Kapitel 1

Hochleistungskommunikation

1.1 Motivation

HLK: leistungsfähige, dienstintegrierende Netze und Systeme
Beispiele: Medizinische Bildverarbeitung, Telekonferenzen, MBone-Tools.
CSCW Computer Supported Cooperative Work
Anwendungsanforderungen: Gigabit-Anwendungen: hohes Datenvolumen & geringe Verzögerung (z.B. Virtual Reality, ABER kein: LAN, HDTV, Distributed Computing, RPC)
Misc zur Motivation:

- Fernsprechen: siehe Telematik1; Minimale Abtastfrequenz 6800Hz (technisch realisiert 8kHz)
- 3-Komponenten-Modell: Netzorientierte Komponente, Transportorientierte Komponente, Anwendungsorientierte Komponente
- Leistungsorientierte Dienstmerkmale:
  - Durchsatz, Burstisness (Verhältnis (max. Bitrate / mittlere Bitrate) eines Bursts.), Verzögerung, Jitter (Schwankung in der Verzögerung von Dateneinheiten), Antwortzeit, Zuverlässigkeitsparameter.
- Dienstklassen: (bei Ressourcenreservierungen benötigt)
  - deterministisch (auch im worst case garantiert)
  - statistisch (mit Wahrscheinlichkeit p garantiert)
  - bestmöglich (keine Garantie)
- Allgemeine Dienstmerkmale:
  - Reihenfolgetreue, Gruppenzustellung, Sicherheitsanforderungen, Synchronisation, Segmentierte Datenzustellung
- Synchroner (exakt definierter Zeitraum) - asynchroner Dienst (zufällig)
- isochroner Dienst (Audio, Video): maximaler Jitter definiert → minimale Verzögerung (kann an unterschiedlichen Dienstschnittstellen erbracht werden!!: isochroner Transportdienst auch ohne isochrone Übertragung möglich.)

- Überblick Vermittlungstechniken: Leitungsvermittlung, Paketvermittlung (Datagramme (verschiedene Wege möglich), virtuelle Verbindungen (ein Weg)) , Multiraten-Leitungsvermittlung (hohe Basisrate (inefizient) vs. niedrige Basisrate (viel zu synchronisieren) (Overlay-Switches(Kontrolle, ISDN, H1, H4 (Bandbreiten), schlechte Ressourcenausnutzung)), Schnelle Leitungsverm., Schnelle Paketverm. (ATM))

1.2 Digitale Übertragungshierarchien

- synchron: einheitliche Grunddatenrate, Taktinfo gilt auch für Teilströme
- plesiochron: s.o. Taktinfo gilt einzeln für jeden Teilstrom und zusätzlich für Gesamtstrom
- asynchron: Datenblöcke werden bedarfsgesteuert den jeweiligen teilströmen zugeordnet

PDH Plesiochrone Digitale Übertragungshierarchie:

- plesio: Schwankungen der Rate durch Stopfbits kompensieren. Zeitmux (Sprachübertragung)
- Problem: Herausfiltern eines niederwertigeren Datenstromes (alle Multiplexstufen durchlaufen)

SONET Synchronous Optical Network
SDH Synchronous Digital Hierarchie

- SDH als Obermenge
- Synchrone Rahmen von 125 μs Länge (Integration von ATM- und STM-basierten Daten, Kompatibilität mit vorhandenem Equipment)
- Grundrate + Vielfache davon Unterstützung unterschiedlicher Übertragungsgesachwindigkeiten

SONET-Architektur:

- Symmetrisch: Endgerät $\rightarrow$ Multiplexer $\rightarrow$ Regeneratoren ...
- Sektion: Teil einer physikalischen Verbindung OHNE Signalregeneratoren
- Leitung: Sequenz von Sektionen
- Pfad: Verbindet zwei Endpunkte eines SONET-Netzes
- Logische SONET-Schichtenarchitektur:
  - physikalische Schicht: Wandlung zwischen Rahmen und elektrischen/optischen Signalen, Übertragung von bits
  - Sektions-Schicht (Sektions-Zusatzinfo): Übertragung von SONET-Rahmen, Rahmenerkennung
  - Leitungs-Schicht (Leitungs-Zusatzinfo): Rahmensynchronisation, Multiplexen von Daten
  - Pfad-Schicht (Pfad-Zusatzinfo): Ende-zu-Ende-Transport von Daten
- Basisstruktur: SONET/SDH-Rahmen

- Tabelle (90 Spalten (3/87) und 9 Reihen (3/6)): Länge 125 μs (=8000Rahmen/s). Rahmen enthalten Nutzlast und Zusatzinfo zur Kontrolle und Synchronisation
- SONET-Rahmen: ersten 3 Spalten Transportzusatzinfo, im Container (87 Spalten) 1 Spalte Pfad-Zusatzinfo, ersten 3 Reihen: Sektions-Zusatzinfo, nächsten 6 Reihen: Leitungs-Zusatzinfo. 1 Element = 1 Byte!! also je STS-1 (Synchronous Transfer Signal Level 1; Basisrahmen) Rahmen 810 Bytes!!. Datenrate 51,84 Mbit/s; Nutzdaten(744 Bytes pro Rahmen)rate: 49,536 Mbit/s.
- Sektions-Zusatzinfo:
  * (pro Kanalbyte 64kbit/s) Rahmenerkennung (2 Byte), Erkennung geMUXter STS-1 Rahmen im STS-N Rahmen, Parität (vorrausgegangener STS-N Rahmen!!), Sprachkanal (Wartung), Sprachkanal (Benutzer), Datenkommunikation (3 Bytes)
- Leitungs-Zusatzinfo:
  * Zeiger bzw. Zeigeraktion (3 Byte), Parität (Leitungs-Zusatzinfo und Nutzdaten des vorrausgeganenen STS-N Rahmens!!), 2 Bytes Signalisierung zwischen Multiplexern (z.B. Fehler), 9 Bytes Datenkommunikation (576 kbit/s (MUX-Management)), 2 Bytes reserviert, 1 Byte Sprachkanal (MUX-Management)
- Pfad-Zusatzinfo:
  * Funktionsüberprüfnungskanal, Parität (Nutzdaten vorrausgegangener STS-N Rahmen!!), Aufbau/Inhalt STS-Nutzdaten (z.B. ATM oder DQDB), Rücktransport von Zustandsinfo über entfernten Pfad-Endpunkt, Benutzerkanal, H4 abhängig von Nutzung (Aufbau/Inhalt s.o.), 3 reservierte Bytes.
- Lokalisierung des Nutzdatenbereichs: (über Zeiger (H1, H2 (= Anzahl der zwischen H3 (=nächstes Byte) und J1 (Funktionsüberprüfung, 1. Byte von Pfad-Zusatzinfo, = direkt an Nutzdatenblock) in Leitungs-Zusatzinfo $\rightarrow$ direkter Zugriff (ohne komplexe MUX/DEMUX-Vorgänge)), Nutzdatenbereich kann 2 STS-1 Rahmen in Anspruch nehmen, H3 als Stopfbyte!))
- Kompensation von geringen Ratenschwankungen: positive(Nutzdatenrate NIEDRIGER als Rahmenrate, Zeigerwert erhöhen, Stopfbyte nach H3)/negative(Nutzdatenrate HÖHER als Rahmendatenrate, H3-Feld: Extrabyte für Rahmen aufnehmen, Zeiger um 1 Byte erniedrigen)Zeigeranpassung
- Daten werden byteweise in STS-N-Rahmen geMUXt. (also N*3 Spalten Transport-Zusatzinfo und N mal Pfad-Zusatzinfo): Zusatzinfo N-mal vorhanden $\rightarrow$ direkter Nutzdatenzugriff möglich.

1.3 B-ISDN

siehe Tele1 (Out of band Signalisierung)

1.4 ATM

Ziele:

- Hohe Flexibilität, Unterstützung unterschiedlicher Bitraten.
- Hohe Datenraten: Komplexität in Endsysteme verlagert, Einfache Vermittlung

Charakteristika:

- Minimale Funktionalität im Netz (Fehler- und Flusskontrolle Ende-zu-Ende, NICHT Link-zu-Link (wie bei X.25!!); Keine Übertragungswiederholung auf Links (Ressourcenreservierung und geeignete Dimensionierung von Warteschlangen))
- verbindungsorientiert! (virtuelle Verbindungen) (Zugehörigkeit der Zellen zu Verbindung im Zellkopf vermerkt, effiziente Vermittlung; Definition eines hierarchischen Verbindungskonzepts)
- Zellen
- Asynchrones Multiplexen
- Übertragung: sychron oder asynchron
- Vermittlung über lokal eindeutige Köpfe

Verbindungskonzept:

- Bündeln zusammengehöriger Datenströme zwischen ATM-Stationen
- Übertragungspfad besteht aus virtuellen Pfaden (virtual path, VP (Bündel virtueller Kanäle mit **gleichen** Endpunkten)), diese beinhalten mehrere virtuelle Kanäle (virtual channels, VC)
- Zurodnung im Zellkopf der ATM-Zelle (feste Länge: 53 bytes (5 Zellkopf, 48 Daten)) durch Identifier: VCI und VPI. Zellkpf enthält Routinginformation!!

Aufbau einer ATM-Zelle:

- (bit): GFC (Generic flow control) (4), VPI (8), VCI (16), PT (3) (payload type), CLP (1) (Cell Loss Priority), HEC (8)
- Obiger Rahmen für UNI (User-to-Network Interface). Bei NNI (Network-to-Network Interface) wird GFC von VPI benutzt (also VPI (12))
- GFC: Kontrolle der Daten an UNI
- VPI,VCI: Dienen der Vermittlung
- PT: Nutzdaten (EFCI-Bit: Stau erfahren, kein Stau erfahren), OAM-Zelle, Zellen zum Ressourcenmanagement
- CLP: Bei Stau Verwerfen von niederprioren Zellen
- HEC: Beheben von 1-Bitfehlern, Erkennen von Mehrbitfehlern und Auffinden von Zellgrenzen

ATM-Zellen werden kontinuierlich gesendet (Leerzellen, wenn nix da).
Übertragung von ATM-Zellen:

- unabhängig von eingesetzter Übertragungstechnik
- sowohl asynchron (Zellstrom) als auch synchron (Integration ATM, STM; Kompatibilität mit vorhandenem Equipment und Signalisierung; Verwendung SONET/SDH)
- synchron: Übertragung in STS-3c-Rahmen (270 Spalten * 9 Reihen, 9 Spalten Transport-Zusatzinfo, 261 Spalten Nutzdaten + Pfad-Zusatzinfos (alles 3mal, da STS-3c-Rahmen)). ATM-Zellen können in synchronen Rahmen übertragen werden. Im Nutzdatenbereich, haben aber keine feste Position und können über Rahmengrenzen hinweggehen. H4-Feld beschreibt Abstand der nächsten Zelle vom H4-Feld.

- asynchron: Kontinuierlicher Zellstrom, Cell Rate Decoupling (Leerzellen), zusätzliche Zellen notwendig (Ablauf, Wartung: OAM-Zellen). OAM-Zellen können durch feste Position (z.B. jede 27. Zelle) Rahmenstruktur generieren.

Vermittlung von ATM-Zellen:

- Routinginfo im Zellkopf. (rel. kurze Kennung; **lokal** eindeutige Kennungen)
- Tabelle im Vermittlungsknoten (Eingang, Ausgang, neue Kennung für Kopf)
- Virtueller Kanal:
  - lokal gültig (=zwischen Vermittlungssystemen)
  - VCI wird in jedem Vermittlungssystem umgesetzt, welches VC vermittelt.
  - VC wird durch Folge von VPI/VCI-Kombinationen identifiziert: (1/3,2/4,7/1)
  - Detailliert: VCC (Virtual Channel Connection (Endsystem-zu-Endsystem)) besteht aus: VCL (Virtual Channel Link): Strecken gleicher VCI
- Virtueller Pfad:
  - lokale Gültigkeit (jeweils für Übertragungsabschnitt (kann also Vermittlungssystem(e) enthalten))
  - Identifikation durch Folge von VPI (**VCI** innerhalb VP bleibt **konstant**)
  - Vermittlungssysteme innerhalb VP setzen nur VPI um.
  - Detailliert: VPC (Virtual Path Connection (Endpunte eines virtuellen Pfades (VCI gleich))) besteht aus VPL (Virtual Path Link).
- zwei Vermittlungsknoten:
  - VP-Vermittlungsknoten: Unterschiedliche Anzahl virtueller Kanäle pro virtuellem Pfad möglich. (Umsetzen der VPI und Permutation der "Pfadschläuche (incl. "Kanaladern")")
  - VP/VC-Vermittlungsknoten: Auswerten un Umsetzen von VPI und VCI. (Zusätzlich Umsetzen der "Kanaladern". VCs aus einem VP können auf verschiedene VPs verteilt werden.)
- Umsetzen der Kennungen: Entfernen der Leerzellen; Austauschen des Zellkopfes; Tabellen für gesamte Dauer der Verbindung aktuell. Aufbau der Tabellen durch Signalisierung bzw. Netzwerkmanagement.
- permanente Vermittlung (PVC: Permanent Virtual Circuit): durch externe Mechanismen etabliert (Netzwerkmanagement (Managementebene)), VPIs und VCIs manuell vergeben (wenig flexibel, zeitaufwendig)
- vermittelte Verbindung (SVC: Switched Virtual Circuit): automatischer Aufbau (Tabelle) während Signalisierungsphase durch Signalisierungsprotokoll (Kontrollebene), standardisierte Signalisierungsprotokolle (reservierter Kanal: VPI=0, VCI=5)
- ATM-Cross-Connect: Vermittlungsschaltung innerhalb Koppelelement nur durch PVC (Managementebene) (für VP-Vermittlung (Backbones großer Netzwerke))
- ATM-Switch: Vermittlungsschaltung innerhalb Koppelelement vorrangig SVCs $\rightarrow$ VP/VC-Vermittlung

- Aufbau ATM-Switch: Kommende Leitungsabschlüsse; Koppelelement unter Steuerung; gehende Leitungsabschlüsse.
  - Klassifizierung nach Koppelelement (Koppelmatrix *(jeder Eingang über Matrix mit jedem Ausgang verbunden; parallele Vermittlung möglich; Blockierungen! (Zwischenpuffer nötig); hoher Verdrahtungsaufwand; auch mehrstufige Matrizen möglich (3-stufiges Banyan-Netzwerk), interne Blockierungen möglich)*, Bus-Struktur *(konfliktfreier Zugriff (Zeitmultiplex); Übertragungskapazität (Summe Eingangskapazitäten); erfordert parallele Bitübertragung auf Bus; kleine räumliche Ausdehnung (nur bis 16 Anschlüsse); Multicast- und Broascast-Unterstützung)*, Ring-Struktur *(konfliktfreier Zugriff duer Zeitmultiplex (slottet Ring); bei statischer Zuordnung Ü-Kapazität so wie bei Bus; Mehrfachbelegung von Zeitslots möglich (unter zusätzlichem Aufwand für Slotfreigabe))*, Gemeinsamer Speicher *(Alle Leitungen mit Gem.Sp. verbunden; deutlich kleiner M; Z=1/2N; Steuerung für parallelen Zugriff nötig)*) und Platzierung von Speicherelementen (zur Konfliktauflösung, Sequentialisierung der Vermittlung) (Eingangsspeicher *(S=1; Z=1/2 (Lesen und Schreiben), FIFO-Puffer, Koppelelement intern blockierungsfrei, Abarbeitungsstrategie: (round-Robin, prio.-gestützt, je nach Pufferfüllung...); Nachteil: HOB (Head-of-Line Blocking), max.Durchsatz unter 75%)*, Ausgangsspeicher *(kleiner Eingangspuffer; Koppelelement blockierungsfrei (intern); S=N, Z=1/(N+1); Max.Durchsatz nahe 100%)*, verteilter Speicher *(Matrix aus Puffern (je eine Eingangsleitung pro Zeile und eine Ausgangsleitung pro Spalte); Z=1/2; S=1; Nachteil: großes M, nur bei Matrixstruktur möglich; Vorteil: KEIN HOB)*)
  - Bewertung der Architekturen: Gesamtspeicherkapazität M, Geschwindigkeitserhöhung (Eingang, Ausgang) S (Faktor), Zykluszeit (Übertragungsdauer einer Zelle) Z.
- Datenfluss in ATM-Netzen: netzinterne Vermittlungsknoten (nur PHY- und ATM-Schicht). Vermittlung auf ATM-Ebene
- PHY-Konvergenzschicht: 3 Zellentypen (Leer, Ungültig, Gültig) (Nur gültige und u.U. Leerzellen werden weitergeschickt)
- PHY-Kontrolle von Fehlern in Zellköpfen: HEC (Polynom: 8,2,1,0); bei Mehrbitfehlern → Verwerfen; Automat: Zustände auf Empfyngsseite: Korrigieren (Einzel- korr., Mehr- erkennen), Erkennen: (Verwerfen aller(!) Zellen mit Fehlern)
- PHY-Erkennen dr Zellbegrenzungen: Automat (HUNT(bitweise), PRE-SYNCH (zellenweise; weil kein Bitstopfne wie bei HDLC), SYNCH); Bei zellbasierter Übertragung: PRE-SYNCH: 8 korrekte HEC → SYNCH; SYNCH 7 inkorrekte HEC → HUNT

DQDB Distributed Queue Dual Bus
SMDS Switched Multi-Megabit Data Service
PVS Permanent virtual Circuit
SVS Switched Virtual Circuit (Verbindungstypen)
HOB Head of Line Blocking (Bei Koppelelement mit Eingangsspeicher) ATM-Schicht: Dienste:

- Kategorien: CBR, rt-VBR, nrt-VBR(burstartig Verkehr, niedrige Fehlerrate), UBR, ABR
- ABR: typisch: Filetransfer: Jitter nicht definiert, keine Echtzeit-Anwendungen, geringe Zellverlustrate gefordert.
- Bandbreitenverteilung: (ABR(UBR(VBR(CBR))))

AAL-Schicht:

- Diensttypen basierend auf: Bitrate (konstant, variabel); Zeitbeziehung (Sender, Empfänger) (mit/ohne); Verbindungsmodus (-orientier/-los)
- Unterteilt in Convergence Sublayer (CS) und SAR. CS bei AAL-3/4 und AAL 5 in CPCS (Common Part CS) und SSCS (Service Specific CS (Über CPCS)) unterteilt.
- AAL1:
  - Eigenschaften: Dienst mit konst. Bitrate (CBR). CS-Schicht: Abbildung zwischen AAL-SDUs auf 47-Byte-Blöcke (wegen 1 Byte SAR-Kopf (CS Indic.(1bit), SN (3), CRC (3), Parität(1))). Anzeigen von Fehlern; SAR-Schicht: Einfügen SN (Seq.Nr.), Absicherung durch CRC und Parität.
  - Datenformat: (ATM-Zellkopf, SAR-Kopf (s.o.: CRC (3,1,0)), SAR-Nutzlast) Parität erkenn Fehler in CRC. CSI: übermittlung von Taktinformation
- AAl2:
  - Eigenschaften: Anwendungsbereich: Wireless ATM. MUX mehrerer Anwendungsverbindungen auf eine AAL2-Verbindung; Keine SAR-Schicht (Benutzerdateneinheiten wesentlich kleiner als ATM-Zellen); CS-Schicht: Benutzerdaten in CPS-Paketen (Common Part Sublayer), Verteilt CPS-Pakete auf ATM-Zellen-Nutzdatenbereich STF: Start-Feld (Anordnung der CPS Packete in ATM-Zellen) (Wireless ATM)
  - Datenformate: CPS-Paketkopf (Channel Identifier, Länge, Payload Types (identifiziert: Nutzdaten oder Management), User-to-User-Information (zwischen AAL-2 Instanzen), HEC), Nutzlast =: CPS-Paket; CPS-PDU in ATM-Nutzdaten: (Offset-Feld (Zeiger auf Beginn des CPS Packetes), Sequenz Nummer (pro CPS-PDU), Parität (hier Startfeld)),CPS-Pakete. OSF+SN+P = STF
- AAL3/4:
  - Eigenschaften: nrt-VBR; CS: CPCS immer, SSCS kann leer sein. Jede Teilschicht eigene Dateneinheiten, Funktionen: reihenfolgegetr. Auslieferung, Fehlererkennung/-anzeige, MUX auf ATM-Verbindung.
  - AAL3/4 vs. AAL5: AAL3/4 10-bit langes MID (MUX-Id) Feld → hoher Protokolloverhead, max. Nutzdatenrate 44/53 = 83%; AAL5: kein MUX nötig, weniger Protokolloverhead, Fehlererkennung durch höhere Protokolle, max. Nutzdatenrate 48/53 = 90,5%.
- AAl5: Datenformate: SSCS: Kopf-Benutzerdaten-Anhang (=A); CPCS: Nutzlast(1-65535bytes) (A)-PAD(0-47)-UU-CPI-Länge-CRC; SAR: PT=xx0-ATM-Nutzdaten, ... , PT=xx1-ATM-Nutzdaten.

SVC Signalling Virtual Channel
SAAL: Signalisierungs-AAL
SSCF Service Specific Coordination Function
SSCOP Service Specific Connection Oriented Protocol Signalisierung:

- Ressourcenresrevierung, Zuweisen von VPI/VCI-Werten, SVC, Sig.-protokoll Q.2931 (höhere Schicht der Kontrollebene), auf AAL-Schicht. SAAL
- SVC:

  - hierüber: Verbindungsaufbau/-abbau
  - 3 Arten von Signalisierungskanälen: Meta-Sig.kanal (VCI=1) (bidirektional, permanent), Punkt-zu-Punkt-Sk (VCI=5) (bidirektional, nur existent, wenn Knoten aktiv), Broadcast-Sk (VCI=3) (unidirektional)

- SAAL:
  - Protokoll Q.2931
  - SSCS-Teilschicht besteht aus: SSCF (Dienstabbildung (Q.2931, SSCOP)) und SSCOP (zuverlässig, verbindungsorientiert, Übertragung der Sig.-nachrichten)
  - CPCS und SAR wie bei AAL5

- Q.2931:
  - Ein Signalisierungskanal (mehrere Sig.-aktivitäten darüber möglich), muss vorgegeben und von Meta-Sig. aufgebaut sein.
  - Aufbau der Sig.-nachrichten: (Protokolldiskriminator, Verbindungsreferenz (welche Sig.-aktivität?), Nachrichtentyp (SETUP, CONNECT,...), Länge (Info.-elemente), m Informationselemente (müssen sich selbst durch Kennzeichnung identifizieren))
  - Nachrichten: lokale(l) oder globale(g) Relevanz: (Verb.aufbau: SETUP, CONNECT (g), CALL PROCEEDINGS(l)..; Verb.aufgabe: RELEASE(g), RELEASE COMPLETE(g); Sonstiges: bei Multicast bzgl Benutzer ADD, DROP)

Bei Infoelementen:
PCR Peak Cell Rate (Maximale Zellrate) (im Informationselement bei SAAL) (Kehrwert der minimalen Zwischenankunftszeit aufeinanderfolgender Zellen (Zellen/s))
SCR Sustainable Cell Rate (Mittlere Zellrate (gesamte Verbindungsdauer))
BT Burst Tolerance (Zeitnterval, in dem Sender mit PCR senden darf)
CDVT Cell Delay Variation Tolerance
MCR Minimum Cell Rate
GCRA Generic Cell Rate Algorithm (TAT: theor. Ankuntszeit. Prüft, ob Zelle konform (wenn TAT $\leq$ Ankuntfszeit + CDVT), erhöht dann TAT um 1/PCR)
CTD Cell Transfer Delay (ausgehandelt) (Ende-zu-Ende-Verzögerung)
CDV Cell Delay Variation (ausgeh.) (Jitter)
CLR Cell Loss Ratio (ausgeh.)
CER Cell Error Ratio (n.ausgeh.)
SECBR Severely Errored Cell Block Ratio (n.ausgeh.)
CMR Cell Misinsertion Rate (n.ausgeh.) (unentdeckte Fehler in Zellkopf)

Dienstgarantien vs. Ressourcenauslastung (Garantien durch Reservierung angeboten, bei burstartigen Verkehr $\rightarrow$ schlechte Ressourcenauslastung) bessere Ressourcenauslastung durch: statisches ZeitMUX (wenn nicht alle gleichzeitig mit PCR senden) $\rightarrow$ reduzierte Ressourcenreservierungen $\rightarrow$ bessere Ressourcenauslastung.
Verkehr:

- Verkehrsvertrag. beinh. Verkehrsparam. (PCR, SCR, BT, CDVT (an UNI), MCR), und Qualitätsparam (CTD, CDV, CLR, CER, SECBR, CMR)

Verkehrskontrolle: (einfach (in RT), effektiv (schnell), fair)
UPC User Parameter Control (Nutzung)
NPC Network Parameter Control (Nutzung)

CAC	Call Admission Control (Zugang)
CLP	Cell Loss Priority

- Verteilung Verkehrskontrolle: Traffi-Shaper an Quelle-priv.Netz, UPC an UNI (priv.Netz - ATM-WAN), CAC auf NPC an NNI, CAC an Senke an UNI.
- Zugangskontrolle: wenn auf SCR basiert → Stau möglich (bei zu langen Bursts). Lässt Verbindung nur zu, wenn bestehende Garantien nicht verletzt werden.
- Nutzungskontrolle: Überprüft VPI/VCI, beobachtet Verkehrsaufkommen. Aktionen bei nichtkonformen Zellen: Verwerfen (nur im Notfall), Kennzeichnen (Tagging mit Hilfe des CLP-Bits, erst bei Stau verwerfen), Verzögern (problematisch bei höherer Verzögerung)
  - Leaky-Bucket: Markengenerator (konst.), Markenpool, nur Senden, wenn noch Marken im Pool, sonst verwerfen (harte Kontrolle), KEINE gleichzeitige Kontrolle von mittlerer und maximaler Zellrate , Wahl von Markenrate zwischen mittlerer und maximaler Zellrate
  - Virtueller Leaky-Bucket: 2 Markenpools (rot und grün), Zelle kriegt rot (CLP-Bit=1), wenn keine grünen mehr da. Bei Überlast, rote Zellen verwerfen. Weiche Kontrolle (Kennzeichnen statt Verwerfen). Problem: Priorität nicht abhängig von Relevanz der Daten
  - Dual Leaky-Bucket: 2 Markengeneratoren (unterschiedliche Raten), 2 Pools (unterschiedlich groß): 1.Pool: Begrenzung PCR (kleiner POOL $\sim$ CDVT, hohe Rate (entspricht PCR)), 2.Pool: Einhaltung SCR und Begrenzung BT (großer Pool $\sim$ BT, Rate entsprechend SCR)
  - fensterbasierte Mechanismen: während Zeitintervall T dürfen n Zellen gesendet werden (zusätzliche verwerfen oder kennzeichnen). Mehrere Varianten:
    - Jumping-Window: nach Ablauf T neues Fenster: nur SCR begrenzt, keine gleichzeitige Begrenzung PCR
    - Moving-Window: T wird kontinuierlich verschoben, bei Zellenankunft Zähler+1, nach T: Zähler-1. Begrenzt auch PCR.

ABR-Flusssteuerung (Staukontrolle, ratenbasiert): (Feedback von Switches)

EFCI-Bit	Explicit Forward Congestion Indication
RM-Zellen	Ressource Management
BECN	Backward Explicit Congestion Notification
PRCA	Proportional Rate Control Algorithm
EPRCA	Enhanced PRCA (bessere Fairness)

Staukontrolle:

- Prioritätskontrolle:
  - Priorität pro Zelle: Durch Benutzer (=Relevanz), durch Verkehrskontrolle (=Konformität). Prio. evtl. auch pro Verbindung.
  - Entscheidungsbasis für: Verwerfen, Verdrängen aus Puffern (!!Prob: reihefolgetreue), Best. der Bearbeitungsreihenfolge
- Reaktive Staukontrolle:
  - Problem: Reaktionszeit. Wichtiger Einsatz: ABR-Verkehr

  - Kreditbasierte Verfahren: zwischen benachbarten ATM-Knoten. Kredit an Vorgänger basiert auf freiem Pufferplatz. Kann auch auf gesamten Link (DigitalsFlowMaster) angewendet werden.
  - Ratenbasierte Verfahren: Ende-zu-Ende-Kontrollverfahren. Feedback Vermittlungssysteme. Sender nutzt Feedback → Anpassen Datenrate.
  - Kreditbasierte vs. Ratenbasierte: Ratenbasierte haben sich durchgesetzt, obwohl Datenverlust möglich und längere Reaktionszeit und mehr Speicheraufwand (dafür einfacher zu implementieren (in Switches))

- Ratenbasierte Staukontrolle ABR-Flussteuerung:
  - EFCI-Bit (in PT-Feld der Datenzellen) =1 bei Stau.
  - RM-Zellen: Ressourcenmanagement (Management: PT-Feld-Kennung)
  - 3 verschiedene Feedbacks:
    * negativ: Stau: Rate erniedrigen; Default: Rate erhöhen
    * positiv: Kein Stau: Rate erhöhen, Default: Rate erniedrigen
    * explizit: Feedback-Info gibt erlaubte Senderate an.
  - ABR-Ende-zu-Ende-Kontrolle:
    * Virtuelle Ziele und Quellen (zwischen Subsystemen) verkürzen Kontrollschleife (=kürzere Reaktionszeiten)
    * ABR-Kontrollschleife muss mit höherschichtigen Kontrollschleifen (z.B. TCP) effizient zusammenarbeiten.
  - Verfahren:
    * EFCI: Sender EFCI=0 → Switch(Stau): EFCI=1 → Ziel kennt Stau, aber nicht WO? → Ziel: periodisch negatives RM zu Quelle → Sender reduziert Datenaufkommen exponentiell (PROBLEM: RM-Zelle im Stau → Kollaps des Systems)
    * BECN: Stauinfo direkt vom Switch an Sender → kürzere Reaktionszeit; aufwendigere Implementierung in den Switches.
    * PRCA: positives Feedback; Sender reduziert Rate bis RM-Zelle mit Feedback eintrifft, dann Erhöhen der Rate, RM-Zellen proportional zur Senderate (Schlecht für Fairness für Verbindung mit langen Wegen und niedriger Datenrate)
    * EPRCA: Sender in Senderichtung proportional zur Datenrate RM-Zellen (expliziter Ratenwunsch (ER)); Switch wertet aus (→ Bestimmung der Lastverteilung); Empfänger REFLEKTIERT; Switch reduziert ER-Wert bei Stau; Sender passt sich an. (bessere Fairness, da jede Verbindung einzeln betrachtet.)

1.5 Internetworking mit ATM

Unterschiede traditionelles LAN - ATM-basierte Netze:

- Dienstabbildung: verbindungslos ⇄ verbindungsorientiert
- Adressumsetzung: IP(32bit, hierarchisch (Netzwerkkennung, Endsystemkennung), identifiziertr benutzer der Vermittlungsschicht)- bzw. MAC-Adressen(16 oder 48bit, herstellerabhängig, flach(keine Lokalisationsinformation)) ⇄ ATM-Adressen(20 Byte, 3 hierarchische Formate (E.164, DCC, ICD))

- Nachbildung der Medienfähigkeiten: Rundruftätigkeit traditioneller lokaler Nezte

ATM-Adressierung:
DCC Data Country Code
ICD International Code Designator
AFI Authority and Format Identifier
SEL Selektor
Bei E.164: AFI- E.164 (ISDN-Rufnummer)-Routing Domain-Area-ESI (Endsystem-ID innerhalb Area)-SEL
Adressierung/Adressauflösung bei traditionellen LANs: ARP/RARP (setzt Rundruftätigkeit mit gemeinsamen Medium vorraus)
LIS Logisches IP-Subnetz

Lösungsansätze: IP über ATM (ATM als Sicherungsschicht, alternatives Netzwerkschicht-protokoll benötigt neue Konvergenzschicht) und LAN-Emulation (ATM als MAC-Schicht, alternatives Netzwerkschichtprotokoll über LLC möglich); Ziele: transparentes ATM-Netz, Schutz getätigter Investitionen; Probleme: Grundlegend unterschiedliche Eigenschaften der Netzwerk-typen, Vielzahl unterschiedlicher Kommunikationsprotokolle; Lösungen: Client/Server-basierte Ansätze (bei LAN-E: MPOA als Weiterentwicklung)
IP über ATM:

- spezifiziert ein logisches IP-Subnetz (LIS)
- lokales ATM-Netz mit eindeutiger IP-Netzadresse, Verbindung der logischen Subnetze über IP-Router, Stationen müssen Adressen abbilden können (ATMARP-Protokoll)
- ATM-Adressauflösung:
  - ATMARP-Server verwaltet innerhalb LIS globale Adressauflösungstabelle (jede Station kennt ATM-Adresse des ATMARP-Servers, bidirektionale Verbindung, Anfrage an ATMARP-Server, falls lokale Adresstabelle kein Eintrag für gewünschte IP-Adresse enthält (KEIN Broadcast)): Client/Server-Modell
- ATMARP-Protokoll:
  - 2 Phasen: **Registrierphase** (Client baut bidirektionale Verbindung auf (SETUP); Server → InATMARP-Request; Client → InATMARP-Reply; Server trägt Abbildung in Tabelle ein.), **Adressauflösung** (Jeder Client verwaltet lokale Adresstabelle. Ist lokale Auflösung nicht möglich → Anfrage an Server; Antwort entweder ATMARP-Reply (→ gesuchte ATM-Adresse) oder ATMARP-NAK (Abbildung konnte nicht durchgeführt werden))
- ATMARP-Server:
  - Adresstabelleneinträge: (IP-Adr., ATM-Adr., Zeitmarke)
  - Zeitmarke = Wann Eintrag. Einträge 20 Minuten gültig (bei Clienttabellen: 15 Minuten)
  - Vor Löschen eines Eintrages → InATMARP-Request an Client; bei Antwort → neue Zeitmarke.

NHRP Next Hop Resolution Protocol:

- Gliederung des Netzes in LIS (Problem: Inter-LIS Verkehr NICHT über ATM, sondern über IP-Router → erhöhte Verzögerung (da mehrere Knoten durchlaufen); NHRP ermöglicht direkte Verbindung, auch wenn Endsysteme in verschiedenen LIS sind)

- Protokollübersicht
  - Registrieren:
    - (NHC = nutzt NHRP-Dienst; NHS = bietet NHRC-Dienst): NHRP-Clients registrieren bei einem NHRP-Server (NHRP Registration Request & NHRP Registration Reply)
  - Adressauflösung:
    - NHS sendet NHRP Resolution Request an NHS, um des Ziels ATM-Adresse zu bestimmen
    - Kann dieser nicht auflösen → Anfrage in Richtung IP-Zieladresse weiterschicken
    - **Jeder** NHS auf Weg prüft, ob Adresse aufl"osen kann
    - NHRP Resolution Request wird zwischen NHS weitergeleitet; NHRP REsolution Reply mit ATM-Zieladresse
  - Direktes Senden zwischen Sender und Ziel (ist vorher bereits indirekt möglich (über Router, LIS-zu-LIS))
  - Koexistenz von ATMARP und NHRP → Migration zu NHRP angestrebt

LEC	LAN-Emulation Configuration Server
LE	LAN-Emulation Server
BUS	Broadcast-and-Unknown-Server

LAN-Emulation (LANE):

- Spezifiziert Emulation zweier Netztypen (Ethernet (802.3) und Token Ring (802.5))
- Cient/Server-Architektur.
  - Client-Komponente: Pro Rechner und pro emuliertem LAN (ELAN) ein LE(LAN Emulation)-Client
  - Server-Komponente: LE-Service zerfällt in drei Teile (LEC-Server, LE-Server und BUS), die zentral oder verteilt sein können.
- LE-Client:
  - LANE-Modul stellt MAC-Schnittstelle zur Verfügung
  - unterhält zur Umsetzung (MAC- ATM-Adr.) eine lokale Tabelle. Diese wird durch Anfragen beim LE-Server aktualisiert.
  - Funktionalität entweder in Hardware (ATM-Adapter) oder Treibersoftware realisiert.
    - Systemstart
    - vierschiedene Initialisierungsphasen (Aufbau der ATM-Verbindungen zwischen LANE-Komonenten), Konfiguration (ATM-Adressen, MAC-Adressen, LAN-Typ, ..)
    - Registrierung beim BUS
    - Datentransfer
    - Abbau der ATM-Verbindungen erfolgt timergesteuert
- LANE-Konfiguration:
  - LEC-Server: Datenbank mit Konfigurationsdaten der ELANs (wird durch Admin initialisiert und über SNMP gewartet)

  - Configuration Direct (bidirektionale verbindung zwischen LE-Client und LEC-Serv., dessen ATM-Adr. der LE-Client während Initialisierung erhält)
  - LE-Client fordert Konfigurationsparameter für das ELAN an

- Registrierung beim LE-Server.
  - jedes ELAn genau einen LE-Server (zuständig für Auflösung der MAC-Adressen)
  - Control Direct (bidirektionale Verbindung zwischen LE-C und LE-S; optional: Control Distribute (Punkt-zu-Mehrpunkt an alle LE-Cs))
  - Registrierung: Senden von "Join Request" zum LE-S (Param.: eigene ATM-Adr. und MAC-Adr.) → Server kann später Adressauflösung durchführen

- LANE-BUS:
  - jedes ELAN genau ein BUS (zuständig für Übermittlung von Multicast- und Broadcast-Nachrichten ("Multicast-Server"))
  - LE-C fordet am LE-S ATM-Adr(BUS) an
  - Multicast Send: Bidirektionale Punkt-zu-Punkt-Verbindung zu BUS zur Übermittlung von Multicast-Nachrichten (von LE-C aus)
  - Multicast Forward: Punkt-zu-Mehrpunkt-Verbindung zu allen LE-Cs zur Verteilung der Multicast-Nachricht (von BUS aus!!)

- LANE-Adressauflösung:
  - LE-C kriegt MAC-Adr. → lokal bekannt?, wenn ja: Abbildung, wenn nein → Server fragen → dort bekannt? wenn ja: Antworten mit gesuchter ATM-Adr, wenn nein: Anfrage an alle LE-Cs ud auf Antwort warten (NAK bei timeout)

- LANE. Unicast-Datentransfer:
  - Abbildung (MAC-Adr. → ATM-Adr.)
  - Während Auflösung kann Sender bereits Daten über BUS senden
  - Sobald ATM-Adr. bekannt: Direkte Verbindung aufbauen bzw. bereits Existierende nutzen

- LANE: Multicast- und Broadcast-Datentransfer
  - Multicast Send → Multicast Forward
  - verschiedene Multicast Nachrichten können nicht zerzahnt werden (nur rein sequentiell)

- Vorteile: sanfte Migration möglich, LANE: Unabhängigkeit gegenüber höheren Protokollen

- Nachteile: viele Leistungsmerkmale von ATM können nicht genutzt werden, keine Dienstgüteunterstützung, keine Ressourcenreservierung, Effizienzeinbüßungen durch mehr Overhead (Adressauflösung, Rundrufnachbildung), über NHRP nur bedingt für große Netze geeignet.

MPOA Multi Protocols over ATM:

- Ziele: effizienter Datentransfer über Subnetze hinweg (ohne Router!), Rahmenwerk für diverse Vermittlungsschichtprotokolle über ATM

- Integration von LANE und NHRP

- Konzepte: virtueller Router trennt Routing und Forwarding; Shortcuts (direkte Verbindungen) zwischen IP-Subnetzen
- Komponenten:
  - MPOA-Client (MPOA-C): definiert Randgeräte (Edge devices (Ethernet-Switches oder Brücken mit ATM-Anschluss und LANE sowie ATM-Endgeräte)), enthält KEINE Routinginformation
  - MPOA-S: logische Komponente eines Routers, welche MPOA-Cs mit Forwarding-Infos versorgt; enthält NHS (Next Hop Server)
- Funktionen:
  - Konfiguration: siehe LANE-Konfig. am LE-S
  - Registrierung: MPOA-C und MPOA-S tauschen Parameter aus (ATM-Adr., Typ (C/S))
  - Adressauflösung: mehrere Pakete an gleiche Schicht-3-Adresse → MPOA Resolution Request (analog NHRP)
  - Datentransfer: über Defaultweg oder (nach Adressauflösung) Shortcut

1.6 ipv6

Probleme IPv4 (Netz zu groß, nicht genug Adressen (wegen A,B,C und deren Verschnitt)) siehe Telematik1ZF..
Neuerungen bei IPv6:

- Flexibles Paketformat (Vereinfachung und Verschiebung von Optionen in flexible Paketkopferweiterungen)
- Erweiterte Adressierung: 128bit Adressen, mehrere Hierarchieebenen, Anycast-Adressen (Kommunikation zu einem Mitglied einer Gruppe)
- Unterstützung von Ressourcenreservierung: Flow Label und Traffic Class pro IPv6-Paket, Nutzung von Ress.reserv.protokollen
- Erweiterung von ICMP: zuvor getrennte Protokolle nun integriert (IGMP, ARP)
- Neighbor Discovery (Adressauflösung, Erkennen nächster Router)
- Automatische Systemkonfiguration
- Unterstützung mobiler Systeme (Adresszuweisung durch autom.Systkonf.; Option "Binding Update" im Destination-Options-Header ermöglicht direkte Weiterleitung der IP-Pakete an aktuellen Standort)
- Berücksichtigung der Sicherheitsaspekte (Authentifizierung, Dienstintegrität)
- Paketformat:
  - Standardkopf (Version(4bit), Traffic Class(8), FlowLabel(20), Nutzlastlänge(16), Nächster Kopf(8), Hop Limit(8), SA(128), DA(128)), optionale Paketköpfe, Daten
  - Keine Kopflänge, KL-Prüfsumme etc.
- Paketkopferweiterungen:

  - unterschiedliche Erweiterungen: Knoten-zu-Knoten-Erweiterungen, Ziel-Optionen, Routing, Fragmentierung, Authentifizierung, Verschlüsselung
  - Jeder Erweiterungstyp darf höchstens nur EINMAL in einer Dateneinheit vorkommen
  - Wird im "Nächster Kopf"-Feld vermerkt (z.B. =Routing, =TCP etc.)

- IPv6 Adressierung:
  - 128 Bit
  - x:x:x:x:x:x:x:x, wobei jedes "x" 16-Bit-Hexa kodiert ist (also 0000 - FFFF)
  - zwischen :: wird mit "0:...:0" aufgefüllt, bis 8 Blöcke vorhanden.
  - Unterscheidung von drei Adressklassen: Unicast, Anycast, Multicast
  - Unterscheidung verschiedener Adresstypen: Link-Local Address, Site-Local Address, Aggregatable Global Unicast Address, IPv4, NSAP-, ...: Typ wird durch Format-Präfix (führende Bits einer IPv6-Adresse) festgelegt

ICMP Internet Control Message Protocol
IGMP Internet Group Message Protocol
Adressen etc.:

- Unicast-Adressen in IPv6:
  - Aggregatable Global Unicast Address (hierarchisch, global gültig im Internet, Felder fester Länge): Präfix 001. Format: 001 (3bit), TLA (Top Level Aggregator (13)), NLA (Next Level Aggregator (32)), SLA (Site Local Aggregator (16)), Interface-ID(64)
  - Link Local Address: nur innerhalb Subnetz gültig, Präfix: 1111.1110.10 und Kennung der Netzwerkkarte, dazwischen Nullen (aufgefüllt)
  - Site Local Address: nur innerhalb Netzwerk ohne Anschluss ans Internet gültig, Präfix: 1111.1110.11 + Kennung Subnetz + Netzkarte, dazwischen Nullen
  - für üUbertragung von IPv4 nach IPv6:
    * IPv4 Mapped Address (identif. IPv4-Rechner): 80*"0", 16*"1", IPv4-Adresse (32)
    * IPv4 Compatible Address (identif. IPv6-fähigen Rechner): 96*"0", IPv4-Adresse (32)
    * Ergeben gleiche Prüfsumme → keine Neuberechnung des TCP-Pseudo-Headers nötig

- Multicast-Adressen in IPv6:
  - identifizieren Menge von Netzkarten
  - spezielles Präfix "FF" (1111.1111).
  - Format: Präfix(8), Flags(4), Scope(4), Gruppenkennung(112)
  - zwei Typen: permanente und transiente Gruppen
  - Scope: Gültigkeitsbereich ("Wo?")
  - einige Multicast-Adressen bereits reserviert ("Alle im Subnetz" etc.)

- Anycast-Adressen in IPv6:
  - werden an (irgend)EIN Mitglied einer Anycast-Gruppe ausgeliefert (nicht an ein bestimmtes!!)

  - entsprechen syntaktisch Unicast-Adressen
  - z.B. zur Lokalisierung von Netz-Ressourcen genutzt

ND Neighbor Discovery
Neighbor Solicitation Address (Abbildung von IPv6 auf Schicht 2 Adressen)
ICMP:

- Koordination und Kontrolle
- ICMP-Nachrichten:
  - Fehlernachrichten
  - Informationsnachrichten (Ping, Verwaltung von Multicast Gruppen)
  - ND (Adressauflösung, Router-Erkennung, Präfix-Erkennung, Automatische Adresskonfiguration)

ND:

- Adressauflösung:
  - Problem: IPv6-Adresse → entspr. Schicht-2-Adresse
  - Lösung: Anfrage mittels Neighbor Solicitation Nachricht (kein Broadcast!)
    * Rechner hören auf SNA (Solicited Nodes Address FF02::1:FF00..) gefolgt von letzten 24bit der Unicast-Adresse
    * Anfrage an SNA des Zielrechners → Zielrechner übermittelt Schicht-2-Adresse an Nachfragenden
- Router-Erkennung:
  - Problem: Welcher Router bringt meine Pakete nach "draussen"?
  - Lösung: Router senden periodisch "Router Advertisement" an "All Hosts(FF02::1)" Diese erhalten: Router Lifetime (bis Info ungültig), Reachability Lifetime (bis ND-Info ungültig) und Prefix Information (Liste Präfixe des Subnetzes)
  - Router Advertisement auch per Router Solicitation anforderbar, dann per Unicast Antwort
- Präfix-Erkennung:
  - Problem: Zielrechner im eigenen Subnetz? Direktes Senden oder über Router?
  - Lösung: Router Advertisement → Prefix Information → logisches UND mit eigener Subnetz-Prefix → wenn identisch → direkt, sonst an Router

DHCP Dynamic Host Configuration Protocol
Automatische Systemkonfiguration:

- Problem: manueller Konfigurationsaufwand, keine Unterstützung für mobile Rechner und "Ad-Hoc"-LANs
- "Plug & Play" wünschenswert
- zwei Arten in IPv6: ND und DHCP
- durch ND:

  - Subnetzpräfix global eindeutig
  - Schicht-2-Adresse innerhalb Subnetz eindeutig
  - → global eindeutige IP-Adresse = Präfix (z.B. 4C00::0/80(Router Advertisement)) + Schicht-2-Adresse
  - wenn kein Router Advertisement gesendet, können Link-Local Adressen gebildet werden (LL-A + Schicht-2-A = IP)
  - Nachteile: korrekte Konf. der Router erforderlich, Feste Zuordnung von IP-Adressen zu Netzkarten
- durch DHCP:
  - Verbesserung: Client/Server Modell (DHCP-Server, DHCP-Relay (Funkschnittstelle))
  - drei Arten der Zuweisung: manuell, automatisch, dynamisch

SPI Security Parameter Index
MD5 Message Digest 5 (Authentifizierung, Schlüssel vorn und hinten, → 128-bit Digest)
ESP Encapsulating Security Payload
CBC Chipher Block Chaining
Sicherheitsaspekte:

- Authentifizierung und Verschlüsselung
- basieren auf Sicherheits-Assoziation (je ein Kontext bei Sender und Empfänger). Sie definiert: Verschlüsselungs-/Authentifikationsalgorithmen, zu verwendende Schlüssel, Lebenszeit dieser bzw. der Assoziation, Sicherheits-Level
- Authentifizierung:
  - Erweiterungskopf in IPv6:
    * Nächster Kopf (8), Länge (8), Reserviert (16), SPI (32), SN(32), Authentifierungsdaten (n*32)
    * SN schützt gegen Kopie
    * Authentifizierungsdaten abhängig von Algorithmus: MD5 (std)
  - MD5:
    * Authentifizierungsschlüssel von Sender (128bit, ggf mit Nullen aufgefüllt) sowohl an Anfang als auch ans Ende des IP-Paketes angehängt..
    * Resultat 128-Bit-Digest, von dem die ersten 96 Bit als Authentifizierungsdaten genutzt werden
    * Vorr. Sender und Empfänger kennen geheimen Schlüssel.
    * Auth. schützt nicht gegen Mithören
- Verschlüsselung:
  - nutzt den ESP-Kopf
  - Default-Algorithmus: CBC
  - zwei Möglichkeiten der Übertragung: Transport, Tunneln
  - Transport-Modus:

    - Nutzdaten + Ziel-Optionen-Erweiterungskopf verschlüsseln und in "Verschlüsselte Nutzdaten" des ESP eintragen
    - Nachteil: Analyse des Datenaufkommens durch Aussenstehende möglich
  - Tunnel-Modus:
    - gesamtes IP-Paket wird verschlüsselt und in unverschlüsseltes IP-Paket gekapselt

Migration: IPv4 → IPv6:

- Problem: Gleichzeitige Umstellug nicht möglich
- Lösung:
  1. Dual Stack (vereinzelte IPv6-Systeme)
  2. Tunneln (alternativ: automatisches Tunneln) (IPv6-Netze über IPv4-Netze verbinden)
  3. Übersetzen der Paketköpfe (überwiegend IPv6-Systeme)

1.7 Gruppenkommunikation

RPF Reverse Path Forwarding
RPF mit Pruning
CBT Core Base Trees
PIM Protocol Independent Multicast
DVMRP Distant Vector Multicast Routing Protocol
RIP Routing Information Protocol
Unicast 1:1
Multicast 1:n
Concast m:1
Multipeer m:n
Grundbegriffe im Kontext von Multicast:

- Übertragungsverzögerung (Zeitspanne zum Übertragen an ALLE Empfänger) auch interessant: mittlere Uverz.
- Ausführzeit (Zeitspanne von Aufruf bis Abschluss einer Multicast-Operation (=Ü-verz. + Ü-wdh + Bearb.-zeit))
- Bandbreitenbedarf: Mittelwert je Übertragungsstrecke (Link)
- Zuverlässigkeit:
  - Zuverlässigkeitsklassen: Multicast:(unzuverlässig, halb-zuverlässig:(statistisch-zuverlässig, k-zuverlässig), zuverlässig:(atomar-zuverlässig, all-zuverlässig, voll-zuverlässig, ...))

Gruppenmanagement: Ideal: Trennung von Gruppenmanagemen und Kommunikationsprotokoll. Anwendung spez. Gruppensemantik an Schnittstelle zum Gruppenmanagement. Dieses steuert dann entspr. Gruppensemantik die verschiedenen Protokollfunktionen.
Multicast-Unterstützung:

- in der Sicherungsschicht:

  - 1. Lösungsansatz: Senden von replizierten Unicast-Nachrichten an alle Gruppenmitglieder → schlecht Skalierbar: Netzlast und übertragungsverzögerung nehmen mit wachsender Gruppe zu.
  - 2. Lösungsansatz: Senden einer Multicast-Nachricht an ALLE angeschlossenen Systeme (Broadcast) → unfair (für Gruppennichtmitglieder (zusätzliche Belastung))
  - optimale Lösung: Senden einer Multicast-Nachricht NUR an alle Gruppenmitglieder (Schicht-2-Multicast-Adresse) → Definition von Multicast-Adressen (Auf Netzwerkadaptern müssen zusätzliche Filter für Multicast-Adressen definierbar sein)
  - → Multicast-Unterstützung in der Sicherungsschicht erforderlich.

- in der Vermittlungsschicht:
  - Ohne Multicast-Unterstützung: mehrere Unicast-Verbindungen, Duplizieren beim Sender
  - Mit Multicast-Unterstützung. einzelne Multicast-Verbindungen; Duplizieren im Router
  - Derzeit IP-Multicast einziger allg. verfügbarer Multicast-Dienst für Weitverkehrsnetze
  - Zugrundeliegendes Prinzip:
    * Sender → Datenpakete an Multicast-Adresse der Gruppe. Er benötigt keine Kenntnis über Gruppenmitgliedschaft
    * Empfänger treten bei (JOIN-Nachricht) und benötigen keine Kenntnisüber aktive Sender
    * Änderungen der Gruppenmitgliedschaft haben auf Übertragungen (an Rest der Gruppe) keine Auswirkungen

IGMP:

- Problem: Woher weiss Router, dass er Multicast-Pakete an angeschlossene Subnetze weiterleiten muss?
- Lösung: Multicast-Empfänger informieren "ihren" Multicast-Router über ihre Gruppenmitgliedschaft(en):
  - MC-Router → periodisch Query-Pakete an MC-Adr. "all hosts" (TTL aber 1)
  - Jeder Empfänger → ein Report-Paket pro Gruppenmitgliedschaft, in welchem MC-Gruppe enthalten ist
  - Vermeidung von Redundanz der Antworten durch zufälliges Verzögern und Unterdrücken der Antwort, fall während Verzögerung schon jemand aus derselben Gruppe antwortet.

Multicast-Routing:

- Aufgaben: Adressierung (Listenbasiert (jeder Empf. explizit angegeben) oder mittels Gruppenadresse (IP-Adr. der Klasse D)), Vermitteln der Datenpakete (Fluten, Spanning Trees, Reverse Path Forwarding, RPF mit Pruning, CBT, PIM)
- Fluten:
  - Jeder Router leitet an alle Anschlüsse (Ausser "Posteingang") weiter. Einschränkungen:
    * "Posteingang"

    - Kein Weitersenden von Duplikaten (Speichern der Kennung bereits empfangener Pakete nötig!)
  - Vorteile: robust, kaum fehleranfällig
  - Nachteile: speicheraufwendig, Belastet alle Verbindungsstrecken

- Spanning Trees.
  - Berechnung Spanning Tree gemäß Dijkstra-Algorithmus (geringste Kosten, iterativ)
  - Algorithmus: Multicast-Pakete an ALLE Anschlüsse weiterleiten (ausser "Posteingang"), die Teil des Spanning Trees sind.
  - Vorteil: robus, wenig komplex, geringer Speicherbedarf (1 bool pro Netzanschluss)
  - Nachteil: Datenverkehr konz. sich auf Spanning Tree, keine Berücksichtigung der Gruppenmitgliedschaft

- Reverse Path Forwarding:
  - nutzt normale Routing-Tabellen für Optimierung der Wegewahl
  - Algorithmus: Wenn lokaler Router nicht auf kürzestem Pfad zwishen Sender und benachbartem Router R liegt, dann leite Paket NICHT an R weiter.
  - $\rightarrow$ unterschiedliche Spaning Trees (in Abhängikeit vom Sender)
  - Vorteile: bessere Verteilung der Netzlast durch unterschiedliche Bäume; garantiert schnellstmögliche Auslieferung von MC-Paketen
  - Nachteil: keine Berücksichtigung der Gruppenmitgliedschaft

- Reverse Path Forwarding mit Pruning:
  - Ziel: Berücksichtigung der Gruppenmitgliedschaft (erstes Multicast-Paket per "flood"; Router ohne Mitgliedschaft leiten NICHT weiter und senden "prune"-Nachricht an vorangegangenen Router weiter)
  - Nachteile: Fluten; Aufbau eines Multicast-Baumes pro Gruppe-Sender-Paar

- Core Based Trees:
  - Pro Multicast-Gruppe ausgehend von einem Core-Router: Aufbau von MC-Baum
  - MC-Sender adressiert Pakete an Core-Router; Erster Router im MC-Baum setzt MC-Adresse als Zieladresse; MC-Pakete entlang des MC-Routingbaums weitergeleitet.
  - Vorteile: Nur *ein* MC-Routing-Baum pro Gruppe (kein Fluten!)
  - Nachteile: Hohes Datenaufkommen in der Nähe des Core-Routers

MC-Unterstützung in der Transportschicht:

- Verbindungsaufbau: (Aufbau einer Verbindung zu mehreren Empfängern, Unterstützung(Gruppenadressen, Namenslisten), Konfliktauflösung bei Aushandlung der Dienstgüte)
- Erweiterung der Dienstschnittstelle: zusätzliche Dienstelemente (so wie ADD und DROP) für Multicast-Kommunikation
- Fehlerbehandlung: Erweiterung Zuverläsigkeitsbegriff, Anpassung Quittungsbetrieb an Gruppe, Fehlerbehebung

- Gruppenverwaltung: Registrierung und ggf. Überprüfung

Beitritt zur Multicast-Verbindung:

- Sender-gesteuert:
  - S erhält AD-Request (von neuem Empfänger) durch übergeordnetes Protokoll → S übermittelt CONNECT-Paket an den Neuen
  - Geeignet für z.B. Konferenzsysteme mit zentraler Verwaltung
- Empfänger-gesteuert:
  - beitrittswilliger Benutzer fordert Beitritt durch JOIN-Request auf → Router schickt JOIN an alle Gruppenmitglieder → nur Sender reagiert auf Beitrittswunsch (Rest ignoriert Paket)
  - geeignet für: "Video on demand"

Sender-Implosion:

- Problem: Sender wird mit Rückmeldungen der Empfänger (Quittungen, Verkehrskontrollnachrichten) überschwemmt → Hohe Netzbelastung sowie hoher Puffer- und Bearbeitungsaufwand auf Senderseite
- Lösung: Lokale Bearbeitung der Quittungen/Kontrollnachrichten → Verteilung der Last, parallele Bearbeitung, Definition ausgez. Empfänger lokaler Quittungen

Fehlerkorrektur: (konventionell: Wdh durch Sender (bei sehr grossen Netzen schlecht), optimiert: lokale Wdh durch anderen Empfänger)
MBone (Multicast Backbone):

- Problem: Multicast erfordert Modifikation *aller* Router im Internet.
- Lösung: Definition eines "virtuellen" Multicast-Netzes im Internet; Tunneln (UC-Kopf an MC-Paket) von non-multicast-Routern; "Overlay"-Netz über das Internet (MBone ist experimentell!! und nicht als Betriebsnetzwerk ausgelegt)
- Leistungsengpässe im MBone, da Multicast-Router häufig auf PC-Basis und MC-Pfade oft nicht auf kürzestem Pfad zwischen Systemen liegen
- Konfiguration: mrouted (Software) (4 Parameter: IP-Adresse, Kosten Tunnel, Threshold (räuml. Einschränkung), Bandbreite)
- Einschränkung der Reichweite: Jedes IP-Paket: TTL-Feld; Jeder Ausgang mit threshold und metric (Kostenfaktor) versehen; TTL um Kostenfaktor reduzieren und Weiterleiten, wenn TTL positiv und threshold $<$ TTL
- Routing im MBone: DVMRP:
  - basiert auf RPF
  - zur Berechnung des RP wird das Distance Vector Multicast Routing Protocol verwendet.
  - Prinzip: Jeder Router: Tabelle (Ziel, Kosten, Anschluss); Router tauschen Informationen (Ziel, Kosten) mit direkten Nachbarn aus.
  - Unterschiede zu Unicast-Routing-Protokollen (z.B. RIP):
    - Handhabung von Tunneln
    - Ziel-Eintrag in Tabelle gibt potentiellen Sender an
  - Schlechte Skalierbarkeit bei großer Teilnehmerzahl

1.8 Neue Dienste im Internet

RSVP Resource Reservation Protocol
ST-II Internet Stream Protocol (IPv5)
ST2+: SCMP Stream Control Message Protocol Ressourcenmanagement:

- Aufgaben:
  - Kommunikationsressourcen (Speicher, Bandbreite der Links) → Ressourcen besitzen endliche Kapazität
  - Nutzung der Ressourcen (exklusiv (z.B. Prozessor) oder gemeinsam (z.B. Speicher))
  - Ressourcenmanagement: Reservierung und Zuteilung basierend auf Verkehrscharakteristika und Dienstanforderungen
  - Muss im verteilten System in allen involvierten End- und Zwischensystemen durchgeführt werden.
- Komponenten: Erweiterte Dienstschnittstelle; Ressourcenverwalter (RV), Reservierungsprotokoll
- Erweiterte Dienstschnittstelle:
  - dortige Interaktionen: Datentransfer, Kontrollfunktionen (Qualitätsverletzungen, Statusmeldungen...)
  - Ressourcenverwaltung vs. Dienstgarantie (s.o.)
- Ressourcenverwalter:
  - Aufgaben: Zugangskontrolltests beim Verbindungsaufbau (genug Kapazität?, fair?); Reservierung der Ressourcen beim Verb.aufbau; Zuteilung der Ressourcen beim Datentransfer; Freigabe danach.
  - Beispiel: Scheduler:
    - legt Reihenfolge der exklusiven Ressourcennutzung durch Scheduling-Algorithmus(FIFO(keine Garantien!), Earliest Deadline First, Hierarchisches Round Robin) fest.
    - die Scheduling-Algorithmen sind unterschiedlich komplex.
    - FIFO:
      - Wenn Puffer voll → neue pakete verwerfen.
      - Verantwortung für Staukontrolle an "Enden" des Netzwerkes verschoben.
      - Heute i.d.R. in IP-Routern;
      - Vorteil: einfache Implementierung;
      - Nachteil: keine Garantien
      - FIFO mit Prioritäten: Warteschlangen verschiedener Prioritäten. Innerhalb Warteschlangen: FIFO (höchste Prio-Warteschlange stets zuerst bedient)
    - Fair Queuing:
      - Grundidee: Jeder Verbindung erhält gleich viel Bandbreite
      - Pro Kommunikationsbeziehung separate Warteschlange. Diese per Round-Robin-Verfahren abarbeiten.
      - Vorteil: Verkehr von unfairer Quelle belastet andere nicht

        - Probleme: Vielzahl von Warteschlangen, Basisalgorithmus reicht nicht aus (z.B. bei unterschiedlicher Paketlänge)
        - Fair Queuing garantiert Mindestbandbreite
    * Weighted Fair Queuing:
        - einzelne Verbindungen erhalten unterschiedlich große Bandbreite. Aushandlung bei Verbindungsaufbau
        - Durch Zeitstempel (Virtual Clocks): 1. Paket → Zeitstempel (geplanter Absendezeitpunkt), jedes weitere Paket → Zeitstempel (alter Zeitstempel + (Paketgröße/Rate))
        - Sendereihenfolge: aufsteigende Zeitstempel
        - Vorteil: einzelne Kommunikationsbeziehungen können individuell gewichtet werden
        - Nachteil: Aufwendige Operationen beeinflussen Leistung
        - WFQ garantiert minimalen Durchsatz und maximale Verzögerung

- Reservierungsprotokolle:
  - Verteilen Dienstanforderungen und Dienstmodifikationen an alle involvierten Systeme (→ **Kontrollprotokolle!**, d.h. KEIN Datentransfer)
  - meist verbindungsorientiert (ausser RSVP!)
  - Verteilen der Dienstspezifikation (FlowSpec): 1. Phase (maximale Reservierungen, Berechnung der Ende-zu-Ende-Parameter), 2. Phase (Adaption der reservierungen)

Dienstgüte:

- wird von mehreren Benutzern bestimmt.
- Aushandlung:
  - potentielle Konflikte bei mehreren Benutzern
  - Konfliktauflösung beim Sender gemäß Gruppensemantik (Abweisen Verbindungsaufbauwunsch; Ablehnen einzelner Empfänger; Aufbau Multicast-Verbindung mit unterschiedlichen Dienstgüten (erfordert entspr. Unterstützung des Kommunikationssystems))
  - unterschiedliche Dienstgüte für Multicast-Verbindungen:
    * Filtern in Zwischensystemen (erfordert hierarchische Kodierung der Benutzerdaten)
    * Filtern ermöglicht effiziente Nutzung der Betriebsmittel und unterstützt Konfliktauflösung

Gemeinsame Nutzung der Ressourcen:

- z.B. Audiokonferenz (mehrere Leute sprechen gleichzeitig)
- Ressourcen einmal reserviert, dann Benutzung durch alle möglich
- Realisiert: "Wildcard"-Filter in RSVP oder Gruppen in ST-II

LRM Local Resource Manager
ST2+:

- verbindungsorientiert

- strikte Trennung: Kontroll-/Datentransfer: SCMP (Stream Controll Message Protocol) für Kontrollnachrichten, ST bietet unzuverlässigen Datentransfer
- Koexistenz von ST2+ und IP in Knoten (Kapselung von ST2+ in IP möglich)
- Stream:
  - ST2+-Instanzen: ST-Agent; Verbindung zwischen ST-Agenten: Hop
  - durch SCMP verwaltet; zuverlässige Übertragung durch Anfrage/Antwort-Modell (lokale Bestätigung + Timer-Überwachung)
  - Datentransfer unidirektional, unzuverl., einfache Mechanismen (KEIN Segmentieren; MTU (maximum transmission unit) an höhere Schichten; sehr wenig Protokollinformation (12 Bytes Paketkopf)
  - ST2+ bietet Zusammenfassung mehrerer Streams in Gruppen an.
- SCMP:
  - Aufgaben: Aufbau, Abbau und Verwaltung der Streams:
  - Aufbau: (2-Wege-Handshake): Verteilung der FlowSpec in CONNECT-Nachricht an Empfänger; Diese: ACCEPT oder REFUSE
  - Modifikation bestehender Streams: Ändern der FlowSpec (CHANGE-Nachricht) durch Sender; Empfängergruppen vergrößern/verkleinern (von beiden initiierbar)
  - Fehlerbehandlung: Überwachung (Anfrage/Antwort), Probleme beim Routig und bei Ressourcenreservierung, Überwachung von Ausfällen (ST-agent, Stream) → sehr komplexe Fehlerbehandlung
- Aufbau eines Streams:
  - involvierte Instanzen: Anwendungsinstanzen als Dienstnutzer; ST-Agenten verteilen FlowSpec an lokale Ressourcenmanager (LRM); RM reservieren maximal auf Hinweg und reduzieren auf Rückweg entsprechend
- FlowSpec von ST2+:
  - Problem: keine Interoperabilität durch Nutzung alternativer FlowSpecs
  - ST2+ erlaubt mehrere FlowSpecs (Versions-Feld!) ("ST2+"-FlowSpec und "Null"-FlowSpec verpflichtend):
    - QoS-Class (QOS-PREDICTIVE (kurze Verletzungen erlaubt, Parameter = mittlere Werte), QOS-GUARANTEED (keine Verletzungen erlaubt, Parameter = Grenzwerte))
    - QoS-Parameter: desired, limit, actual für: maximale Nachrichtengröße, Nachrichtenrate und Ende-zu-Ende-Verzögerung
    - Precedence: Wichtigkeit der Verbindung (wird für Verdrängung gebraucht)
- Ändern der Empfängergruppe:
  - Durch Sender (CONNECT → ACCCEPT); Durch Empfänger (JOIN → (ST-Agent)CONNECT → ACCEPT → (an Sender)NOTIFY)
  - Mehrere Typen von Streams: senderorientiert (Sender bestimmt), empfängerorientiert (Empfänger kann sich anshliessen), hybrid (Sender bestimmt Teilmenge der Empfängergruppe)

  - Strema mit unterschiedlichen Autorisierungslevel:
    * Level 0: JOIN verboten
    * Level 1: JOIN erlaubt, ST-Agent reagiert (CONNECT), an Sender NOTIFY (von ST-Agent)
    * Level 2: JOIN immer erlaubt, ohne NOTIFY

RSVP:

- Kontrollprotokoll
- Datentransfer durch IP (arbeitet verbindungslos! → unterschiedliche Wege)
- RSVP reserviert Ressourcen entlang eines Weges (durch Routing-Protokoll bestimmt) → IP-Datagramme einer verbindung folgen diesem Weg
- RSVP unterstützt unidirektionale Multipeer(m:n)-Verbindungen
- Konzepte:
  - Session: Menge von Datenströmen mit gleichem Ziel (ident. durch IP-Zieladr., Protokoll-Id, Ziel-Port)
  - Dienstbeschreibung durch FlowDescriptor (besteht aus: FlowSpec (Qualitätsparameter und Verkehrscharakteristik) und FilterSpec (wer darf Reservierung nutzen?)) → Trennung von Reservierung und Nutzung der Ressourcen!!
  - Empfängerorientierte Reservierung: Empfänger initiiert.
- RSVP-Architektur:
  - RSVP-Systeme benötigen:
    * Zugangskontrolle (kann Knoten Snorderungen erfülen?)
    * Policy-Kontrolle (darf Anwendung Reservierungen machen?)
    * Classifier: teilt Pakete QoS-Klasse zu (entsprechend FilterSpec)
    * Packet Scheduler: teilt Ressourcen zu (entspr. FlowSpec) (Verwaltung Bandbreite, welches Paket als nächstes?)
- RSVP: Nachrichtenaustausch:
  - Sender periodisch Path-Nachrichten → Baum
  - Empfänger periodisch Resv-Nachrichten retrograd durch Baum
  - → Reservierungen effizienter genutzt (abh. von mehreren Reservierungsmodi)
  - → "soft states" in Knoten: Zustandsinformationen periodisch aktualisiert
- initialer Protokollablauf: Empfänger tritt Gruppe bei (über IGMP); Sender: Path-Nachricht mit (Senderbeschreibung, Verkehrscharakteristik(Sender), optional: Aushandlungsinfo für Empfänger); Empfänger sendet Resv-Nachricht mit seinen Anforderungen (nach Erhalt von Path-N.); Sender beginnt; → erst Daten, wenn Reservierung aufgebaut; → Empfänger wartet auf Path-Nachricht BEVOR er reserviert
- Zusammenfassen von Reservierungen (Merging):

  - Behandlung der resv-Nachrichten in jedem Knoten (Zugangs-/Policy-Kontrolle: wenn negativ. Ablehung, ResvErr an Empf.; Merging: nur Obermenge wird reserviert; Reduktion Kontrollverkehr: Resv-Anfrage nur an Sender, wenn höhere reservierung vorgenommen wurde)
  - Empfänger kann Bestätigung anfordern → !!! Bestätigungen geben keine Garantien!! (Problem: Reservierung wird in späteren Knoten abgelehnt)

- Aushandlung der Dienstqualität. in RSVP nicht möglich. Keine garantierten Dienste (Änderungen in der Route möglich (→ neuer Baum)) in RSVP
- Reservierungsmodi: Sender: speziell für jeden oder gemeinsam für alle; Kontrolle: explizit (Sendermenge) oder alle dürfen senden. Durch Filter im Knoten:
  - Fixed Filter: dedizierte reservierung, ein spezieller Sender
  - Shared-Explicit Filter: gemeinsame Ressourcennutzung durch ausgewählte Sendermenge (z.B. Audioübertragung)
  - Wildcard-Filter: gemeinsame Res.nutzung durch alle Sender
- "Killer Reservation"-Problem:
  - Problem: Reservierung X1 ok; X2 > X1 scheitert in späterem Knoten; oder Y2 scheiter in späterem Knoten; Y1 < Y2 könnte erfolgreich durchgeführt werden.
  - Lösung: bestehende Resevierungen nicht direkt löschen, wenn sie in späteren Knoten scheitern (da Empfänger kontinuierlich pollt (resv in allen Knoten) und gewünschte Resv in möglichst vielen Knoten haben will)
  - Realisierung: gescheiterte Resv → ResvErr an alle Empf.; bestehende Resv nach Empfang von ResvErr "blockaded" markieren; diese nicht mergen und nach bestimmter Zeit löschen
- Fehlerbehandlung durch periodischen Nachrichtenaustausch

TCA Traffic Conditioning Agreement
"Integrated Services":

- Architektur: Unterstützung von Realzeit-Applikationen (→ mehr QoS, Dienstgarantien pro Datenstrom) → Modifikation der Internet-Struktur erforderlich
- Elemente: Integriertes Dienstgütemodell, Implementierungs-Rahmenwerk
- Integriertes Dienstgütemodell:
  - Dienste: best effort (nrt), garantiert (rt), kontrollierte Last (rt)
  - Ressourcen müssen explizit verwaltet werden
  - Schlüssel-Bausteine (Ressourcenreservierung, -zuteilung, Zugangskontrolle); Änderung in Routern (Zustandsdaten PRO Verbindung)
  - Vergabe von Ressourcen unter Beachtung von: Authentifikation von Benutzern, Authentifikation von Dateneinheiten
  - Taxonomie von Applikationen: Applikationen(Realzeit(Tolerant!kontrollierte Last!(adaptiv(Verzögerung, Rate), nicht adaptiv), Intolerant!garantiert!), Elastisch!best effort!(Interactive Bursts!Telnet!, Interactive Bulk!FTP!, Asynchron!E-Mail!))

  - garantierter Dienst (Policing (Konformität), Re-Shaping (verzögert nicht konforme Pakete));

- Implementierungsrahmenwerk:
  - Zugangskontrolle, Classifier (Paket kann auf unterschiedlichen Routern unterschiedichen Klassen zugeteilt werden), Packet Scheduler, RSVP
  - Router-Architektur: Routing-Agent (Routing-Datenbank), RSVP-Agent (Zugangskontrolle + Kontroll-Datenbank), Management-Agent (Kontr.DB); Eingangstreiber, Classifier, Ausgangstreiber (Packet Scheduler)
  - Probleme der IS-Architektur:
    * Skalierbarkeit: Info pro Datenstrom, Weiterleiten durch Klassifizierung → großer Aufwand bei Hochleistungsnetzen
    * Qualitätsparameter frei wählbar → Packet-Scheduling komplex
    * → Differential Services (Prinzip KISS (Keep it Simple and Stupid))

RED Random Early Detection
RIO RED with In and Out
"Differential Services":

- Ziel: bessere Skalierbarkeit (Verkehrsklassen statt pro Datenstrom, Komplexität an Rand der Netze, Definition von DS-Domains)
- verschiedene Dienste: Premium Service, Assured Service. Identifikation durch DS-Byte (IPv4: TOS; IPv6: TCF)
- Architekturmodell:
  - alle Knoten einer DS-Domain: gleiches Per Hop Behavior (bei Weiterleiten)
  - DS-Domain handelt TCA aus (langfristige Lösung!, nicht dynamisch!!)
  - Randknoten: Sicherstellen, dass Verkehr TCA-konform
  - Innere Knoten: nur Forwarden
- Premium Service.
  - Virtual Leased Lines (feste Bitrate). PS im höchsten Knoten belegt nur geringen %-satz der verfügbaren Netzkapazität; nicht genutzte Bandbreite: best-effort
  - Router erbringen den Dienst: First-Hop-Router (Überprüfen Verkehrsvertrag, klassifizieren Verkehr (P-Bit), Traffic-Shaping), Innere-Router (Weiterleiten, Premium-Verkehr höhere Prio.), Border-Router (Policing und Re-Shaping)
- Assured Service:
  - Profiles (Verkehrsvertrag), konform (statistische Garantie), nicht konform (best effort) → keine garantierte Dienstqualitäten!!
  - First-Hop-Router: wie oben (konform: A-Bit), Innere-R: Weiterleiten mit RIO-Queue-Management, Border-R: Policing
- RED und RIO:
  - RED: Best. akt. Warteschlangenlänge → Wahrscheinlichkeit zum Löschen ienes Paketes; → schnelle Reaktion, kürzere Warteschlangen, niedrigere Verlustraten

  - RIO: Konforme Pakete seperat betrachtet, frühes Löschen nicht konformer Pakete

- First-Hop-Router:
  - Vergleich Verkehrsmuster (Datenstrom) mit Profil und klassifiziert Pakete (setzt A/P-Bit), Warteschlange mit Premium-Paketen besitzt höhere Prio.

- Vergleich: IS - DS: DS: keine Garantien pro Datenstrom, einfachere Routerimplementierungen

1.9 Moderne Transportsysteme

Ratenkontrolle mittels Paketmindestabstand oder zeitgesteuertem Byte-Zähler. XTP Xpress Transport Protocol:

- Ziel: effizient. Unterstützung unterschiedlicher Dienste
- Entwurfsprinzipien: Spezifikation (Menge endlicher Automaten), Def. von Paketköpfen fester Länge, Trennung von Daten-/Kontrollfunktionen(Fluss-, Raten-, Fehlerkontrolle) abhängig vom Dienst (Trennung Fluss- und Fehlerkontr., verschiedene Fehlerkontrollmodi, Entlastung des Empfängers), Unterstützung Multicast-Kommunikation, Aushandlung der Dienstgüte
- Dienste:
  - zuverlässiger, verbindungsorientierter Unicast Dienst (in zwei Richtungen, Fluss- Ratenkontr., gesicherter Abbau) (entspr. TCP)
  - Datagramm-Dienst (unbestätigt (entspr. UDP), bestätigt (Empfangsbest.,Fehlerkontr.))
  - Transaction-Dienst: 2 Phasen Datenaustausch nur für Unicast (Anfrage/Antwort/Bestätigung Antwort)
  - verbindungsorientierte Multicast-Dienste: unbestätigt (unzuverl., in eine Richtung, unbek. Gruppe, keine Rückmeldung), bestätigt (zuverlässig, Kontrollverkehr in andere Richtung, Fluss- Ratenkontr., bek. Gruppe)
- Verbindungsaufbau:
  - Input-Primitiv, wartet auf eingehende Verbindung, Kontext wird erzeugt → Listen
  - Output-Primitiv initiiert Verbindungsaufbau, Kontext wird erzeugt, FIRST-Paket
  - Empf. von FIRST überprüft, ob Adr.info einem Kontext im Listen-Zustand zugeordnet werden kann, wenn ja: Assoziation (impliziter Verbindungsaufbau)
- Fluss-/Ratenkontrolle
  - fensterbasierte Flusskontrolle und Ratenkontrolle bei XTP
  - Flusskontrolle: durch NOFLOW-Bit (im Paketkopf) vom Sender aktiviert/deaktiviert, Empf. sig. Sender alloc(höchste erlaubte Sequenznummer) und rseq(höchste empfangene Seq.nr +1)
  - Ratenkontrolle: reguliert traffic spec durch: bursts (Anzahl der Bytes, die in einem burst gesendet werden dürfen), rate (mittlere Datenrate Byte/Sec) → credit wird nach Ablauf von RTIMER auf burst gesetzt und beim Senden um Paketgröße dekrementiert (solange > 0) RTIMER: Intervall=burst/rate

TCP in Hochleistungsnetzen:

- nur bedingt geeignet (3-Wege-HS, Fenstergröße zu klein (bei hoher Pfadkapazität), go-Back-N, Berechnung der Umlaufzeit zu ungenau)
- derzeit nur problemspezifische Lösungen: Fensterskalierung, Zeitstempel, selektive Quittierung, T/TCP (Umgehung 3-Wege-HS) (T=Transaction)
- Fensterskalierung:
  - Problem: limitierte Fenstergröße (16-bit) nicht mehr ausreichend
  - Lösung: Skalierung in SYN-Dateneinheiten, Basiseinheiten in 2^x. Skalierungsfaktor auf 2^14 beschränkt (sonst Sequenznummerüberlauf)
- Zeitstempel.
  - Problem: Paketumlaufzeit (Berechnung zu ungenau)
  - Lösung: Sender setzt akt. Timerwert in Optionsfeld TSval eine Dateneinheit; Empfänger kopiert diesen in TSecr der entspr. Quittung; Sender bildet Differenz mit neuem aktuellen Timerwert
- Weiterentwicklungen: Unterstützung von Gruppenkommunikation, Straffung etc.

RTP	Real Time Transport Protocol
RTCP	Real Time Control Protocol
ALF	Application Level Framing
SSRC	Synchronization Source Identifier (RTP Datenformat)
CSRC	Contributing Source Identifier (1. - letzter)

- Unterstützt Ende-zu-Ende Realzeitdaten (durch Zeitstempel)
- Keine Unterstützung von Fehlerkontr. und Ressourcenreserv. → Dienstgüteunterstützung ggf. von darunterliegender Protokollschicht (z.B. RSVP)
- Komponenten: RTP und RTCP
- Daten- und Kontrollfluss:
  - getrennt
  - per Multicast
  - Empf. werten Kontrolldaten anderer Grupenmitgleider aus (Anpassung Statusrate an Gruppengröße, Monitoren der Ü-qualität, Kenntnis über weitere Gruppenmitglieder (ID, Empfangsqualität))
- Kontrolle einer RTP-Übertragung:
  - unterstützt lediglich Transfer → separates RTCP
  - periodischer Austausch zwischen allen Teilnehmern
  - UDP-Port für RTCP um 1 höher als entspr. RTP-Port
  - 5 verschiedene nachrichtentypen (Sender Report, Receiver Report, Source Description, Mye, Application Specific)
  - Funktionalität: (Rückmelden der akt. Dienstqualität, Kontrolle der durch RTP belegten Bandbreite (wegen Skalierbarkeit, 5% für RTCP), Sitzungskontrolle)

- Protokollarchitektur:
  - ALF
  - kein vollständiges Protokol, nur Rahmenwerk
  - Bereitstellung Basisfunktionalität (IP; UPD; RTP; (MPEG, H.261, JPEG,...))
  - Anwendungen ergänzen spezifischen Details
  - Implementierungstechnik: Programmbibliothek anstelle von Protokoll-Serverprozess (effizientere Schnittstellenkommunikation)
- Unterstützung heterogener Gruppen:
  - Problem: Empf. unterschiedlich leistungsfähig
  - Lösung: Einsetzen intelligenter Zwischensysteme (Mixer oder Translator)
  - Mixer:
    - Kombinieren mehrere Datenströme in einen einzelnen Datenstrom (z.B. Audiokonferenz mit mehreren Sendern)
    - Datenformat wird beibehalten
    - NICHT geeignet fÜr Video-Übertragung
  - Translator:
    - Wandeln Daten zwischen zwei unterschiedlichen Formaten
    - $\rightarrow$ können unterschiedliche Dienstqualitäten unterstützen
    - Erlaubt Anschluss von Endgeräten mit unterschiedlicher Leistungsfähigkeit
- Datenformat:
  - Version; Padding; Extension; Contributer Count; Marker; Payload Type; SN; SSRC(32bit), 1. CSRC (32), ..., letzter CSRC, Anwendungsdaten
  - Extension: (normaler Paketkopf folgt, erweiterter Paketkopf folgt)
  - Marker: Dient Trennung von Nachrichten
  - SSRC bzw. CSRC: Dienen Identifikation des Synchronisationspunktes bzw. des ursprünglichen Absenders der Dateneinheit
- Source Identifier:
  - SSRC identifiziert Sender der Daten (legt SN und Zeitstempel fest)
  - Translator verändern SSRC NICHT
  - Mixer verändern SSRC!!! und tragen die Sender der ursprünglichen Nachricht als CSRC ein.