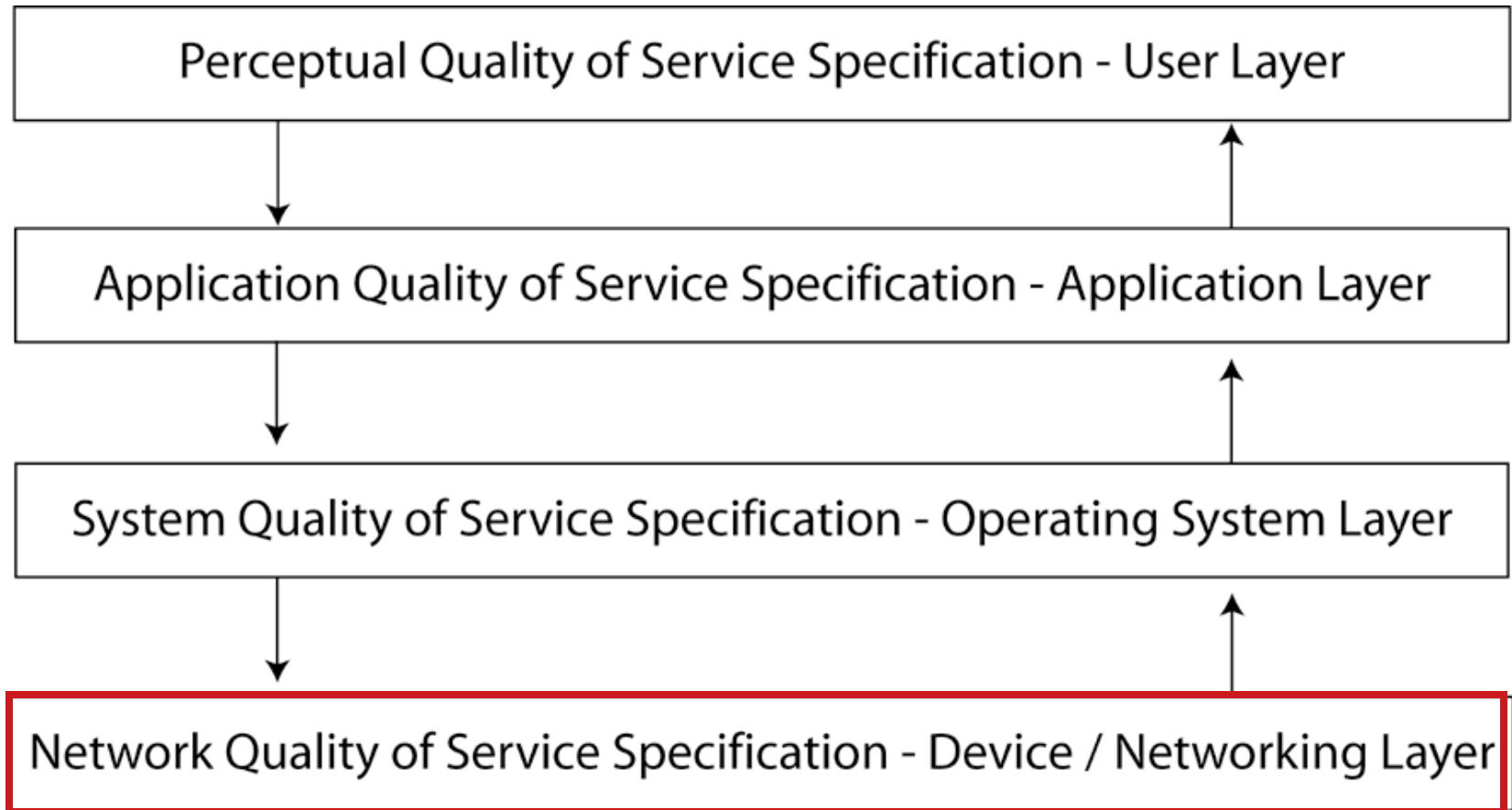


Quality of Service in Multimedia Networking

- The QoS Problem in Packet Networks
- Network QoS Operations
 - Shaping
 - Queuing & Dropping
- Architectures: DiffServ & IntServ
- Traffic Engineering
 - Multi Protocol Label Switching



QoS – Layered Model



Problem Statement

- o The standard Internet is 'Best Effort' service
 - Re-routing - Change of link properties (wireless!)
 - Heterogeneous link transitions - Congestion
- o New sensitive applications
 - Interactive media streams (for medical treatment ...)
 - Remote real-time controls
 - 'Synchronous' IP (I-SCSI)
- o ISPs want to sell special services
 - ★ Use bandwidth effectively ★ Avoid congestion collapse



Recall: VoIP/VCoIP Real-Time Requirements

- ! Latency $\approx < 100$ ms
- ! Inter-stream Latency $\approx < 30/40$ ms audio ahead/behind
- ! Jitter $\approx < 50$ ms
- ! Packet loss $\approx < 1$ %
- ! Interruption: 100 ms ≈ 1 spoken syllable
- ! Packet reordering may cause loss & jitter



Critical Issue: Jitter

Main Jitter Sources

- ⇒ Processing & multiplexing at end systems
 - Under user / end system control
- ⇒ Statistical multiplexing at (physical) network devices
 - Mainly LAN controlled
- ⇒ Random queuing delays at routers
 - Accumulate in (unknown) wide area transport



Jitter Source: End Systems

- o Adjust processing complexity and load
- o Introduce Jitter-hiding buffers/delays

- Fixed Buffer
- Adaptive Buffer:

If $p_i =$ Time of playout for the i -th packet (of timestamp t_i)

Then for appropriate K (e.g. 4 like in TCP)

$p_i = t_i + d_i + K J_i$ is an appropriate over estimator

But: playout delays may be only adjusted between spurts

▽ Playout delays distract interactivity



Jitter Source: Network - Statistical Multiplexing



- ▶ Packet delays are added randomly
- ▶ Sensitive to instantaneous load (UDP bursts)
- ▶ Timing 'out of control', even in over provisioned networks
- ▶ L2 Approaches: - 802.1p packet prioritisation,
- 802.1AVB



Ethernet 802.1Q/p - Tagging

8 bytes	6 bytes	6 bytes	2	46 - 1500 bytes	4 bytes
preamble	destination address	source address	type	data / pad	CRC

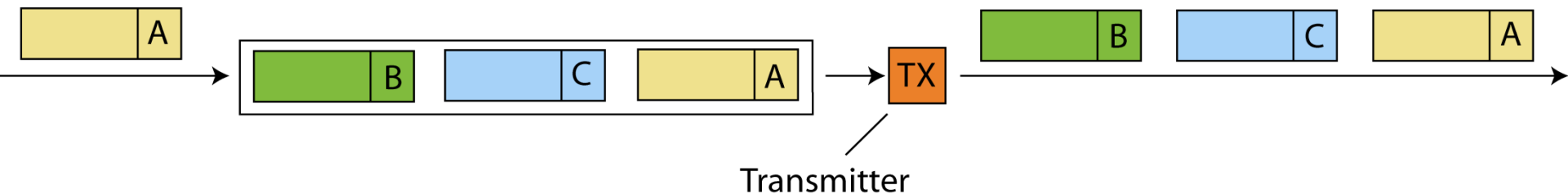
7 bytes	1	6 bytes	6 bytes	2	1	46 - 1492 bytes	4 bytes
preamble	S O F	destination address	source address	length	8 0 2	data / pad	CRC

16 bits	3 bits	1	12 bits
TPID	Priority	C F I	VLAN-ID

Tag Protocol Identifier=0x8100
Priority Tagging for 802.1p

Canonical Format Identifier
VLAN ID: 802.1Q Mapping

Jitter Source: Routing - Queuing Delays



- ▶ Queuing time in FIFO depends on queue length & loss strategy
- ▶ Load adds random delays
- ▶ Insufficient buffer space results in packet discarding
- ▶ May remain bound in over provisioned networks ?



The Nature of Internet Traffic

Internet traffic is mainly the sum of congestion controlled TCP flows with sudden bursts (UDP sources ... viruses/worms)

- o Bursts are uncontrolled and unlimited by the transport layer

- o 'Regular' TCP traffic is self-similar, not Poissonian

- Peaks add up on fractional time scales
- No i.i.d. 'Ups and Downs'
- Overflow probabilities decrease very slowly, not exponentially

⇒ There is no reliable *and* no reasonable Internetwork resource bound



What can a Network do?

Shaping & Selecting:

- Control network entry points
- Prevent bursts / overloads entering the network

Priority Queuing:

- Forward packets at different priorities

Buffering or dropping:

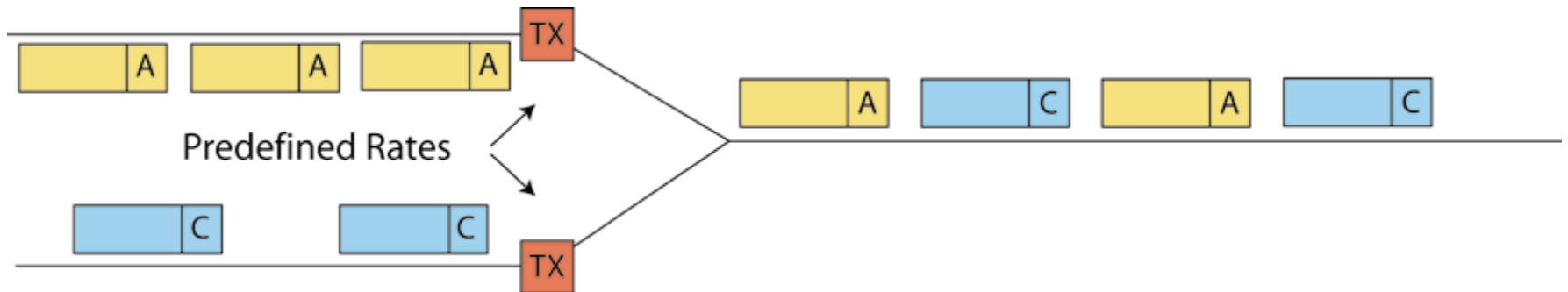
- Buffer queues add delay, no 'reasonable' length
- Rule of thumb in use: link capacity \times \langle RTT \rangle flows
- 'Blind' dropping can be harmful
- ➔ Try to use selective mechanisms

Traffic engineering:

- Balance traffic flows according to network resources



Traffic Shaping



- ▶ Simple á priori macro control: **Leaky Bucket**
- ▶ **Traffic shaping**: controlled distribution across network (per port, per protocol or per flow)
- ▶ May limit **average rates**, **peak rates** and **burst sizes**
- ▶ Fairly static: needs continuous monitoring
- ▶ Problem: network resources unused?



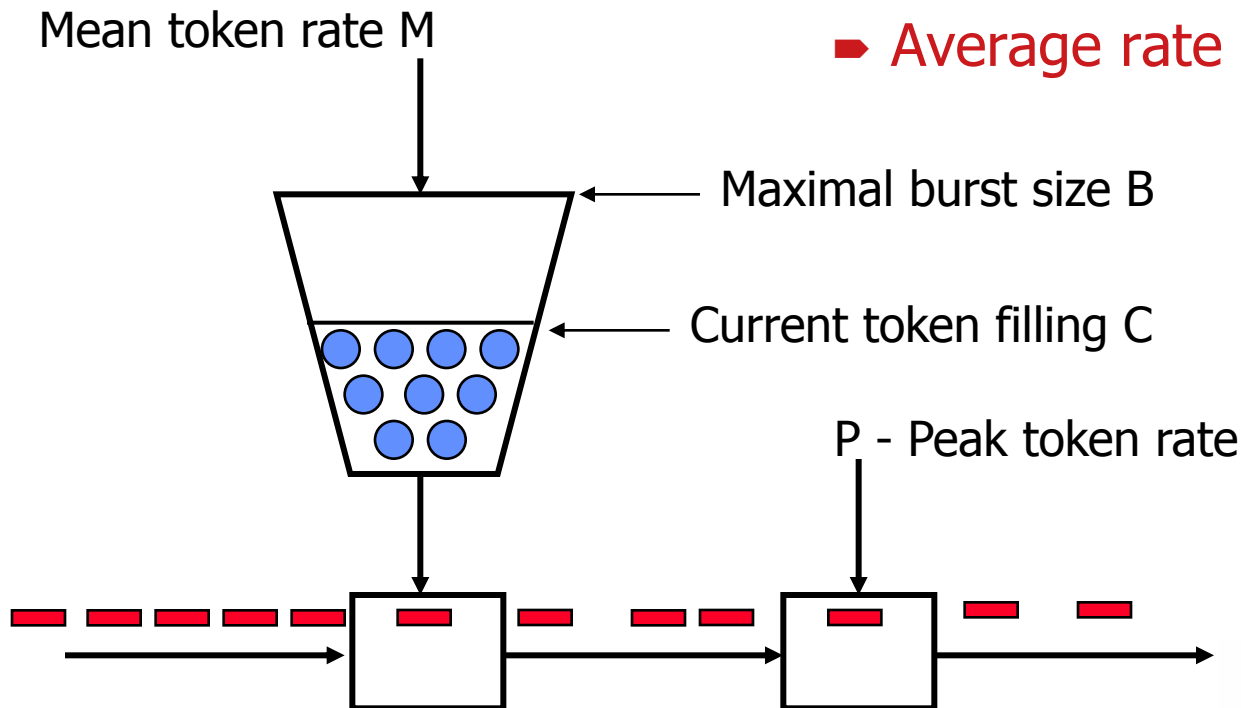
Leaky Token Bucket (Dual)

► Shape traffic to predefined limits:

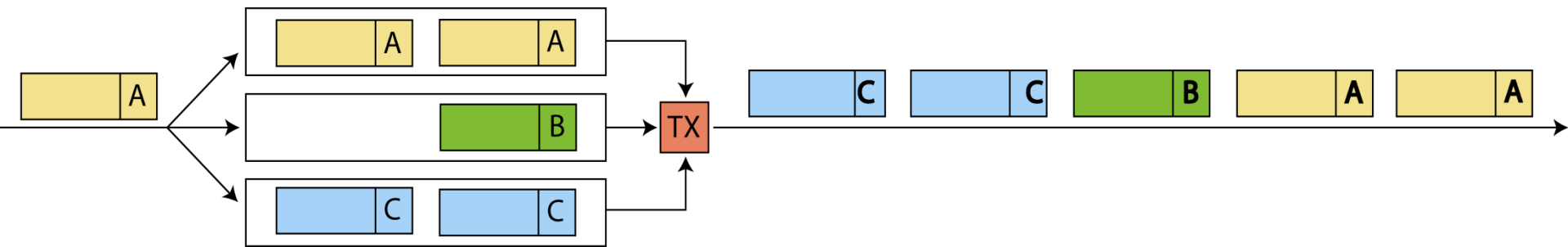
► Maximal burst size: B

► Peak rate P

► Average rate M below P



Priority Queuing



- Identified traffic assigned to different queues
- Needs scheduling:
 - Weighted Round Robin
 - Class Based Queuing
 - Weighted Fair Queuing



Queuing

Class Based Queuing - CBQ:

- Transmits packets from highest nonempty queue first

(Weighted) Round Robin - WRR:

- Visits queue after queue in round robin fashion
- Picks 1 (N_i) packets from queue i
- Problem: does not account for packet lengths

Weighted Fair Queuing - WFQ:

- Visits queues in round robin fashion
- Donates a predefined data rate to each queue



Dropping

Old better than new (WINE):

- On overload drop newest packet first (TCP-like)

New better than old (MILK):

- On overload drop oldest packet first (Real-time data)

Random Early Detection (RED):

- Start discarding packets prior to overload
- Observe watermarks of queue lengths
- Idea: TCP will slow down on packet loss
- Problem: UDP – some ideas of selective discards



Example: Balanced Network with Maximal Delay

- Suppose a traffic flow enters a network through a leaky bucket with **average rate M** and **burst limit B**
- Suppose routers with balanced links of transmission capacity **T** and WFQ forward this flow with rate **$T\omega$**
- Furthermore **$M \leq T\omega$** , then:

$\frac{B}{T\omega}$ is the maximum queue delay for any packet.



Traffic Classification

How to identify packets for QoS treatments?

- Per port (simple & rough)
- Per TOS/Traffic Class field
 - Labelling from application or at network entry point
- Per flow

Identifying Quintuple in IPv4

- Source & Destination Address
- Transport Protocol
- Source & Destination Port
- Problem: Packet fragmentation, header compression, encryption

IPv6: Flow Label



Policy-based Routing

- Policy defines
 - Forwarding and queuing strategies
 - Call admission control rules
 - Leaky bucket parameters
 - Dropping conditions
- Policy might depend on
 - Type of traffic (classification)
 - Overall resource consumption (metering results)
 - Externals like time of day, authenticated user, ...
- Automatic Policy Distribution: COPS (RFC 2748 + 4261)
 - A server actively installs policies into devices



IntServ – Integrated Service Architecture

Ambitious Solution (RFCs 2205-2212) with

- ▶ Per-flow resource reservation & queuing at all routers
- ▶ Quality of service for sessions (end-to-end)
- ▶ Hard guarantees desired

Two service types defined:

- ▶ **Guaranteed Service**: guaranteed bandwidth, firm bounds on end-to-end queuing delays
- ▶ **Controlled Load**: approximates congestion-free network

But:

- ▶ High complexity - Vulnerable to flow state attacks
- ▶ Needs support of all routers - Low scalability



IntServ

- Provide mechanisms to reserve resources (link bandwidth, buffers) at routers along the path of each flow.
- Flow context used to drive a token bucket
- Initial call setup to implement QoS states at routers:
 - ▶ Requested QoS – Rspec
 - ▶ Traffic characteristic – Tspec
- Signalling process with Resource reSerVation Protocol (RSVP)
- Initiates virtual queues at routers: one for each flow

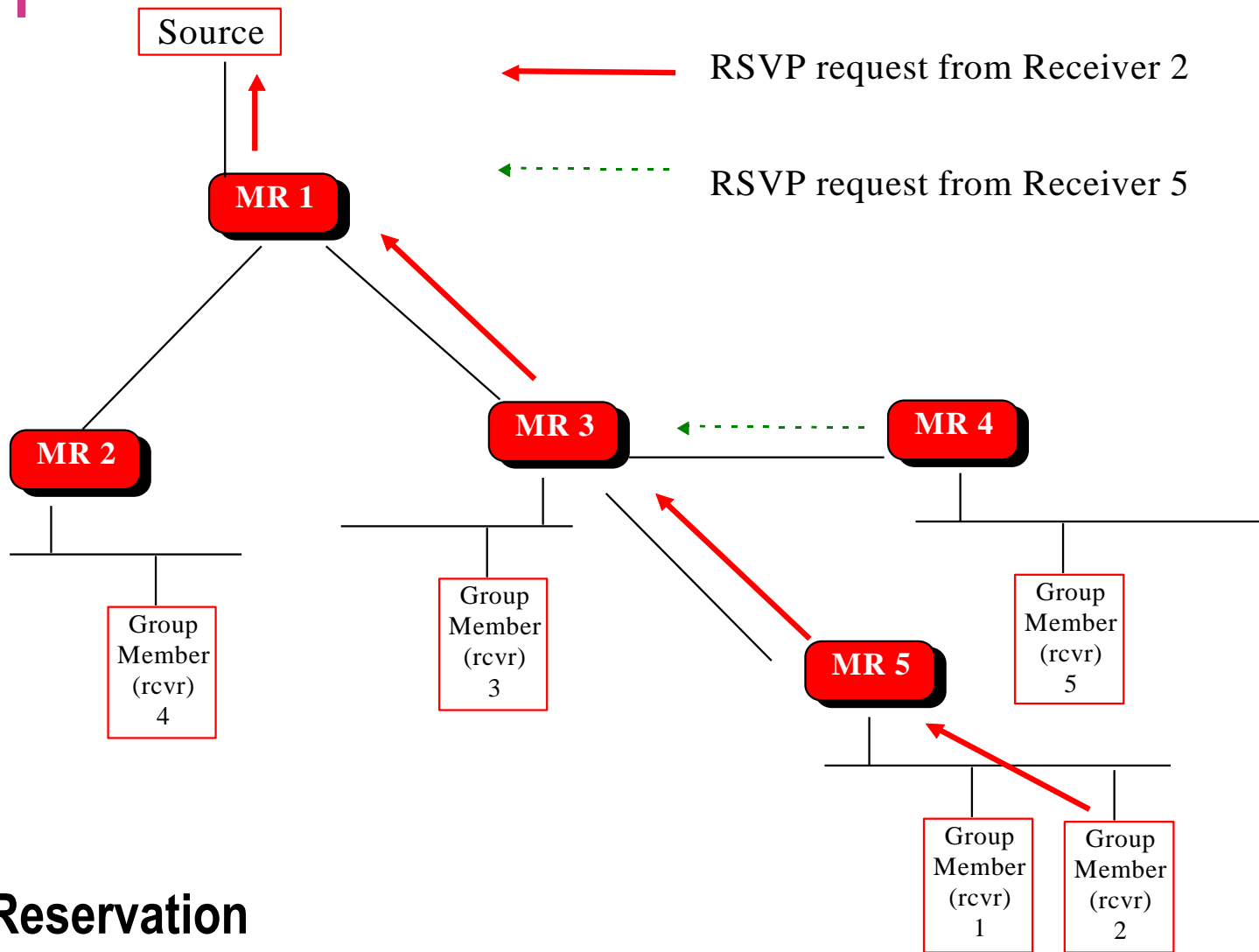


Resource reSerVation Protocol (RSVP)

- Signalling protocol to reserve router resources along a path
- RFC 2205 (Zhang et al, 1997)
- Resource reservation for multicast distribution trees (including unicast)
- Destination oriented reservations
 - Sender pushes periodically PATH messages (establish router states)
 - Receiver answers with RESV packets
 - Routers interpret these along the paths
- Involves applications and all intermediate devices
- Soft-State-Concept: reservation states with lifetime



RSVP

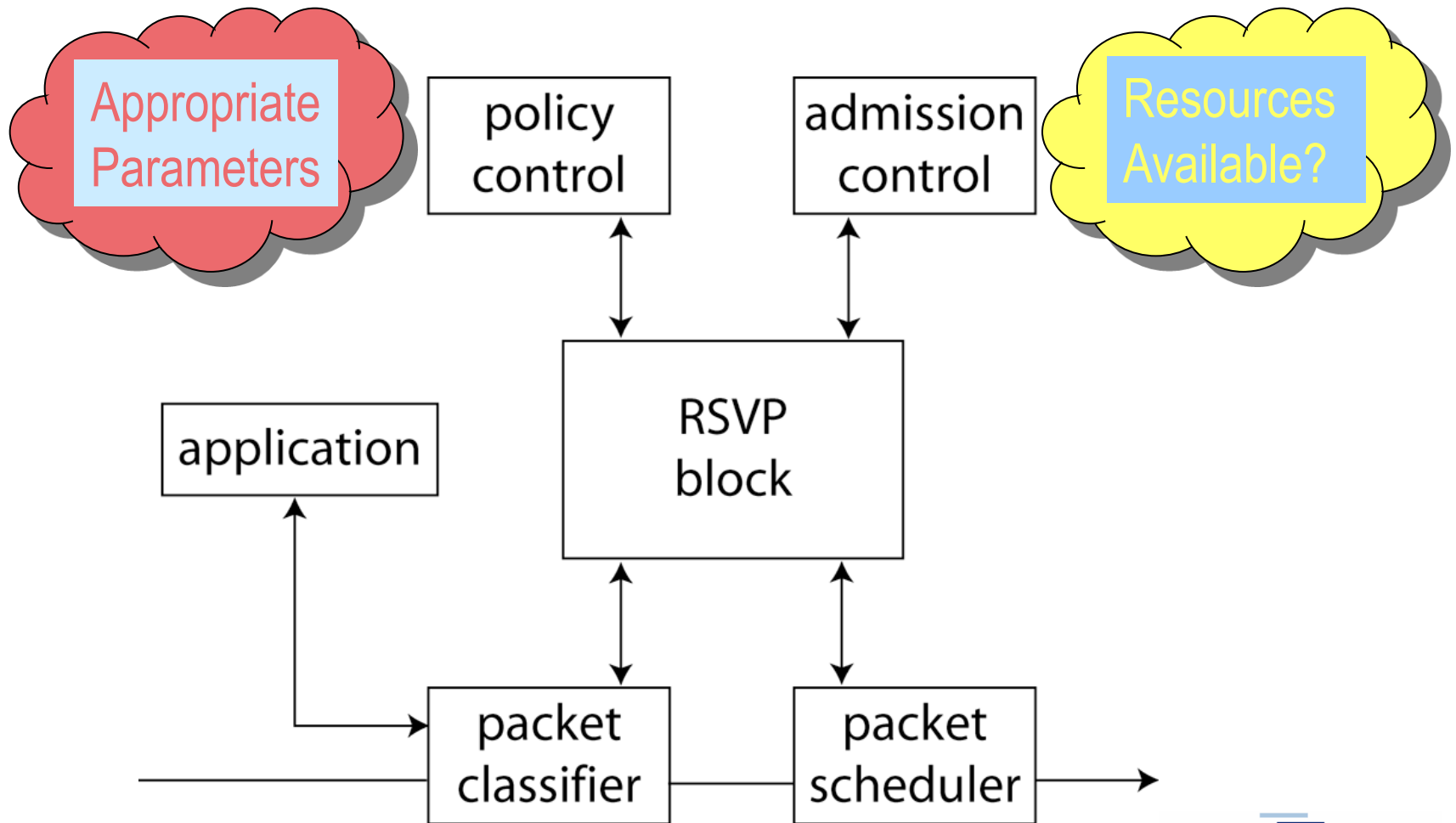


Group Reservation

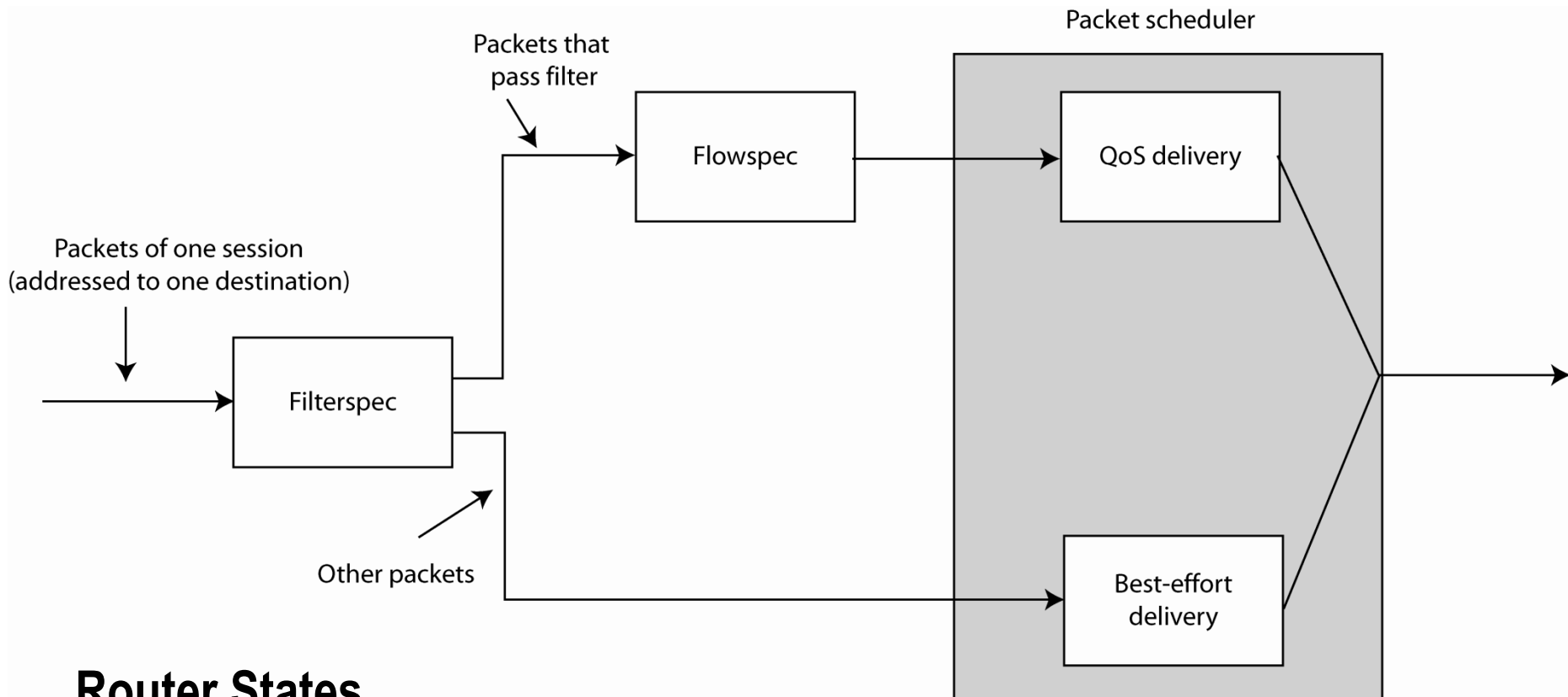
RSVP defines **QoS paths** from receiver (to specific source)

Resource reservations are merged when possible (on flow identification)

RSVP Functional Blocks



RSVP per Router Scheduling



Router States

Filterspec: defines packets of flows with QoS reservation

Flowspec: defines QoS parameters per flow for scheduler



DiffServ- Differentiated Service Architecture

Less ambitious solution (RFC 2475,3260) with

- ▶ Different services for different classes of traffic
- ▶ No guaranteed quality of service (end-to-end), but
- ▶ Controlled **Per-Hop Behaviour (PHB):**
Expedited / Assured Service Groups

Using

- ▶ Traffic classification (ToS/Traffic Class = DiffServ field)
- ▶ Per-class queuing (no distinctive flows)

Aiming at scalable, efficient, easy-to-deploy QoS services



Differentiated Services: Components & Terminology

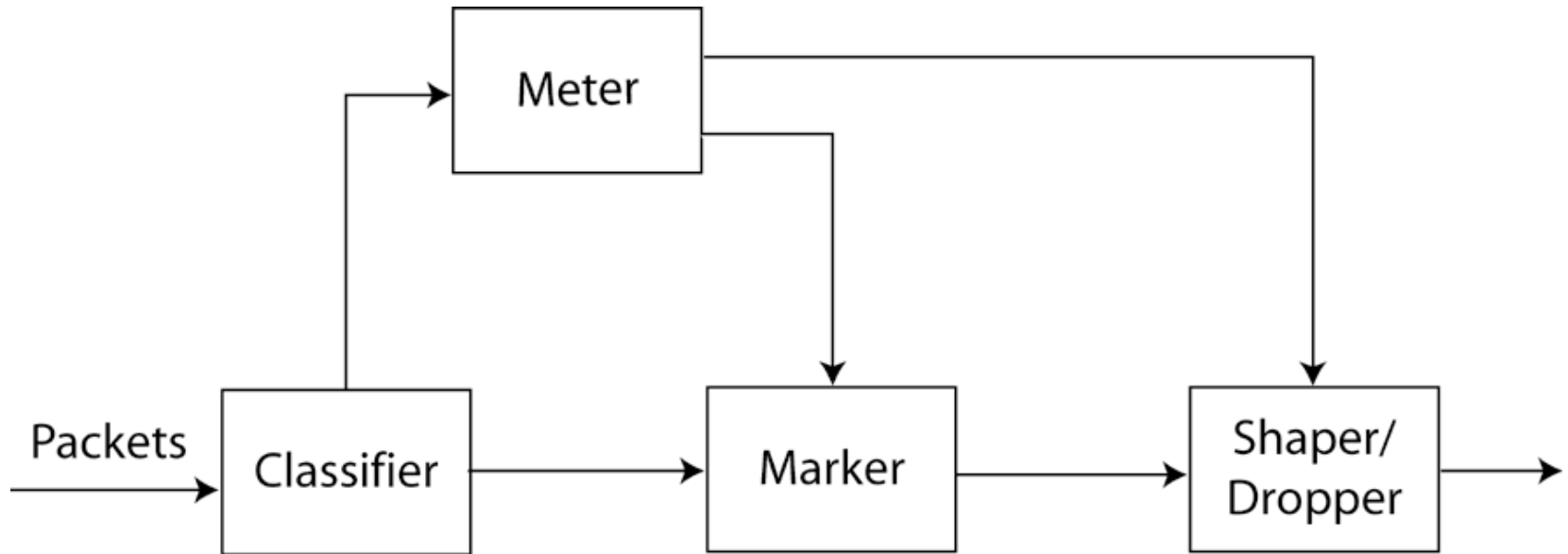
- **Service Level Specification (SLS)**: a set of parameters/values, which together define the service offered by a DS domain
- SLS is based on **Traffic Condition Specification (TCS)**: a set of parameters specifying classifier rules and a traffic profile
- **Classifying, metering and marking** at boundary nodes, no application dependence

At Router

- ▶ Queuing and forwarding based on **DiffServ Codepoints**
- ▶ Traffic aggregation according to Codepoints
- ▶ No connection states



Diffserv: Traffic Conditioner



- **Classifier:** Separate packets into classes
- **Meter:** Measure submitted traffic for conformance profile
- **Marker:** Polices by (re-)marking packets with codepoints
- **Shaper/Dropper:** Delays / discards packets

DiffServ: Service Details

- To attain “Network Services”, isolated per-hop behaviours must be coordinated to PHB groups:
- Expedited Forwarding Behaviour (EF):
 - “Virtual leased line” service
 - Simple service model for small delay/real time apps
 - Aggregated flows bound by peak bandwidth
 - Ingress router: policing/dropping – Egress router: shaping
- Assured Forwarding Behaviour (AF):
 - Complex service type with support for bursty flows
 - Defines different classes with independent resources as AF instances
 - Three drop precedences for each class (“Bronze”, “Silver”, “Gold”)



Resource Allocation

Resources are allocated by marking IP packets with appropriate DiffServ **Codepoints** at boundary nodes (also network transition points):

- **Static**: Mark packets by IP-address and/or protocol port
- **Bandwidth Broker** (RFC 2638): Unit to configure resources from network-wide policy table (at ingress+egress routers)
- **Dynamic with BB**: Router states are monitored by BB to optimise network resource utilisation and performance (dynamic TCSs).
- **QoS signalling**: Common Open Policy Service Protocol (COPS, RFC 2748)



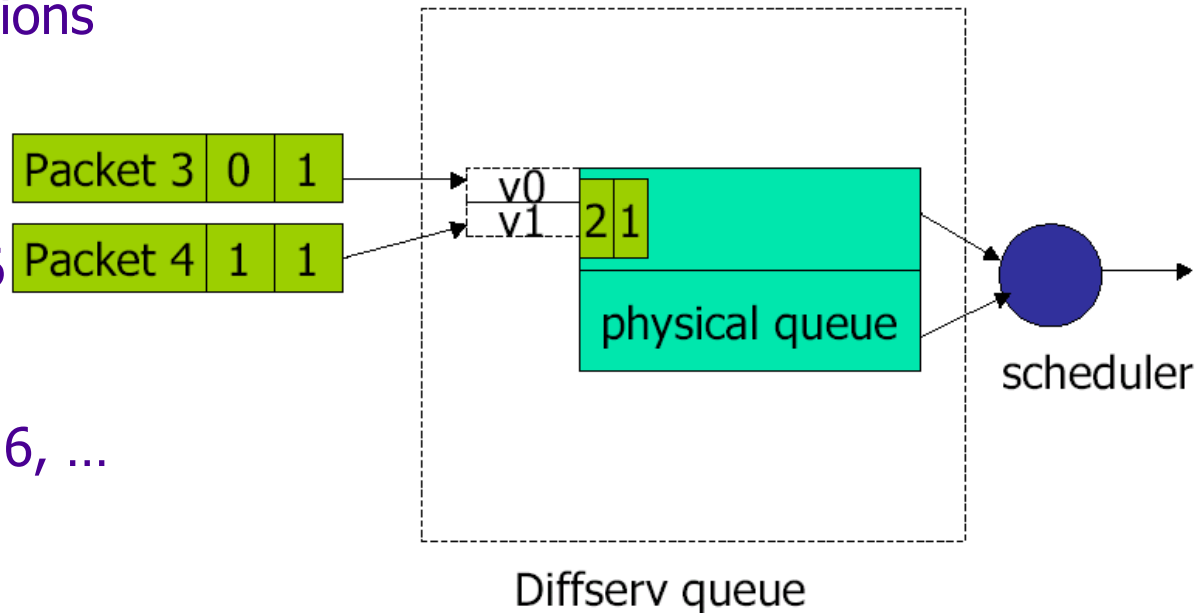
DiffServ Field: Codepoints

- Defined in RFC 2474 ++
- General form: xxxxxxRR (= 64 possible Codepoints)
- Standard Assignment: xxxxx0 (Default: 000000)
- IPv4 compatibility: xxx000
Queue-Service and Congestion Control as in RFC 1812
- Assured Forwarding as in RFC 2597: Four classes, each with three drop precedences – AF1x, AF2x, AF3x, AF4x, x= 1 ... 3:
- Expedited Forwarding as in RFC 3248: 101111
- Experimental: xxxxx1

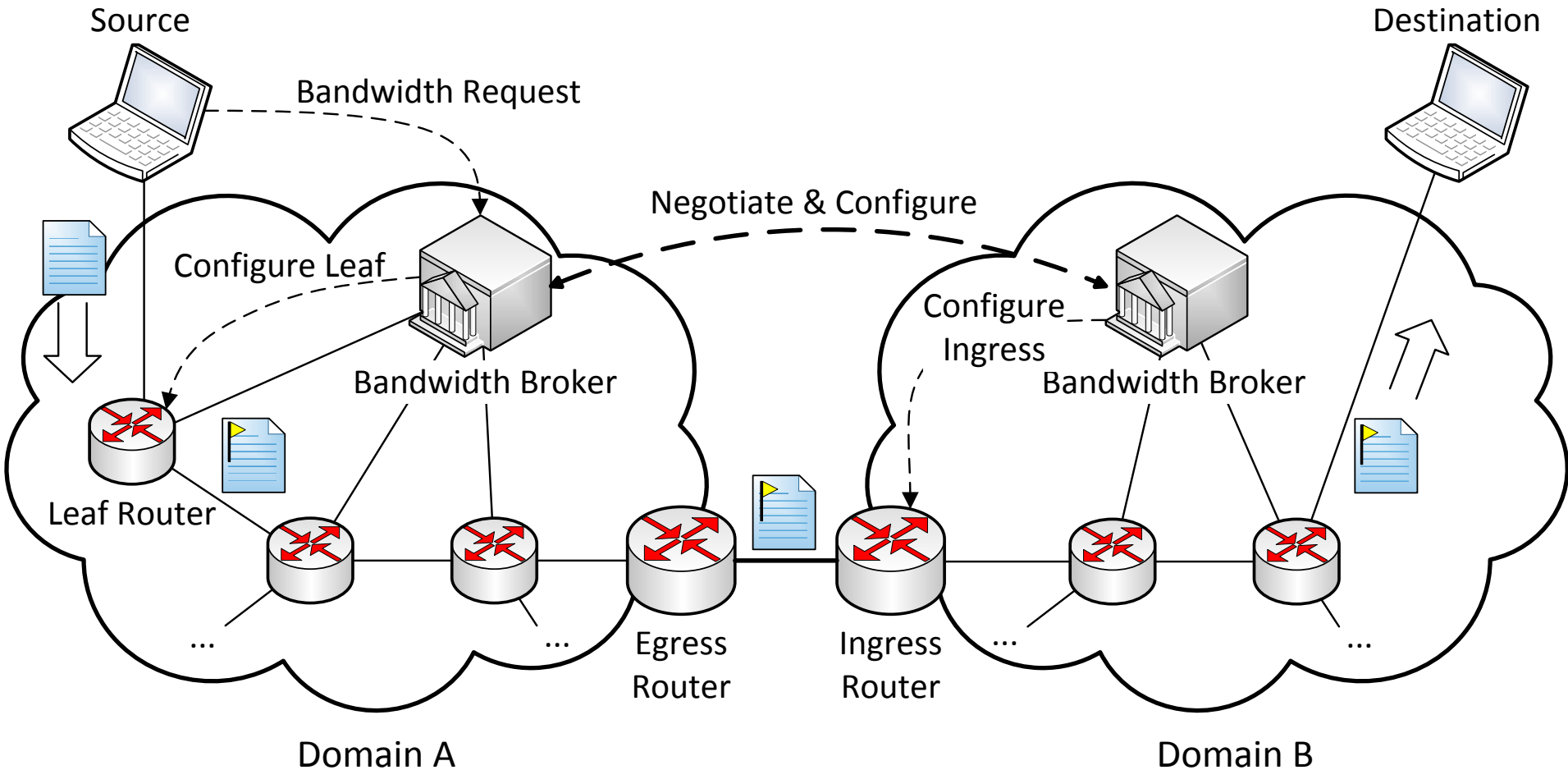
Drop Prec:	Class 1	Class 2	Class 3	Class 4
Low	001010	010010	011010	100010
Medium	001100	010100	011100	100100
High	001110	010110	011110	100110

DiffServ Virtual Queues: Mapping Problem

- DiffServ does not define implementation details (separation of forwarding & control)
- Problem: Mapping of logical to physical resources
- L3 virtual to physical queues: Vendor implementations
- LAN resources (e.g. 802.1p): IEEE & RFC 2814-16
- WLAN resources: IEEE 802.11e, 802.16, ...



DiffServ Architecture



IntServ vers. DiffServ, Quo vadis QoS ?

IntServ: Flexible, granular, application oriented service
but: does not scale

DiffServ: Scalable, provider oriented, easy deployable service
but: application-ignorant

→ **Approach:** IntServ (edges) over DiffServ (core)

General Issues (RFC2990 from IAB):

- ▶ State versus statelessness in QoS?
- ▶ Inter-Domain signalling?
- ▶ Which mechanisms will form an end-to-end QoS architecture?
- ▶ Transport layer issues – what to do with TCP?
- ▶ Security and accounting open ...

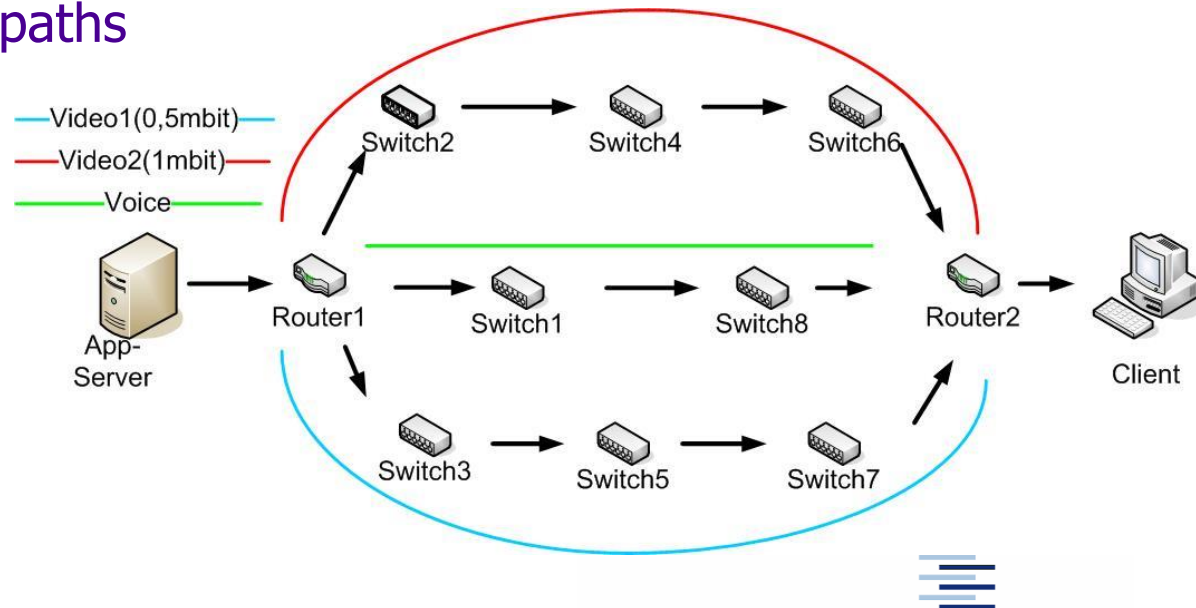


Traffic Engineering

Problem: IP routing traditionally follows shortest paths. This may lead to overloaded links, while the physical infrastructure offers meshes

Traffic engineering is concerned with

- ▶ discovering current traffic load
- ▶ discovering alternate paths
- ▶ directing traffic



Traffic Engineering

- Simple Approach: **Equal Cost Multipath routing (ECMP)**
 - Local decision at branch router
- Discovery of on-local network utilization:
Explicit Congestion Notification – ECN
 - ECN Codepoints in Traffic Class field
- Problem: Route overlays according to L2 properties or QoS requirements?
 - Initially: Exploit ATM VCs
 - IP: Source Routing or IP in IP tunnelling
- IETF's answer: Simplified 'tunnel' tag (label)
 - Inserted below IP
 - Multi Protocol Label Switching (RFC 3031 ++)

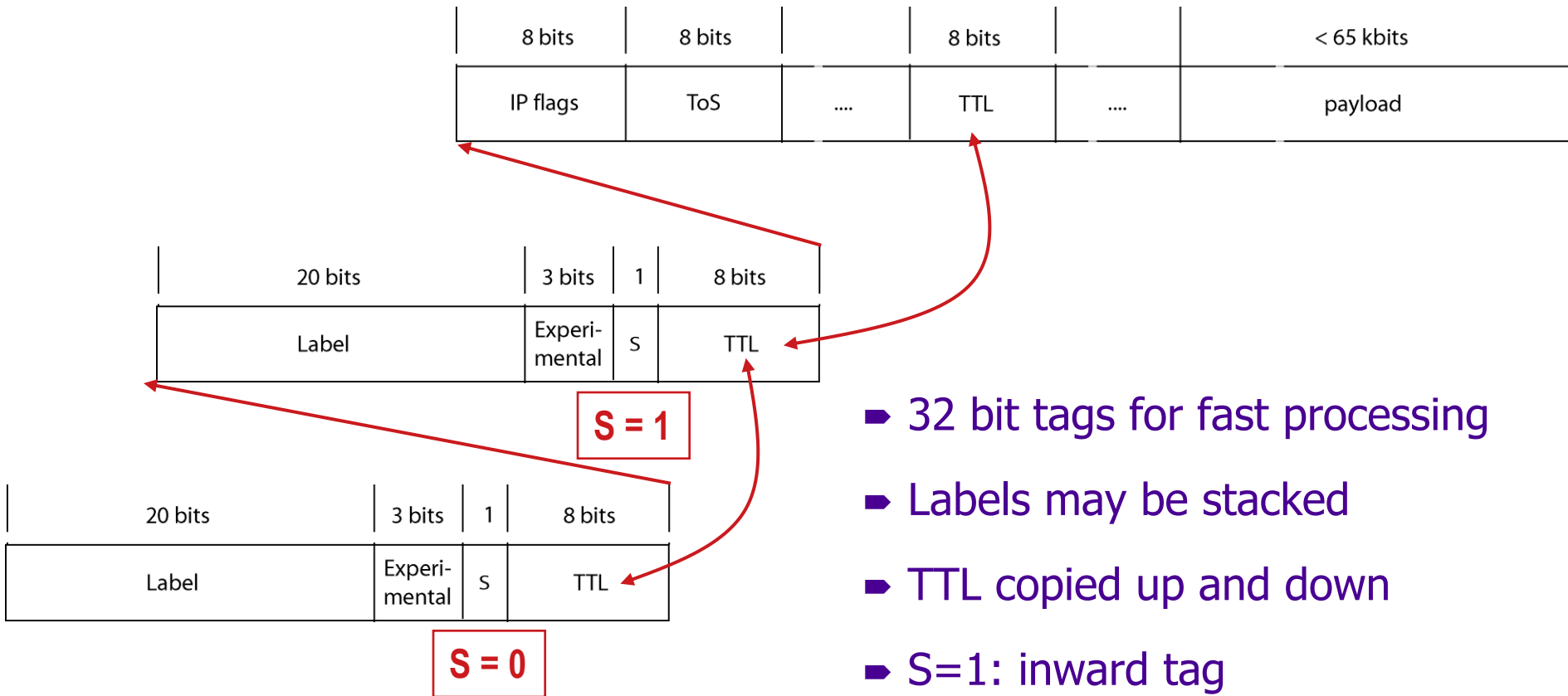


Multi Protocol Label Switching - MPLS

- Shim header to label packets
- Label data limited to forwarding plane
- Label switching routers (LSR) forward on label switching paths
- Instruction Table: Label Forwarding Information Base (LFIB)
- Insert / remove labels at edge routers (LER)
- Label distribution via Label Distribution Protocol (LDP)

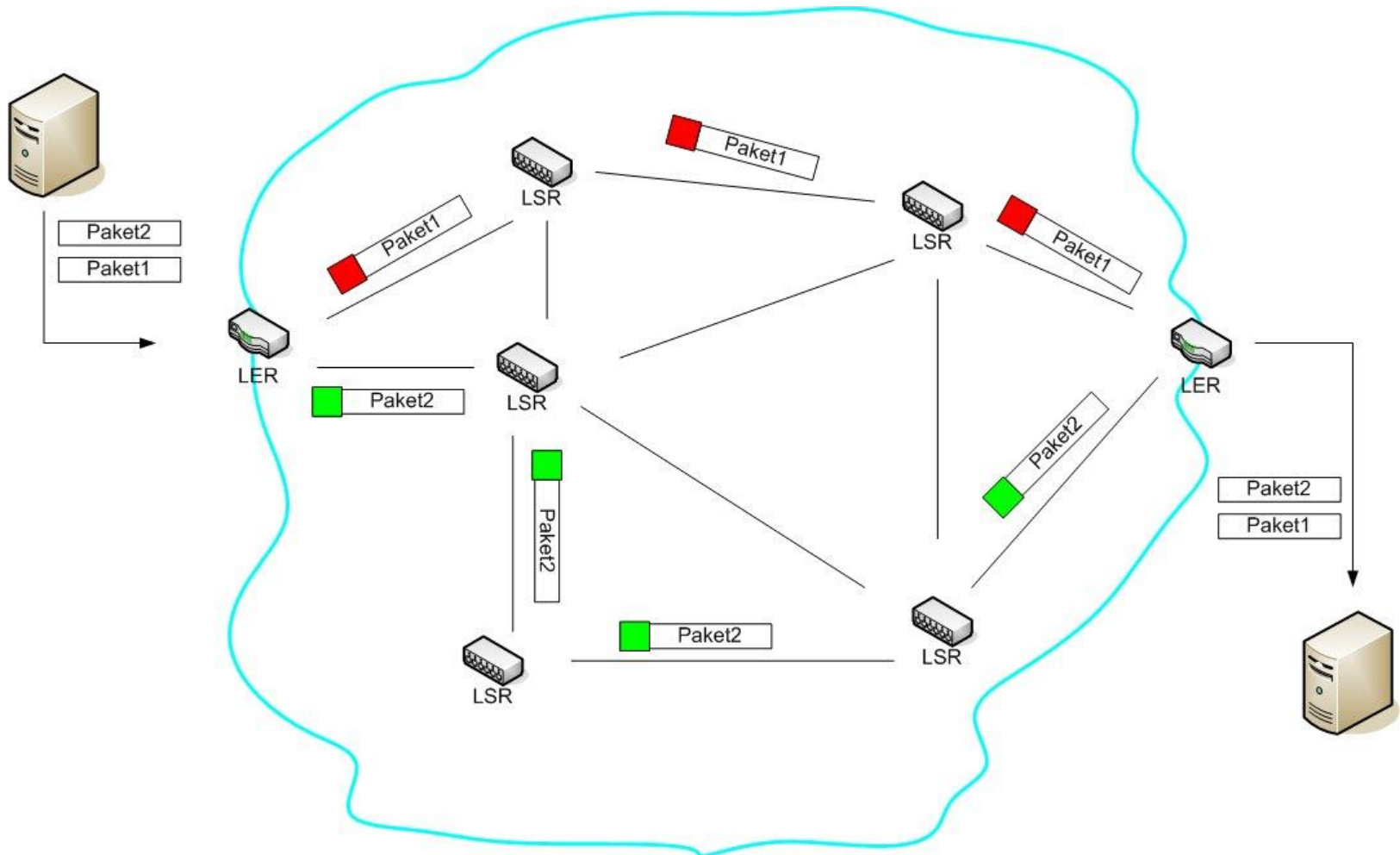


MPLS Tagging



- ▶ 32 bit tags for fast processing
- ▶ Labels may be stacked
- ▶ TTL copied up and down
- ▶ S=1: inward tag
- ▶ Routing: push/pop/swap

Label Switched Paths



Label Distribution Protocol (LDP)

- Functions of LDP
 - Discovery of adjacent LDP peers
 - Control negotiations on capabilities and options
 - Label advertisement and withdrawal
- LDP peers establish sessions after Hello multicast messages that announce a label space
- Label distribution in downstream direction
 - Unsolicited, or
 - On Demand



Multi Protocol λ Switching - MP λ S (GMPLS)

- Basis: Wavelength (λ) Division Multiplexing (WDM)
 - Optical packet switching (based on colours)
- Option to route IP over λ s
 - Needs IP layer decision at branches
- Easier and more efficient:
 - MPLS overlays represented as λ s (λ = label)
- But: heavy layer violation!



QoS via MPLS

- IntServ over MPLS
 - Set up a label switched RSVP tree
 - Extension to RSVP: RSVP-TE (RFC 3209, 3936), Label request/reserve
- DiffServ over MPLS
 - Constraint-based LS-Path setup using LDP (RFC 3212, 3468)
 - Group packets according to Codepoints
 - Differing approaches (E-LSP, L-LSP) on EF and AF service treatment



Deployment Practice:

- (G)MPLS is a Success Story
 - ▶ Widely deployed at provider level
 - ▶ Some deployment across providers (e.g., tagged transit)
- IP-layer Technologies Hesitant to Spread
 - ▶ Some commercial DiffServ / Expedited Forwarding offers
 - ▶ IntServ bound to 'Walled Gardens'
- Congestion Control & Resource Pooling
 - ▶ Tendency to treat congestion on Transport layer (e.g., ECN in TCP)
 - ▶ Increasing activities to support multipath Transport



Reading

- Michael Welzl: **Network Congestion Control**, Wiley, Chichester, UK, 2005.
- Adrian Farrel: **The Internet and Its Protocols**, Morgan Kaufmann, 2004.
- J.Shin, D. Lee, C.Kuo: **Quality of Service for Internet Multimedia**, Prentice Hall, Upper Saddle River, NJ, 2004.
- Rao, Bojkovic, Milovanovic: **Multimedia Communication Systems**, Prentice Hall, Upper Saddle River, NJ, 2002.
- G. Huston: **Next Steps for the IP QoS Architecture**, RFC 2990, November 2000.
- IETF Documents: www.rfc-editor.org
- IEEE Documents: www.ieee.org

