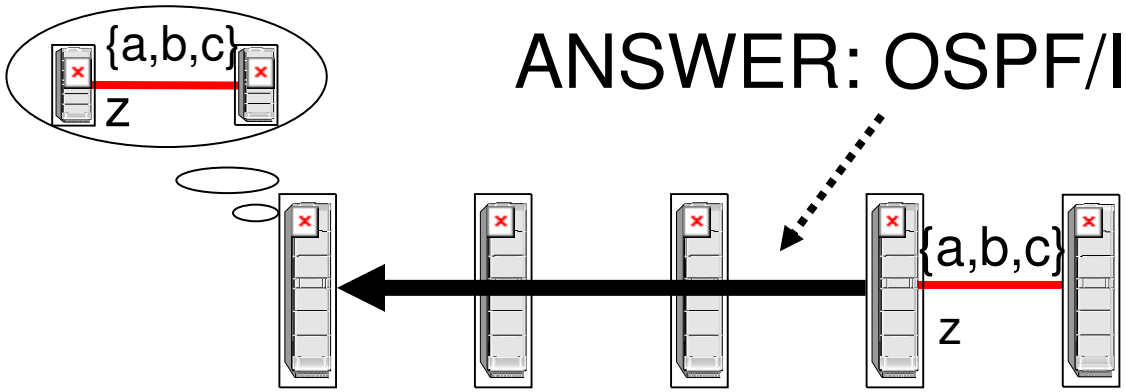


NORTEL NETWORKS™
How the world shares ideas.

Pieces Required for Constraint Based Routing

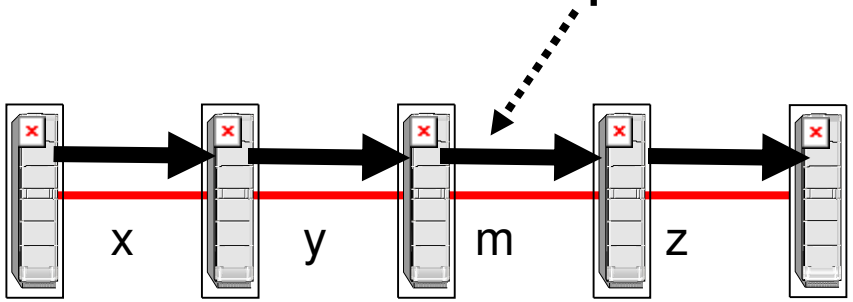
NANOG

1) A topology database that knows about link attributes.

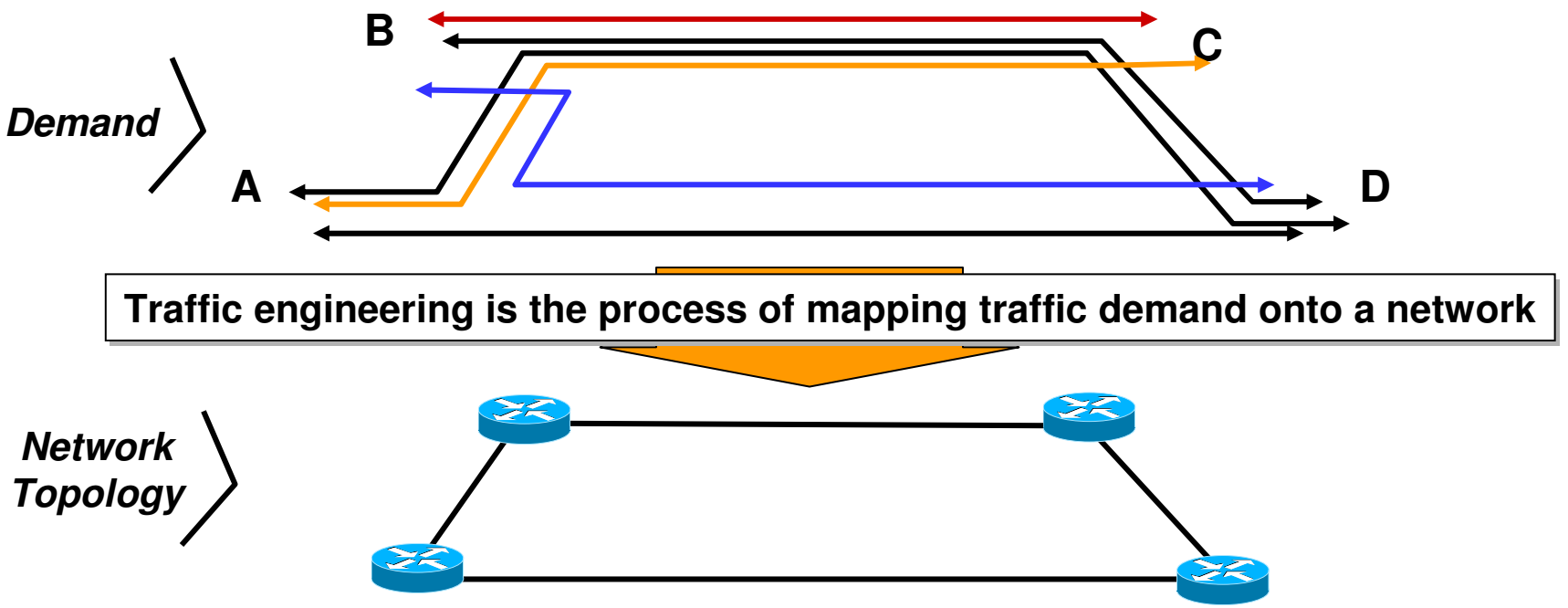


2) A label distribution protocol that goes where it's told.

ANSWER: LDP + Explicit Route{x,y,m,z}



Traffic Engineering



Purpose of traffic engineering:

- Maximize utilization of links and nodes throughout the network
- Engineer links to achieve required delay, grade-of-service
- Spread the network traffic across network links, minimize impact of single failure
- Ensure available spare link capacity for re-routing traffic on failure
- Meet policy requirements imposed by the network operator

Traffic engineering key to optimizing cost/performance

MPLS Traffic Engineering Methods

- MPLS can use the source routing capability to steer traffic on desired path
- Operator may manually configure these in each LSR along the desired path
  - analogous to setting up PVCs in ATM switches
- Ingress LSR may be configured with the path, RSVP used to set up LSP
  - some vendors have extended RSVP for MPLS path set-up
- Ingress LSR may be configured with the path, LDP used to set up LSP
  - many vendors believe RSVP not suited
- Ingress LSR may be configured with one or more LSRs along the desired path, hop-by-hop routing may be used to set up the rest of the path
  - a.k.a loose source routing, less configuration required
- If desired for control, route discovered by hop-by-hop routing can be frozen
  - a.k.a “route pinning”
- In the future, constraint-based routing will offload traffic engineering tasks from the operator to the network itself

Tutorial Outline

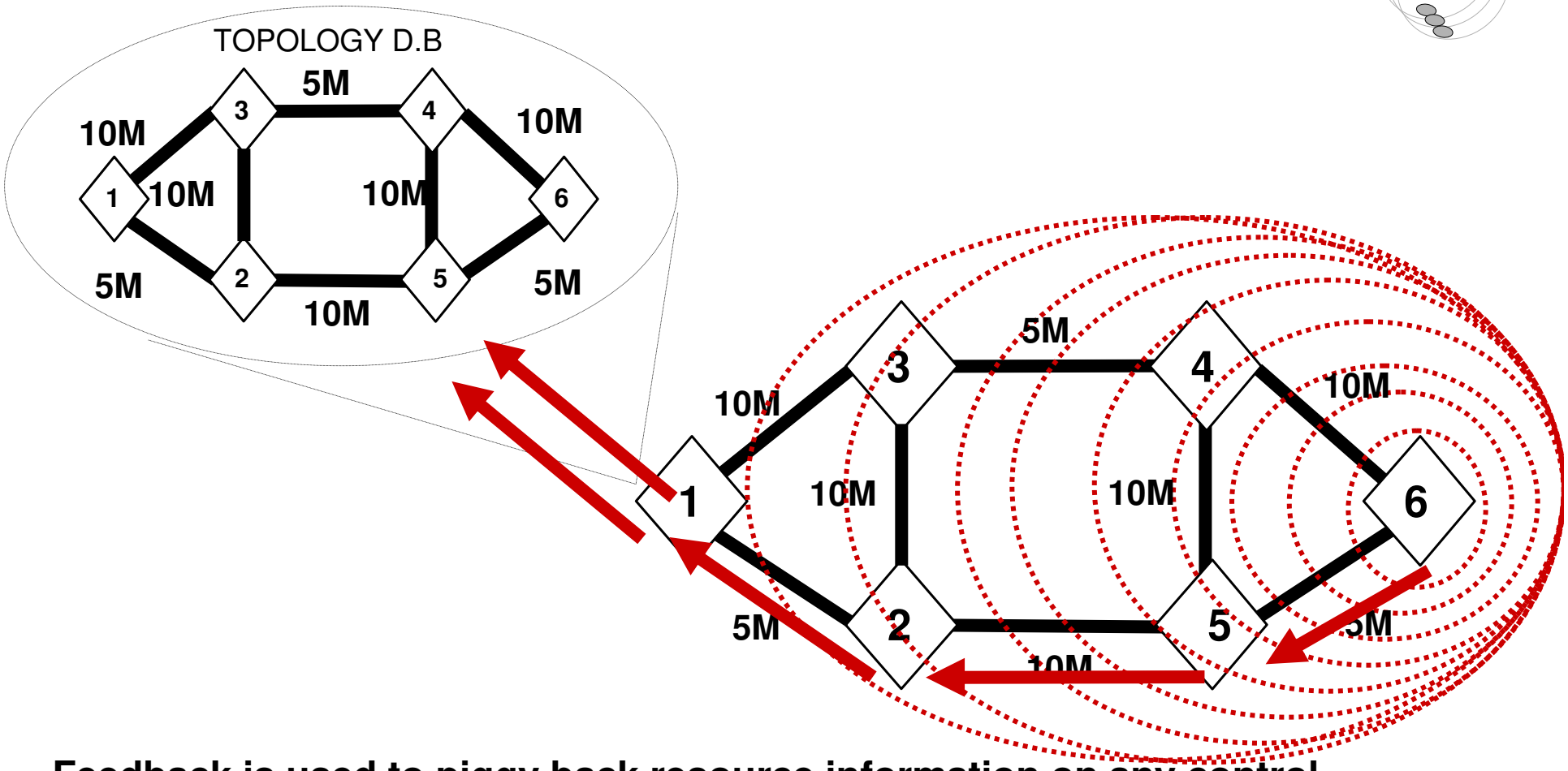
- Overview
- Label Encapsulations
- Label Distribution Protocols
- MPLS & ATM
- Constraint Based Routing with CR-LDP
- **Operational Experiences with Similar Protocols.**
- Summary

OPERATIONAL EXPERIENCES WITH SIMILAR PROTOCOLS (PORS)

- Feedback required to get acceptable blocking and improved rerouting times/accuracy and bigger flat networks.
- Load Spreading requires Preemption be supported together.
- Optimization is required and must be done as a hot swap.
- Region to Region routing is possible with local segment optimization/rerouting.

RESOURCE FEEDBACK

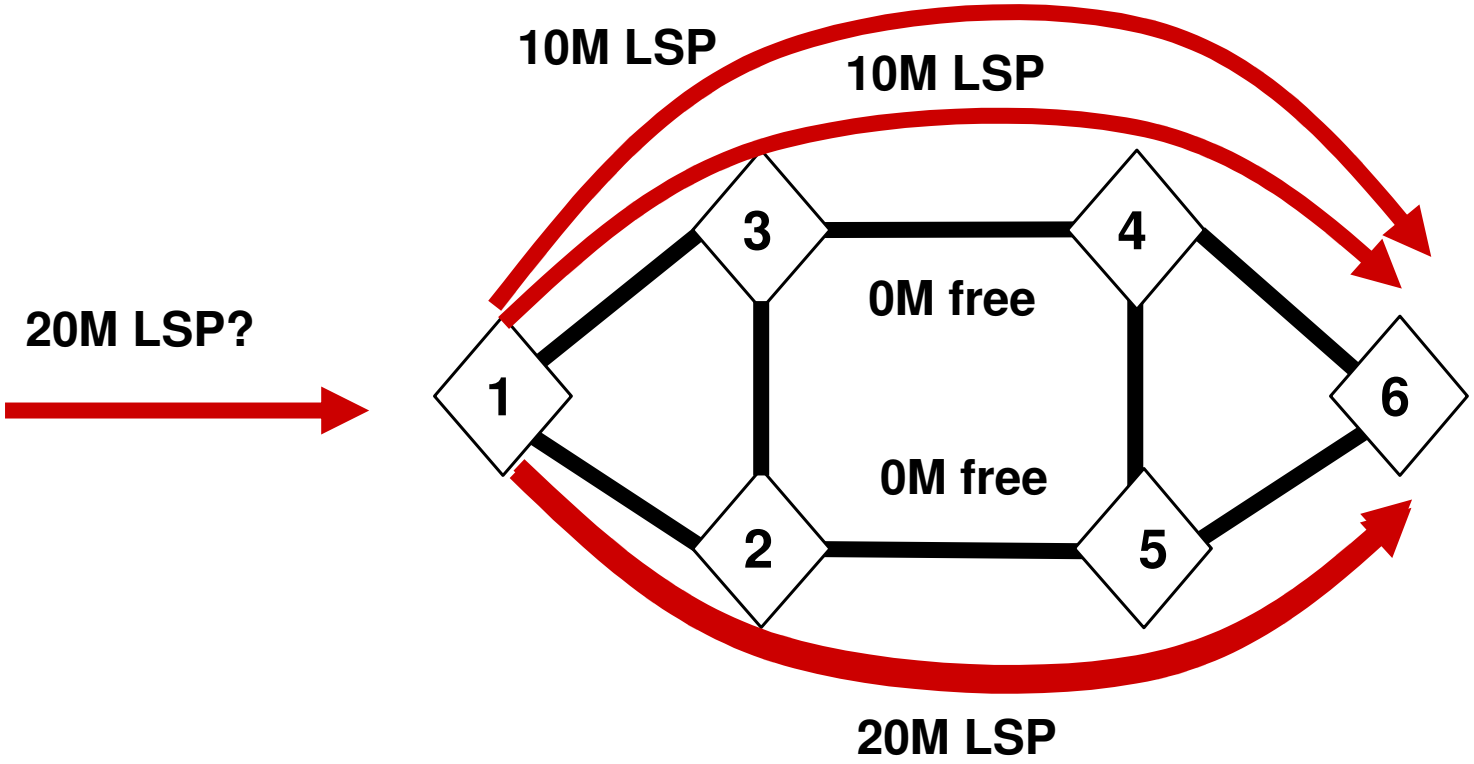
NANOG



Feedback is used to piggy back resource information on any control messages such as setup, release, notification. This reduces the time required for the database to synchronize and allows rerouting BEFORE the floods arrive. This decreases blocking time, reduces flood intervals and allows larger flat topologies.

NORTEL NETWORKS™
How the world shares ideas.

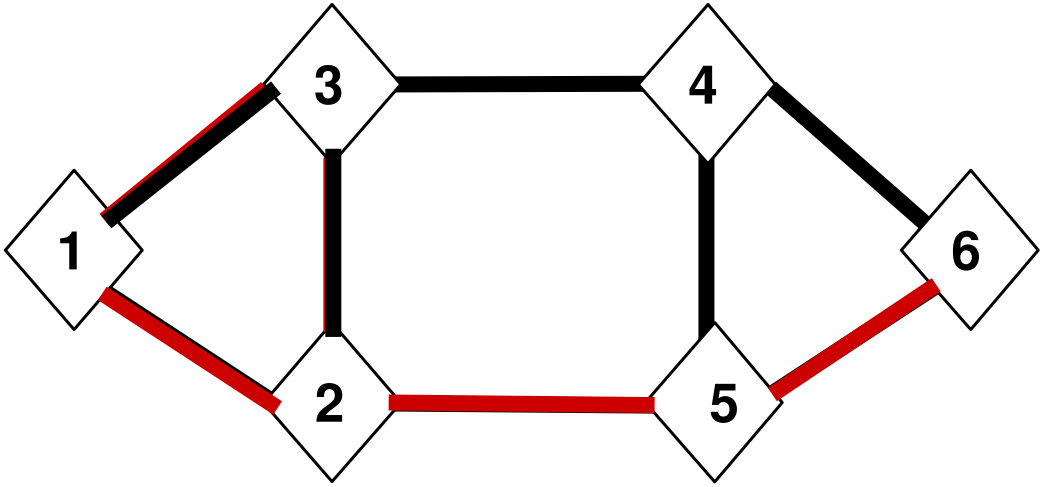
LOAD SPREADING REQUIRES PREMPTION



If you spread load you will leave lots of small bandwidth holes which individually may not be enough to satisfy new requests but taken together would be able. Therefore if you do spread load you need a way to move that load around to free up larger holes of bandwidth, I.e. you need preemption if you do spreading.

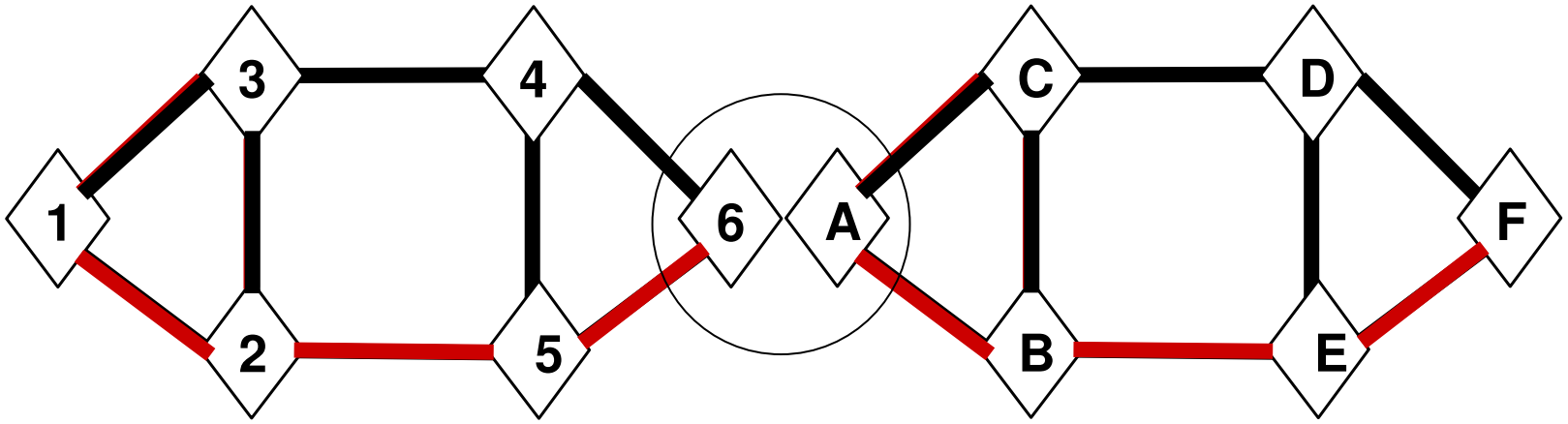
HOT OPTIMIZATION IS REQUIRED

NANOG



Just as a connectionless network will react to the discovery of a better route by using it, so should a path oriented routing system. An MPLS LSP must detect the presence of a better route and switch to it with the minimum of loss. This means it must do it hot, I.e. establish the new LSP, then move traffic to it. It must also do this without double booking bandwidth on common sub segments.

REGION TO REGION WITH LOCAL OPTIMIZATIONS/REPAIR



It is possible to do optimizations and repair within a flat topology region. This means that the gateway remains fixed but that segments between the gateways that cross an AS can move around independently of each other.

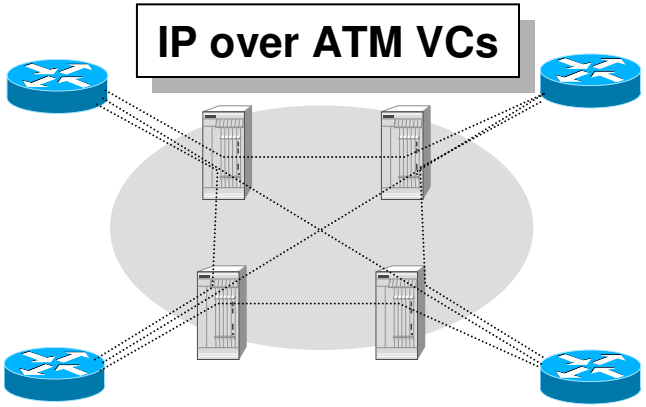
Tutorial Outline

- Overview
- Label Encapsulations
- Label Distribution Protocols
- MPLS & ATM
- Constraint Based Routing with CR-LDP
- Operational Experiences with Similar Protocols.
- Summary

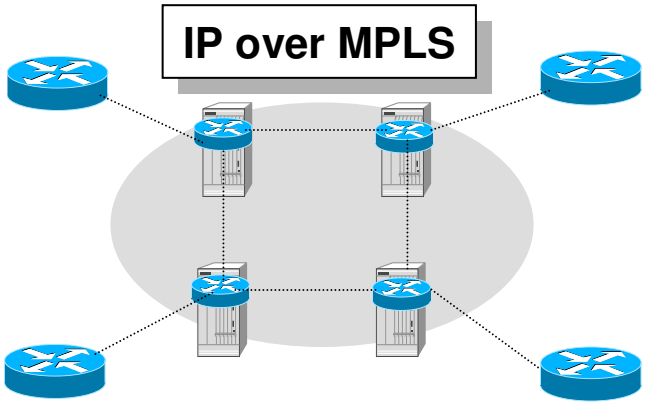
Summary of Motivations for MPLS (not just fast forwarding)

- Simplified forwarding based on exact match of fixed length label
  - initial drive for MPLS was based on existence of cheap, fast ATM switches
- Separation of routing and forwarding in IP networks
  - facilitates evolution of routing techniques by fixing the forwarding method
  - new routing functionality can be deployed without changing the forwarding techniques of every router in the Internet
- Facilitates the integration of ATM and IP
  - allows carriers to leverage their large investment of ATM equipment
  - eliminates the adjacency problem of VC-mesh over ATM
- Enables the use of explicit routing/source routing in IP networks
  - can be easily used for such things as traffic management, QoS routing
- Promotes the partitioning of functionality within the network
  - move granular processing of packets to edge; restrict core to packet forwarding
  - assists in maintaining scalability of IP protocols in large networks
- Improved routing scalability through stacking of labels
  - removes the need for full routing tables from interior routers in transit domain; only routes to border routers are required
- Applicability to both cell and packet link-layers
  - can be deployed on both cell (eg. ATM) and packet (eg. FR, Ethernet) media
  - common management and techniques simplifies engineering

IP and ATM Integration



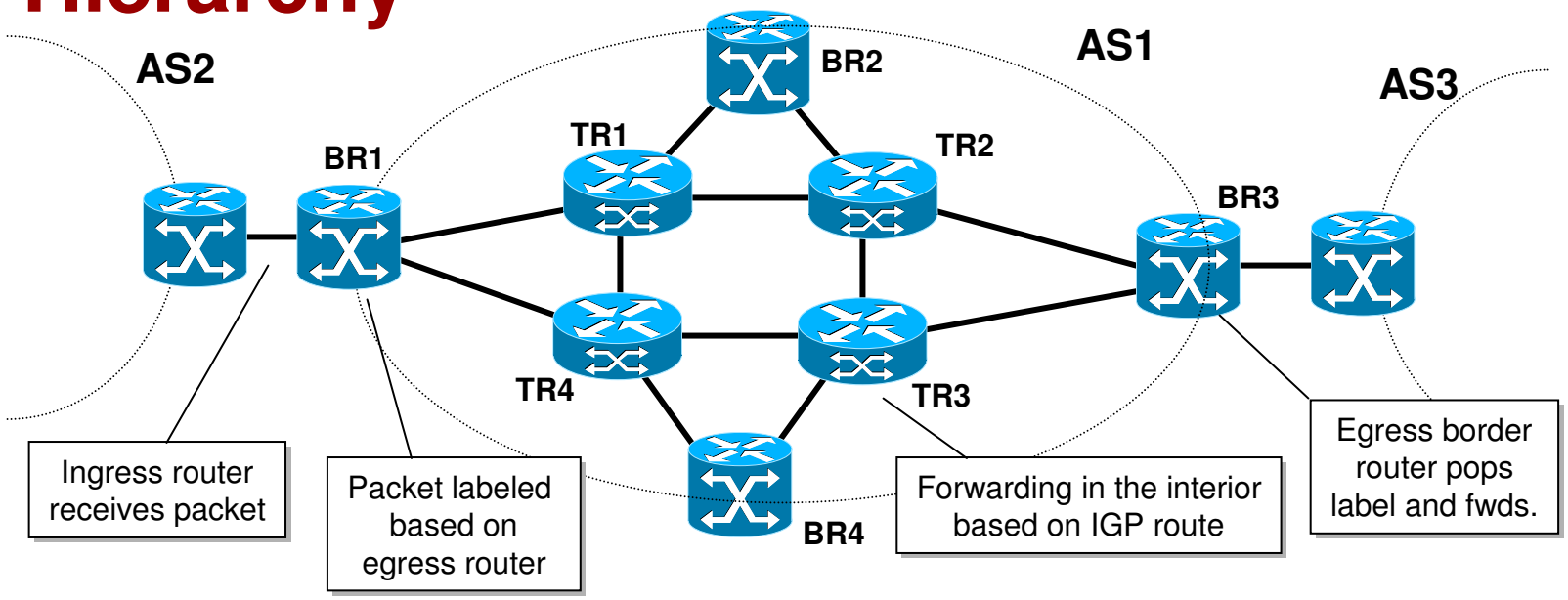
- ATM cloud invisible to Layer 3 Routing
- Full mesh of VCs within ATM cloud
- Many adjacencies between edge routers
- Topology change generates many route updates
- Routing algorithm made more complex



- ATM network visible to Layer 3 Routing
- Singe adjacency possible with edge router
- Hierachical network design possible
- Reduces route update traffic and power needed to process them

MPLS eliminates the "n-squared" problem of IP over ATM VCs

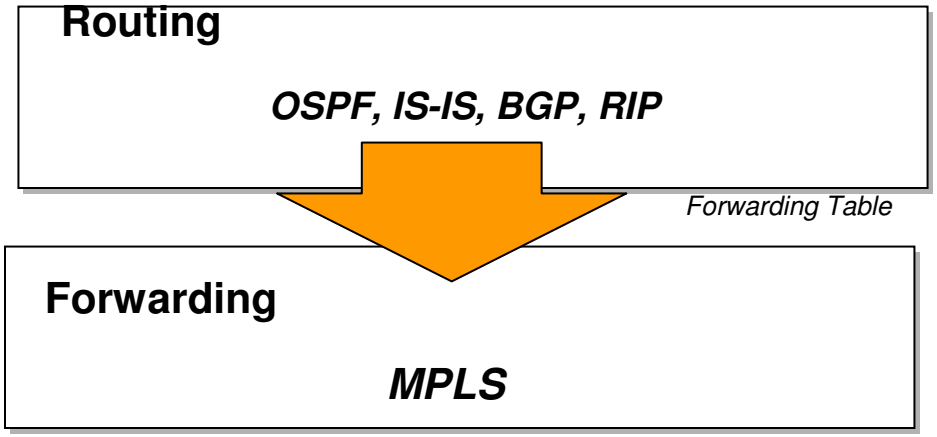
MPLS: Scalability Through Hierarchy



- Border routers BR1-4 run an EGP, providing inter-domain routing
- Interior transit routers TR1-4 run an IGP, providing intra-domain routing
- Normal layer 3 forwarding requires interior routers to carry full routing tables
  - transit router must be able to identify the correct destination ASBR (BR1-4)
- Carrying full routing tables in all routers limits scalability of interior routing
  - slower convergence, larger routing tables, poorer fault isolation
- MPLS enables ingress node to identify egress router, label packet based on interior route
- Interior LSRs would only require enough information to forward packet to egress

MPLS increases scalability by partitioning exterior routing from interior routing

MPLS: Partitioning Routing and Forwarding



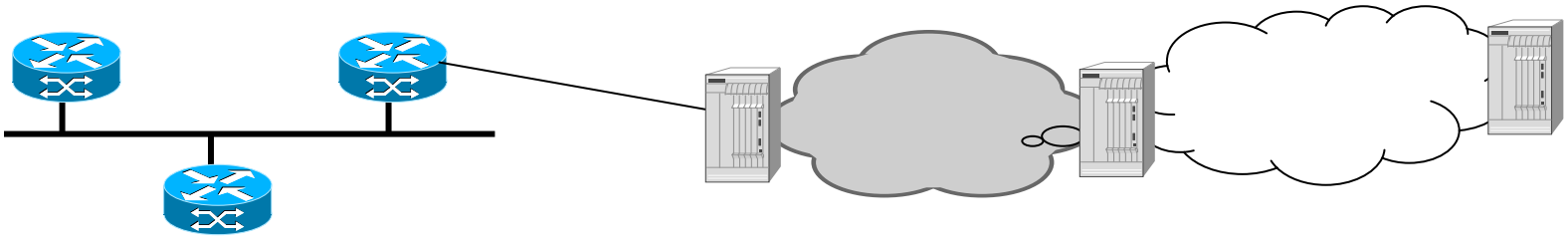
Based on:
Classful Addr. Prefix?
Classless Addr. Prefix?
Multicast Addr.?
Port No.?
ToS Field?

Based on:
Exact Match on Fixed Length Label

- Current network has multiple forwarding paradigms
  - class-ful longest prefix match (Class A,B,C boundaries)
  - classless longest prefix match (variable boundaries)
  - multicast (exact match on source and destination)
  - type-of-service (longest prefix. match on addr. + exact match on ToS)
- As new routing methods change, new route look-up algorithms are required
  - introduction of CIDR
- Next generation routers will be based on hardware for route look-up
  - changes will require new hardware with new algorithm
- MPLS has a consistent algorithm for all types of forwarding; partitions routing/fwding
  - minimizes impact of the introduction of new forwarding methods

MPLS introduces flexibility through consistent forwarding paradigm

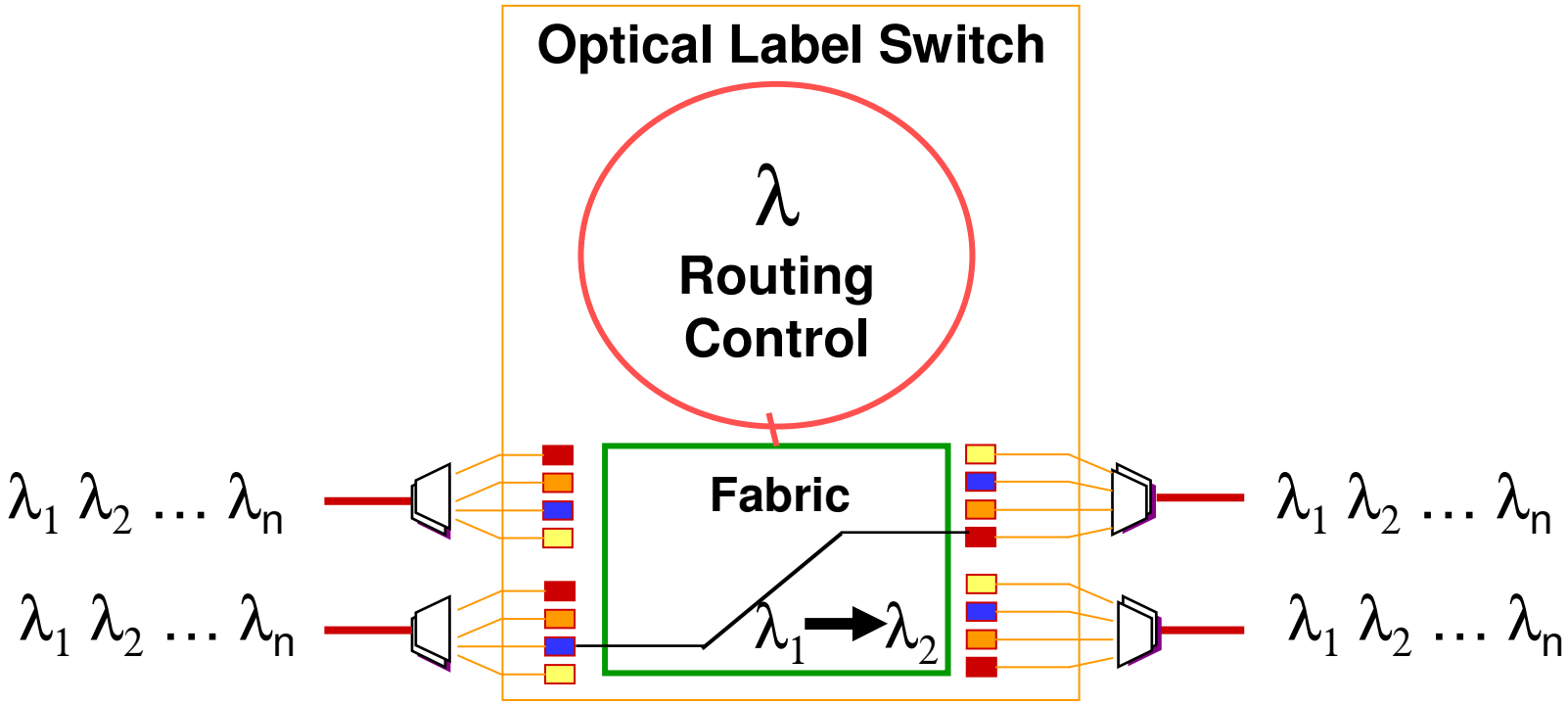
Upper Layer Consistency Across Link Layers



- MPLS is "multiprotocol" below (link layer) as well as above (network layer)
- Provides for consistent operations, engineering across multiple technologies
- Allows operators to leverage existing infrastructure
- Co-existence with other protocols is provided for
  - eg. "Ships in the Night" operation with ATM, muxing over PPP

MPLS positioned as end-to-end forwarding paradigm

PROBABLY THE ONLY OPTION FOR ROUTING AT LIGHT SPEEDS



When we get to true frequency to frequency switching there is no way to route and LDP will be required to setup OSPF routes. CR-LDP will be required to engineer. λ is just another label to distribute. No new protocols required.

Summary

- MPLS is an exciting promising emerging technology.
- Basic functionality (Encapsulation and basic Label Distribution) has been defined by the IETF.
- Traffic engineering based on MPLS/CR-LDP is just round the corner.
- MPLS/LDP/CR-LDP have been recommended by the ITU for IP transport on ATM in public networks.
- Convergence is one step closer ……