

MPLS

Traffic Engineering & QoS / CoS

Materialien: Gabriele Schrenk
Vortragender: Herbert Almus

EANTC AG

Dezember 2004

<http://www.eantc.com>

Global Crossing Tightens MPLS SLAs

Quelle: Converge! Network Digest, 29.11.2004

Global Crossing announced new Service Level Agreements (SLAs) for its converged IP services, including IP VPNs, VoIP and IP video.

The upgraded SLAs include:

- end-to-end availability of 99.999%, compared to the previous guarantee 99.9%.
- packet delivery of 99.999%, compared to the previous guarantee of 99.9%.
- jitter of 5 milliseconds (ms), compared to the previous guarantee of 15 ms.
- a new service guarantee ensures that service performance issues trigger availability guarantees and reparations.
- a new performance guarantee equates service degradation with unavailability, and offers reparations to customers experiencing severe latency, jitter or packet delivery issues.

Global Crossing said its IP network has been operating at 99.999% packet delivery and availability, and less than 5 ms of jitter during 2004. **Routing is optimized across the network using MPLS.**

The company also noted that it is currently carrying up to 2.5 billion minutes per month over its private, global VoIP platform, deployed four years ago.

Why yet Another Technology?

IP is the dominant protocol today

people want to do everything over IP (EOIP) as well as
IP over everything (IPOE)

but:

IP *lacks* Class / Quality of Service (CoS / QoS)

MPLS promises to

enable Traffic Engineering for providers
realize end-to-end Quality of Service for IP networks
be independent of a specific layer 2 technology
have the ability to scale in real large networks
provide higher performance with less costs

MPLS the Concept

Route at the edge, switch in the core

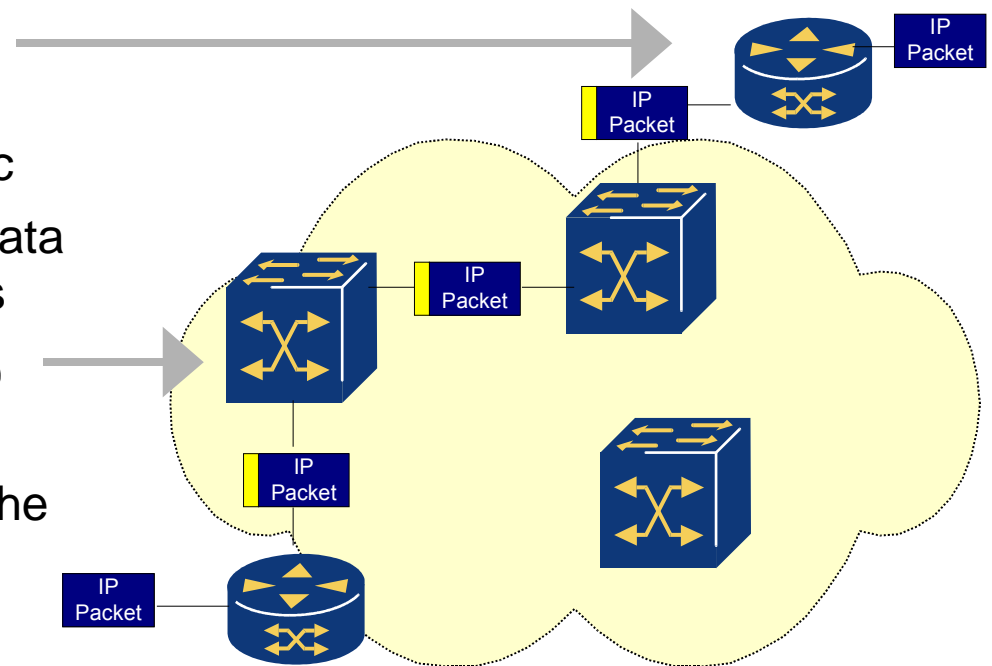
Prefix packets with labels at the network boundary

Label Edge Router (LER)

- analyse and classify traffic
- request/assign labels to data at the network boundaries

Label Switch Router (LSR)

- label swapping and forwarding in the core of the MPLS network



Terms: FEC

Label

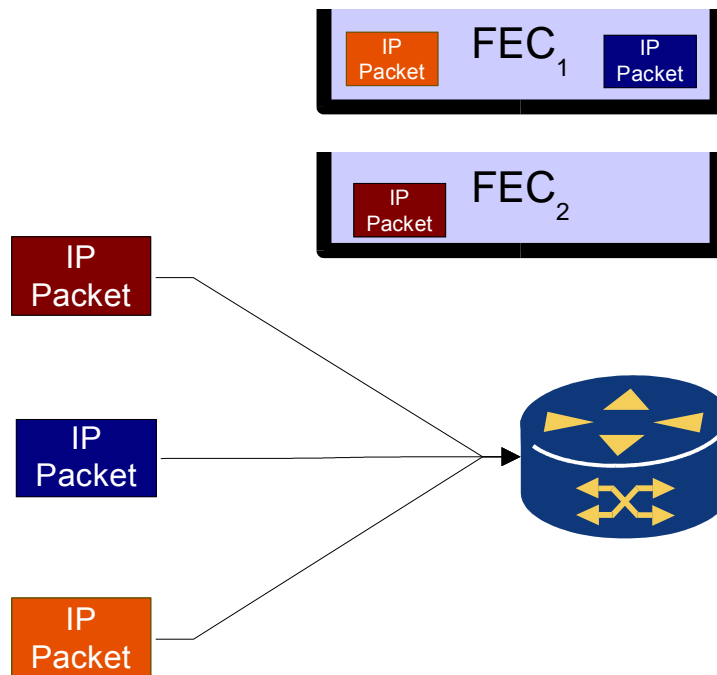
- Short, fixed-length value prefixing the packet
- Associates a **Forward Equivalence Class** (FEC)

Forward Equivalence Class (FEC)

a group of L3 packets which are forwarded in the same manner
e.g., over the same path, with the same forwarding treatment

'FEC is therefore the set of L3 packets which could safely be mapped to the same label'

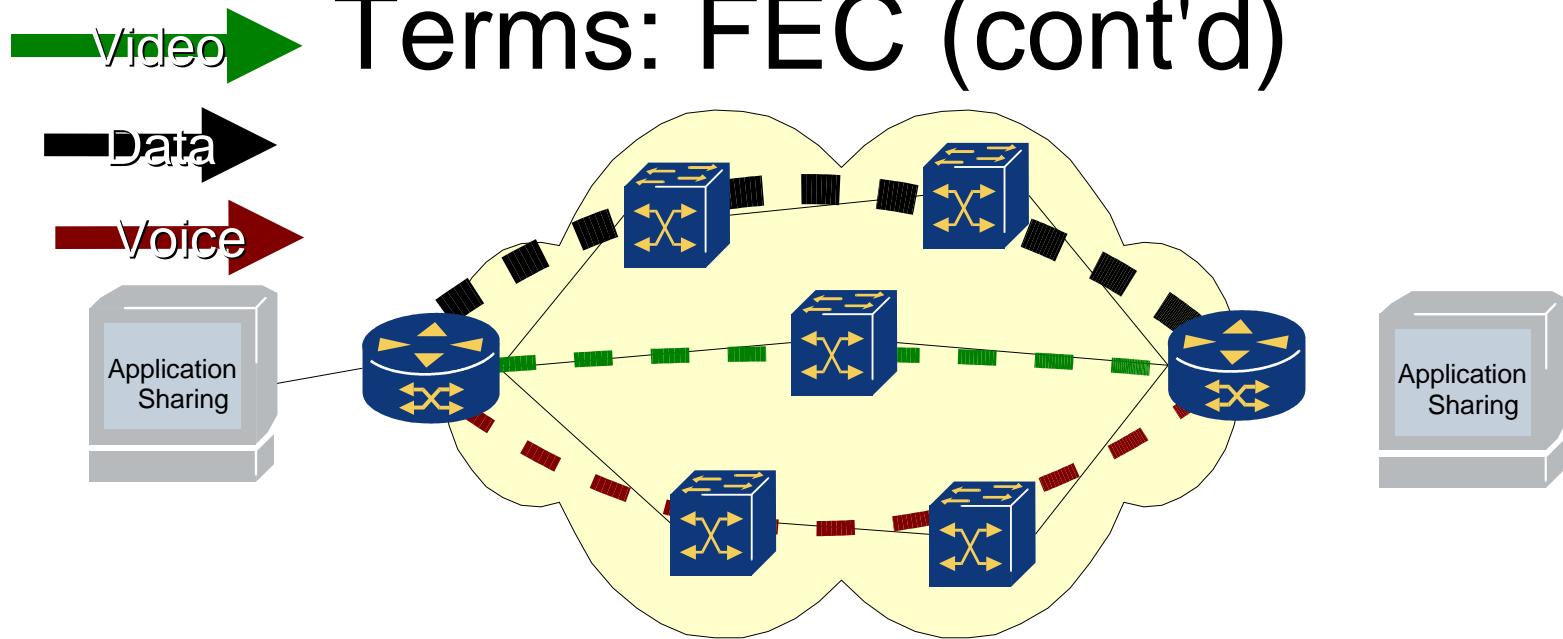
Terms: FEC (cont'd)



Classify traffic at the edge
assign packets to a given FEC

Currently standardised FECs:
Destination host
Destination network

Terms: FEC (cont'd)



Normally bundles several packets of different sources and destinations but:

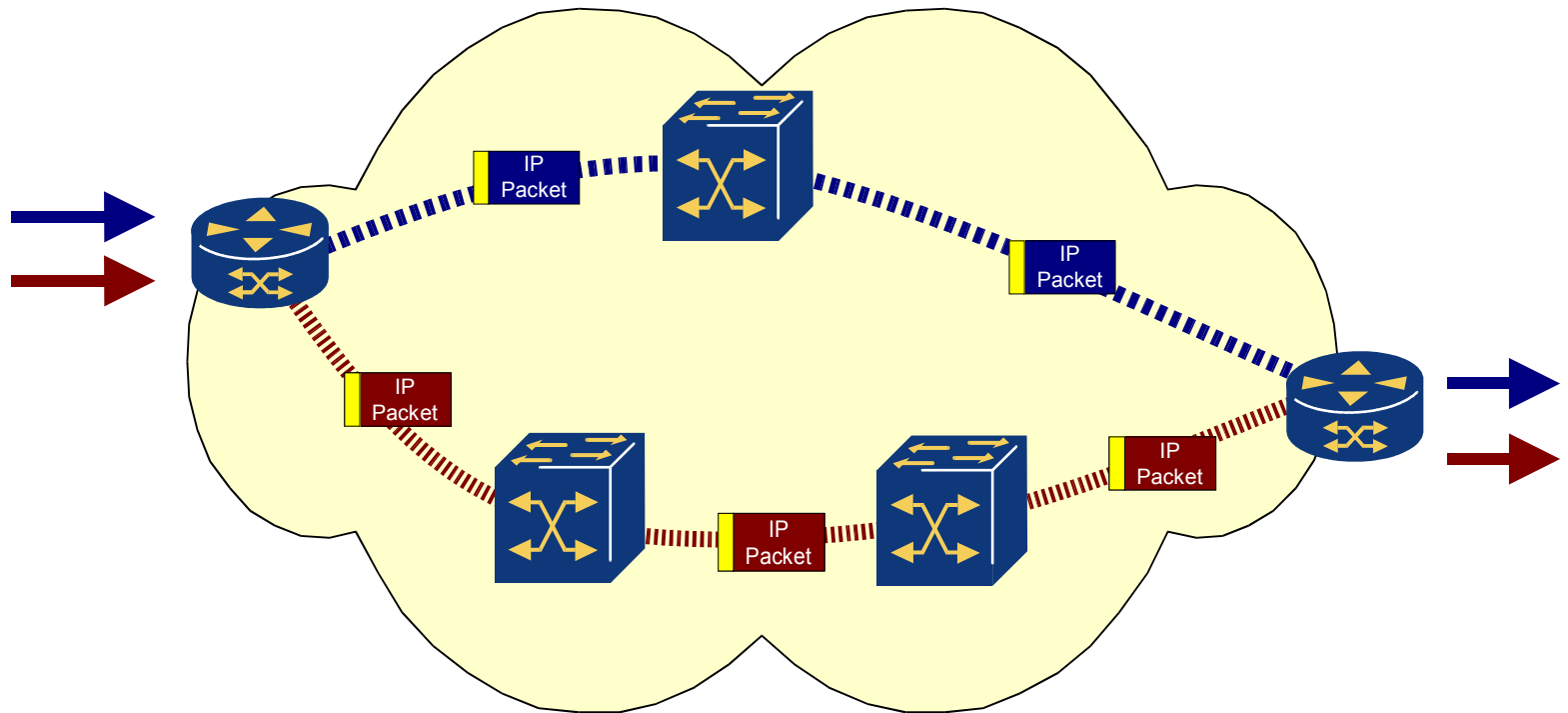
- can also split traffic from one source to one destination by QoS requirements

Aggregates traffic and enables scaling

Terms: LSP

Label Switched Path

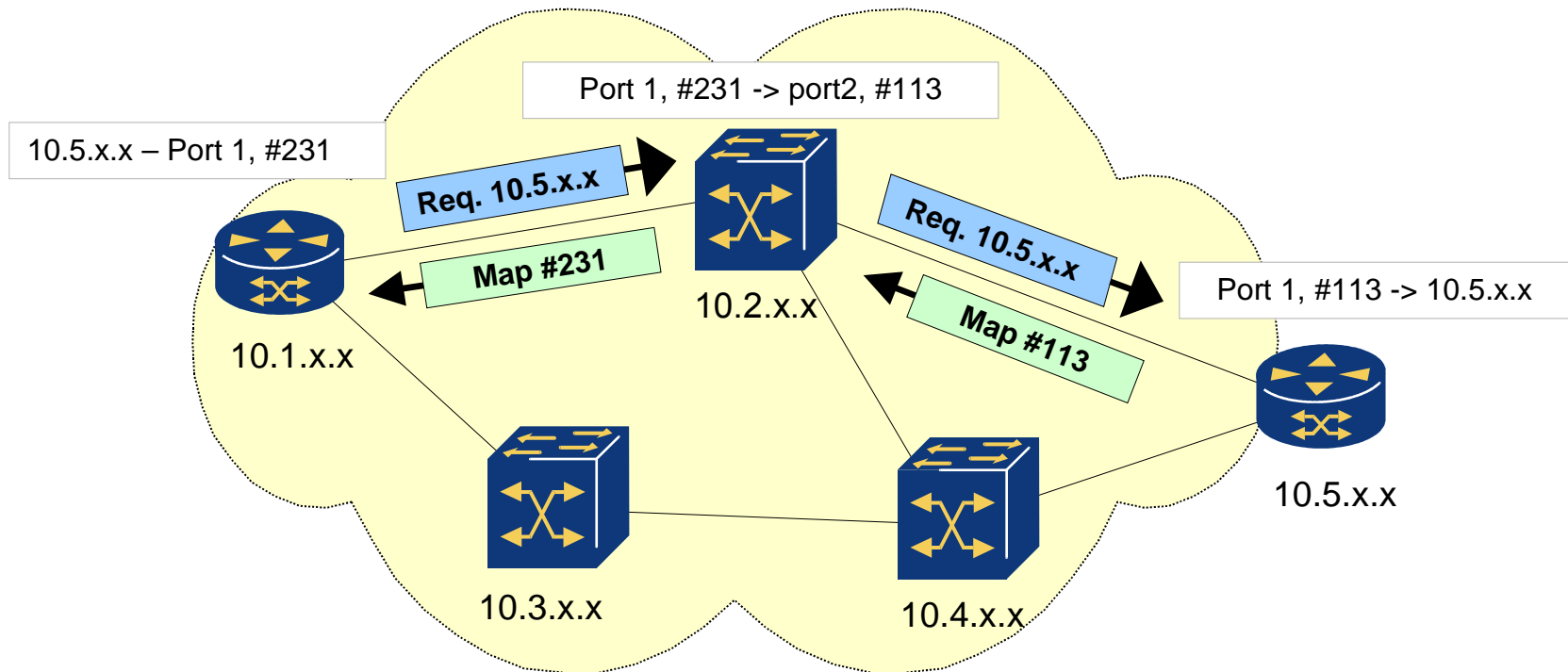
a virtual path (along which a packet is travelling) connecting LSRs through a cloud of MPLS switches



LSP establishment

Labels must be exchanged between LSR's:

Requiring a signalling Protocol which is responsible for the label distribution between the LSR's

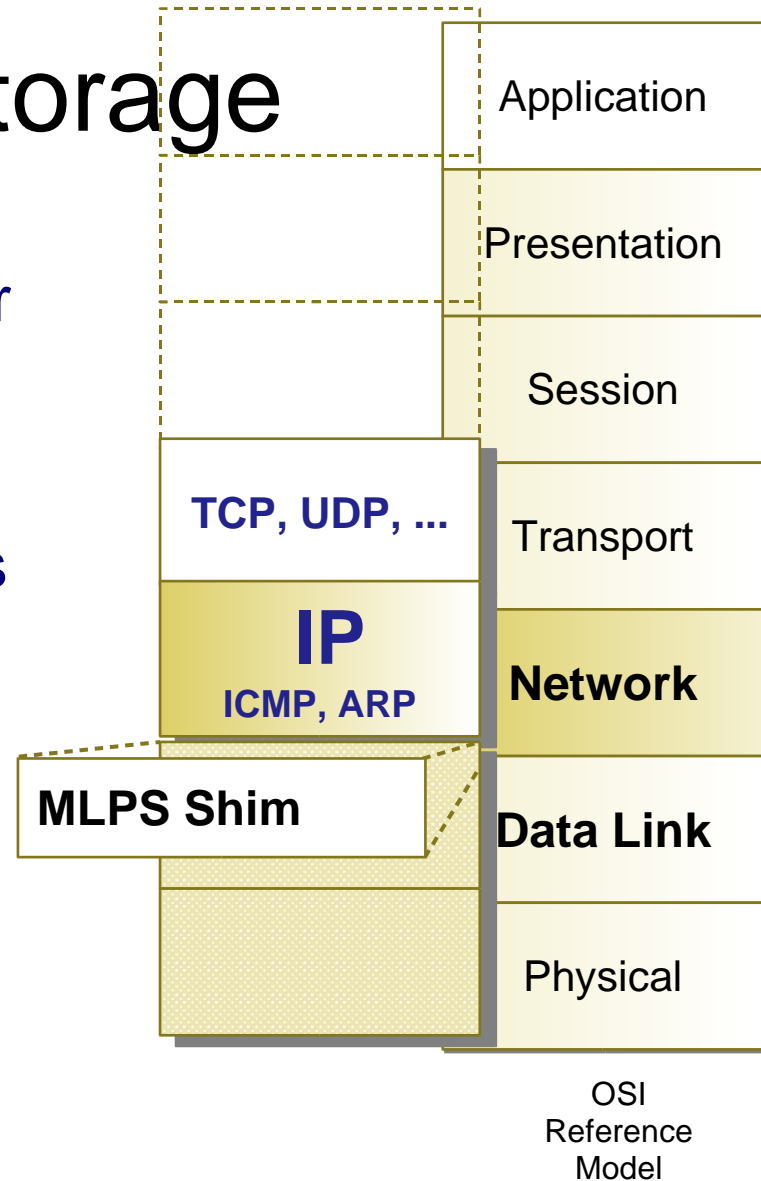


Label storage

MPLS is intended to run over multiple link layers

Labels can be understood as shim (thin) layer between L2 and L3 hiding L2 differences

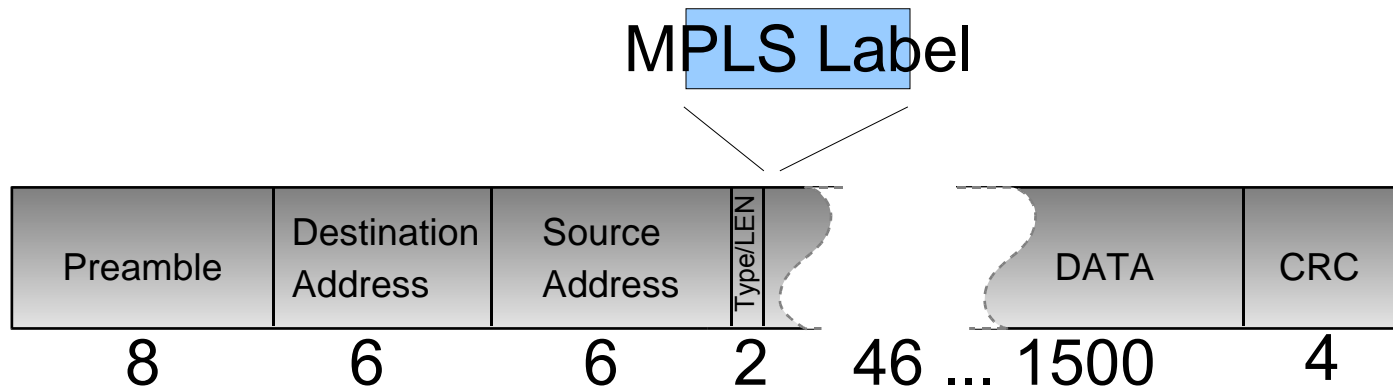
Labels should be integrated with L2 specific features



Link Layer Differences

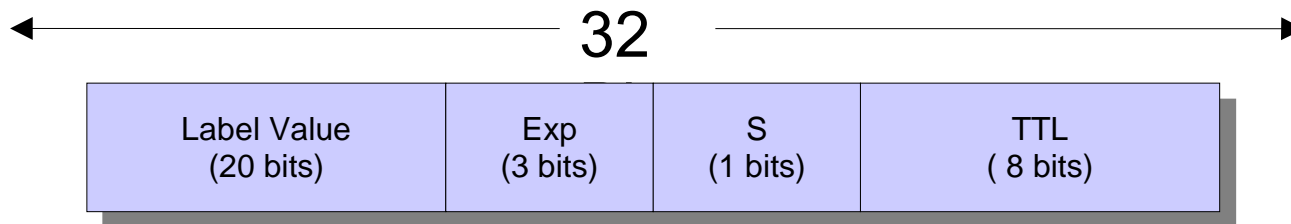
Some L2 technologies have no native information fields where an MPLS label could be placed

Introduction of a Shim Header prefixing the original L3 PDU.



Labels

Standard Shim Header consists of four fields



- Label: Label Value, 20 bits
- Exp: Experimental Use, 3 bits
- S: Bottom of Stack, 1 bit
- TTL: Time to Live, 8 bits

Experimental Bits

Reserved for future use in combination with other technologies, i.e. Differentiated Services

Different approaches to use this field:

DiffServ uses this field to store Class of Service information

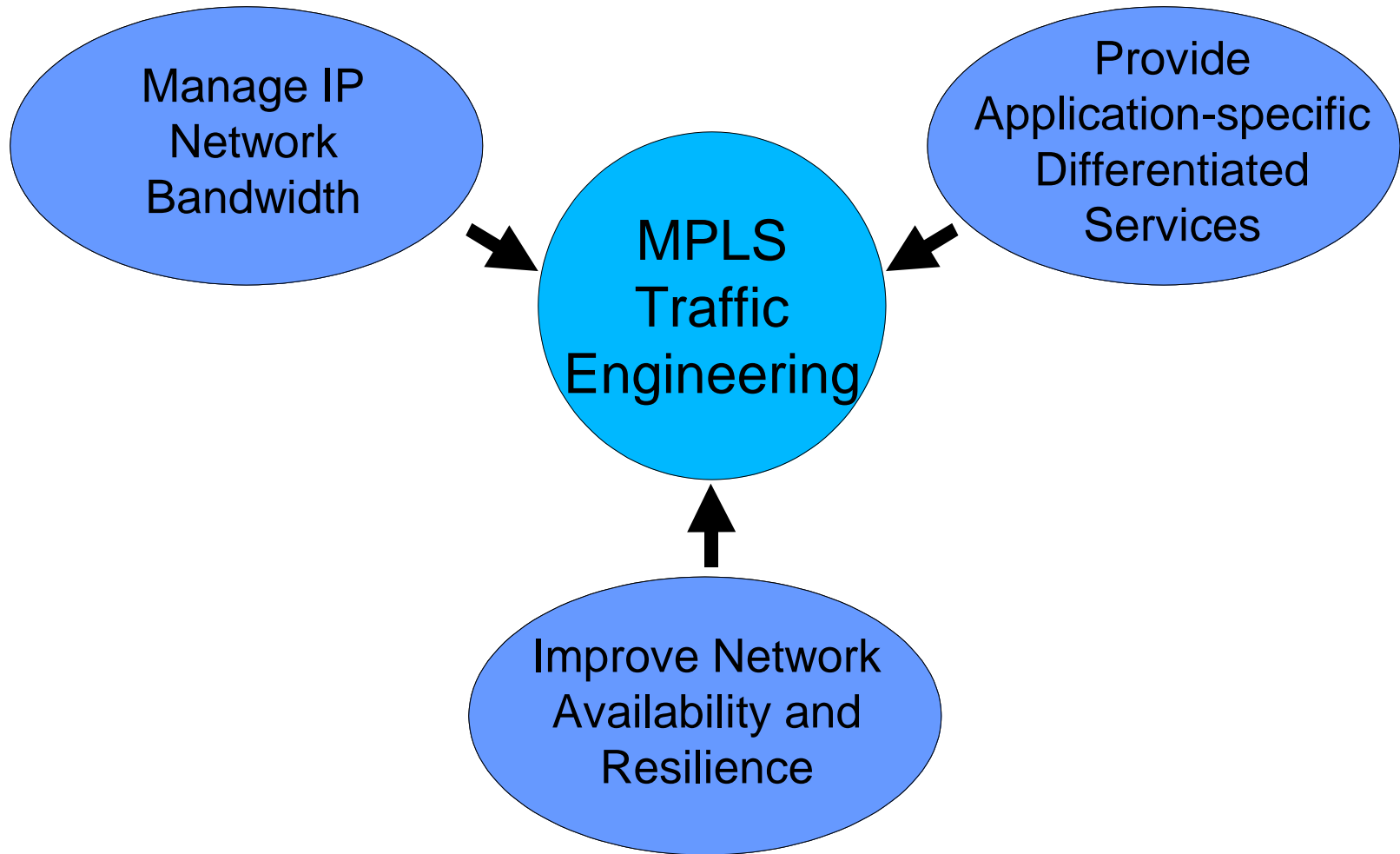
- It is possible to transport several service classes in a single Path

Provides a simple method to segregate flows within a label

MPLS

Traffic Engineering

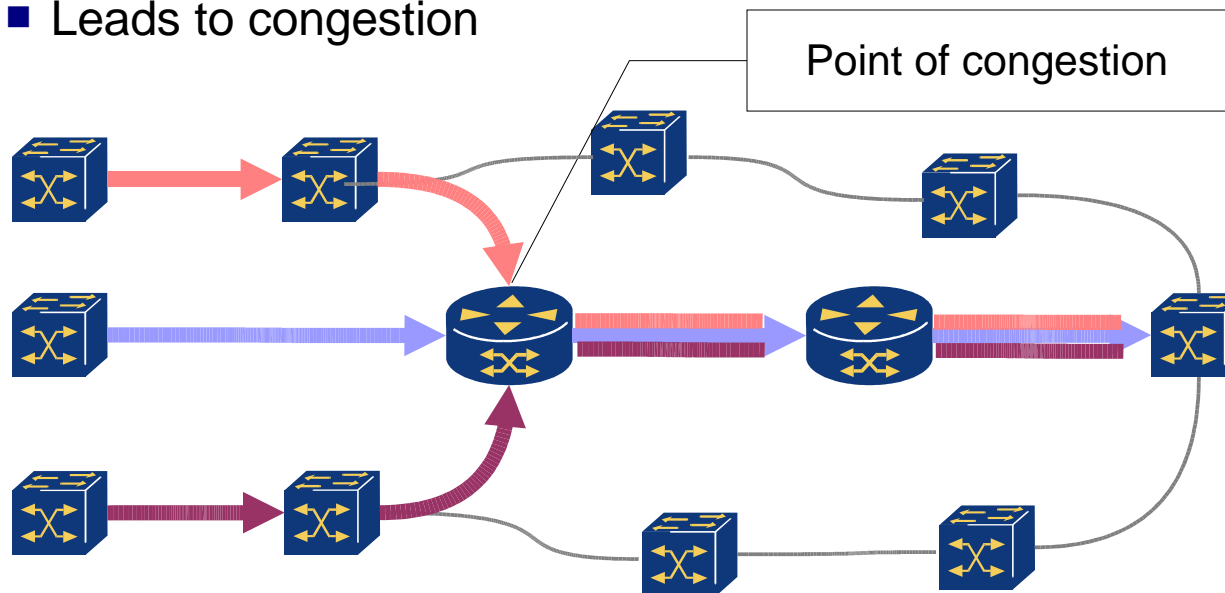
Why use Traffic Engineering?



Reasons to use TE

Destination-based routing protocols aggregate and concentrate traffic

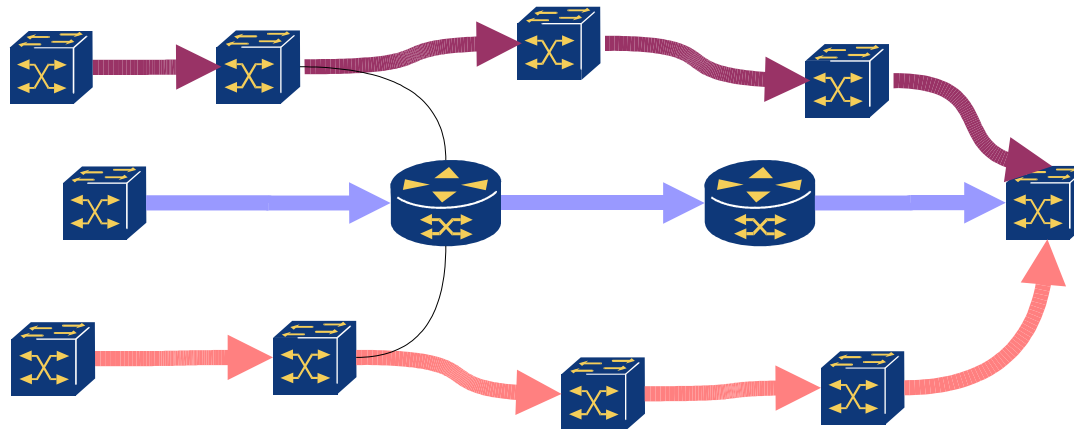
- Leads to congestion



Reasons to use TE (2)

Solution:

- Reduce traffic on over-utilized sections
- Re-route flows to under-utilized sections



- Segregate traffic instead of aggregating it!
(separate traffic flows going to the same destination)
- Requires non-destination based (“explicit”) routing

Application-Specific Services

Some IP applications need specific network service:

- Delay-sensitive applications
- Bandwidth-sensitive applications
- Bulk applications

Multi-service converged networks need more differentiated services:

- Any kind of legacy data service (e.g. Internet, IP VPNs, Frame Relay)
- Real-time services (ATM, TDM voice)

Multiple types of service classes regarding delay and jitter

=> Traffic engineering necessary to handle flows differently

Improve Network Resilience

Traffic Engineering helps to improve service resilience to 99.9 % and more:

Protocol mechanisms supporting

- Link, Node, and Path Protection
- Fast Rerouting

“Real” Service Level Management means:

- Service levels are guaranteed by MPLS protocols for every second
- (most SLAs today are based on long-term statistical analysis)

How Far is MPLS TE Ready For Deployment?

Let's have a look ...

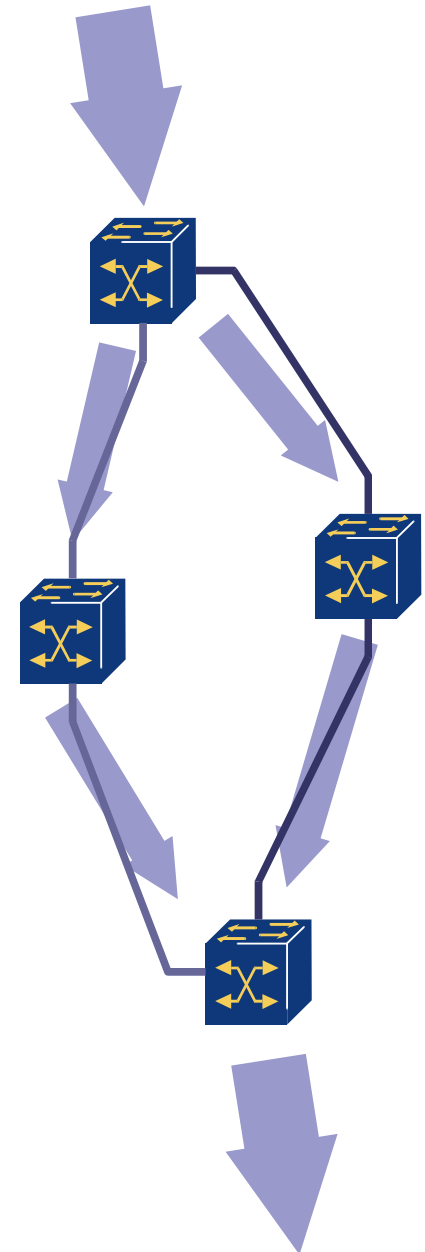
- RFC2702 is available since September 1999!
- Vendors have implemented most aspects of MPLS traffic engineering
- Drafts become more stable now (e.g. in the area of fast rerouting)

The following slides provide an introduction to the technology:

Definition: Traffic Trunk

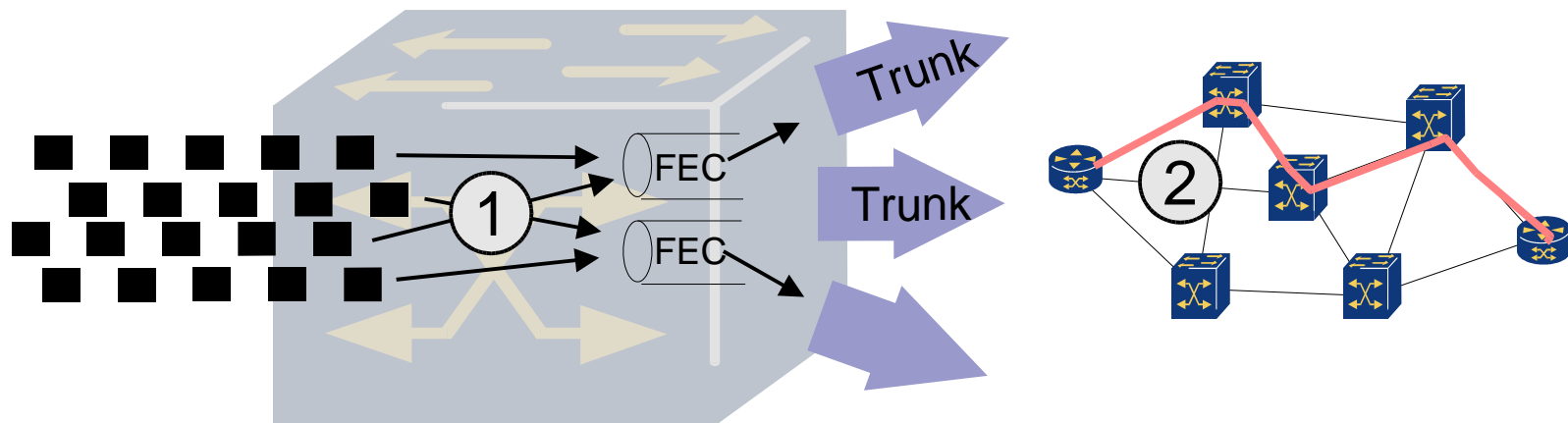
- Aggregation of traffic flows of the same class (usually one single FEC)
- Unidirectional
- Carries traffic from an ingress LSR to an egress LSR
- Placed in one Label Switched Path (LSP) at one time
- Distinct from the LSP itself: Trunks can be moved from one LSP to another (by protection switching / rerouting)

(According to RFC2702)



Overview TE Tasks

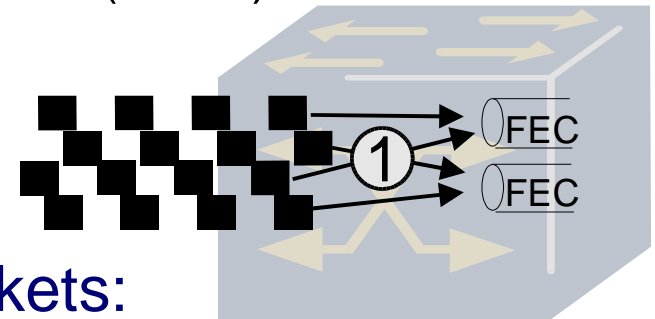
1. Map packets to FECs¹: Define flows
2. Map trunks to physical network topology: Implement constraint-based routing



1) FEC = Forward Equivalence Class

TE Task 1: Map packets to FECs

- To classify packets and identify flows, they need to be mapped to Forward Equivalence Classes (FECs)
- Internal task of the switch, manufacturer dependent



Many different ways to classify packets:

- IP addresses, TCP/UDP ports
- TOS / DiffServ code points
- Edge multi-service type (Frame Relay, ATM, TDM, ...)
- VPN membership
- ...

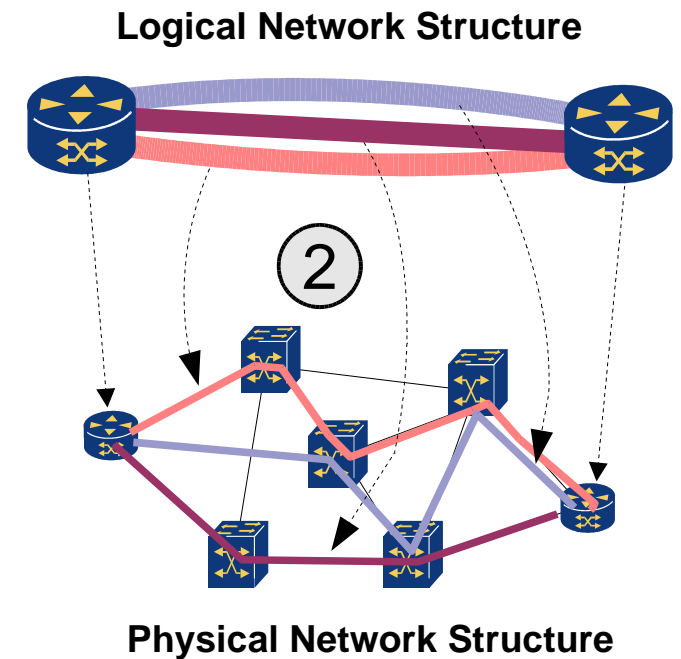
TE Task 2: Map Trunks to Physical Topology

Major task
of MPLS Traffic Engineering!

Take all trunks,
find suitable paths through
the physical network for each
of the trunks

Requires:

- Traffic engineering extensions to signalling:
RSVP-TE, CR-LDP
- Constraint-based routing
(to find routes through the network)



Constraint-based Routing

Constraint-based routing allows

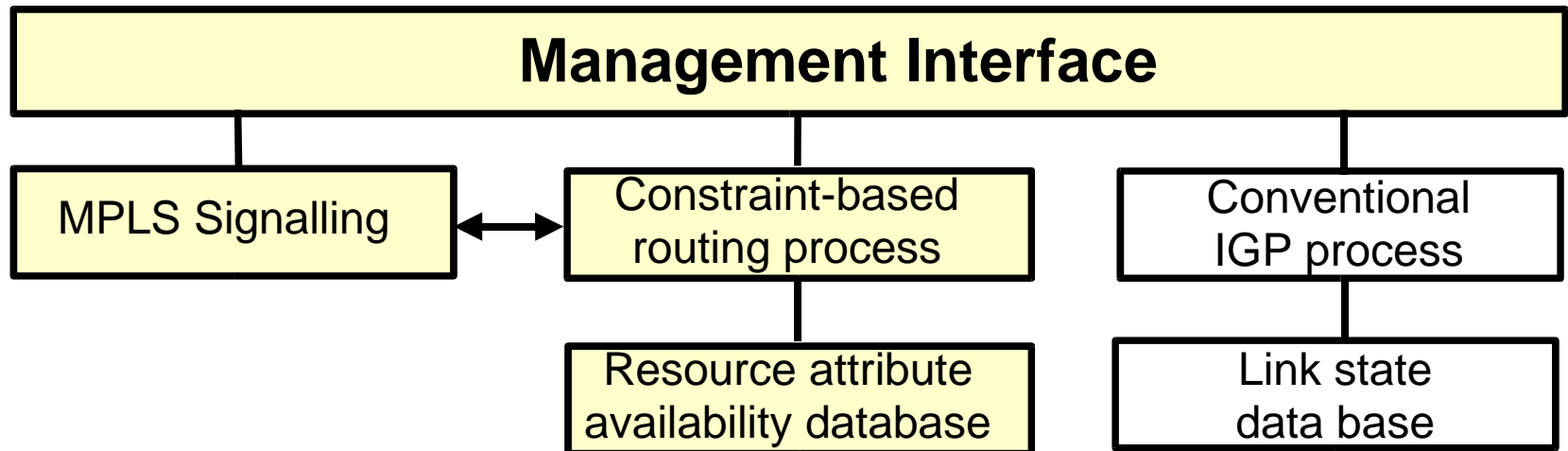
- Explicit route selection
- Bandwidth parameters and recovery mechanisms
- Enhancements to signalling and routing protocols
 - OSPF -> OSPF-TE
 - ISIS -> ISIS-TE
 - RSVP -> RSVP-TE
 - LDP -> CR-LDP

Benefit:

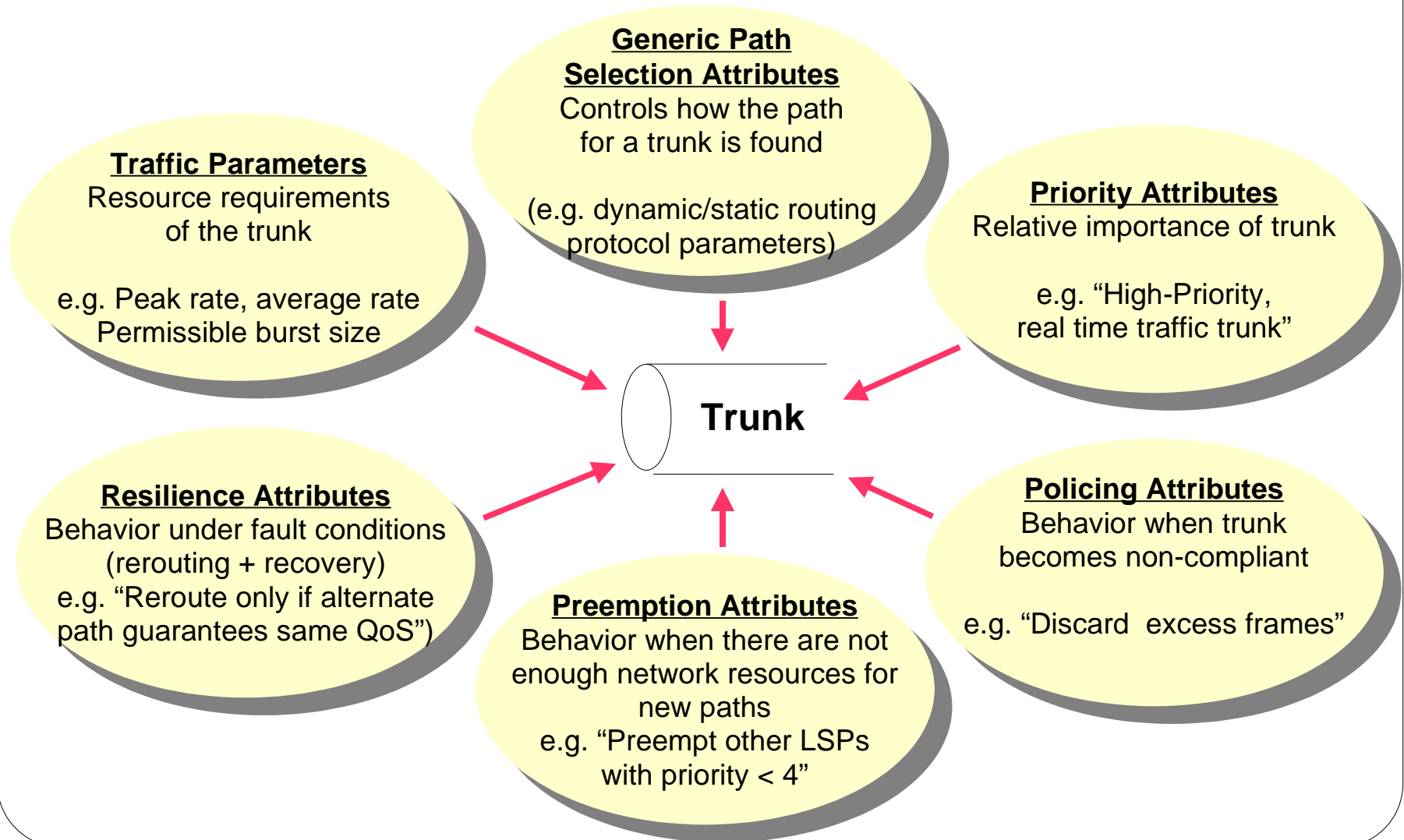
- Greatly reduces the level of manual configuration
- Traffic-engineered paths are routed dynamically according to their constraints
- BUT: Will carriers trust it?

Constraint-based Routing: Router Implementation

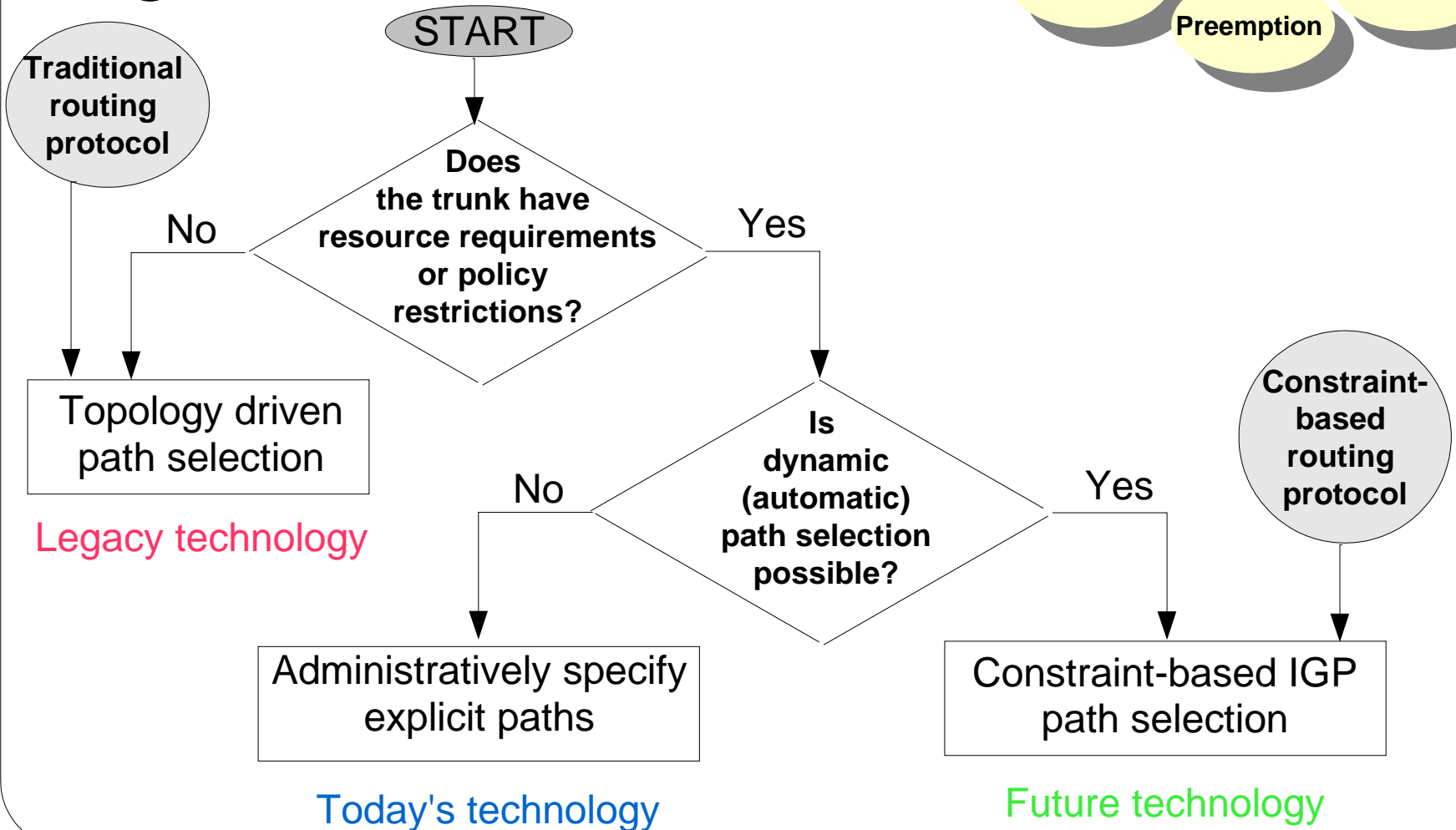
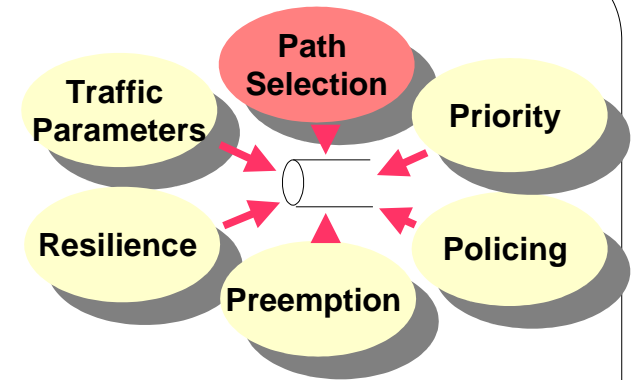
- Extend the current routing protocols to allow constraint-based routing, e.g. OSPF -> OSPF-TE
- Add a constraint-based routing process to each router which can co-exist with current IGPs.
- Let signalling and routing processes communicate.



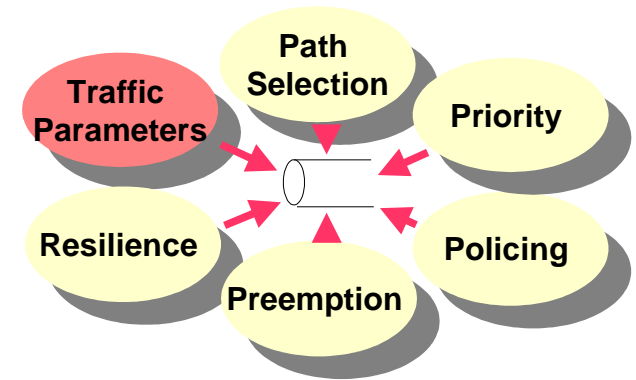
What Trunk Attributes exist?



Path Selection Algorithm



QoS versus CoS (1)



Quality of Service (QoS)

Hard allocation of resources

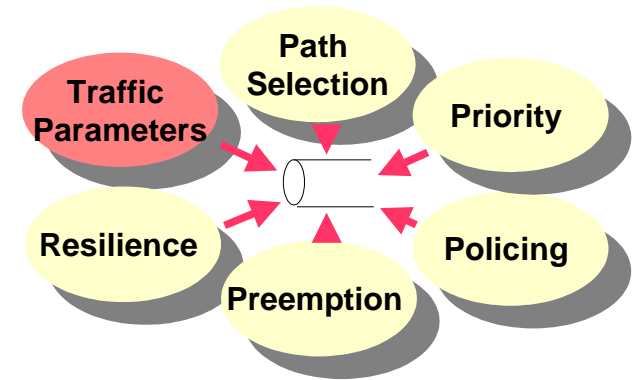
What you pay for is what you get

- Example: Connection set-up in an ATM network which allocates a fixed amount of resources

QoS is the ability to guarantee transmission characteristics end-to-end

- Bandwidth, maximum end-to-end delay, maximum delay variation
- Signaling and end-to-end connections are required
- Granularity and state information on a per flow basis

QoS versus CoS (2)



Class of Service (CoS)

Groups of flows with same level of treatment

Service-level depends on state of the network and relative priority

- Example: 802.1p (Ethernet): No resources allocated per call, but differentiation of service level through priority. Relative priority, i.e. If network is congested service WILL degrade

CoS is the ability of network devices to prioritize best-effort traffic

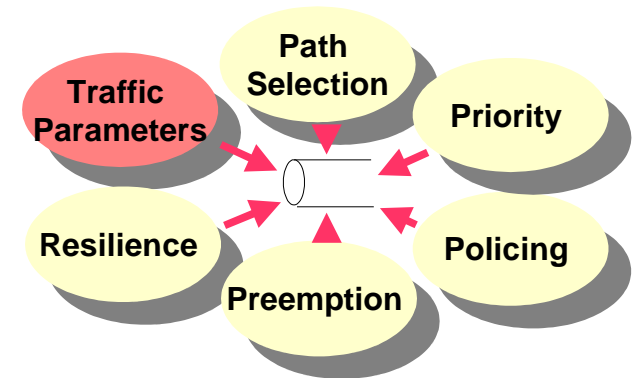
No guarantees, highest class traffic get priority over other traffic

No connections maintained end-to-end

No per-flow state information maintained

Traffic Parameters

Two alternative methods:



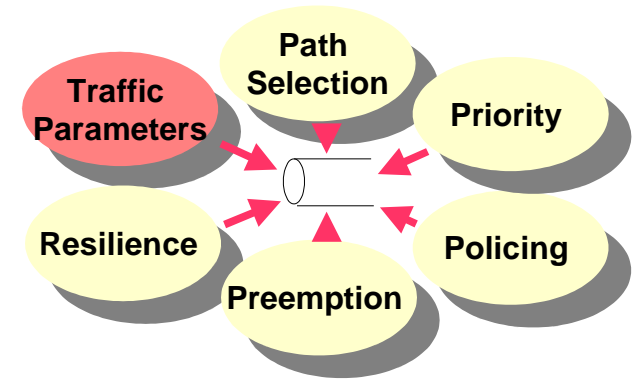
Differentiated Services (DiffServ)

- Implements a number of Classes of Service defining relative priorities
- Each packet in the traffic flow is fitted to one of the classes – no specific signalling used

Integrated Services (Intserv)

- RSVP's end-to-end quality of service model has been extended to cover MPLS label switched paths:
RSVP-TE
- Traffic parameters can be defined for LSPs:
Peak and committed bandwidth, delay prioritization, ...

Traffic Parameter IntServ Example



Task: Establish two LSPs from LER A to LER D

- First LSP with CDR 70 Mbit/s
- Second LSP with CDR 40 Mbit/s

Resource Information Base (RIB) in LER A

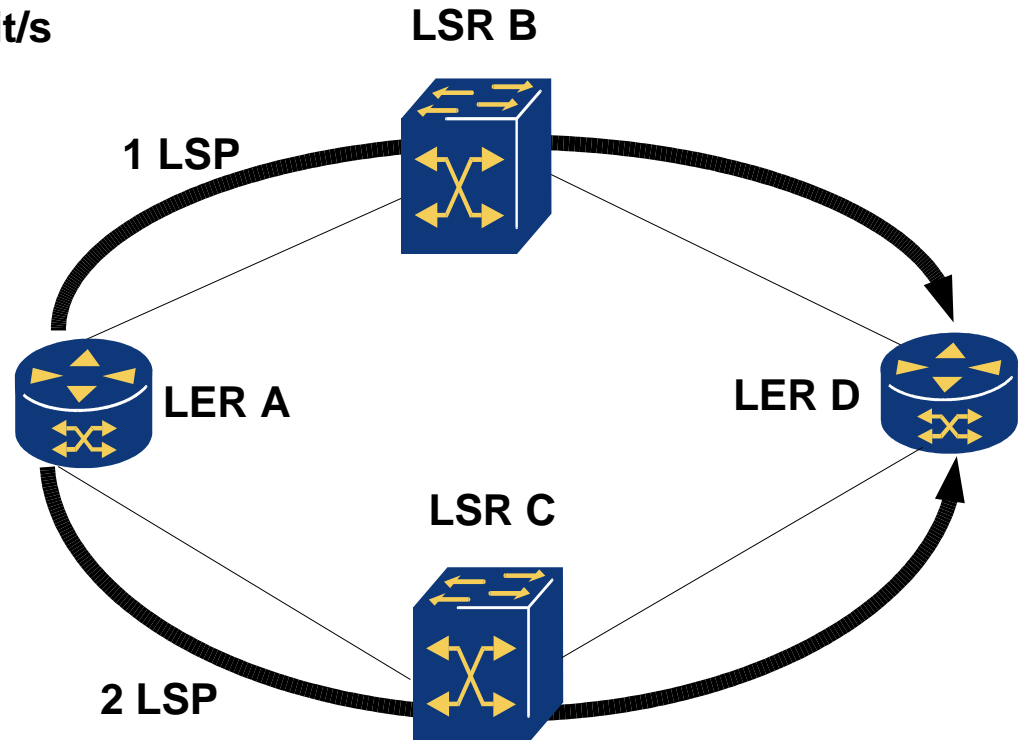
100 Mbit free through LSR B
50 Mbit free through LSR C

RIB after one LSP is established

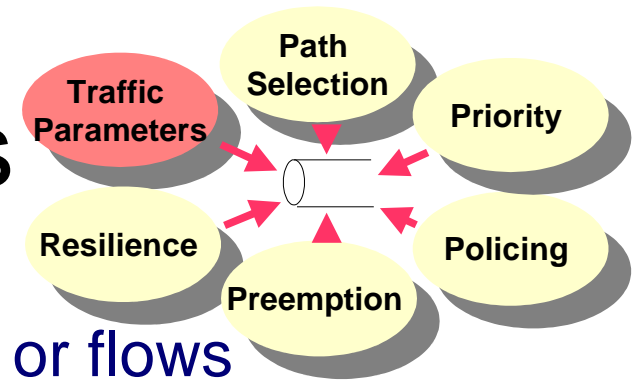
30 Mbit free through LSR B
50 Mbit free through LSR C

RIB after two LSPs are established

30 Mbit free through LSR B
10 Mbit free through LSR C



IntServ Characteristics



Introduces the model of connections or flows

Requires:

- Known QoS requirements
- Signalling protocol (i.e. RSVP-TE)

Significant enhancements on network device:

- Admission control
- Policy control
- Packet classification
- Packet scheduling and queuing
- Packet dropping policy

Priority Attributes

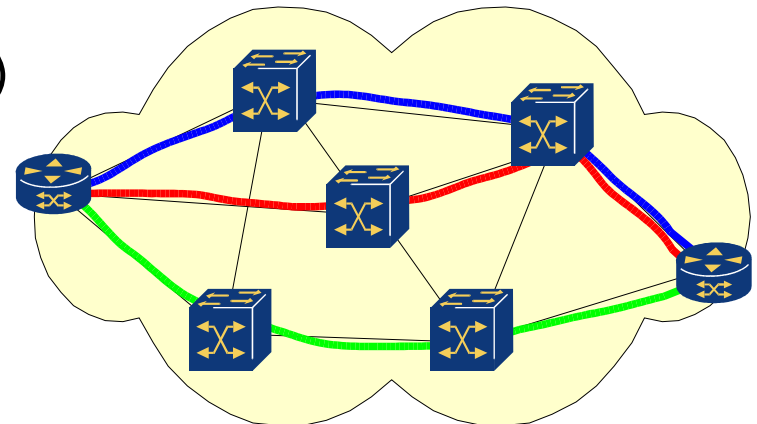
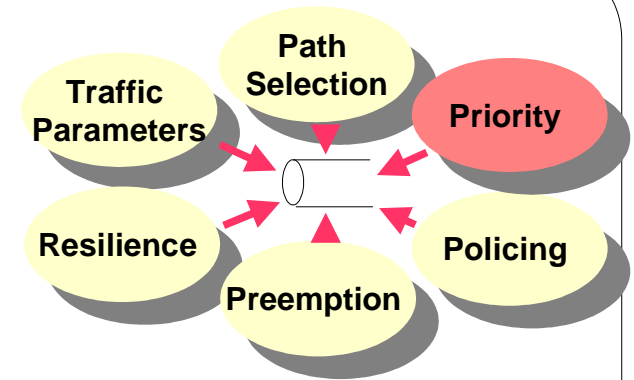
Defines the relative importance of traffic trunks

Used to determine which path (of N) is selected for a given trunk

Priority values: 0, 1, ..., 7, highest priority is 0

Two kinds of priority:

- Setup priority
 - valid at path establishment time
- Hold priority
 - valid throughout path lifetime



- Constraint Routed LSP (highest priority)
- Best Effort LSP (medium priority)
- Best Effort LSP (low priority)

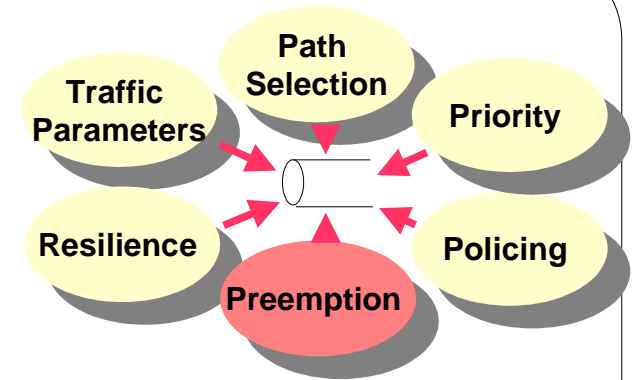
Preemption Attributes

Regulates whether a trunk may preempt another trunk

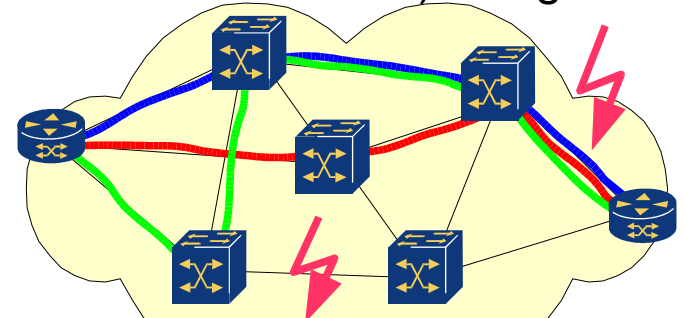
Trunks can be

- Preemptable or not
(may be released if higher priority trunks request new resources)
- Allowed to preempt or not
(may disconnect lower priority trunks for own new resource requests)

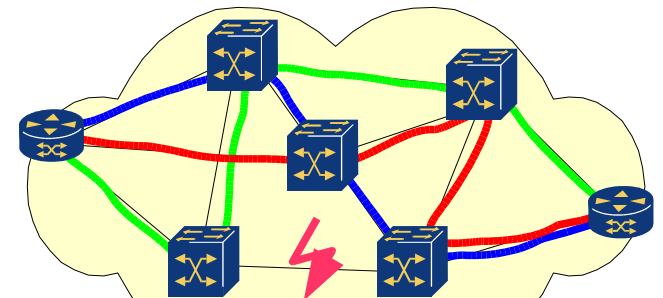
Priority Attributes and Preemption Attributes need to work closely together



2) Congestion

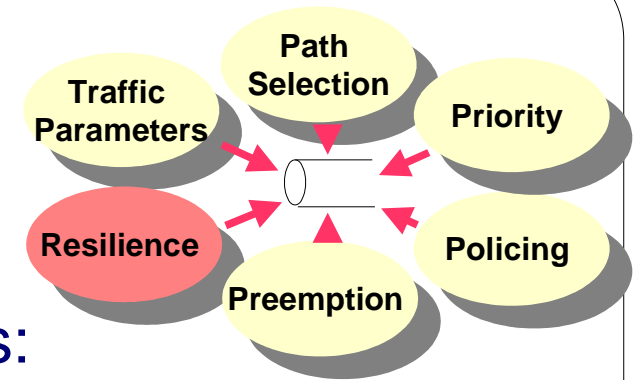


1) Rerouting after link failure



3) Trunk preemption

Resilience Attributes



Define behavior under fault conditions:

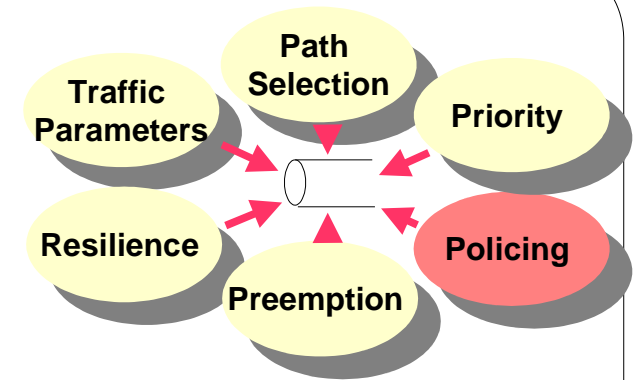
Find backup paths around failing nodes and links
with Fast Rerouting (RFC 3469)

Goal:

Substitute layer 1 protection by layer 3 protection

Improve level of protection (link – node)

Policing Attributes



Control behavior when trunk becomes non-compliant

- Exceeds committed / peak bandwidth

Possible actions

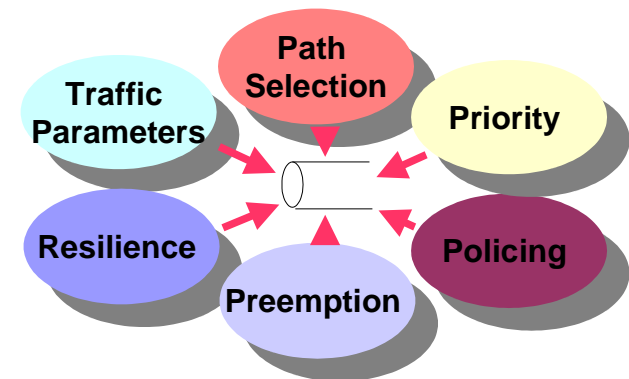
- limit rate (optionally with shaping - smooth data stream)
- tag (remark as “red” packet class) and forward
- just forward without any policing action

Summary

- MPLS Traffic Engineering provides solutions for different application requirements
- TE modules may be used independent of each other
- TE enables MPLS networks to carry legacy data, driving down costs
- MPLS Traffic Engineering is available to improve service quality and network availability

Major task for routers:

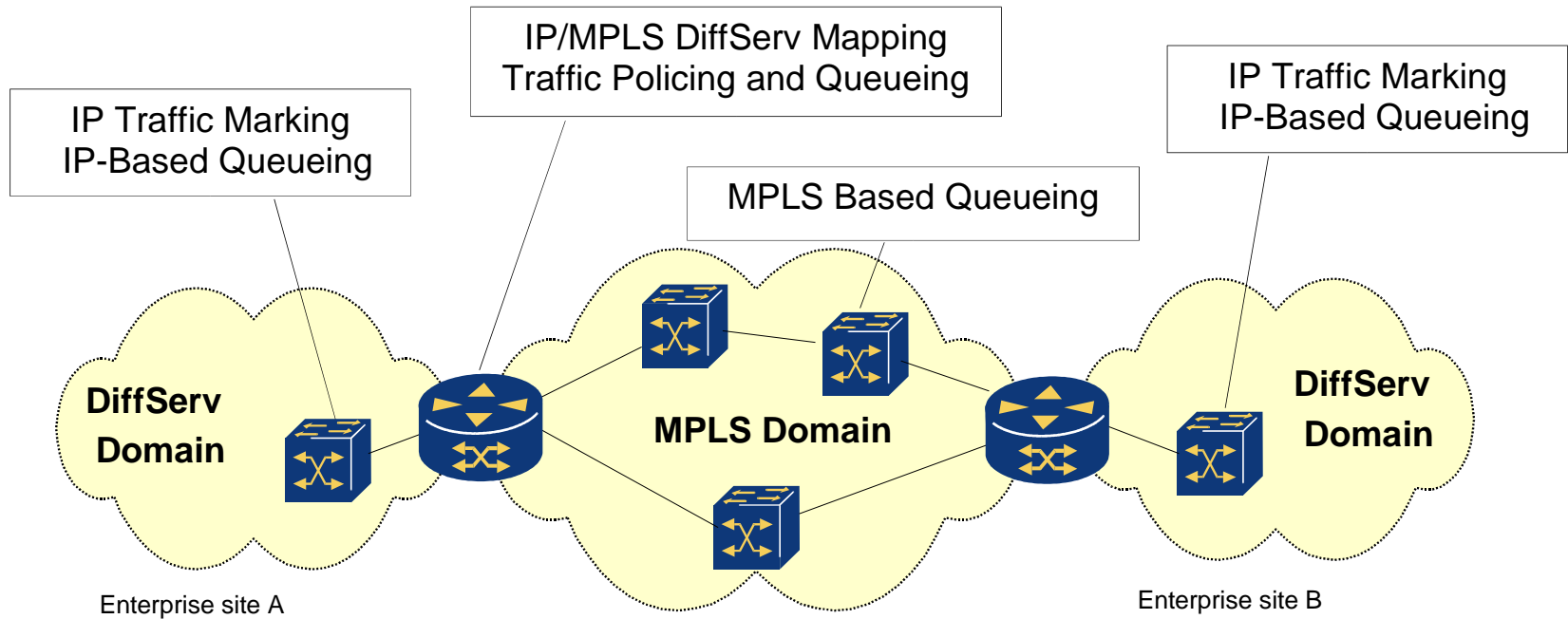
- Implement a complex framework of inter-dependent parameter sets
- None of the parameters is absolutely mandatory in the standard (RFC2702) -
- This is still an area of active development.



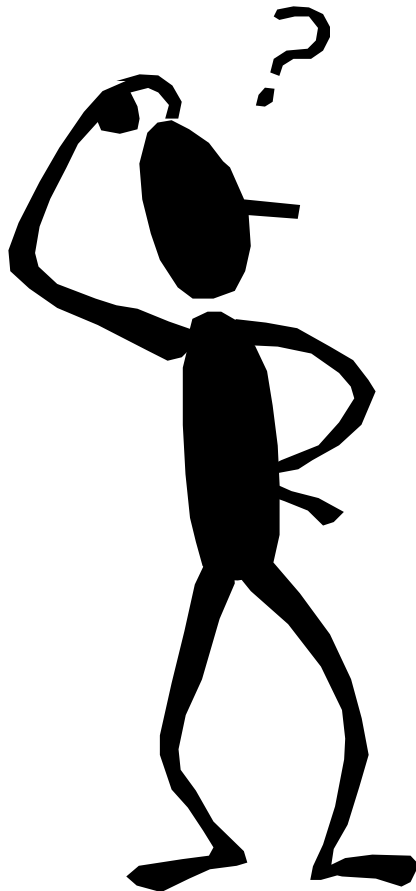
MPLS

And Differentiated Services

Overview: DiffServ in IP and MPLS



Next Pages



IP Differentiated Services

Short introduction

- RFC 2474 / 2475

DiffServ and MPLS

- RFC 3270
- RFC 3564

IP DiffServ Concept

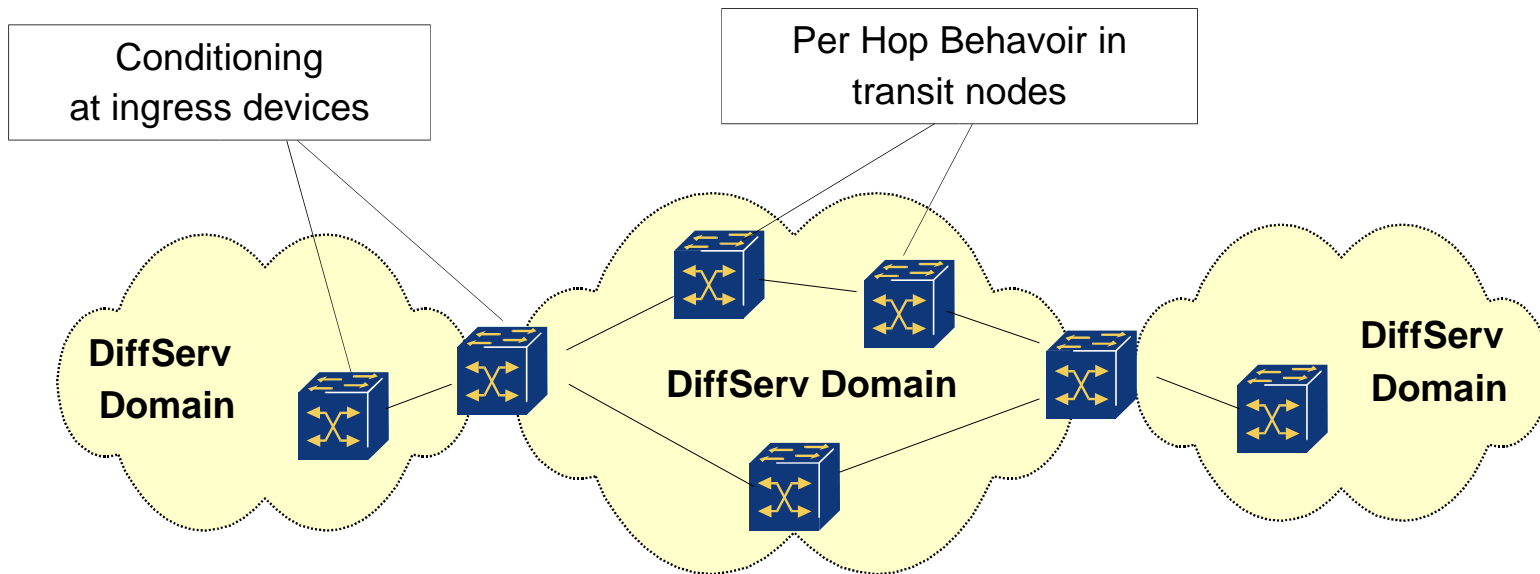
What's the general idea?

An architecture for implementing scalable, stateless service differentiation

A service defines significant characteristics of packet transmission in one direction across a set of one or more paths in the network

- Examples of characteristics: Delay, Jitter, Packet loss
- Relatively simple method to provide different Classes of Service
- Packet flows are classified (marked) at the network boundary
- Forwarding treatment in the network according to specific class
- Service agreements are a local matter of each router

DiffServ Service Concept



Characteristics of Per Hop Behaviors defined per DiffServ Domain

DiffServ Service Classes or Per Hop Behaviors (PHB)

Describes the forwarding behavior applied to an aggregate of flows

The means by which a network-node allocates resources to meet a behavior aggregate

Per Hop Behaviors are implemented via

- Queue management and scheduling (today usually 4-8 queues)
- Buffer size, Queue depth, Over-subscription policy
- Scheduling
- Scheme to determine which queue to service when link is available
- Congestion management and avoidance
- Optimize resource utilization

Where DiffServ Is Active

IP Edge Routers

- Traffic conditioning and policing
- Classification to different service classes
- Marking

IP Core Routers

- Traffic forwarded using simple differentiated treatments

Realized with multiple queuing mechanisms at the network devices

How DiffServ Realizes CoS

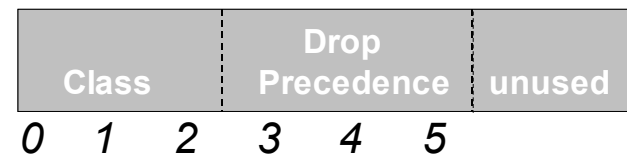
DiffServ stores Class of Service attributes in the IP header and therefore needs

- no signalling
- no additional protocols
- no flow-based resource reservation

VERS	HL	TOS	Fragment Length
Datagram ID		FLAG	
TTL	Protocol	Header Checksum	
Source Address			
Destination Address			
Options (if any)			
Data			

Internet Protocol Version 4 Header

DiffServ field (6 bits)



IP DiffServ Classes of Service

48 Traffic classes in three common groups:

EF = Explicit Forwarding : premium service

- Bandwidth assigned, low loss rate, low latency and low jitter.

AF = Assured Forwarding

- Ensure packet forwarding with high probability
- Three different groups AF1, AF2, AF3 for different loss probability classes

Best Effort

- The common best effort forwarding behavior available in all routers
- Network will deliver these packets whenever resources available
- Packets with an unidentified DSCP should also receives this PHB

DiffServ - Conclusion

What DiffServ provides...

- ✓ Implements prioritization of IP traffic to provide basic Class of Service features for applications
- ✓ No complex protocols, no signaling, simple implementation
- ✓ Smooth integration of existing core networks without having to invest in new hardware
- ✓ Scalability for large networks

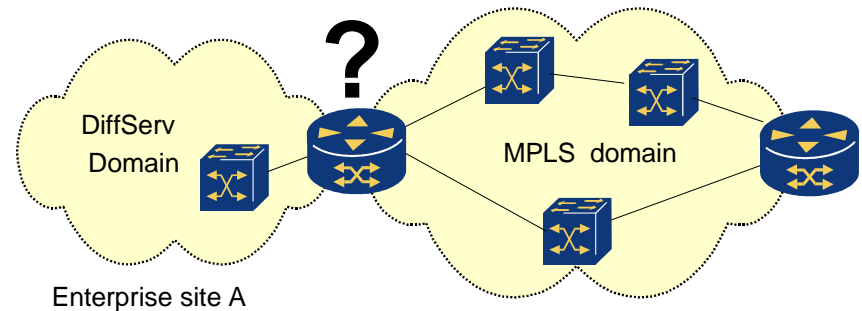
... and what not

- ✗ a solution to provide well-defined end-to-end QoS
- ✗ a technology covering all QoS aspects
- ✗ does not introduce changes to traditional IP routing such as source routing
- ✗ end-to-end service definitions

DiffServ in MPLS Networks

How to differentiate traffic?

- MPLS uses common labels for the whole data stream
- IP DiffServ uses tags in each packet
- BUT: IP packet contents are not accessible inside an MPLS network



Task:

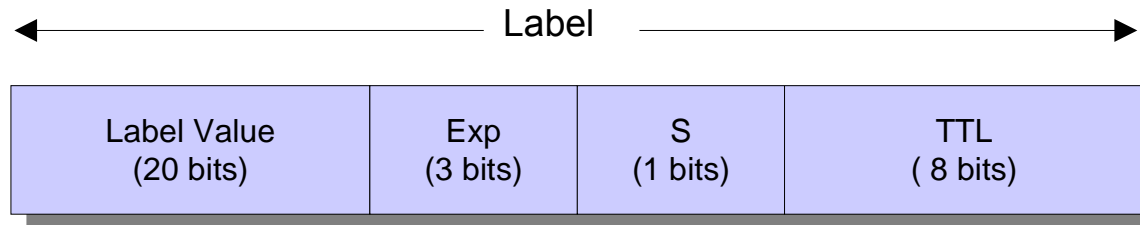
Map IP DiffServ classes to MPLS label switched paths

Mapping DiffServ to MPLS (1)

Mapping between DS Codepoints and MPLS Labels

Two alternatives:

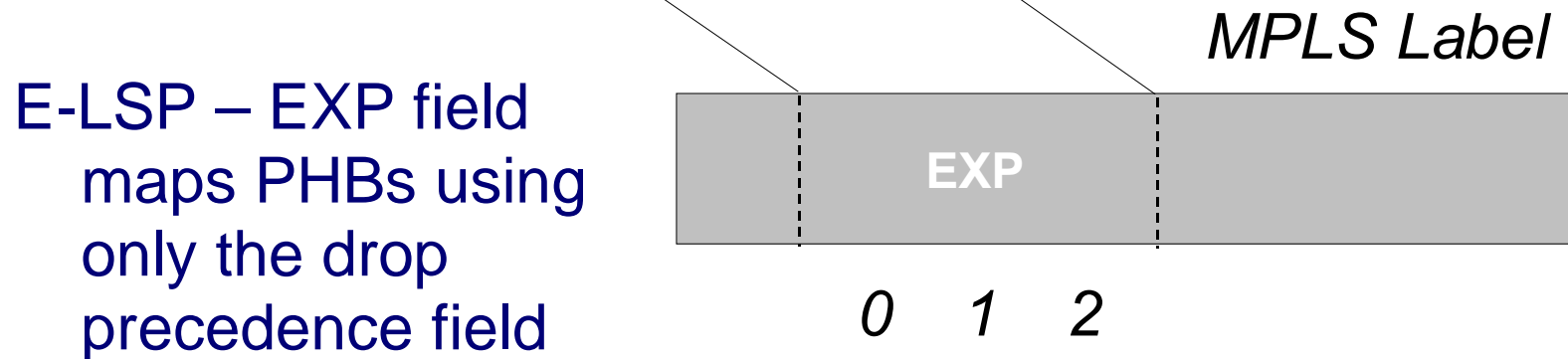
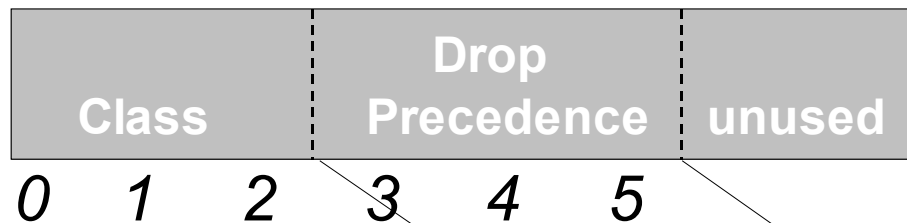
- E-LSPs: DiffServ bits are mapped into the EXP field
 - One LSP is used for all DiffServ classes
 - Experimental (EXP) field in the label transports class information
 - DiffServ classes are mapped into 8 MPLS classes



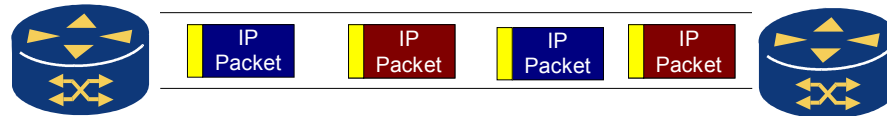
- L-LSPs:
 - Separate LSPs for each DiffServ class

Mapping DiffServ to MPLS (2)

Class – specifies the PHB that packet is to receive
 Drop Precedence - marks relative importance of a packet within a given class



DiffServ - EXP Mapping



First alternative to map DiffServ to MPLS:

E-LSP (EXP-Inferred-Per Hop Behavior Scheduling Class)

- The different ordered aggregates are transported in the Experimental (EXP) bits of an MPLS shim header.
- The next hop interprets scheduling rules (per hop behavior) by inspection of the EXP bits.
- Dominant alternative in IP-router worlds
- **Advantage:**
Label space conservation
- **Disadvantage:**
DiffServ policies need to be advertised to MPLS core backbone

DiffServ – EXP Mapping (2)

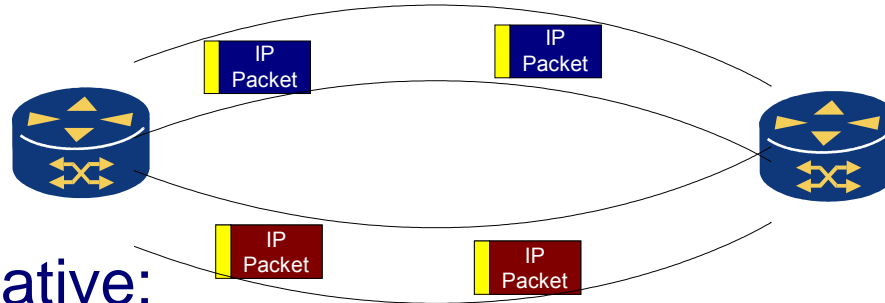
Local vs. Global DiffServ bit interpretation

- DiffServ policies are born in the edge domains
- EXP mapping is also performed at the provider edge
- BUT: EXP interpretation needs to be known throughout the MPLS backbone

Solution:

- New RFC 3270
- Implements RSVP-TE DiffServ object transporting the EXP bit to DiffServ class mapping

DiffServ - LSP Mapping

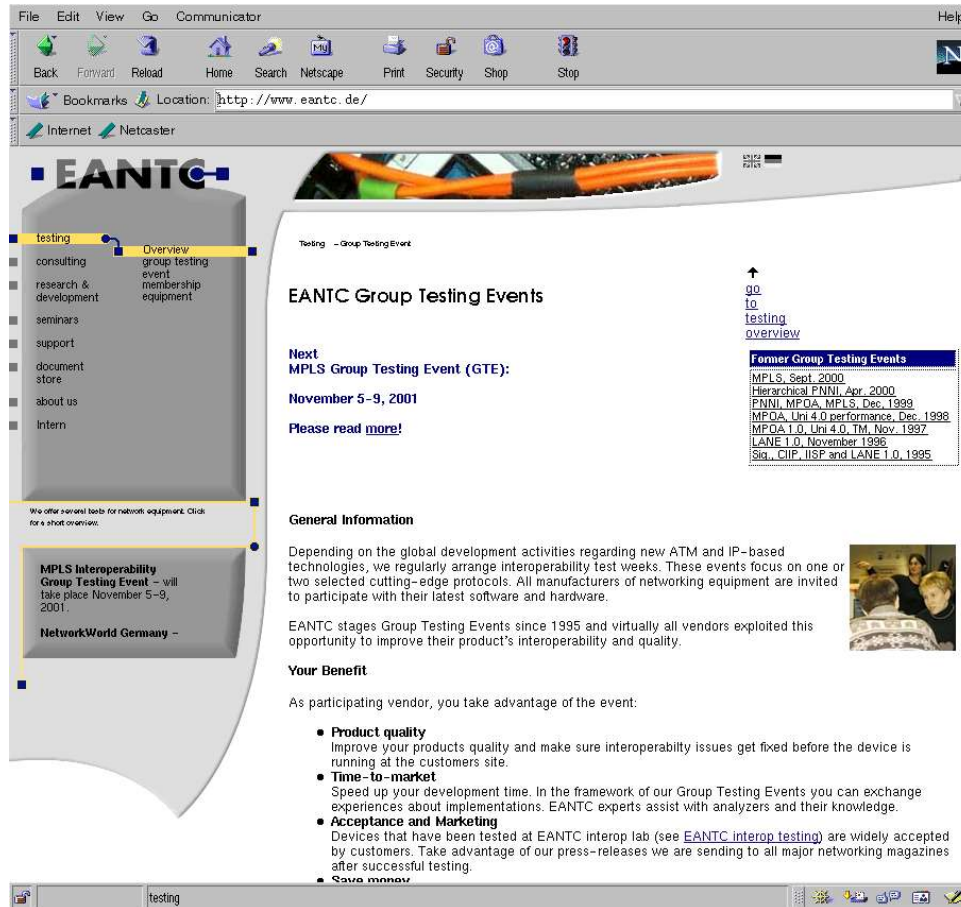


Second Alternative:

L-LSP (Label-Only-Inferred-Per Hop Behavior Scheduling)

- Separate LSPs are established for each single traffic class
- The forwarding treatment is inferred only by the label value, using MPLS RSVP-TE traffic engineering mechanisms
- Preferred by some ATM / Frame Relay switch vendors and carriers
- **Advantages:**
 - Usage of underlying queuing mechanisms per LSP
 - Traffic Engineering per LSP possible
- **Disadvantage:**
 - Needs more LSPs

Thanks! Any Questions?



File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Shop Stop

Bookmarks Location: http://www.eantc.de/

Internet Netcaster

EANTC

testing Overview group testing event membership equipment consulting research & development seminars support document store about us Intern

We offer several tools for network equipment. Click for a short overview.

MPLS Interoperability Group Testing Event - will take place November 5-9, 2001.

NetworkWorld Germany

Testing - Group Testing Event

EANTC Group Testing Events

↑ go to testing overview

Next
MPLS Group Testing Event (GTE):
November 5-9, 2001
 Please read [more!](#)

Former Group Testing Events

MPLS, Sept. 2000
Hierarchical PNNI, Apr. 2000
PNNI, MPOA, MPLS, Dec. 1999
MPOA, Uni 4.0 performance, Dec. 1998
MPOA 1.0, Uni 4.0, TM, Nov. 1997
LANE 1.0, November 1996
Sis, CIP, IISF and LANE 1.0, 1995

General Information

Depending on the global development activities regarding new ATM and IP-based technologies, we regularly arrange interoperability test weeks. These events focus on one or two selected cutting-edge protocols. All manufacturers of networking equipment are invited to participate with their latest software and hardware.

EANTC stages Group Testing Events since 1995 and virtually all vendors exploited this opportunity to improve their product's interoperability and quality.

Your Benefit

As participating vendor, you take advantage of the event:

- **Product quality**
 Improve your products quality and make sure interoperability issues get fixed before the device is running at the customers site.
- **Time-to-market**
 Speed up your development time. In the framework of our Group Testing Events you can exchange experiences about implementations. EANTC experts assist with analyzers and their knowledge.
- **Acceptance and Marketing**
 Devices that have been tested at EANTC interop lab (see [EANTC interop testing](#)) are widely accepted by customers. Take advantage of our press-releases we are sending to all major networking magazines after successful testing.
- **Save money**

Please visit our homepage to read more about high speed networking:

www.eantc.com