Advanced Networking Concepts
Quality of Service (QoS)

Contents - Advanced Networking - QoS Concepts

- Introduction
- Components of an ideal QoS Architecture
- Current QoS Concepts in IP networks
  - DiffServ
  - IntServ
  - MPLS QoS
- Further QoS Aspects
  - Cross-layer QoS
  - Cross-domain QoS
  - QoS optimization during network operation

Introduction

Motivation

- The internet is increasingly getting part of business infrastructure:
  - support of Virtual Private Networks (VPN)
  - common access for customers (e.g. home banking)
  - commercial services (music download, VoIP, IPTV, games, …)

- Problems:
  - missing support of QoS requirements of services/applications (e.g. guaranteed throughput, real-time requirements: limited delay and delay variation) → QoS is heavily dependent on current network load
  - lack of resources (capacity) in the Internet

What is Quality of Service?

- A service is associated with specific requiremenzs on network quality - this covers various aspects
- The Quality of Service is described by a set of (Quality of Service) parameters characterizing the service requirements in detail:
  - performance, operational or functional parameters
  - level of assurance per parameter
    - guaranteed (deterministic)
    - statistical
    - no guarantee (best effort)

QoS Definitions

- ITU-T definition: 3 aspects
  - **Quality of Service (QoS):**
    Network performance from point of view of a service user (e.g. quality of a voice or video connection) – see also ITU-T standard E.800 ("QoS is the collective effort of service performances which determine the degree of satisfaction of a user of the service")
  - **Grade of Service (GoS):**
    Part of QoS, which directly depends on network dimensioning and mechanisms inside the network (example: how often occurs a call blocking event in case of telephone calls)
  - **Network Performance:**
    Ability of a network to provide a requested or agreed service (examples: bitrate/throughput, max. paket loss probability)
- IETF definition:
  - No specific distinction as with ITU-T – common term "QoS" for everything
- ISO definition:
  - "A set of qualities related to the collective behavior of one or more objects"

Quality of Service (QoS) vs. Quality of Experience (QoE)

- Quality of Service:
  - derived by **objective** (quantitative) QoS analysis
  - verifiable by measurement (mean value and confidence interval)
  - example: QoS analysis by checking all required QoS parameters (bitrate, packet delay, packet loss probability, call blocking probability, etc.)

- Quality of Experience:
  - derived by **subjective** (qualitative) QoS analysis
  - possibly verifiable by comparison (relative statement), influenced by subjective human perception
  - example: QoE analysis of a speech communication by means of the Mean Opinion Score (MOS) method (ITU-T P.800)

Quality of Service Influencing Factors

- The Quality of Service perceived by a user may depend on many factors:
  - data manipulation of the application (e.g. audio or video coding)
  - **network performance**
  - end system performance
  - protocol properties (all layers) – in particular protocol mechanisms which entail adaptive reactions in other layers
- In this lecture we consider:
  - **Quality of Service concepts regarding the (IP-)network layer**
    - influencing the following **performance related** Quality of Service parameters:
      - bit rate/throughput
      - packet delay
      - packet delay variation (jitter)
      - packet loss
    - influencing the following **operational** Quality of Service parameters:
      - availability (resilience)

Quality of Service Parameters

- Performance related Quality of Service parameters:
  - bitrate / throughput
  - end-to-end delay
  - delay variation (delay jitter)
  - Reliability (transmission errors, packet loss, doubled packets, …)
  - Delay and error probability of connection establishment
- Operational Quality of Service parameters:
  - availability
  - coverage / reach
  - ...
- General (functional) Quality of Service parameters:
  - level of security
  - synchronisation quality
  - ...

Quality of Service Assurance Levels

- No guarantee (best possible assurance = best effort):
  - common in the Internet today
  - no resource allocation
  - deal with the available resources as effectively (and fast) as possible
  - resource conflicts may occur frequently
- Statistical assurance:
  - Quality of Service parameters are only guaranteed in a statistical sense, i.e. with a specific probability, e.g. 80% of packets suffer a delay < 100 ms
  - resources are over-allocated to a certain extent (overprovisioning)
  - resource conflicts are possible (however conflicts decrease with growing probability of assurance)
- Deterministic assurance:
  - Quality of Service parameters are always met (rigorous assurance)
  - resources are exclusively allocated to the users
  - no conflicts occur but requests might be blocked if no resources are available

Applications and Quality of Service Requirements

- Classification of applications:
  - traditional (TCP based) applications → elastic traffic
  - real-time applications → stream-type traffic
- These classes have different requirements on the Quality of Service provided by a communication network

“elastic” traffic

Characteristics:
- file transfer
- bitrate depends on the network load (TCP rate adaption: 0 to max. transmission rate)

Main QoS requirements:
- throughput during file transfer

Application examples:
- Web browsing, FTP, Email

“stream-type” traffic

Characteristics:
- continuous transmission
- bitrate depends on the application

Main QoS requirements:
- minimum bitrate/throughput
- delay, jitter
- packet loss probability

Application examples:
- voice telephony, video telephony, audio/video streaming

Applications with Elastic Traffic Characteristic

- In general, no fixed upper time limit for data delivery
  - data is not unusable in case of longer transmission times
  - these applications can be realized without problems in the Internet today
- But: some elastic applications may demand for a minimum bitrate (throughput) during data transmission (to limit the transmission time)
- The delay requirements of elastic applications may vary heavily – following three categories (with increasing tolerance wrt. delay) can be distinguished:
  - Interactive (e.g. telnet): little tolerance towards delay
  - Interactive, bursty (e.g. FTP): quite tolerant towards delay
  - Asynchronous (e.g. Email): very tolerant towards delay

⇨ **Usually less sensitive wrt. packet loss, delay and jitter; Quality of Service is proportional to throughput during data transmission**

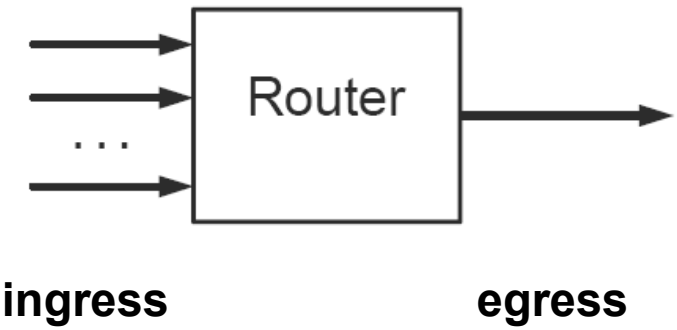
Applications with Stream-Type Traffic Characteristic

- Hard real-time requirements for data delivery
  - example: audio and video applications (VoIP, video streaming)
- Requirements wrt. minimum bitrate and upper limit for packet loss
- Possibilities to adapt to QoS fluctuations:
  - inclusion of redundant information into the data stream
    - by that a certain tolerance concerning packet loss is achieved
    - this also may help against late arrival of packets (which are regarded as lost)
  - adaptivity of the application:
    - adaptation wrt. packet delay: adjustment of the jitter puffer size
    - adaptation wrt. throughput: adjustment of the required bitrate (e.g. by use of adaptive codecs)

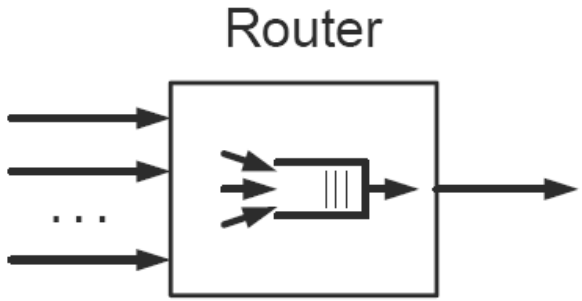
⇨ **Usually sensitive towards packet loss, delay and jitter if no application-layer redundancy mechanisms (e.g. robust codecs) or adaptation mechanisms (e.g. adaptive codecs, adaptive jitter buffer) are applied;**
A minimum bitrate/throughput has to be guaranteed!
Problem: in the current Internet no guarantees are possible

Fundamental Issues with QoS in IP Networks

- No overload as long as the capacity of the routers' egress interface is not exceeded
  - no QoS problem in this case



- Short-term overloads can be compensated by packet buffers (at the routers' egress interface)
  - typically in the range of some ms
  - this leads to slight packet delays and delay variations (as the buffer fill level fluctuates)



- Long-term overloads lead to congestion situations
  - major packet delays occur due to long waiting times in the buffer
  - also packet loss may occur due to buffer overflow (as the buffer size is limited)

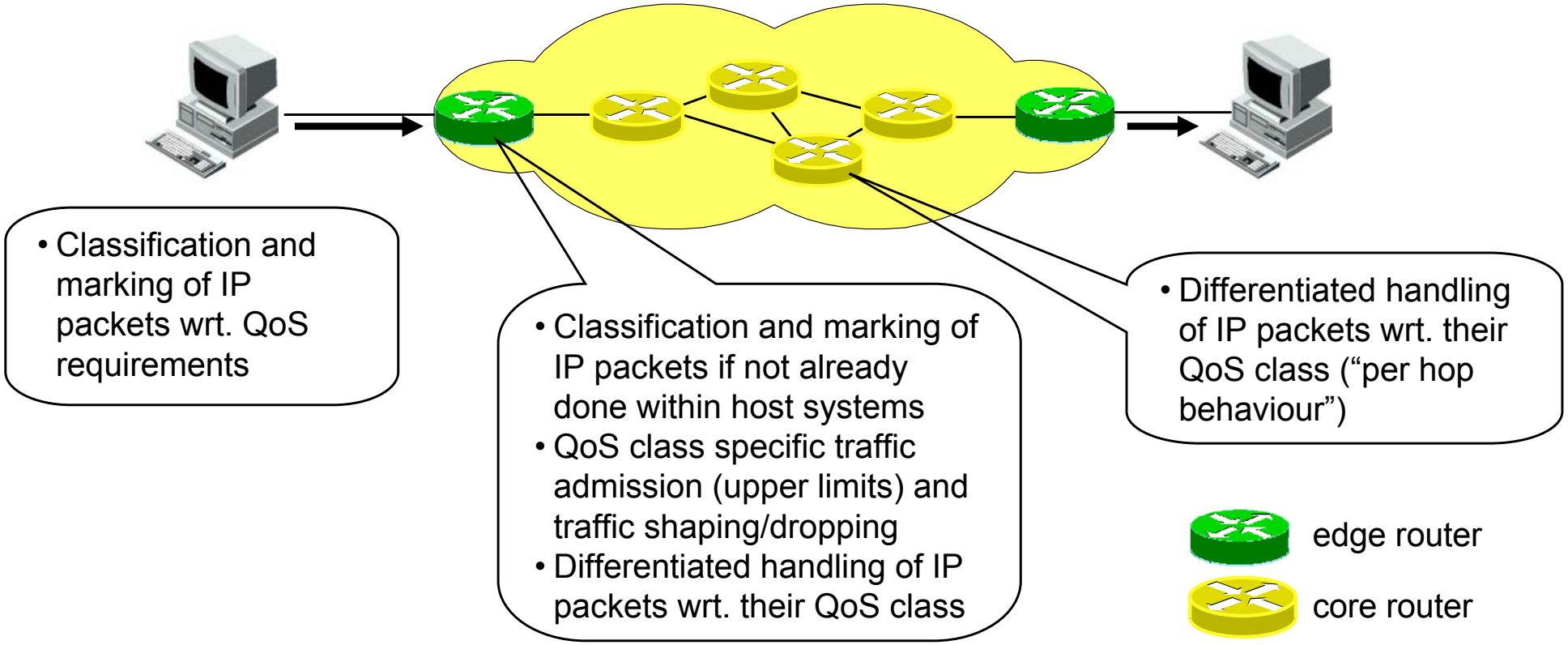
Measures to guarantee QoS in IP Networks

- Adaptive and robust applications: adaptation to the network condition
  - problem: nevertheless a minimum QoS has to be provided by the network
- Over-provisioning: congestion situations occur never or with a very small probability
  - problems:
    - Internet traffic flow patterns and traffic growth are hard to predict (but are necessary as input for dimensioning)
    - overprovisioning is not economical
- Traffic classification and differentiated packet handling in routers
  - problem: no resource reservation → does not help if congestion is caused by traffic of same priority (only relative QoS guarantee)
- Resource reservation and admission control (AC)
  - problem: high effort

⇨ **These measures might be applied in combination; practical QoS concepts for IP networks (DiffServ concept, IntServ concept) use a mixture of these measures**

Traffic Classification / Differentiated Packet Handling

- Realized by buffer scheduling strategies in routers
- Taking into account different QoS requirements (packet delay and packet loss) of the traffic streams
- Basic architecture:



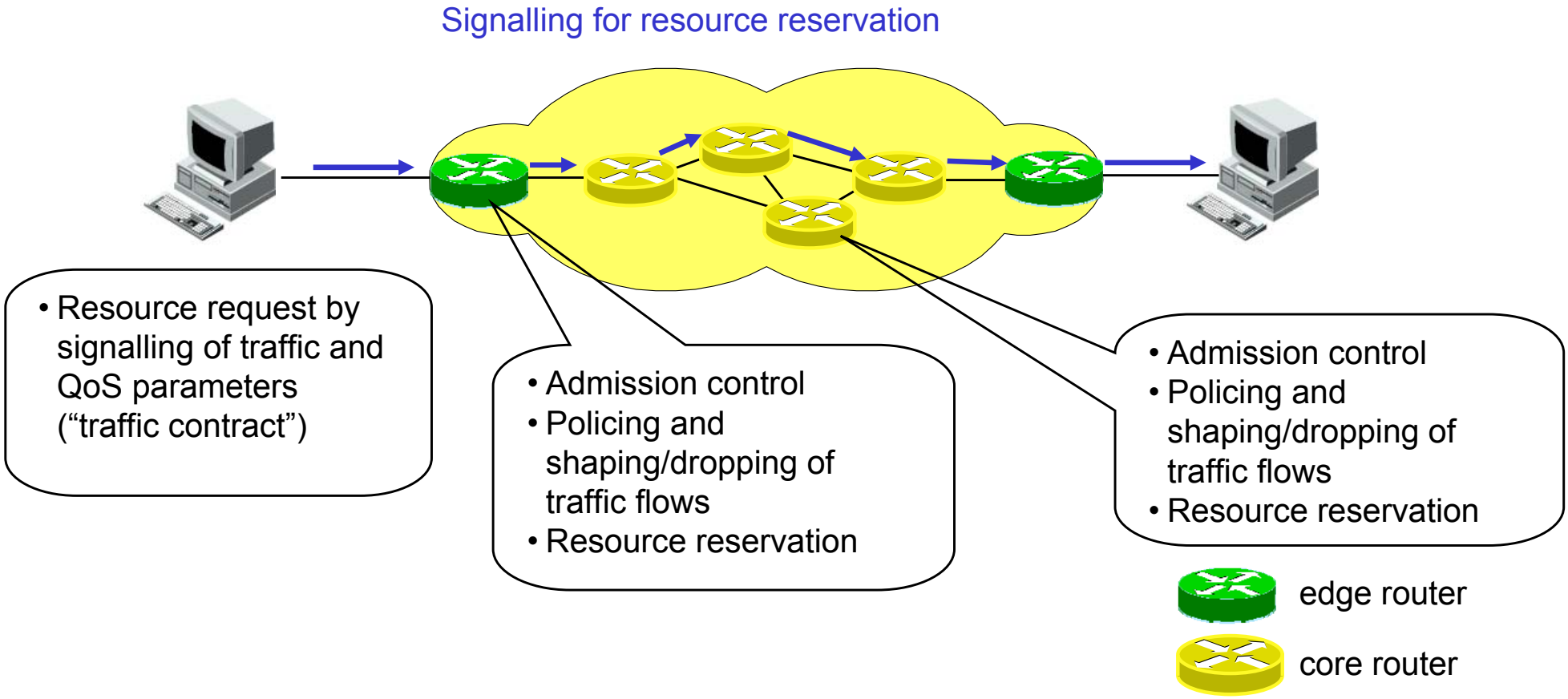
Resource Reservation and Admission Control (AC)

- Traffic flow specific resource reservation in order to guarantee QoS
- Resources:
  - network resources: capacity
  - system resources: buffer space, computation time
- The admission control (AC) function blocks new traffic flows in case of resource shortage and therefore prevents QoS degradation of already existing traffic flows
- Signalling protocol (for resource reservation and admission control) necessary
- Many possible variants:
  - resource reservation: for each traffic flow or for aggregated traffic; a priori reservation or on-demand reservation via signalling
  - locations of admission control: at the network edge (only at ingress routers or coordinated between ingress and egress) or at each router interface
  - control plane: in routers or outside; centralized or decentralized realization

Resource Reservation and Admission Control (AC)

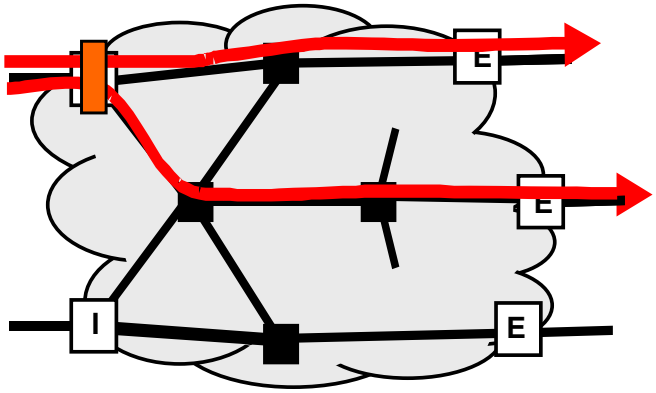
- Basic architecture:

  Example: - resource reservation per traffic flow
  - admission control on each link

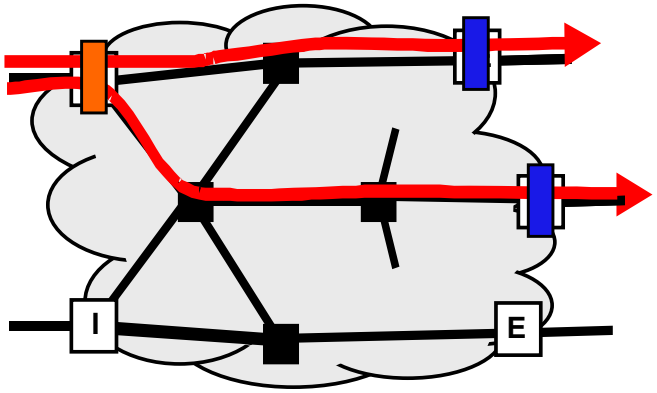


Admission Control - Placement of Admission Control Points

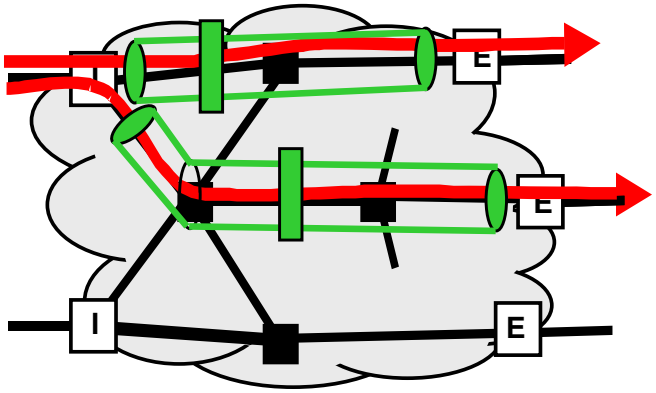
AC at ingress



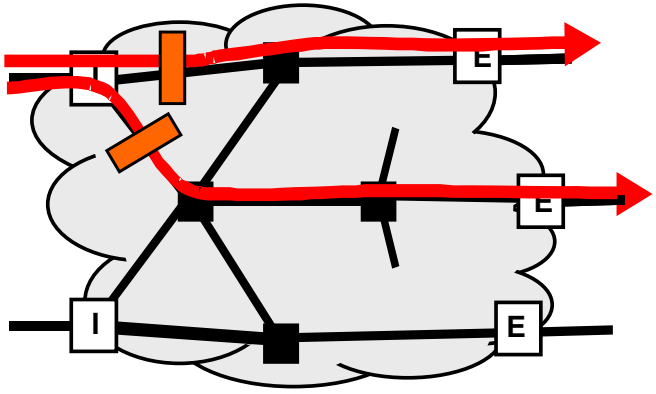
AC at ingress and at egress



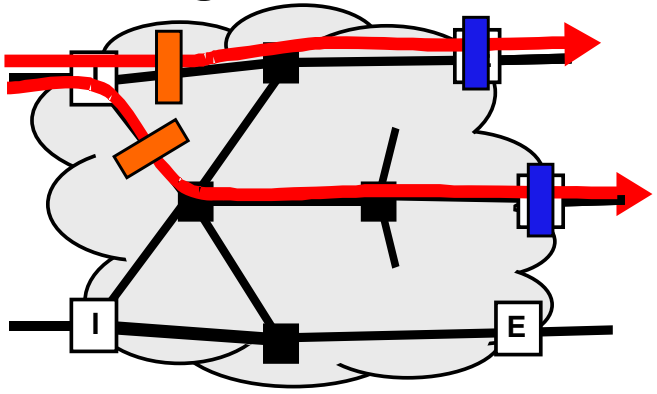
AC per ingress-egress pair



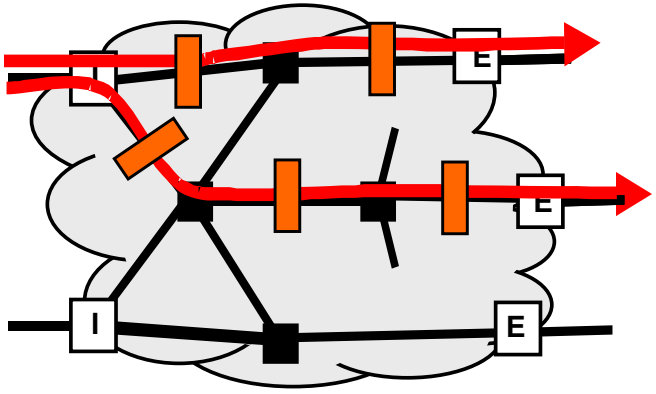
AC at ingress per core interface



AC at ingress per core interface and at egress



AC per link



Resource Reservation / AC vs. Over-Provisioning

- Over-provisioning:
  - + simple and easy approach
  - – a lot of a capacity is required to ensure the QoS of existing traffic flows in all (even in rare) situations (e.g. link failures, load imbalances)

- Resource reservation / admission control:
  - + absolute QoS guarantee for existing traffic flows
  - + robustness against unplanned high loads, failures (of nodes and links) and load imbalances
  - + fairness: fair resource allocation can be enforced
  - + capacity savings compared to over-provisioning; however strong dependency on the type of admission control
  - – higher complexity compared to over-provisioning
  - – exact description of resource requirements required (difficult for elastic applications)

Application Scope of QoS Mechanisms

- QoS specific handling of data packets within network nodes (buffer scheduling) but no QoS specific routing (QoS independent next-hop selection)
  - QoS independent routes

- QoS specific handling of data packets in network nodes and QoS specific routing (QoS specific next-hop selection)
  - different (QoS specific) routes for traffic with different QoS requirements – several (QoS specific) routing tables necessary

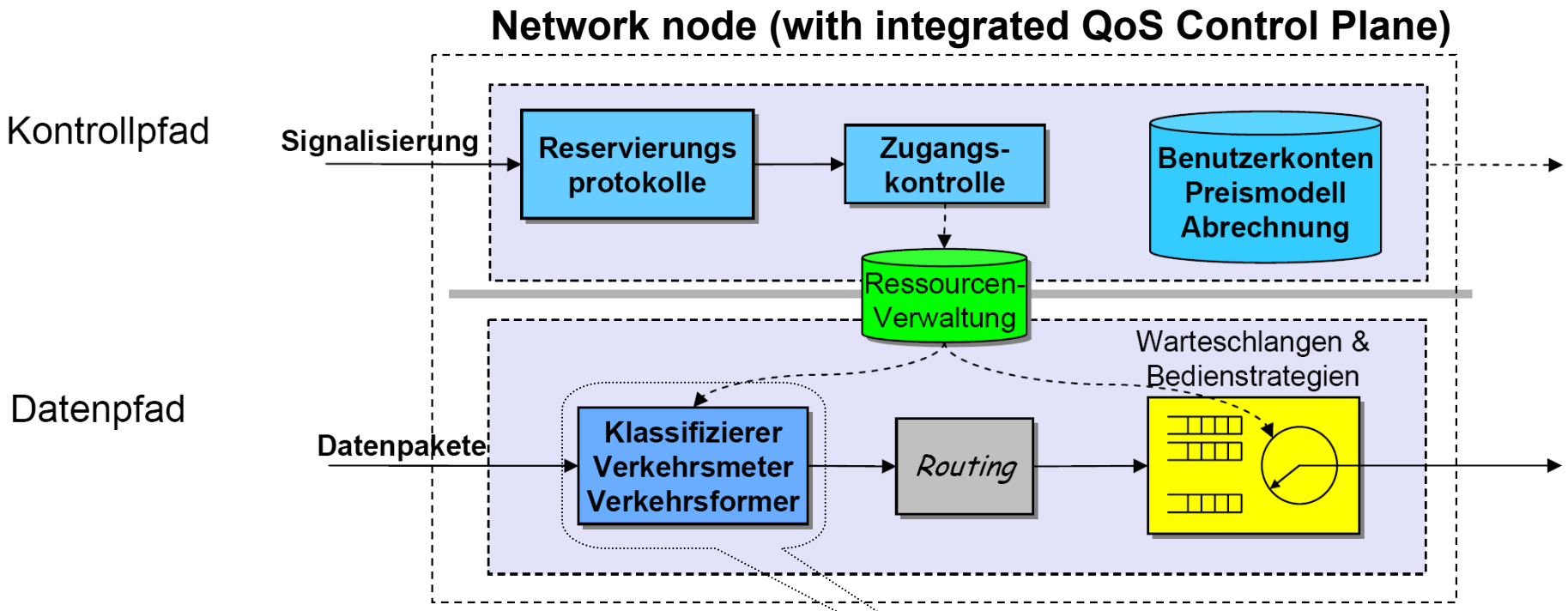
- QoS specific traffic aggregation (e.g. in QoS specific MPLS tunnels) and traffic aggregate (tunnel) specific routing and handling of data packets within network nodes
  - different (QoS specific) routes for traffic aggregates with different QoS requirements

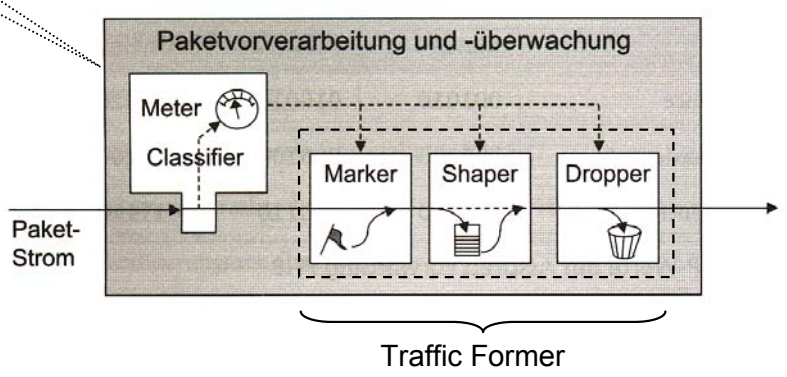
Components of an ideal Quality of Service Architecture

General Quality of Service Provisioning Taks

- Specification of Quality of Service requirements and traffic characteristics
  - via Quality of Service and traffic parameters
- QoS negotiation and agreement with the network (traffic contract, Service Level Agreement)
  - the admission control function has to check whether enough network resources are available
- Resource allocation within the network
- QoS monitoring and maintenance
  - QoS monitoring is necessary because the network operator normally has to provide evidence about his of QoS compliance
  - QoS maintainance during network operation via traffic/network engineering mechanisms
  - notification of customers, if QoS cannot be fulfilled anymore → a new QoS negotiation might be required
- Monitoring of traffic characteristics (Policing)
  - traffic Policing is normally performed at the ingress (and if necessary also inside the network)

QoS Components within Network Nodes - Overview





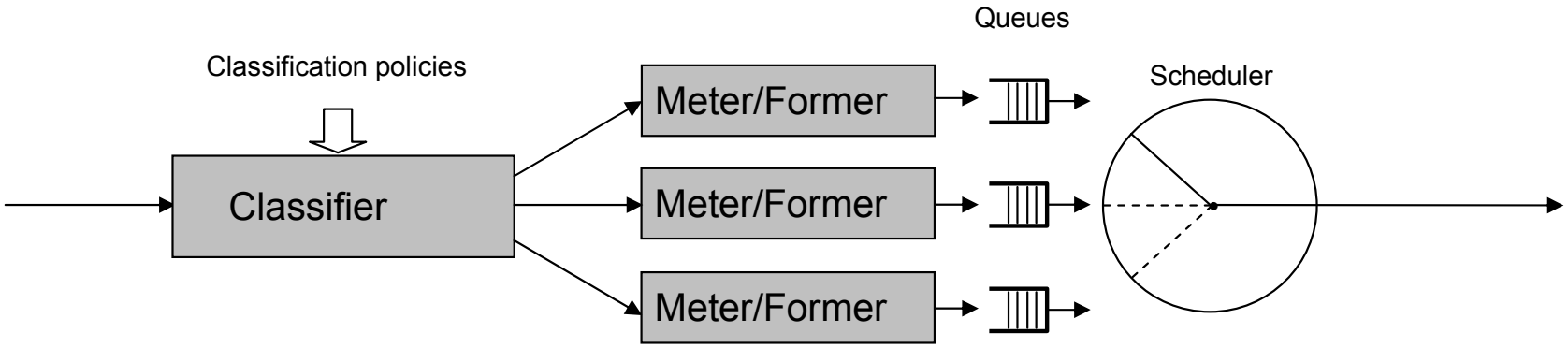
- **Components that are in the data path**
  - traffic classifier and meter, traffic former
  - packet handling in buffer
- **Components that are in the control path**
  - resource control
    - admission control
    - resource allocation (reservation protocols)
  - control of tariffing / billing, user accounts

Components in the Data Path - Overview

- Traffic classifier
  - maps packets to traffic classes with certain QoS properties/requirements
- Traffic meter
  - checks whether the traffic is conform to the traffic contract
- Traffic former
  - modifies the characteristic of a traffic stream/aggregate so as to keep the traffic conform to the traffic contract
- Packet handling in buffer
  - packet enqueueing
    - sorting of packets into different (virtual) queues (depending on their QoS)
  - queue management
    - packet storage and retrieval strategies
    - active / passive queue management
  - packet dequeueing (scheduling)
    - determines the queue from which the next packet is retrieved

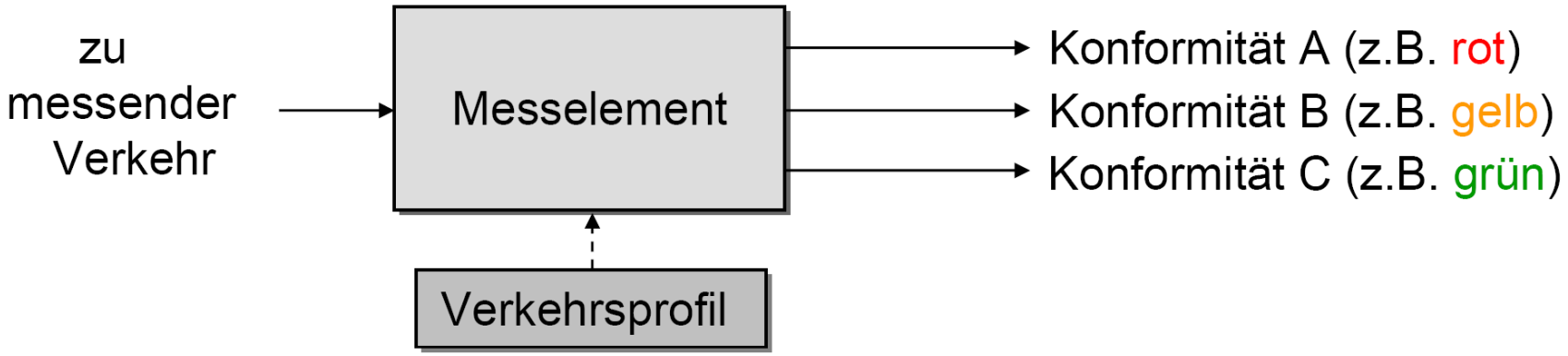
Traffic Classifier

- The traffic classifier maps packets to traffic classes with certain QoS properties/requirements
- Classification policies (examples):
  - classification derived from marking in packet header (Aggregate Classifier), e.g. DiffServ Code Point (DSCP)
    - suitable for aggregated traffic – many traffic flows experience the same handling
    - very easy and fast classification
  - classification derived from multiple packet header fields (Multi-Field Classifier), e.g. IP address, IP protocol field, port number
    - very complicated and complex classification
  - classification derived from ingress interface



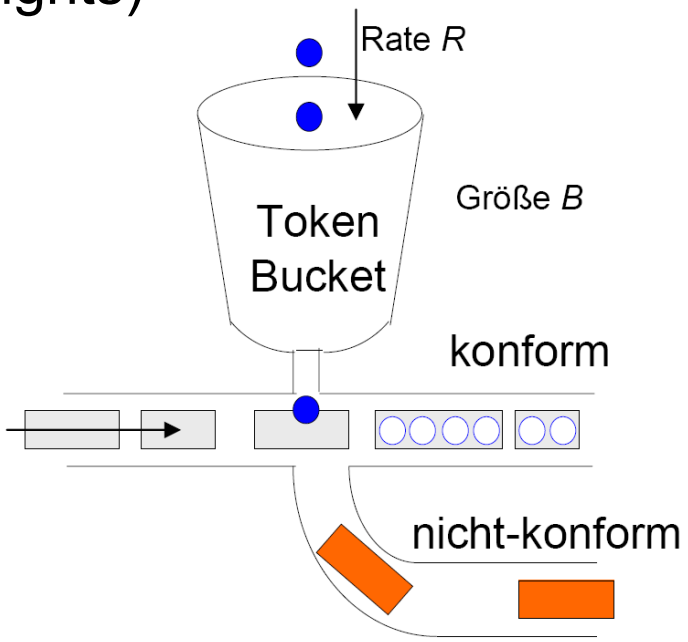
Traffic Meter

- The traffic meter monitors the traffic:
  - it checks whether the traffic is conform to the traffic contract
  - no active manipulation of traffic
- Inputs:
  - traffic profile (traffic parameters) as specified in the traffic contract
  - traffic flow
- Outputs:
  - conformity decision: conform, not conform
  - multiple conformity classes are possible e.g. {0,..,n} or {red, yellow, green}
- Traffic metering methods (examples):
  - Token Bucket method
  - average rate meter (moving window) method



Traffic Meter - Token Bucket Method

- The bucket can contain max. B token (sending rights)
- Token drop with rate R into the bucket

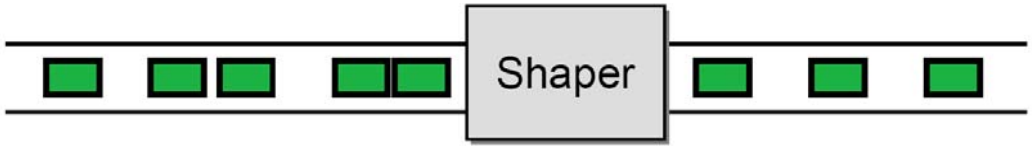


Checks conformity of traffic flows wrt. bitrate R (bit/s) and (burst) tolerance B (byte)

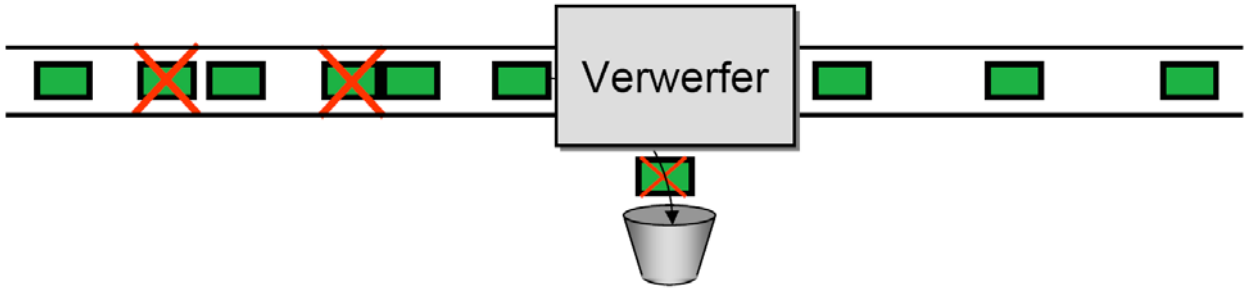
- A packet is marked to be conform, if there are still token in the bucket after the packet arrival (often: 1 byte = 1 token) - otherwise the packet will be marked as non-conform
- The token bucket only monitors the traffic, no shaping is performed; therefore packet bursts remain (if enough token are available)

Traffic Former (Shaper / Dropper / Marker)

- Traffic shaper
  - flattens the traffic stream so as to keep conformity to the traffic contract



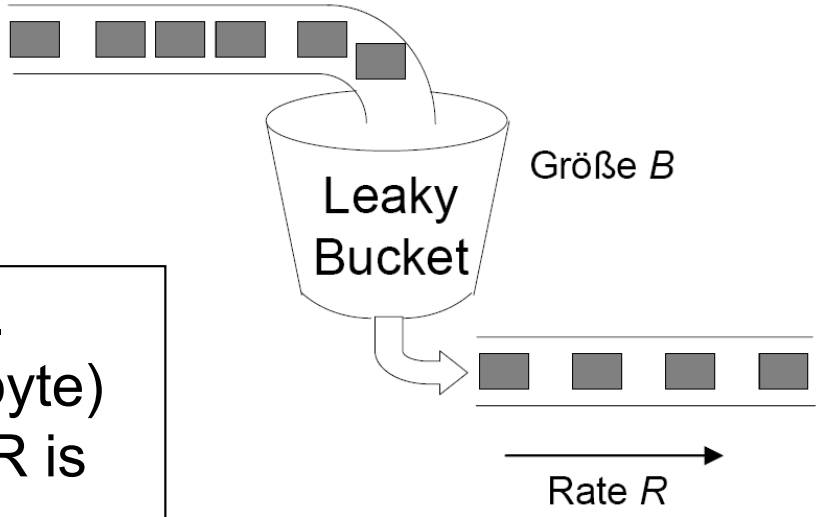
- Traffic dropper
  - selects non-conform packets and drops them



- Traffic marker
  - selects non-conform packets and marks them as lowest priority

Traffic Former - Leaky Bucket Method

- The bucket contains max. B Bytes
- The packets enter the bucket and leak out with desired rate R (bit/s)



Checks the conformity of traffic flows wrt. bitrate R (bit/s) and (burst) tolerance B (byte) and flattens the traffic so that the bitrate R is not exceeded

- A packet of size L byte is sent, if L bytes are leaked out of the bucket; if a new incoming packet doesn't fit into the bucket, it will be dropped or marked as non-conform
- Packets in the bucket are delayed until the transmission is conform to the bitrate (but this causes additional delay!)

Packet Handling in Buffer - Enqueueing & Queue Mngmt.

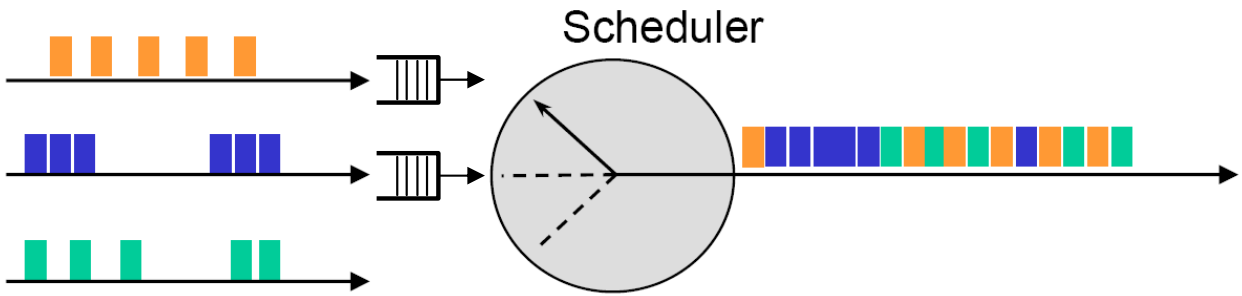
- Depending on their class, packets are sorted into different (virtual) buffers (queues); for each queue different storage and retrieval strategies might be applied:
  - wrt. the order of arrival (**First In First Out, FIFO**)
  - wrt. temporal priority (**Earliest Deadline First, EDF**)
- Passive Queue Management:
  - packets are dropped, if the buffer is full
  - variants: drop tail (new arriving packets will be dropped); drop-from-front (first packet in queue will be dropped); push-out (last packet in queue will be dropped in favour of a new arriving one)
- Active Queue Management:
  - preventive, load-dependent dropping of packets before the maximum buffer fill level is reached
  - variants: Random Early Discard (RED), Weighted RED (WRED), AVQ, REM

Packet Handling in Buffer - Passive vs. Active Queue M.

- Active queue management allows a better congestion control compared to passive queue management:
  - prevention of simultaneous rate reductions (and subsequent rate increases) of many traffic flows (**Global Synchronisation Problem with TCP**) by means of probabilistic dropping of packets of different traffic flows
  - prevention of **Lock-Out** and **Full-Queue** problems
- Problem:
  - both mechanisms rely on the fact, that nodes or end systems in downstream direction react on the increasing packet loss with rate reduction (**implicit congestion control**); Examples: TCP protocol, rate-adaptive applications
- Alternative solutions:
  - **explicit control by signalling the congestion** to neighbor nodes in upstream direction (up to end systems); Example: **Explicit Congestion Notification (ECN)**
  - admission control and resource allocation for individual traffic flows or aggregated traffic + appropriate network dimensioning → a priori prevention of overload and congestion (no further congestion control/adaption necessary)

Packet Handling in Buffer - Scheduling Strategies

- The scheduler determines the queue from which the next packet is retrieved (i.e. is send to the egress interface):
  - by that, the resource (bandwidth) allocation for different traffic classes can be controlled
- Scheduling strategy examples:
  - Round Robin:
    - all queues are served one after another – no queue is preferred
  - Simple Priority Queueing / Strict Priority Queueing:
    - always the highest priority queue is served first (as long as it contains packets)
    - permanent preference of the highest priority queue
  - Weighted Fair Queueing:
    - Round Robin with weighting
    - queues with high weight are served more often



Components in the Control Path - Overview

- Resource control
  - monitoring of the resource occupancy (transmission capacity, buffer space)
  - control of resource usage: admission and usage control
    - admission control (before allocation and usage)
    - usage control (during usage)
  - resource allocation
    - deterministic reservation vs. statistical allocation
    - variants of deterministic reservations:
      - on-demand reservations after the successful negotiation of a traffic contract (remark: the negotiation of a traffic contract might be accelerated by suggesting acceptable parameter ranges or by announcing lists of acceptable parameter ranges; the network answers with the actually applied parameter values)
      - long-term reservations
      - preliminary reservations
- Control of tariffing / billing, user accounts

Current QoS Concepts in IP Networks

Challenges with Introduction of QoS in IP Networks

- The introduction of QoS mechanisms in IP networks is difficult and expensive
- Challenges:
  - replacement of routers → expensive
  - reservation of network resources → complex
  - differentiated charging according to QoS → unclear
  - scalability of the QoS architecture → unclear
  - protocol changes in all end systems → practically impossible
- But:
  - better QoS is required for new real-time critical applications
  - support of QoS creates new revenue opportunities for network operators

Basic Approaches for QoS Provisioning in IP Networks

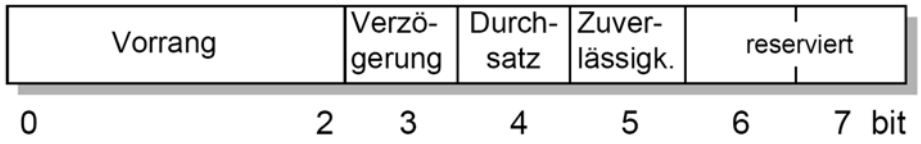
- IP Type of Service (ToS) marking (within IP packet header)
  - based on packet classification (via ToS) and differentiated packet handling
- Integrated Services concept (IntServ)
  - based on resource reservation and admission control
    - soft-state based reservation
    - receiver-oriented resource reservation signalling protocol RSVP
  - 3 services: Guaranteed Service, Controlled Load Service, Best-Effort
  - support of multicast
- Differentiated Services concept (DiffServ)
  - based on packet classification (via DSCP) and differentiated packet handling
    - support of max. 64 different service classes
    - service class marking in IP packet header via DiffServ Code Points (DSCP)
  - multicast is currently not considered
  - applicable today, but still open issues (e.g. DiffServ management architecture)
- MPLS with defined QoS for label-switched paths

ToS Marking - Old Definition

- Original QoS approach for IP (RFC 791)
- RFC 791:

  „The use of the Delay, Throughput and Reliability indications may increase the cost (in some sense) of the service. In many networks better performance for one of these parameters is coupled with worse performance on another.“

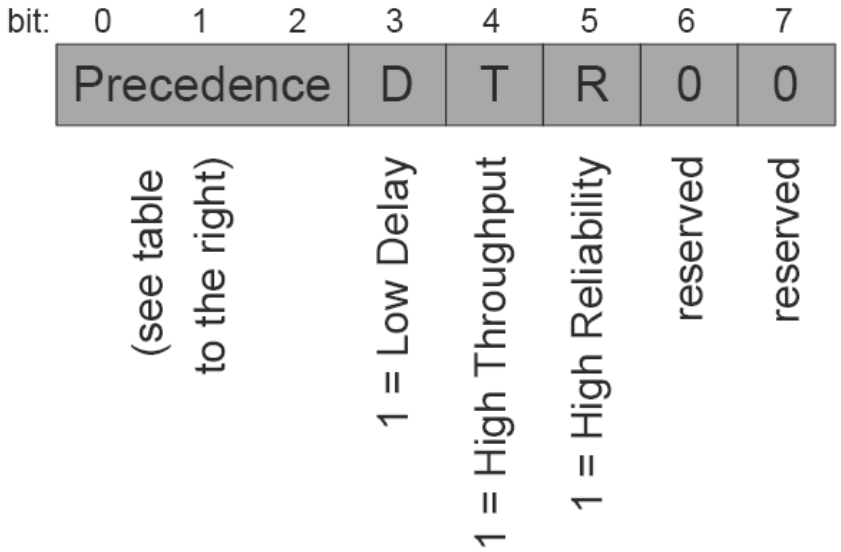
- Type of Service (ToS) field:
  - Precedence: different priority levels
  - Delay: normal (0), low (1)
  - Throughput: normal (0), high (1)
  - Reliability: normal (0), high (1)



- Problem: with the ToS field alone it is not possible to provide hard QoS guarantees

ToS Marking - Old vs. new Definition

- Old ToS definition (RFC 791):



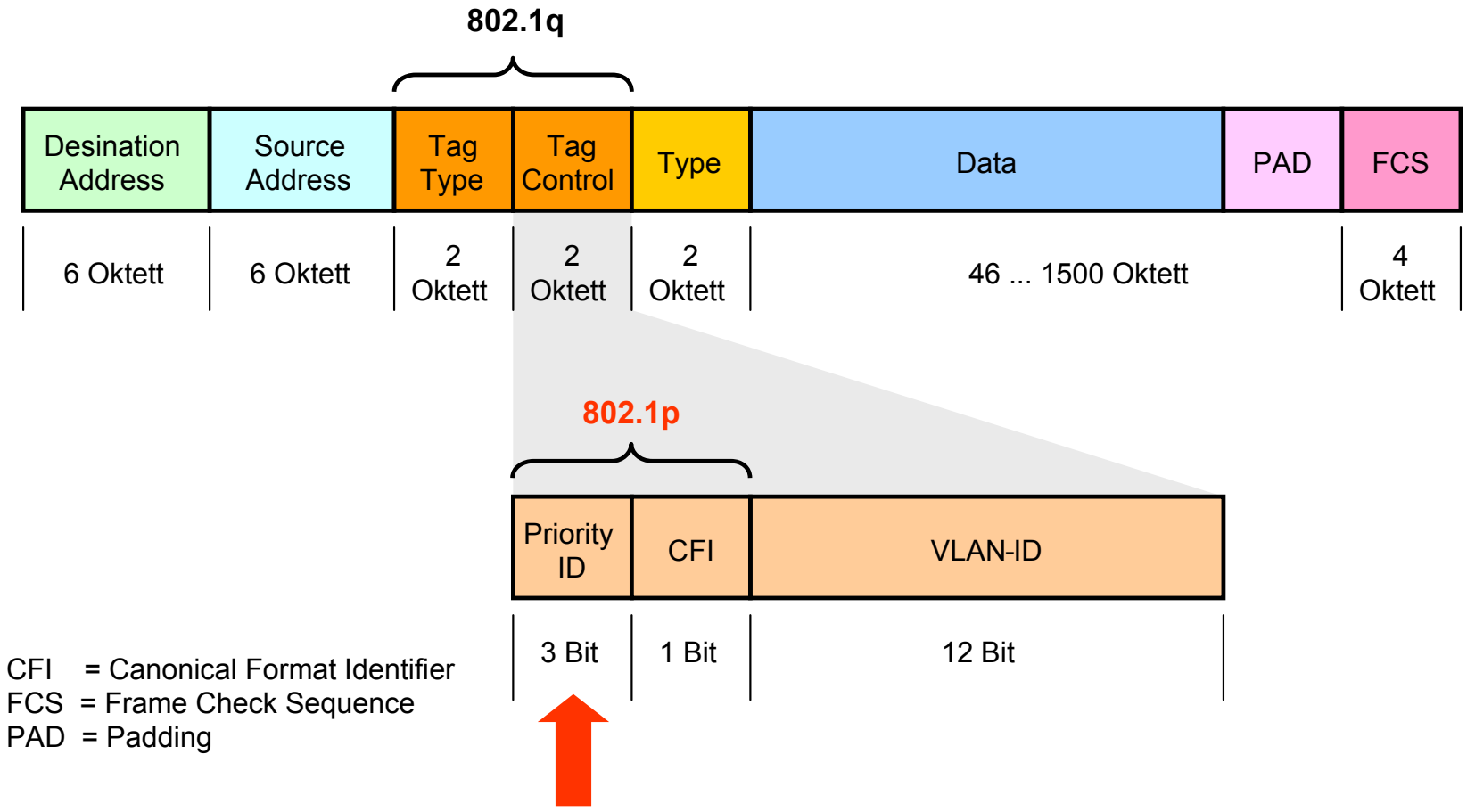
Precedence Values

111	Network Control (inside a network)
110	Internetwork Control
101	CRITIC / ECP
100	Flash Override
011	Flash
010	Immediate
001	Priority
000	Routine

- New ToS definition (according to DiffServ, RFCs 2474, 2475):
  - reallocation of the Precedence, Delay, Throughput and Reliability fields into a 6 bit DSCP (DiffServ Code Point) field
  - since 2001 bits 6 and 7 can be used (according to RFC 3168) for ECN (Explicit Congestion Notification = congestion signalling at IP layer)

ToS Marking - Remarks

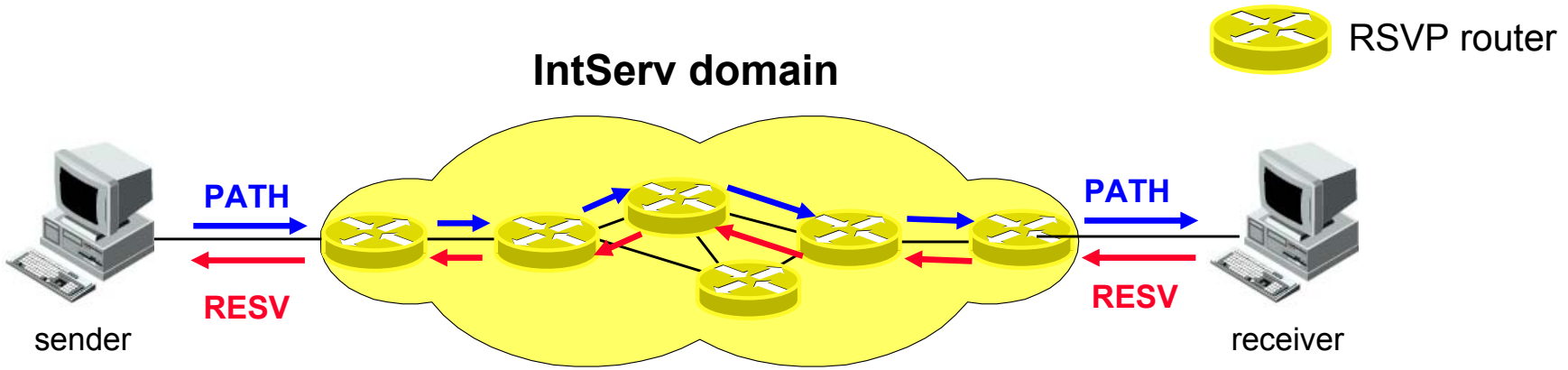
- With the modified frame format (IEEE 802.1q) also Ethernet allows for a QoS marking within the header field of the MAC frame: **3 bit Priority ID** (according to IEEE 802.1p)



IntServ Concept

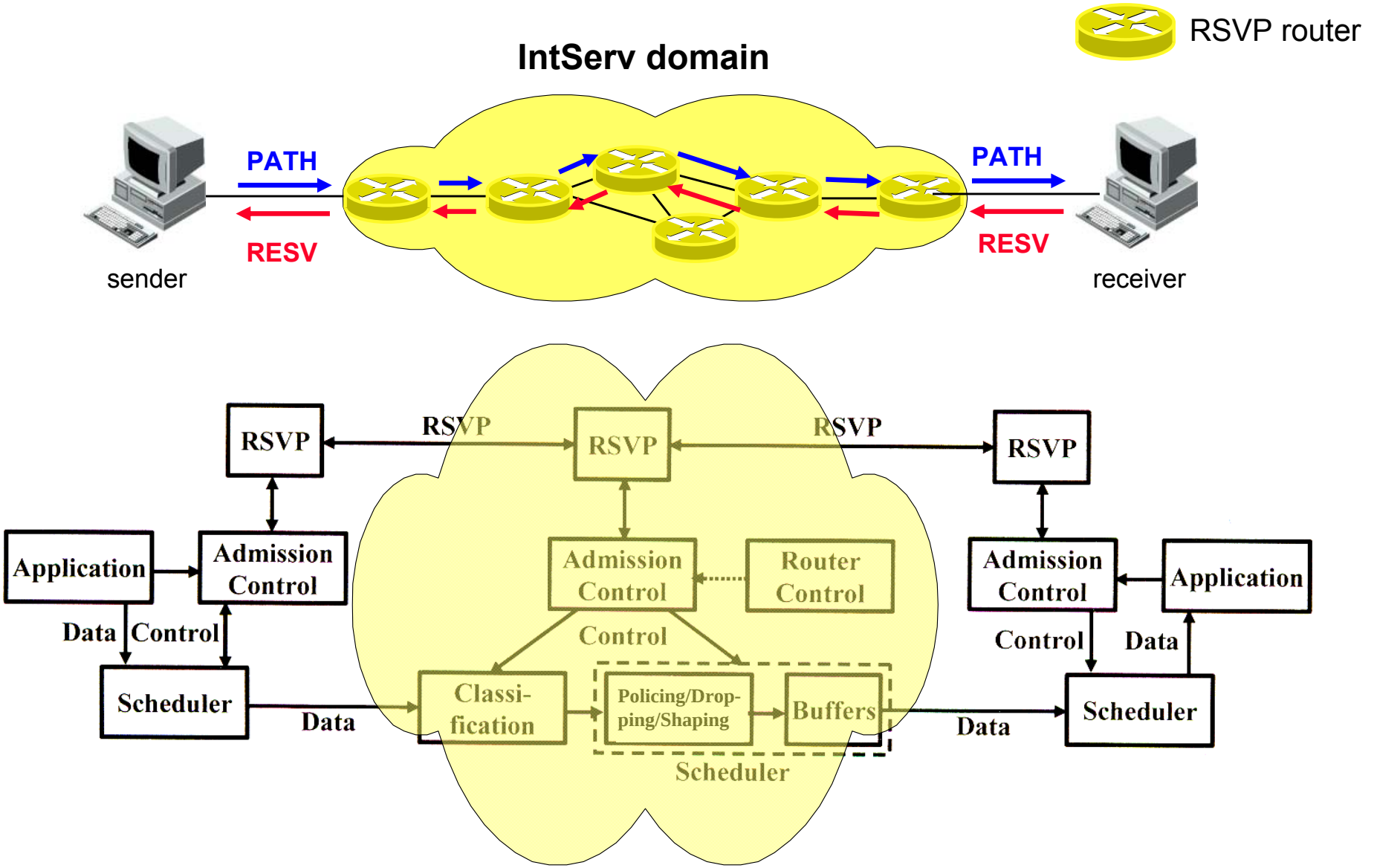
- The IntServ concept provides an architecture for integrated services in IP networks
  - support of multimedia applications (e.g. video conferences)
  - integration of group-communication applications
- Basic principle:
  - resource allocation (per traffic flow) using a special signalling protocol (RSVP)
    - signalling of traffic specification and requested QoS
  - routers retain a (soft) state per traffic flow
    - packet handling according to traffic specification and requested QoS
    - increased robustness due to soft states
  - admission control (per traffic flow)
- Problem: enhancement of the existing Internet architecture neccessary

IntServ Concept - IntServ Basic Principle

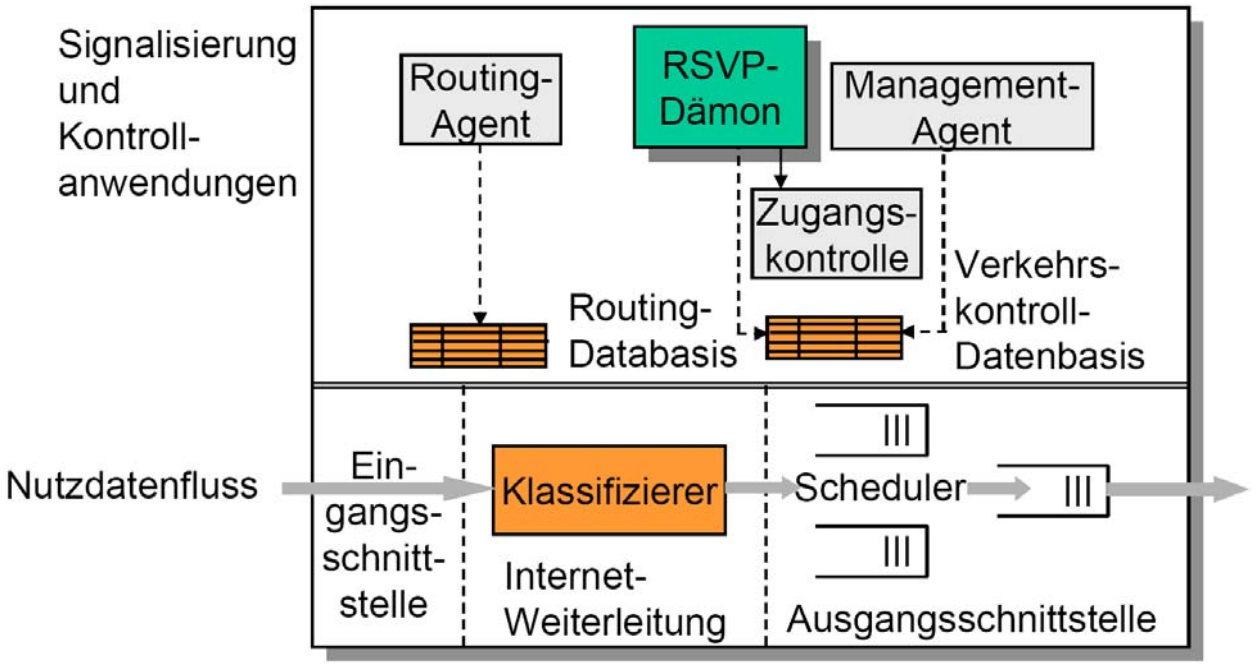


- Resource reservation per traffic flow/connection via Resource Reservation Protocol (RSVP):
  - PATH messages contain the traffic specification of the flow/connection
  - RESV messages invoke the reservation of capacity/buffer space along the data path
- Traffic flow/connection specific admission control and traffic parameter monitoring (policing) is performed on each link

IntServ Concept - IntServ System Components



IntServ Concept - IntServ System Components Details

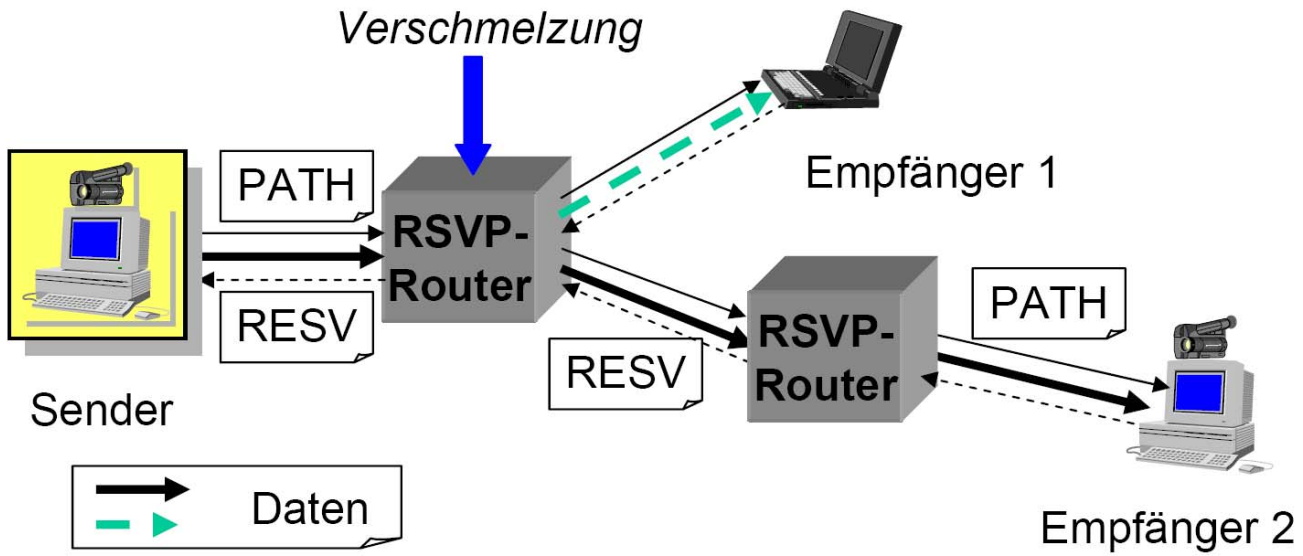


- Components in RSVP-capable systems (routers)
  - RSVP daemon: not involved in the user plane (user traffic), but only in the control plane (RSVP signalling)
  - Admission Control: checks, if the router can fulfill the resource requests
  - Classifier: marks the packets, which should benefit from the reservation
  - Scheduler: allocates the resources (e.g. the egress link capacity) and determines which packet should be send next

IntServ Concept - RSVP Signalling (RFC 2205)

- Purpose:
  - signalling of reservation requests in IP based networks
- Concept:
  - receiver-based signalling of reservation requests for an unidirectional traffic flow
  - support of heterogeneous quality of service
  - support of multicast communication
  - support of different reservation types
  - soft-state principle:
    - periodic renewal of states by end systems
    - the receiver obtains no positive acknowledgement of the reservation
    - no explicit teardown of reservations necessary: a timer is assigned to each reservation, upon expiry the reservation is deleted

IntServ Concept - RSVP Signalling (RFC 2205)



- Sender periodically send PATH messages and thereby establish a distribution tree; the PATH messages contain information about the traffic stream: QoS (QoS parameters) and traffic specification (traffic parameters)
- Receiver periodically send RESV messages for resource allocation in backward direction along the distribution tree (which has been set up via the PATH messages)
- RSVP routers merge reservation requests of different receivers into a new reservation request and forward it to the sender

IntServ Concept - RSVP Reservations

- RSVP reservation request (RESV message) = Flow Descriptor
  - **FlowSpec**
    - describes the requested QoS
    - **RSpec (Requirement Spec):** specification of the QoS (via QoS parameters) that is requested from the network
    - **TSpec (Traffic Spec):** specification of the traffic properties (via traffic parameters) of the traffic stream
  - **FilterSpec**
    - describes, which data units of a RSVP session are allowed to use the reservations
- Remark: definition of a RSVP session
  - a RSVP session is a set of traffic streams with equal destination (uni- or multicast)
  - a RSVP session is defined by the following triple: destination IP address, destination protocol ID, destination port (remark: in IPv4 no direct identification of traffic streams (e.g. via flow labels) is provided)

IntServ Concept - RSVP Message Format

- RSVP directly runs over IP (protocol ID 46) or optionally over UDP
- Message format:

4 bit Version	4 bit Flags	Message Type 8 bit	RSVP Checksum (16 bit)
Send TTL (8 bit)		Reserved 8 bit	RSVP Length (16 bit)
Objekte			

- Example of an object with RSVP FlowSpec:

Version 0	reserved		message lenght (7)
service header service 5 (controlled load	0	reserved	length of controlled load data (6 words)
parameter ID 127 (token bucked TSpec)	flags (none set)		length of parameter (5 words)
token bucket rate (r)			
token bucket size (b)			
peak data rate (p)			
minimum policed unit (m)			
maximum packet size (M)			

IntServ Concept - IntServ Service Types

- **Guaranteed Service** service type:
  - for traffic streams with hard real time requirements
    - hard packet delay and delay variation (jitter) thresholds
    - very low packet loss probability
    - guaranteed peak bit rate/throughput
- **Controlled Load Service** service type:
  - for traffic streams where statistical QoS guarantees are sufficient
    - no packet delay and delay variation thresholds
    - very low packet loss probability
    - guaranteed mean bit rate and burst tolerance
  - approximately corresponds to the QoS which is experienced by a traffic stream in an unloaded IP network without any QoS mechanism (in case the traffic stream is compliant to the traffic parameters)
- **Best Effort Service** service type:
  - for traffic streams with no QoS guarantees at all

IntServ Concept - Admission Control Rules

- Admission control rules (i.e. rules for the acceptance/rejection decision of a resource reservation request) in RSVP systems shall be aligned within the IntServ domain of a network operator
- For that so-called **Policy Server** are provided, which might be queried by admission control points
- As query protocol the so-called **Common Open Policy Services (COPS) Protocol** is used; the COPS data format is compatible to the RSVP data format
- Policy Server = **Policy Decision Point/Function (PDP/PDF)**
- Admission control point = **Policy Enforcement Point/Function (PEP/PEF)**

IntServ Concept - Advantages and Disadvantages

- Advantages:
  - absolute QoS guarantee
  - also suitable for multicast connections
- Disadvantages:
  - lack of scalability:
    - signalling via RSVP: routers have to cope with high signaling load
    - traffic stream/connection specific admission control and traffic parameter monitoring on each link
    - state information (traffic and QoS parameter, timer, sender and receiver addresses) are kept in routers → router performance decreases at high amount of reservations
    - packet forwarding is complex because of the classification of each single packet → challenging for high bitrates
  - due to the free choice of QoS parameters RSVP routers have to support a large variety of QoS (which also changes dynamically) → sophisticated packet scheduling

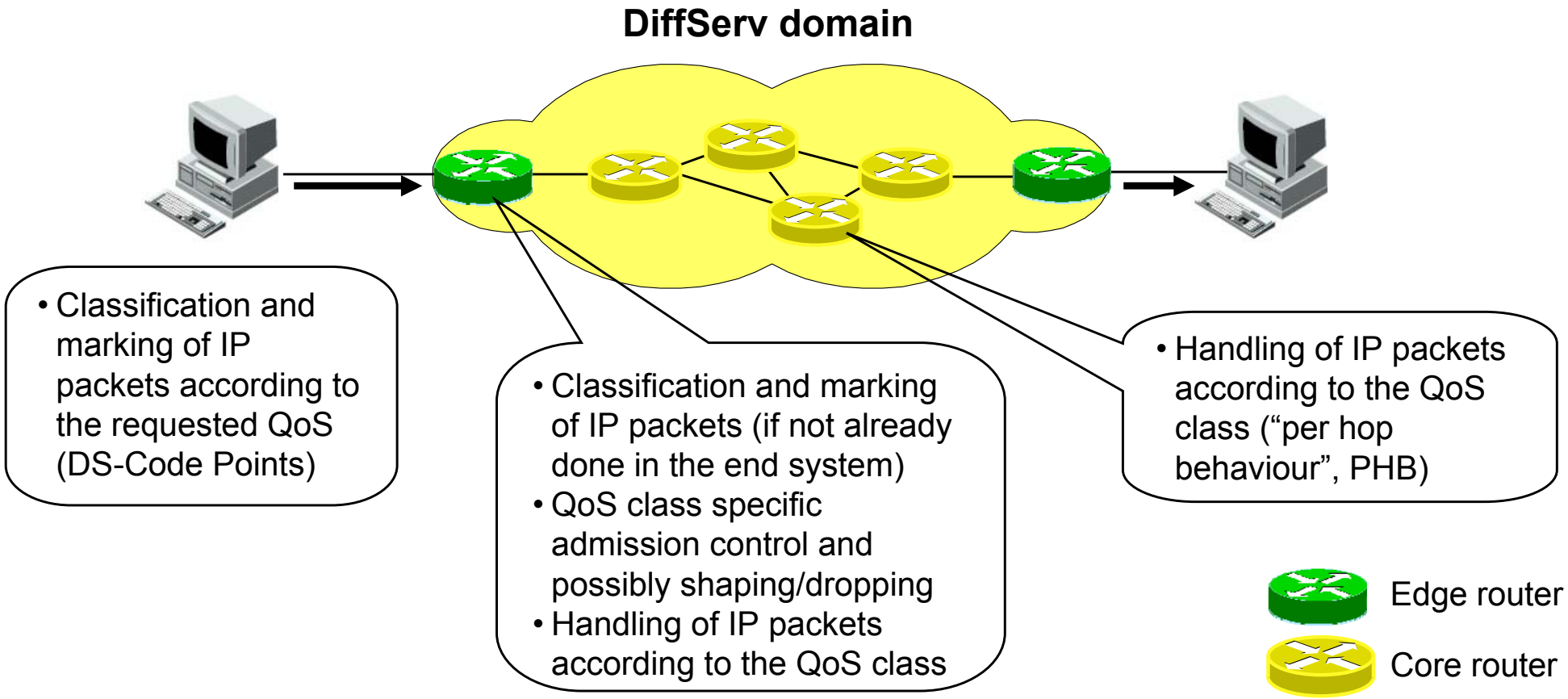
DiffServ Concept (RFC 2475)

- Objectives of DiffServ:
  - enabling QoS by simple, scalable mechanisms
  - reduction of (the network-internal) complexity: less states, lower functionality
  - compatibilty to existing applications and end systems (→ rapid implementation possible)
- DiffServ basic concept:
  - traffic classification and differentiated packet handling
  - traffic class specific limitation of ingress traffic at the network edge
- DiffServ history:
  - 1997: first proposals by David Clark and Van Jacobson in order to provide scalable quality of service in the Internet
    - original document “A Two-Bit Architecture” - see RFC 2638
  - beginning of 1998: work group “Differentiated Services” in IETF (until 2003)
    - objective: definition of base mechanisms only, but no services

DiffServ Concept (RFC 2475)

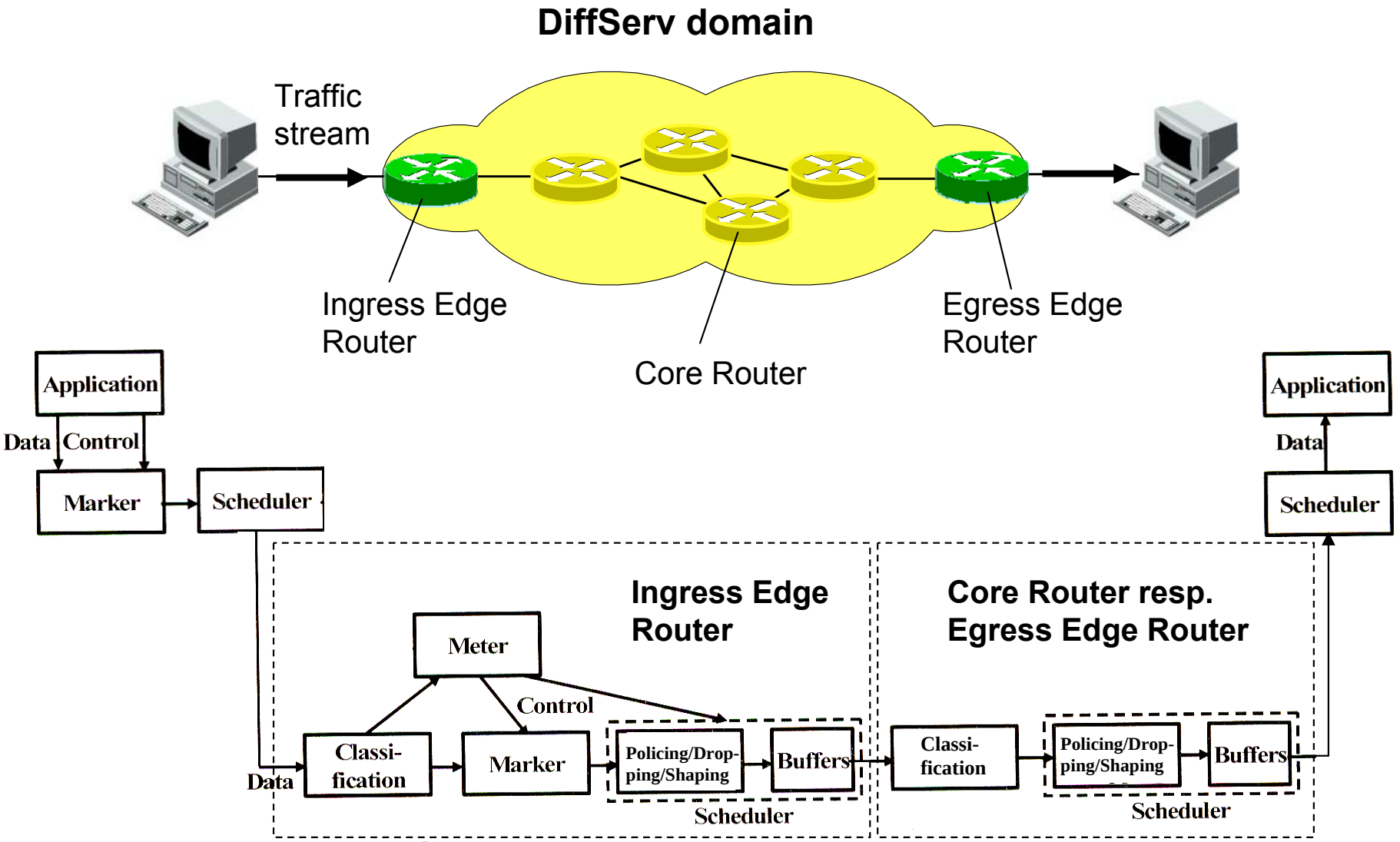
- Complexity reduction by:
  - classification of traffic (into few classes) within the network: packets belonging to the same service class carry the same marking
  - prevention of end-to-end traffic stream (micro-flow) specific or user specific states within the network
  - more complex functions like classification, marking and monitoring takes place only at the network edge
- Marking of the class specific per-hop-behaviour (PHB) in the IP packet header (via DiffServ Codepoints, DSCP)
- Separation of QoS specific forwarding mechanisms and QoS management mechanisms (e.g. configuration, resource management and allocation)
- Improvement of the ToS approach:
  - explicit definition of the edge router functions and traffic influencing mechanisms
  - the PHB model is more flexible than relative priorities or QoS markings

DiffServ Concept - DiffServ Basic Principle



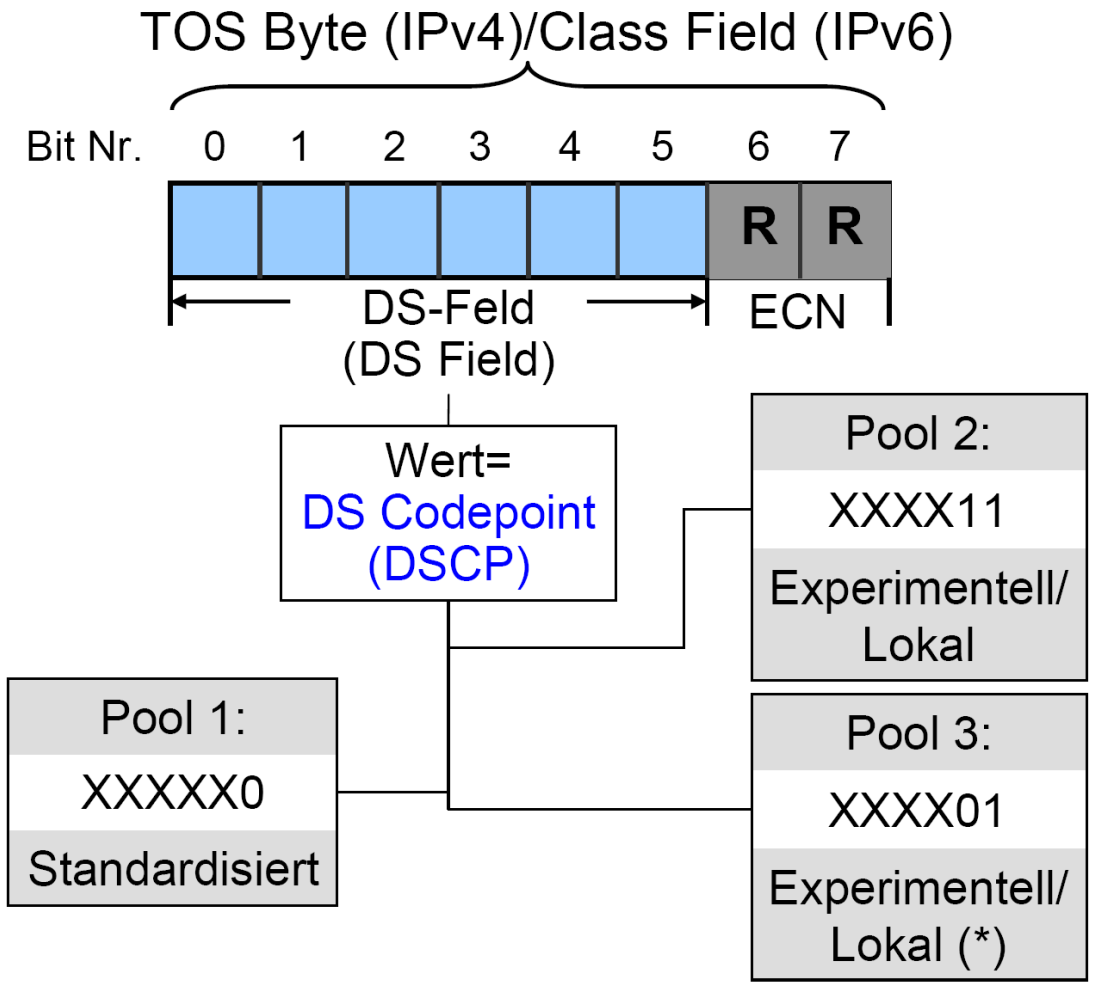
- Definition of QoS classes: marking of IP packets wrt. their QoS class
- QoS class specific handling of IP packets in routers
- QoS class specific ingress traffic load limitation at the network edge

DiffServ Concept - DiffServ System Components



DiffServ Concept - DiffServ Codepoints and PHBs

- Basically the DS field (in the IPv4 and IPv6 packet header) is unstructured
- By the DiffServ Codepoint (DSCP) a corresponding PHB is uniquely assigned
- Multiple DSCPs may refer to the same PHB
- There can be more PHBs than DSCPs
- The DSCP normally has no influence on the route selection



(*): Bereich kann bei Bedarf zur Standardisierung weiterer Werte herangezogen werden

DiffServ Concept - Standardized PHBs

- Default PHB (RFC 2474)
  - DSCP = 000000
  - complies with the current best effort behaviour
- Class Selector (CS) PHBs (RFC 2474)
  - DSCP = XXX000
  - compatible to the IP precedence field in the ToS concept
  - a higher value means higher priority → 8 relative priority classes (incl. the Default PHB)
- Expedited Forwarding (EF) PHB (RFC 3246)
  - DSCP = 101110
- Assured Forwarding (AF) PHB (RFC 2597)
  - 12 DSCPs (4 AF classes with 3 dropping priorities each)
- Lower Effort (LE) PHB (RFC 3662)
  - DSCP = 001000

DiffServ Concept - Expedited Forwarding PHB

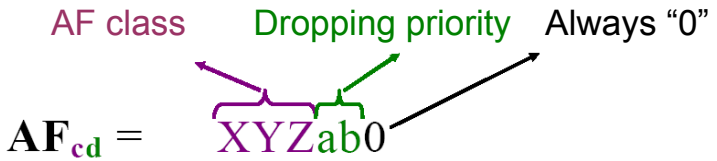
- Expedited Forwarding PHB is suitable for services with:
  - guaranteed bitrate, low and limited jitter, low delay, low packet loss
- Packet handling for Expedited Forwarding PHB:
  - arriving packets “see” only an empty queue, hence packets are transmitted without significant waiting time (delay)
  - implementation e.g. via Simple Priority Queueing or WFQ
- Prerequisite:
  - the overall bitrate of Expedited Forwarding traffic at an egress interface (of a router) must be smaller than the egress rate of the interface (otherwise overload within same class and packet loss occurs) → admission control necessary
- The original specification of the Expedited Forwarding PHB in RFC 2598 is imprecise; therefore a new definition is provided in RFC 3246 (and additional comments in RFC 3247)

DiffServ Concept - Assured Forwarding PHB

- Assured Forwarding PHB is suitable for services with:
  - guaranteed bitrate below a agreed transmission rate
    - the use of more capacity (above the agreed rate) is possible, if currently available → statistical multiplexing of Assured Forwarding PHB traffic streams
    - in case of congestion packets (above the requested transmission rate) are dropped
  - bursty traffic characteristics
    - longer burst have a higher dropping probability
- Packet handling for Assured Forwarding PHB:
  - via active queue management mechanims (e.g. Random Early Discard, RED)

DiffServ Concept - Assured Forwarding PHB

- Assured Forwarding PHB Group (RFC 2597):
  - definition of the attributes for Assured Forwarding PHBs
  - definition of so-called AF classes
- AF class:
  - an AF class is a group of m PHBs, each with a different dropping probability (e.g. m=3, meaning dropping priorities low, medium, high)
  - an AF class is completely independent of other AF classes
- Proposal: 4 independent AF classes with standardised DSCPs



Different dropping priorities (realized via Active Queue Management mechanisms, e.g. RED, WRED)

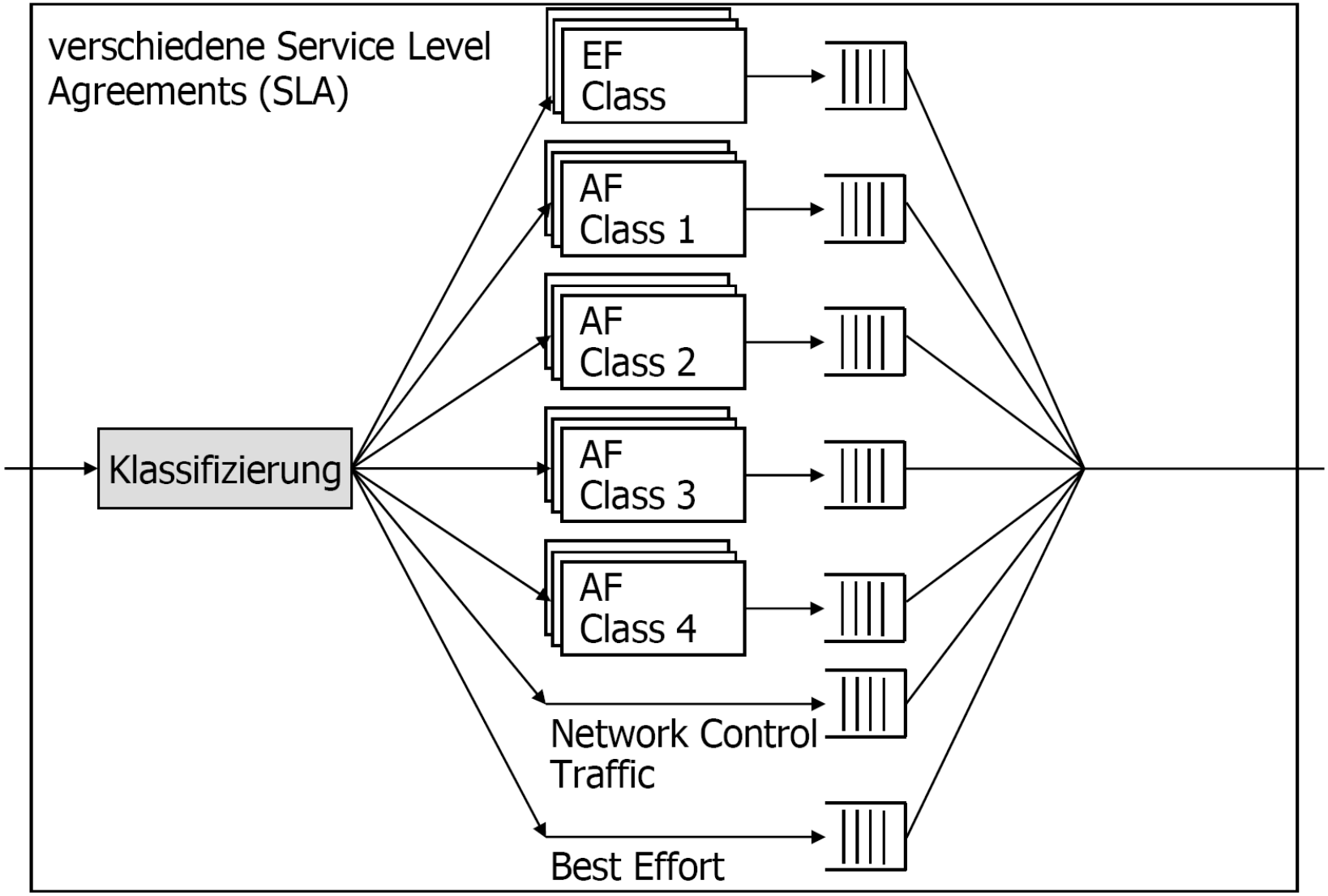
Verwurfspriorität (y)	AF 1y	AF 2y	AF 3y	AF 4y
Niedrig (1) ●	001010	010010	011010	100010
Mittel (2) ●	001100	010100	011100	100100
Hoch (3) ●	001110	010110	011110	100110

AF-Klasse

DiffServ Concept - Overview of Standardized PHB

	DSCP	PHB
1	**000 000**	**Default PHB / CS0**
2	**001 000**	**LE / CS1**
3	001 010	AF11
4	001 100	AF12
5	001 110	AF13
6	010 000	CS2
7	010 010	AF21
8	010 100	AF22
9	010 110	AF23
10	011 000	CS3
11	011 010	AF31
12	011 100	AF32
13	011 110	AF33
14	100 000	CS4
15	100 010	AF41
16	100 100	AF42
17	100 110	AF43
18	101 000	CS5
19	**101 110**	**EF**
20	110 000	CS6
21	111 000	CS7

DiffServ Concept - Realization of PHBs



DiffServ Concept - Resource Management (1)

- Resource consumption
  - each DiffServ marked packet occupies resources of a node which are assigned to its respective PHB
  - monitoring of individual end-to-end traffic streams is performed (if at all) only within the first edge node (ingress edge router)
    - the sender (end system) or the first edge node decides on the packet marking (and by that on the further resource consumption)
    - the first edge node monitors (by means of the traffic profile) the resource consumption of the traffic stream
    - network internal nodes normally have no knowledge about traffic profiles, thus the resource consumption of individual traffic streams can not be monitored there
  - at subsequent edge nodes in upstream direction (multi domain case) only the behaviour of traffic class aggregates can be monitored, but not the behaviour of individual traffic streams (micro flows)

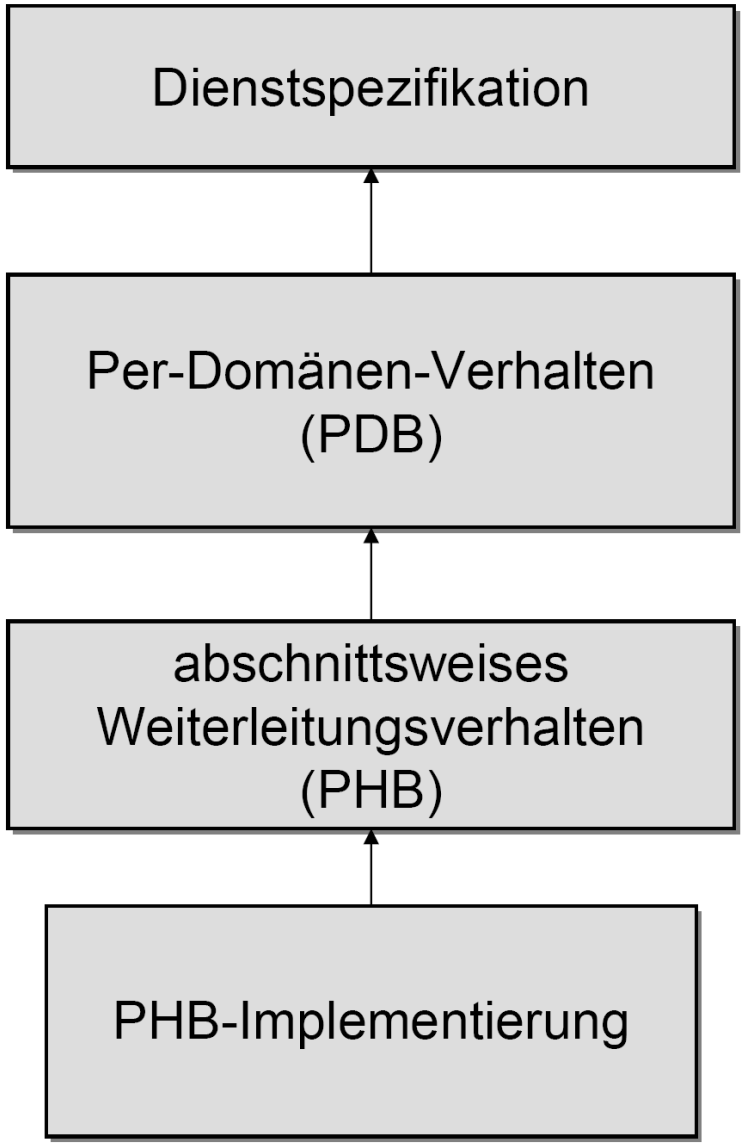
DiffServ Concept - Resource Management (2)

- Resource allocation
  - usually performed per traffic class (aggregate)
  - mostly static allocation via manual configuration by the network operator
    - by that an early adoption of DiffServ is possible
    - later: dynamic resource allocation (on demand) - this however requires signalling
  - static resource partitioning (wrt. different PHBs) within the nodes
  - nodes do not have to monitor the assigned/available resources
  - the release of network resources is performed indirectly by storing the related traffic profile in the first edge node (the profile might be downloaded from a so-called policy server)

DiffServ Concept - DiffServ Services

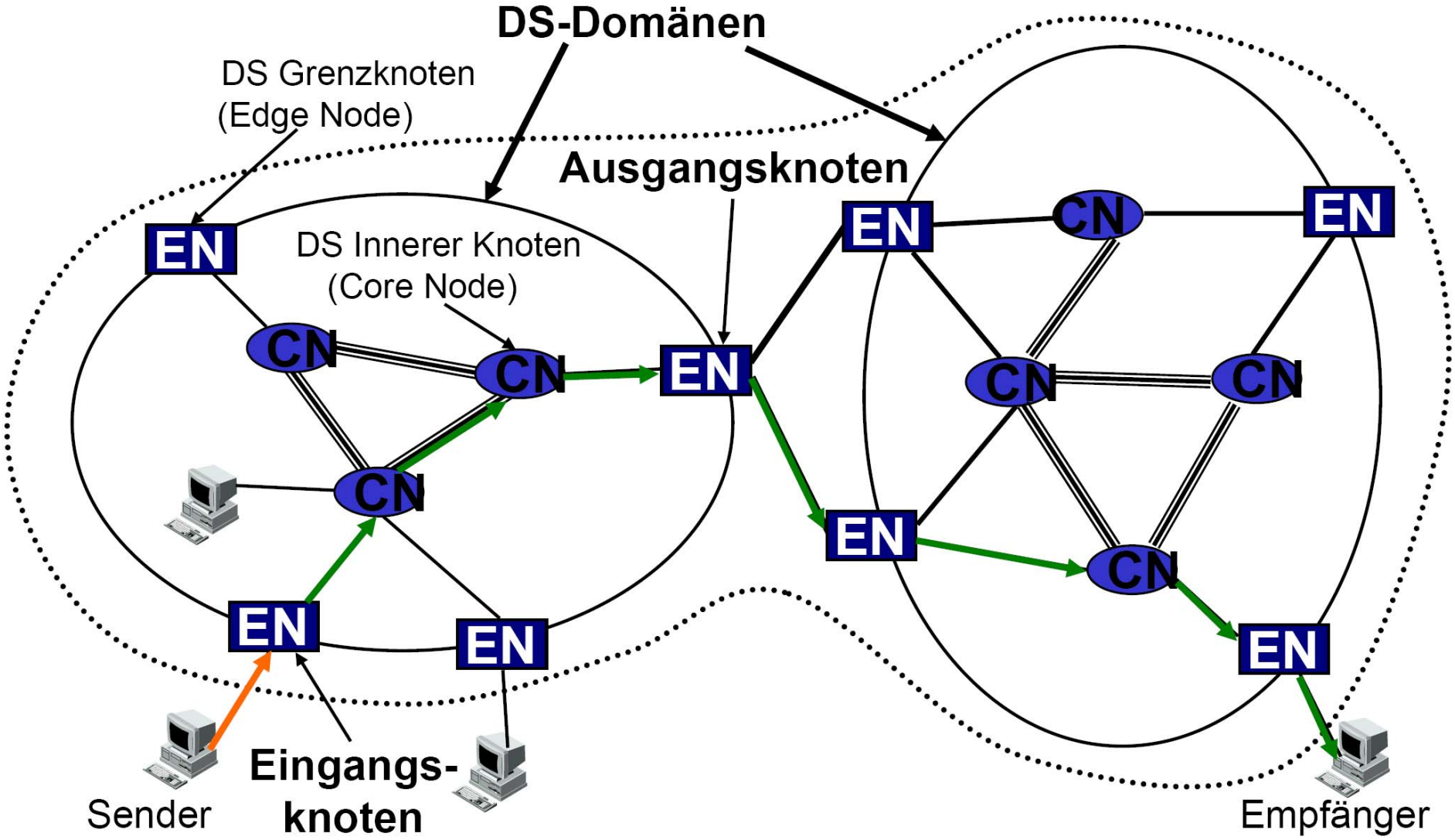
- A DiffServ service is characterized by:
  - definition of the per-hop forwarding behaviour (PHB)
  - definition of the domain-specific forwarding behaviour (per-domain behaviour, PDH) via concatenation of PHBs; the PDH is created by queue management and scheduling mechanisms together with traffic manipulation rules (classification, meter, former) within the domain
  - definition of the end-to-end service via concatenation of PDBs (across multiple domains); hierarchy: PHB → PDB → service
- Service Level Agreement (SLA):
  - service contract between customer and operator to agree on a specific forwarding service
- Building blocks required to realize a DiffServ service:
  - PHB implementation (concrete realization of a PHB)
  - traffic manipulation mechanisms
  - (charging model)

DiffServ Concept - DiffServ Services



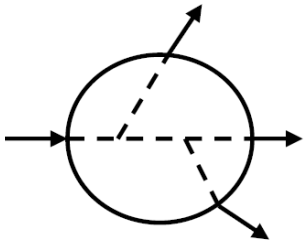
- Specification of a end-to-end service
  - via Service Level Agreement (SLA)
  - composed of PDBs of the traversed domains
- Domain specific behaviour
  - PHB specification + traffic manipulating mechanisms (Traffic Conditioning)
- Per hop (packet forwarding) behaviour
  - packet processing within one node (based on the DSCP)
- PHB implementation
  - realized via queueing and scheduling mechanisms

DiffServ Concept - Multiple DiffServ Domains



DiffServ Concept - Per Domain Behaviour (PDB, RFC 3086)

- A Per-Domain-Behaviour describes the behaviour, which packets experience on their way through a DiffServ domain:
  - arbitrary ingress to arbitrary egress (One-to-One)
  - arbitrary ingress to any egress (One-to-Any)
  - arbitrary ingress to certain egresses (One-to-Few)



- Rules for traffic manipulation (traffic conditioning) and for applying appropriate PHBs have to be specified
- Traffic aggregation effects have to be considered
- Attributes which characterize the PDB must be measurable and quantifiable

DiffServ Concept - Service Level Agreements (SLAs)

- Service Level Agreements (SLAs) are bilateral service contracts with static or dynamic service specification between neighboring ISPs
- Technical service specification:
  - performance parameters, e.g.:
    - throughput
    - packet delay
    - packet loss probability
  - range of validity
  - traffic profile (traffic parameters, e.g. bitrate and burstiness)
  - marking and traffic forming rules
  - additional general parameters, such as:
    - availability
    - reliability
    - route restrictions
- SLAs may also contain non-technical aspects, e.g. business and economical aspects (e.g. conditions for regress payments)

DiffServ Concept - SLA Example

Availability, Reliability, Error Handling
Security, Encryption
Routing (Constraints)
Authentication
Traffic Measurement
Responsible Person (e.g. E-Mail)
Charging
Traffic Conditioning Specification (TCS)
Performance, e.g.: - Throughput, Bit Rate - Packet Loss, Drop Probability - Packet Delay, Latency
Traffic Profile, e.g.: - Token Bucket Parameter
Handling of Excess Traffic, e.g.: - Marking - Shaping

DiffServ Concept - Advantages and Disadvantages

- Advantages:
  - good scalability:
    - no QoS signalling required - only marking
    - no traffic-stream/connection specific traffic specification and admission control - only overall load limitation for ingress traffic per QoS class at the network edge
    - no storage of states within routers necessary
- Disadvantages:
  - no absolute but only relative QoS guarantee - a risk of network internal overload persists

    possible solution: combination of DiffServ with traffic-stream/connection specific admission control via central resource manager which knows the load situation within the network

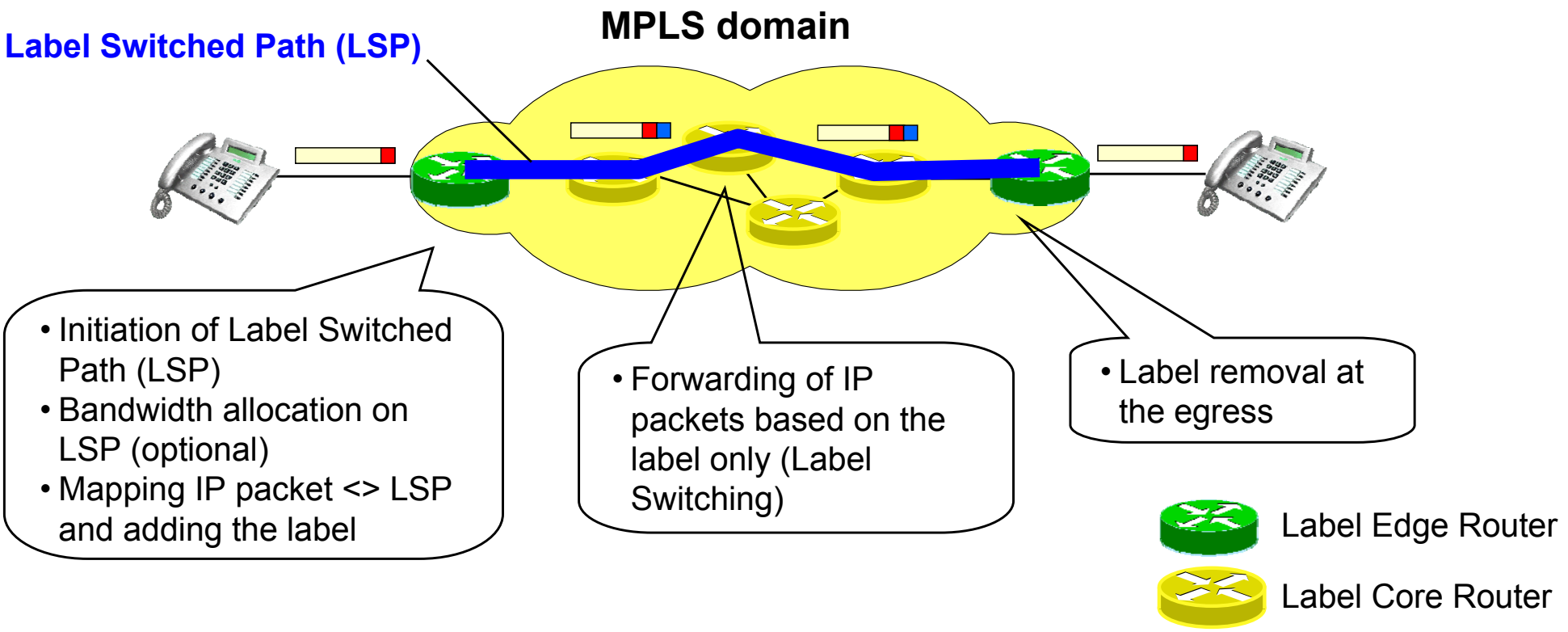
DiffServ vs. IntServ Concept

	Best-Effort	IntServ	DiffServ
QoS guarantee	none	per traffic-stream (micro-flow)	per aggregated traffic (traffic class)
configuration	none	per traffic-stream	per aggregated traffic between domains
type of guarantee	none	per traffic-stream	per aggregated traffic (traffic class)
duration of guarantee	none	short-lived (duration of traffic stream)	long(er)-term
states	none	per traffic-stream	per aggregated traffic (traffic class)
signalling	none	RSVP protocol	none (not yet defined)
multicast support	IP multicast	receiver-oriented, heterogeneous	IP multicast (but problems with resource allocation)

MPLS Concept

- MPLS enables virtual path connections (label switched paths, LSPs) for individual connections/traffic-streams or for aggregated traffic
- For QoS support a combination of MPLS with DiffServ or IntServ is possible:
  - combination with IntServ: explicit resource allocation along the LSP via the signalling protocol RSVP-TE (used for LSP setup and resource allocation); the QoS for the LSP / RSVP session is associated with the MPLS label
  - combination with DiffServ: no resource allocation for LSPs; two possible methods for mapping of DiffServ classes and MPLS:
    - DiffServ classes (incl. dropping priority) are derived from the first 3 bits of the EXP field in the MPLS header: **EXP-inferred LSP (E-LSP)**
    - Diffserv classes are associated with the MPLS label (and the dropping priority is derived from the EXP field): **Label-inferred LSP (L-LSP)**

MPLS Concept - Basic Principle



- MPLS enables virtual path connections (LSPs) with/without capacity allocation for individual connections/traffic-streams or for aggregated traffic
- LSPs can be routed independently from the IP routing protocol (explicit routing - can be applied for traffic engineering)

MPLS Concept - Advantages and Disadvantages

- Advantages:
  - explicit routing of LSPs (independently from IP routing) → can be applied for traffic enginnering
  - QoS support in combination with IntServ or DiffServ possible
  - Label switching allows faster forwarding than routing (hence leading to lower packet delay)
- Disadvantages:
  - higher effort compared to pure DiffServ:
    - signalling for setup of LSPs
    - state information per LSP within label switsched routers (LSRs)
    - admission control per LSP
  - no end-to-end QoS solution because MPLS is not intended for end systems; the ingress label edge router (LER) has to map the QoS of the non-MPLS domain to MPLS QoS
  - guaranteed QoS for connections/traffic-streams within a LSP is only achievable by use of admission control on the LSP

QoS Concepts for IP Networks - Summary

- Differentiated Services (DiffServ) concept:
  - definition of QoS classes; marking of IP packets according to their QoS class
  - QoS class-specific queue management within routers
  - coarse-grained admission control: limitation of the total incoming traffic per QoS class at the network edge
- Integrated Services (IntServ) concept:
  - capacity allocation per traffic-stream/connection via RSVP protocol
  - traffic-stream/connection individual admission control at each link
- Multi Protocol Label Switching (MPLS) concept:
  - virtual path connections (Label Switched Paths, LSPs) with or without assigned capacity for individual traffic-streams/connections or aggregated traffic
  - QoS support possible by combination with IntServ or DiffServ

QoS Concepts for IP Networks - Remarks

- The deployment of Quality of Service within IP networks a quite complex matter
- Challenges:
  - end-to-end QoS guarantee:
    - end-to-end within a domain: end-to-end guarantee of certain QoS parameters such as delay, jitter is difficult
    - end-to-end beyond multiple domains / operator AS: even more difficult because different QoS mechanisms might be used in the domains → goal: uniform QoS interworking and QoS signalling: **IETF NSIS WG**
  - layering of multiple technologies e.g. IP over Ethernet or IP over ATM:
    - all layer below IP have to be considered
    - no problem, in case of IP over “circuit-switched” networks (like TDM-/SDH-/WDM-based networks) with fixed transmission capacities
    - problem, in case of IP transport over packet/frame-switched networks that allow statistical multiplexing (e.g. Frame Relay, ATM, Ethernet networks); here an alignment of the IP QoS mechanisms and the QoS mechanisms used in the lower layer network is necessary (Inter-Layer QoS problem)
  - application of traffic engineering (TE) and/or resilience mechanisms:
    - QoS has to be maintained during TE and restoration

QoS Concepts for IP Networks - Practical Issues

- Several network operators state that they don‘t need QoS support:
  - a well dimensioned network (overprovisioning) guarantees a low load and minimal delays
  - thus packet transmission is possible without high packet loss and with delay guarantee
- Today's routers are equipped with DiffServ mechanisms; therefore some big network operators already apply DiffServ
- The application of DiffServ mechanisms at the network edge (i.e. within access networks) may lead to cost savings:
  - separation of data traffic and real-time critical voice traffic (realized e.g. via using the EF class for voice traffic)
  - the cost savings result from the integration of different services with different QoS requiements (e.g. voice and data) on the same (router) platform (no separate networks required)

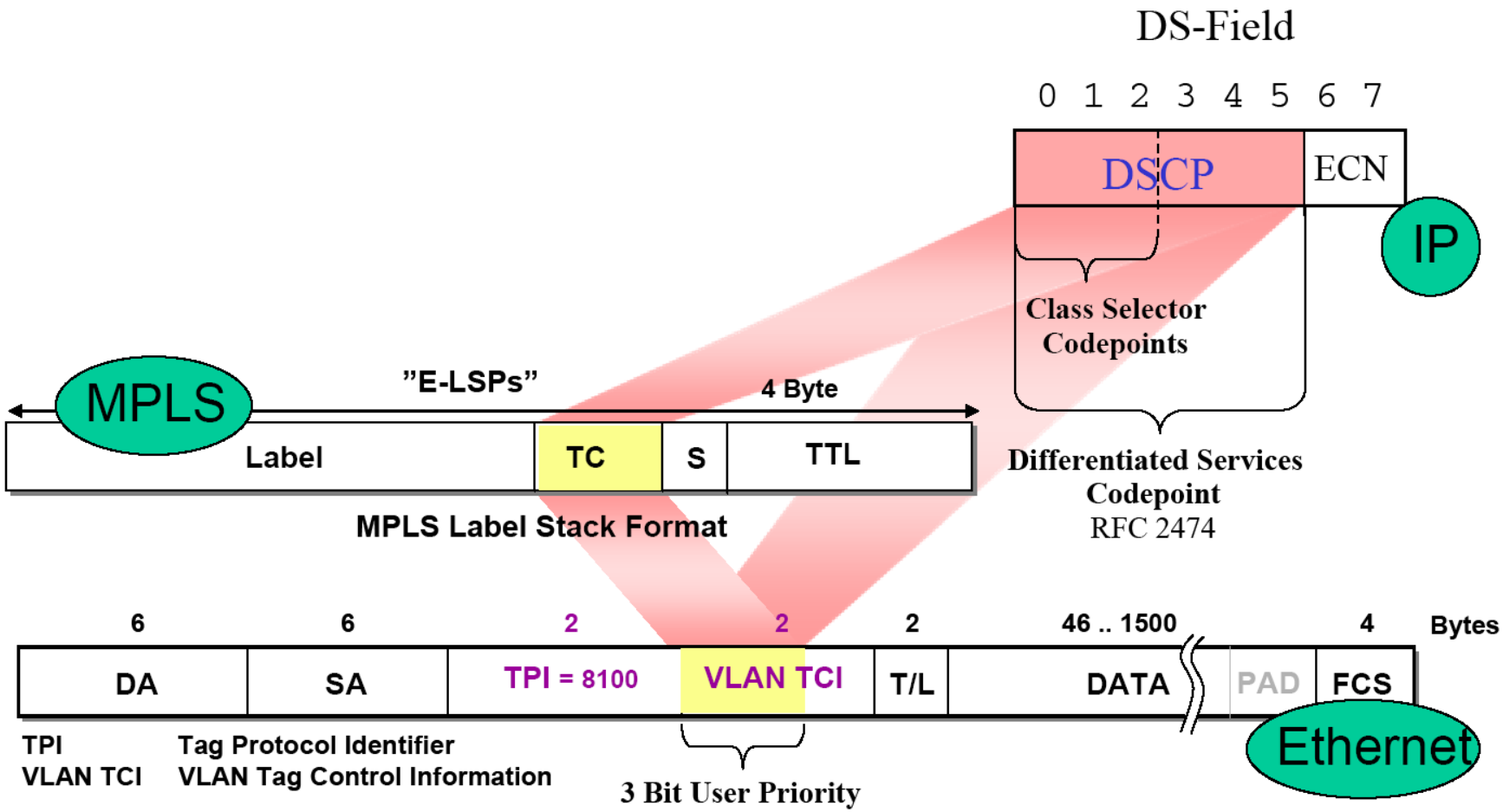
Further QoS Aspects

Cross-Layer Quality of Service

- An alignment of QoS concepts between network layers is necessary if IP traffic is transported over (layer 2) packet-based networks with own QoS mechanisms
- Examples:
  - IP over Frame Relay (FR)
    - QoS support at FR: rate agreement (CIR, EBS), Discard Eligibility (DE) bit
  - IP over ATM
    - QoS support at ATM: traffic contract, Cell Loss Priority bit
  - IP over Ethernet-VLAN
    - QoS support at Ethernet: 3 bit Priority ID (Priority Code Point, PCP)
  - IP over MPLS
    - QoS support at MPLS: 3 bit Traffic Class Field
- In general cross-layer QoS alignment is not standardized

Cross-Layer Quality of Service

- Example: cross-layer QoS mapping IP ↔ MPLS and IP ↔ Ethernet



Cross-Domain Quality of Service

- An alignment of QoS concepts between domains (Autonomous Systems, AS) is necessary for ensuring IP traffic QoS across multiple domains
- Problems:
  - each network operator (ISP) may apply a specific QoS concept within its AS (e.g. DiffServ with n classes) - there is no obligation to inform other AS about this
  - for cross-domain QoS alignment no standardized solution exists
- Possible solutions:
  - set-up of Inter-AS MPLS paths (via RSVP-TE)
  - Inter-AS signalling of QoS class sets (e.g. via BGP extensions)

QoS Optimization during Network Operation

- Classification criteria:
  - optimization variables (“what is changed”), constraints and objective
  - dynamic behavior (“how fast are the changes") and time scale
  - offline (proactive calculation, static) vs. online (reactive calculation and execution, dynamic)
- Variants of network/QoS optimization:
  - **Traffic Engineering:**
    - definition: "put the traffic where the bandwidth/capacity is"
    - characteristics: online, dynamic (ms range), objective: good network utilization
  - **Network Engineering (≈ expansion planning):**
    - definition: "put the bandwidth/capacity where the traffic is"
    - characteristics: online or offline, dynamic or static (day-month range), optionally also routing changes, objective: bottleneck prevention
  - **Network Planning:**
    - definition: "put the bandwidth/capacity where the traffic is forcasted to be"
    - characteristics: offline, static (1-20 years range, according to the planning horizon), objective: cost minimization

Traffic Engineering as Optimization Problem

- Problem:
  - optimization of traffic-streams wrt. the optimization objective and the QoS constraints (while keeping the network capacity unchanged)
- Possible optimization objectives:
  - best possible network utilization (maximization of the network throughput)
  - best possible service quality (better than required)
  - maximization of revenues from network operation
- Optimization variable:
  - routing of traffic-streams

    issues:
    - definition of routing paths and number of routes (route set)
    - load distribution on the routes
- Constraints:
  - network capacity
  - QoS guarantees

Traffic Engineering (TE) Variants in IP Networks* (1)

*here considered only: intradomain case

- Manipulation of routing paths:
  - variation of link-cost-metrics (in case of traditional IP routing)
    - calculation of optimum link-metrics (centralized calaculation e.g. in NMS)
  - application of modified routing protocols that allow the consideration of QoS constraints: constraint-based routing
    - examples: OSPF-TE or ISIS-TE; problem: overall optimum solution for the whole network?
  - explicit routing:
    - IP Source Routing options (often not supported by routers):
      - strict Source Routing = IP Option 9 (RFC 791); definition of fixed routing paths by determination of up to 9 consecutive routers that have to be passed
      - loose Source Routing = IP Option 3 (RFC 791); determination of up to 9 routers which have to be along on the routing path
    - MPLS:
      - centralized calculation of optimum LSPs (in NMS) → Path Computation Element (PCE) concept (RFC 4655)
      - or: application of CR-LDP or RSVP-TE (consideration of QoS constraints in explicit routing); issue: overall optimum solution for the whole network?

Traffic Engineering (TE) Variants in IP Networks* (2)

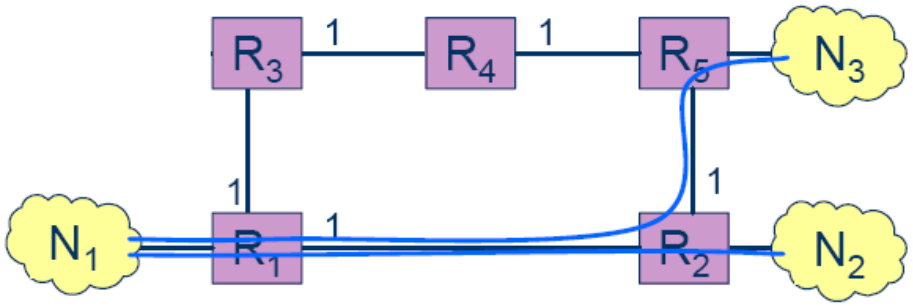
*here only considered: intradomain case

- Manipulation of routing granularity:
  - traditional IP routing: routing based on destination IP address (destination prefix)
  - generation of individual routing paths considering further criteria (DSCP code points, labels, port no. etc.) → finer granularity
    - example: individual MPLS LSPs per FEC class (hereby QoS specific routes can be realized)
- Multipath routing with (dynamic) traffic distribution:
  - up to now: multipath IP routing (to the same destination) is only possible with equal cost paths and equal traffic distribution
    - example: **Equal Cost Multi-Path (ECMP) Routing**
  - challenge: dynamic, network load dependent, traffic-stream specific traffic distribution on multiple (possibly not cost-equal) routing paths; (remark: such a routing scheme is already implemented in telephone networks, see e.g. schemes like ODR, ADR, DNHR, DCR, etc.)

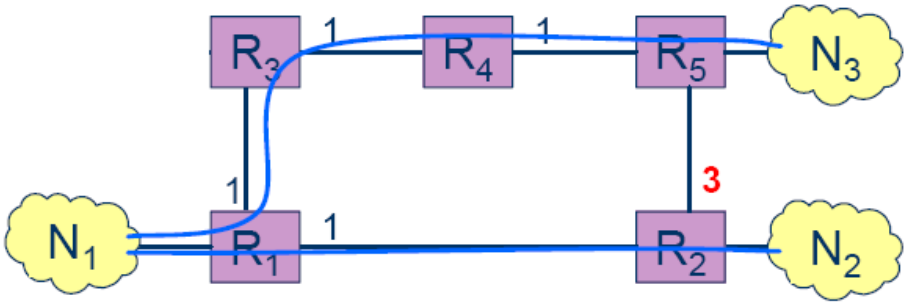
Example: TE by changing the Routing Metrics

- Example:

Routing based on hop-count metrics:



Routing based on modified metrics:

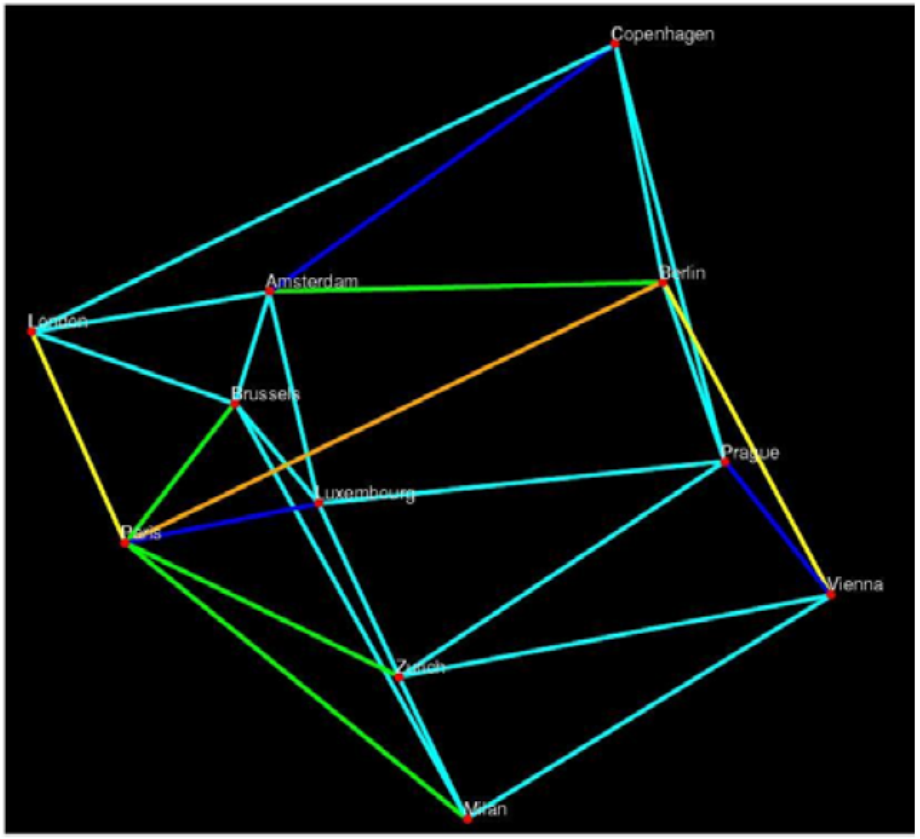


- Disadvantages of this method compared to MPLS-TE:
  - not every traffic distribution can be realized (underdetermined problem: less variables (link metrics) than routes)
  - nonlinear relation between link metrics and link traffic loads: nonlinear optimization problem (hard to solve) – solution e.g. via heuristic or hybrid optimization methods, e.g. greedy-search, genetic algorithms, simulated annealing

Example: TE by changing the Routing Metrics

Hop-count based routing

- mean link load = 24.38%
- max. link load = 71.35%



Routing with optimized metrics

- mean link load = 24.69%
- max. link load = 38.75%