



Seminar 2: MPLS Overview & Applications

Tony Bogovic
tjb@research.telcordia.com
(973) 829-4348

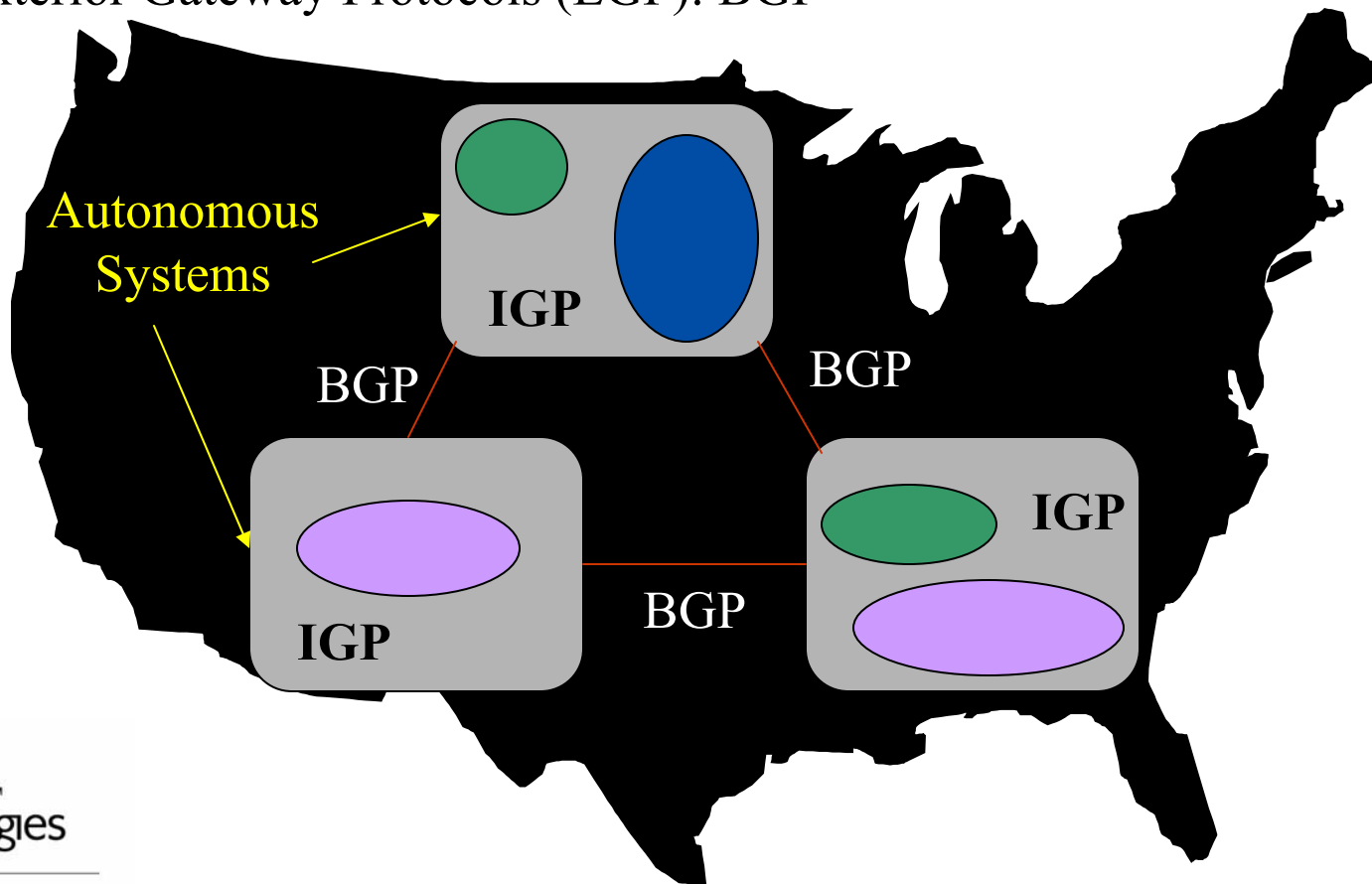
February 25, 2002

Outline

- IP Routing Review
- MPLS
 - motivating factors
 - functionality
 - signaling protocols
 - applications
 - standards
 - LSR Implementations
 - SP Deployments
- Summary

Interior and Exterior Gateway Protocols

- Internet (IP) routing is adaptive and distributed
- Interior Gateway Protocols (IGPs): RIP, OSPF, IS-IS, etc.
- Exterior Gateway Protocols (EGP): BGP



IP Routing Review

- IP routing can be partitioned into two basic components:
 - 1) The control component:
 - is responsible for construction and maintenance of the routing table
 - each IP hop runs its own instance of the routing algorithm
 - the link metrics most IGPs use for deciding what path to send traffic on are either
 - hop count, or
 - administrative weight
 - 2) The forwarding component:
 - forwards packets from input to output based on the information carried in the packet itself and a routing table maintained by a router
 - IP forwarding is done independently at every hop
 - for the most part, forwarding in IP networks is currently based solely on destination address

IP Forwarding

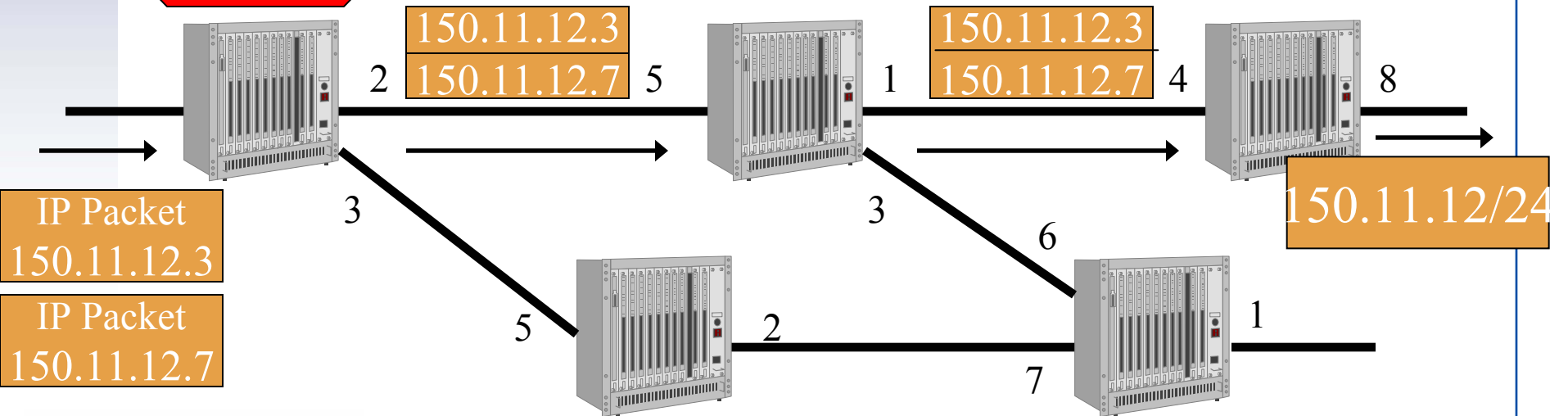
- FEC (Forwarding Equivalence Class)
 - a group of IP packets which are forwarded in the same manner
 - e.g., over the same path with the same forwarding treatment
- IP packets are classified into FECs at each hop in conventional routing
 - in MPLS, only classified once at the ingress
- IP packet forwarding works by
 - assigning a packet to a FEC
 - determining the next-hop of each FEC

IP Forwarding continued...

Address Prefix & mask	Next-hop	I/F
150.11.12/24	150.11.12.1	2
150.11.13/24	150.11.13.1	3
...

Address Prefix & mask	Next-hop	I/F
150.11.12/24	150.11.12.2	1
150.11.13/24	150.11.13.2	3
...

Address Prefix & mask	Next-hop	I/F
150.11.12/24	150.11.12.5	8
150.11.13/24	150.11.13.4	3
...



- IP packets with different destination addresses but same FEC are forwarded along the same route, thus same output interface and same next-hop

Traditional IP

Three Important Questions...

- 1 Q: What field in the IP header is used to make the forwarding decision?
 - A: The destination IP address
- 2 Q: When this field is used as an index into the Routing table, what is looked up?
 - A: The next hop IP address
- 3 Q: What other vital piece of information does the Routing Table contain?
 - A: The output interface

Multiprotocol Label Switching



Formerly Bellcore...
Performance from Experience

Initial MPLS Motivating Factors

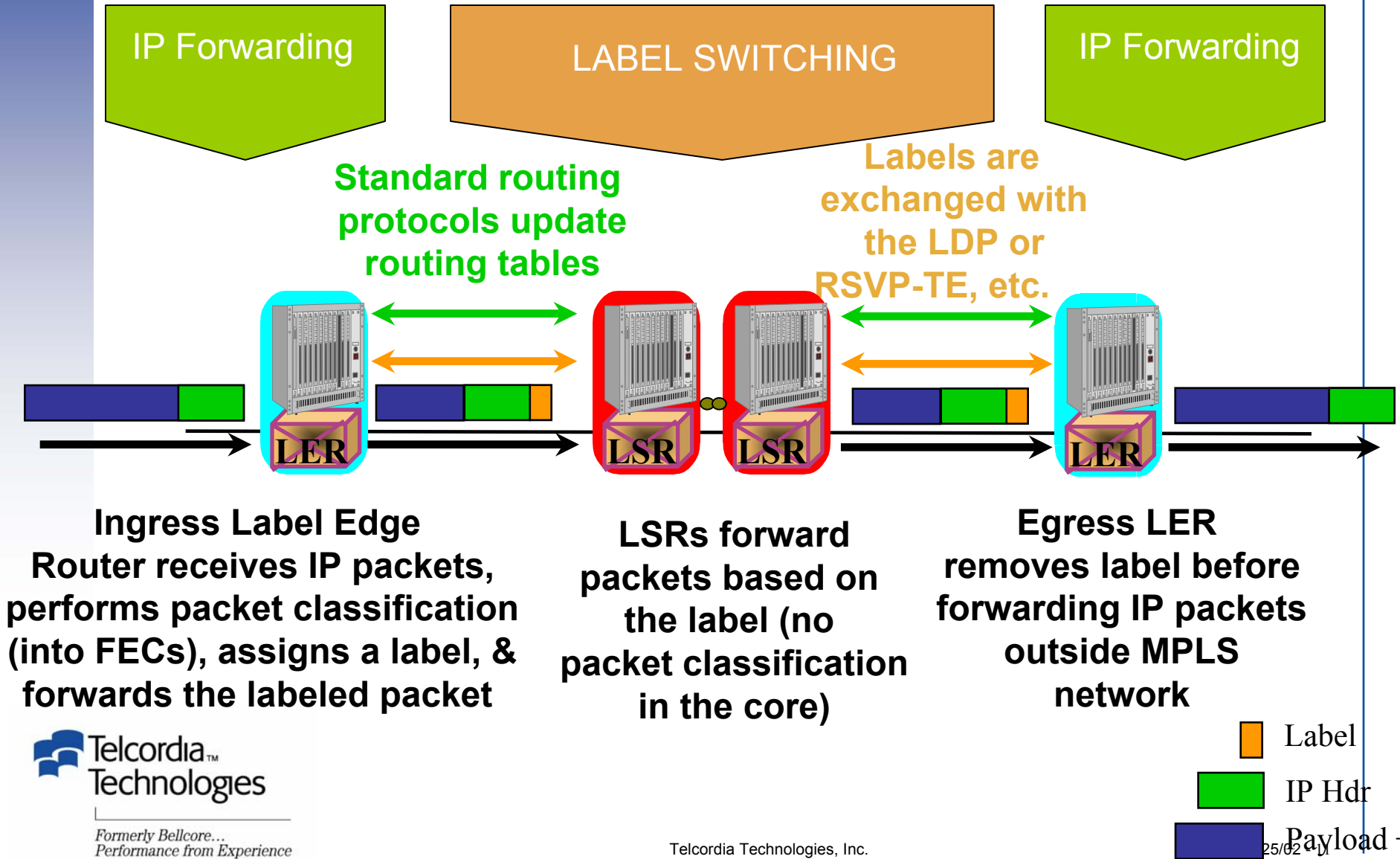
- Scalability
 - due to the growth in the number of Internet users and user bandwidth requirements, higher performance equipment was needed
- Extend routing capabilities of the Internet
 - routing functionality was difficult to evolve due, in part, to the close coupling between control and forwarding in routers
 - e.g., difficulty in adapting existing code for Classless InterDomain Routing (CIDR)
- Price and Performance
 - ATM switches tended to have greater port densities and greater throughputs at lower cost than IP routers, but less so today
- IP over ATM integration
 - due to price, performance and traffic mgmt reasons, ATM is being used in the Internet backbone for forwarding IP traffic but has scaling issues
- In 1997, traffic engineering became *the* motivating factor for MPLS

Multiprotocol Label Switching

What is it?

- MPLS is a combination of:
 - A forwarding mechanism based on label switching
 - i.e., MPLS forwards IP packets based on a label swapping paradigm
 - Label Switched Path (LSP) set-up protocols such as LDP, CR-LDP, and RSVP-TE
 - mapping definitions onto Layer 2 technologies such as ATM, Frame Relay, Ethernet, and PPP
 - MPLS integrates IP and link layer technologies
- MPLS brings connection-oriented functionality into a connectionless IP paradigm
- Terminology:
 - Label: a short, fixed length identifier which is used to identify a FEC
 - LER: Label Edge Router
 - LSR: Label Switch Router

How does MPLS work?



Ingress Label Edge Router receives IP packets, performs packet classification (into FECs), assigns a label, & forwards the labeled packet

LSRs forward packets based on the label (no packet classification in the core)

Egress LER removes label before forwarding IP packets outside MPLS network



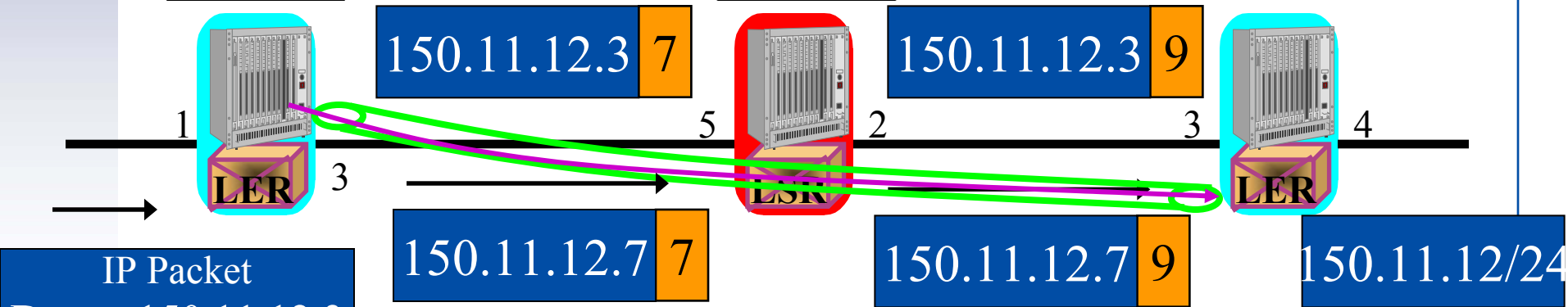
Formerly Bellcore...
Performance from Experience

MPLS Forwarding Example

Input		Address Prefix	Output	
I/F	Label		I/F	Label
1	None	150.11.12/24	3	7
...

Input		Address Prefix	Output	
I/F	Label		I/F	Label
5	7	150.11.12/24	2	9
...

Input		Address Prefix	Output	
I/F	Label		I/F	Label
3	9	150.11.12/24	4	None
...



IP Packet
Dest.= 150.11.12.3

IP Packet
Dest.= 150.11.12.7

- Ingress LER classifies IP pkts into a FEC & assigns label
- LSR forwards labeled packets based on label value
 - no further packet classification into FEC is done
- Egress LER removes label

MPLS

Three Important Questions...

- 1 Q: What field on the labeled packet is used to make the forwarding decision?
 - A: The outermost label
- 2 Q: When this field is used as an index into the Label Information Base (LIB), what is looked up?
 - A: The outbound label value
- 3 Q: What other vital piece of information does the LIB contain?
 - A: The output interface

Forwarding Equivalence Class Granularity

- A FEC is used to define the level of flow aggregation
- A range of granularity levels can be defined for an FEC:
 - *finest granularity level*: application flow (entire host IP address) - most appropriate for local/campus networks
 - *medium granularity level*: IP address prefix - best suited for enterprise networks
 - *coarsest granularity level*: set of IP prefixes - most appropriate for the core/backbone
- Multiple FEC granularities can be used within the same network
- Every LSP is associated with a FEC



Telcordia™
Technologies

Formerly Bellcore...
Performance from Experience

FECs are determined by the network operator, not equipment vendors

MPLS Classification

- As the packet enters the MPLS network, packet classification is performed at the ingress LER (or Edge-LSR)
- Packet classification is done only once at the edge
- Classification mechanism may be complex, since it can rely on:
 - IGP
 - Layer 2 information
 - QoS
 - VPN
 - Traffic Engineering, etc.
- The, potentially, complex packet classification at the edge does not affect packet forwarding performance in the core
 - information required to do packet classification does not need to be

Label Distribution Mechanisms

- All LSRs use a label distribution protocol
 - not necessarily the same mechanism in all LSRs in a MPLS network
- Label Distribution Mechanisms include:
 - static assignment/configuration
 - routing protocols
 - signaling protocols
- Label Distribution via routing
 - Border Gateway Protocol v4 (BGP4)
 - assigns labels to BGP routes

Label Distribution Signaling Mechanisms

- Label Distribution Protocol (LDP)
 - provides mappings from FECs to labels
 - Basic LDP mechanisms include:
 - LDP neighbor detection, session initiation, maintenance and termination
- Constraint-based routing with LDP (CR-LDP)
- Resource Reservation Protocol w/ extensions - RSVP-TE
- RSVP-TE or CR-LDP are used for establishing TEed LSPs
 - most vendors are implementing both signaling mechanisms
 - Some key characteristics:
 - supports explicitly routed LSPs
 - supports LSP set up with QoS parameters
- For most applications, label distribution options in MPLS are richer than necessary

Label Distribution Protocol

- LDP defines a set of procedures by which one LSR informs another LSR of the label bindings it has made
- Does not support
 - multicast, QoS
- Labels are exchanged between LDP Peers
 - two LSRs use an LDP Session to exchange label mapping information
 - peering between non-directly connected LSRs is also supported
- LDP provides a number of protocol control functions
 - peer discovery
 - session management
 - notification

Label Distribution Protocol Message Types

- Four categories of LDP messages are defined:
 - Discovery messages:
 - used to announce and maintain the presence of an LSR in a network
 - Session messages:
 - used to establish, maintain, and terminate sessions between LDP peers
 - Advertisement messages:
 - used to create, change, and delete label mappings for FECs
 - Notification messages:
 - used to provide advisory information and to signal error info.
- Message Transport

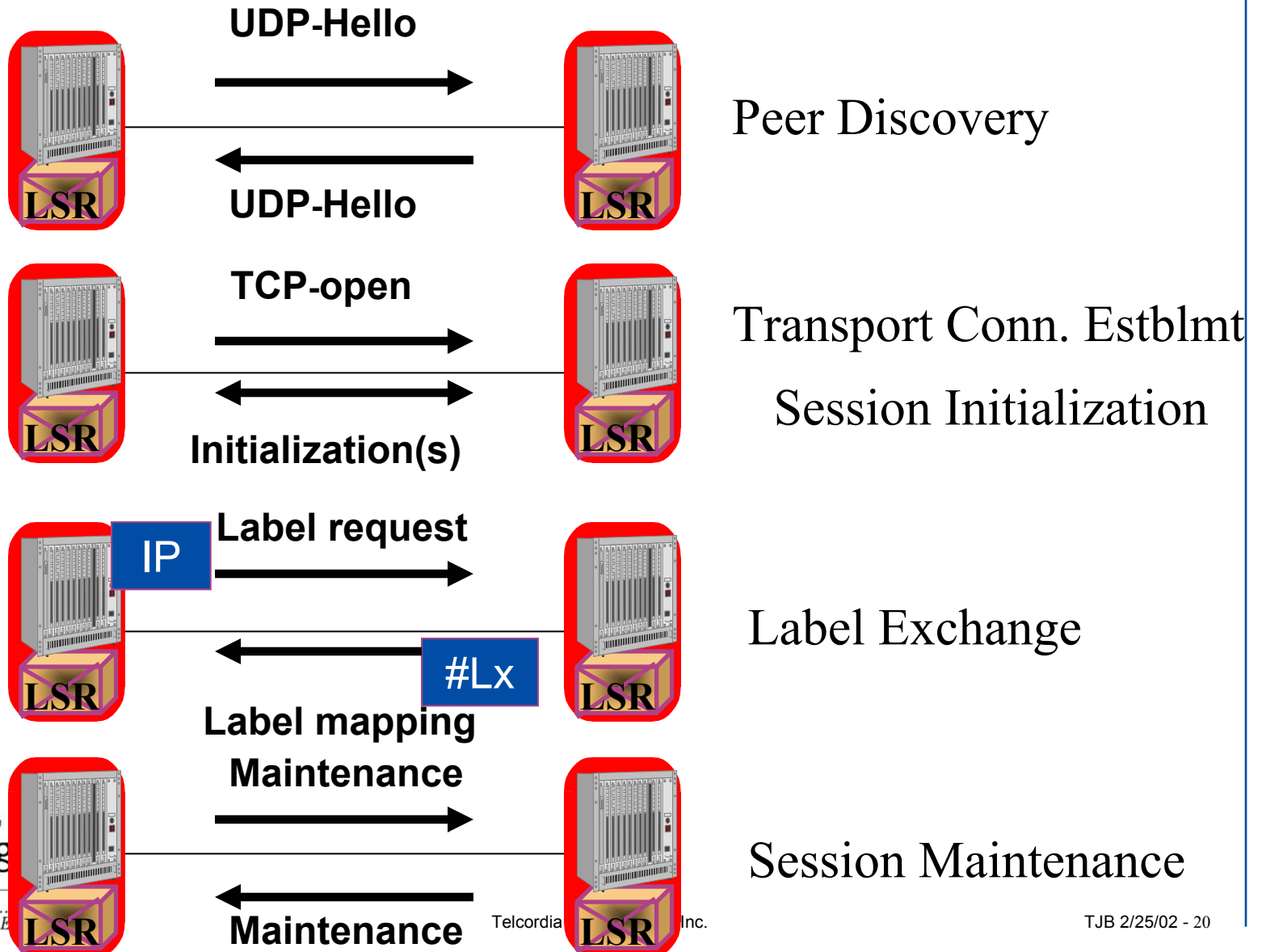


Telcordia™
Technologies

Formerly Bellcore...
Performance from Experience

— All other messages use TCP

Phases of Label Distribution Protocol Operation




Major MPLS Applications

- Transition from IP over ATM to IP/ MPLS
 - Embedded ATM networks carrying IP traffic are migrating to IP/MPLS networks
- Traffic Engineering
 - Optimizes the use of network resources
 - Explicit and policy routing
 - Fast Restoration
- Services
 - IP VPNs (RFC 2547bis: BGP/MPLS VPN)
 - Layer 2 VPNs
 - Layer 2 Transport: ‘Foo’ over MPLS, Foo = ATM, FR, Ethernet, etc.
 - Voice over IP over MPLS and Voice over MPLS (VoMPLS)
- For Optical Networks: Generalized MPLS (GMPLS)
 - Extend MPLS control plane to optical domain



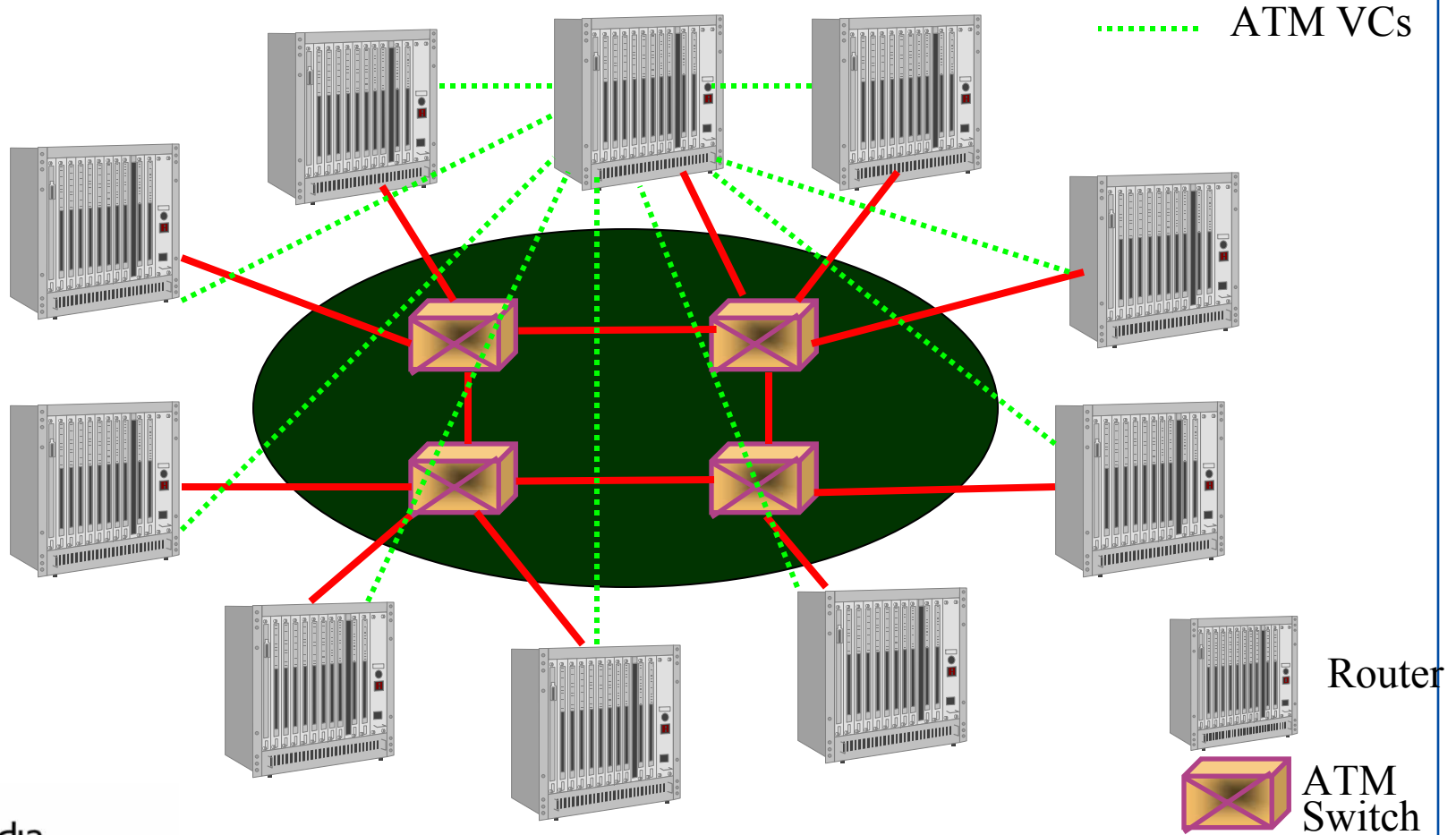
Transition from IP over ATM to IP/ MPLS

- Expensive to maintain two networks
- IP routers can now keep up with ATM switches
 - IP Gigarouters and Terarouters are capable of wire-speed performance
- Why per-hop routing?
 - Answer: IP over ATM
 - investment was already made in ATM, yet growth is in IP traffic
- MPLS is envisioned to provide graceful migration of ATM switches in Internet backbone networks

 Telcordia™
Technologies leverage existing ATM hardware

Overlay network

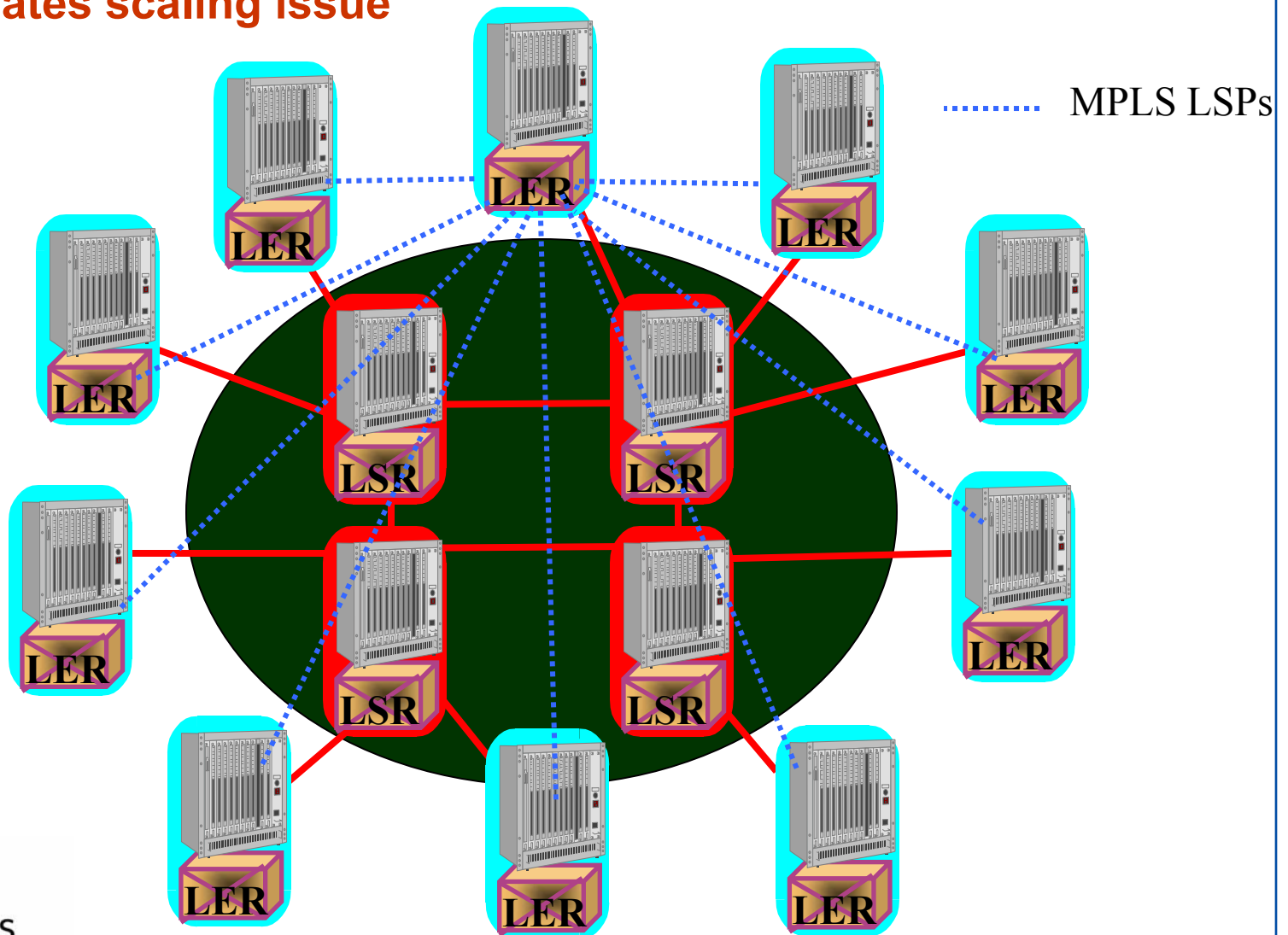
Scaling issue



- IGP routing doesn't scale for full meshes $\rightarrow O(n^3)$, $n = \text{routers}$
- More complex network management \rightarrow 2-level network

Label Switching Routers


Alleviates scaling issue



- IGP routing in MPLS is not dependent on full mesh
- Telcordia Technologies, Inc.

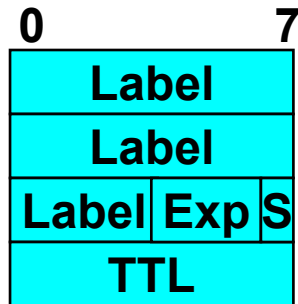
ATM Switches as Label Switching Routers

- MPLS forwarding is similar to that of ATM switches
 - both employ label swapping mechanism
 - ATM switches use input port, VPI, VCI values and map them to output port, VPI, VCI values
- Three methods of encoding labels in the ATM cell header include:
 - Switched Virtual Circuit encoding
 - VPI/VCI field is used to encode the label
 - no label stack operations
 - Switched Virtual Path encoding
 - VPI field to encode the top label; VCI field to encode the second label
 - permits the use of ATM ‘VP-switching’
 - Switched Virtual Path multipoint encoding
 - VPI field to encode the top label; part of the VCI field to encode the 2nd label on the stack, and use the remainder of the VCI field to identify the LSP ingress

 All use, e.g., LDP as the ATM ‘signaling’ protocol
no ATM Forum routing and signaling protocols are used

Other MPLS Encapsulations

- Label format and length depend on encapsulation used
- MPLS is not tied to any particular encapsulation method,
 - e.g. Packet-over-SONET (POS) utilizes IP over PPP over SONET with MPLS shim header



MPLS Shim Header

Label = 20 bits

Exp = Experimental = 3 bits

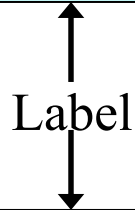
S = Bottom of stack = 1 bit

TTL = Time to live = 8 bits

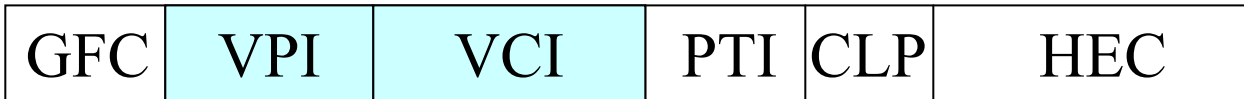
MPLS Labels



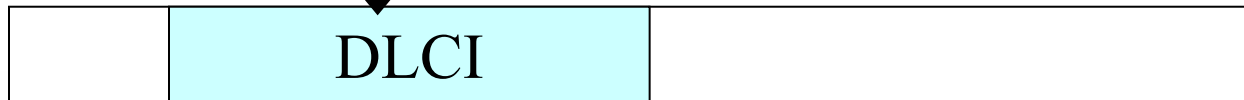
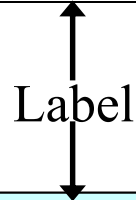
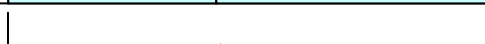
**PPP Header
(POS)**



**LAN MAC
Header**



**ATM Cell
Header**



**Frame Relay
Header**

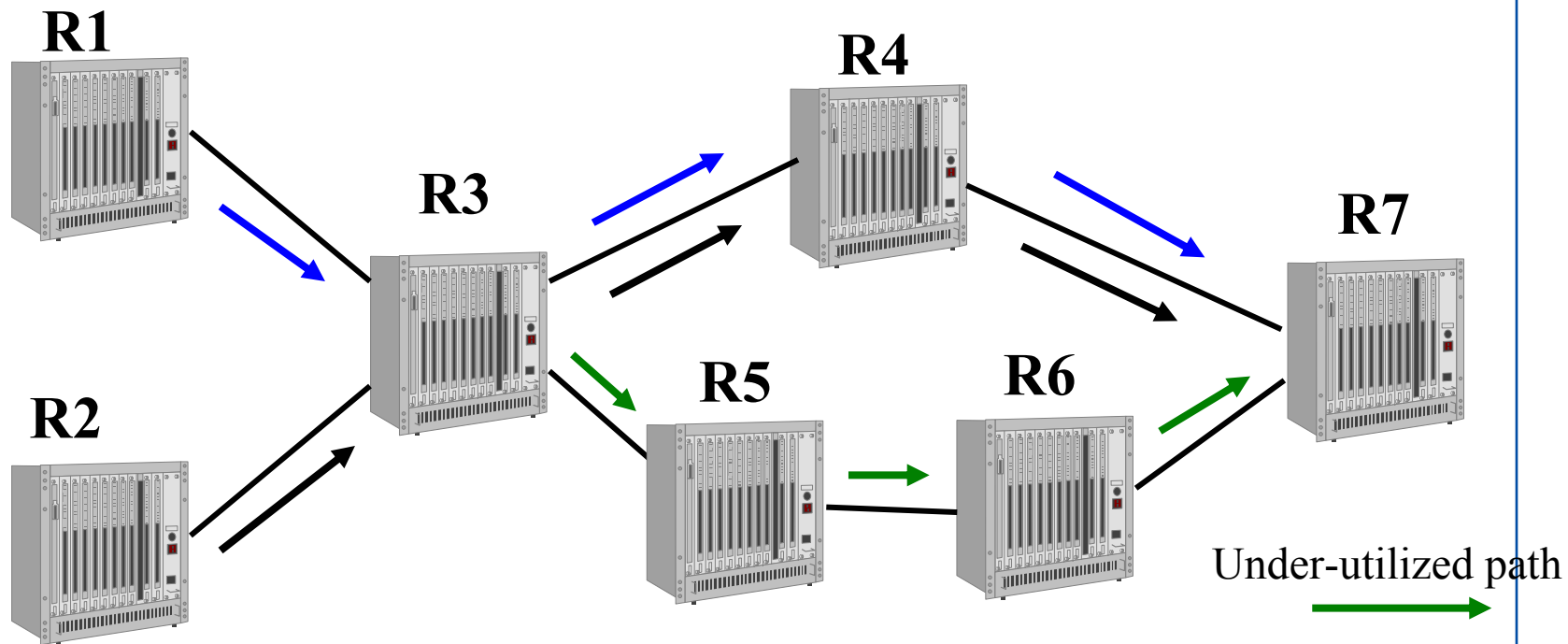


Traffic Engineering

- The goal of traffic engineering is to optimize the utilization of network resources, thus, the performance of operational networks by moving traffic efficiently and reliably through the network
 - reducing congestion & improving network throughput
 - more cost-effective
 - efficiency gained through load balancing
- Other TE Mechanisms (besides MPLS):
 - Excess Capacity / Over provisioning
 - Overlay networks: IP over ATM or FR
 - primary drawbacks include: 2-level network mgmt and scalability
 - Layer 3: path computation based solely on IGP metric is not sufficient
 - operationally difficult tinkering with L3-only metrics in large networks
 - trial-error approach
 - prone to oscillations
 - thus, depending on IGP routing for TE is not sufficient

Traffic Engineering

The Hyper-aggregation or 'Fish' Problem



- IP employs shortest path destination based routing
 - there are other paths available besides the shortest path
- Shortest path may be over-utilized while alternate path may be under-utilized

MPLS as a solution

- MPLS provides better support for routing in the traffic engineering context
 - supports explicit routes based on constraints other than destination address, e.g. available bandwidth
 - supports priorities for pre-empting existing paths and for holding onto resources
 - supports resource class affinities that allow/disallow certain “colored” links from the path of the traffic trunk
 - supports load balancing for parallel paths
 - supports better fault recovery procedures for rerouting and restoring paths upon failure

Components for MPLS Traffic Engineering

- Terminology: Traffic Trunk - aggregation of flows that are:
 - forwarded along a common path within a SP network
 - primarily from a POP to another POP
 - share a common QoS requirement
- Trunk Attributes
- Information Distribution
 - distributes resources/constraints pertaining to links
- Path Selection
 - computes paths that obey constraints
- Signaling
 - establishes path

Trunk Attributes

- These attributes are configured at the ingress LER
- Bandwidth
- Priorities
 - setup priority: priority for taking a resource
 - holding priority: priority for holding on to a resource
- Resource Class Affinity
 - in addition to QoS-based routes, routes can be based on policy
 - supports the ability to exclude/include certain links for specific traffic trunks based on policy
 - LSP Tunnel is characterized by a
 - 32-bit resource-class affinity bit string
 - 32-bit resource-class mask



Telcordia™
Technologies

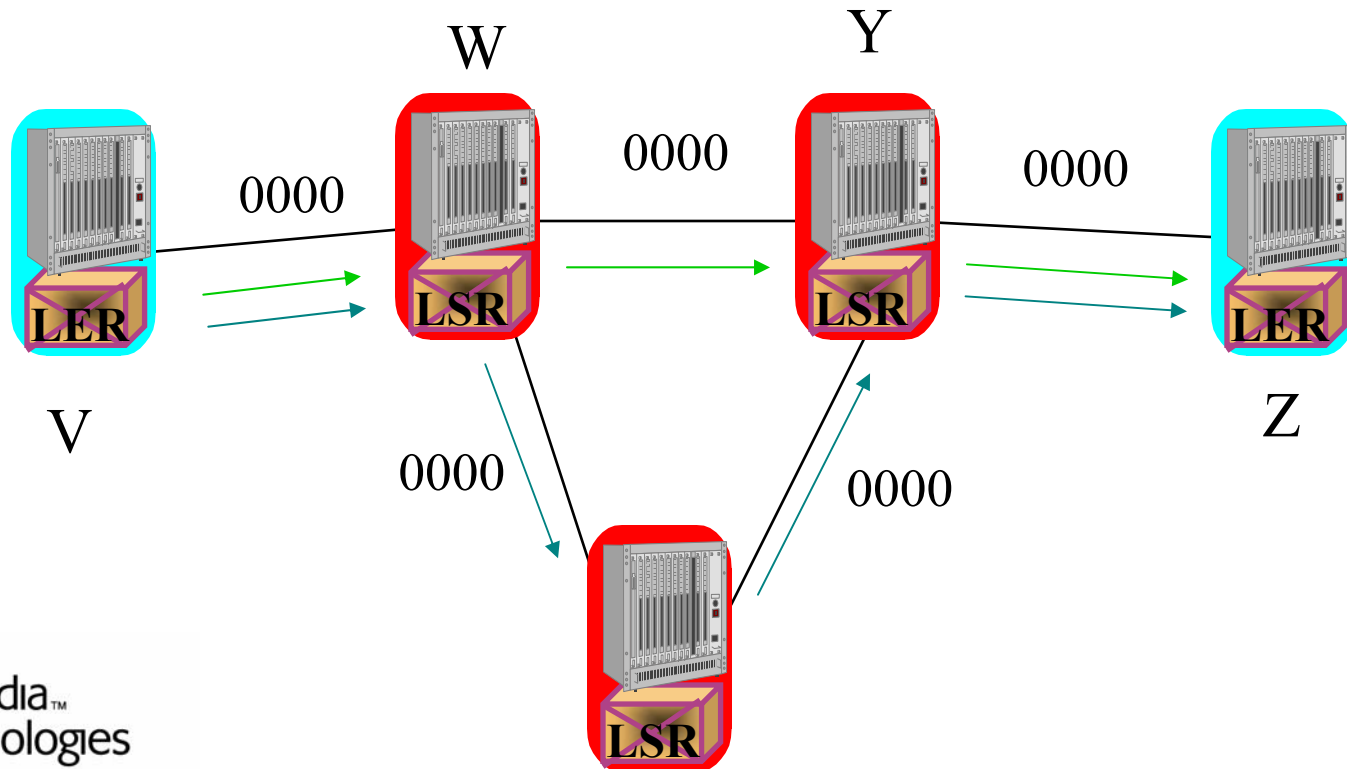
Formerly Bellcore...
Performance from Experience

– 0 = don't care & 1 = care

link is characterized by a 32-bit resource class attribute string

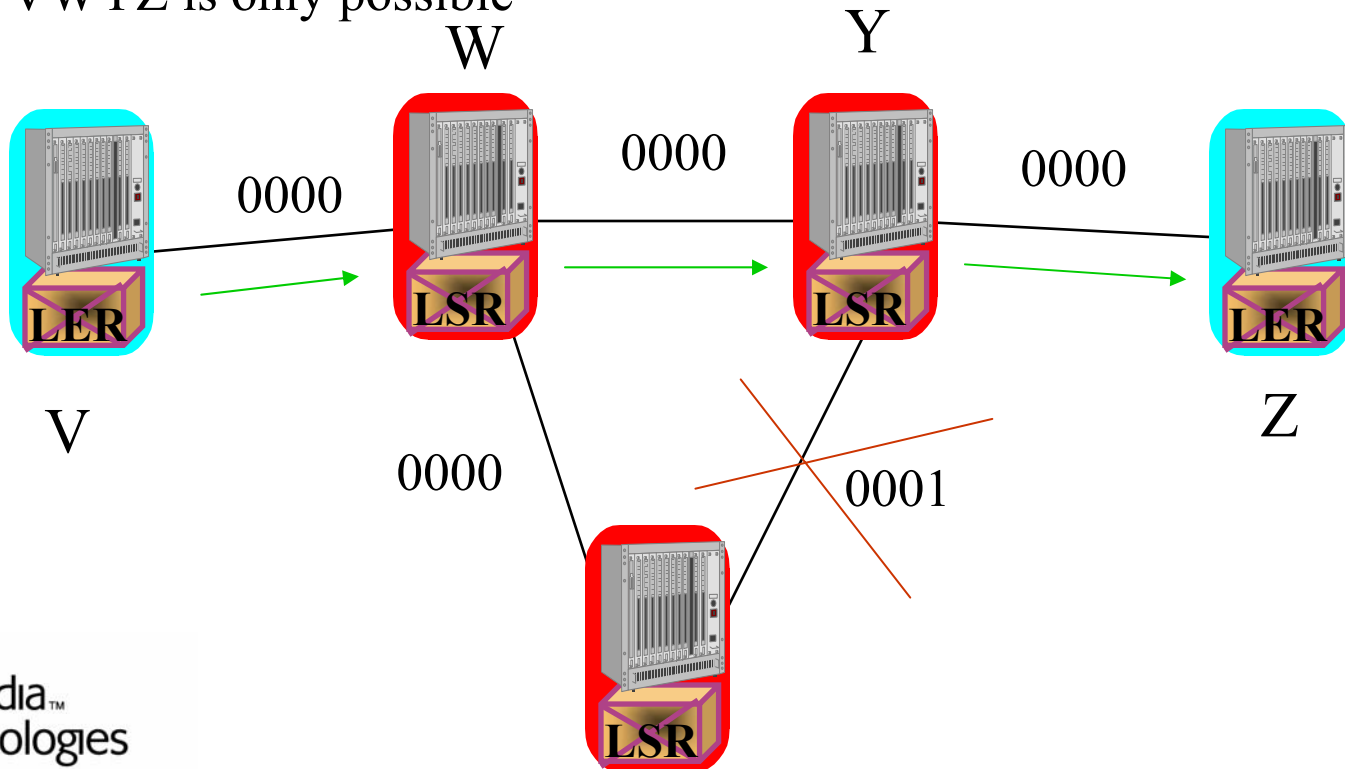
Policy Example 1

- Trunk V to Z:
 - tunnel = 0000, t-mask = 0011
- VWYZ and VWXYZ are both possible



Policy Example 2

- Setting X-Y link bit pushes all tunnels off the link
- Trunk V to Z:
 - tunnel = 0000, t-mask = 0011
- VWYZ is only possible



Information Distribution

- TE requires detailed knowledge about network topology and resources
- The flooding service from link-state IGP is re-used
 - opaque LSA for OSPF-TE
 - new TLV for IS-IS-TE
- TE extensions include
 - link bandwidth
 - maximum reservable link bandwidth
 - available bandwidth
 - traffic engineering metric
 - link color

Path Selection

- May be a combination of on-line and off-line procedures
 - active area of research
- Constrained Shortest Path First
 - on-line mechanism
 - takes into account specific restrictions when calculating the shortest path
- Offline procedure is needed to optimize traffic engineering globally
 - pre-determines LSPs

Path Selection

- ‘Problem Statement’
 - Given network information:
 - Connectivity
 - Link capacities
 - Demand between each pair of nodes
 - Route demands to optimize capacity use:
 - Two decisions for each demand:
 - 1) What are the LSPs?
 - 2) How is flow allocated among LSPs?
- The “Optimization Problem”
 - Constraints on decisions:
 - We have to route all of the offered demand
 - We can’t exceed the available capacity on any link
 - Optimization goals
 - Delay?
 - Congestion?
 - Path length?

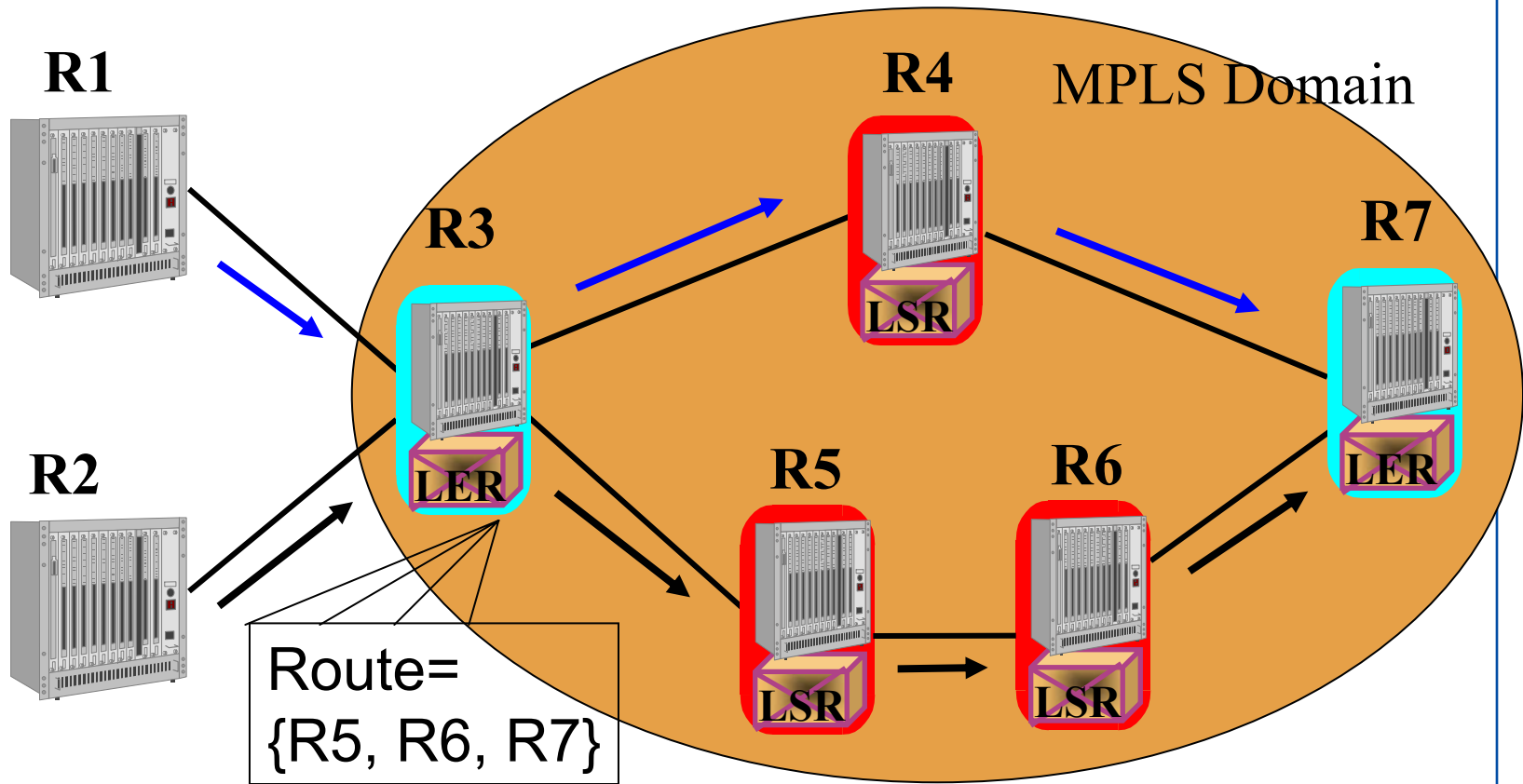
Signaling

- Establishes forwarding state and performs label distribution
 - path is not known if workable until the LSP is established
- RSVP-TE or CR-LDP are used for establishing LSPs
 - most vendors are implementing both signaling mechanisms
- Some characteristics:
 - supports explicit and record route functions
 - supports QoS
 - Preemption
 - supports make-before-break
 - Neighbor failure detection

Explicitly Routed LSPs

- MPLS allows traffic to be forwarded on paths other than those that are indicated by network layer routing
 - efficiency, reliability, and optimization
- ‘Explicit Routing’ (a.k.a., ‘source routing’)
 - builds a path from source to destination for a particular FEC
 - essentially a unidirectional VC
 - MPLS supports ‘strict’ or ‘loose’ modes
 - may be manually or automatically provisioned
 - QoS, policy, plus other constraints may be used to determine ER
 - Backup paths may be pre-provisioned for rapid restoration

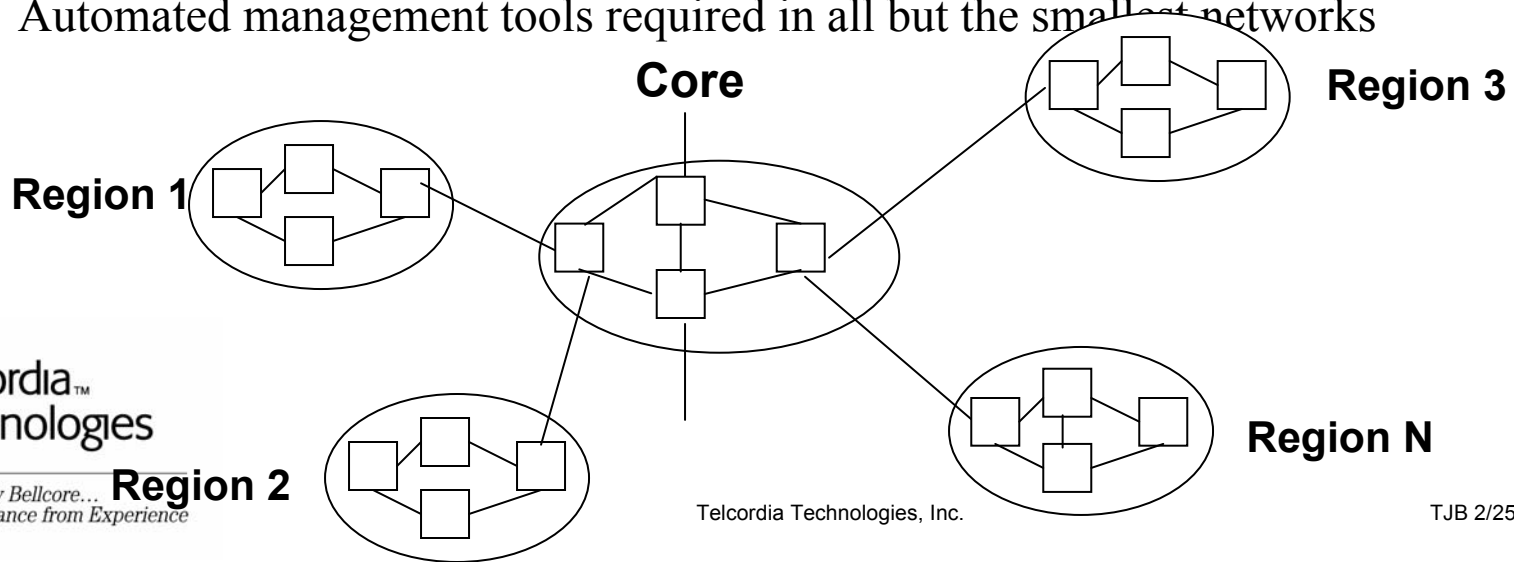
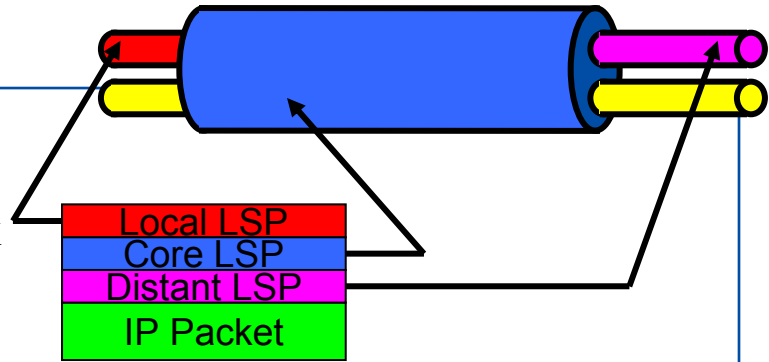
MPLS Solution to the Hyper-aggregation Problem



- Blue path -> produced by LDP, follows normal IP routing
- Black path -> ER-LSP follows route that ingress LER selects

Hierarchical MPLS Network

- MPLS lends itself to the hierarchical network
- Full mesh of MPLS LSPs is not scalable
 - e.g., 5K nodes, yields ~ 25M paths
- Splitting the MPLS network into core and regional networks makes network management simpler
 - full mesh within each regional network - 9900 LSPs
 - full mesh within the core to interconnect regions - 2450 LSPs
 - total LSPs is $9900 * 50 + 2450 = 497,450$
 - better scalability
 - Only LSPs in the region affected when node is added
 - Task of TE tools is simpler
- Automated management tools required in all but the smallest networks



Virtual Private Networks

- Virtual Private Networks provide interconnection of customer sites over a shared network infrastructure
 - the shared infrastructure could be the “Internet” or a Service Provider’s (SP) backbone network
- VPNs provide a cost effective solution
 - savings in network infrastructure hardware
 - savings in management of the network infrastructure
- Key issues for VPNs:
 - private IP addresses: non-unique, overlapping address spaces
 - data security: authentication, integrity, privacy
 - quality of service assurances: bandwidth, latency
 - scalability

VPN Solutions

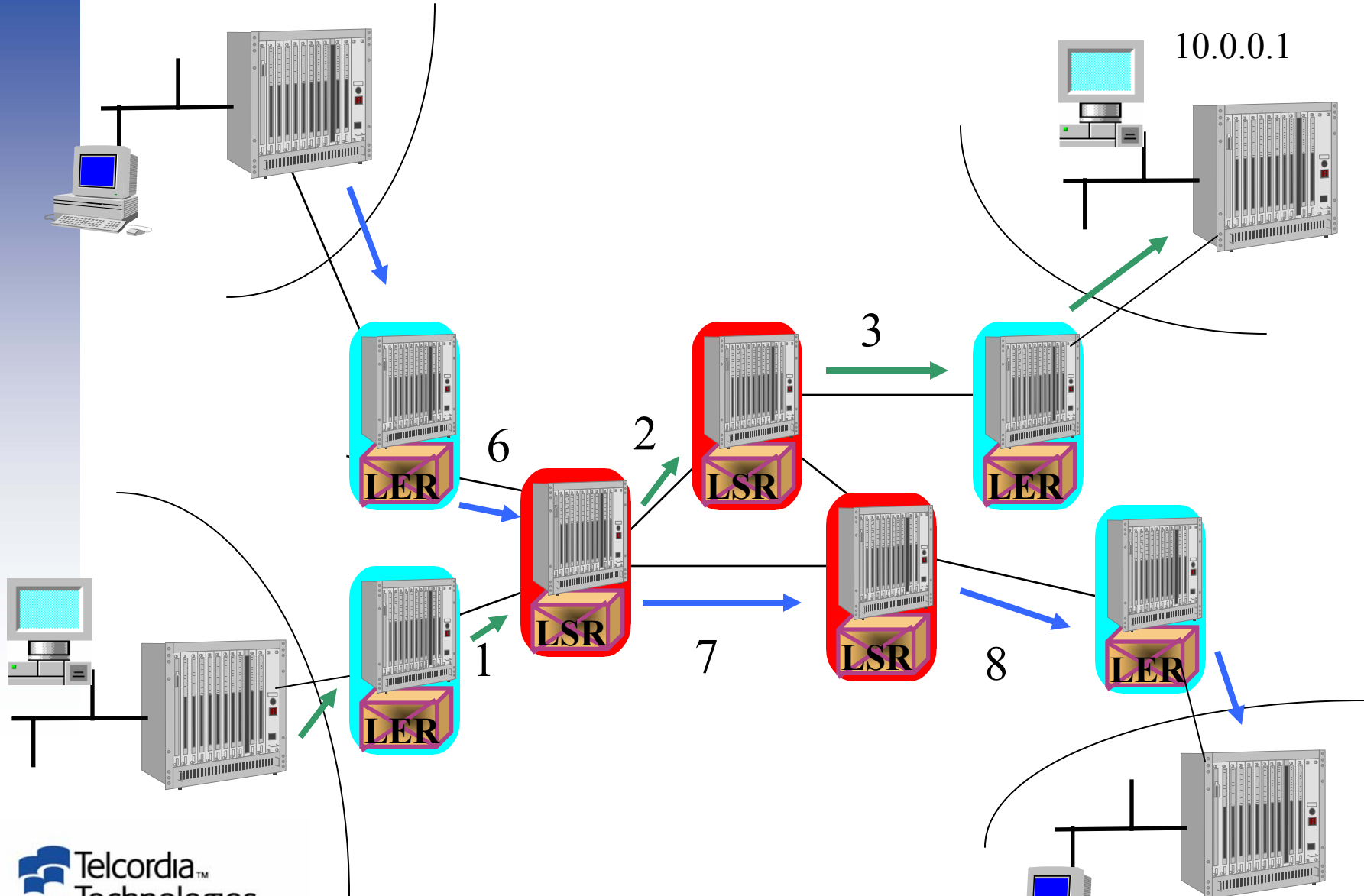
- A multitude of VPN solutions exist
 - CPE-based VPNs:
 - e.g., GRE, L2TP, PPTP, IPSec
 - Virtual Leased Line (VLL) VPNs:
 - WAN connectivity through leased line or switched circuit
 - Service Provider (SP) does not examine Network Layer Reachability Information (NLRI) of VPN data packets; e.g., Frame Relay, ATM, MPLS
- MPLS VPNs can also be Network-based (or Provider Provisioned) Virtual Private Routed Networks
 - based on NLRI
 - SP participates in the management and provisioning of the VPN

How can MPLS help?

- Due to the ability of MPLS to de-couple the context of a packet's IP header via a label, it provides a straightforward solution to hide private addresses
 - creates tunnels (via encapsulation)
 - Tunnels extend only as far as MPLS extends
- Provides adequate security
 - 'ATM grade' security
 - strong security requires IPsec tunnels inside MPLS tunnels
- Quality of Service
 - provides signaling of bandwidth and QoS requirements
 - Connectionless IP appears as connection-oriented

Enterprise ABC

Enterprise XYZ



10.0.0.1

10.0.0.1



Formerly Bellcore...
Performance from Experience

Enterprise XYZ

Enterprise ABC

Telcordia Technologies, Inc.

MPLS VPNs

- There is no standards based MPLS VPN solution
 - however, the IETF and ITU are trying to work towards that goal
 - definition, requirements, and scope of VPNs being developed
- Each vendor has their own proprietary MPLS VPN scheme
 - e.g., Cisco's BGP/MPLS VPN, Nortel's MPLS-based Virtual Router, Lucent's Virtual Router
- Being deployed in a number of ISPs

MPLS Industry Fora and Consortia

- The Internet Engineering Task Force (IETF)
 - Developed MPLS protocols, encapsulations, etc.
- MPLS Forum
 - focusing on work items that accelerates MPLS deployment
 - e.g., interoperability and VoMPLS
- International Telecommunications Union (ITU)
 - Specifies MPLS architectures and equipment requirements
- Among others...

IETF MPLS Standardization Status

- IETF MPLS standardization
 - working group began in early 1997
 - lots of interest as gauged by the attendance/participation at MPLS WG meetings
- RFCs issued:
 - RFC 2702: ‘Requirements for Traffic Engineering Over MPLS’
 - Standards track RFCs: 3031-3038, 3063, among others
- Over the last year, PPVPN (Provider Provisioned Virtual Private Network) working group in the IETF was created
 - part of ‘sub-IP’ pseudo-area that the IESG created

 **Work in progress:**
Telcordia™
Technologies – Generalized Multiprotocol Label Switching

Label Switching Router Implementations

- Cisco Systems
- Juniper Networks
- ~~Cascade's~~ ~~Ascend's~~ Lucent's IP Navigator
- Nortel
 - ~~Bay's~~ Nortel's Versalar Backbone Node routers, Passport Switch
- Ericsson's AXI 530 switch product family
- ~~Fore Systems~~ Marconi
- Lots of start-up vendors

SPs that announced MPLS-based VPN Services

- AT&T
- Global Crossing
- Level 3 Comm.
- UUNET
- among others...
- Bell Canada
- British Telecom
- France Telecom
- Swisscom
- Telenor
- among others...

Summary

- MPLS will play a key role in future network architectures
- Network Element support for MPLS is pervasive
- Service Providers
 - are deploying MPLS in their operational networks
 - are pushing MPLS in directions that enable them to more easily grow their networks
- MPLS is currently mainly a core technology; access part being worked
- MPLS is being used to provide VPN service
- Holds a lot of potential for dealing with some real problems such as traffic engineering
- Accelerated MPLS deployments in operational networks are anticipated this year

Thank You!

Questions/Comments?

MPLS reference:

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



*Formerly Bellcore...
Performance from Experience*