Dienstzugangspunkte in der TCP/IP-Architektur[1]

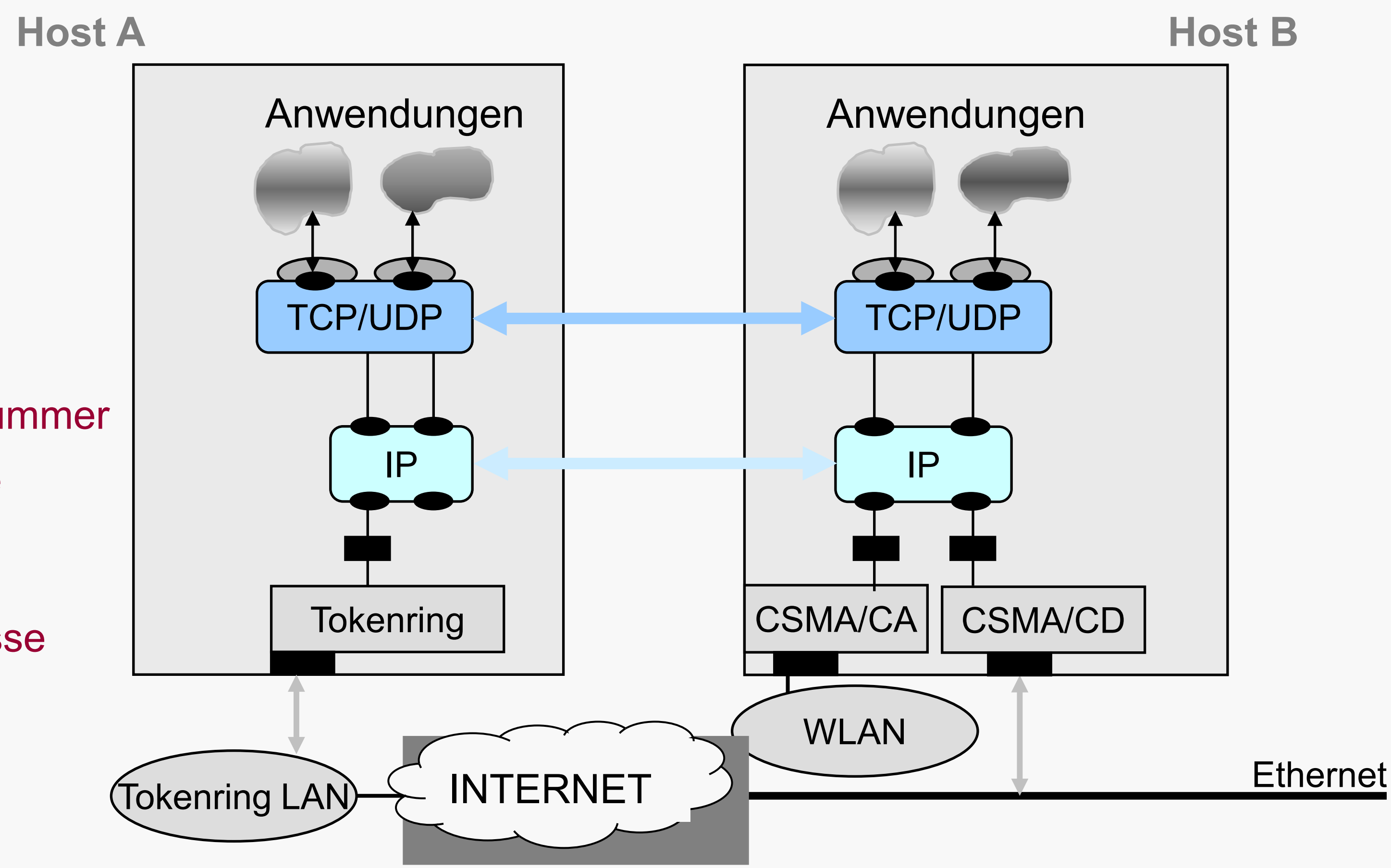
Sockets/
Ports

Protokollnummer

IP-Adresse

IP-Version

MAC-Adresse



1) nach Stainov, R.: IPng Das Internet-Protokoll der nächsten Generation. Thompson Publishing, 1997

Assuming that an organization has registered the OUI of AC-DE-48 and that the organization has created the MAC-48 value of AC-DE-48-23-45-67 by concatenating the extension identifier 23-45-67, this MAC-48 identifier has the following binary transmission order:

```
|              OUI              |     extension identifier     | field
|  1st   |   2nd    |  3rd      |   4th    |   5th    |   6th    | octet
| C    A  |  E    D  |  8    4  |  3    2  |  5    4  |  7    6  | hex
0011 0101 0111 1011 0001 0010 1100 0100 1010 0010 1110 0110 bits
|       | |       | |       | |       | |       | |       |
lsb   msb lsb   msb lsb   msb lsb   msb lsb   msb lsb   msb
```

Source: Wikipedia

Structure of MAC Adresses

Assignment of MAC Adress-Space to Organisations (Excerpt)

Registry:

Hex	Name	Reference
00000C	Cisco	
00000E	Fujitsu	
00000F	NeXT	
000010	Sytek	
00001D	Cabletron	
000020	DIAB (Data Intdustrier AB)	
000022	Visual Technology	
00002A	TRW	
000032	GPT Limited (reassigned from GEC Computers Ltd)	
00005A	S & Koch	
00005E	IANA	
000065	NetScout Systems, Inc.	[Singhal]
00006B	MIPS	
000077	Interphase Corporation	
00007A	Ardent	
000080	Cray Communications A/S	
000089	Cayman Systems Gatorbox	
000093	Proteon	
00009F	Ameristar Technology	
0000A2	Wellfleet	
0000A3	Network Application Technology	
0000A6	NetScout Systems, Inc. (internal assignment, not for products)	[Singhal]
0000A7	NCD X-terminals	
0000A9	Network Systems	
0000AA	Xerox Xerox machines	
0000B3	CIMLinc	
0000B7	Dove Fastnet	
0000BC	Allen-Bradley	
0000C0	Western Digital	
0000C5	Farallon phone net card	
0000C6	HP Intelligent Networks Operation (formerly Eon Systems)	
0000C8	Altos	

Registry:

Hex	Name	Reference
0000C9	Emulex Terminal Servers	
0000D0	Develcon	
0000D7	Dartmouth College (NED Router)	
0000D8	3Com? Novell? PS/2	
0000DD	Gould	
0000DE	Unigraph	
0000E2	Acer Counterpoint	
0000EF	Alantec	
0000FD	High Level Hardvare (Orion, UK)	
000102	BBN BBN internal usage (not registered)	
0010D1	BlazeNet	
001700	Kabel	
0020AF	3COM ???	
0020C9	Victron	
002094	Cubix	
00802B	IMAC ???	
00802D	Xylogics, Inc. Annex terminal servers	
008037	Ericsson	[Johnson]
008064	Wyse Technology / Link Technologies	
00808C	NetScout Systems, Inc.	[Singhal]
0080C2	IEEE 802.1 Committee	
0080D3	Shiva	
00A03E	ATM Forum	
00AA00	Intel	
00DD00	Ungermann-Bass	
00DD01	Ungermann-Bass	
020701	Racal InterLan	
020406	BBN BBN internal usage (not registered)	
026086	Satelcom MegaPac (UK)	
02608C	3Com IBM PC; Imagen; Valid; Cisco	
02CF1F	CMC Masscomp; Silicon Graphics; Prime EXL	
080002	3Com (Formerly Bridge)	

Source: http://www.iana.org/assignments/ieee-802-numbers

Besondere IPv4-Adressen nach RFC 3330:

CIDR-Adressblock	Adressbereich	Beschreibung	RFC
0.0.0.0/8	0.0.0.0 bis 0.255.255.255	aktuelles Netz (nur als Quelladresse gültig)	RFC 3232 (ersetzt RFC 1700)
10.0.0.0/8	10.0.0.0 bis 10.255.255.255	Netzwerk für den privaten Gebrauch	RFC 1918
100.64.0.0/10	100.64.0.0 bis 100.127.255.255	Mehrfach benutzter Adressbereich für Provider-NAT	RFC 6598
127.0.0.0/8(1)	127.0.0.0 bis 127.255.255.255	Localnet	RFC 3330
169.254.0.0/16	169.254.0.0 bis 169.254.255.255	Zeroconf	RFC 3927
172.16.0.0/12	172.16.0.0 bis 172.31.255.255	Netzwerk für den privaten Gebrauch	RFC 1918
192.0.0.0/24	192.0.0.0 bis 192.0.0.255	reserviert, aber zur Vergabe vorgesehen	
192.0.2.0/24	192.0.2.0 bis 192.0.2.255	Dokumentation und Beispielcode (*TEST-NET-1*)	RFC 5737 (ersetzt RFC 3330)
192.88.99.0/24	192.88.99.0 bis 192.88.99.255	6to4-Anycast-Weiterleitungspräfix	RFC 3068
192.168.0.0/16	192.168.0.0 bis 192.168.255.255	Netzwerk für den privaten Gebrauch	RFC 1918
198.18.0.0/15	198.18.0.0 bis 198.19.255.255	Netz-Benchmark-Tests	RFC 2544
198.51.100.0/24	198.51.100.0 bis 198.51.100.255	Dokumentation und Beispielcode (*TEST-NET-2*)	RFC 5737
203.0.113.0/24	203.0.113.0 bis 203.0.113.255	Dokumentation und Beispielcode (*TEST-NET-3*)	RFC 5737
224.0.0.0/4	224.0.0.0 bis 239.255.255.255	Multicasts (früheres Klasse-D-Netz)	RFC 3171
240.0.0.0/4	240.0.0.0 bis 255.255.255.255	reserviert (früheres Klasse-E-Netz)	RFC 3232 (ersetzt RFC 1700)
255.255.255.255[2)]	255.255.255.255	Broadcast	

Source: Wikipedia

Reserved IP-Address Spaces

Source NAT [Bearbeiten]

Bei jedem Verbindungsaufbau durch den Client wird die Quell-IP-Adresse durch eine des Routers ersetzt. Außerdem wird der Quellport durch einen freien Port des Routers ersetzt, der dadurch belegt wird. Diese Zuordnung wird in der NAT-Table des Routers gespeichert. Der Vorgang wird als PAT (Port and Address Translation) bezeichnet.

lokales Netz (LAN)

Quelle	Ziel
192.168.0.2:4701	170.0.0.1:80
192.168.0.3:5387	170.0.0.1:80
192.168.0.4:1001	170.0.0.1:23

Router

===== = =====>

NAT

öffentliches Netz (WAN)

Quelle	Ziel
205.0.0.2:6787	170.0.0.1:80
205.0.0.2:8709	170.0.0.1:80
205.0.0.2:4806	170.0.0.1:23

Source-NAT und IP-Routing am Beispiel [Bearbeiten]

In diesem Beispiel nutzt das private Netz die IP-Adressen 192.168.0.0/24. Zwischen diesem Netz und dem öffentlichen Internet befindet sich ein *Source-NAT-Router* mit der öffentlichen Adresse 205.0.0.2/32.

Source: Wikipedia

Illustration of Network Address Translation

...

Functional numbering and location dependent addressing General

There is a requirement to be able to address communications to a ‘functional number’ rather than a telephone number. Such numbers are generally only associated with a user for a limited period of time. (I)

The addressing scheme can be divided into two areas: (I)
− functional addressing;
− location dependent addressing.

Functional addressing
Principles
An addressing scheme shall be provided which permits users to be identified by numbers corresponding to their functional roles rather than by numbers tied to the terminal equipment that they are using. (M)

...

Definition of Functional Numbers (Eirene Project / GSM-R)

Synchronisation

In bestimmten Situationen der Protokollausführung ist es erforderlich, dass sich die kommunizierenden Instanzen in definierten Zuständen befinden, um die Konsistenz bestimmter Protokollabläufe sicherzustellen.

- **Beispiele:**
  - Verbindungsaufbau
    - gewährleisten, dass beide Seiten über den Aufbau der Verbindung informiert sind und von gleichen Anfangswerten bei den Parameter-einstellungen (z. B. Sequenznummer, Fenstergröße) ausgehen
  - Verbindungsabbau
    - Verbindung muss wirklich auf beiden Seiten abgebaut sein
    - nicht zufällig einseitig !!!
- **Problem der Synchronisation**
  - nur über den Austausch von Nachrichten möglich
    - können verloren gehen !!!

Handshake-Verfahren

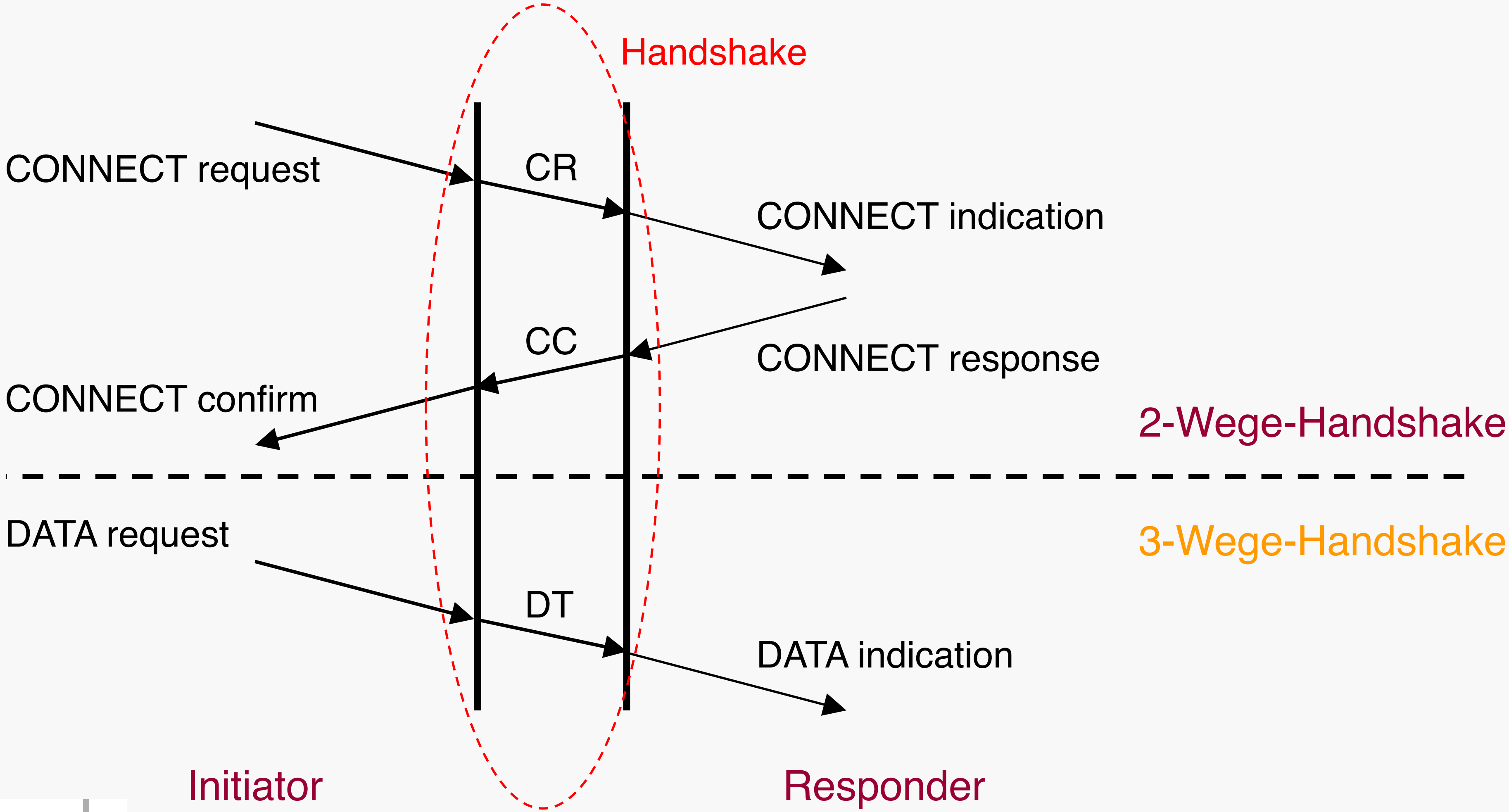
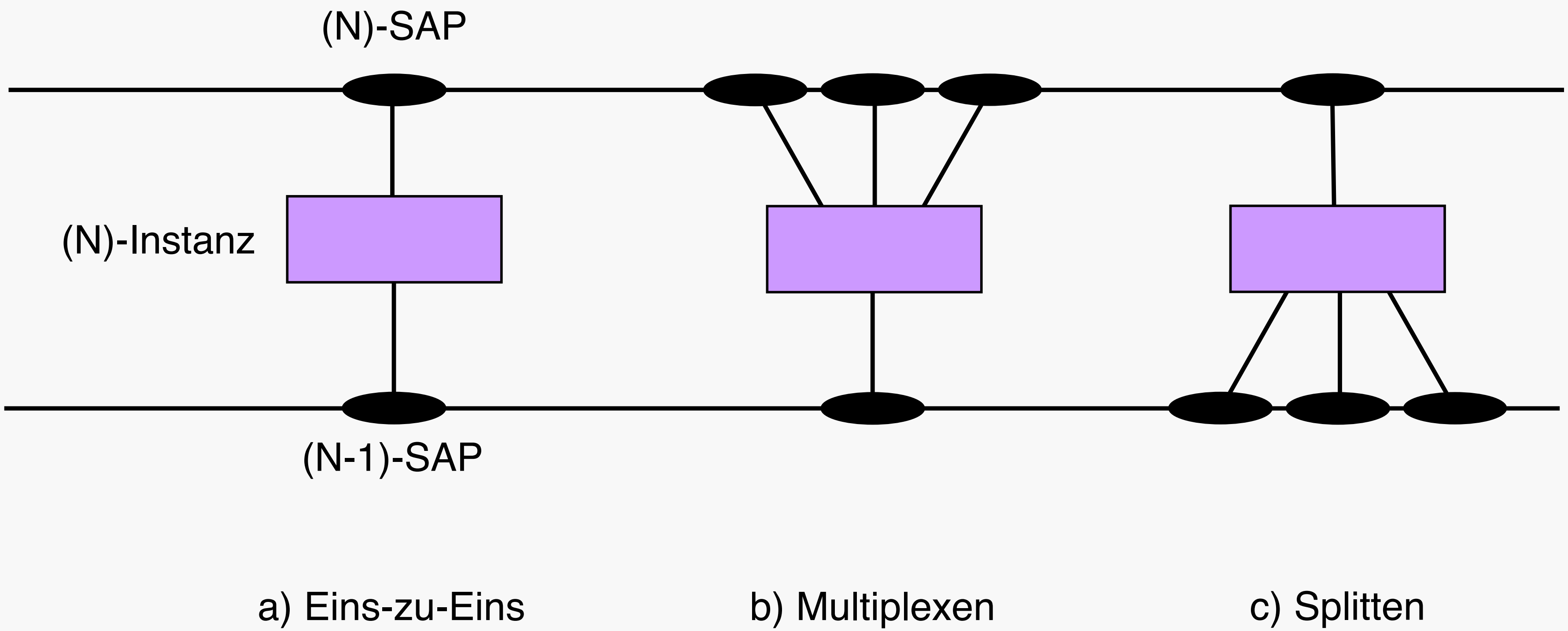


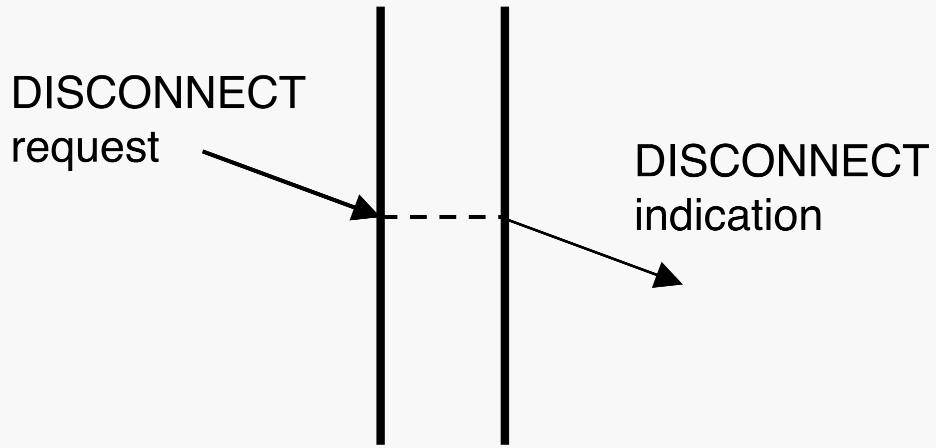
Abbildung von N- auf (N-1)-Verbindungen

Eine weitere Aktivität des Verbindungsmanagements ist die Abbildung der Verbindung auf die (N-1)-Schicht.

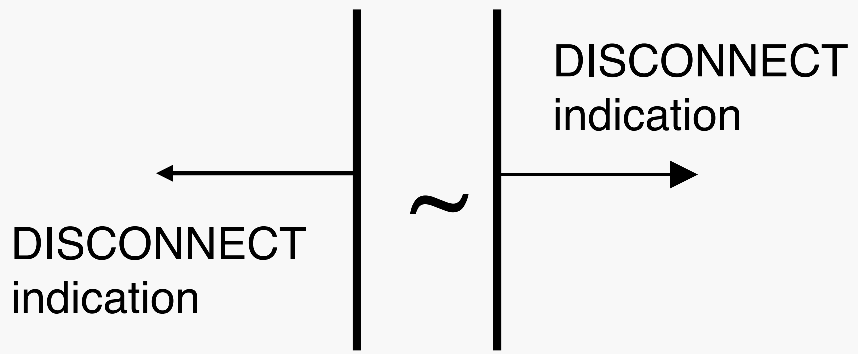


Verbindungsabbau

- Ein Verbindungsabbau wird in der Regel durch einen Dienstnutzer ausgelöst, wenn die Datenübertragung beendet ist oder er einen Fehler in der Übertragung festgestellt hat.



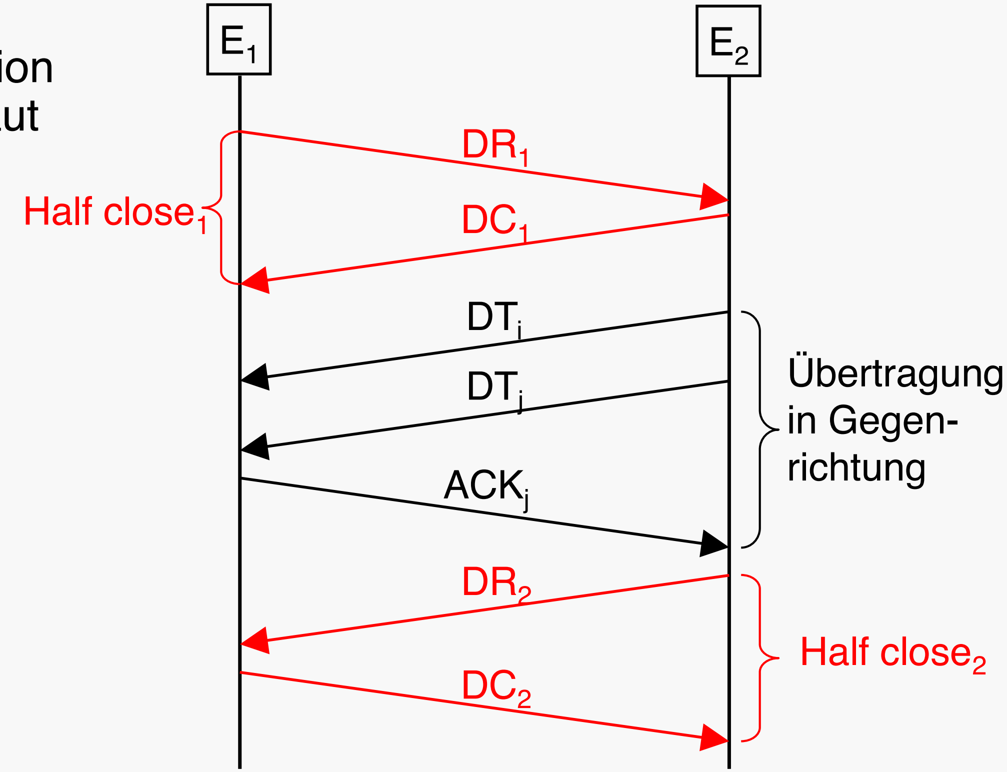
- Diensterbringer kann ebenfalls die Verbindung beenden, wenn er die Übertagung nicht mehr gewährleisten kann.



Verbindungsabbau bei Duplexkommunikation

- Verbindungen mit Duplexkommunikation müssen in beiden Richtungen abgebaut werden

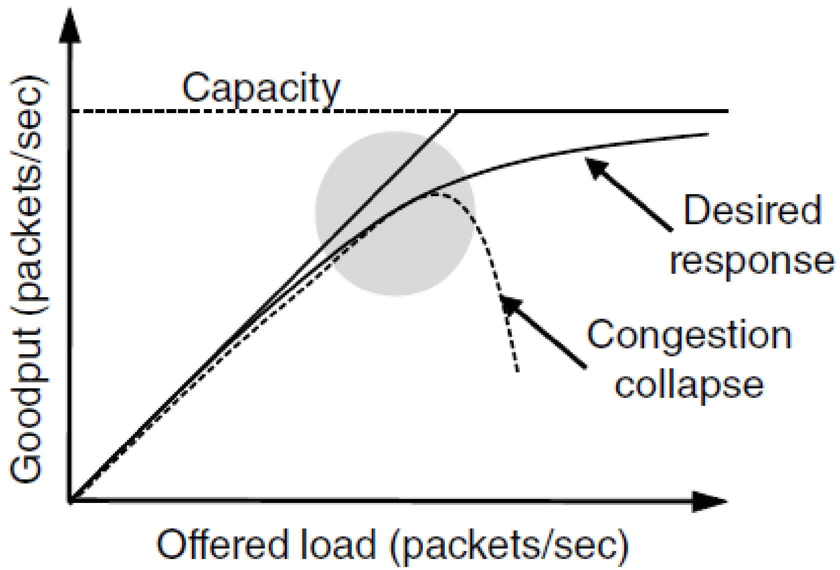
- Lösungsvarianten:
  - separater Abbau pro Richtung
    - Half close
    - Datenübertragung in andere Richtung kann weiterlaufen
  - gleichzeitiger Abbau möglich
    - zwei 2-Wege-Handshakes oder
    - ein 3-Wege-Handshake

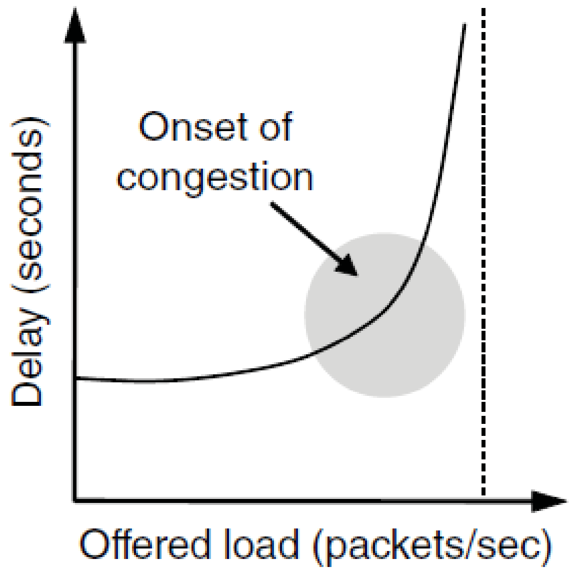


DR – DISCONNECT request PDU
DC – DISCONNECT confirm PDU

Desirable Bandwidth Allocation (1)

Efficient use of bandwidth gives high goodput, low delay





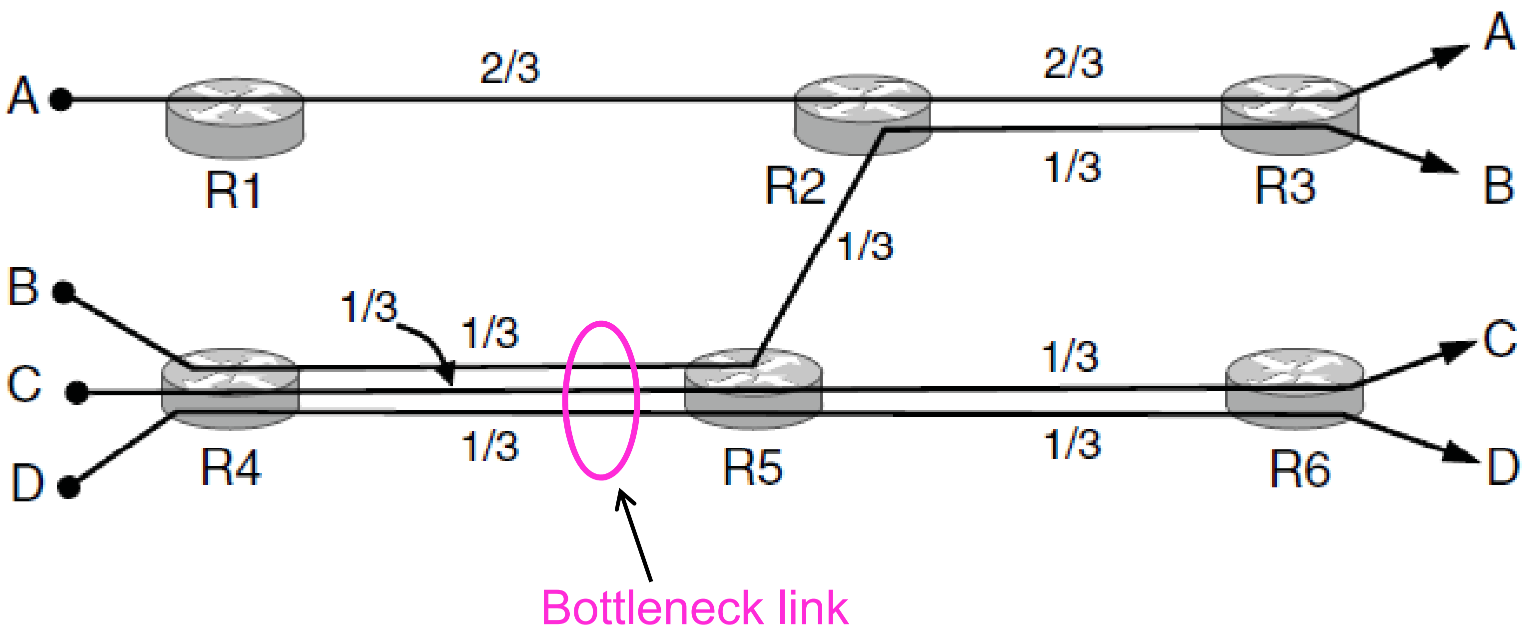
Goodput rises more slowly than load when congestion sets in

Delay begins to rise sharply when congestion sets in

Desirable Bandwidth Allocation (2)

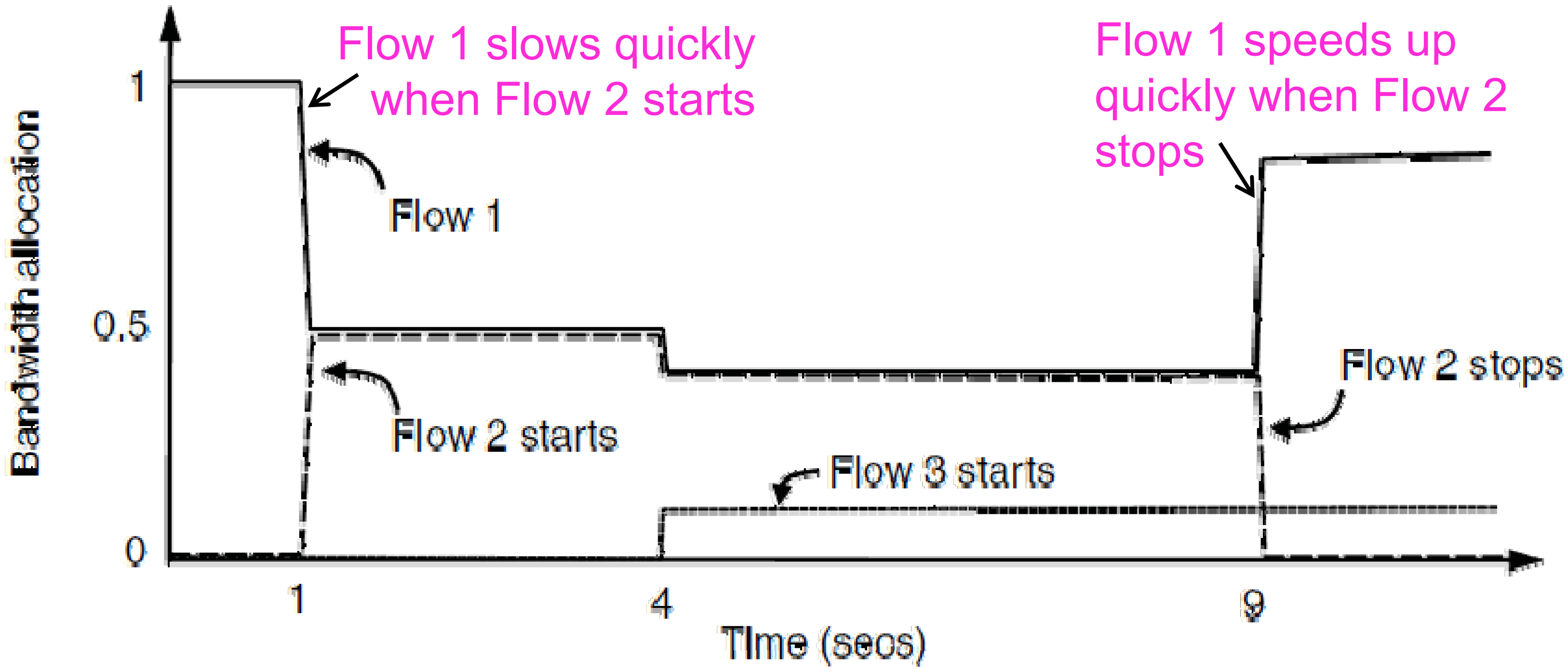
Fair use gives bandwidth to all flows (no starvation)

- Max-min fairness gives equal shares of bottleneck



Desirable Bandwidth Allocation (3)

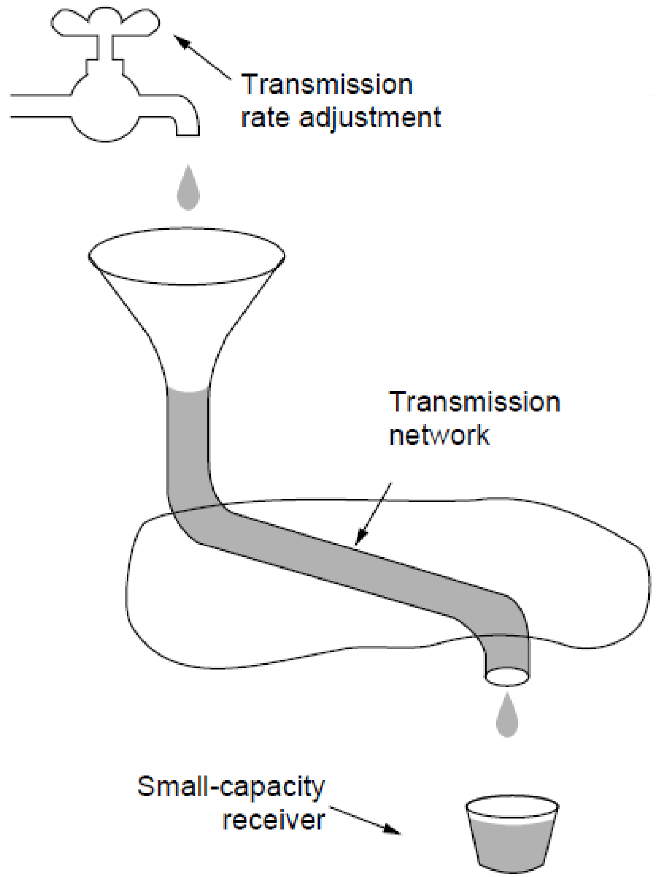
We want bandwidth levels to converge quickly when traffic patterns change



Regulating the Sending Rate (1)

Sender may need to slow down for different reasons:

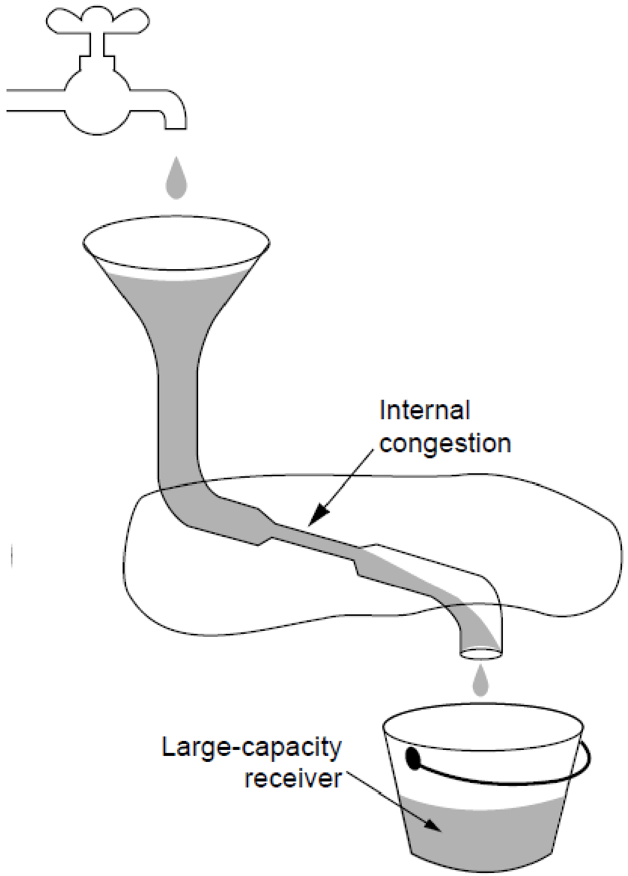
- Flow control, when the receiver is not fast enough [right]
- Congestion, when the network is not fast enough [over]



A fast network feeding a low-capacity receiver
→ flow control is needed

Regulating the Sending Rate (2)

Our focus is dealing with this problem – congestion



A slow network feeding a high-capacity receiver
→ congestion control is needed

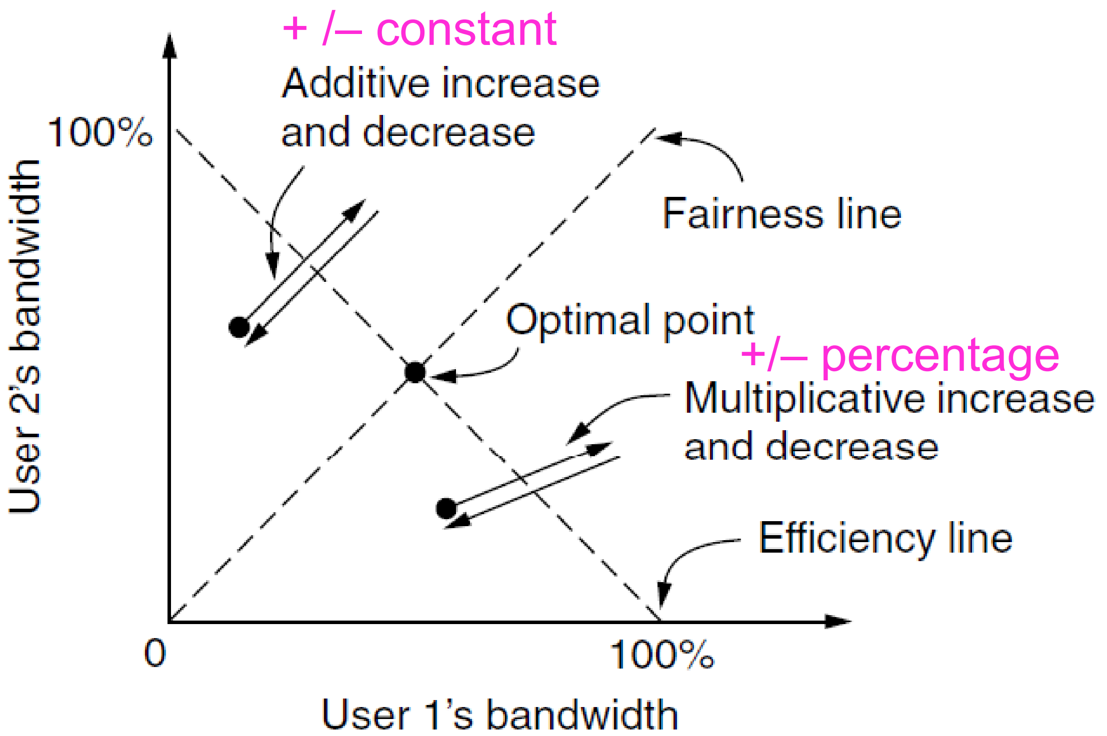
Regulating the Sending Rate (3)

Different congestion signals the network may use to tell the transport endpoint to slow down (or speed up)

Protocol	Signal	Explicit?	Precise?
XCP	Rate to use	Yes	Yes
TCP with ECN	Congestion warning	Yes	No
FAST TCP	End-to-end delay	No	Yes
CUBIC TCP	Packet loss	No	No
TCP	Packet loss	No	No

Regulating the Sending Rate (3)

If two flows increase/decrease their bandwidth in the same way when the network signals free/busy they will not converge to a fair allocation

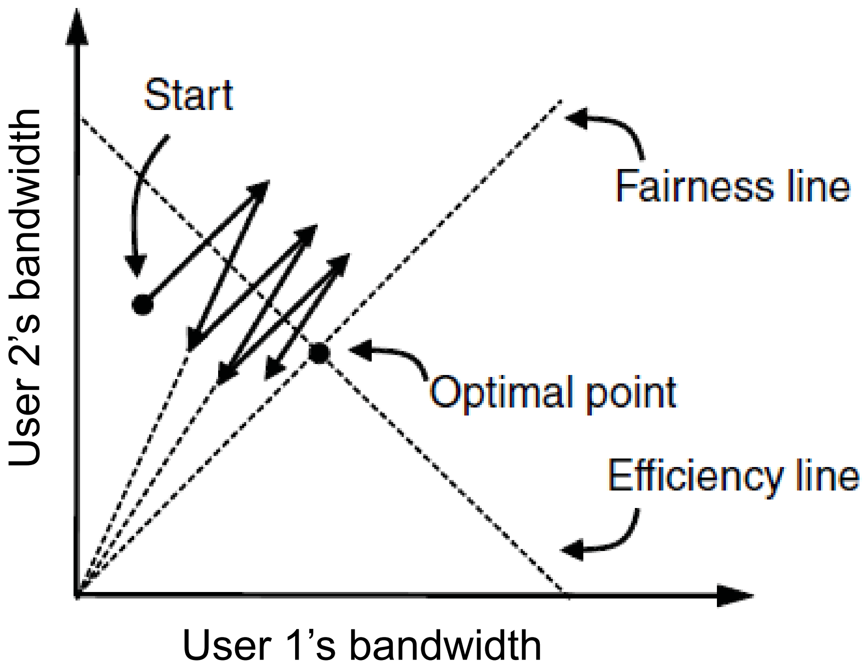


CN5E by Tanenbaum & Wetherall, © Pearson Education-Prentice Hall and D. Wetherall, 2011

Regulating the Sending Rate (4)

The AIMD (Additive Increase Multiplicative Decrease) control law does converge to a fair and efficient point!

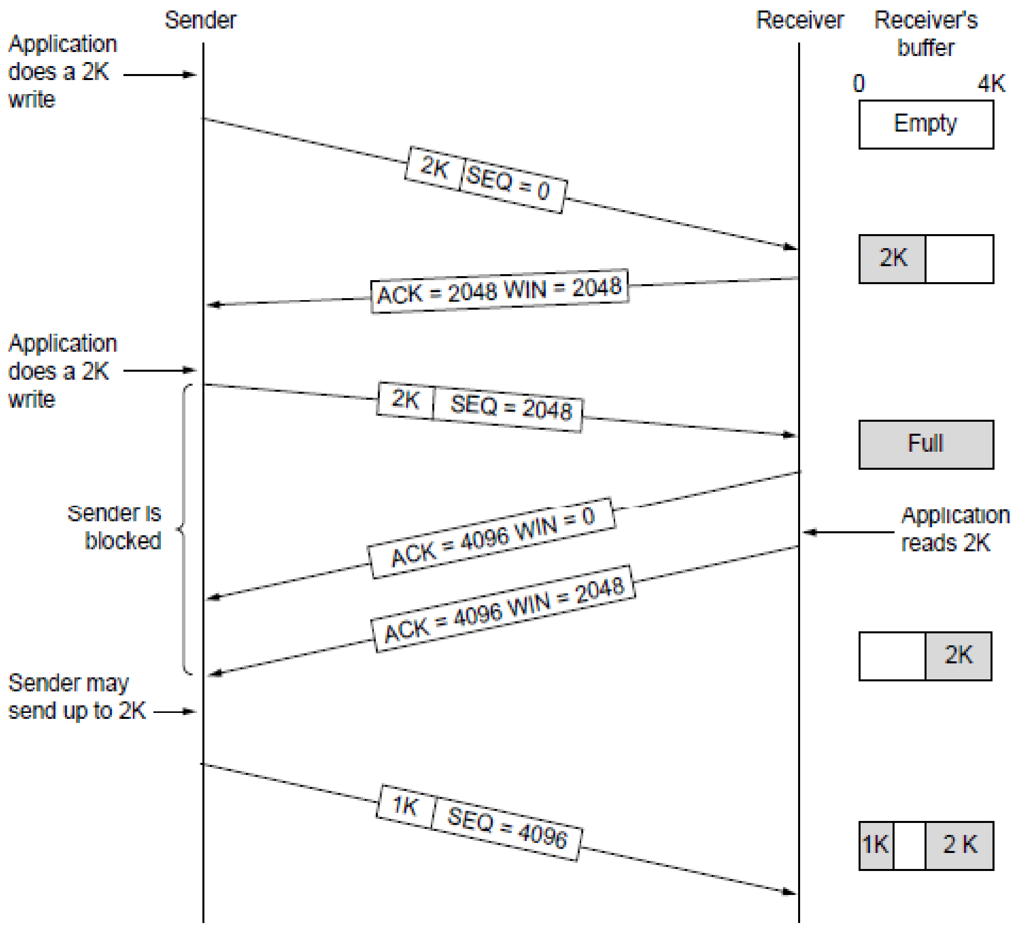
- TCP uses AIMD for this reason



TCP Sliding Window (1)

TCP adds flow control to the sliding window as before

- ACK + WIN is the sender's limit

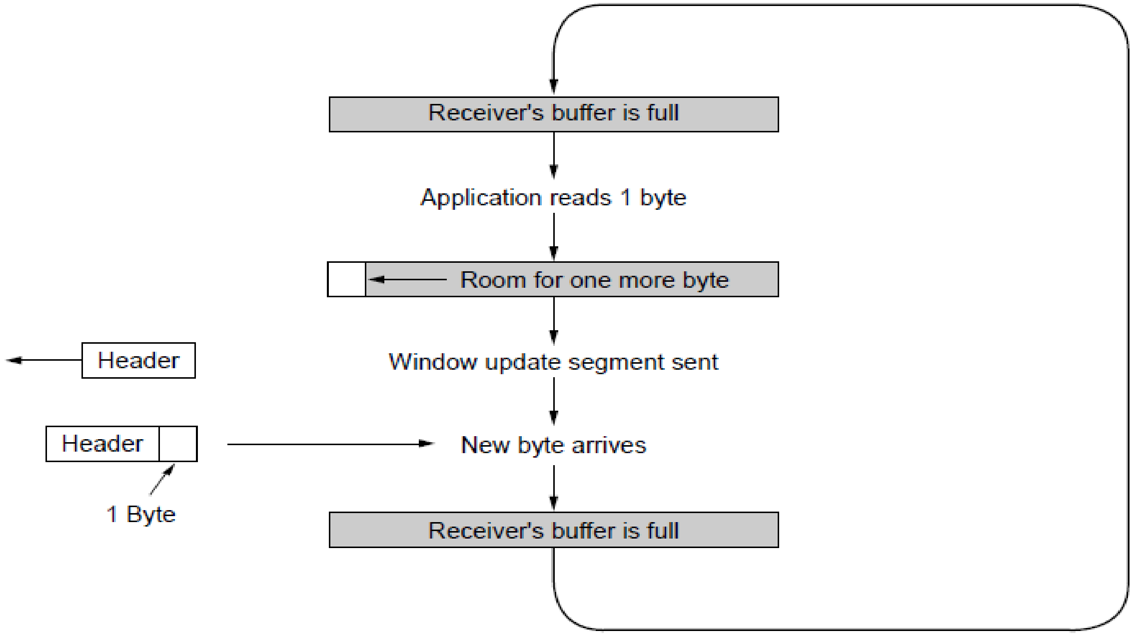


CN5E by Tanenbaum & Wetherall, © Pearson Education-Prentice Hall and D. Wetherall, 2011

TCP Sliding Window (2)

Need to add special cases to avoid unwanted behavior

- E.g., silly window syndrome [below]

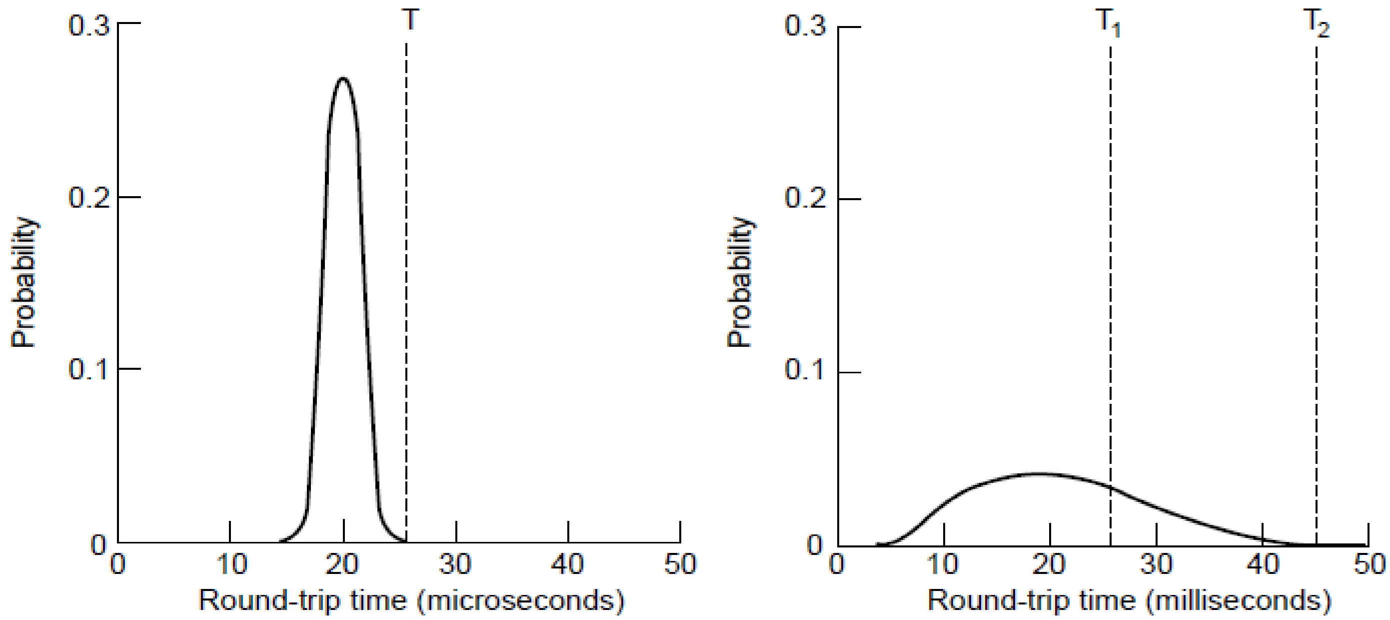


Receiver application reads single bytes, so sender always sends one byte segments

TCP Timer Management

TCP estimates retransmit timer from segment RTTs

- Tracks both average and variance (for Internet case)
- Timeout is set to average plus 4 x variance



LAN case – small, regular RTT

Internet case – large, varied RTT

TCP Congestion Control (1)

TCP uses AIMD with loss signal to control congestion

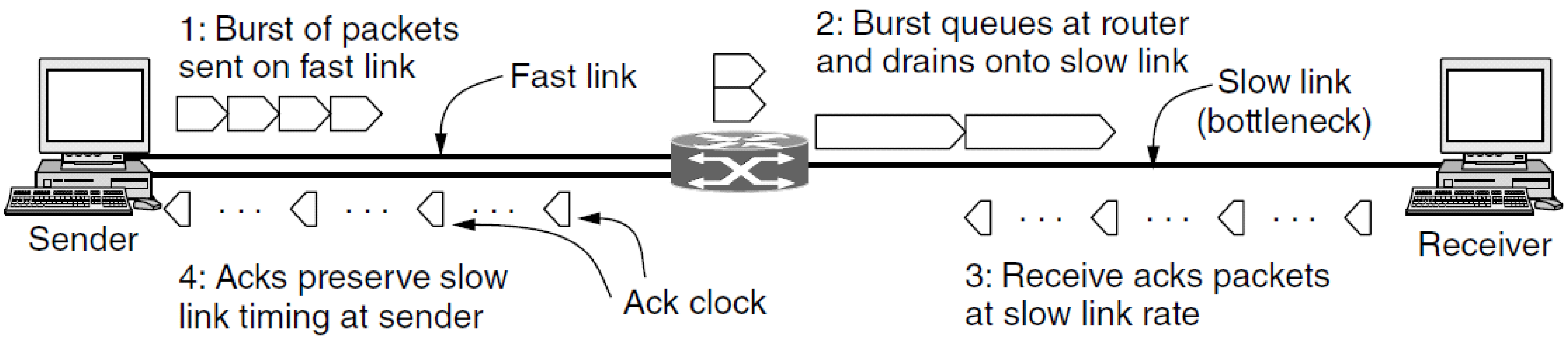
- Implemented as a congestion window (cwnd) for the number of segments that may be in the network
- Uses several mechanisms that work together

Name	Mechanism	Purpose
ACK clock	Congestion window (cwnd)	Smooth out packet bursts
Slow-start	Double cwnd each RTT	Rapidly increase send rate to reach roughly the right level
Additive Increase	Increase cwnd by 1 packet each RTT	Slowly increase send rate to probe at about the right level
Fast retransmit / recovery	Resend lost packet after 3 duplicate ACKs; send new packet for each new ACK	Recover from a lost packet without stopping ACK clock

TCP Congestion Control (2)

Congestion window controls the sending rate

- Rate is cwnd / RTT; window can stop sender quickly
- <u>ACK clock</u> (regular receipt of ACKs) paces traffic and smoothes out sender bursts

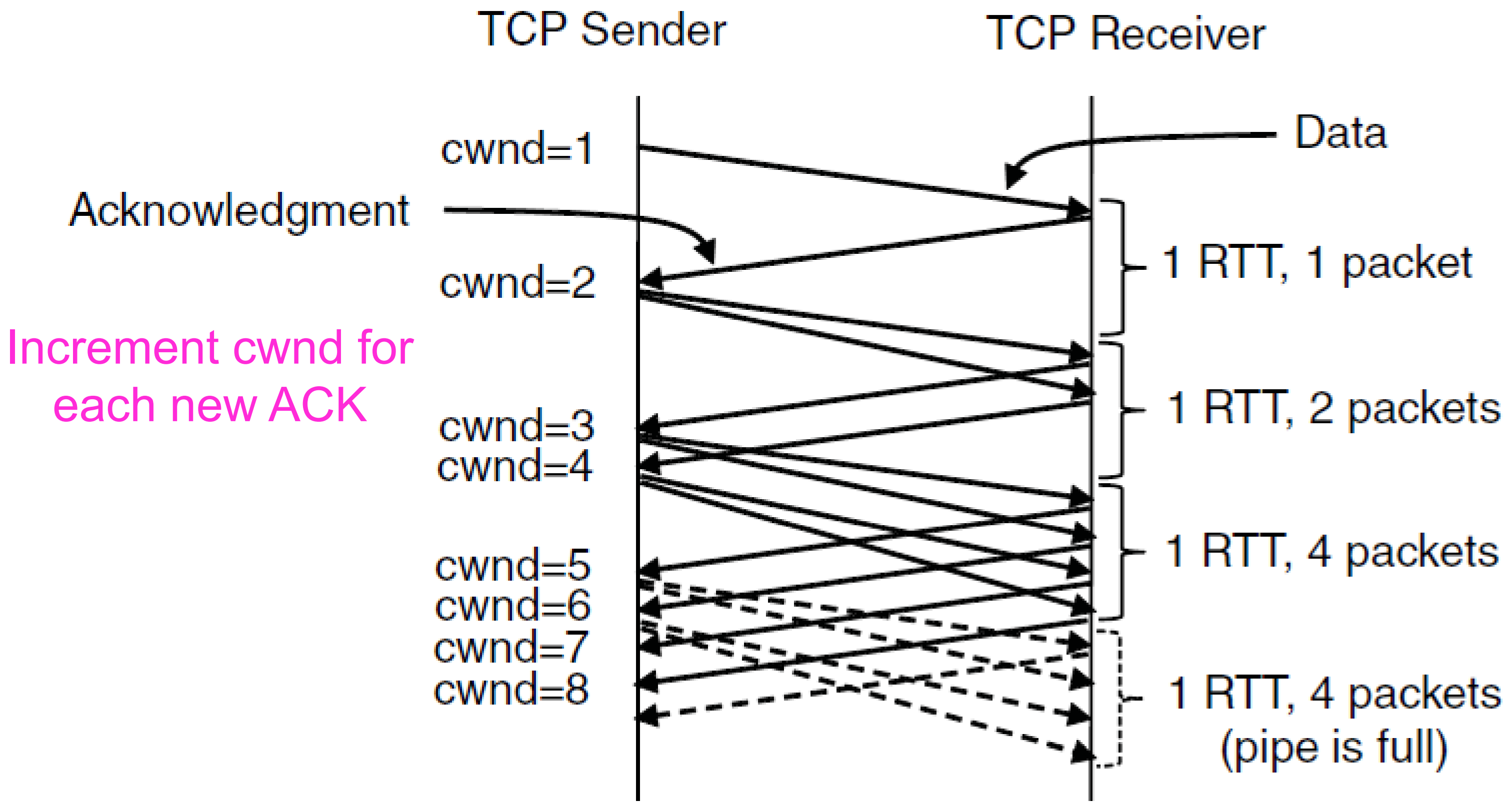


ACKs pace new segments into the network and smooth bursts

TCP Congestion Control (3)

Slow start grows congestion window exponentially

- Doubles every RTT while keeping ACK clock going

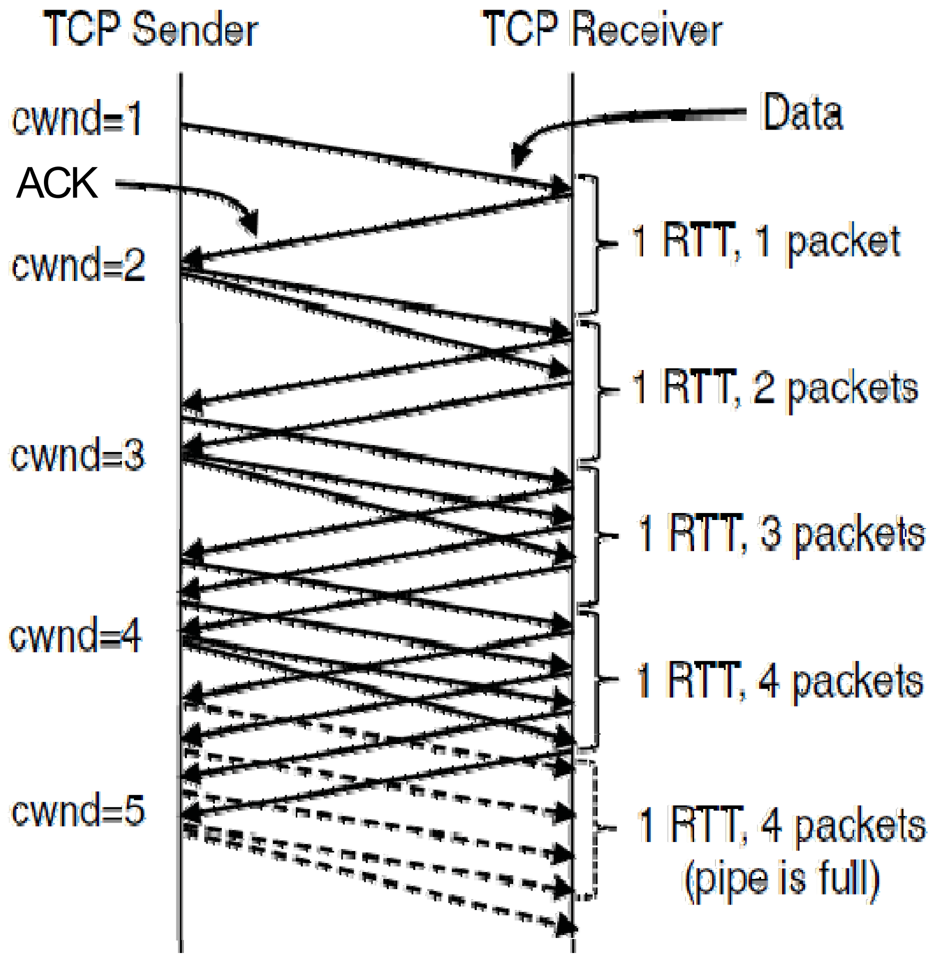


CN5E by Tanenbaum & Wetherall, © Pearson Education-Prentice Hall and D. Wetherall, 2011

TCP Congestion Control (4)

Additive increase grows cwnd slowly

- Adds 1 every RTT
- Keeps ACK clock



CN5E by Tanenbaum & Wetherall, © Pearson Education-Prentice Hall and D. Wetherall, 2011

Quality of Service

- Application requirements »
- Traffic shaping »
- Packet scheduling »
- Admission control »
- Integrated services »
- Differentiated services »

Application Requirements (1)

Different applications care about different properties

- We want all applications to get what they need

Application	Bandwidth	Delay	Jitter	Loss
Email	Low	Low	Low	Medium
File sharing	High	Low	Low	Medium
Web access	Medium	Medium	Low	Medium
Remote login	Low	Medium	Medium	Medium
Audio on demand	Low	Low	High	Low
Video on demand	High	Low	High	Low
Telephony	Low	High	High	Low
Videoconferencing	High	High	High	Low

"High" means a demanding requirement, e.g., low delay

Application Requirements (2)

Network provides service with different kinds of QoS (Quality of Service) to meet application requirements

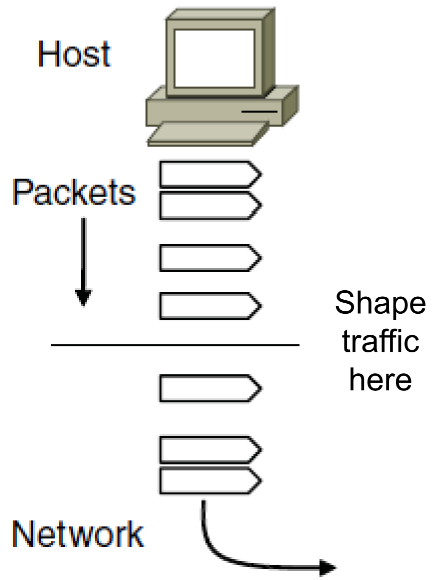
Network Service	**Application**
Constant bit rate	Telephony
Real-time variable bit rate	Videoconferencing
Non-real-time variable bit rate	Streaming a movie
Available bit rate	File transfer

Example of QoS categories from ATM networks

Traffic Shaping (1)

Traffic shaping regulates the average rate and burstiness of data entering the network

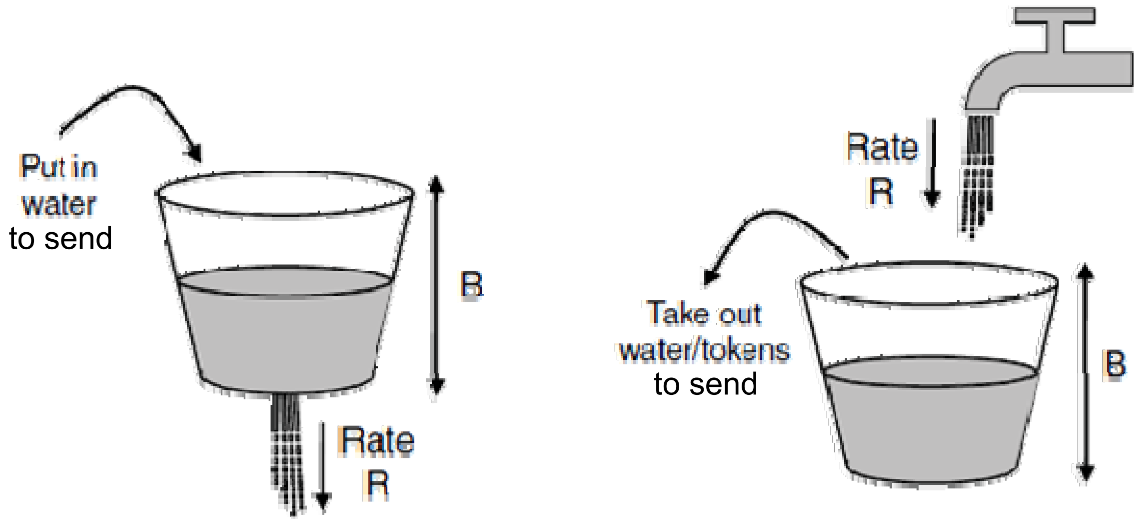
- Lets us make guarantees



Traffic Shaping (2)

Token/Leaky bucket limits both the average rate (R) and short-term burst (B) of traffic

- For token, bucket size is B, water enters at rate R and is removed to send; opposite for leaky.

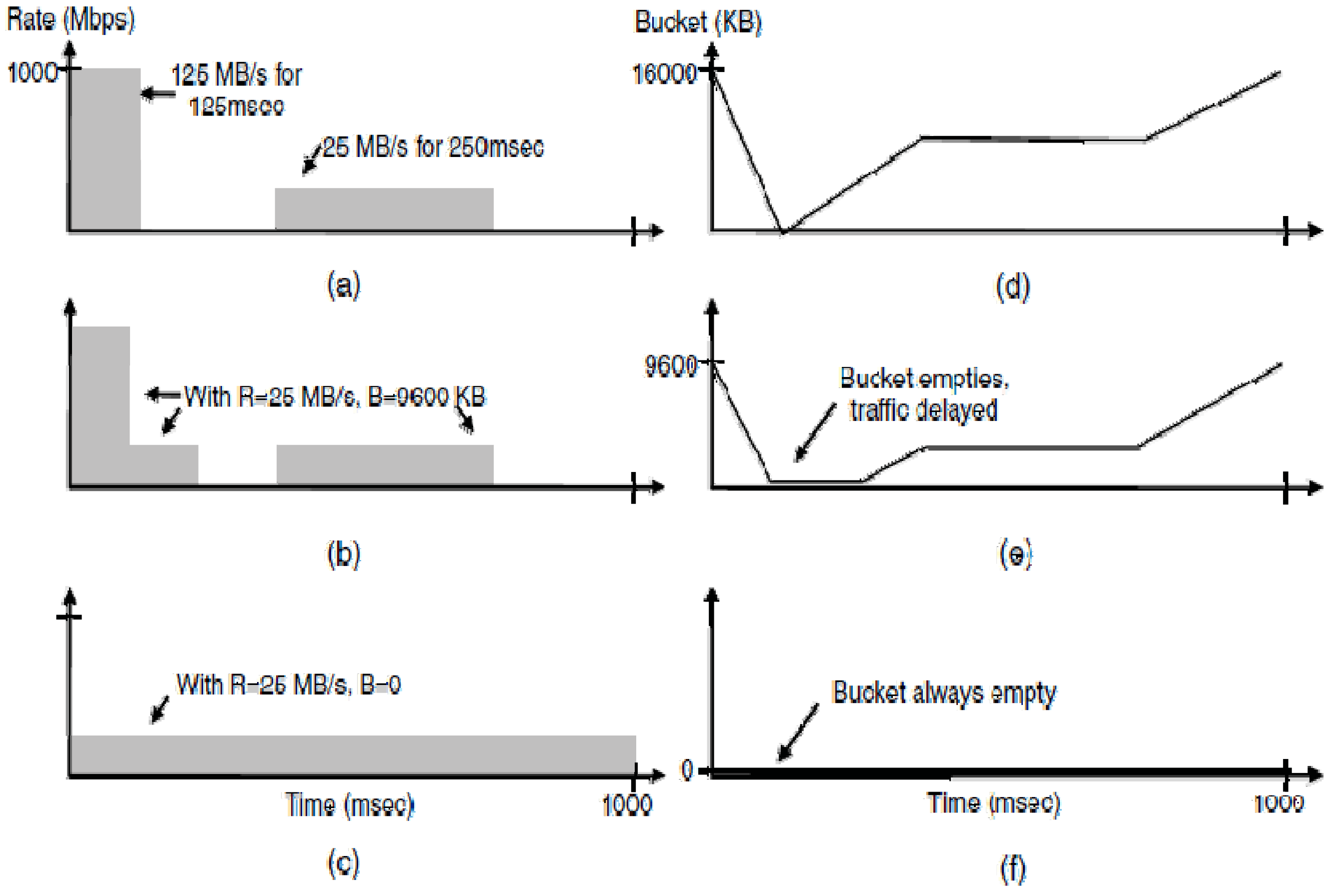


Leaky bucket
(need not full to send)

Token bucket
(need some water to send)

Traffic Shaping (3)

Host traffic
R=200 Mbps
B=16000 KB
↓
Shaped by
R=200 Mbps
B=9600 KB
↓
Shaped by
R=200 Mbps
B=0 KB

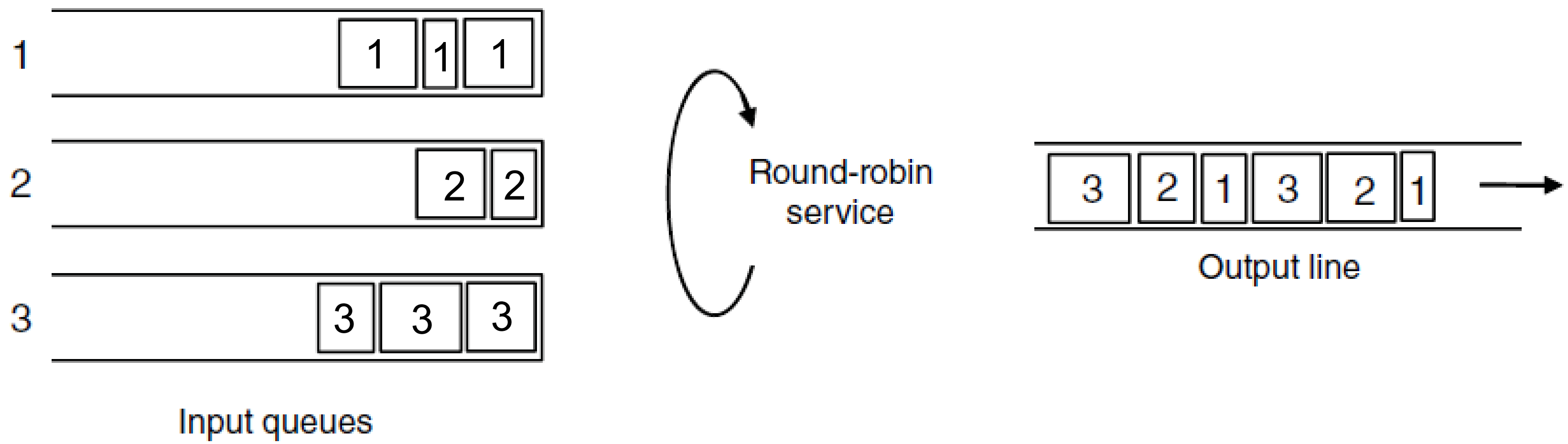


Smaller bucket size delays traffic and reduces burstiness

CN5E by Tanenbaum & Wetherall, © Pearson Education-Prentice Hall and D. Wetherall, 2011

Packet Scheduling (1)

Packet scheduling divides router/link resources among traffic flows with alternatives to FIFO (First In First Out)



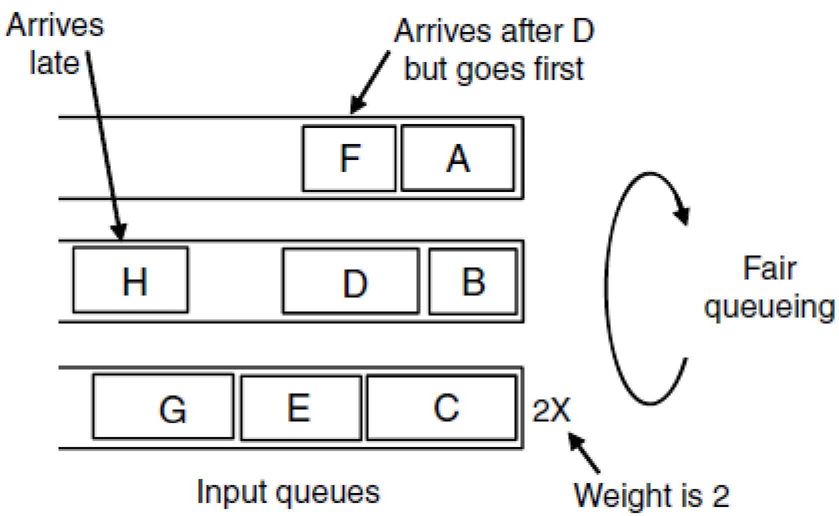
Example of round-robin queuing

CN5E by Tanenbaum & Wetherall, © Pearson Education-Prentice Hall and D. Wetherall, 2011

Packet Scheduling (2)

Fair Queueing approximates bit-level fairness with different packet sizes; weights change target levels

- Result is WFQ (Weighted Fair Queueing)



Packet	Arrival time	Length	Finish time	Output order
A	0	8	8	1
B	5	6	11	3
C	5	10	10	2
D	8	9	20	7
E	8	8	14	4
F	10	6	16	5
G	11	10	19	6
H	20	8	28	8

$$F_i = \max(A_i, F_{i-1}) + L_i/W$$

Packets may be sent out of arrival order

Finish virtual times determine transmission order

Admission Control (1)

Admission control takes a traffic flow specification and decides whether the network can carry it

- Sets up packet scheduling to meet QoS

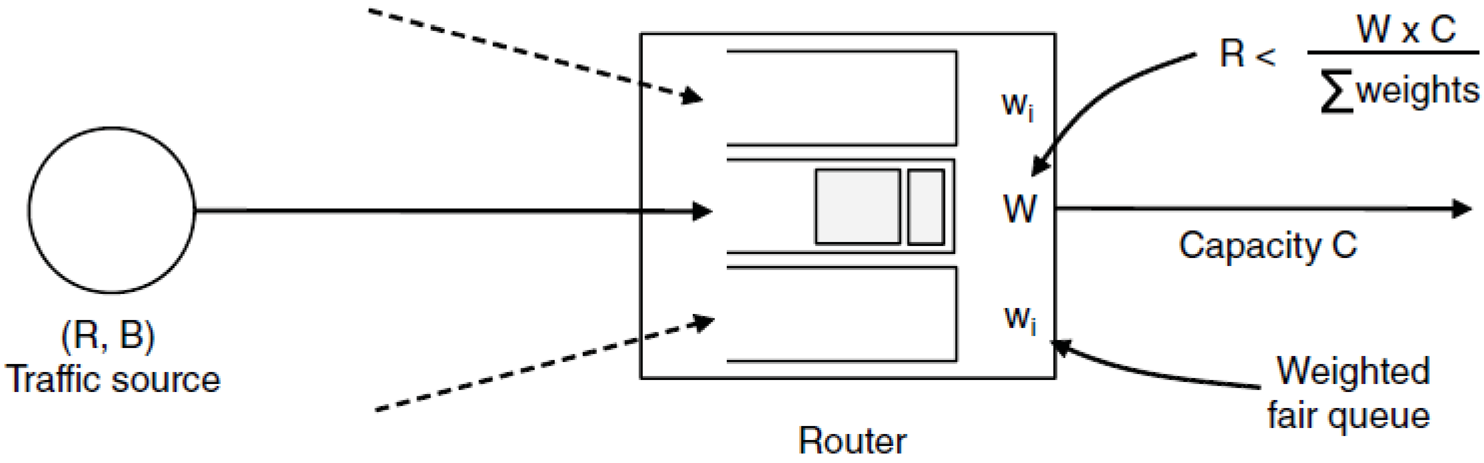
Parameter	Unit
Token bucket rate	Bytes/sec
Token bucket size	Bytes
Peak data rate	Bytes/sec
Minimum packet size	Bytes
Maximum packet size	Bytes

Example flow specification

Admission Control (2)

Construction to guarantee bandwidth B and delay D:

- Shape traffic source to a (R, B) token bucket
- Run WFQ with weight W / all weights > R/capacity
- Holds for all traffic patterns, all topologies



CN5E by Tanenbaum & Wetherall, © Pearson Education-Prentice Hall and D. Wetherall, 2011

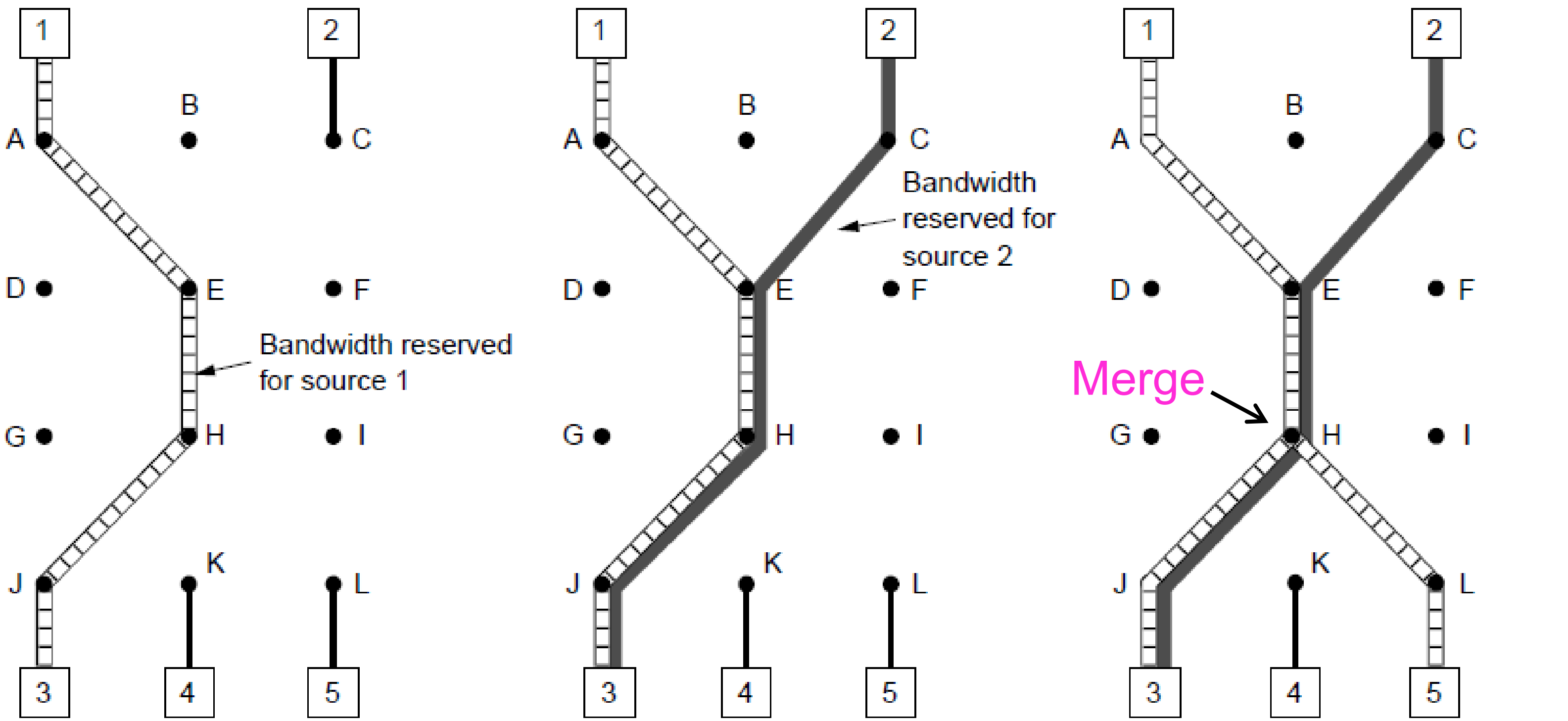
Integrated Services (1)

Design with QoS for each flow; handles multicast traffic.

Admission with RSVP (Resource reSerVation Protocol):

- Receiver sends a request back to the sender
- Each router along the way reserves resources
- Routers merge multiple requests for same flow
- Entire path is set up, or reservation not made

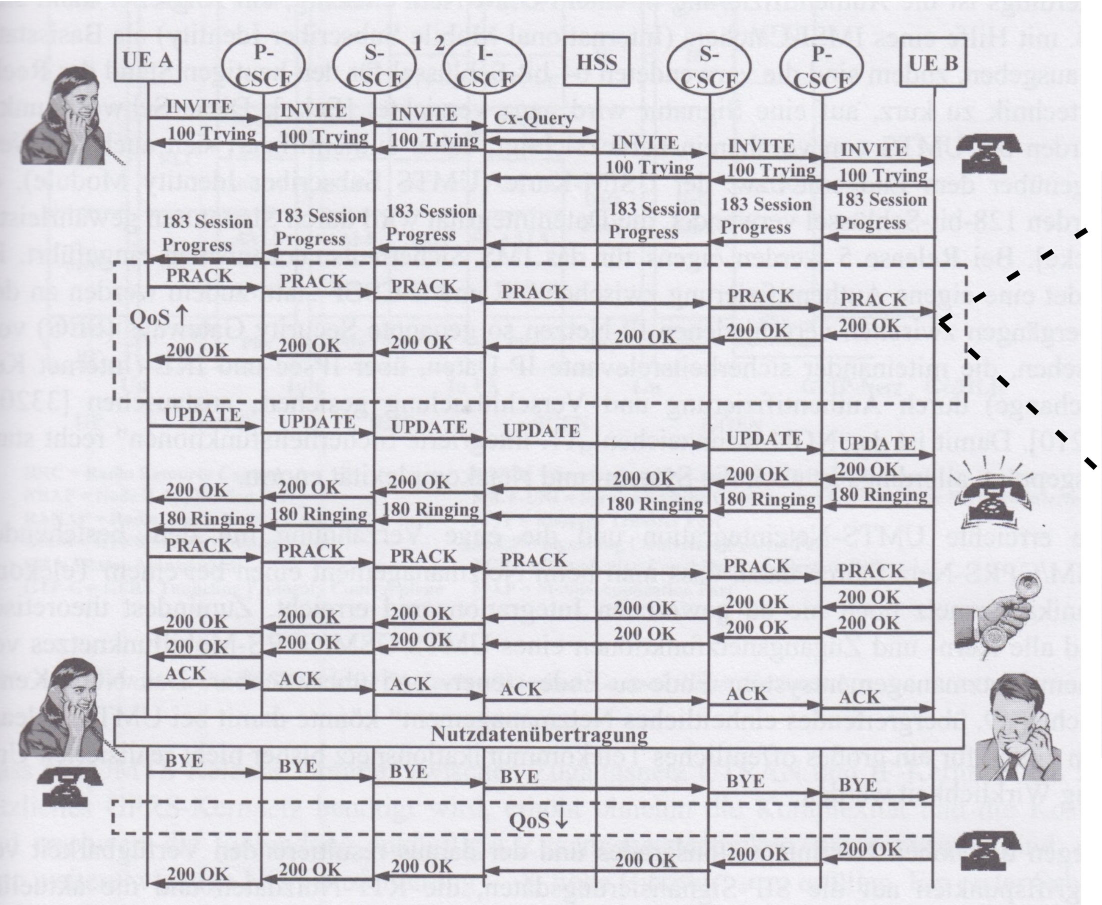
Integrated Services (2)

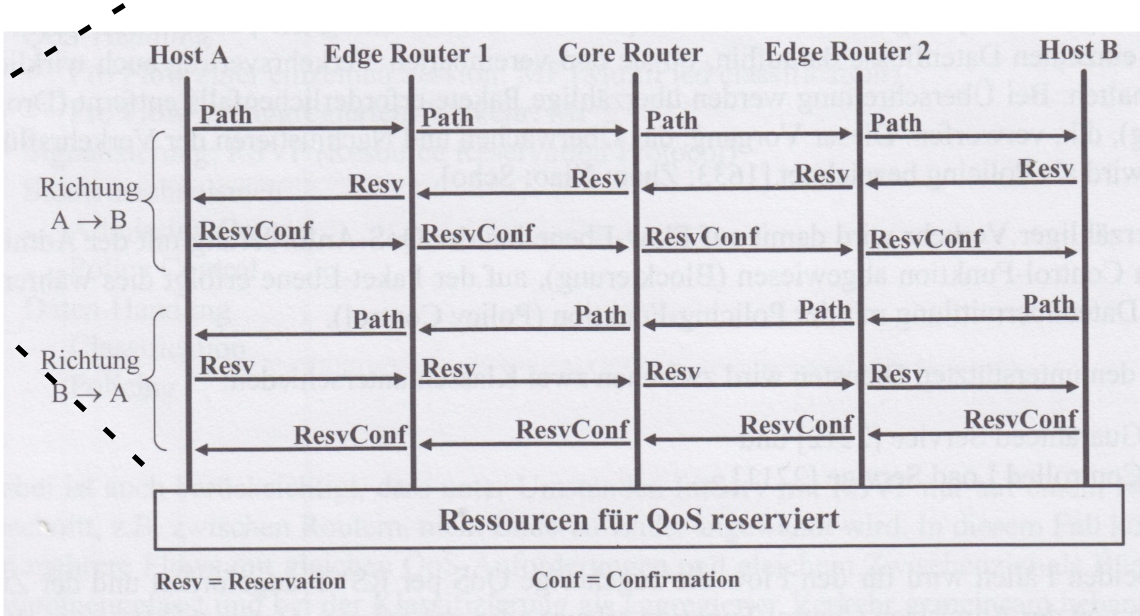


R3 reserves flow from S1

R3 reserves flow from S2

R5 reserves flow from S1; merged with R3 at H



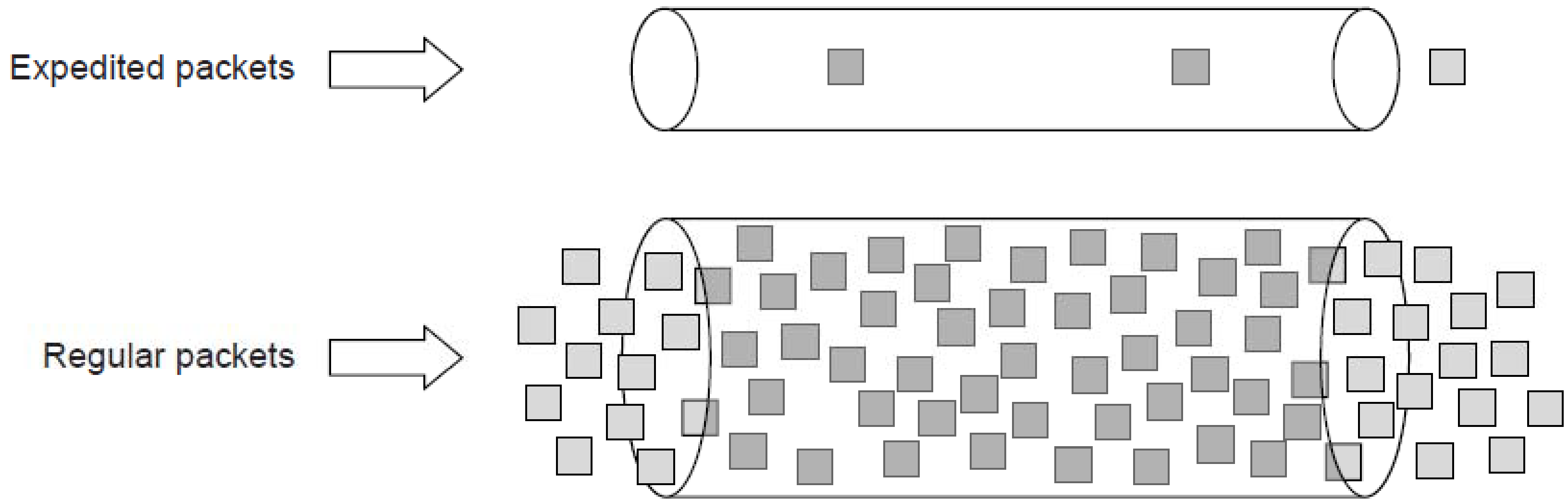


Use of RSVP for IntServ QoS Setup in VoIP (IMS, UMTS Release 5)

Differentiated Services (1)

Design with classes of QoS; customers buy what they want

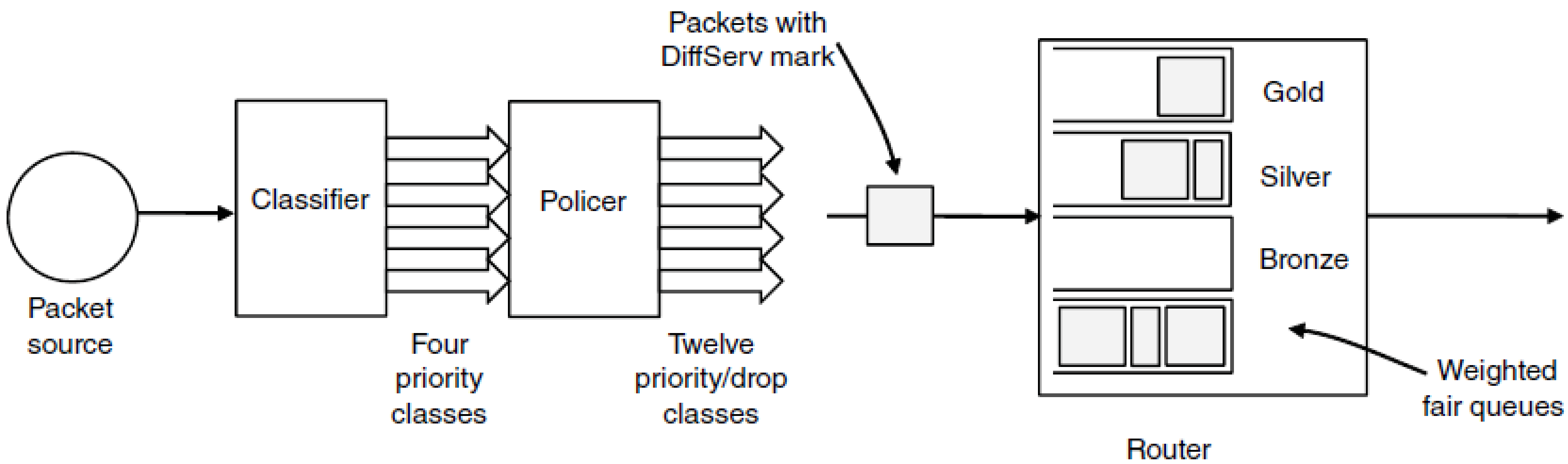
- Expedited class is sent in preference to regular class
- Less expedited traffic but better quality for applications

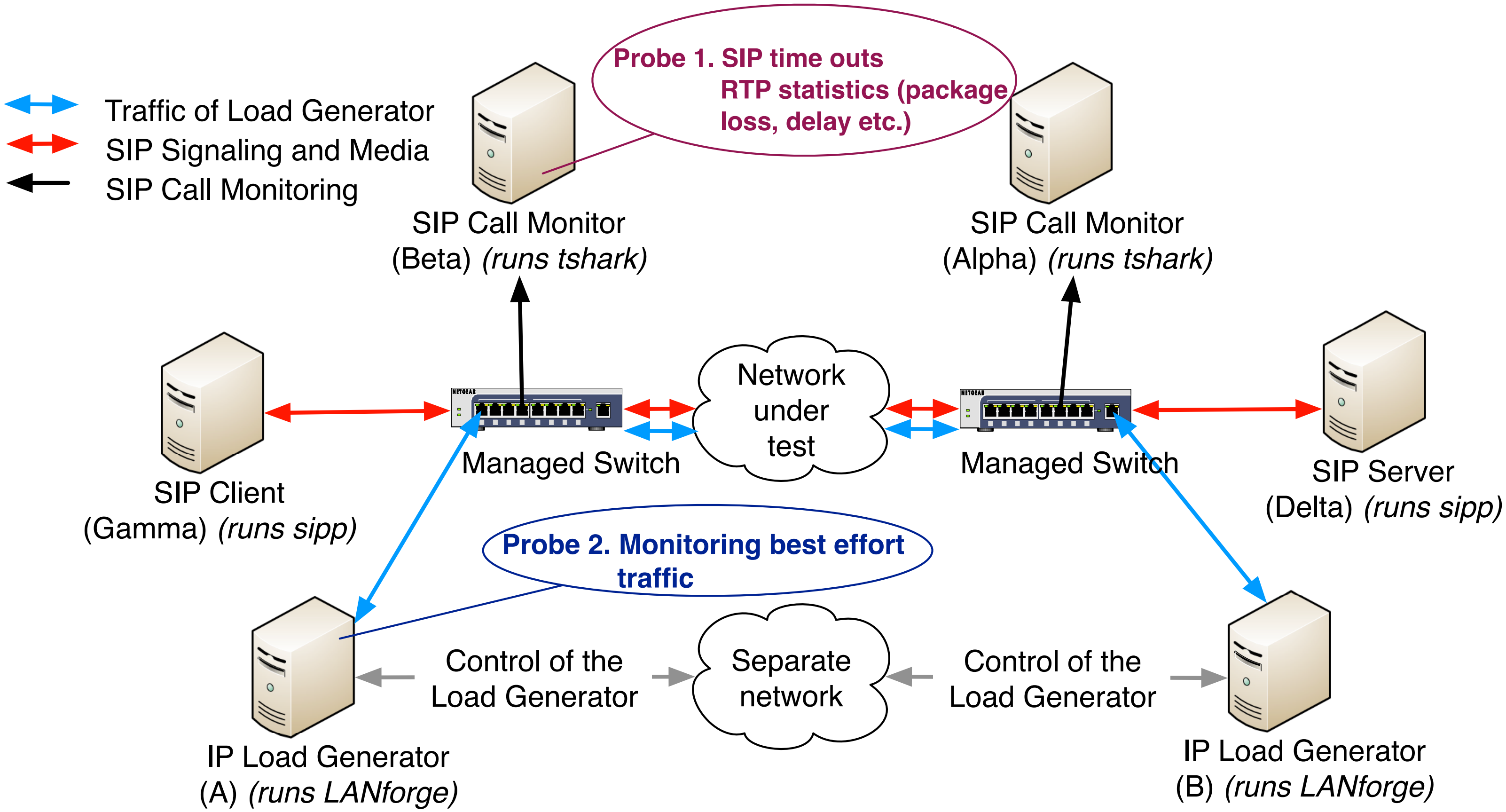


Differentiated Services (2)

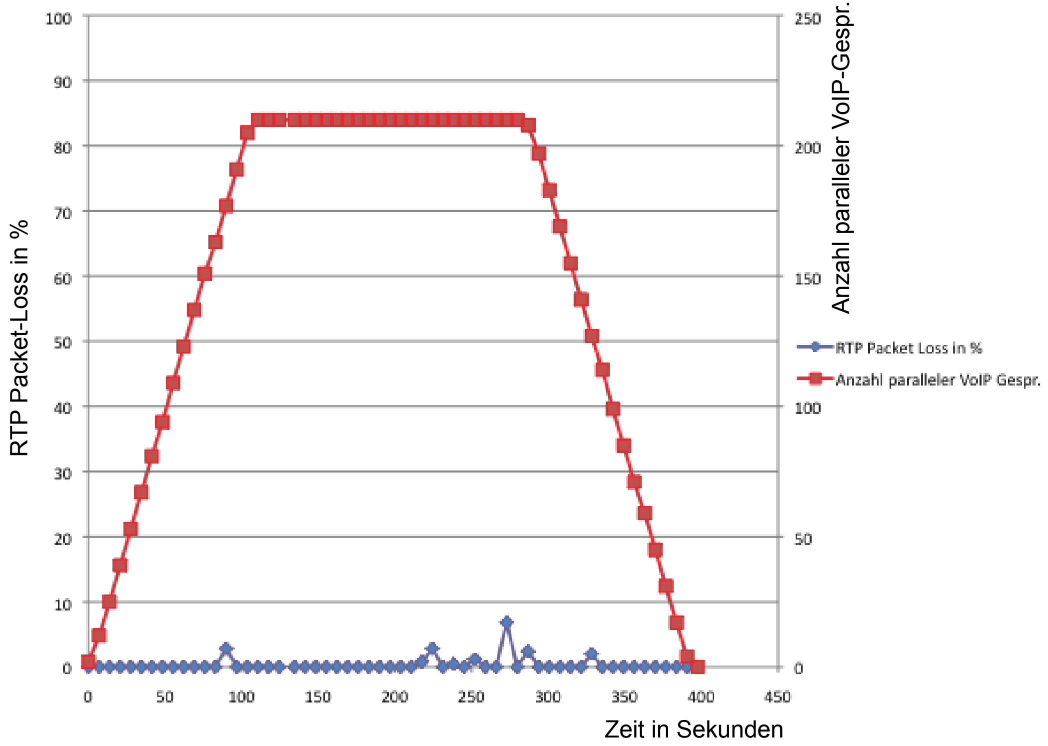
Implementation of DiffServ:

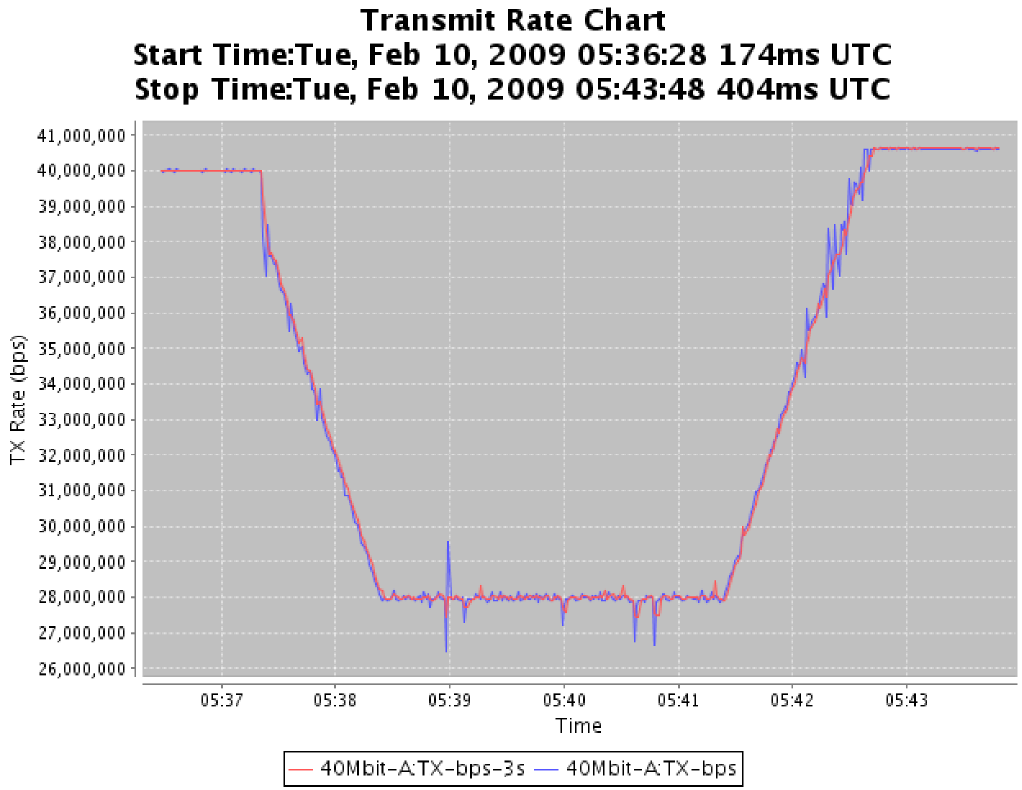
- Customers mark desired class on packet
- ISP shapes traffic to ensure markings are paid for
- Routers use WFQ to give different service levels





Configuration Measuring QoS-Relevant Behaviour of Network under Test





Results of QoS Measurement

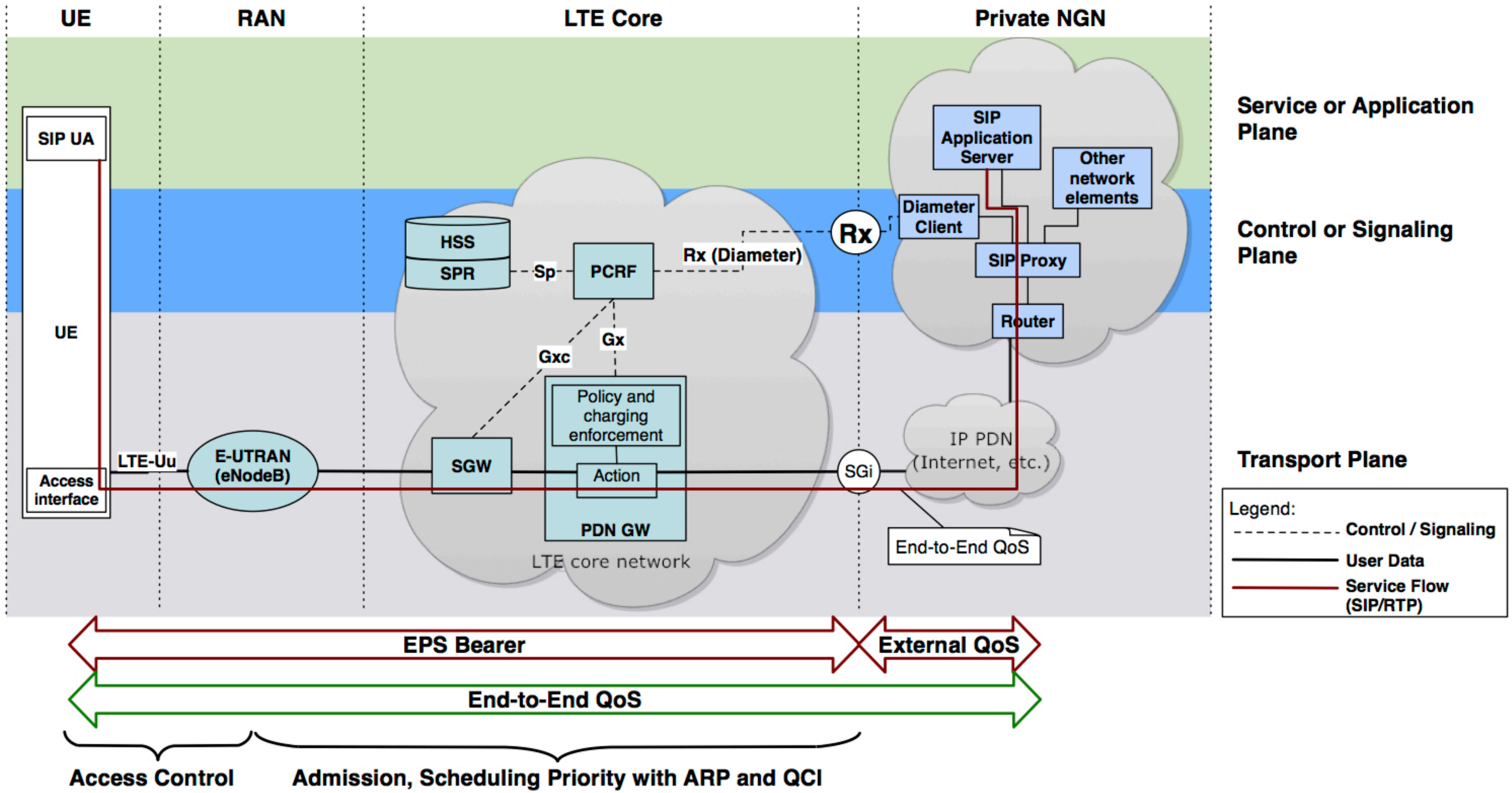
Access Class	Description
0 – 9	Randomly allocated to commercial subscribers
10	Reserved for emergency access to the system (911, 112 etc.)
11	Reserved for the MNO
12	Security services
13	Public Utilities
14	Emergency services
15	Reserved for the MNO

Example for Precedence: Access Classes defined for LTE Air-Interface

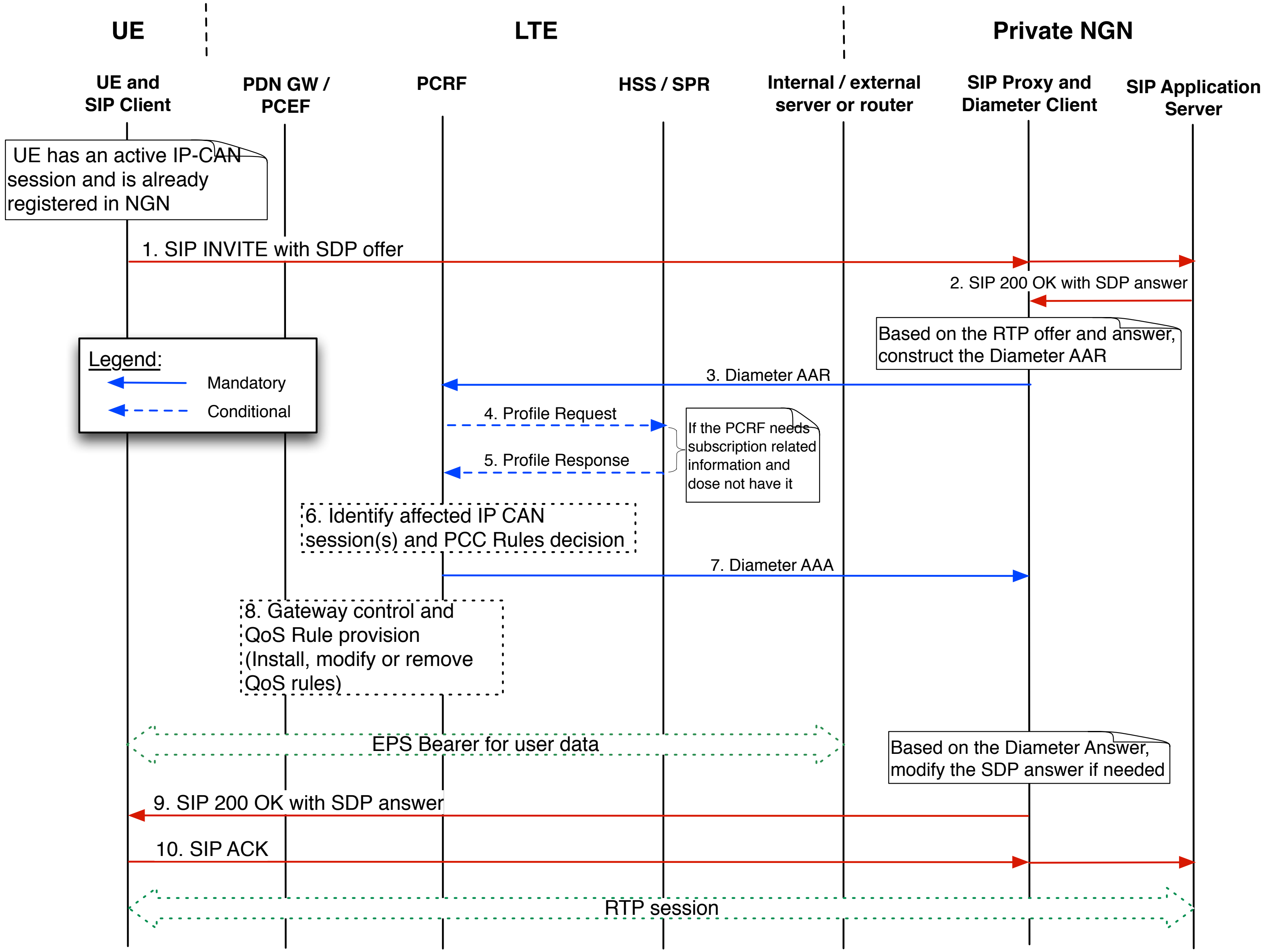
Priority (ARP)	Priority class	GBR Bearers		Non-GBR Bearers	
		Can pre-empt	Can be pre-empted	Can pre-empt	Can be pre-empted
1	Emergency	Yes	No	Yes	No
2-5	High	Yes	No	Yes	No
6	High	No	Yes	No	No
7-11	Medium	No	Yes	No	No
12-15	Low	No	Yes	No	Yes

- Allocation Retention Class (ARP)
- Pre-emption is performed based on ARP value
- ARP is stored in HSS

Example for Pre-emption: Admission Control in LTE



Example: Implementation of MLPP in LTE



Example: Implementation of MLPP in LTE