MPLS

Multi-Protocol Label Switching

Agenda

- **Review ATM**
- **IP over WAN Problems (Traditional Approach)**
- **MPLS Principles**
- **Label Distribution Methods**
- **MPLS Details (Cisco)**
- **RFCs**

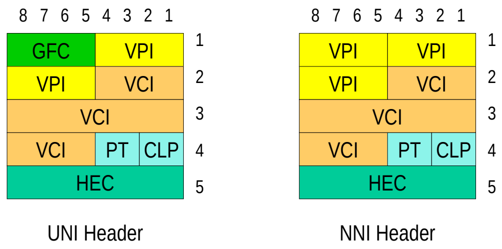
Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

ATM Principles

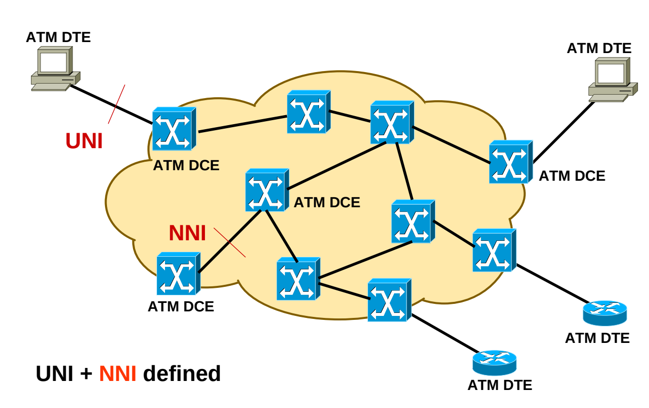
- **ATM**
  - Asynchronous Transfer Mode
  - Based on asynchronous TDM
    - Hence buffering and address information is necessary
    - Variable delay (!)
- **Cell switching technology**
  - Based on store-and-forward of cells
  - Connection-oriented type of service with PVC and SVC
  - But no error recovery (!)
- **ATM cell**
  - Small packet with constant length
  - 53 bytes long (5 bytes header + 48 bytes data)

Cell Format



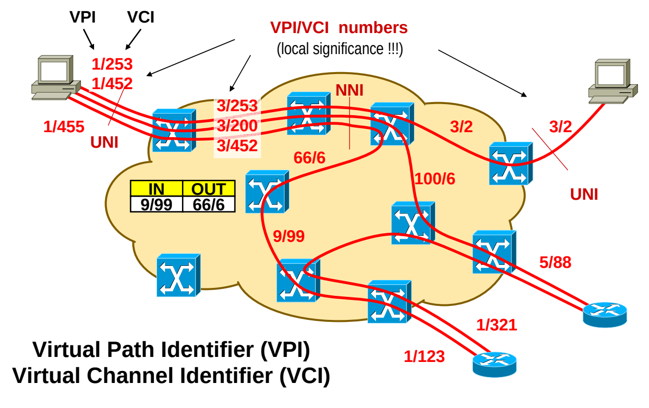
- **Two slightly different formats**
  - UNI … 8 bits for VPI
  - NNI … 12 bits for VPI

ATM Network: Physical Topology



© 2012, D.I. Lindner MPLS v4.6 5

ATM Network: Virtual Circuits
Local Connection Identifiers and Logical Channels

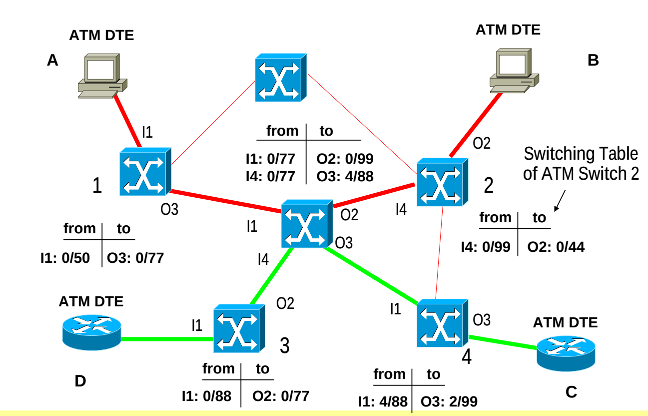


© 2012, D.I. Lindner MPLS v4.6 6

Datenkommunikation 384.081 - SS 2012

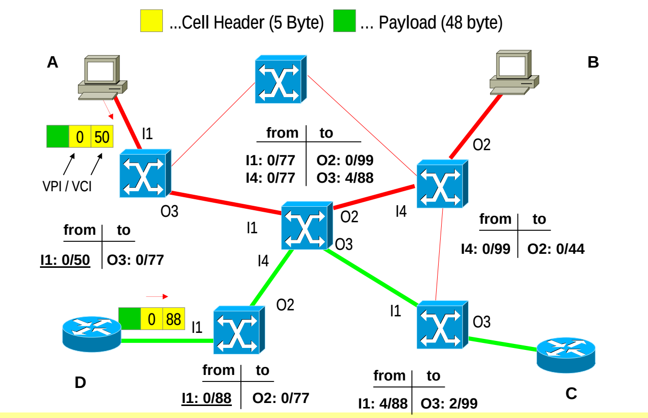
Appendix 3 - MPLS (v4.6)

ATM Switching Tables



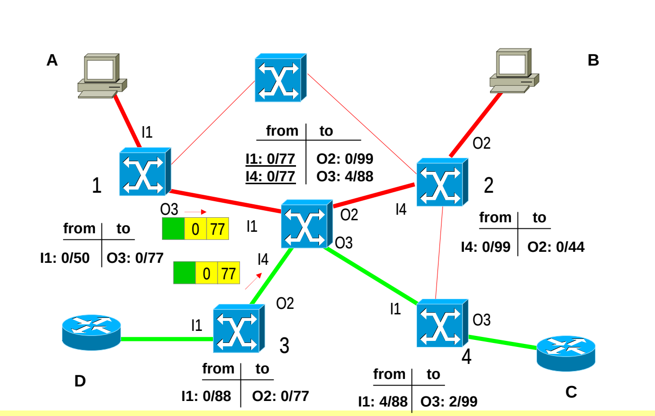
© 2012, D.I. Lindner MPLS v4.6 7

Cell Forwarding / Label Swapping 1

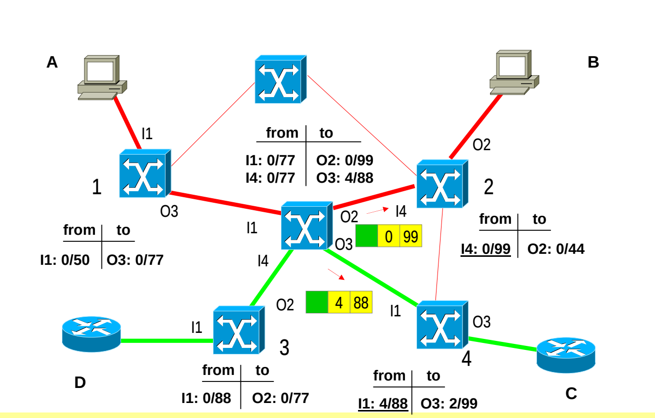


© 2012, D.I. Lindner MPLS v4.6 8

Cell Forwarding / Label Swapping 2



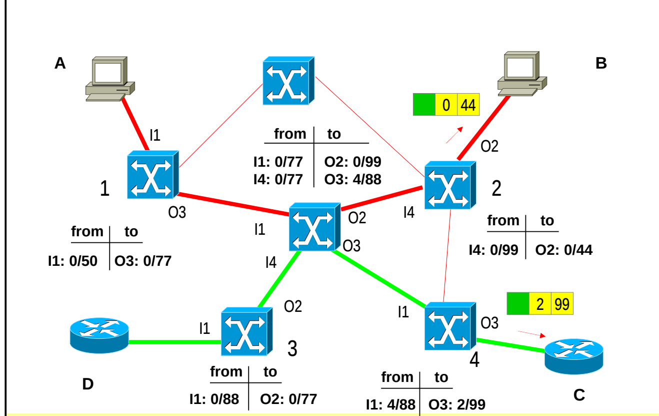
Cell Forwarding / Label Swapping 3



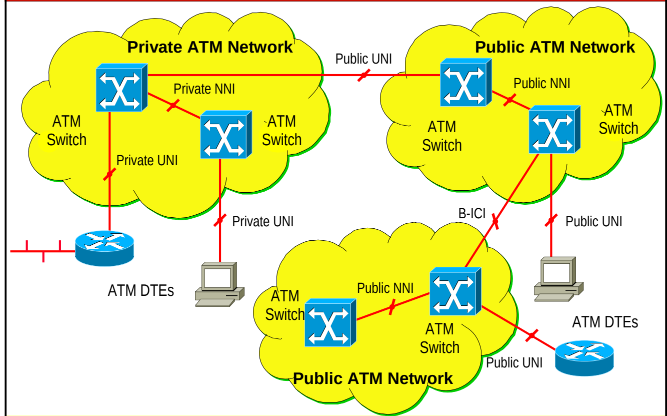
Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

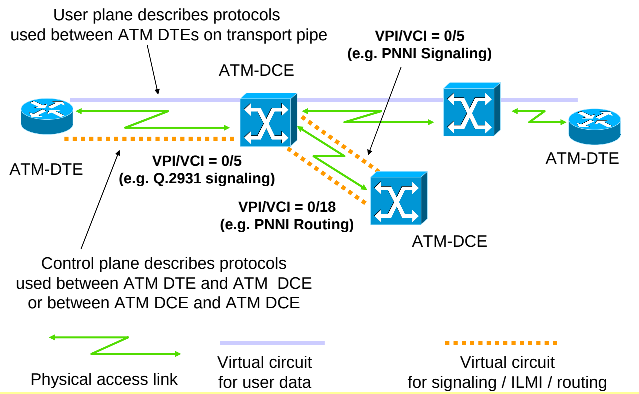
Cell Forwarding / Label Swapping 4



UNI and NNI Types



Control Plane <-> User Plane



Service Classes

Guaranteed Service "Bandwidth on Demand"	CBR	Constant Bit Rate Circuit Emulation, Voice
	VBR	Variable Bit Rate Full Traffic Characterization Real-Time VBR and Non Real-Time VBR
"Best Effort" Service	UBR	Unspecified Bit Rate No Guarantees, "Send and Pray"
	ABR	Available Bit Rate No Quantitative Guarantees, but Congestion Control Feedback assures low cell loss

Traffic Contract per Service Class

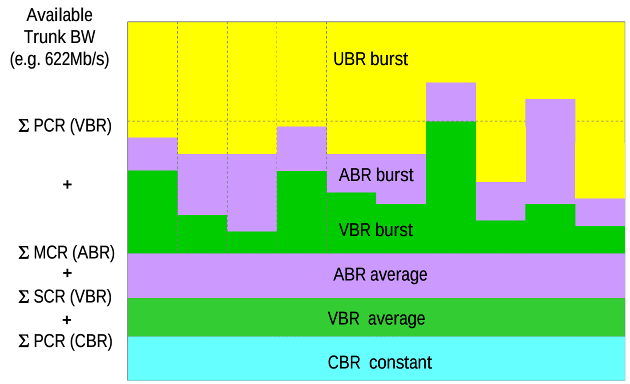
- **Specified for each service class**

ATTRIBUTE	CBR	rt-VBR	nrt-VBR	ABR	UBR
PCR & CDVT	Specified			Specified	
SCR, MBS, CDVT	n/a	Specified		n/a	
MCR	n/a			Specified	n/a
max CTD & ptp CDV	Specified		Unspecified	Unspecified	
CLR	Specified			Optional	Unspecified

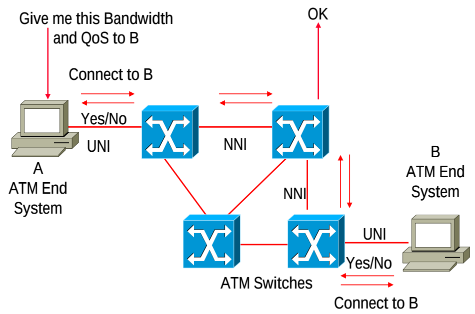
CLR = Cell Loss Ratio
CTD = Cell Transfer Delay
CDV = Cell Delay Variation
MBS = Maximum Burst Size

PCR = Peak Cell Rate
CDVT = CDV Tolerance
SCR = Sustainable CR
MCR = Minimum CR

ATM as an Intelligent Bandwidth Management System



ATM Goal: Bandwidth on Demand with QoS Guarantees



ATM Routing in Private ATM Networks

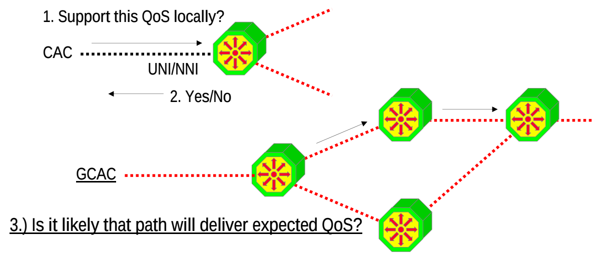
- **PNNI is based on Link-State technique**
  - like OSPF
- **Topology database**
  - Every switch maintains a database representing the states of the links and the switches
  - Extension to link state routing !!!
  - Announce status of node (!) as well as status of links
    - Contains dynamic parameters like delay, available cell rate, etc. versus static-only parameters of OSPF (link up/down, node up/down, nominal bandwidth of link)
- **Path determination based on metrics**
  - Much more complex than with standard routing protocols because of ATM-inherent QoS support

Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

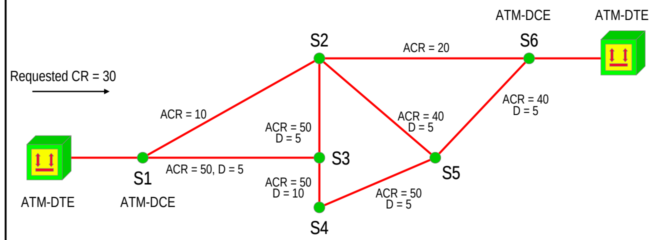
PNNI Routing

- **Generic Connection Admission Control (GCAC)**
  - Used by the source switch to select a path through the network
  - Calculates the expected CAC (Connection Admission Control) behavior of another node



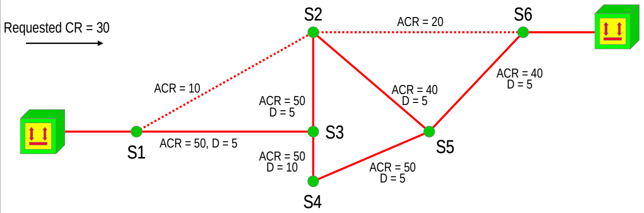
PNNI Routing (Simple QoS -> ACR only)

- **Operation of the GCAC**
  - CR … Cell Rate
  - ACR … Available Cell Rate
  - D … Distance like OSPF costs



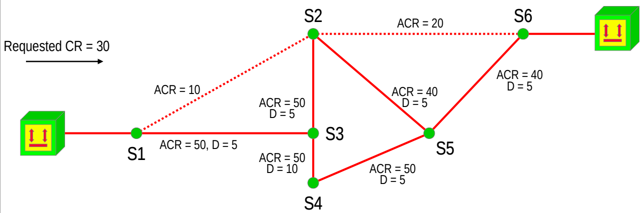
PNNI Routing

- **Operation of the GCAC**
  - 1) Links not supporting requested CR are eliminated ->
    - Metric component -> ACR value used



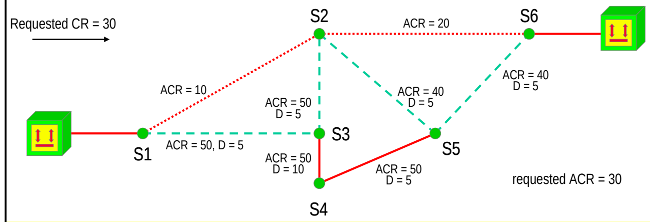
PNNI Routing

- **Operation of the GCAC**
  - 2) Next, shortest path(s) to the destination is (are) calculated
    - Metric component -> Distance value used



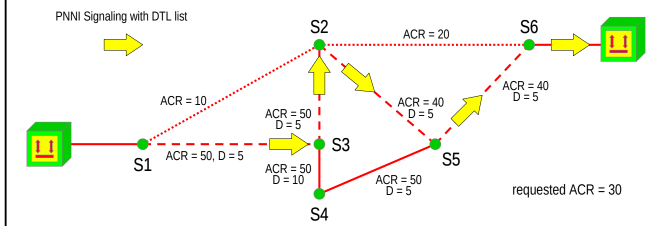
PNNI Routing

- **Operation of the GCAC**
  - 3) One path is chosen and source node S1 constructs a Designated Transit List (DTL) -> source routing -->
    - Describes the complete route to the destination



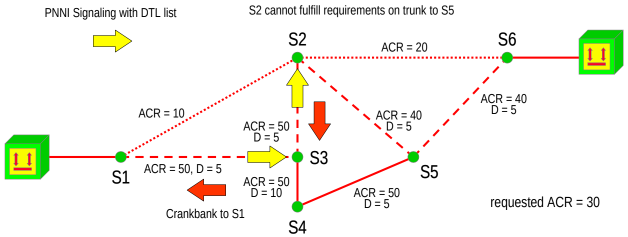
PNNI Routing - Source Routing

- **Operation of the GCAC**
  - 4) DTL is inserted into signaling request and moved on to next switch
  - 5) After receipt next switch perform local CAC
    - 5a) if ok -> pass PNNI signaling message on to next switch of DTL
  - 6a) finally signaling request will reach destination ATM-DTE -> VC ok



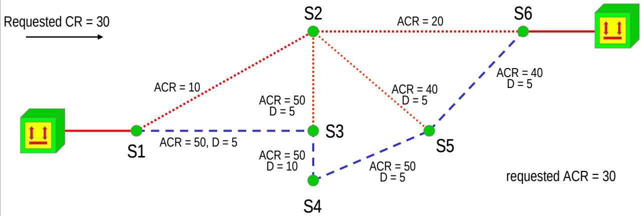
PNNI Routing - Crankbank

- **Operation of the GCAC**
  - 5) After receipt next switch (S2) perform local CAC
    - 5b) if nok -> return PNNI signaling message to originator of DTL
  - 6b) S1 will construct alternate source route



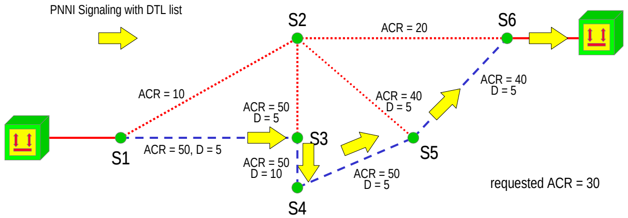
PNNI Routing - New Trial

- **Operation after Crankbank**
  - 7b) The other possible path is chosen - source node constructs again a new Designated Transit List (DTL)



PNNI Routing - Source Routing

- **Operation of the GCAC**
  - 8b) DTL is inserted into signaling request
  - 9b) After receipt next switch perform local CAC
    - if ok -> pass PNNI signaling message on to next switch of DTL
  - 10b) finally signaling request will reach destination ATM-DTE -> VC ok



Agenda

- **Review ATM**
- **IP over WAN Problems (Traditional Approach)**
  - Introduction, Base Problem 1
    - Non-NBMA-View
    - NMBA-View
  - Base Problem 2, Solution
- **MPLS Principles**
- **Label Distribution Methods**
- **MPLS Details (Cisco)**
- **RFCs**

IP Overlay Model - Scalability

- **Base problem Nr.1**
  - IP routing separated from ATM routing because of the normal IP overlay model
  - no exchange of routing information between IP and ATM world
  - leads to scalability and performance problems
    - many peers, configuration overhead, duplicate broadcasts
  - note:
    - IP system requests virtual circuits from the ATM network
    - ATM virtual circuits are established according to PNNI routing
    - virtual circuits are treated by IP as normal point-to-point links
    - IP routing messages are transported via this point-to-point links to discover IP neighbors and IP network topology

IP Performance

- **Base problem Nr.2**
  - IP forwarding is slow compared to ATM cell forwarding
    - IP routing paradigm
    - hop-by-hop routing with (recursive) IP routing table lookup, IP TTL decrement and IP checksum computing
    - destination based routing (large tables in the core of the Internet)
  - Load balancing
    - in a stable network all IP datagram's will follow the same path (least cost routing versus ATM´s QoS routing)
  - QoS (Quality of Service)
    - IP is connectionless packet switching (best-effort delivery versus ATM´s guarantees)
  - VPN (Virtual Private Networks)
    - ATM VC´s have a natural closed user group (=VPN) behavior

Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

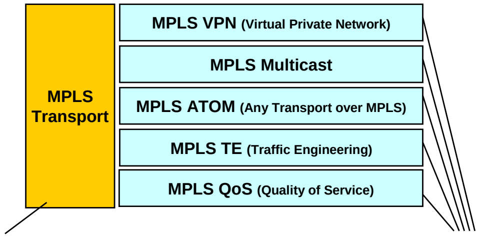
Basic Ideas to Solve the Problems

- **Make ATM topology visible to IP routing**
  - to solve the scalability problems
  - a classical ATM switch gets IP router functionality
- **Divide IP routing from IP forwarding**
  - to solve the performance problems
  - IP forwarding based on ATM´s label swapping paradigm (connection-oriented packet switching)
  - IP routing based on classical IP routing protocols
- **Combine best of both**
  - forwarding based on ATM label swapping paradigm
  - routing done by traditional IP routing protocols

MPLS

- **Several similar technologies were invented in the mid-1990s**
  - IP Switching (Ipsilon)
  - Cell Switching Router (CSR, Toshiba)
  - Tag Switching (Cisco)
  - Aggregated Route-Based IP Switching (ARIS, IBM)
- **IETF merges these technologies**
  - MPLS (Multi Protocol Label Switching)
    - note: multiprotocol means that IP is just one possible protocol to be transported by a MPLS switched network
  - RFC 3031

MPLS Building Blocks



You always need this!
MPLS Transport solves most of the mentioned problems (scalability / performance)

If you need "Advanced Features" like VPN or Multicast support you optionally may choose from these building blocks riding on top of a MPLS Transport network

Agenda

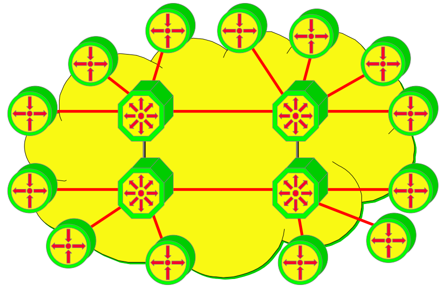
- Review ATM
- IP over WAN Problems (Traditional Approach)
  - Introduction, Base Problem 1
    - Non-NBMA-View
    - NMBA-View
  - Base Problem 2, Solution
- MPLS Principles
- Label Distribution Methods
- MPLS Details (Cisco)
- RFCs

Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

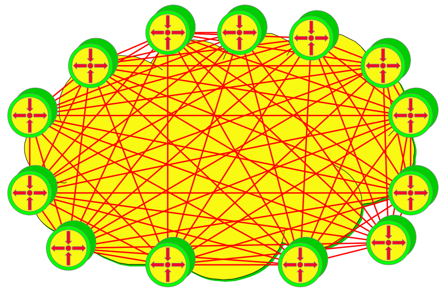
A Simple Physical Network ...

Physical wiring

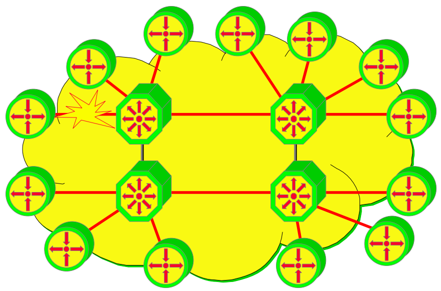


IP Data Link View (Non-NBMA)

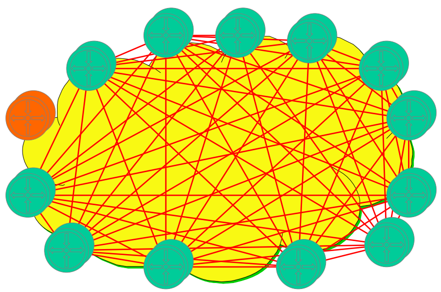
Every virtual circuit has its own IP Net-ID (subinterface technique)



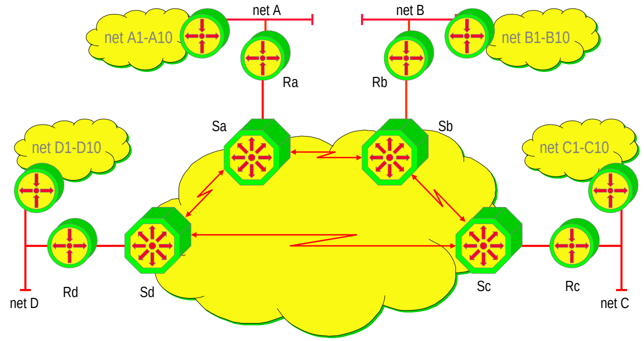
A Single Network Failure ...



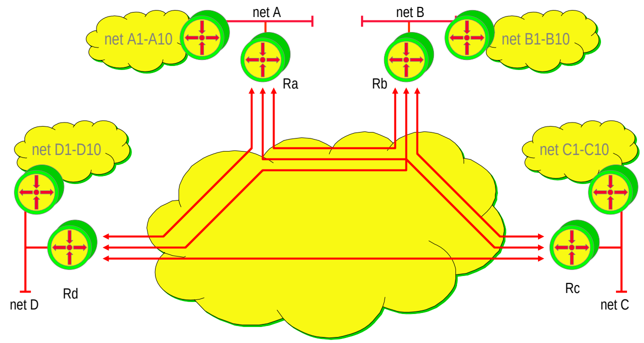
Causes Loss of Multiple IP Router Peers !!!



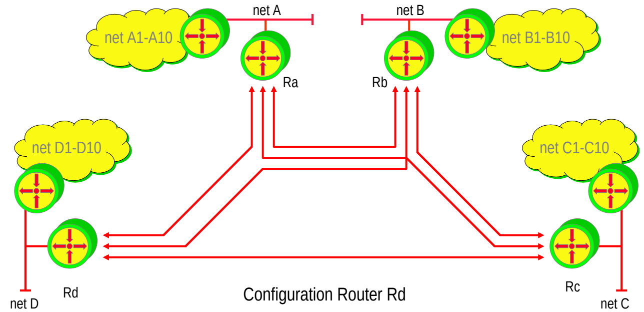
Example - Physical Topology



IP Connectivity through Full-mesh VC´s



Static Routing/No Routing Broadcasts



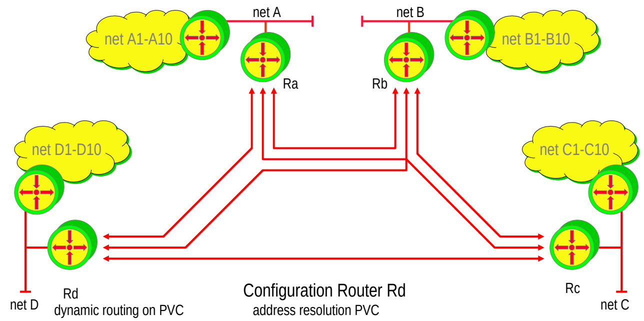
Configuration Router Rd

static routing
net A via next hopRa
net B via next hopRb
net C via next hopRc
every remote network listed here!

address resolution PVC
Ra map VPI/VCI Rd ⇨ Ra
Rb map VPI/VCI Rd ⇨ Rb
Rc map VPI/VCI Rd ⇨ Rc

address resolution SVC
Ra map ATM addr. Ra
Rb map ATM addr. Rb
Rc map ATM addr. Rc

Dynamic Routing/Routing Broadcasts



Configuration Router Rd

dynamic routing on PVC
VPI/VCI Rd ⇨ Ra broadcast
VPI/VCI Rd ⇨ Rb broadcast
VPI/VCI Rd ⇨ Rc broadcast

address resolution PVC
Ra map VPI/VCI Rd ⇨ Ra
Rb map VPI/VCI Rd ⇨ Rb
Rc map VPI/VCI Rd ⇨ Rc

note: SVCs may be possible if Cisco neighbor command is specified for Cisco routing process because no automatic neighbor discovery is possible in this case

Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

Observations

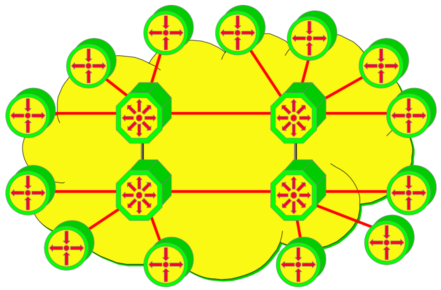
- **This clearly does not scale**
- **Switch/router interaction needed**
  - peering model
- **Without MPLS**
  - Only outside routers are layer 3 neighbors
  - one ATM link failure causes multiple peer failures
  - routing traffic does not scale (number of peers)
- **With MPLS**
  - Inside MPLS switch is the layer 3 routing peer of an outside router
  - one ATM link failure causes one peer failure
  - highly improved routing traffic scalability

Agenda

- **Review ATM**
- **IP over WAN Problems (Traditional Approach)**
  - Introduction, Base Problem 1
    - Non-NBMA-View
    - NMBA-View
  - Base Problem 2, Solution
- **MPLS Principles**
- **Label Distribution Methods**
- **MPLS Details (Cisco)**
- **RFCs**

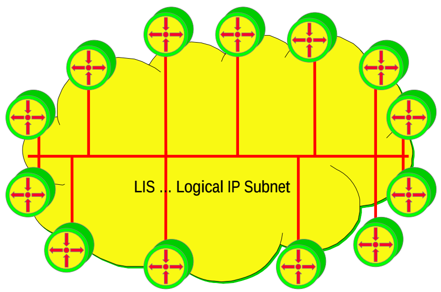
A Simple Physical Network ...

Physical wiring and NBMA behavior



IP Data Link View (NBMA)

Routers assume a LAN behavior because all interfaces have the same IP Net-ID but LAN broadcasting to reach all others is not possible

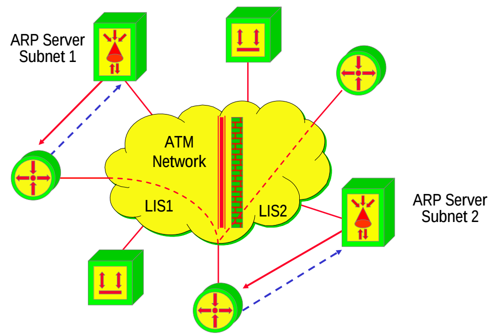


Some Solutions for the NBMA Problem

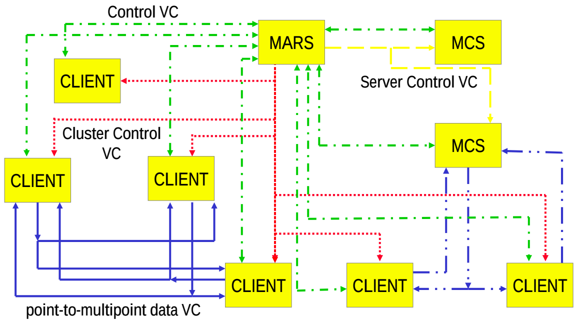
- ARP (Address Resolution Protocol) Server
  - keeps configuration overhead for address resolution small
  - but does not solve the routing issue (neighbor discovery and duplicate routing broadcasts on a single wire)
- MARS/MCS (Multicast Address Resolution Server / Multicast Server)
  - additional keeps configuration overhead for routing small
  - and does solve broadcast/multicast problem with either full mesh of point-to-multipoint circuits or by usage of MCS server
- LANE (LAN Emulation = ATM VLAN´s)
  - simulates LAN behavior where address resolution and routing broadcasts are not a problem
- All of them
  - require a lot of control virtual circuits (p-t-p and p-t-m) and SVC support of the underlying ATM network

RFC 2225 Operation (Classical IP over ATM)

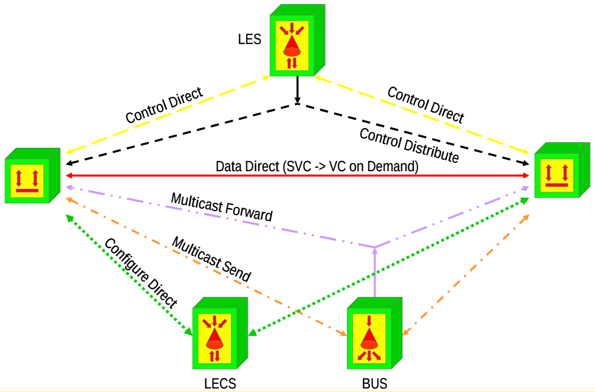
- **ARP server for every LIS**
  - multiple hops for communication between Logical IP Subnets



MARS/MCS Architecture



LANE Connections



Datenkommunikation 384.081 - SS 2012

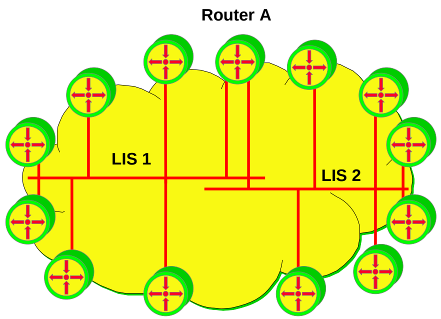
Appendix 3 - MPLS (v4.6)

Scalability Aspects

- **Number of IP peers determines**
  - number of data virtual circuits
  - number of control virtual circuits
  - number of duplicate broadcasts on a single wire
- **Method to solve the broadcast domain problem**
  - split the network in several LIS (logical IP subnets)
  - connect LIS's by normal IP router (ATM-DCE) which is of course outside the ATM network
- **But then another problem arise**
  - traffic between to two systems which both are attached to the ATM network but belong to different LIS´s must leave the ATM network and enter it again at the connecting IP router (-> SAR delay)

IP Multiple LIS´s in case of ROLC (Routing Over Large Clouds)

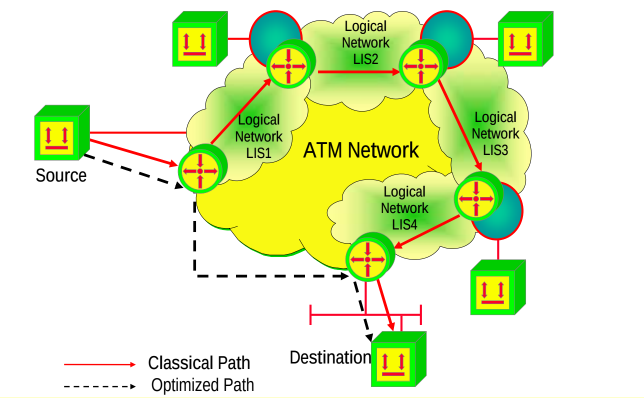
IP router A connects LIS1 and LIS2



Some Solutions for the ROLC Problem

- **NHRP (Next Hop Resolution Protocol)**
  - creates an ATM shortcut between two systems of different LIS´s
- **MPOA (Multi Protocol Over ATM)**
  - LANE + NHRP combined
  - creates an ATM shortcut between two systems of different LIS´s
- **In both methods**
  - the ATM shortcut is created if traffic between the two systems exceeds a certain threshold -> data-flow driven
  - a lot of control virtual circuits (p-t-p and p-t-m) is required

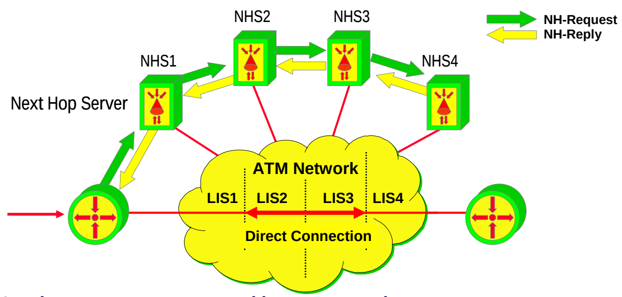
Wish for Optimized Connectivity



Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

Next Hop Resolution Protocol (RFC 2332)



- **Next hop requests are passed between next hop servers**
  - Next hop servers do not forward data
- **NHS that knows about the destination sends back a NH-reply**
  - Allows direct connection between logical IP subnets across the ATM cloud
  - Separates data forwarding path from reachability information

Agenda

- **Review ATM**
- **IP over WAN Problems (Traditional Approach)**
  - Introduction, Base Problem 1
    - Non-NBMA-View
    - NMBA-View
  - Base Problem 2, Solution
- **MPLS Principles**
- **Label Distribution Methods**
- **MPLS Details (Cisco)**
- **RFCs**

IP Performance

- **Base problem Nr.2**
  - IP forwarding is slow compared to ATM cell forwarding
    - IP routing paradigm
    - hop-by-hop routing with (recursive) IP routing table lookup, IP TTL decrement and IP checksum computing
    - destination based routing (large tables in the core of the Internet)
  - Load balancing
    - in a stable network all IP datagram's will follow the same path (least cost routing versus ATM´s QoS routing)
  - QoS (Quality of Service)
    - IP is connectionless packet switching (best-effort delivery versus ATM´s guarantees)
  - VPN (Virtual Private Networks)
    - ATM VC´s have a natural closed user group (=VPN) behavior

Basic Ideas to Solve the Problems

- **Make ATM topology visible to IP routing**
  - to solve the scalability problems
  - a classical ATM switch gets IP router functionality
- **Divide IP routing from IP forwarding**
  - to solve the performance problems
  - IP forwarding based on ATM´s label swapping paradigm (connection-oriented packet switching)
  - IP routing based on classical IP routing protocols
- **Combine best of both**
  - forwarding based on ATM label swapping paradigm
  - routing done by traditional IP routing protocols

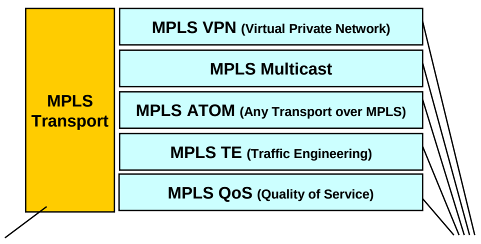
Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

MPLS

- **Several similar technologies were invented in the mid-1990s**
  - IP Switching (Ipsilon)
  - Cell Switching Router (CSR, Toshiba)
  - Tag Switching (Cisco)
  - Aggregated Route-Based IP Switching (ARIS, IBM)
- **IETF merges these technologies**
  - MPLS (Multi Protocol Label Switching)
    - note: multiprotocol means that IP is just one possible protocol to be transported by a MPLS switched network
  - RFC 3031

MPLS Building Blocks



You always need this! MPLS Transport solves most of the mentioned problems (scalability / performance)

If you need "Advanced Features" like VPN or Multicast support you optionally may choose from these building blocks riding on top of a MPLS Transport network

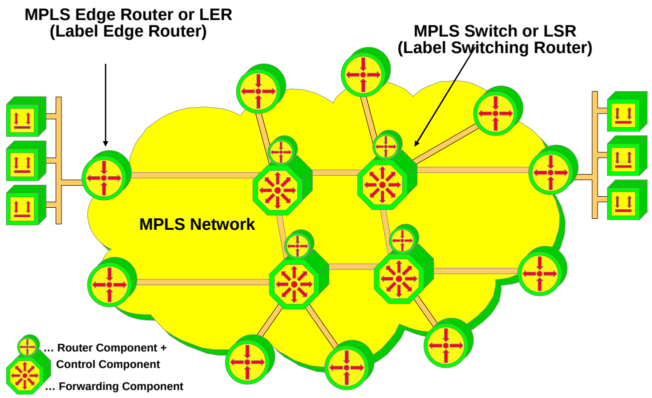
Agenda

- Review ATM
- IP over WAN Problems (Traditional Approach)
- <u>MPLS Principles</u>
- Label Distribution Methods
- MPLS Details (Cisco)
- RFCs

MPLS Approach

- **Traditional IP uses the same information for**
  - path determination (routing)
  - packet forwarding (switching)
- **MPLS separates the tasks**
  - L3 addresses used for path determination
  - <u>labels</u> used for switching
- **MPLS Network consists of**
  - MPLS Edge Routers and MPLS Switches
- **MPLS Edge Routers and MPLS Switches**
  - exchange routing information about L3 IP networks
  - exchange forwarding information about the actual usage of labels

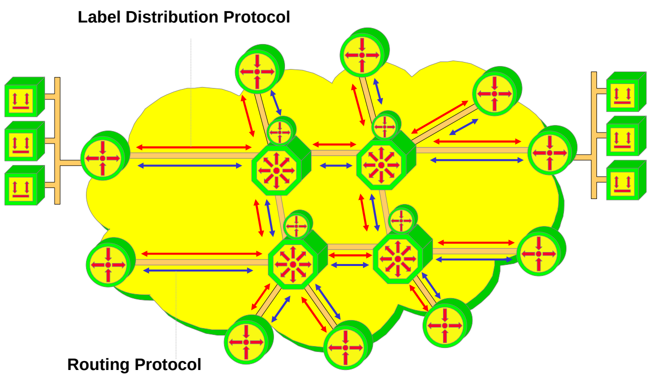
MPLS Network



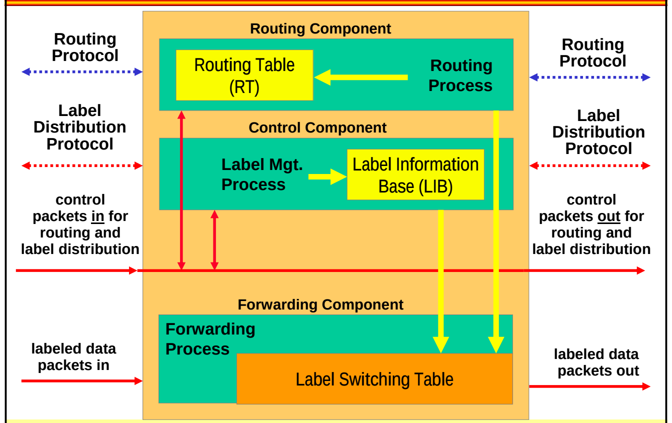
MPLS LSR Internal Components

- **Routing Component**
  - still accomplished by using standard IP routing protocols creating <u>routing table</u>
- **Control Component**
  - maintains correct label distribution among a group of label switches
  - <u>Label Distribution Protocol</u> for communication
    - between MPLS Switches
    - between MPLS Switch and MPLS Edge Router
- **Forwarding Component**
  - uses labels carried by packets plus label information maintained by a label switch (classical VC <u>switching table</u>) to perform packet forwarding -> <u>"label swapping"</u>

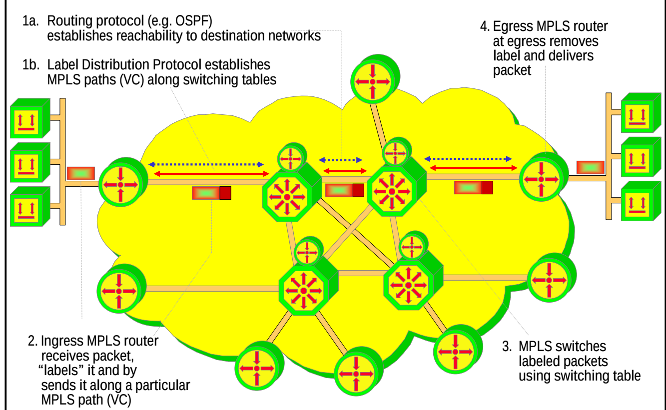
MPLS Control Communication



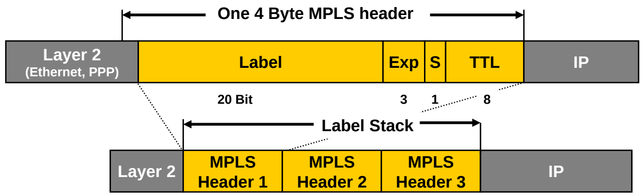
Generic Overview of MPLS LSR Internal Processes and Communication



MPLS Basic Operations



MPLS Header: Frame Mode



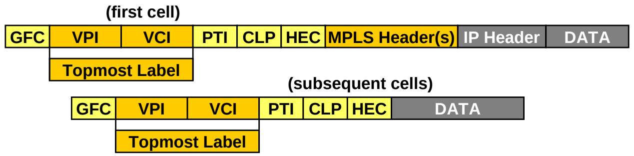
- **"Layer 2.5" can be used over Ethernet, 802.3 or PPP links**
  - note: 2.5 means 32 bit
  - 20-bit MPLS label (Label)
  - 3-bit experimental field (Exp)
    - could be copy of IP Precedence -> MPLS QoS like IP QoS with DiffServ Model based on DSCP
  - 1-bit bottom-of-stack indicator (S)
    - Labels could be stacked (Push & Pop)
    - MPLS switching performed always on the first label of the stack
  - 8-bit time-to-live field (TTL)

MPLS Header: Cell Mode

ATM Convergence Sublayer (CS):

- **ATM Switches can only switch VPI/VCI—no MPLS labels!**
  - Only the topmost label is inserted in the VPI/VCI field

ATM Segmentation and Reassembling Sublayer (SAR):



Labels and FEC

- **A label is used to identify a certain subset of packets**
  - which take the same MPLS path or which get the same forwarding treatment in the MPLS label switched network
  - The path is so called Label Switched Path (LSP)
    - "The MPLS Virtual Circuit"
- **Thus a label represents**
  - a so called Forwarding Equivalence Class (FEC)
- **The assignment of a packet to FEC**
  - is done just once by the MPLS Edge Router, as the packet enters the network
  - most commonly this is based on the IP network layer destination address

Label Binding

- **Two neighboring LSRs R1 and R2**
  - may agree that when R1 transmits a packet to R2, R1 will label with packet with label value L if and only if the packet is a member of a particular FEC F
- **They agree**
  - on a so called "binding" between label L and FEC F for packets moving from R1 to R2
- **As a result**
  - L becomes R1´s "outgoing label" or "remote label" representing FEC F
  - L becomes R2´s "incoming label" or "local label" representing FEC F

Creating and Destroying Label Binding 1

- **Control Driven (favored by IETF-WG)**
  - creation or deconstruction of labels is triggered by control information such as
    - OSPF routing, IS-IS routing
    - PIM Join/Prune messages in case of IP multicast routing
    - IntSrv RSVP messages in case of IP QoS IntSrv Model
    - DiffSrv Traffic Engineering in Case of IP QoS DiffSrv Model
  - hence we have a pre-assignment of labels based on reachability information
    - and optionally based on QoS needs
  - also called Topology Driven

Creating and Destroying Label Binding 2

- Data Driven
  - creation or deconstruction of labels is triggered by data packets
    - but only if a critical threshold number of packets for a specific communication relationship is reached
    - may have a big performance impact
  - hence we have dynamic assignment of labels based on data flow detection
  - also called Traffic Driven

Some FEC Examples for Topology Driven

- FECs could be for example
  - a set of unicast packets whose network layer destination address matches a particular IP address prefix
    - MPLS application: Destination Based (Unicast) Routing
  - a set of multicast packets with the same source and destination network layer address
    - MPLS application: Multicast Routing
  - a set of unicast packets whose network layer destination address matches a particular IP address prefix and whose Type of Service (ToS) or DSCP bits are the same
    - MPLS application: Quality of Service
    - MPLS application: Traffic Engineering or Constraint Based Routing

Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

Label Distribution

- **MPLS architecture allows an LSR to distribute bindings to LSRs that have not explicitly requested them**
  - “Unsolicited Downstream" label distribution
  - usually used by Frame-Mode MPLS
- **MPLS architecture allows an LSR to explicitly request, from its next hop for a particular FEC, a label binding for that FEC**
  - “Downstream-On-Demand" label distribution
  - must be used by Cell-Mode MPLS

Label Binding

- **The decision to bind a particular label L to a particular FEC F**
  - is made by the LSR which is DOWNSTREAM with respect to that binding
  - the downstream LSR then informs the upstream LSR of the binding
  - thus labels are "downstream-assigned“
  - thus label bindings are distributed in the "downstream to upstream“ direction
- **Discussion were about if**
  - labels should also be “upstream-assigned“
  - not any longer part of current MPLS-RFC

Label Retention Mode 1

- **A LSR may receive a label binding**
  - for a particular FEC from another LSR, which is not next hop based on the routing table for that FEC
- **This LSR then has the choice**
  - of whether to keep track of such bindings, or whether to discard such bindings
- **A LSR supports "Liberal Label Retention Mode"**
  - if it maintains the bindings between a label and a FEC which are received from LSR´s which are not its next hop for that FEC

Label Retention Mode 2

- **A LSR supports "Conservative Label Retention mode "**
  - If it discards the bindings between a label and a FEC which are received from LSR´s which are not its next hop for that FEC
- **Liberal Label Retention mode**
  - allows for quicker adaptation to routing changes
  - LSR can switch over to next best LSP
- **Conservative Label Retention mode**
  - requires an LSR to maintain fewer labels
  - LSR has to wait for new label bindings in case of topology changes

Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

Independent versus Ordered Control

- **Independent Control:**
  - each LSR may make an independent decision to assign a a label to a FEC and to advertise the assignment to its neighbors
  - typically used in Frame-Mode MPLS for destination based routing
  - loop prevention must be done by other means (-> MPLS TTL) but there is faster convergence
- **Ordered Control:**
  - label assignment proceeds in an orderly fashion from one end of a LSP to the other
  - under ordered control, LSP setup may be initiated by the ingress (header) or egress (tail) MPLS Edge Router

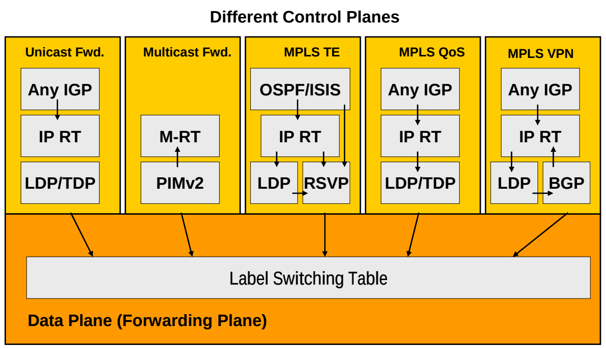
Ordered Control - Egress

- in case of egress method the only LSR which can initiate the process of label assignment is the egress LSR
- a LSR knows that it is the egress for a given FEC if its next hop for this FEC is not an LSR
- this LSR will sent a label advertisement to all neighboring LSRs
- a neighboring LSR receiving such a label advertisement from a interface which is the next hop to a given FEC will assign its own label and advertise it to all other neighboring LSRs
- inherent loop prevention
- slower convergence

Ordered Control - Ingress

- in case of ingress method the LSR which initiates the process of label assignment is the ingress LSR
- the ingress LSR constructs a source route and pass on requests for label bindings to the next LSR
- this is done until LSR which is the end of the source route is reached
- from this LSR label bindings will flow upstream to the ingress LSR
- used for MPLS Traffic Engineering (TE)

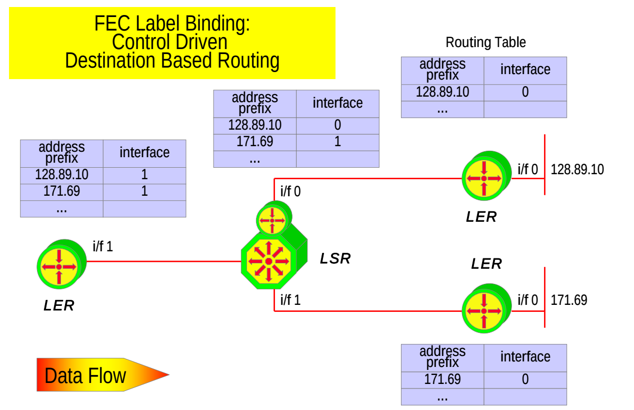
MPLS Applications and MPLS Control Plane



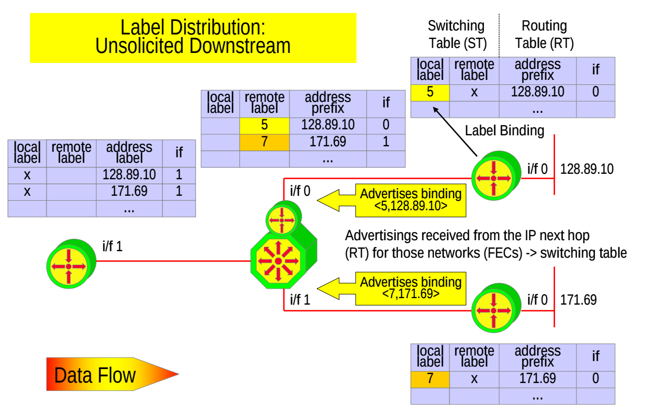
Agenda

- **Review ATM**
- **IP over WAN Problems (Traditional Approach)**
- **MPLS Principles**
- **Label Distribution Methods**
  - Unsolicited Downstream
  - Downstream On Demand
  - MPLS and ATM, VC Merge Problem
- **MPLS Details (Cisco)**
- **RFCs**

Routing Table Created by Routing Protocol



Labels Sent by LDP



local label	remote label	address label	if
x		128.89.10	1
x		171.69	1
		...	

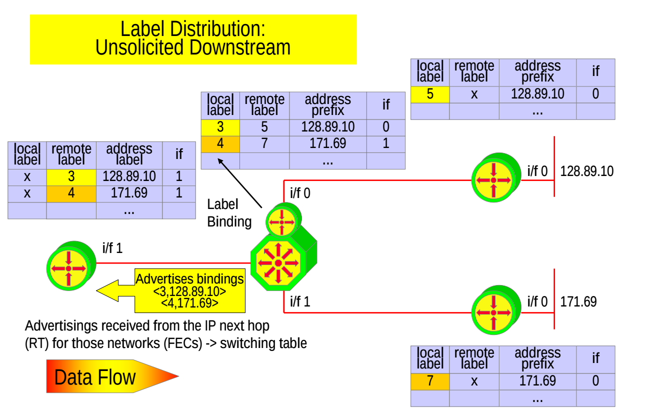
local label	remote label	address prefix	if
	5	128.89.10	0
	7	171.69	1
		...	

local label	remote label	address prefix	if
5	x	128.89.10	0
		...	

local label	remote label	address prefix	if
7	x	171.69	0
		...	

© 2012, D.I. Lindner MPLS v4.6 85

Labels Sent and Switching Table Entry Created by MPLS Switch



local label	remote label	address label	if
x	3	128.89.10	1
x	4	171.69	1
		...	

local label	remote label	address prefix	if
3	5	128.89.10	0
4	7	171.69	1
		...	

local label	remote label	address prefix	if
5	x	128.89.10	0
		...	

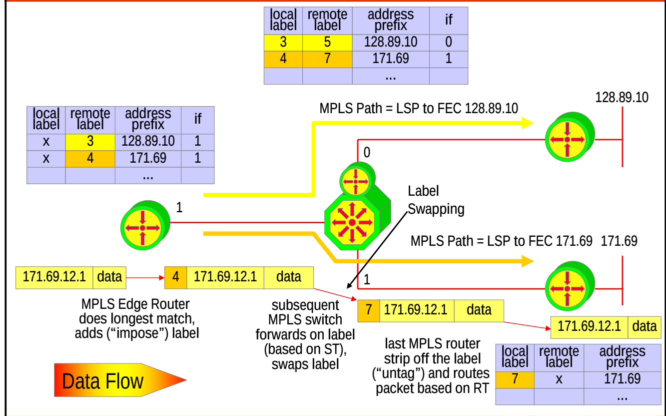
local label	remote label	address prefix	if
7	x	171.69	0
		...	

© 2012, D.I. Lindner MPLS v4.6 86

Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

MPLS Switched Packets



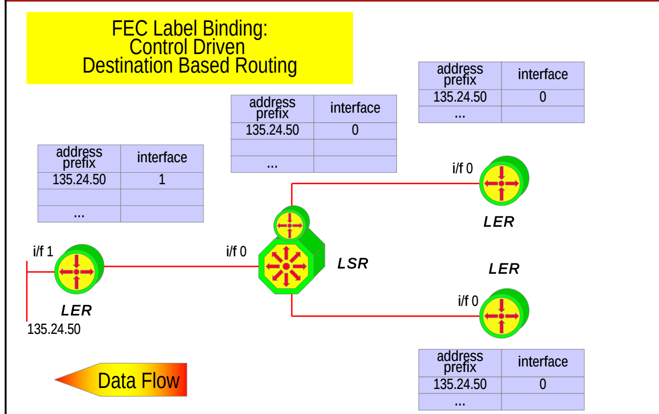
local label	remote label	address prefix	if
3	5	128.89.10	0
4	7	171.69	1
		...	

local label	remote label	address prefix	if
x	3	128.89.10	1
x	4	171.69	1
		...	

local label	remote label	address prefix	if
7	x	171.69	
		...	

© 2012, D.I. Lindner MPLS v4.6 87

Routing Table Created by Routing Protocol



address prefix	interface
135.24.50	1
...	

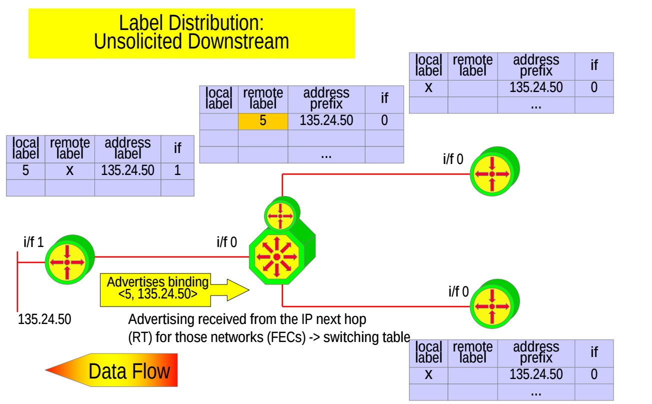
address prefix	interface
135.24.50	0
...	

address prefix	interface
135.24.50	0
...	

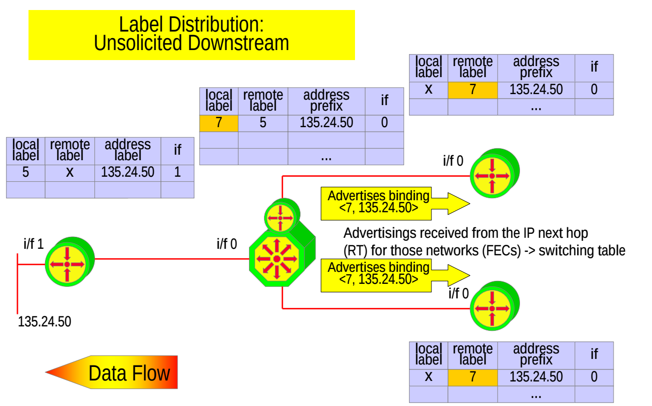
address prefix	interface
135.24.50	0
...	

© 2012, D.I. Lindner MPLS v4.6 88

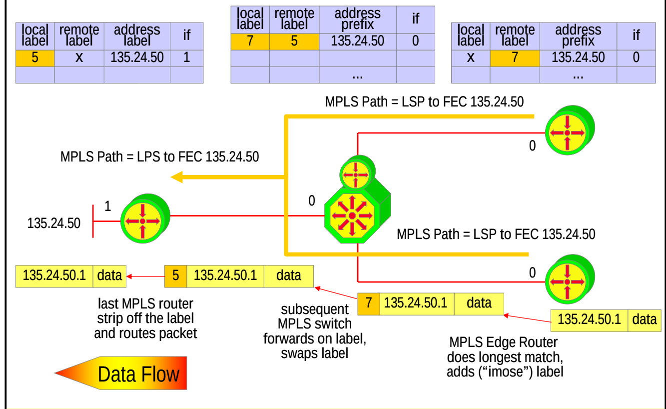
Labels Sent by LDP



Labels Sent and Switching Table Entry Created by MPLS Switch

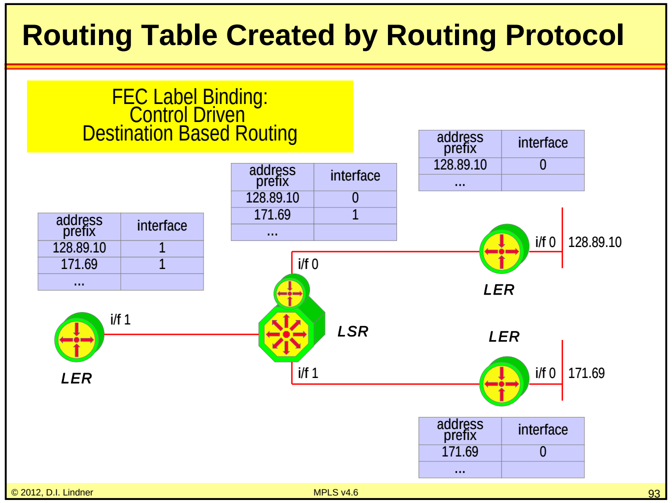


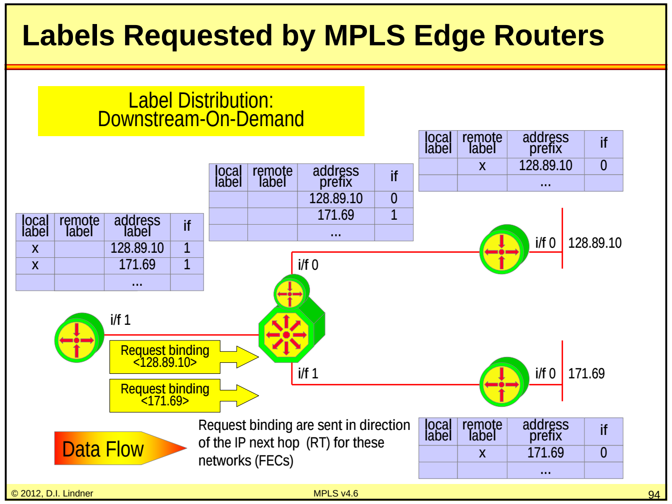
Label Merging - LSP Merging

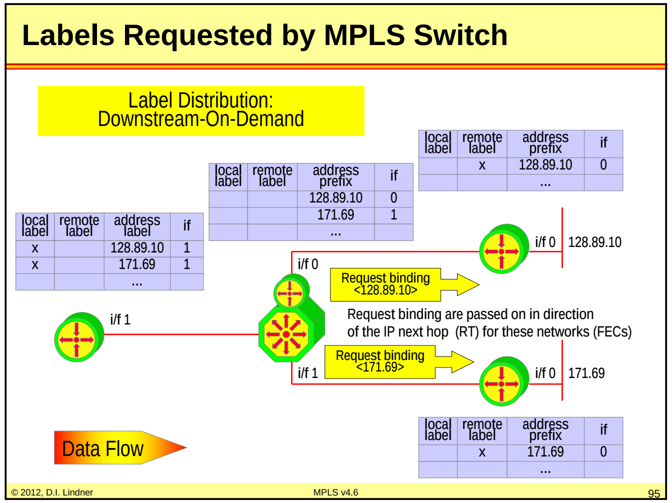


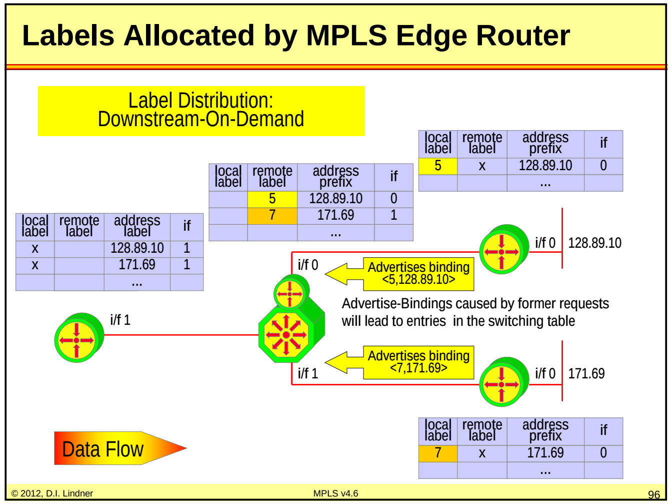
Agenda

- **Review ATM**
- **IP over WAN Problems (Traditional Approach)**
- **MPLS Principles**
- **Label Distribution Methods**
  - Unsolicited Downstream
  - Downstream On Demand
  - MPLS and ATM, VC Merge Problem
- **MPLS Details (Cisco)**
- **RFCs**

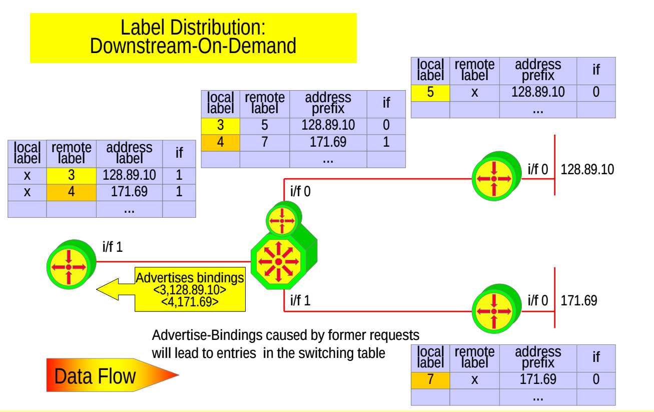

Routing Table Created by Routing Protocol
FEC Label Binding:
Control Driven
Destination Based Routing
address prefix | interface
128.89.10 | 1
171.69 | 1
...
address prefix | interface
128.89.10 | 0
171.69 | 1
...
address prefix | interface
128.89.10 | 0
...
address prefix | interface
171.69 | 0
...
i/f 1
i/f 0
i/f 1
i/f 0
i/f 0
128.89.10
171.69
LER
LSR
LER
LER
© 2012, D.I. Lindner
MPLS v4.6
93


Labels Requested by MPLS Edge Routers
Label Distribution:
Downstream-On-Demand
local label | remote label | address label | if
x | | 128.89.10 | 1
x | | 171.69 | 1
...
local label | remote label | address prefix | if
| | 128.89.10 | 0
| | 171.69 | 1
...
local label | remote label | address prefix | if
| x | 128.89.10 | 0
...
local label | remote label | address prefix | if
| x | 171.69 | 0
...
i/f 1
i/f 0
i/f 1
i/f 0
i/f 0
128.89.10
171.69
Request binding <128.89.10>
Request binding <171.69>
Request binding are sent in direction of the IP next hop (RT) for these networks (FECs)
Data Flow
© 2012, D.I. Lindner
MPLS v4.6
94


Labels Requested by MPLS Switch
Label Distribution:
Downstream-On-Demand
local label | remote label | address label | if
x | | 128.89.10 | 1
x | | 171.69 | 1
...
local label | remote label | address prefix | if
| | 128.89.10 | 0
| | 171.69 | 1
...
local label | remote label | address prefix | if
| x | 128.89.10 | 0
...
local label | remote label | address prefix | if
| x | 171.69 | 0
...
i/f 1
i/f 0
i/f 1
i/f 0
i/f 0
128.89.10
171.69
Request binding <128.89.10>
Request binding <171.69>
Request binding are passed on in direction of the IP next hop (RT) for these networks (FECs)
Data Flow
© 2012, D.I. Lindner
MPLS v4.6
95


Labels Allocated by MPLS Edge Router
Label Distribution:
Downstream-On-Demand
local label | remote label | address label | if
x | | 128.89.10 | 1
x | | 171.69 | 1
...
local label | remote label | address prefix | if
| 5 | 128.89.10 | 0
| 7 | 171.69 | 1
...
local label | remote label | address prefix | if
5 | x | 128.89.10 | 0
...
local label | remote label | address prefix | if
7 | x | 171.69 | 0
...
i/f 1
i/f 0
i/f 1
i/f 0
i/f 0
128.89.10
171.69
Advertises binding <5,128.89.10>
Advertises binding <7,171.69>
Advertise-Bindings caused by former requests will lead to entries in the switching table
Data Flow
© 2012, D.I. Lindner
MPLS v4.6
96

Labels Allocated and Switching Table Built by MPLS Switch



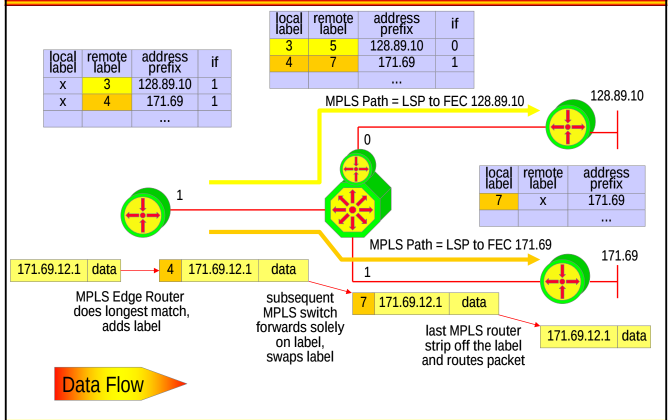
local label	remote label	address label	if
x	3	128.89.10	1
x	4	171.69	1
		...	

local label	remote label	address prefix	if
3	5	128.89.10	0
4	7	171.69	1
		...	

local label	remote label	address prefix	if
5	x	128.89.10	0
		...	

local label	remote label	address prefix	if
7	x	171.69	0
		...	

MPLS Switched Packets

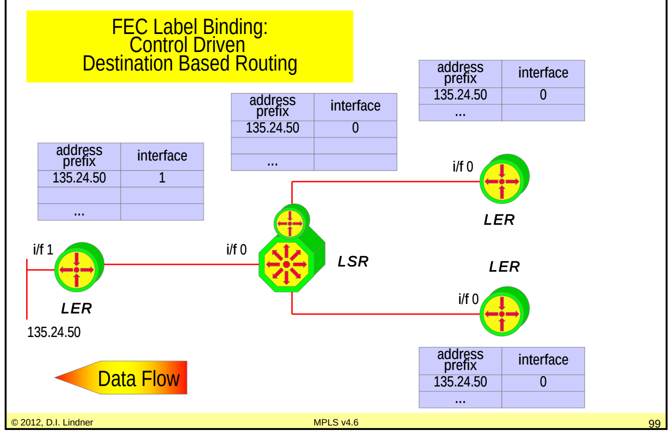


local label	remote label	address prefix	if
x	3	128.89.10	1
x	4	171.69	1
		...	

local label	remote label	address prefix	if
3	5	128.89.10	0
4	7	171.69	1
		...	

local label	remote label	address prefix
7	x	171.69
		...

Routing Table Created by Routing Protocol



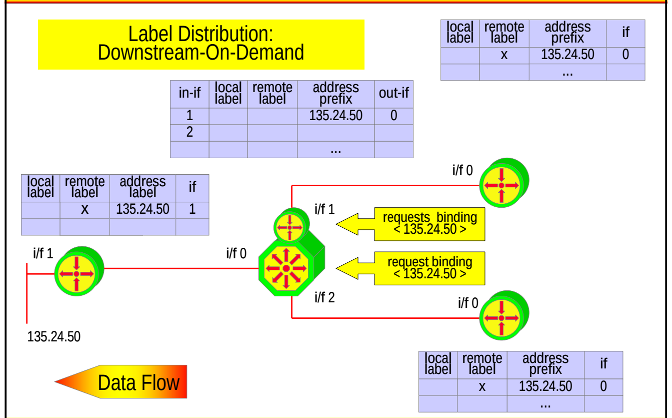
address prefix	interface
135.24.50	1
...	

address prefix	interface
135.24.50	0
...	

address prefix	interface
135.24.50	0
...	

address prefix	interface
135.24.50	0
...	

Labels Requested by MPLS Edge Routers

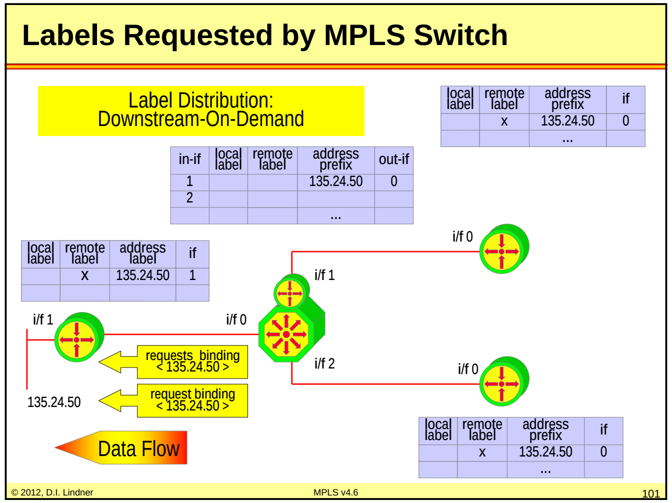


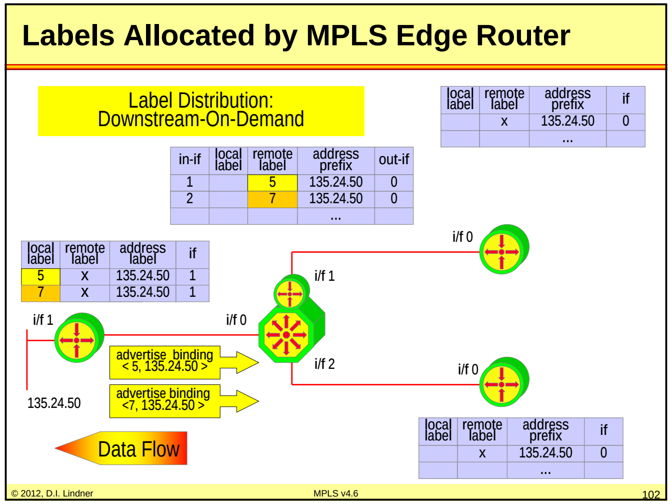
local label	remote label	address prefix	if
	x	135.24.50	0
		...	

in-if	local label	remote label	address prefix	out-if
1			135.24.50	0
2				
			...	

local label	remote label	address prefix	if
	x	135.24.50	1

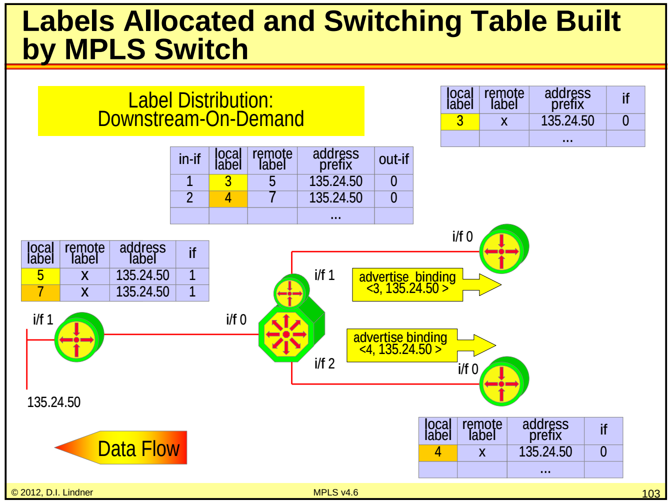
local label	remote label	address prefix	if
	x	135.24.50	0
		...	

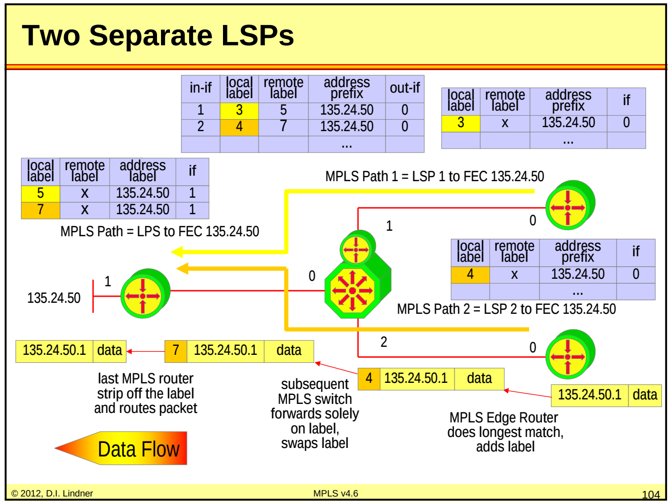




Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)





Agenda

- **Review ATM**
- **IP over WAN Problems (Traditional Approach)**
- **MPLS Principles**
- **Label Distribution Methods**
  - Unsolicited Downstream
  - Downstream On Demand
  - MPLS and ATM, VC Merge Problem
- **MPLS Details (Cisco)**
- **RFCs**

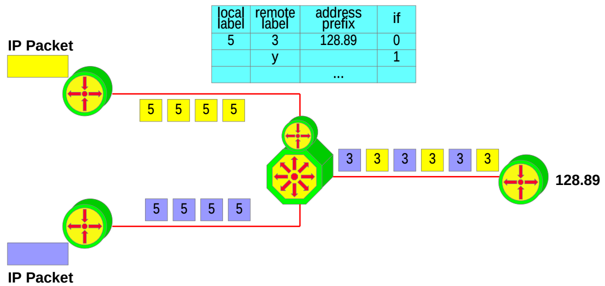
Label Switching and ATM

- **Can be easily deployed with ATM because ATM uses label swapping**
  - VPI/VCI is used as a label
- **ATM switches needs to implement control component of label switching**
  - ATM attached router peers with ATM switch (label switch)
    - exchange label binding information
- **Differences**
  - how labels are set up
    - label distribution -> downstream on demand allocation
  - label merging
    - in order to scale, merging of multiple streams (labels) into one stream (label) is required

Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

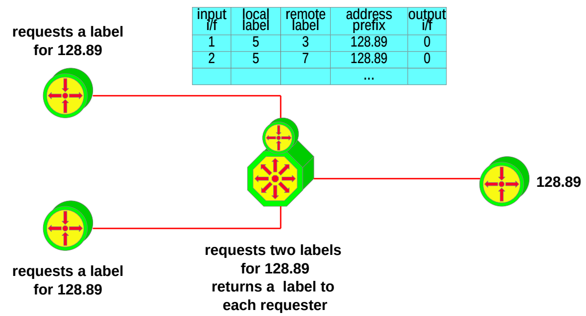
Label Switching and ATM



local label	remote label	address prefix	if
5	3	128.89	0
	y		1
		...	

ATM switch interleaves cells of different packets onto same label.
That is a problem in case of AAL5 encapsulation.
No problem in case of AAL3/AAL4 encapsulation because of AAL3/AAL4´s inherent multiplexing capability.

Label Distribution Solution for ATM

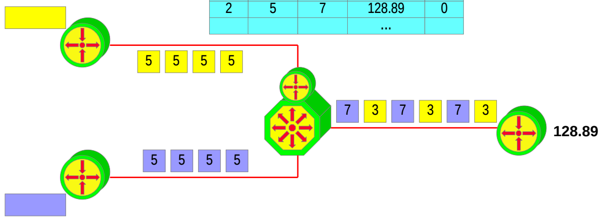


input i/f	local label	remote label	address prefix	output i/f
1	5	3	128.89	0
2	5	7	128.89	0
			...	

- **"Downstream On Demand" Label Distribution**

Label Distribution Solution for ATM

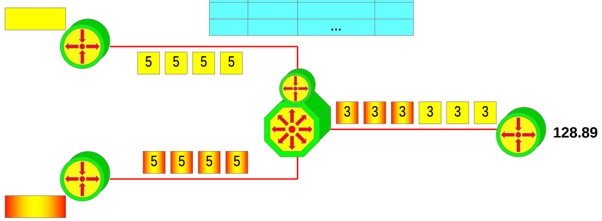
input i/f	local label	remote label	address prefix	output i/f
1	5	3	128.89	0
2	5	7	128.89	0
			...	



- **Downstream On Demand label distribution is necessary**
  - multiple labels per FEC may be assigned
  - one label per (ingress, egress) router pair
- **Label space can be reduced with VC-merge technique**

VC Merge Technique

local label	remote label	address prefix	if
5	3	128.89	0
		...	



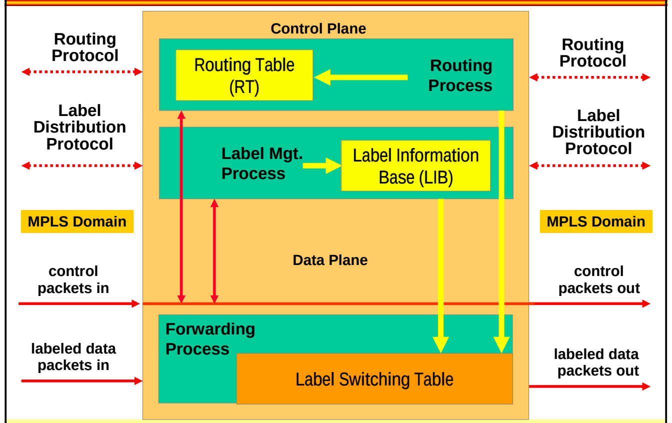
- **ATM switch avoids interleaving of frames**
  - VC Merge technique
  - looking for AAL5 trailers and storing corresponding cells of a frame until AAL5 trailer is seen

Datenkommunikation 384.081 - SS 2012

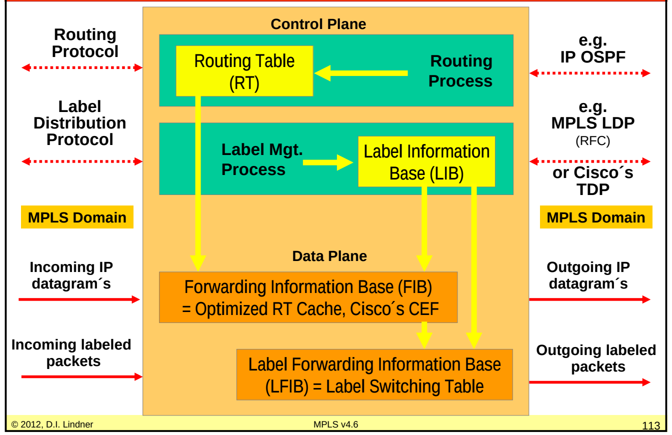
Appendix 3 - MPLS (v4.6)

Agenda

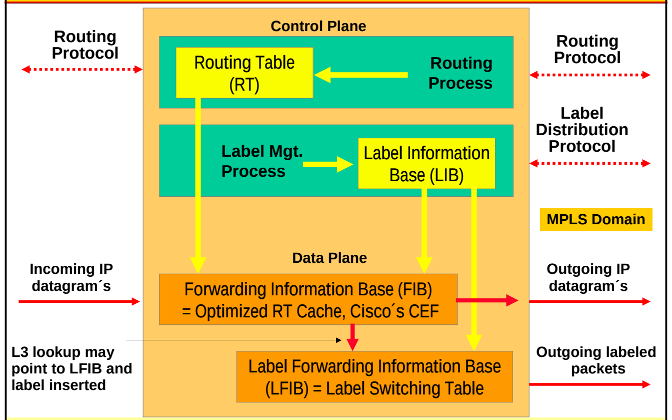
Generic MPLS Control and Data Plane



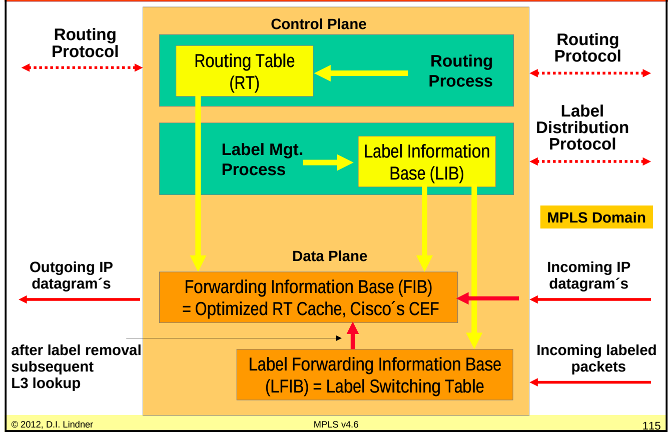
Frame Mode MPLS for IP at LSR (Cisco)



Frame Mode MPLS for IP at Edge (LER) 1



Frame Mode MPLS for IP at Edge (LER) 2



Important Databases

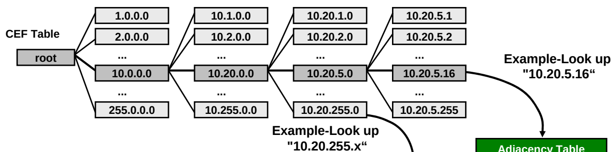
- **FIB**
  - Forwarding Information Base
  - This is the CEF database at Cisco routers
  - Contains L2/L3 headers, IP addresses, labels, next hop, metric
    - The routing table is only a subset of the FIB
- **LIB**
  - Label Information Base
  - Contains all labels and associated destinations
- **LFIB**
  - Label Forwarding Information Base
  - Contains selected labels used for forwarding
    - Selection based on FIB

Cisco Express Forwarding (CEF)

- **Requirement for MPLS**
  - Forwarding information (L2-headers, addresses, labels) are maintained in FIB for each destination
  - Newest and fastest IOS switching method
  - Critical in environments with frequent route changes and large RT's: The Internet backbone!
- **Invented to overcome Fast Switching problems:**
  - Originally Hash table, since 10.2 2-way radix-tree
  - No overlapping cache entries
  - Any change of RT or ARP cache invalidates route cache
  - First packet is always process-switched to build route cache entry
  - Inefficient load balancing when "many hosts to one server"

How CEF Works

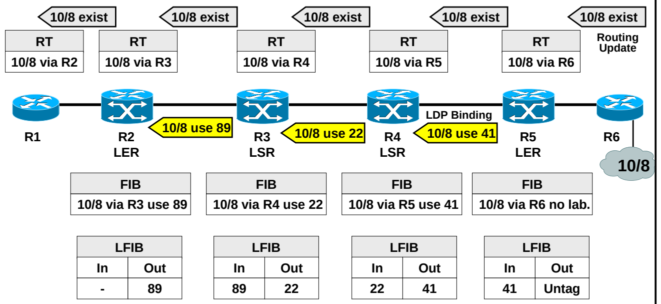
- CEF "Fast Cache" consists of
  - CEF table: Stripped-down version of the RT (256-way mtrie data structure)
  - Adjacency table: Actual forwarding information (MAC, interfaces, ...)
- CEF cache is pre-built before any packets are switched
  - No packet needs to be process switched
- CEF entries never age out
  - Any RT or ARP changes are immediately mapped into CEF cache



CEF Table is built directly from the RT
Adjacency Table is built directly from the ARP cache in case of LAN

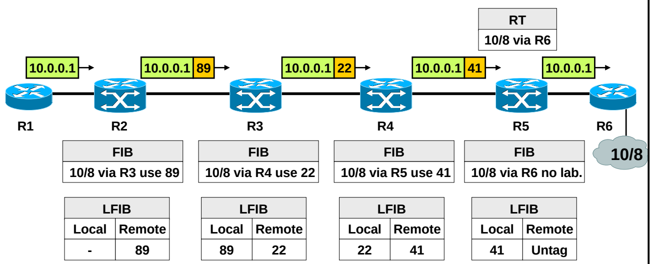
Attention: For an IP-Prefix the pointer to the Adjacency Table will start earlier in the structure

Label Distribution



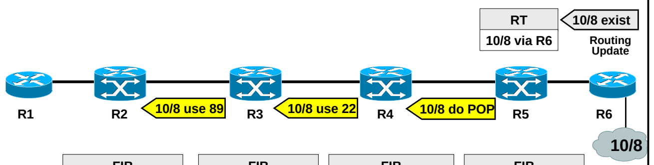
- **Both routing updates and LDP/TDP distribute reachability information**
- **"in" = local label, "out" = remote label**

Label Switching



- **R5 must perform double lookup:**
  - LFIB tells "remove the label"
  - FIB tells "use next hop R6"
- **Label should be removed on hop earlier (by R4) !!!!**

Penultimate Hop Popping 1



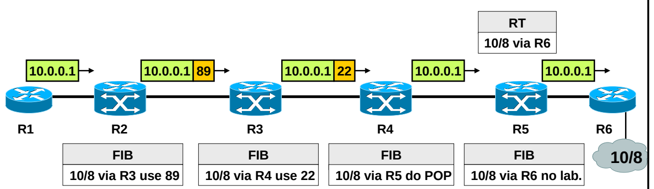
FIB	FIB	FIB	FIB
10/8 via R3 use 89	10/8 via R4 use 22	10/8 via R5 do POP	10/8 via R6 no lab.

LFIB In	LFIB Out	LFIB In	LFIB Out	LFIB In	LFIB Out	LFIB In	LFIB Out
-	89	89	22	22	POP	implicit null	-

- **Last hop router (R5) tells penultimate router (R4) to remove label**
  - "**Penultimate Hop Popping**" (PHP)
  - Also called "Implicit Null Label"

© 2012, D.I. Lindner MPLS v4.6 121

Penultimate Hop Popping 2



FIB	FIB	FIB	FIB
10/8 via R3 use 89	10/8 via R4 use 22	10/8 via R5 do POP	10/8 via R6 no lab.

LFIB In	LFIB Out	LFIB In	LFIB Out	LFIB In	LFIB Out	LFIB In	LFIB Out
-	89	89	22	22	POP	implicit null	-

- **R5 only performs single lookup in FIB**
- **Note: PHP does not work with ATM**
  - VPI/VCI cannot be removed

© 2012, D.I. Lindner MPLS v4.6 122

Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

Cisco IOS Standard Behavior 1

- Routers with packet interfaces (Frame-Mode MPLS)
  - Per-platform Label Space !!!
    - a label assigned by an LSR to a given FEC is used on all interfaces in advertisements of this LSR
  - Unsolicited Downstream Label Distribution
    - label distribution is done unsolicited
  - Liberal Label Retention Mode
    - received labels which are not used by a given LSR are still stored in the LIB
    - allows faster convergence of LSP after rerouting
  - Independent Control
    - labels are assigned by LSR independently from each other

© 2012, D.I. Lindner MPLS v4.6 123

Cisco IOS Standard Behavior 2

- Routers with ATM interfaces (Cell-Mode MPLS)
  - Per-interface Label Space
    - a different label for the same FEC is used on each single interface in advertisements of this LSR
  - Downstream On Demand Label Distribution
    - label distribution is done on request
  - Conservative or Liberal Label Retention Mode
    - received labels which are not used by a given LSR are not stored in the LIB in case of conservative mode
  - Independent Control

© 2012, D.I. Lindner MPLS v4.6 124

Cisco IOS Standard Behavior 3

- ATM switches (Cell-Mode MPLS)
  - Per-interface Label Space
  - Downstream On Demand Label Distribution
  - Conservative Label Retention Mode
  - Ordered control
    - labels are assigned by LSR in a controlled fashion from egress to ingress

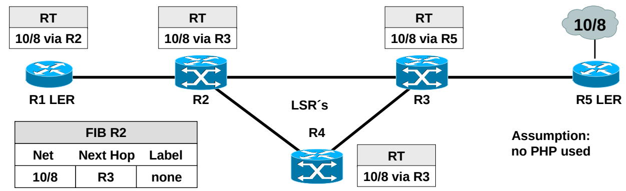
Agenda

- **Review ATM**
- **IP over WAN Problems (Traditional Approach)**
- **MPLS Principles**
- **Label Distribution Methods**
- **MPLS Details (Cisco)**
  - Internal Components
  - MPLS in Action
  - TDP, LDP
  - TTL
  - Traffic Engineering
  - MPLS and BGP
- **RFCs**

Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

Building Routing Tables



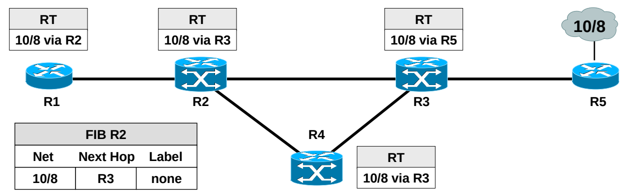
FIB R2		
Net	Next Hop	Label
10/8	R3	none

LIB R2		
Net	Label	Type
-	-	-
-	-	-
-	-	-

LFIB R2		
Label	Action	Next Hop
-	-	-

- **Routing Protocol**
  - establish routing tables RT in all routers
  - best path based on metric is stored in RT
- **RT**
  - contains next hop information (outgoing interface)
- **FIB**
  - additionally contains outgoing label information (which label can be used towards next hop)

Allocating Labels



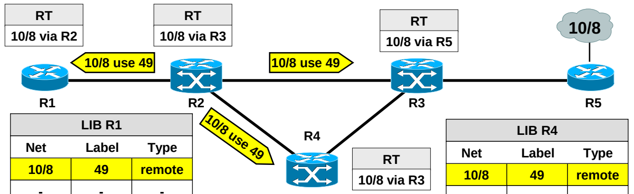
FIB R2		
Net	Next Hop	Label
10/8	R3	none

LIB R2		
Net	Label	Type
10/8	49	local
-	-	-
-	-	-

LFIB R2		
Label	Action	Next Hop
49	untag	-

- **R2**
  - allocates label 49 to FEC 10/8
  - stored in LIB with type local
  - stores action untag in LFIB because no other router has advertised a label for that FEC
- **Every MPLS router**
  - allocates labels for all IP destinations found in the routing table
  - this is done independently from each other
  - a label has only local significance

Advertising and Receiving Labels via LDP



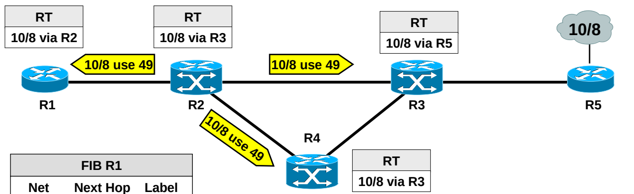
LIB R1		
Net	Label	Type
10/8	49	remote
-	-	-

LIB R4		
Net	Label	Type
10/8	49	remote
-	-	-

LIB R3		
Net	Label	Type
10/8	49	remote
-	-	-

- **R2**
  - advertises label 49 for FEC 10/8 to all neighbor routers
- **Per platform label allocation**
  - same label on all interfaces
  - LFIB may not contain an incoming interface (next HOP) field at that moment
- **Every neighbor MPLS router**
  - stores received label for IP destination 10/8 in the corresponding LIB

Actions on Receiving Labels on R1



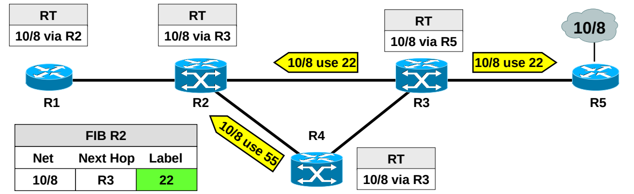
FIB R1		
Net	Next Hop	Label
10/8	R2	49

LIB R1		
Net	Label	Type
10/8	49	remote
-	-	-

LFIB R1		
Label	Action	Next Hop
-	49	R2

- **R1**
  - receives label 49 for FEC 10/8
  - label is advertised by router which is the next hop in the routing table -> therefore populates the FIB
  - LFIB is adapted to use label 49 for FEC 10/8 towards R2
  - action in LFIB has the meaning of outgoing label or remote label
  - label in LFIB has the meaning of incoming label or local label

Actions on Receiving of Labels from R3 and R4 on Router R2



FIB R2		
Net	Next Hop	Label
10/8	R3	22

LIB R2		
Net	Label	Type
10/8	49	local
10/8	22	remote
10/8	55	remote

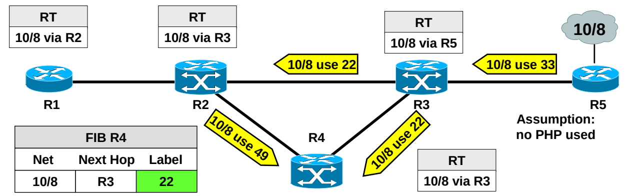
LFIB R2		
Label	Action	Next Hop
49	22	R3

- **R2**

  receives label 22 for FEC 10/8
  - this label is advertised by router which is the next hop in the routing table -> therefore populates the FIB
  - LFIB is adapted to use (swap) label 22 for FEC 10/8 towards R3

  receives label 55 for FEC 10/8
  - this label is advertised by router which is not the next hop in the routing table but will be still stored in the LIB -> liberal retention mode

Receiving of Labels from R2 and R3 on Router R4



FIB R4		
Net	Next Hop	Label
10/8	R3	22

LIB R4		
Net	Label	Type
10/8	55	local
10/8	22	remote
10/8	49	remote

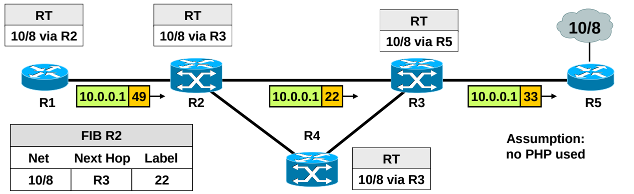
LFIB R4		
Label	Action	Next Hop
55	22	R3

- **R4**

  receives label 22 for FEC 10/8
  - this label is advertised by router which is the next hop in the routing table-> therefore populates the FIB
  - LFIB is adapted to use label 22 for FEC 10/8 towards R3

  already received label 49 for FEC 10/8
  - this label is advertised by router which is not the next hop in the routing table but will be still stored in the LIB

Label Switching



FIB R2

Net	Next Hop	Label
10/8	R3	22

LIB R2

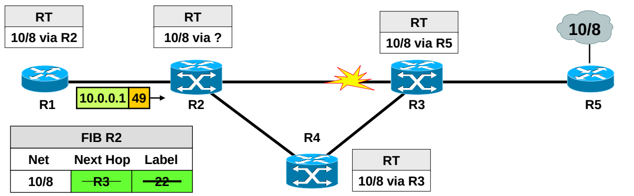
Net	Label	Type
10/8	49	local
10/8	22	remote
10/8	55	remote

LFIB R2

Label	Action	Next Hop
49	22	R3

- **Packets of FEC 10/8 will follow the corresponding Label Switched Path**

Link Failure R2 <-> R3



FIB R2

Net	Next Hop	Label
10/8	~~R3~~	~~22~~

LIB R2

Net	Label	Type
10/8	49	local
10/8	~~22~~	~~remote~~
10/8	55	remote

LFIB R2

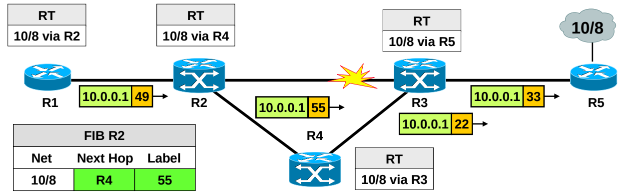
Label	Action	Next Hop
49	~~22~~	~~R3~~

- **Routing protocol neighbors and LDP neighbors are lost after failure**
  - corresponding entries in FIB, LIB and LFIB are removed
- **Traffic**
  - to FEC 10/8 will not be forwarded until routing table converges

Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

Routing Protocol Convergence



FIB R2

Net	Next Hop	Label
10/8	R4	55

LIB R2

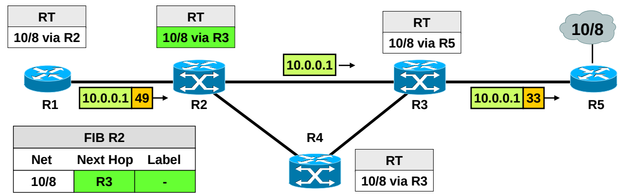
Net	Label	Type
10/8	49	local
10/8		
10/8	55	remote

LFIB R2

Label	Action	Next Hop
49	55	R4

- **After routing protocol convergence**
  - R2 can switch over immediately to other LSP if alternative label advertisements were stored in LIB and labeled packets will flow again
  - Otherwise R2 must wait for new bindings and can forward packets only based on IP address in the meantime (action untag in LFIB)
- **Packets of FEC 10/8 will follow the new Label Switched Path via R4**

Link Failure Repair 1



FIB R2

Net	Next Hop	Label
10/8	R3	-

LIB R2

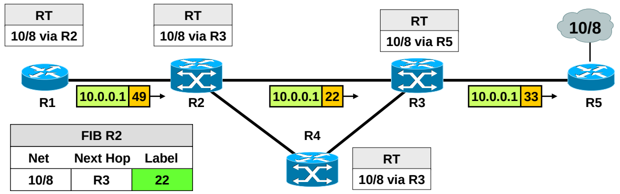
Net	Label	Type
10/8	49	local
10/8		
10/8	55	remote

LFIB R2

Label	Action	Next Hop
49	untag	-

- **After link repair**
  - Routing protocol neighbor detection and routing table adaptation
  - R2 must wait for new bindings and can forward packets only based on IP address in the meantime (action untag in LFIB)

Link Failure Repair 2



FIB R2

Net	Next Hop	Label
10/8	R3	22

LIB R2

Net	Label	Type
10/8	49	local
10/8	22	remote
10/8	55	remote

LFIB R2

Label	Action	Next Hop
49	22	R3

- **After LDP session to R3 is up and binding for FEC 10/8 from R3 received**
  - Packets of FEC 10/8 will follow the corresponding Label Switched Path again

Agenda

- **Review ATM**
- **IP over WAN Problems (Traditional Approach)**
- **MPLS Principles**
- **Label Distribution Methods**
- **MPLS Details** (Cisco)
  - Internal Components
  - MPLS in Action
  - TDP, LDP
  - TTL
  - Traffic Engineering
  - MPLS and BGP
- **RFCs**

Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

TDP Key Facts

- **Tag Distribution Protocol (TDP)**
  - invented by Cisco
  - for distributing <label, prefix> bindings
  - enabled by default
- **Session establishment: UDP/TCP port 711**
  - Hello messages via UDP
  - destination address -> 224.0.0.2
    - well-known multicast address for all subnet routers
  - TDP session via TCP, incremental updates
- **Not compatible with LDP**
  - but can co-exist as long as two peers use the same protocol

LDP Key Facts

- **Label Distribution Protocol**
- **IETF standard RFC 3036**
  - descendent of Cisco's proprietary TDP
- **Same concept but port 646**
- **LDP-Identifier**
  - Router ID (4 bytes)
  - Label Space ID (2 bytes)
    - in case of per-platform label space this field is set to zero
    - note: in ATM you need a per-interface label space
- **TCP session initiated from router with highest address**

LDP Message Types

- **Four basic types:**
  - Discovery (UDP):
    - getting into contact with neighbor LSR´s
  - Adjacency (TCP):
    - Initialization, Keepalive and Shutdown of LDP sessions
  - Label Advertisement (TCP):
    - Label Binding - Advertisement, - Request, - Withdrawal, - Release
  - Notification (TCP):
    - Signal of Error Information, Advisory Information
- **TLV (Type/Length/Value)**
  - encoding is used for easy extension of the protocol

Discovery Message

- **Basic discovery of directly connected LSRs:**
  - Hello Message with targeted bit set to 0
    - UDP to port 646
    - IP multicast address “all routers on this subnet” (224.0.0.2)
- **Extended discovery of non-directly connected LSR’s:**
  - Hello Message with targeted bit set to 1 (Targeted Hello)
    - UDP to port 646
    - IP unicast address of neighbor
  - used e.g. in case of MPLS Traffic Engineering
- **After discovery**
  - LDP session is created running on top of TCP
    - well known port 646

Adjacency Messages

- **Adjacency**
  - Initialization
    - negotiates
      - protocol version (current version = 1)
      - label advertisement discipline
        - » Unsolicited Downstream = 0
        - » Downstream-on-Demand = 1
      - keepalive time
  - Keepalive
    - maintains LDP session

Label Advertisement Messages

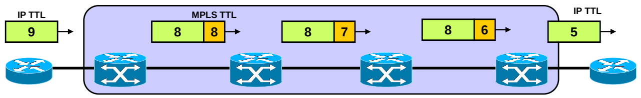
- **Label Advertisement**
  - Label Mapping
    - advertise a binding between a FEC and a label
  - Label Withdrawing
    - reverse the mapping process
    - e.g. if FEC is not longer valid because address prefix has been removed from the routing table
  - Label Release
    - issued by a LSR which has previously received a label mapping and no longer has a need for that mapping
  - Label Request / Label Request Abort
    - for Downstream-on-Demand method
    - abort is used to revoke a request before it has been satisfied

Agenda

- **Review ATM**
- **IP over WAN Problems (Traditional Approach)**
- **MPLS Principles**
- **Label Distribution Methods**
- **MPLS Details (Cisco)**
  - Internal Components
  - MPLS in Action
  - TDP, LDP
  - TTL
  - Traffic Engineering
  - MPLS and BGP
- **RFCs**

Normal TTL Usage

- **Loop detection**
  - LDP and TDP basically rely on IGP loop detection, therefore no additional tasks are necessary for MPLS control packets
  - Additionally a TTL field in the MPLS header prevents endless routing of MPLS data packets
- **TTL Propagation:**
  - IP TTL is copied into MPLS header
  - Done by Ingress LSR (LER)
  - MPLS TTL decremented by every LSR
  - Egress LSR copies MPLS-TTL back to IP TTL
  - Enabled by default on Cisco routers

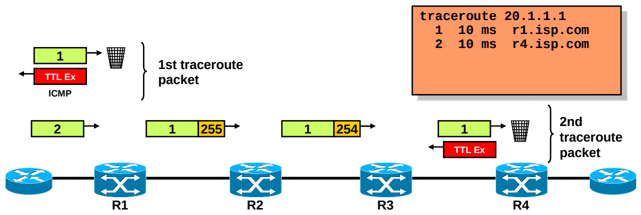


Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

Disable TTL Propagation

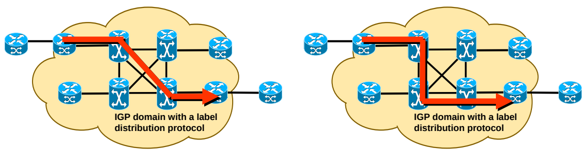
- **No TTL copying between IP and MPLS header**
- **Ingress router assigns MPLS TTL 255**
- **Core routers are hidden**
  - E. g. traceroute fails to show them



Agenda

- **Review ATM**
- **IP over WAN Problems (Traditional Approach)**
- **MPLS Principles**
- **Label Distribution Methods**
- **MPLS Details (Cisco)**
  - Internal Components
  - MPLS in Action
  - TDP, LTP
  - TTL
  - Traffic Engineering
  - MPLS and BGP
- **RFCs**

Label Switch Path (LSP)



LSP follows IGP shortest path

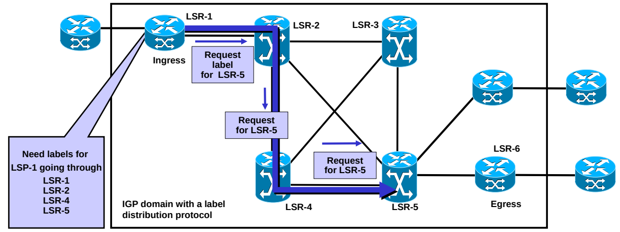
LSP diverges from IGP shortest path

- **Normal MPLS Destination Based Routing**
  - FEC is determined in LSR-ingress
  - LSP's derive from IGP routing information
- **If LSPs should diverge from IGP shortest path**
  - LSP Explicit Routing (LSP Tunnel) is necessary
  - MPLS Traffic Engineering

Traffic Engineering via LSP - Tunnels

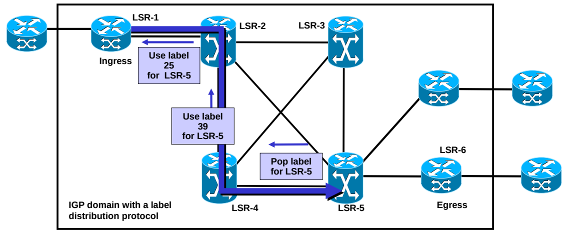
- **Explicit Routing:**
  - Source Routing
  - Constraint-Based Path Selection Algorithm
    - similar to ATM PNNI
  - OSPF / IS-IS extension for flooding of resources / policy information
    - traffic class, resource requirements and the available network resources (bandwidth)
  - RSVP as the mechanism for establishing LSP's
    - uses new RSVP objects in PATH and RESV messages
      - Explicit-Route (ERO) in Path, Label found in RSV
  - Usage of ER-LSPs in the forwarding table
    - label stack

Explicit Routing 1



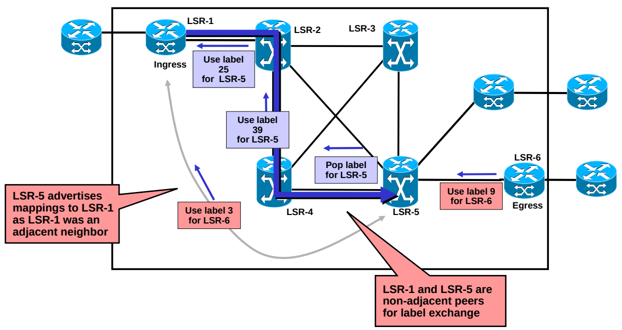
- **LSR-1 request an explicit LSP to LSR-5:**
  - LSR-1, LSR-2, LSR-4, LSR-5
- **The request travels hop-by-hop**
  - using RSVP PATH messages

Explicit Routing 2

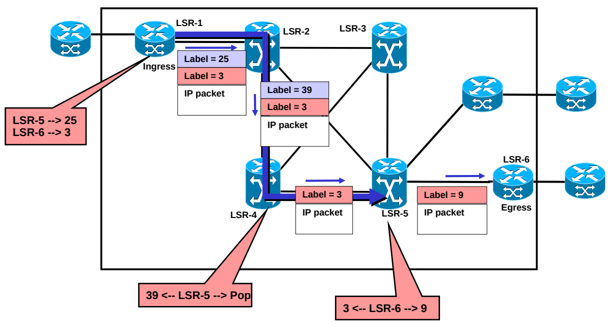


- **When the request reaches the egress point labels are advertised back to the ingress LSR**
  - via RSVP RESV messages

Explicit Routing 3



Explicit Routing 4



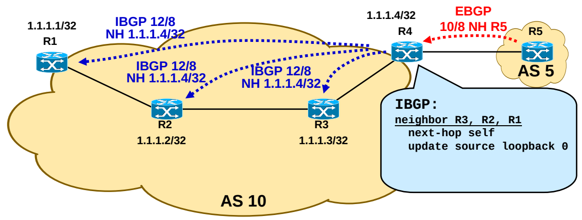
Datenkommunikation 384.081 - SS 2012

Appendix 3 - MPLS (v4.6)

Agenda

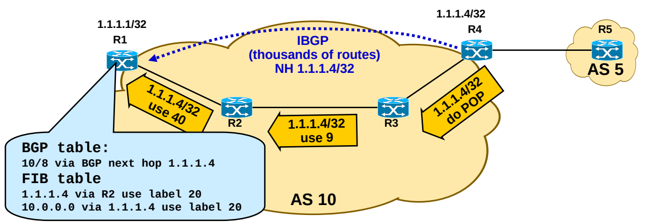
- Review ATM
- IP over WAN Problems (Traditional Approach)
- MPLS Principles
- Label Distribution Methods
- MPLS Details (Cisco)
  - Internal Components
  - MPLS in Action
  - TDP, LTP
  - TTL
  - Traffic Engineering
  - MPLS and BGP
- RFCs

BGP Standard Behavior



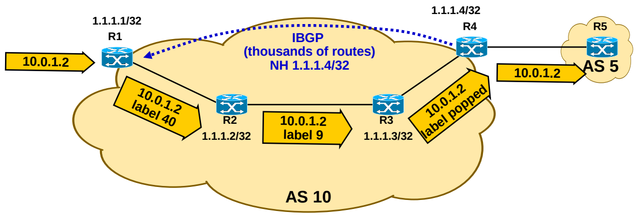
- Good style: Use loopback addresses and next hop self
  - BUT: Full mesh IBGP !!!
  - BUT: Each router has full routing table !!!
- IGP is used to propagate loopback addresses
  - 1.1.1.1/32, 1.1.1.2/32, 1.1.1.3/32, and 1.1.1.4/32
- Note: BGP Synchronization Off
  - Otherwise IBGP routes would never be copied into the routing table
  - IBGP updates would only be propagated by PE-router if this network is reachable via IGP

MPLS and BGP 1

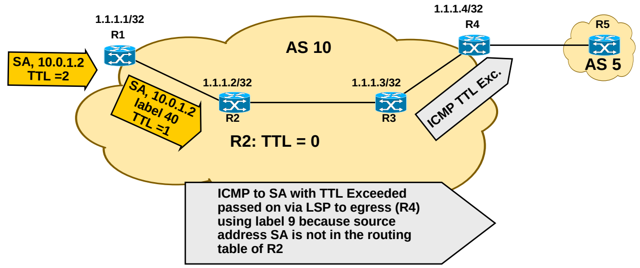


- **FEC = Next Hop**
  - Only EBGP routers must learn all external routes
  - Internal routers do not require the external networks to be in the routing table
  - packets to external networks are labeled with the label to reach the BGP next hop
- **IBGP sessions only between PE-routers**

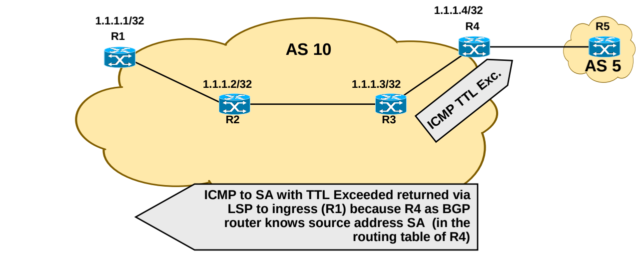
MPLS and BGP 2



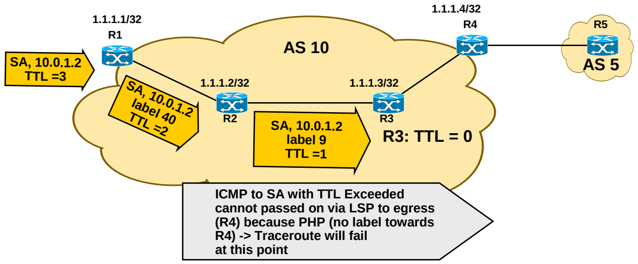
Traceroute Behavior in case of MPLS-BGP 1



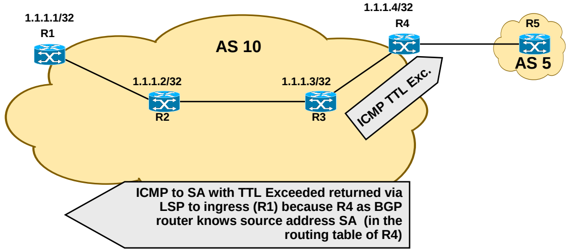
Traceroute Behavior in case of MPLS-BGP 2



Traceroute Behavior in case of MPLS-BGP 3



Traceroute Behavior in case of MPLS-BGP 2



Agenda

- Review ATM
- IP over WAN Problems (Traditional Approach)
- MPLS Principles
- Label Distribution Methods
- MPLS Details (Cisco)
- RFCs

RFC References 1

- **RFC 3031**
  - Multiprotocol Label Switching Architecture
- **RFC 3032**
  - MPLS Label Stack Encoding
- **RFC 3036**
  - LDP Specification
- **RFC 3063**
  - MPLS Loop Prevention Mechanism
- **RFC 3270**
  - MPLS Support of Differentiated Services

RFC References 2

- **RFC 3443**
  - Time To Live (TTL) Processing in MPLS
- **RFC 3469**
  - Framework for Multi-Protocol Label Switching (MPLS)-based Recovery
- **RFC 3478**
  - Graceful Restart Mechanism for Label Distribution Protocol
- **RFC 3479**
  - Fault Tolerance for the Label Distribution Protocol (LDP)

