

Kapitel 8

Medien- und Echtzeitkommunikation

8.1	Motivation und Grundbegriffe	2
8.2	Anwendungen mit Bedarf an Medien- und Echtzeitkommunikation	6
8.3	Endbenutzerorientierte Anforderungen bei verteilten multimedialen Anwendungen	8
8.4	Dienstgüte aus Kommunikationsnetz-Sicht	11
8.5	Das “QoS-Mapping“ - Problem	15
8.6	Implikationen von Echtzeitkommunikationsanforderungen für kommunizierende Endsyste	17
8.7	Verbesserung vorhandener und Gewährleistung geforderter Dienstgüte aus Endbenutzersicht	19
8.8	Zur Beurteilung der Dienstgüte aus Endbenutzersicht	41

8. MEDIEN- UND ECHTZEIT-KOMMUNIKATION

8.1 Motivation und Grundbegriffe

**Def. (Technische) Medienkommunikation, auch :
Multimedialkommunikation :**

Austausch unterschiedlicher Informationsarten
(insbesondere Daten, Festbilder, Audio-/Videoströme)
mittels eines Kommunikationssystems/-netzes dergestalt, dass
Qualitätsanforderungen der Endbenutzer berücksichtigt
(und möglichst gut erfüllt) werden. ■

nota bene :

zahlreiche andere Def. für *Medienkommunikation* möglich
(insbesondere auch Begriffsverwendung außerhalb der Informatik).



Def. Echtzeitkommunikation :

Kommunikation (über ein technisches Kommunikationssystem) bei der
Echtzeitbedingungen berücksichtigt werden, insbesondere mit dem Ziel,
Anforderungen aus Endbenutzersicht hinsichtlich Übertragungs-
verzögerungen zu erfüllen. ■

➤ **Harte versus weiche Echtzeitbedingungen (hard vs. soft real-time requirements) :**

- **harte Echtzeitbedingungen :**
 - Echtzeitanforderungen nahezu immer zu erfüllen
→ Bsp.: PDV (Prozessdatenverarbeitung)
- **weiche Echtzeitbedingungen :**
 - Echtzeitanforderungen mit hinreichender großer Wahrscheinlichkeit zu erfüllen
→ Bsp.: paketisierte Sprachübertragung



➤ *Nebenbei :*

Garantien mit Wahrscheinlichkeit 1 überhaupt möglich ? (bei Echtzeitkommunikation? bei irgendwelchen technischen Systemen?)

➤ Unabdingbar bei Echtzeitkommunikation (und -verarbeitung) :
Kontrolle der (auftragsgenerierenden) Systemumgebung,
z.B. menschliche Benutzer, technische Prozesse.

Dienstgüte bei Medien- und Echtzeitkommunikation

➤ **Grobklassifikation von (anwendungsorientierten) Diensten** in Rechnernetzen :

- **Kommunikationsdienste** (IPC, Mail, ...)
- **Kooperationsdienste** (RPC, Load Balancing, ...)

→ wesentlich für Dienstleistung :

- spezifizierter Funktionsumfang erfüllt
- spezifizierte Qualitätsmerkmale erfüllt

Häufige Forderung bei Dienstleistung :

Realisierung einer (möglichst "garantierten") angemessenen Dienstqualität



➤ **Def. Dienstgüte/Dienstqualität** (Quality-of-Service, QoS) :

a) **ISO – Def.:**

Dienstgüte stellt eine Menge von Güten dar, die auf das gesamte Verhalten eines oder mehrerer Objekte bezogen sind
[nota bene: Objekte \equiv an Dienstleistung beteiligte Komponenten]

b) **ITU-T- Def.:**

Dienstgüte ist definiert als der kollektive Effekt der Dienst-Leistungen (service performances), die den Grad der Zufriedenheit des Benutzers eines Dienstes bestimmen. ■

➤ Mögliche Konsequenzen nicht-erfüllter bzw. zu geringer Dienstqualität, z.B.

- verspätete Dienstleistung
- Unverfügbarkeit von Dienst
- evtl. nutzlose Dienstleistung
- Bezahlung nicht-erbrachter (Netz-) Dienstleistungen

➤ **QoS-Management :**

Maßnahmen zur Realisierung der Dienstgüteverwaltung in einem Kommunikations-/Rechnernetz.

8.2 Anwendungen mit Bedarf an Medien- und Echtzeitkommunikation

➤ Grobklassifikation von Anwendungen mit benötigter Medienkommunikation:

- single-media

- nur 1 Mediendatenstrom (z.B. Sprachübertragung)

- multi-media

- Überlagerung von ≥ 2 Mediendatenströmen (z.B. Bildtelefonie)

➤ Klassifikation multimedialer Anwendungen in : [Hal 01] F. Halsall:

Multimedia Communications, Addison-Wesley 2001,

nota bene : single- sowie multi-media (gemäß obiger Grobklassifikation) umfassend

❖ Wichtige Trends im Internet : Voice over IP (VoIP), IPTV, Audio-/Video-Streaming, Video on Demand (VoD), ...

Kategorie	Medien	Anwendungs- beschreibungen
Kommunikation zwischen Personen ("interpersonal communications")	Sprache	Telefonie; Voice -Mail; Telekonferenzen
	Festbild	Fax
	Text	E-Mail
	Text und (Fest-)Bilder	Computer-supported cooperative working (CSCW)
	Sprache und Video	Videotelefonie; Video -Mail; Videokonferenz
	Text, Bilder, Audio/Video (A/V)	Multimedia E-Mail; Gruppen-Videospiele
Interaktive Anwendungen über das Internet	Text, Bilder, Audio/ Video	Informationsabfrage (z.B. Neueste Nachrichten, Wetter, Bücher, Zeitschriften, Videospiele, ...); E-Commerce
Unterhaltungs- dienste	Audio und Video	Audio-/CD-on-Demand; (Near) Movie-/Video-on-Demand; Analoges bzw. Digitales TV / IPTV (im Broadcast- bzw. Multicast-Verfahren); Interaktives TV

8.3 Endbenutzerorientierte Anforderungen bei verteilten multimedialen Anwendungen

- Welche **allg. Anforderungen** für Endbenutzer wichtig ?
 - Übertragung verschiedener Verkehrsarten mit jeweils spezifischen Güteanforderungen, abhängig von
 - *Verkehrsklasse* → Bsp.: Sprache vs. Bewegtbild und
 - *Anwendungstyp* → Bsp.: Spiele- vs. medizinische Anwendung
 - Synchronisation von Strömen (z.B. Sprache mit Bewegtbild)
 - Abhörsicherheit
 - Multicast-Funktionalität
 - angemessene Kosten, etc., etc.

- Welche **Dienstgüteanforderungen** für Endbenutzer wichtig (abhängig von Verkehrsklassen) ?
- **Daten**kommunikation
 - *Verzögerung*, im allg. $\leq 1-10$ sec (ca.)
 - *Durchsatz*, im allg. stark variabel
 - *Zuverlässigkeit*, i.a. sehr hoch
 - **Sprach**kommunikation
 - *Verzögerung*, i.a. $\leq 100-500$ ms (ca.)
 - *Durchsatz*, i.a. relativ gering (z.B. 64 kb/s)
 - *Zuverlässigkeit*, i.a. eher gering (z.B. Paketverlustwahrscheinlichkeit bei digitalisierter Sprache $< 5\%$)
 - **Audio-/Video**kommunikation
 - *Verzögerung*, i.a. 10-100 msec (ca.), bis zu 1 sec bei Verteildiensten
 - *Verzögerungsschwankung*, i.a. stark zu limitieren
 - *Durchsatz*, i.a. relativ hoch (z.B. 1-100 Mb/s), evtl. Komprimierung zur Durchsatzreduktion; bei IPTV: SDTV versus HDTV (Standard/High Definition TV)
 - *Zuverlässigkeit*, i.a. eher gering (abhängig vom Komprimierungsfaktor)
 - **Festbild**kommunikation
 - *Durchsatz*, i.a. kurzfristig hoch



➤ Aus endbenutzerorientierten Anforderungen resultierende **Probleme für Dienstanbieter**

- Anforderungen typischerweise
 - qualitativ
 - (evtl. sehr) subjektiv
 - schwer präzise spezifizierbar (da "fuzzy")
- Anforderungen aus Sicht des Dienstanbieters z.T. widersprüchlich, z.B.
 - hoher Durchsatz bei kleiner Ende-zu-Ende-Verzögerung
 - hohe Qualität bei geringen Kosten
(*nota bene* : Kostenerhebung für Anforderungen hochwertiger Kommunikationsdienste unerlässlich zur Steuerung des Benutzerverhaltens)
- Transformation der qualitativen Anforderungen in quantifizierbare
- Transformation der endbenutzerorientierten Anforderungen in solche, die für typische Kommunikationsnetz-Schnittstellen aussagekräftig (vgl. "QoS-Mapping" in Abschn. 8.5)

→ **Beispiel :**

Welche Implikation hat Forderung nach qualitativ hochwertiger Übertragung von medizinischen Bewegtbildsequenzen oder nach Lippen-synchronisation bei A/V-Kommunikation für das Kommunikationsnetz ?

8.4 Dienstgüte aus Kommunikationsnetz-Sicht

➤ Beurteilung von Dienstgüte

- an konkreter Schnittstelle zwischen Dienstbringer und -benutzer
- seitens Dienstbenutzer (im Hinblick auf die durch sie/ihn gestellten QoS-Anforderungen)
- evtl. durch Messungen
- durch Mitberücksichtigung der benutzerseitig angebotenen Last bzw. der übergebenen Auftragssequenz (z.B. von Übertragungsaufträgen)

➤ Typische Forderungen an Kommunikations-/Rechnernetze :

„Garantien“ bzgl.



- Verzögerungszeiten (“*delay*“) → nicht messbar bei Sender
- Verzögerungsschwankungen (“*delay jitter*“) → exakte Definition ?!
- Durchsatz (“*throughput*“), NICHT : Bandbreite (“*bandwidth*“) !
→ erledigte Übertragungsaufträge pro sec, z.B. Pakete/s, bit/s, ...
- Verlustraten (“*loss rate*“), z.B. für Pakete, Videoframes, ...
→ Problem der Def. von Durchsatz und Verlustraten :
zugrundegelegtes Beobachtungsintervall unverzichtbar für Def. !

QoS-Garantien

➤ Was wird "garantiert" ?

- Max.
- Min.
- Mittelwert
- Varianz
-
-
-



➤ Was heißt "garantiert" ?

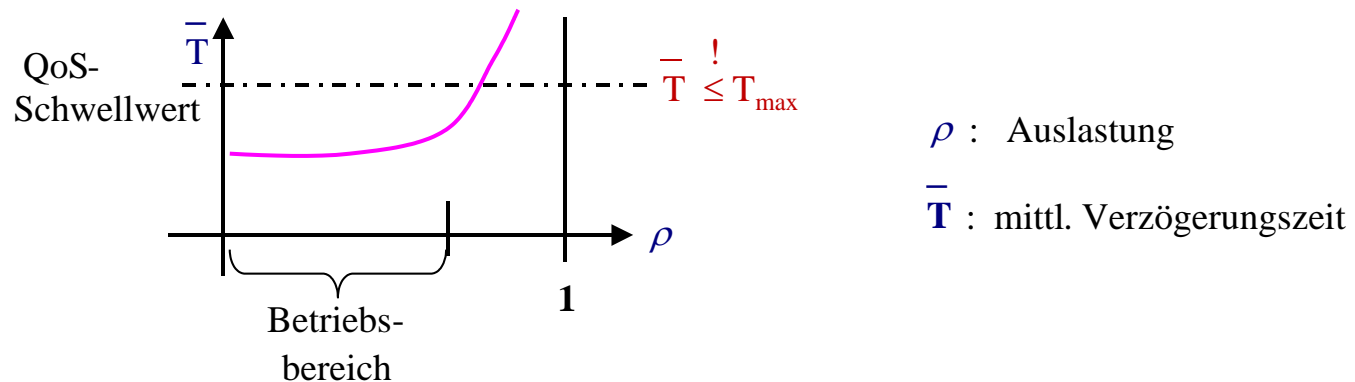
- (nahezu) mit Wahrscheinlichkeit 1
- mit großer Wahrscheinlichkeit
(evtl. ermittelt über relative Häufigkeit, z.B. in 95% aller Fälle, **ABER** wiederum : in welchem Beobachtungsintervall ?)

➤ *Beispiel* für wichtige Klasse von Kommunikationsnetzen mit QoS-Garantien: vgl. ATM-Architektur, -Dienste und -Protokolle in Abschn. 6.7

➤ **Möglichkeiten für QoS-Garantien** in Kommunikationsnetzen

- **"dumb fat pipe"** → Überdimensionierung ("Pipe" im Sinne einer virtuellen Verbindung)

nota bene: wegen Komplexitätsreduktion nicht notwendigerweise teurer als die Verfahren mit Priorisierung oder Betriebsmittelreservierung (s.u.)



- **Priorisierung**, z.B. zu übertragender Dateneinheiten (Pakete, Datagramme, ATM-Zellen, ...)

→ ABER : Wie zu hohes Verkehrsaufkommen auf höchster Prioritätsebene verhindern ?

(ergo : Zugangskontrolle unverzichtbar !)

Beispiele :

Verkehrspriorisierung durch Modifikation des Ethernet, des Internet (vgl. DiffServ)

Möglichkeiten für QoS-Garantien in Kommunikationsnetzen (Forts.) :

- **Betriebsmittelreservierung**, z.B. bei Verbindungsaufbau (im allg.: je starrer desto bessere QoS-Garantien und desto weniger Verwaltungsaufwand)
 - ABER :
 - evtl. Betriebsmittel nur gering ausgelastet,
 - d.h. Ineffizienzen und höhere Blockierungswahrscheinlichkeit bei Verbindungsaufbauwünschen
 - (nota bene : hier Zugangskontrolle bei Verbindungsaufbau)

Beispiele :

Betriebsmittelreservierung, z.B. in Routern, im Internet (vgl. IntServ)



- **Gegenwärtiger Stand** (2013/2014) :
 Zahlreiche Kommunikationsnetze noch ohne QoS-Garantien, sog. "best effort"-Netze, d.h. "Man tut, was man kann !" ☺

Beispiele :

- traditionelles Internet (insbes. das IP-Protokoll), Intranets
- traditionelles Ethernet, Fast Ethernet, Gbit-Ethernet, ...
- Mobilfunknetze (z.B. WLANs ≠ IEEE 802.11e), u.v.a.m.

⇒ Ausweg ?? ("dumb fat pipe" ?! oder unzufriedene Benutzer ?! oder FuE ?!)

8.5 Das „QoS-Mapping“-Problem

"QoS-Mapping" :



Notwendigkeit der **Abbildung** ("mapping") **von**

- den **benutzerseitigen Dienstanforderungen**, incl. der erwarteten QoS (z.B. qualitative Anforderungen beschrieben mit allg. Bewertungskriterien)

auf

- **die netzseitig zu erfüllenden Dienstgütemerkmale** (quantitative Anforderungen spezifiziert z.B. unter Benutzung der speziellen Leistungs- und Zuverlässigkeitskenngrößen für Netze)

→ "QoS-Mapping" als eine der Aufgaben von QoS-Managementsystemen.

Nota bene :

Resultat des "QoS-Mapping" i.a. nicht direkte und eindeutig bestimmte Konsequenz aus den gestellten Benutzeranforderungen, \rightarrow vgl. nachfolgendes

Beispiel :

Benutzeranforderung : "Garantie von Nutzdurchsatz von D_0 [kbit/s]"

wahlweise erbracht durch :

1. Netzseitigen Datendurchsatz von $D_0 + \Delta D$ (wobei ΔD hinreichend zur Übertragung der bei Dienstleistung entstehenden zusätzlichen Kontrollinfo) *und* vernachlässigbar geringe Paketfehler-/verlustraten
2. Netzseitigen Datendurchsatz von $D_x = (D_0 + \Delta D) \cdot (1 + x)$, $x > 0$ *und* Paketverlustrate ε_p unter Nutzung von FEC, wobei $x = f(\varepsilon_p)$ dergestalt, dass D_x den zur FEC erforderlichen, zusätzlichen Durchsatz beinhaltet.

8.6 Implikationen von Echtzeitkommunikationsanforderungen für kommunizierende Endsysteme

Adäquate Qualität bei Erbringung der Netzdienste

$\neg \Rightarrow$ adäquate Qualität aus Endbenutzersicht (\neg : *NOT*)

denn :

Engpässe in den Endsystemen möglich !

Ergo insbesondere notwendig : Echtzeitfähiges Betriebssystem

Ansatz :

- Priorisierung von Prozessen bzw. Tasks
- Möglichkeit der Unterbrechung von Prozessen
- Nutzung geeigneter Scheduling-Algorithmen zur CPU-Vergabe

Beispiele für Echtzeit-Scheduling-Algorithmen :

– Earliest Deadline First (EDF) :

Tasks priorisiert;

Task mit frühester Deadline (bzgl. Ausführungsende) erhält CPU

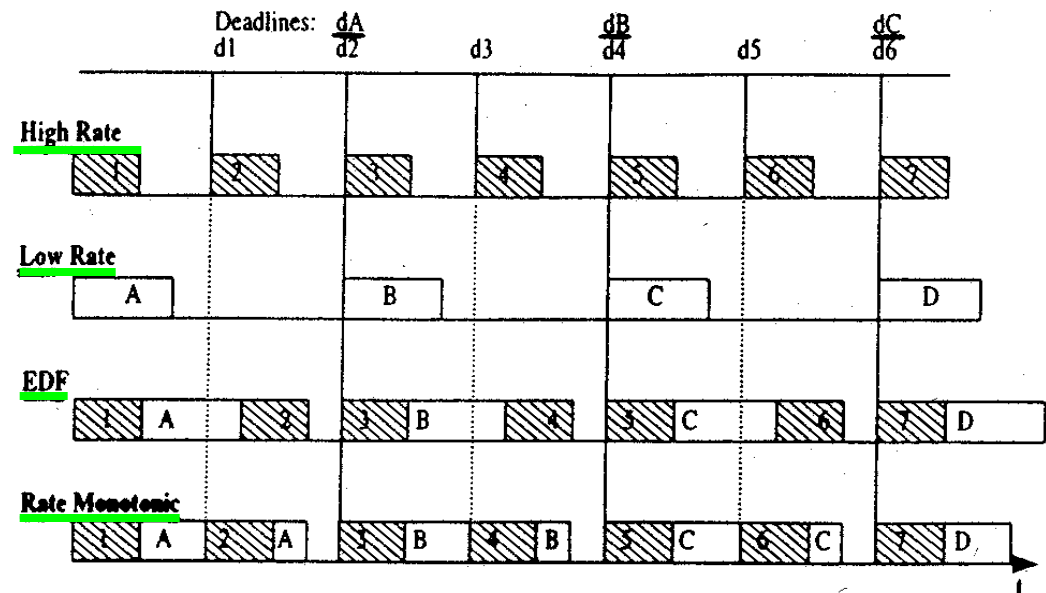
→ Scheduling bezogen auf höchste vertretene Prio-Klasse;

bei Entstehen von neuer Task T_N : evtl. Unterbrechung der z.Zt. ausgeführten Task und direkte Ausführung von T_N , sofern Deadline für und Prio. von T_N dies rechtfertigt (d.h. liegt vor allen anderen Deadlines).

– Rate Monotonic Scheduling :

statischer, prioritätsgesteuerter Schedulingalgor. für gegenseitig, unabhängige, unterbrechbare, periodische Tasks mit jeweils konstanten Ausführungszeiten.

Exemplarische Illustration :



8.7 Verbesserung vorhandener und Gewährleistung geforderter Dienstgüte aus Endbenutzersicht

QoS-Verbesserungen in "Best Effort" -Kommunikationsnetzen :



- Nutzung **schnellerer Kommunikationshardware und -software**, wie z.B. schnellere Übertragungsmedien, Vermittlungsrechner, Netzadapter, Prozessoren in Endsystemen u.a.m.
- **Beschleunigung** der **Algorithmen** zur Realisierung der Netzdienste (z.B. schnellere und/oder lastadaptive Routingalgorithmen)
- **Erweiterung** von **Protokollen/Diensten** um Funktionalität für die Gewährleistung von Dienstgüte, wie z.B. – bezogen auf das Internet –
 - a) Betriebsmittelreservierung in Routern (gemäß RSVP)
 - b) Priorisierung von Datagrammen→ Details zu Protokoll-/Dienstenerweiterung, s.u.

- **Erweiterung** von **Protokollen/Diensten** um Funktionalität für die Gewährleistung von Dienstgüte, wie z.B. – bezogen auf das Internet –
 - a) **Betriebsmittelreservierung in Routern** (gemäß **RSVP**)
 - Idee : Erkennen von Verbindungen (hier auch "flows") mit spez. QoS-Anforderungen, dedizierte Reservierung von Ressourcen in RSVP-fähigen Routern, vgl. 8.4
 - b) **Priorisierung von Datagrammen**
 - Idee : Nutzung des "**Type-of-Service** (ToS) "-Feldes von IPv4 zur Spezifikation von QoS-Anforderungen (als Prioritätsangabe interpretierbar), vgl. 8.4
- ⇒ **Vergleich** der Maßnahmen a) versus b) :
Nachteile von a) :
 - viel Zustandsinfo erforderlich
 - relativ komplex, d.h. teuer zu implementieren
[nota bene : schnelle Router unverzichtbar im heutigen Internet !]
- **"Information Dispersal"**, d.h. Aufsplitten des zu übertragenden Verkehrs auf möglichst (Knoten- und Leitungs-) disjunkte Verbindungen /Kommunikationspfade
→ z.B. Nutzung verschiedener ISPs
- **Kommunikationsnetzinterne Nutzung** von **FEC**
u.v.a.m.

Reduktion der QoS-Anforderungen



- **Reduktion der Durchsatzanforderungen**, u.a. durch Einsatz von Kompressionsverfahren, z.B. bei Videokommunikation, zur Eliminierung der
 - **zeitlichen Redundanz** (Abhängigkeit zwischen Folgebildern)
 - **räumlichen Redundanz** (mögl. Ähnlichkeiten zwischen benachbarten Bildbereichen)in Bewegtbildsequenzen, vgl. Standards
 - MPEG-1, -2, -4, ...,
 - H.261, H.263, H.264, etc.
- **Reduktion der Zuverlässigkeitsanforderungen**
 - vgl. Bereinigung von Fehlern/Verlusten durch FEC (erkauft durch erhöhten Durchsatz)
- **Reduktion der zeitlichen Anforderungen**
 - z.B. Nutzung von „Video Display Buffer“ zum empfangsseitigen Ausgleich von Verzögerungsschwankungen (absichtliche Verzögerung der empfangsseitigen Videopräsentation zu Beginn der Videosequenz, d.h. Puffer erst partiell gefüllt mit Videodaten); vgl. “*Video Streaming*“

Einschub zu Video-Streaming :

Eine der frühesten Publikationen zu Streaming

Most Cited Computer Science Articles generated from documents in the CiteSeerX database as of February 26, 2008. This list is automatically generated and may contain errors. The list is generated in batch mode and citation counts may differ from those currently in the CiteSeerX database, since the database is continuously updated.

[All Years](#) | [1990](#) | [1991](#) | [1992](#) | [1993](#) | [1994](#) | [1995](#) | [1996](#) | [1997](#) | [1998](#) | [1999](#) | [2000](#) | [2001](#) | [2002](#) | [2003](#) | [2004](#) | [2005](#) | [2006](#) | [2007](#) | [2008](#)

... Auszug aus der Liste der Publikationen des Jahres 1991:

175. 91 M A Shwe, B Middleton, D E Heckerman, M Henrion, E J Horvitz, H P Lehman, G F Cooper. *Probabilistic diagnosis using a reformulation of the INTERNIST-1/QMR knowledge base.*

I. The probabilistic model and inference algorithms. Methods of Information in. Medicine, , 1991.

176. 91 J Y Halpern, Y Shoham. *A propositional modal logic of time intervals.* Journal of the ACM, , 1991.

177. 89 X Lai, J Massey, S Murphy. *Markov Ciphers and Differential Cryptanalysis.* Proceedings of Eurocrypt 91, LNCS, , 1991.

178. 89 Michał Muzalewski. *Construction of rings and left-, right-, and bi-modules over a ring.* Formalized Mathematics, , 1991.

179. 89 B Wolfinger, M Moran. *A Continuous Media Data Transport Service and Protocol for Real-Time Communication in High Speed Networks.* Proc. Second Int'l. Workshop on Network and Operating System Support for Digital Audio and Video, , 1991.

180. 89 N Sadeh. *Look-ahead techniques for micro-opportunistic job shop scheduling.* ph.d. diss. , 1991.

181. 88 Alex Lascarides, Nicholas Asher. *Discourse relations and defeasible knowledge.* In Proceedings of the 29th Annual Meeting, , 1991. ...

Einschub zu Video-Streaming :

Eine der frühesten Publikationen zu Streaming

B. E. Wolfinger, M. Moran. *A Continuous Media Data Transport Service and Protocol for Real-Time Communication in High Speed Networks*. Proc. Second Int'l. Workshop on Network and Operating System Support for Digital Audio and Video, 1991.

Abstract:

An important class of applications with real-time data transport requirements is defined by applications requiring transmission of data units at regular intervals. These applications, which we call continuous media (CM) clients, include video conferencing, voice communication, and high-quality digital sound. The design of a data transport service for CM clients and its underlying protocol (within the XUNET II project) is presented in this paper. The service makes use, in particular, of an a priori characterization of future data transmission requests by CM clients. First, we will give a few examples of CM clients and their specific data transmission needs. From these clients, we then extract a generalized list of data transport requirements for CM and describe the basic features of a service designed to meet these requirements. This service provides unreliable, in-sequence transfer (simplex, periodic) of so-called stream data units (STDUs) between a sending and a receiving client, with ...

Reduktion der QoS-Anforderungen (Forts.)

➤ Reduktion der Sicherheitsanforderungen

→ z.B. durch

- “*Information Dispersal*“ in Verbindung mit geeigneter Codierung (s.o.) : Rekonstruktionsmöglichkeit für Daten nur bei Verfügbarkeit der Übertragungen mehrerer (oder gar sämtlicher) ISPs
- netzexterne Verschlüsselung

➤ Ablehnung von Verbindungsaufbauwünschen

→ Reduktion der Gesamtmenge der Benutzeranforderungen

Steuerung des "Umgebungsverhaltens"



Umgebung (hier):

Menge der auftragserzeugenden Endbenutzer (bezogen auf den/die betrachteten Dienst(e) zur Medienkommunikation).

a) Maßnahmen zur **Beeinflussung des individuellen Benutzerverhaltens:**

- A priori Abschluss eines "Traffic Contracts" zwischen DÜ-Diensterbringer und -benutzer bei Verbindungsaufbau (!) : Benutzer charakterisiert die später während der Verbindung zu erwartende Last ("worst case" aus Sicht von Dienstbringer), vgl. LBAP-Lastcharakterisierung, s.u.
- Überprüfung(en) während Verbindung, ob "Traffic Contract" eingehalten → Messungen
oder alternativ:
 Zugangskontrolle für Verkehr eines Benutzers (z.B. "Leaky Bucket", "Token Bucket", s.u.) → Erzeugung einer möglichst konstanten Senderate
- Verkehrsglättung ("smoothing") : Benutzer verteilt Sendewünsche möglichst gleichmäßig über sog. "Glättungsintervall" der Länge T → durch Glättung evtl. Erhöhung der zeitlichen Anforderungen bei DÜ durch das Netz, aber weniger stoßweise Belastung

Steuerung des "Umgebungsverhaltens" (Forts.)

b) Maßnahmen zur **Beeinflussung der Gesamtmenge der Benutzer** :

- Ablehnung neuer Verbindungsaufbauwünsche oder gar Abbrechen existierender Verbindungen
- Reduktion der Qualitätsansprüche von Benutzern, z.B. Router mit zwangsweiser Veränderung benutzerseitig erfolgter Prioritätseinstufungen für Pakete, u.a.m.

Das “LBAP-Modell“

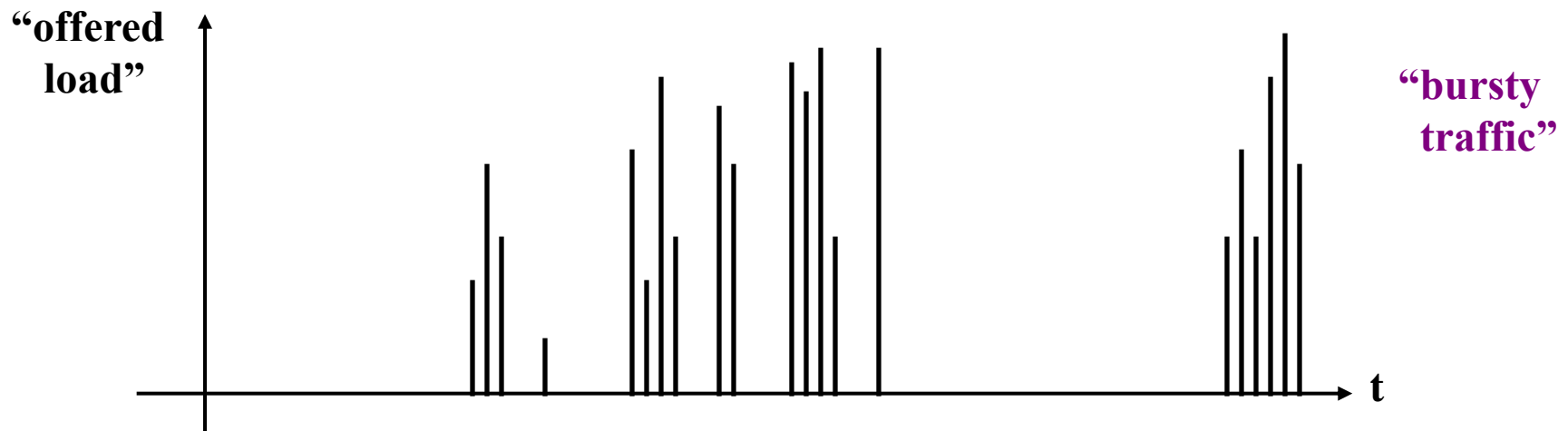


Ziel:

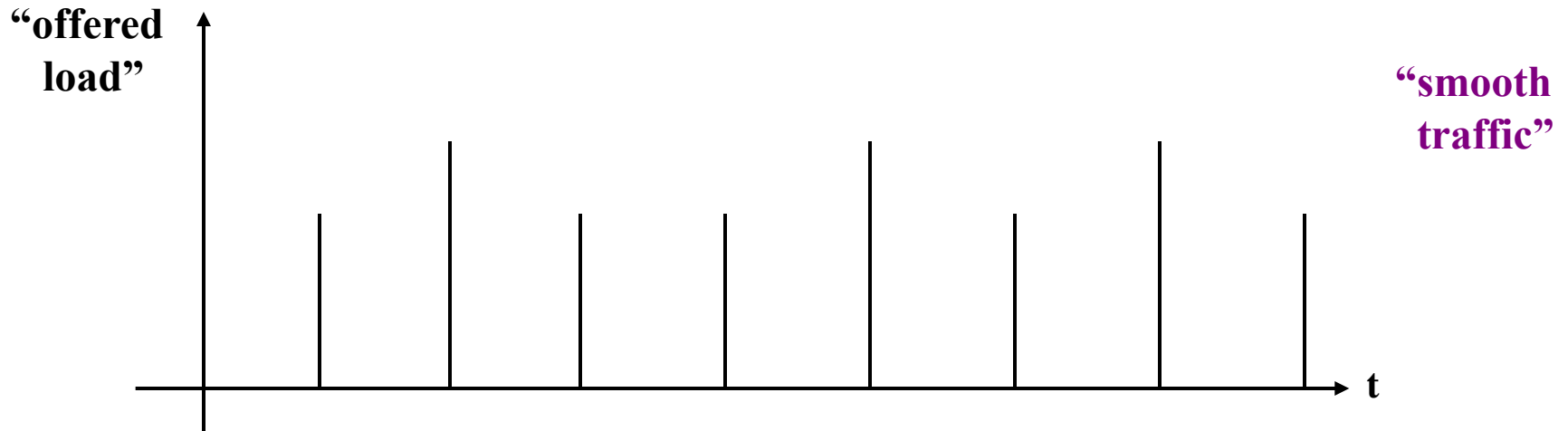
A priori - Beschränkung der während einer Verbindungsdauer seitens eines Benutzers von Übertragungsdiensten an das Kommunikationsnetz KN übergebenen Datenmengen pro Zeit.

Beispiele für Verkehrsaufkommen :

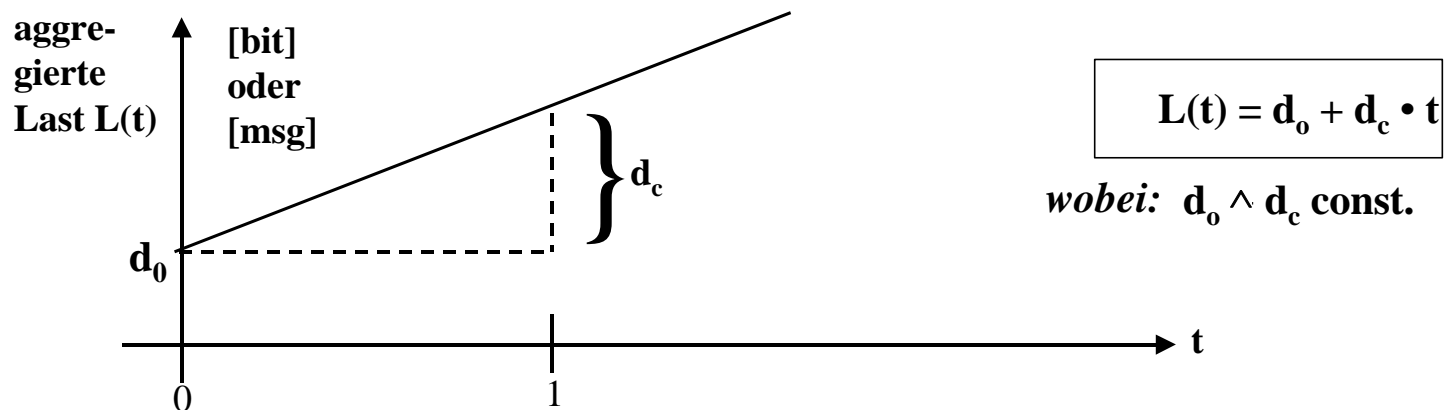
- "unschön", z.B.



- Wünschenswertes Verkehrsaufkommen, z.B.



⇒ Das LBAP-Modell (**LBAP** – Linear Bounded Arrival Process)



Das LBAP-Modell : Benutzerkonformität zum Modell

Benutzerverhalten konform zu **LBAP – Modell**

\Leftrightarrow Für alle mögl. Beobachtungszeitpunkte t_B gilt :

für beliebiges $t < t_B$ ist die Ungleichung

$$l[t, t_B] \leq d_o + d_c \bullet (t_B - t) \text{ erfüllt,}$$



wobei $l[t, t_B]$ die seitens eines Benutzers U im Intervall $[t, t_B]$ an das KN übergebene Datenmenge bezeichne;

d_o z.B. in Bit (oder Byte) und d_c z.B. in Bit/s (oder Byte/s).

Nota bene:

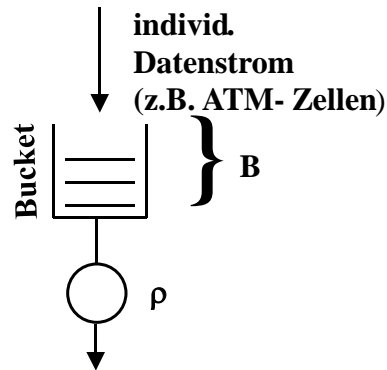
Bei konstantem Durchsatz von d_c (z.B. Leitungsvermittlung) ist Puffer der Größe d_o für KN ausreichend; überdies QoS-Garantien stark erleichtert.

“Leaky Bucket“ versus “Token Bucket“

Ziel : Verkehrsglättung und Lastbeschränkung durch ratenbasierte Zugangskontrolle bei Inanspruchnahme eines Übertragungsdienstes.

➤ **“Leaky Bucket“ :**

→ vgl. Eimer mit Loch, der dynamisch mit Wasser gefüllt wird



B : “Bucket“- Größe
 (“Bucket ”≡ Behälter, z.B. Puffer)

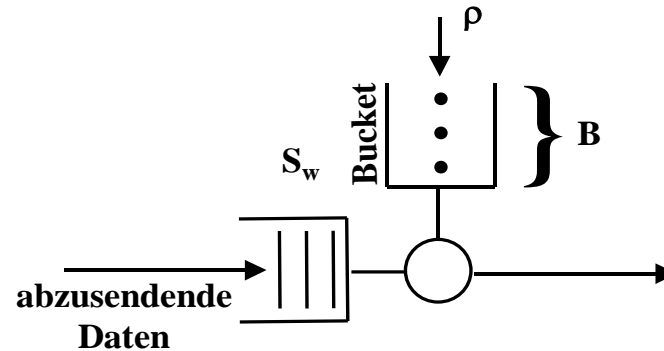
ρ : Senderate (konstant, sofern
 B nicht-leer), z.B. [Byte/s]



Prinzip :

- Abzusendende Dateneinheiten (z.B. ATM-Zellen, IP-Datagramme, o.ä.) in Puffer/Bucket übergeben; Datenverlust sofern freier Pufferbereich zu klein für neu ankommende Dateneinheit (DE)
- Absenden von 1. wartender DE der Länge l [Byte] nach l/ρ Zeiteinheiten.

➤ “Token Bucket“ :



Prinzip :

- Generierung von "Tokens" mit konstanter Rate ρ , sofern Bucket-Belegung $< B$ (ansonsten Tokengenerierungspause)
- jeder Token entspricht Kredit für Absenden von 1 Byte,

ergo : auf Absenden in Sendewarteschlange S_w wartende DE der Länge l [Byte] genau dann abgesendet, sofern Bucket-Belegung $\geq l$, und dabei Bucketgröße um l Tokens reduziert

⇒ Wirkungsweise : Menge der in Zeitintervall der Länge T abgesendeten Daten limiert durch

$$B + T \cdot \rho,$$

vgl. LBAP-Modell !

Nota bene: Leaky Bucket ist Sonderfall von Token Bucket für $B = 0$

“Smart Applications“

Annahme : Kommunikationsnetze ohne hinreichende QoS-Garantien

⇒ Ausweg :

Ausgleich (so gut wie möglich!) **der** eventuell starken **Qualitätsschwankungen** des Netzes **auf Anwendungsebene bzw. in** sog. **“Middleware“-Schichten** zwischen

Netzdiensten und kommunizierenden Anwendungsprozessen

- *bei Netzkonfigurierung* :

kein bzw. wenig Einfluss des QoS-Managements

- *bei lfd. Betrieb* :

“Lernen“ über (Momentan-) Zustand des Netzes und Erhöhung der Dienstqualität “außerhalb“ des Netzes, d.h. Realisierung adaptiver Anwendungen bzw. sog.

“smart applications“.

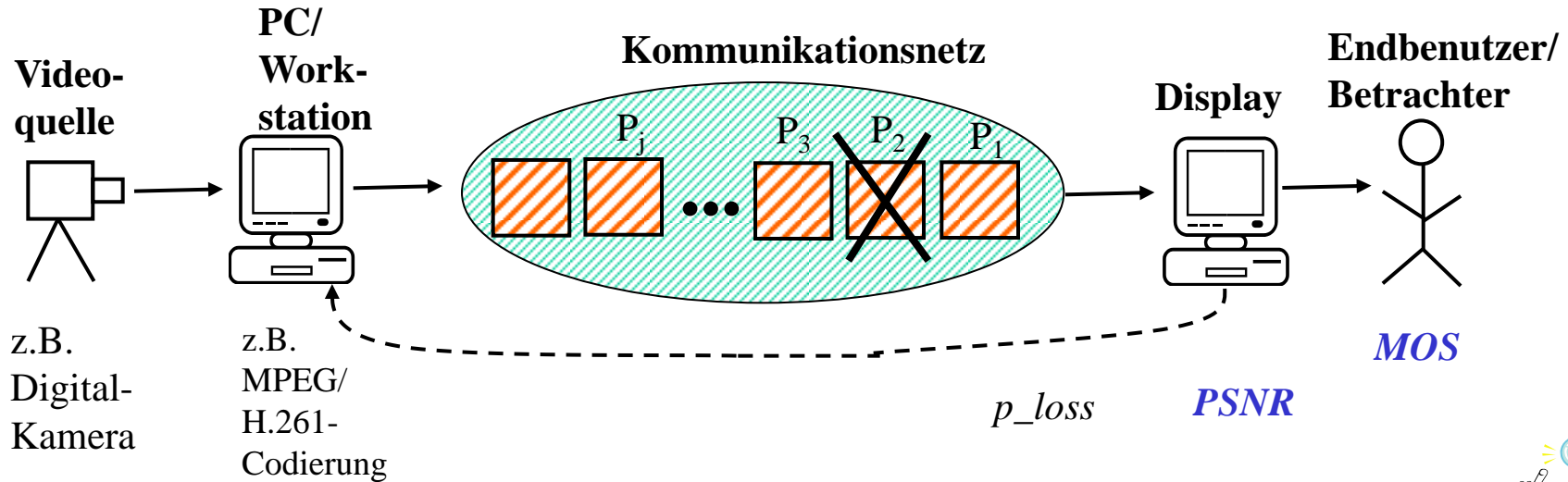
Beispiel zu „Smart Applications“:

Videokommunikation über traditionelles Internet

- Momentanzustand des Netzes, z.B. charakterisiert durch beobachtete Paketverlustraten in geeignet gewählten Zeitintervallen
- Adaptionsmaßnahmen der "smart applications" bei beobachteter signifikanter Erhöhung der Paketverlustraten, z.B.
 - häufigeres Senden unabhängig codierter Bilder zur Begrenzung der Fehlerfortpflanzung
 - Änderung der Quantisierung zur Reduktion des Verkehrsaufkommens mit erhoffter Reduktion der Paketverlusthäufigkeiten)
 - Fragmentierung in kleinere Pakete mit im allg. geringerem Verlustrisiko

o. ä.

Procedere zur adaptiven Videocodierung



Notation:

P_j : (IP-) Pakete, ATM-Zellen, o.ä.; pro Videoframe $\rightarrow \geq 1$ Paket / ATM-Zelle

p_{loss} : Paketverlusthäufigkeit, wobei Paketverlust Verlust von ≥ 1 Videoframes impliziert

Bewertungsmaße (*quantitativ* – wie bei PSNR versus *qualitativ/subjektiv* – wie bei MOS)

PSNR: **P**eak **S**ignal to **N**oise **R**atio

MOS: **M**ean **O**pinion **S**core

Idee: Adaptive Codierung abhängig von empfangsseitig beobachteter Paketverlusthäufigkeit (um QoS über einer vorgegebenen Qualitätsschranke zu halten)



Zur gegenseitigen Beeinflussung von verteilten (Echtzeit-)Anwendungen und Kommunikationsnetzen

➤ Netz beeinflusst verteilte Anwendung

Bsp.: Paketverlust [direkt oder da Paketverzögerung $\tau_p > T_{max}$, wobei T_{max} den gerade noch akzeptablen Verzögerungsschwellwert bezeichne]
 \Rightarrow (direkter) Verlust eines Videoframes oder von Teilen davon
 \Rightarrow evtl. indirekter Verlust weiterer Videoframes (Fehlerfortpflanzung)
 \Rightarrow QoS-/QoE-Einbuße

Verbesserungen u.a.:

- Reduziere p_{loss} (besseres Netz, FEC, ...)
- Reduziere τ_p (besseres Netz, weniger „Bursts“ bei generierter Last, ...)
- Akzeptiere größeres T_{max} seitens Empfänger, d.h. auf Anwendungsebene
- Begrenze Fehlerauswirkung bei Empfänger

➤ Verteilte Anwendung beeinflusst Netz

Verbesserungen u.a.:

- Reduziere Durchsatzforderungen (stärkere Komprimierung)
- Reduziere „Burstartigkeit“ bei generierter Last („Smoothing“)
- Nutze adaptive Codierung

Ergo: Konfigurierungs-/Dimensionierungsaufgaben, u.a.

- Etabliere Netz mit noch akzeptablem p_{loss} und τ_p
- Wähle Videocodierung, die gegebenes p_{loss} und τ_p noch „verkräftet“
- Passe ggf. Echtzeitschranke T_{max} an Netzcharakteristika an \Rightarrow s. Tool **MedienExplorativ**

Erkennung von QoS-Verletzungen

Dienstgüteüberprüfungen :

- *periodisch*
- *ereignisgesteuert*

(im allg. häufigere Überprüfungen bei “best effort”-Netzen)

Definition:

Quality-of-Experience (QoE) bezeichnet die aus Benutzersicht empfundene Dienstgüte.

Methoden der Dienstgüteermittlung :

a) **QoE-Beurteilung aus Benutzersicht**

(z.B. explizite Befragung, permanente QoS-/QoE-Überwachung, ...)

- + : • Dienstgüte aus Benutzersicht letztendlich ausschlaggebend
- : • subjektive Bewertung
 - evtl. hoher Aufwand für Benutzer bei häufiger Befragung
 - evtl. verspätete Reaktion auf QoS-/QoE-Reduktion (menschl. Reaktionszeit)

Methoden der Dienstgütermittlung (Forts.):

b) QoS-Beurteilung mittels Messungen

b1) Messungen in den kommunizierenden Endsystemen

→ netzexterne Messungen, d.h. indirekte Beobachtung des Kommunikationsnetzes (auch für “Black-Box Netze” brauchbar)

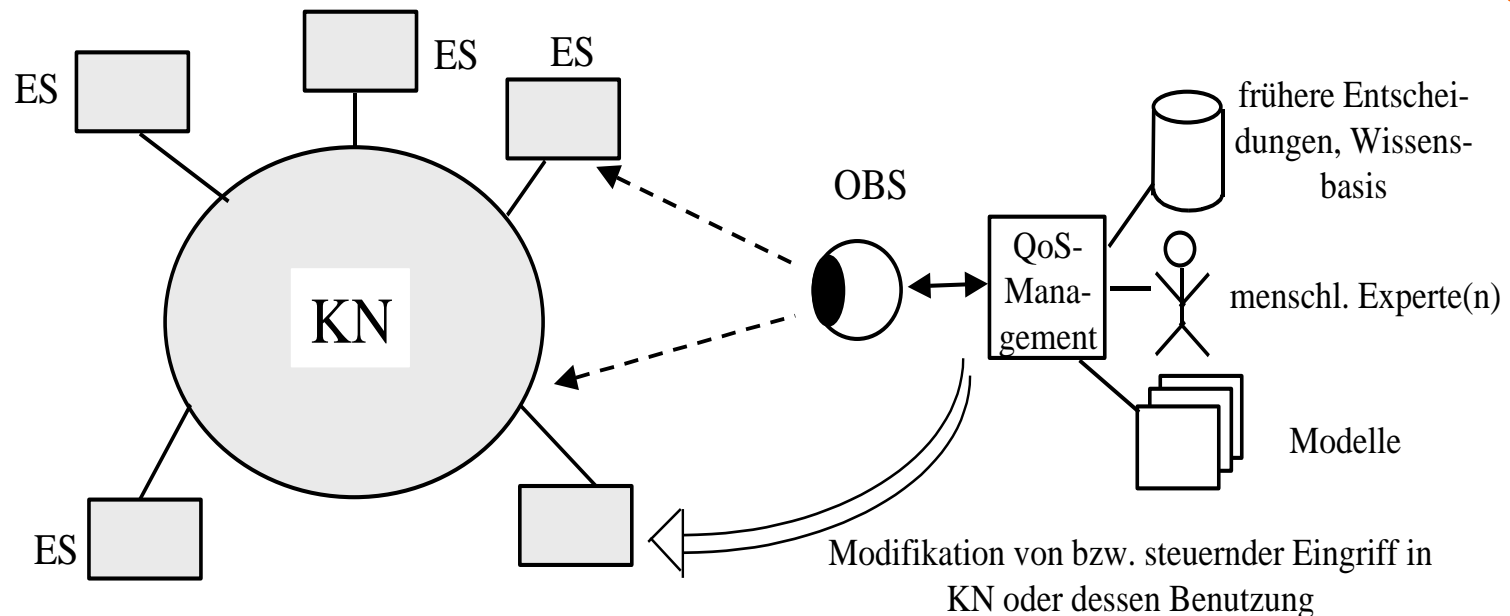
b2) Messungen direkt im Kommunikationsnetz

→ netzinterne Messungen mit verbesserten Möglichkeiten der Engpasserkennung

- + :
 - objektive Bewertung der QoS
 - schnelle Reaktionsmöglichkeit auf QoS-Reduktion durch automatisierte Messdatenerfassungen und -auswertung
 - QoS-Überprüfung transparent für Endbenutzer
- :
 - nicht gemessene QoS sondern QoE aus Benutzersicht relevant
 - evtl. Beeinträchtigung der Vorgänge im Netz durch Messungen
 - “Momentanzustand” für Netz evtl. schwierig zu definieren (z.B. bei Momentanauslastung, Momentandurchsatz, ...)
 - beobachteter Netzzustand : veraltete Sicht, nur partielle Sicht (auf Subsystem)

Grobarchitektur für modellgestütztes QoS-Management

Komponenten eines (modellgestützten) QoS-Managementsystems



Notation :

ES ≡ Endsystem mit Benutzer(n)

KN ≡ Kommunikationsnetz (evtl. gekoppelte Transitnetze)

OBS ≡ Beobachtungskomponente(n) [„Observer“-Komponente]

➤ Rolle der Hauptkomponenten :

- **OBS :** Komponenten zur Gewinnung von Zustandsinfo (zur Unterstützung von QoS-Entscheidungen, incl. Parametrisierung von Prognosenmodellen),

z.B. Messwerkzeug(e) für Messungen

- im Kommunikationsnetz (z.B. Auslastungen von Netzkomponenten, Paketverlusthäufigkeiten auf Pfaden)
und/oder
- in den Endsystemen (z.B. gegenwärtiges Verkehrsaufkommen, derzeitige QoE aus Endbenutzersicht)

- **QoS - Manager :**

Komponente zur Realisierung des QoS-Managements gestützt auf

- erfasste Zustandsinfo
- frühere QoS-Managemententscheidungen
- Expertenmeinungen/-urteil
- Prognosemodelle bzgl. zu erwartendem Netzverhalten als Konsequenz entsprechender Managemententscheidungen.

➤ **Procedere des QoS-Managers**

Schritt 1 :

periodischer oder ereignisgesteuerter **Empfang von Zustandsinfo** seitens der Beobachtungskomponente(n)

- u.a. Momentanzustand der Betriebsmittelbelegung, gegenwärtiges Niveau der Dienstqualität, Momentan-Belastung des Kommunikationsnetzes und Belastungstrends

Schritt 2 :

Vorbereitung einer QoS-Managemententscheidung, z.B.

- nach Ablauf von Δt
- nach Feststellen einer QoS-Verletzung
- bei Wunsch nach Aufbau einer Verbindung mit QoS-Garantie
- bei Komponentenausfällen (Reduktion der verfügbaren Betriebsmittel)

Schritt 3 :

Modellbasierte QoS-Prognose (Leistungs-/Zuverlässigkeitsprognose) für mögliche Managemententscheidungen

- Suche der optimalen Entscheidung D_{opt}

Schritt 4 :

sofern Eingriff in das Kommunikationsnetz (KN) selbst bzw. in die Art der KN-Benutzung notwendig :

steuernder Eingriff (z.B. Beeinflussung der Betriebsmittelbelegung); weiter bei Schritt 1.

8.8 Zur Beurteilung der Dienstgüte aus Endbenutzersicht

Möglichkeiten der Beurteilung:

- **Benutzer fragen**, z.B. *Mean Opinion Scores* (MOS)
- **Quantitative Maße zur benutzerorientierten Qualitätsbewertung**, z.B. *Peak Signal to Noise Ratio* (PSNR) für Videosequenzen



⇒ Grundsätzliche Probleme :

- Aufwand für Bewertung
- Subjektivität der Endbenutzer
- Hohe Qualität nicht kostenlos (Abrechnungsmodell notwendig)
- Echtzeitfähigkeit mit zu bewerten (evtl. problematisch für Modelle)

MedienExplorativ: Ein eLearning-Werkzeug zur Bewertung von Audio-/Videokommunikation

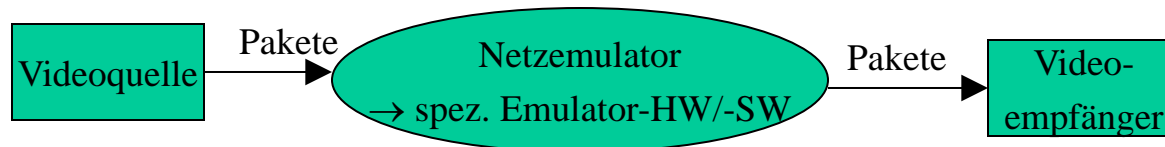
Videokommunikation über verlustbehaftete Netze

- Realisierungsvarianten -

I. Ausschließlich Realkomponenten



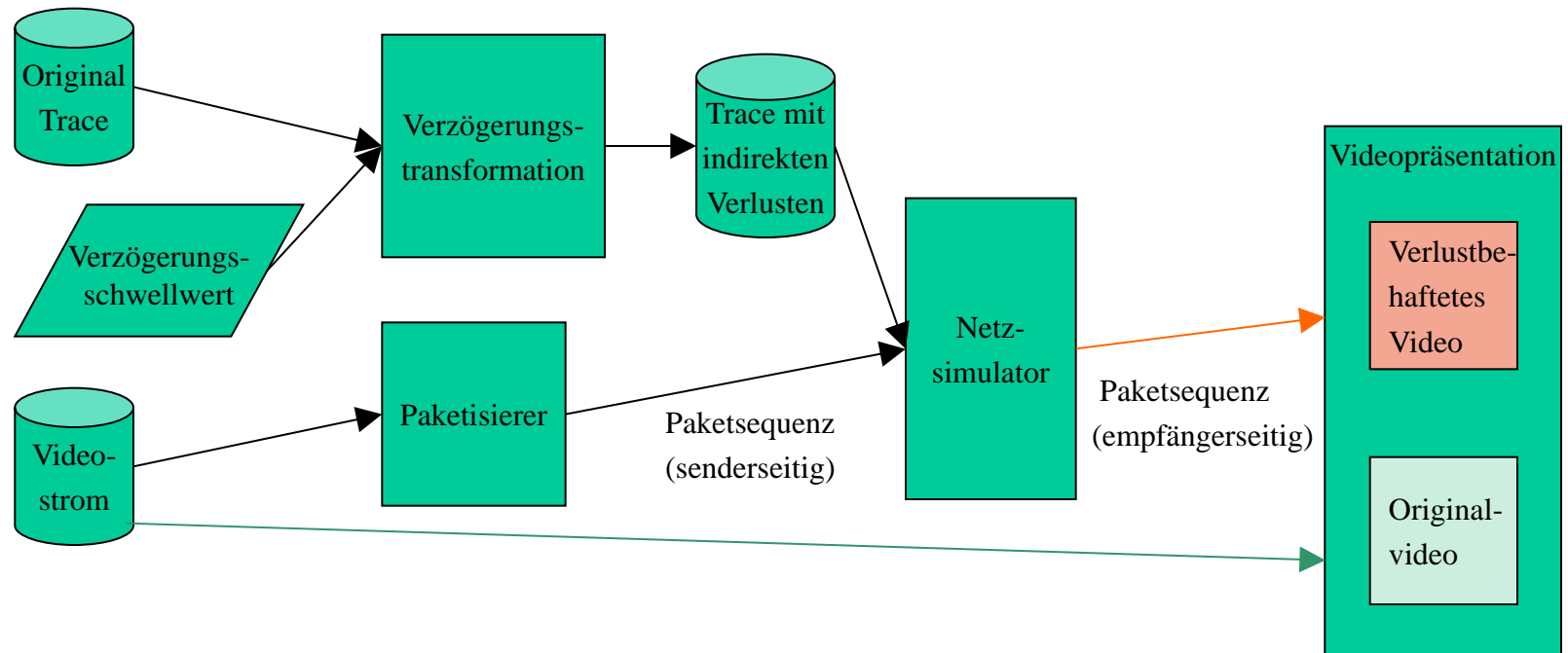
II. Reale Videoanwendung; emuliertes Netzverhalten



III. Gewinnung von Traces zur Charakterisierung von Paketverzögerungen / -verlusten im Netz; Simulation zur zeitl. Charakterisierung des Videostroms an der Schnittstelle zum Videoempfänger

Architektur für *VideoExplorativ*

(Videokomponente des Tools *MedienExplorativ* –
entsprechende Komponente verfügbar für Audio)



Hauptkomponenten und Ablaufschema von *VideoExplorativ*

[Literatur: Fiolka K., Heidtmann K., Wolfinger B.E., Ein eLearning-Werkzeug zur Videokommunikation über simulierte verlustbehaftete Netze, GI-Workshop "Elektronische Unterstützung der Präsenzlehre" im Rahmen der GI-Jahrestagung INFORMATIK 2004, Ulm, September 2004]

Möglichkeiten zur Gewinnung valider Traces für Paketverzögerungen in Netzen

- a) **Direkte Messungen in realen Netzen** (z.B. Internet, WLANs)
 - + Trace valide (für betrachtetes Szenario)
 - Übertragbarkeit; Messaufwand (auch: Wahl geeigneter Schnittstellen)
- b) **Messungen unter Nutzung von Simulator**
 - + relativ aufwandsarme Gewinnung von Traces für unterschiedliche Szenarien
 - Validationsprobleme für Simulator; evtl. aufwändige Erstellung von Simulator
- c) **Erstellung von Traces gemäß Wahrscheinlichkeitsverteilung**
 - + sehr aufwandsarme Gewinnung von Traces
 - beträchtl. Validationsproblem (welche Verteilungen, mit welchen Parametern?)
- d) **Erstellung von Traces „von Hand“**
 - + sehr flexible Möglichkeiten (z.B. Untersuchungsmögl. von Extremsituationen)
 - gute Kenntnisse bzgl. Netzverhalten benötigt; Aufwand bei „großen“ Traces; evtl. werden völlig unrealistische Traces erzeugt

Exemplarische Nutzung von VideoExplorativ

Kurzdemo (evtl. in Vorlesung sowie bei Praktikumsversuchen): experimentelle Randbedingungen

Experimentserie

- *Gemessener Trace* für speziellen Pfad durch das Internet:
“Uni HH (FB Informatik) → Kairo“ (u.a. Messung vom Sept. 2004)
- *Verzögerungsschranke* (Endbenutzervorgabe) :
 - Ohne Verzögerungsschranke (ergo: nur direkte Paketverluste des Internet)
 - 220 ms (ergo: auch indirekte Verluste durch die gegenseitigen Abhängigkeiten der Videoframes)

Ziel: Veranschaulichung der Implikationen auf

- a) Video “*Claire*“, MPEG-Video (geringe Bewegungsintensität)
- b) Video “*Foreman*“, MPEG-Video (mittlere Bewegungsintensität)

→ vgl. auch Nutzung von VideoExplorativ/MedienExplorativ im DKR-Praktikum !