Chair for Network Architectures and Services – Prof. Carle
Department for Computer Science
Technische Universität München

Master Course Computer Networks IN2097

Prof. Dr.-Ing. Georg Carle
Christian Grothoff, Ph.D.
Dr. Nils Kammenhuber

Chair for Network Architectures and Services
Institut für Informatik
Technische Universität München
http://www.net.in.tum.de

Technische Universität München

Chapter 4.5: Between Network Layer and Link Layer

MPLS

- Multi-Protocol Label Switching
- "Layer 2.5":
  - Below IP (Layer 3), but above Link Layer (Layer 2)
  - Borrows a lot of information from IP Layer
  - Borrows a lot of concepts from ATM (Layer 2)
  - "A compromise/marriage between IP world and ATM world"
- Mixture of packet switching and circuit switching
  - Establish virtual circuits (LSPs) between endpoints
  - Send labelled packets along these LSPs
- Used by many, but not all, large ISPs

Why a new protocol? — Deficiencies of IP

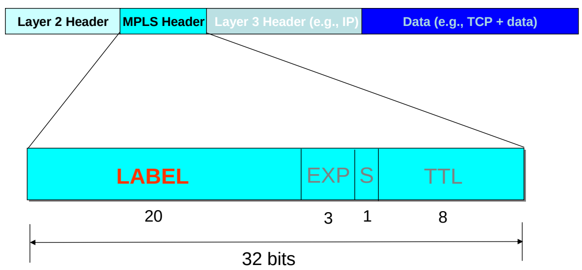
- IP forwarding = longest prefix match on address = expensive
  - < 1K gateways to other ASes; >> 100K interdomain prefixes!
  - Longest prefix match in *every* router on path through our network!
- IP forwarding = destination-based
  - Not all paths possible, cf. exercise #8
  - Would be nicer for traffic engineering
- IP header is long complex
  - Destination, TTL [, QoS bits] at different byte/bit offsets
  - Expensive to parse in hardware
- Traffic of different VPN customers may disturb each other
  - Not visible to each other, but overloads on common links
- IP routing = slow: OSPF convergence 300ms to *X* seconds
  - Routing loops etc. during convergence → packet losses
  - Think of VoIP, videoconferencing, games, telesurgery, …

Why should we use MPLS?

- Original motivation:
  - Switching is faster than routing
  - Build cheaper high-speed "routers" (which *switch* MPLS)
- Today's motivation:
  - Separation of virtual circuits; MPLS-VPNs
  - Multiservice networks: not only IP
  - Arbitrary paths, better for traffic engineering
  - Fast reroute mechanisms (protection switching): 50ms
  - Better control over routing: More deterministic, more predictable, better for QoS service level agreements (SLAs)

MPLS ideas (1): Short and simple header

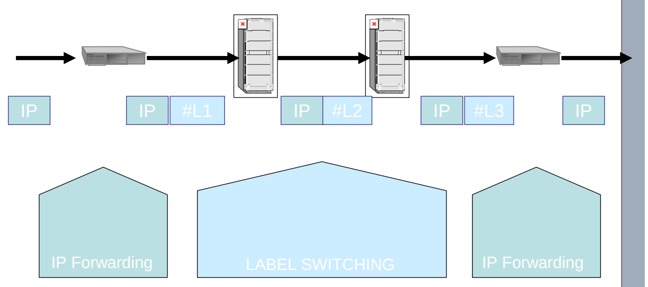
- Easy to parse in hardware



Basic Concepts and Terms

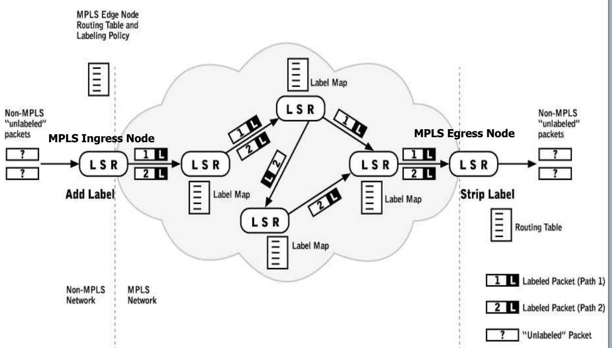
- Label Switching Routers (LSRs): Any router supporting MPLS
- Label
  - A fixed-length (20-bit) address
  - Label semantics are *local* to a router: One label ≠ one path!
  - Labels may be swapped at each router
  - Labels may be stacked: “MPLS in MPLS”
- Label Switched Paths (LSPs)
  - An MPLS virtual circuit: Like a tunnel through the network
  - LSPs are unidirectional
- Forwarding Equivalence Classes (FECs):
  All packets that are to to forwarded….:
  - To the same next hop
  - Out the same interface
  - [With the same forwarding treatment (CoS)]

MPLS ideas (2a): Route at edge, switch in core

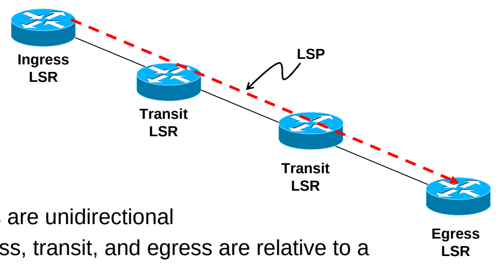


MPLS ideas (2b): Forwarding through the network

- IP packets: Labelled at ingress, label stripped at egress
- Within network: Forwarding by label, not by IP address!



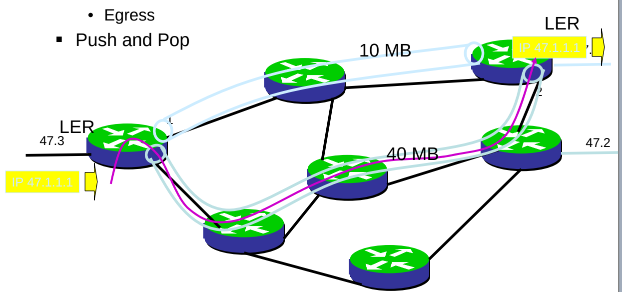
Basic Concepts and Terms



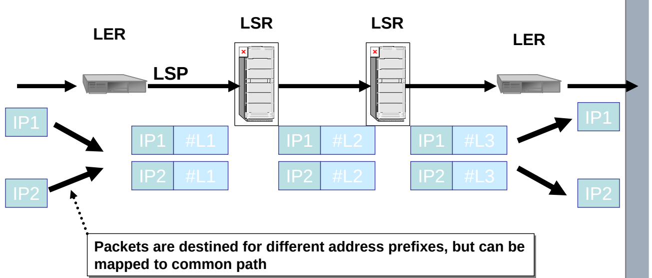
- LSPs are unidirectional
- Ingress, transit, and egress are relative to a given LSP
- A given router can be ingress, egress, and transit for different LSPs

Label Edge Routers

- Label Edge Router (LER)
  - A tunnel (LSP) endpoint
    - Ingress
    - Egress
  - Push and Pop



Forwarding Equivalence Classes

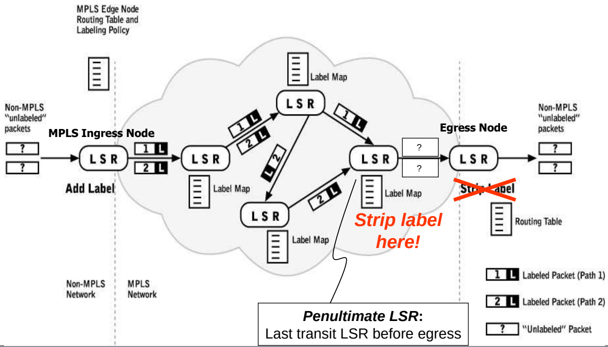


Packets are destined for different address prefixes, but can be mapped to common path

- 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.

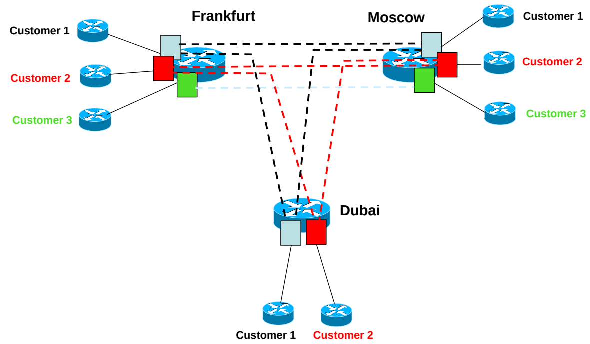
Penultimate Hop Popping

- Egress has to apply IP routing anyway
- → Can remove MPLS label one hop *before* egress



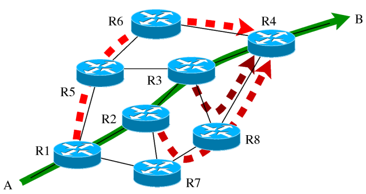
MPLS VPNs

- Virtual Private Networks: Make customers feel as if they have a direct and private connection



MPLS Fast Reroute (FRR)

- Configuration of backup paths in network
  - Many (local) backup paths for each primary LSR
- Upon detection of a failure
  - LSR immediately switches to its local backup path
  - No need to wait for signalling upstream! (in our example: R1)
- Very fast reaction speed: 50ms



Configuring labels

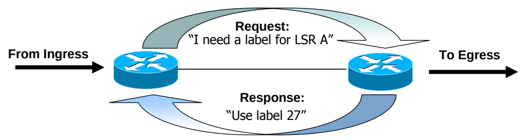
- Labels: *local* to each router
- How do routers get to know labels and their semantics?
  a) Manual configuration: does not scale
  b) Signalling: Using some label distribution protocol
     - Set of procedures by which one LSR informs another LSRs of the bindings (label/FEC) it has made

Label Assignment and Distribution

- Decision to bind a particular label L to a particular FEC F
  - made by LSR which is downstream (with respect to that binding)
- Downstream LSR informs upstream LSR of the binding
- Direction
  - Labels are 'downstream assigned'
  - Label bindings are distributed in 'downstream to upstream' direction.

Label Distribution

- Requests for labels flow *downstream*
  - From Ingress to Egress (like the MPLS packets)
  - Because ingress is the LSR that establishes the LSP
- Assignment of labels (label binding) flows *upstream*
  - From Egress to Ingress
  - Because LSRs need to map *incoming* labels to some action (Push, Swap, Pop)



Label Distribution Protocols

- Label Distribution Protocol (LDP)
  - Hop-by-hop label distribution
  - Follows IGP best path: No traffic engineering capabilities
  - Highly scalable: Best suited for apps using thousands of LSPs (VPNs)
- Resource Reservation Protocol with Traffic Engineering Extensions (RSVP-TE)
  - End-to-end LSP signaling
  - Enables specification of path constraints
  - Less scalable, LSRs maintain soft state: Best suited for traffic engineering in the core

Label Distribution: RSVP

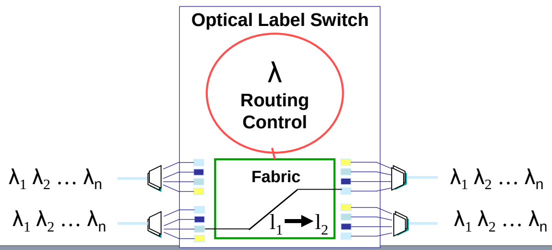
- End-to-end *constrained* path signaling
- Enabled by OSPF or IS-IS with TE extensions
  - Extended IGPs flood TE interface parameters, e.g.:
    - Maximum Reservable Bandwidth
    - Unreserved Bandwidth
    - …
- Interface parameters used to build *Traffic Engineering Database* (TED)
- *Constrained Shortest Path First* (CSPF): Calculates best path based on specified constraints

Label Distribution Protocols: Less used

- Constraint-Based Routed LDP (CR-LDP)
  - TE-capable LDP
  - Never widely deployed
- MP-BGP
  - Best suited for inter-AS VPNs
  - Inter-AS MPLS is a pain in the neck…

Generalized MPLS (GMPLS) for optical media

- Optical networks:
  - Switch fabric ≈ mirrors that reflect light beams
  - One glass fibre, multiple wavelenghts: $\lambda_1\ \lambda_2 \ldots \lambda_n$
  - Problem: Keep same wavelength λ_i through entire network!
- λ_i = just another label to distribute! No new protocols required.



Why not to use MPLS

- Complexity
  - MPLS + some LDP = complex
  - Intradomain IP routing + MPLS + some LDP + intelligent Link Layer = very complex
- Higher complexity means…
  - Hard to debug
  - More administration overhead, and administrators are expensive
- Inter-AS MPLS only works in theory
  - Intradomain routing + Interdomain routing + MPLS + own LDP configuration + LDP configuration of peer ASes + intelligent Link Layer + intelligent Link Layers of other ASes = unmanageable

MPLS: Summary

- Sits between IP (L3) and Link Layer (L2)
- Switching instead of routing
- Aribtrary paths in network (LSPs)
- Setup of LSPs: Label distribution protocols
- GMPLS for optical networks

THANK YOU