

## 9 QoS Concepts

### 9.1 Network Transmission Quality

- Ursachen von congestion (Stau)
  - Aggregation (viele Links → Uplink)
  - große Bandbreite → kleine Bandbreite (speed mismatch, LAN to WAN)
- congestion → queue → delay
- Priority Queue
- Sources of Delay
  - Code delay: compress data at the source (fixed time)
  - Packetization delay: encapsulate packet (fixed)
  - Queuing delay (variable)
  - Serialization delay: transmit frame onto wire (fixed)
  - Propagation delay: travel between source and destination (variable)
  - De-jitter delay: buffer flow of packets, send out in evenly spaced intervals (fixed)
- Jitter: variation of delay of received packets
- Playout delay buffer: send de-jittered stream (packet drop due to excessive jitter)

### 9.2 Traffic Characteristics

	Voice	Video	Data
Smooth (constant rate)	✓	-	✓
Bursty	-	✓	✓
Benign (freundlich)	✓	-	✓
Greedy (gierig)	-	✓	✓
Drop sensitive	✓	✓	-
Delay sensitive	✓	✓	-
UDP priority	✓	✓	-
TCP retransmits	-	-	✓
One-Way Requirements			
Latency	≤ 150ms	≤ 200-400ms	-
Jitter	≤ 30ms	≤ 30-50ms	-
Loss	≤ 1%	≤ 0,1-1%	-
Bandwidth	30 - 128 kb/s	384 kb/s - 20 Mb/s	-

### Factors to Consider for Data Delay

Factor	Mission Critical	Not Mission Critical
Interactive	Prioritize for the lowest delay, 1-2s response time	benefit from lower delay
Not interactive	supply minimum bandwidth	lowest priority

## 9.3 Queuing Algorithms

- FIFO (First In first Out): keine Priorisierung, Pakete werden direkt weitergeleitet
- WFQ (Weighted Fair Queuing)
  - automatische Zuordnung der Pakete anhand ToS (IPv4-Header) bzw Traffic Class (IPv6-Header)
  - Hohe Priorität: low-volume, interactive traffic (Telnet, Voice)
  - Niedrige Priorität: high-volume traffic (FTP)
  - Funktioniert nicht mit Tunneling und Encryption
- CBWFQ (Class Bases Weighted Fair Queuing):
  - user-defined traffic classes (protocol, ACLs, input interface)
  - Ein FIFO-Queue für jede Class
  - user-defined bandwidth, weight, and maximum packet limit für jede Class
  - keine Class hat absolute Priorität
- LLQ (Low Latency Queuing)
  - PQ (priority queuing): absolute Priorität für delay-sensitive Pakete (Voice)
  - Andere Pakete werden mit CBWFQ behandelt

## 9.4 QoS Models

- Best-effort model: kein QoS, Paket ist wie Brief
- IntServ (Integrated Services Model)
  - Hard QoS: guarantees bandwidth, delay, and packet-loss rates, from end to end
  - connection-oriented (wie Telefon)
  - application requests a specific kind of service
  - RSVP (Resource Reservation Protocol): signal QoS needs along devices in the end-to-end path
  - Resource intensive, not scalable to large implementations such as the internet
- DiffServ (Diff Services Model)
  - Soft QoS: Router classify traffic hop-by-hop
  - low-latency guaranteed service for voice or video
  - best-effort for web traffic or file transfers
  - kein end-to-end service, kein signaling, scalable

## 9.5 QoS Implementation Techniques

### Avoiding Packet Loss

- Increase link capacity
- Increase buffer space
- Drop lower-priority packets

### QoS Tools

- 1. Classification and 2. Marking tools
  - Layer 2: Ethernet (802.1Q, 802.1p): CoS (Class of Service) (3 bits)
  - Layer 2: Wi-Fi (802.11): TID (3 bits)
  - Layer 2: MPLS: EXP (3 bits)
  - Layer 3: IPv4 alt: IPP (IP Precedence:) (3 bits im ToS)
  - Layer 3: IPv4 and IPv6 neu: DSCP (Differentiated Services Code Point) (6 bits)
  - Trust boundaries: Trusted Endpoints klassifizieren Traffic (IP-Phone, L2-Switch, L3-Switch)
- 3. Congestion avoidance tools (Drop packets): WRED (Weighted random early detection)
  - Je voller der Queue ist, desto mehr % der Pakete werden gedroppt.
  - WRED: weighted random early detection (nur für TCP, Window-size?)
- Congestion management tools: 4. Queuing: CBWFQ, LLQ, 5. Scheduling, shaping
  - Shaping traffic (outbound): Pakete zwischenspeichern und gleichmäßig aussenden
  - Policing (inbound): Provider kappt bei Überschreitung der CIR (customer information rate)

### QoS Policy Guidelines

- Enable queuing at every device in the path between source and destination
- Classify and mark traffic as close to the source as possible.
- Shape and police traffic flows as close to their sources as possible.