

Next Generation Internet

Kapitel 7: Multicast-Transport

INSTITUT FÜR TELEMATIK



KIT – Universität des Landes Baden-Württemberg und
nationales Forschungszentrum in der Helmholtz-Gemeinschaft

www.kit.edu

Überblick Kapitel 7

I. Einführung

1. Einführung

II. Internet-Architektur

2. Internet-Architektur
3. NAT & IPv6
4. Dienstgüte

III. Multicast

5. Grundlagen
6. Multicast Routing
7. Multicast Transport

- 7.1 Multicast-Transport: Eigenschaften
- 7.2 Transportprotokolle für Multicast:
Anforderungen und Mechanismen
- 7.3 Multicast-Transportprotokolle:
Ausgewählte Beispiele

IV. Flexible Dienste und Services

8. Neuere Transportprotokolle
9. Aktive Netze
10. Peer-to-Peer

2

Next Generation Internet SS2010 – 7. Multicast Transport (R0)

Institut für Telematik, Fakultät für Informatik
<http://tm.kit.edu/>

7.1 Gruppen-Transportdienste



- Traditionelle Transportdienste für Punkt-zu-Punkt-Kommunikation: **unzuverlässig** (UDP), **zuverlässig** (TCP)
- Erweiterung auf Gruppenkommunikation
 - Bei unzuverlässigen Diensten
 - Nutzung von UDP möglich
 - Verwendung von Multicast-Adressen als Zieladressen
 - Beitritt zu Gruppen und Verlassen von Gruppen
 - Bei zuverlässigen Diensten zusätzlich erforderlich:
 - Verbindungsauf- und -abbau
 - Fehlerkontrolle, Quittungsbehandlung, Sendewiederholungen, Staukontrolle
 - Sicherheit (z.B. Denial-of-Service-Schutz)
 - Ein oder mehrere Sender:
 - **1:n** → meist ohne menschliche Interaktion, z.B. SW-/Datenverteilung,
 - **m:n** (Multipeer dann meist m:m) → oft interaktiv, z.B. verteilte Spiele/Simulationen, Tele-/Videokonferenzen

3

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
<http://tm.kit.edu/>

• TCP kann nicht einfach mit Multicast-Zieladressen betrieben werden, da es nur für 1:1-Kommunikation entwickelt wurde.

Problem



- Traditionelle Mechanismen nicht ohne Weiteres anwendbar/übertragbar
 - Nur für Punkt-zu-Punkt-Kommunikation definiert
 - Skalierbarkeitsprobleme bzgl. Anzahl der Empfänger
 - Kontrolldaten
 - Menge der übertragenen Nachrichten (z.B. Anzahl Quittungen)
 - Anzahl der zu verarbeitenden Nachrichten
 - Zustandshaltung
- Es zeichnet sich keine einheitliche Lösung ab
 - Gründe hierfür sind u.a.:
 - Aufgabenbereiche für Transport variieren stark
 - Erfahrungen mit Gruppenkommunikation sind noch gering
 - Daher werden **modulare Lösungen** entwickelt

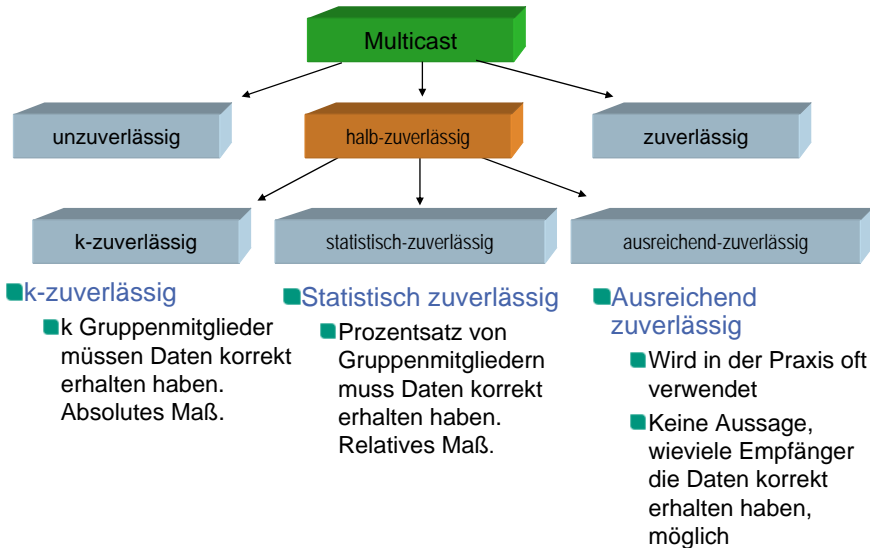
4

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
<http://tm.kit.edu/>

Zuverlässigkeitsklassen



5

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
<http://tm.kit.edu/>

- Unzuverlässiger Gruppendienst
 - Es werden keine Aussagen über die Auslieferung der Daten getroffen
- Halbzuverlässiger Gruppendienst
 - Nicht alle Teilnehmer müssen alle Daten korrekt erhalten. Weitere Unterscheidung auf der Basis einzelner Dateneinheiten
 - Statistisch zuverlässig
 - Prozentsatz von Gruppenmitgliedern muss Daten korrekt erhalten haben. Relatives Maß.
 - k-zuverlässig
 - k Gruppenmitglieder müssen Daten korrekt erhalten haben. Absolutes Maß.
 - Ausreichend zuverlässig
 - Keine Aussage, wieviele Empfänger die Daten korrekt erhalten haben, möglich - weil häufig die Gruppe anonym ist - daher ähnlich dem unzuverlässigen Gruppendienst. Wird in der Praxis oft verwendet.

Zuverlässiger Gruppendienst



- **Zuverlässigkeit:** Alle Empfänger erhalten Daten
 - fehlerfrei
 - ohne Duplikate
 - in der korrekten Reihenfolge (Sendereihenfolge)
- **Atomar zuverlässig**
 - Korrektheit, keine Duplikate, Auslieferung an alle oder keinen
- **Globale Ordnungserhaltung**
 - **Ordnungserhaltend**
 - atomar zuverlässig
 - Daten eines Senders kommen reihenfolgetreu an
 - **Total Ordnungserhaltend** (nur Multipeer)
 - ordnungserhaltend
 - alle Daten aller Sender erreichen alle Empfänger in der gleichen Reihenfolge

6

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



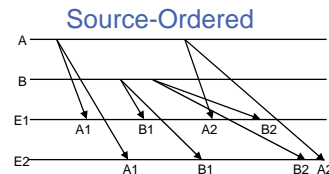
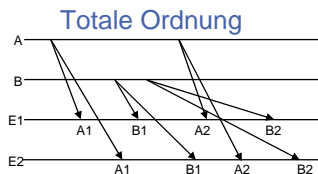
Institut für Telematik, Fakultät für Informatik
<http://tm.kit.edu/>

- Alle Empfänger erhalten Daten fehlerfrei, in der korrekten Reihenfolge und ohne Duplikate. Korrekte Reihenfolge bezieht sich auf die Sendereihenfolge der Daten. Hinsichtlich der **globalen Ordnungserhaltung** sind die folgenden Begriffe wichtig.
 - Atomar zuverlässig
 - Duplikate von Daten werden abgefangen. Korrekte Übertragung der Daten wird gewährleistet, ansonsten Abbruch der Kommunikationsbeziehung.
 - Darüber hinaus wird gewährleistet, dass entweder alle Teilnehmer die Daten erhalten oder sie an keinen der Teilnehmer ausgeliefert werden.
 - Ordnungserhaltend
 - Die Kommunikation ist atomar zuverlässig
 - Darüber hinaus können sich die Daten eines Senders nicht überholen und kommen bei allen Empfängern in der gleichen Reihenfolge an.
 - Total ordnungserhaltend (Multipeer)
 - Atomar zuverlässig und ordnungserhaltend
 - Alle Daten aller Sender erreichen alle Empfänger in der gleichen Reihenfolge

Ordnungserhaltung



- **Multicast:** Entspricht Unicast-Kommunikation
- **Multipeer:** Berücksichtigung unterschiedlicher Sender
- **Ordnungskriterien**
 - **Globale Ordnung:**
 - Daten werden in genau der Reihenfolge ausgeliefert, in der sie gesendet wurden
 - **Totale Ordnung:**
 - Daten werden bei allen Empfängern in der gleichen Reihenfolge ausgeliefert
 - **Source-Ordered:**
 - Daten werden nur bezogen auf eine Quelle in der richtigen Reihenfolge ausgeliefert



7

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
http://tm.kit.edu/

- Multicast: Entspricht Unicast-Kommunikation
- Multipeer: Berücksichtigung unterschiedlicher Sender
- **Ordnungskriterien**
 - Globale Ordnung: Daten werden in genau der Reihenfolge ausgeliefert, in der sie gesendet wurden
 - Setzt global synchronisierte Uhren voraus
 - Ist in der Regel nicht erforderlich
 - Totale Ordnung: Daten werden bei allen Empfängern in der gleichen Reihenfolge ausgeliefert
 - Schwächere Anforderung, reicht in der Regel aber aus
 - Source-Ordered: Daten werden nur bezogen auf eine Quelle in der richtigen Reihenfolge ausgeliefert
- Globale Ordnung im Beispiel wäre: A1,B1,B2,A2 (B2 wird vor A2 gesendet und muss auch so bei den Empfängern ankommen)

7.2 Transportprotokolle für Multicast: Allgemeines



- Anforderungen aus Sicht der IETF [RFC3048]
 - **Staukontrolle**
 - Keine Überlastung durch **Zusatzverkehr (Redundanz)** oder Verkehr zur Behebung von Fehlersituationen (sog. **Reparatur-Verkehr** – „Repair Packets“)
 - Erreichen einer guten Link-Auslastung
 - Keine Ausgrenzung konkurrierender Datenströme („Verhungern“)
 - **Skalierbarkeit**
 - Einsetzbar unter verschiedenen Bedingungen
 - Verschiedene Netztopologien
 - Verschiedene Link-Geschwindigkeiten
 - Verschiedene Empfängergruppen
 - Gutes Verständnis erforderlich, wann Protokoll einsetzbar bzw. nicht geeignet
 - **Sicherheit**
 - Authentisierung des Senders
 - Bekämpfung von Denial-of-Service-Attacken

8

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
http://tm.kit.edu/

- Die folgenden Fragestellungen leiten die Auswahl eines geeigneten Gruppendienstes und damit die Wahl eines geeigneten Transportprotokolls
 - Muss jeder Empfänger alle Daten erhalten?
 - Müssen Unterschiede zwischen Empfängern berücksichtigt werden?
 - Ist Skalierbarkeit für große Gruppen erforderlich?
 - Ist „komplette“ Zuverlässigkeit erforderlich?
 - Ist Ordnungserhaltung (globale Ordnung oder andere) erforderlich?
 - Müssen Daten mit kurzer Verzögerung ausgeliefert werden?
 - Müssen Daten in einer vorgegebenen Zeitschranke ausgeliefert werden?
 - Sind viele interagierende Sender vorhanden?
 - Muss die Anwendung ohne einen Rückpfad funktionieren?
 - Welche Sicherheitsaspekte sind relevant?

Multicast-Transport-Unterstützung innerhalb des Netzes



■ Keine Unterstützung im Netz

- Router leiten Daten nur weiter
- Mechanismen zur Multicast-Transport-Unterstützung komplett in Sender und Empfängern (Ende-zu-Ende-Prinzip)
- Geschichtete Datenströme (Layered Streams)
 - Daten werden in unterschiedliche Datenströme aufgeteilt, die jeweils einer Gruppe zugeordnet werden
 - Empfänger treten den Gruppen einzeln bei, beispielsweise je nach individuell gewünschter Dienstqualität

■ Server-basiert

- Dedizierte Systeme im Netz unterstützen Sender und Empfänger
- Systeme sind nicht notwendigerweise im Datenpfad

■ Router-basiert

- Router unterstützen Sender und Empfänger
- Router müssen Ressourcen hierfür bereitstellen

9

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
<http://tm.kit.edu/>

- Zuverlässige bzw. halbzuverlässige Gruppendienste erfordern Protokollmechanismen zur Fehlerkontrolle und -behebung etc. Folgende Varianten zur Multicast-Unterstützung sind möglich

- Keine Unterstützung im Netz

- Router leiten Daten nur weiter
- Mechanismen zur Multicast-Transport-Unterstützung komplett in Sender und Empfängern (Ende-zu-Ende-Prinzip)

- Geschichtete Datenströme (Layered Streams)

- Daten werden in unterschiedliche Datenströme aufgeteilt, die jeweils einer Gruppe zugeordnet werden
 - Typisches Beispiel: Videoströme
 - Abhängig von zu übertragenden Daten, ob hierarchische Codierung möglich
 - Möglich aber auch: eine Gruppe für Grunddaten, eine Gruppe für Redundanzdaten (FEC)
- Empfänger treten den Gruppen einzeln bei, beispielsweise je nach individuell gewünschter Dienstqualität oder Stausituation

- Server-basiert

- Dedizierte Systeme im Netz unterstützen Sender und Empfänger
- Systeme sind nicht notwendigerweise im Datenpfad

- Router-basiert

- Router unterstützen Sender und Empfänger, z.B. durch Filtern von Quittungen etc.
- Router müssen Ressourcen hierfür bereitstellen
 - Gutes Beispiel für den Einsatz aktiver und programmierbarer Netze (s. NGI-Kapitel: Aktive Netze)

Fehlerbehandlung



■ Zuverlässigkeit erfordert Fehlerbehandlung

■ Multicast-Fehlerkontrolle ist grundsätzlich aufwändiger als bei Punkt-zu-Punkt-Kommunikation

- Fehlererkennung wird in der Regel vom Empfänger durchgeführt
- Fehlerbehebung erfolgt oftmals durch Übertragungswiederholung
- Quittungen werden zum Sender geschickt

■ Skalierbarkeit

- Mechanismen müssen so ausgelegt werden, dass sie bei großen und bei geographisch weit verstreuten Gruppen einsetzbar sind

10

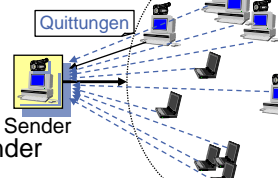
Next Generation Internet SS2010 – 7. Multicast Transport (R0)

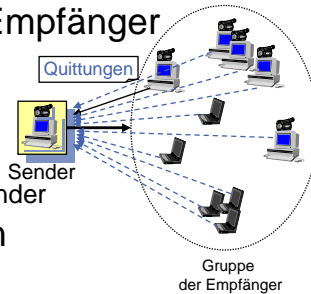


Institut für Telematik, Fakultät für Informatik
<http://tm.kit.edu/>

- Zuverlässigkeit erfordert die korrekte Auslieferung aller Dateneinheiten in der richtigen Reihenfolge → Fehlerbehandlung erforderlich

Quittungsverarbeitung (1)

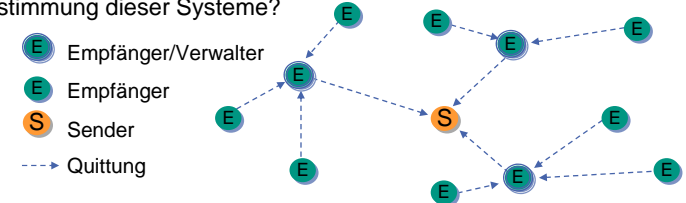
- Sender muss Quittungen vieler Empfänger behandeln
 - Problem der **Quittungs-Implosion**
 - Hohe Netzbelastung
 - Hoher Pufferbedarf beim Sender
 - Hoher Bearbeitungsaufwand beim Sender
 - Es ist zu unterscheiden zwischen
 - Multicast **ohne Empfängerliste**
 - Quittungen werden der gesamten Gruppe zugeordnet
 - Reaktion des Senders auf Quittung kann sofort oder verzögert erfolgen
 - Multicast **mit Empfängerliste**
 - Quittungen können einzelnen Empfängern zugeordnet werden
 - Sender besitzt detaillierte Information über jedes einzelne Gruppenmitglied
- 



- Verarbeitung von Quittungen im Multicast-Fall weicht erheblich von der Bearbeitung im Unicast-Fall ab
- Sender muss Quittungen von einer Vielzahl von Empfängern behandeln
 - Führt zu dem Problem der Quittungs-Implosion
 - Hohe Netzbelastung
 - Hoher Pufferbedarf beim Sender
 - Hoher Bearbeitungsaufwand beim Sender
- Es ist zu unterscheiden zwischen
 - Multicast ohne Empfängerliste
 - Quittungen werden der gesamten Gruppe zugeordnet.
Es ist lediglich ein halbzuverlässiger Dienst möglich. Weshalb?
 - Reaktion des Senders auf Quittung kann sofort oder verzögert erfolgen
 - Welchen Vorteil hat die verzögerte Reaktion?
 - Multicast mit Empfängerliste
 - Quittungen können einzelnen Empfängern zugeordnet werden
 - Sender besitzt detaillierte Information über jedes einzelne Gruppenmitglied. Es kann ein k-zuverlässiger und ein zuverlässiger Dienst realisiert werden.

Quittungsverarbeitung (2)

- Ziel: Vermeiden der Quittungs-Implosion
- Verfahren
 - Lokale Bearbeitung von Quittungen für Teilgruppen in Servern bzw. Routern
- Vorteil
 - Verteilung der Bearbeitungslast auf mehrere Kommunikationssysteme, parallele Abarbeitung möglich.
 - Ausgezeichnete Empfänger bzw. spezielle Zwischensysteme bearbeiten Quittungen bzw. Kontrolldateneinheiten
 - Bestimmung dieser Systeme?



- Ziel
 - Vermeiden von Quittungs-Implosion
- Verfahren
 - Lokale Bearbeitung von Quittungen für Teilgruppen in Servern bzw. Routern
- Vorteil
 - Verteilung der Bearbeitungslast auf mehrere Kommunikationssysteme, parallele Abarbeitung möglich.
 - Ausgezeichnete Empfänger bzw. spezielle Zwischensysteme bearbeiten Quittungen bzw. Kontrolldateneinheiten
 - Bestimmung dieser Systeme?
 - Hier muss eine automatische Auswahl erfolgen und bei Ausfall für Ersatz gesorgt werden
 - Die Auswahl dieser Systeme hat Einfluss auf die Leistung

Übertragungswiederholung



- Grundsätzlich dieselben Mechanismen einsetzbar wie für Unicast

- Go-Back-N
- Selektive Übertragungswiederholung

- Möglichkeiten zur Wiederholung der Daten

■ Per Multicast

- Erfordert nur eine Wiederholung, auch für mehrere Empfänger
- Belastet alle Empfänger und involvierte netzinterne Zwischensysteme
- Verbesserung durch **Subcasting** (s. NGI-Kapitel: Reliable Multicast Transport Protocol)

■ Per Unicast

- Erfordert eine dedizierte Wiederholung pro Empfänger, d.h., die Empfänger müssen bekannt sein
- Ineffizient, wenn Daten von mehreren Empfängern angefordert werden

13

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



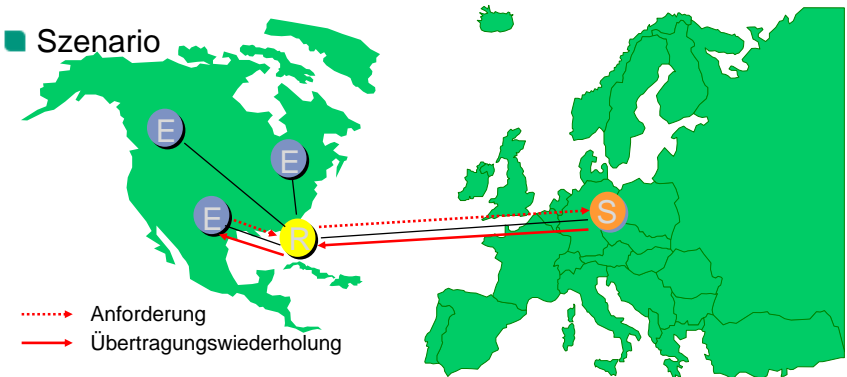
Institut für Telematik, Fakultät für Informatik
http://tm.kit.edu/

- Grundsätzlich dieselben Mechanismen einsetzbar wie für Unicast
 - Go-Back-N
 - Führt jedoch zu erheblicher Netzbelastung
 - Leichte Realisierbarkeit
 - Geringe Pufferkapazitäten bei den Empfängern erforderlich
 - Selektive Übertragungswiederholung
 - Geringere Netzbelastung als bei Go-Back-N
 - Hohe Pufferkapazitäten bei Empfängern notwendig
- Unterschiedliche Mechanismen zur Wiederholung der Daten
 - Per Multicast
 - Erfordert nur eine Wiederholung, auch für mehrere Empfänger
 - Belastet alle Empfänger und involvierte netzinterne Zwischensysteme
 - Verbesserung durch Subcasting (s. NGI-Kapitel: Reliable Multicast Transport Protocol)
 - Per Unicast
 - Erfordert eine dedizierte Wiederholung pro Empfänger, d.h., die Empfänger müssen bekannt sein
 - Ineffizient, wenn Daten von mehreren Empfängern angefordert werden

Beispiel: Geographisch verteilte Gruppen



■ Szenario



■ Basisverfahren

- Empfänger fordern fehlende Daten stets direkt beim Sender an

■ Problem

- Eventuell hohe Belastung beim Sender und lange Laufzeiten, da keine Berücksichtigung der Gruppenstruktur und der Netztopologie

14

Next Generation Internet SS2010 – 7. Multicast Transport (R0)

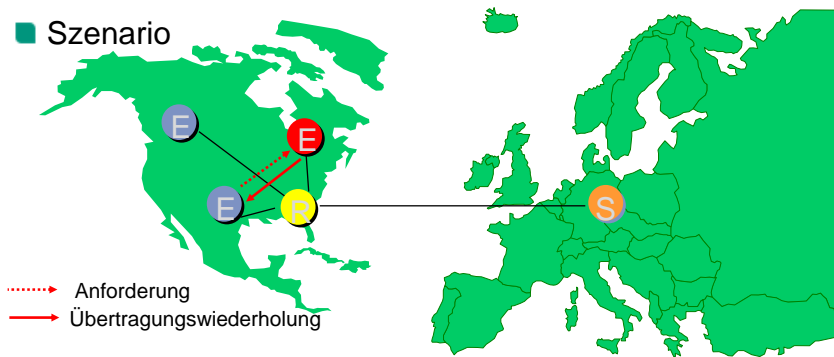


Institut für Telematik, Fakultät für Informatik
http://tm.kit.edu/

Optimierte Fehlerbehebung durch dedizierte Systeme



Szenario



Optimierung

- Übertragungswiederholungen durch dedizierte Systeme, hier „lokale“ Empfänger

Vorteil

- Minimierung der (globalen) Netzbelastung und der durchschnittlichen Übertragungsverzögerung durch lokale Übertragungswiederholungen

15

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
<http://tm.kit.edu/>

7.2.1 Verfahren zur Multicast-Fehlerbehebung



- Die folgenden Varianten der Fehlerbehebung können im Kontext von Multicast-Transportprotokollen unterschieden werden

- Sender-gesteuert
- Empfänger-gesteuert
- Ring-basiert
- Baum-basiert
- Zeitgeber-gesteuert
- Vorwärtsfehlerkorrektur

16

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



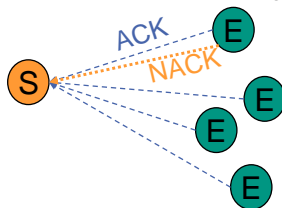
Institut für Telematik, Fakultät für Informatik
<http://tm.kit.edu/>

Sender- vs. Empfänger-gesteuert



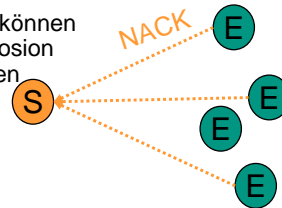
Sender-gesteuert

- Empfänger bestätigen korrekten Empfang mit **ACKs**
- Sender stößt Übertragungswiederholung an
- Bekannte Gruppe erforderlich
- Problem der Quittungs-Impllosion inhärent
- Geringer Durchsatz
- Vollzuverlässiger Dienst realisierbar mit endlichem Speicher
- Mischform mit NACKs möglich



Empfänger-gesteuert

- Empfängerseitige Fehlerkontrolle
- Empfänger fordern im Fehlerfall mit **NACKs** Übertragungswiederholungen an
- Gruppenmitglieder müssen nicht bekannt sein
- Sender muss keine Information über Zustand der Empfänger verwalten
- Höherer Durchsatz als bei Sender-gesteuerten Verfahren erreichbar
- Freigabe der Sende-Puffers aus NACKs nicht ableitbar
- NACKs allein können Quittungsimpllosion nicht vermeiden



17

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
http://tm.kit.edu/

Sender-gesteuert:

- Ablauf und Eigenschaften
 - Empfänger bestätigen korrekten Empfang mit ACKs
 - Sender stößt Übertragungswiederholung nach Ausbleiben der Quittung an
 - Mischform mit NACKs möglich
 - Reduktion der Übertragungsverzögerung
 - Bekannte Gruppe erforderlich
 - Problem der Quittungs-Impllosion inhärent
 - Aggregation mehrerer Quittungen in einer Dateneinheit hilft nur eingeschränkt
 - Geringer Durchsatz
 - Vollzuverlässiger Dienst realisierbar mit endlichem Speicher
 - Sender kann Puffer freigeben, wenn alle Quittungen empfangen wurden

EMPFÄNGER-GESTEUERT:

- Ablauf und Eigenschaften
 - Fehlerkontrolle wird von den Empfängern durchgeführt
 - Empfänger fordern im Fehlerfall mit NACKs Übertragungswiederholungen an
 - Gruppenmitglieder müssen nicht bekannt sein
 - Sender muss keine Information über Zustand der Empfänger verwalten
 - Höherer Durchsatz als bei Sender-gesteuerten Verfahren erreichbar
 - Freigabe der Sende-Puffers aus NACKs nicht ableitbar
 - Praktisch nur halb-zuverlässiger Dienst realisierbar
 - NACKs allein können Quittungsimpllosion nicht vermeiden
 - Beispiel: MTP

Empfänger-gesteuert mit NACK-Vermeidung



■ Ziel

- Der Sender (bzw. ein entsprechender anderer Knoten) soll **nur ein NACK** für eine fehlende Dateneinheit erhalten
- Die Empfänger, die eine Sendewiederholung veranlasst haben, sollen nur eine Kopie der Daten erhalten

■ Beispiele:

■ Zeitgeber

- Empfänger senden NACK im Fehlerfall erst nach Ablauf eines Zeitgebers
 - Zeitgeber-Dimensionierung schwierig
- Trifft NACK für die gleiche Dateneinheit ein, wird der Zeitgeber erneut gestartet
- NACKs und Übertragungswiederholungen belasten die gesamte Gruppe

■ Router-Unterstützung

- Router unterdrücken doppelte NACKs

18

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
http://tm.kit.edu/

• Ziel

- Der Sender (bzw. ein entsprechender anderer Knoten) soll nur ein NACK für eine fehlende Dateneinheit erhalten
- Die Empfänger, die eine Sendewiederholung veranlasst haben, sollen nur eine Kopie der Daten erhalten

• Beispiel: Zeitgeber

- Empfänger senden NACK im Fehlerfall erst nach Ablauf eines Zeitgebers (Abwarten, ob anderes System nicht erst ein NACK schickt)
 - Zeitgeber-Dimensionierung schwierig
- Trifft ein NACK für die gleiche Dateneinheit ein, wird der Zeitgeber erneut gestartet
- NACKs und Übertragungswiederholungen belasten die gesamte Gruppe
 - „Crying baby problem“

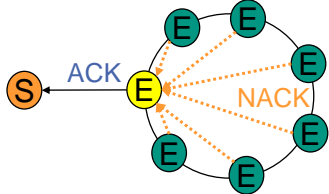
• Beispiel: Router-Unterstützung

- Router unterdrücken doppelte NACKs

Ring- vs. Baum-basiert

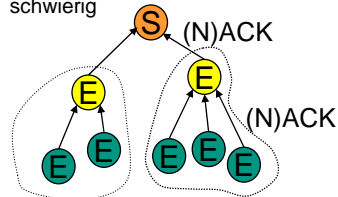
Ring-basiert

- Gruppe bildet einen virtuellen Ring, in dem ein **Token** zirkuliert
- Empfänger senden NACKs an den aktuellen **Token-Halter**
- Token-Halter sendet ACK an Sender
- Vermeidung der Quittungsimplosion
- Totale Ordnungserhaltung realisierbar
- Sender kann Speicher nach endlicher Zeit freigeben
- Durchsatz hängt von Gruppengröße ab, skaliert daher nicht gut
- Quittungs-Implosion beim Token-Halter möglich (→ Hierarchische Ringe)



Baum-basiert

- Hierarchie von lokalen Teilgruppen mit ausgezeichnetem **Group Leader**
- Quittungen werden an jeweils zuständigen Group Leader gesendet
- Aggregation von Quittungen** in den Group Leaders
- Quittungs-Implosion wird vermieden
- Mit aggregierten ACKs ist „komplett“ zuverlässiger Dienst möglich (mit endlichem Speicher)
- Erhöhte Ende-zu-Ende-Verzögerung
- Aufbau und Verwaltung des Baums schwierig



19

Next Generation Internet SS2010 – 7. Multicast Transport (R0)

• Ring-basiert: Ablauf und Eigenschaften

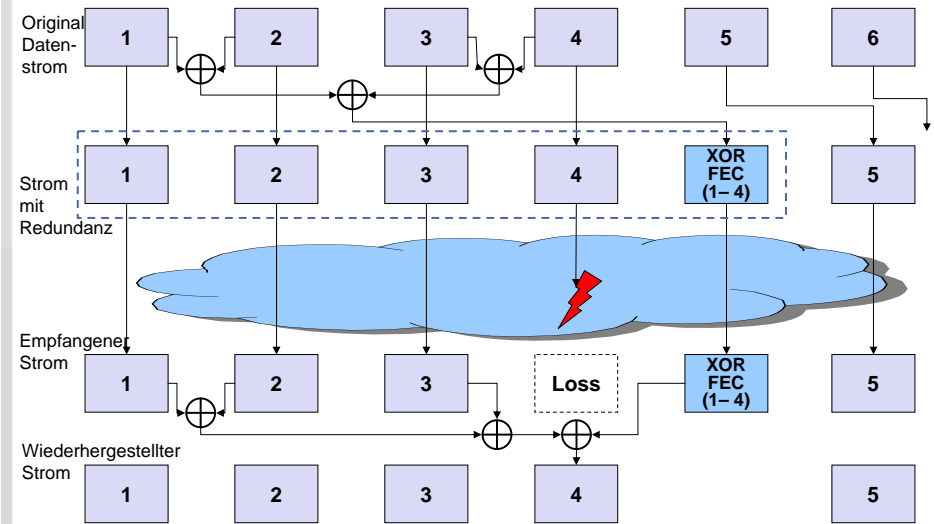
- Gruppe bildet einen virtuellen Ring, in dem ein Token zirkuliert
- Empfänger senden NACKs an den aktuellen Token-Halter, falls Daten fehlen
- Token-Halter sendet ACK an Sender
- Vermeidung der Quittungsimplosion
- Totale Ordnungserhaltung realisierbar
- Sender kann Speicher nach endlicher Zeit freigeben
- Durchsatz hängt von Gruppengröße ab, skaliert daher nicht gut
- Quittungs-Implosion beim Token-Halter möglich (→ Hierarchische Ringe)
- Beispiel: RMP (Reliable Multicast Protocol)

• Baum-basiert: Ablauf und Eigenschaften

- Hierarchie von lokalen Teilgruppen mit ausgezeichnetem Group Leader
- Quittungen werden an jeweils zuständigen Group Leader gesendet
- Aggregation von Quittungen in den Group Leaders
- Quittungs-Implosion wird vermieden
- Mit aggregierten ACKs ist „komplett“ zuverlässiger Dienst möglich (mit endlichem Speicher)
- Erhöhte Ende-zu-Ende-Verzögerung
- Aufbau und Verwaltung des Baums schwierig
- Beispiel: RMTP (s. NGI-Kapitel: Reliable Multicast Transport Protocol)

• Szenario

Vorwärtsfehlerkorrektur



20

Next Generation Internet SS2010 – 7. Multicast Transport (R0)

• Beispiel

- Zusammenfügen einiger Dateneinheiten zu einer Gruppe
- XOR-Operation über diese Gruppe von Dateneinheiten zur Erstellung einer neuen Dateneinheit, die ebenfalls gesendet wird
- Fehlerhafte bzw. verlorengegangene Dateneinheit kann rekonstruiert werden, falls die durch XOR-Operation erzeugte Dateneinheit korrekt empfangen wurde

• Gliederung in

- Proaktive Vorwärtsfehlerkorrektur (Redundanz von vorneherein)
- Reaktive Vorwärtsfehlerkorrektur (Redundanz nach Bedarf hinzufügen)

• Eigenschaften

- „Reparatur“-Verkehr
 - kann je nach Bedarf reduziert werden
 - kann als separater Strom übertragen werden
- Ende-zu-Ende-Verzögerung erhöht sich
 - Bedingt durch Gruppierung der Dateneinheiten
 - Beispielsweise bei interaktiven Anwendungen problematisch

7.2.2 Verfahren zur Multicast-Staukontrolle



■ Sender-kontrolliert

- mit einer Gruppe
 - Empfänger liefern Feedback-Information an Sender
 - Sender stellt Senderate auf Empfänger mit langsamsten Pfad ein
- mit mehreren Gruppen
 - initiale Gruppe wird dynamisch in Teilgruppen gegliedert
 - angepasste Senderate für verschiedene Teilgruppen

■ Empfänger-kontrolliert

- mit einer Gruppe
 - Empfänger verlassen die Gruppe, falls Verlustrate zu hoch
- mit geschichteter Organisation (Layered Multicast)
 - Sender verteilt Daten auf mehrere Multicastgruppen
 - Empfänger treten den jeweiligen Multicastgruppen bei bzw. verlassen diese in Abhängigkeit der aktuellen Netzlast
 - Empfänger „hinter“ einem Engpass im Netz müssen koordiniert den Multicastgruppen beitreten bzw. diese verlassen

21

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
<http://tm.kit.edu/>

- Ausgangssituation
 - Internet basiert auf Best-Effort-Dienst zur Auslieferung der Daten
 - Anwendungen sollen sich adaptiv verhalten und an die Netzlast anpassen
- Verfahren zur Multicast-Staukontrolle
 - Sender-kontrolliert mit einer Gruppe
 - Sender-kontrolliert mit mehreren Gruppen
 - Empfänger-kontrolliert mit einer Gruppe
 - Empfänger-kontrolliert mit geschichteter Organisation
 - Router-basiert
- Sender-kontrolliert:
 - Eine Gruppe
 - Empfänger liefern Feedback-Information an Sender
 - Sender stellt Senderate auf den Empfänger mit langsamsten Pfad ein
 - Mehrere Gruppen
 - Die initiale Gruppe wird dynamisch in Teilgruppen gegliedert
 - Ausschlaggebend sind Stauunkte im Netz
 - Den verschiedenen Teilgruppen werden die Daten mit jeweils angepasster Rate gesendet
 - Relay auf Anwendungsebene erforderlich zur Pufferung
- Empfänger-kontrolliert:
 - Eine Gruppe
 - Empfänger verlassen die Gruppe, falls die Verlustrate für sie zu hoch ist
 - Geschichtete Organisation
 - Sender verteilt Daten auf mehrere Multicastgruppen („Layered Multicast“, Layered Coding)
 - Beispiel: MPEG I-, P-, B-Frames jeweils in eine eigene Gruppe senden
 - Empfänger treten den jeweiligen Multicastgruppen bei bzw. verlassen diese in Abhängigkeit der aktuellen Netzlast
 - Empfänger „hinter“ einem Engpass im Netz müssen koordiniert den Multicastgruppen beitreten bzw. diese verlassen

Router-basierte Multicast-Staukontrolle



■ Router-basiert

- zusätzliche Mechanismen zur Staukontrolle in Routern
Beispielsweise:
 - Bedingte Beitritte zu einer Gruppe
 - Oberhalb einer angegebenen Verlustrate verweigert der Router den Beitritt
 - Filtern von Daten (Verwerfen von Paketen)
 - Router filtert Daten, die eine derzeit vernünftige Datenrate übersteigen
- Grundsätzlich erfordern Router-basierte Verfahren zusätzliche **Zustandshaltung in den Routern**
 - Werden derzeit deshalb nicht im Backbone zu erwarten sein
 - Typische Anwendung für aktive/programmierbare Netze

22

Next Generation Internet SS2010 – 7. Multicast Transport (R0)

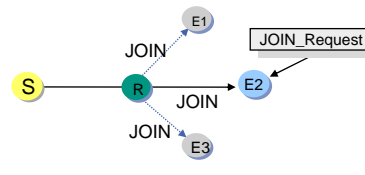
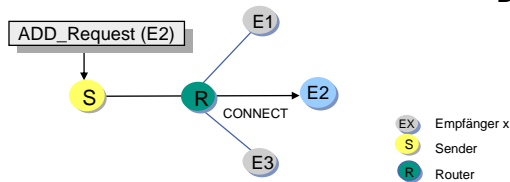


Institut für Telematik, Fakultät für Informatik
<http://tm.kit.edu/>

7.2.3 Multicast-Verbindungsaufbau



- **Sender-gesteuerter Beitritt** zur Multicast-Verbindung
 - Sender erhält die Aufforderung zur Aufnahme eines weiteren Empfängers (ADD_Request)
 - Sender übermittelt eine Verbindungsaufbaunachricht
- Geeignet z.B. für Konferenzsysteme mit einer zentralen Verwaltung der Kommunikationsteilnehmer
- **Empfänger-gesteuerter Beitritt** zur Multicast-Verbindung
 - Beitrittswillige Empfänger senden JOIN_Request
 - Sender übermittelt unter Nutzung der Multicast-Adresse den Beitrittswunsch (JOIN) an alle Gruppenmitglieder
 - Nur der Empfänger reagiert auf den Beitrittswunsch
- Geeignet z.B. für „Video on Demand“ oder „Internet Radio“



23

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
<http://tm.kit.edu/>

- Sender-gesteuerter Beitritt zur Multicast-Verbindung
 - Sender erhält (durch ein übergeordnetes Protokoll) die Aufforderung zur Aufnahme eines weiteren Empfängers (ADD_Request)
 - Sender übermittelt eine Verbindungsaufbau-Dateneinheit (CONNECT) an den beitrittswilligen Empfänger
 - Hierfür sind implizite und Handshake-basierte Verfahren möglich
- Geeignet z.B. für Konferenzsysteme mit einer zentralen Verwaltung der Kommunikationsteilnehmer
- Empfänger-gesteuerter Beitritt zur Multicast-Verbindung
 - Beitrittswilliger Empfänger fordert den Beitritt zur Multicast-Kommunikation an (JOIN_Request)
 - Sender übermittelt unter Nutzung der Multicast-Adresse den Beitrittswunsch (JOIN) an alle Gruppenmitglieder
 - Nur der Empfänger reagiert auf den Beitrittswunsch; die restlichen Kommunikationsteilnehmer ignorieren die JOIN-Dateneinheit
- Geeignet z.B. für „Video on Demand“ oder „Internet Radio“

7.3 Multicast-Transportprotokolle: Ausgewählte Beispiele



- Es gab bereits vor einiger Zeit eine Reihe von experimentellen Transportprotokollen
Beispiele sind
 - MTP (Multicast Transport Protocol)
 - RMP (Reliable Multicast Protocol)
 - RMTP (Reliable Multicast Transport Protocol)
 - PGM (Pragmatic General Multicast)
- Zuverlässigkeit steht beim Entwurf der meisten Protokolle im Mittelpunkt
- Im Folgenden werden einige Beispiele präsentiert. Konkret etabliert hat sich heute noch kein Protokoll.

24

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
<http://tm.kit.edu/>

7.3.1 Multicast Transport Protocol (MTP)



Eigenschaften von MTP

- Halbzuhverlässiger Multipeer-Transportdienst mit globaler Ordnungserhaltung
- Empfänger-gesteuert

Anforderungen

- Unicast- und Multicast-fähige Vermittlungsschicht

Grundlegendes Schema

- Token regelt Sendeberechtigung
- Multicastgruppe wird bei MTP als **Web** bezeichnet und für die Mitglieder werden drei Rollen unterschieden
 - **Master** – einmal pro Gruppe, kontrolliert das Verhalten des Webs → dient zur Ordnungserhaltung
 - **Producer** – Sender von Nutzdaten, erhält Kontrolldateneinheiten
 - **Consumer** – Agiert nur als Empfänger

25

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
http://tm.kit.edu/

- Grundlegendes Schema
 - Token regelt Sendeberechtigung
 - Nur diejenige Station, die im Besitz des Tokens ist, darf senden
 - Multicastgruppe wird bei MTP als **Web** bezeichnet und für die Mitglieder werden drei Rollen unterschieden
 - Master
 - Existiert genau einmal pro Gruppe
 - Kontrolliert das Verhalten des Webs. Sorgt für die globale Reihenfolgeerhaltung.
 - Producer
 - Kann als Sender von Nutzdaten an die Multicastgruppe fungieren
 - Erhält Kontrolldateneinheiten von den Empfängern
 - Multipeer-Kommunikation wird unter Verwendung mehrerer Producer in einem Web realisiert
 - Consumer
 - Agiert nur als Empfänger, nicht als Sender
- Vor dem Austausch von Daten muss bei MTP zunächst das Web aufgebaut werden
- Vergabe der Master-Rolle
 - Der erste Teilnehmer fungiert als Master
 - Es wird mit speziellen Kontrolldateneinheiten überprüft, dass tatsächlich auch nur ein Master aktiv ist
- Neue Teilnehmer
 - Teilnehmer sendet JOIN-REQUEST an die Multicast-Adresse der Gruppe
 - Parameter: Rolle des Teilnehmers, Zuverlässigkeitsklasse (unzuverlässig, zuverlässig), Kommunikationsform (m:n, 1:n)
 - Master antwortet per Unicast
 - Kann Beitritt ablehnen, beispielsweise wenn in ein 1:n-Web ein weiterer Producer beitreten will
- Nachricht
 - Grundlegende Einheit bei MTP, die von der Anwendung übergeben wird
 - Kann für den Datenaustausch zwischen MTP-Instanzen in mehrere Dateneinheiten segmentiert werden
 - Jede Nachricht wird mit einer Sequenznummer versehen
- Producer fordert periodisch Token zum Senden einer Nachricht beim Master an
- Zustände eines Token
 - IDLE: Bisher wurde für die Sequenznummer kein Token ausgewiesen und sie fällt nicht in den durch den Statusvektor ausgewiesenen Bereich
 - Keine Vergabe des Tokens möglich
 - PENDING: Token wurde vergeben, Master hat keine passende Dateneinheit gesehen
 - Token könnte verloren gegangen sein, daher erfolgt erneute Zuweisung
 - BUSY: Master hat Dateneinheiten mit der entsprechenden Nachrichten-Sequenznummer erhalten
 - Neue Anforderung wird verworfen; Producer ist bereits im Besitz eines Tokens
- Producer gibt Token am Ende der Nachricht wieder frei

Datentransfer



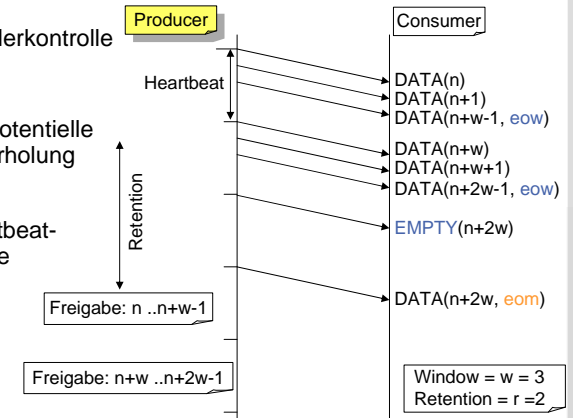
Drei grundlegende Parameter

- **Heartbeat**
 - Für Raten- und Fehlerkontrolle
- **Window**
- **Retention**
 - Speicherdauer für potentielle Übertragungswiederholung → halb-zuverlässig

Ablauf

- Producer muss im Heartbeat-Intervall mindestens eine Dateneinheit senden
 - **EMPTY**, falls sonst nichts zu senden
- **Tokenfreigabe**
 - Durch **eom** angezeigt
- Fenster ausgeschöpft
 - Durch **eow** signalisiert

Schema



26

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
http://tm.kit.edu/

- Drei grundlegende Parameter
 - **Heartbeat**
 - Zeitintervall in Millisekunden
 - Bei allen Teilnehmern gleich
 - Für Raten- und Fehlerkontrolle
 - **Window**
 - Anzahl der Dateneinheiten, die Producer während Heartbeat-Intervall maximal senden darf
 - **Retention**
 - Zeitintervall, in dem der Producer nach der Übertragung Daten für eine eventuelle Wiederholung bereithalten muss
 - Halbzuhverlässig!

Fehlerkontrolle



Ablauf

- Empfänger erkennt Verlust anhand der Sequenznummer, die höher ist als erwartet
- Negative, selektive Quittungen
- Wiederholungen unterliegen der Ratenkontrolle
 - Sind Daten nicht mehr vorhanden, antwortet der Producer mit NACK-DENY

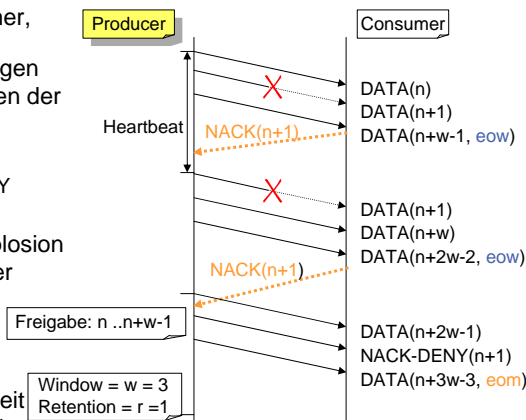
Bewertung

- Problem der Quittungs-Implosion
- Gefahr, dass alle Consumer quasi gleichzeitig reagieren

Partitionierung des Netzes

- Consumer erhält während Heartbeat keine Dateneinheit
 - Abbruch der Kommunikation

Schema



27

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
http://tm.kit.edu/

- Ablauf
 - Empfänger erkennt Verlust anhand der Sequenznummer, die höher ist als erwartet
 - Negative, selektive Quittungen
 - Wiederholungen unterliegen der Ratenkontrolle
 - Sind Daten nicht mehr vorhanden, antwortet der Producer mit NACK-DENY
- Bewertung
 - Problem der Quittungs-Implosion
 - Gefahr, dass alle Consumer quasi gleichzeitig reagieren
- Partitionierung des Netzes
 - Consumer erhält während Heartbeat keine Dateneinheit
 - Abbruch der Kommunikation
- Auslieferungsreihenfolge
 - Wird durch die Nachrichten-Sequenznummer bestimmt
 - Gültige Nachrichten sind nur solche, die der Master vollständig erhalten hat
- Zustände einer Nachricht
 - ACCEPTED: Master hat Nachricht vollständig erhalten
 - Darf an Anwendung weitergegeben werden
 - PENDING: Beim Master stehen Dateneinheiten einer Nachricht aus und Producer ist noch aktiv
 - REJECTED: Beim Master stehen Dateneinheiten einer Nachricht aus und er nimmt an, dass der Producer nicht mehr erreichbar ist
 - Nachricht muss verworfen werden und darf nicht an Anwendung ausgeliefert werden
- Statusvektor
 - Enthält den Zustand von maximal 12 aufeinanderfolgenden Nachrichten

27

7.3.2 Reliable Multicast Protocol (RMP)



Eigenschaften

- Ring-basiertes Protokoll
- Zuverlässige und ordnungserhaltende m:n-Kommunikation

Anforderungen

- Basiert auf UDP und IP

Grundlegendes Schema

- Modifikation des Token-Passing-Protokolls

Token-Ring

- Grundlegende Einheit zur Gruppenkommunikation
- Basis zum Ordnen der Dateneinheiten innerhalb einer Gruppe
- Alle Mitglieder einer Gruppe sind im gleichen Token-Ring angeschlossen
- Gleichzeitige Mitgliedschaft in mehreren Token-Ringen ist möglich

28

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
http://tm.kit.edu/

- Senderecht
 - Jede Station im Ring kann Daten senden
- Dateneinheiten gekennzeichnet durch
 - Sender-ID
 - Daten-Sequenznummer bezogen auf den Sender
 - Gewünschter Grad der Zuverlässigkeit
- Quittungen
 - Negative Quittung zur Anforderung von Wiederholungen
 - Positive Quittung von ausgezeichnetem Empfänger (Token-Halter) per Multicast an gesamte Gruppe
- Globale Sequenznummer
 - Stellt Basis für globale Ordnung dar
 - Wird mit der Quittung assoziiert und kontinuierlich inkrementiert
 - Ist also auch erst nach dem Empfang einer Quittung bekannt
- Gemeinsame Sicht der Gruppenmitgliedschaft: Erforderlich zur Bereitstellung eines komplett zuverlässigen Multicast- bzw. Multipeer-Dienstes
- Liste der Gruppenmitglieder ist in jedem Knoten verfügbar
- Ringmanagement
 - Fehler wird angenommen, falls für eine gewisse Zeit keine Quittung empfangen wurde
 - Trägt im Wesentlichen zur Komplexität bei

28

7.3.3 Reliable Multicast Transport Protocol (RMTP)



■ Eigenschaften

■ Halb-zuverlässiger Multicast

- Zielanwendung von RMTP sind Verteildienste (z.B. Software-Versionen oder Aktienpreise)

■ Grundlegendes Schema

- Aufbau eines **Baums**, in dem sich die Empfänger befinden
- Lokale Übertragungswiederholungen, um Skalierbarkeit zu erhöhen
 - Auswahl sogenannter **Designated Receiver**
 - Hierfür werden ausgesuchte Empfänger herangezogen
 - Aspekte der Sicherheit in Bezug auf Designated Receiver werden nicht angesprochen
 - Verarbeiten Quittungen für den „darunter“ liegenden Teilbaum
- Zustandshaltung ist unabhängig von der Gruppengröße
- Mechanismen zur Fluss-, Raten- und Staukontrolle werden bereitgestellt

29

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
http://tm.kit.edu/

Fehlerkontrolle



■ Empfängerorientiert

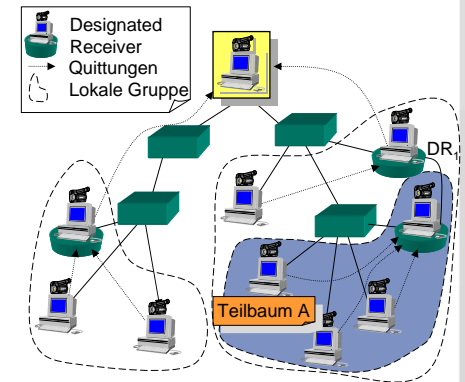
- Periodische positive, selektive Quittungen
 - Anzahl der Quittungen steigt linear mit Anzahl der Empfänger
 - Quittungen werden an Designated Receiver gesendet
 - DR wiederholt Daten (per Multicast oder Unicast)
- Designated Receiver kann Dateneinheit anfordern, falls er sie nicht besitzt
- Sendewiederholungen sind auf entsprechenden Teilbaum limitiert

■ Subcasting

- Weiterleiten der Daten nur auf Teilbaum, der logisch unterhalb des Routers liegt
- In Multicast-IP nicht verfügbar, deshalb wird Tunneling verwendet

■ Designated Receiver

- Dynamische Auswahl erforderlich, wegen Ausfall oder Netzpartitionierung



30

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
http://tm.kit.edu/

- Ratenkontrolle
 - Legt fest, in welchem Zeitintervall ein gesamtes Flusskontrollfenster gesendet werden darf
- Fensterbasierte Flusskontrolle
 - Weiterschalten des Fensters bei Empfang einer Quittung
 - Langsame Empfänger werden berücksichtigt
- Staukontrolle
 - Verlust von Dateneinheiten wird als Indikator für eine Stausituation angesehen
 - In welchem „traditionellen“ Transportprotokoll wird dies ebenso gehandhabt?
 - Staukontrollfenster wird verwaltet

7.3.4 Pragmatic General Multicast (PGM)



- Eigenschaften
 - Halbzuhverlässiges Multipeer-Protokoll
 - Ordnungserhaltung liegt beim Empfänger: Source-Ordered oder nicht geordnet
- Anforderungen
 - Unzuverlässiger Multicastdienst, z.B. Multicast-IP
- Grundlegendes Schema
 - Empfänger-basierte Fehlerkontrolle mit NACKs
 - Router sind aktiv an der Bereitstellung des Multicastdienstes beteiligt (NACK-Vermeidung)
 - Aufhebung der strikten Protokollebenen-Trennung
 - Gruppenmitglieder müssen nicht bekannt sein
- Bemerkung
 - An der Entwicklung war ein namhafter Routerhersteller beteiligt, trotzdem konnte sich das Protokoll bis jetzt nicht durchsetzen.

31

Next Generation Internet SS2010 – 7. Multicast Transport (R0)

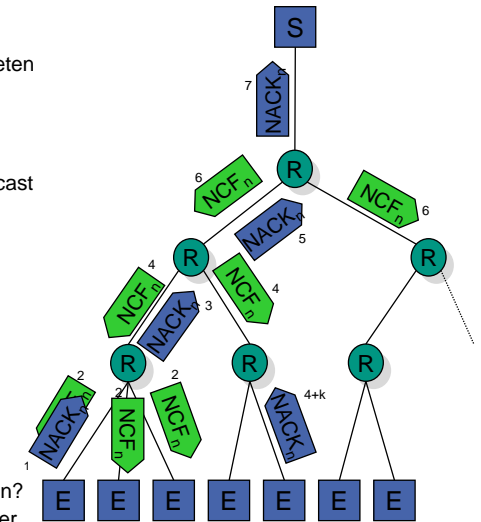


Institut für Telematik, Fakultät für Informatik
http://tm.kit.edu/

Fehlerkontrolle



- Negative Quittungen
 - NACKs per Unicast an übergeordneten Router
 - Verteilbaum muss bekannt sein
 - Periodische Source Path State Messages (SPM) vom Sender aus
 - NACK Bestätigung (NCF) per Multicast auf Interface, über das NACK empfangen wurde
 - Vorteile beim Endsystem? Beim Router?
- Unterdrückung gleicher NACKs im Teilbaum soll Quittungs-Implosion vermeiden
 - Statushaltung in den Routern pro empfangenem NACK
 - Welche NACKs schon empfangen?
 - Auf welchen Interfaces gleiche NACKs empfangen?
- Wiederholungen werden vom Sender per Multicast an die Gruppe gesendet



32

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
http://tm.kit.edu/

- Erweiterung
 - Neben Sender können Wiederholungen von ausgezeichneten Endsystemen gesendet werden

Flusskontrolle



■ Sender-gesteuert

■ Flusskontrollfenster

- Durch verschiedene Kontrollschemas zu beeinflussen
 - Keines fest implementiert
- Definiert durch die Menge an Daten, die sich der Sender für mögliche Übertragungswiederholungen merken muss
- Gegeben durch die Zeit, die eine Dateneinheit gespeichert bleibt

■ Bandbreitenregulierung durch Ratenkontrolle

- Maximale Übertragungsrate definiert

33

Next Generation Internet SS2010 – 7. Multicast Transport (R0)





Institut für Telematik, Fakultät für Informatik
<http://tm.kit.edu/>

7.3.5 IETF Reliable Multicast Transport (RMT) Working Group



■ Ziele

- Zunächst ausschließlich 1:n-Multicast
- Zwei Protokolldefinitionen für unterschiedliche Anforderungen
 - NACK Oriented Reliable Multicast (NORM)
 - Asynchronous Layered Coding (ALC)
- Definition von Building-Blocks für RMT
 - Für mehrere Protokolle
 - Forward Error Correction Building Block  [RFC5052]
 - Basic Schemes & verschiedene Verfahren (z.B. Reed-Solomon)
 - NACK Building Block  [RFC3941]
 - Layered Coding Transport (LCT) Building Block
 - Basis für ratenadaptive Transportprotokolle
 - Building Blocks für Ratenkontrolle und TCP-freundliche Staukontrolle
- Generelle Diskussionen (z.B. Security für RMT)

34

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
<http://tm.kit.edu/>

- Definitionen für 1:n Multicast
- Protokolldefinitionen:
 - NORM: NACK-basiert
 - ALC: mit FEC-Mechanismen, kein Feedback vom Empfänger nötig
- Building-Blocks:
 - grobgranulare Komponenten (Bausteine), die von mehreren Protokollen genutzt werden können
 - Abstrakte API-Definition
 - FEC Building Blocks (teilweise neue Drafts verfügbar):
 - Basic Schemes: Design Rahmen für FEC-Building-Blocks
 - Konkrete FEC Verfahren (z.B. Reed-Solomon)
 - NACK Building Block: Basis für NACK-basierte Protokolle
 - LCT Building Block: Basis für ratenadaptive Transportprotokolle

NACK Oriented Reliable Multicast (NORM)



[RFC3940, RFC3941]

- Ende-zu-Ende-Transport
 - Zwischenknoten benötigen nur Standard-IP-Multicast Support
- TCP-kompatible Staukontrolle
- Nutzt FEC und NACK Building Blocks
- Selektive, negative Quittungen
- NACK-Unterdrückung
 - Skalierbarkeit: für zehntausende Teilnehmer
 - Empfänger: Zustand pro Sender
 - Sender: Zustand pro Empfänger, der Staukontrollinformationen liefert
- Soll auch in dynamischen, fehlerbehafteten Umgebungen gut funktionieren
 - Mit hohen Verlustraten und Delays
 - z.B. in mobilen/drahtlosen Netzen

35

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
http://tm.kit.edu/

- Staukontrolle: nur eine Übertragungsrate für die gesamte Gruppe
- Das NORM Protokoll und der Multicast NACK Building Block werden aktuell weiter überarbeitet → Standards Track

Asynchronous Layered Coding (ALC)



[RFC3450]

- Nutzt
 - Layered Coding Transport Building Block
 - Multi-Rate Congestion Control Building Block
 - FEC Building Block
- Skalierbarkeit:
 - Millionen von Empfängern
 - Hunderte von Gigabytes pro Session
 - Minimales Protokoll
 - „Bootstrapping“ nicht definiert
 - Feedback an Sender nicht vorgesehen
- Ansätze basierend auf ALC (Drafts):
 - FLUTE: File Delivery over Unidirectional Transport
 - Beispiel-Protokoll für die Nutzung von ALC
 - Erlaubt Signalisierung von Dateieigenschaften (Meta-Daten) und Multiplexing von Dateien
 - FCAST: Scalable Object Delivery
 - Nutzung von ALC für beliebige Objekte (beliebige Meta-Daten)

36

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
http://tm.kit.edu/

- Skalierbarkeit:
 - Bootstrapping: Bezug der Sitzungsinformationen, die der Empfänger benötigt, um den richtigen Kanälen beitreten zu können
 - Feedback an Sender: Nicht vorgesehen, da schlecht für die Skalierbarkeit – wenn die Anwendung ein solches Feedback für Statistiken benötigt, dann kann sie diese nicht über ALC beziehen
- Auf ALC basierende Ansätze müssen fehlende Funktionalitäten (Bootstrapping, Empfänger-Feedback, Austausch von Meta-Daten etc.) nachrüsten: Bsp. FLUTE, FCAST
- Aktuell weitere Überarbeitung, um Experimental-Status in Standards Track zu überführen

Layered Coding Transport (LCT)



■ Multi-Raten Zustellung

- Mehrere Kanäle mit unterschiedlichen Raten pro Multicast-Session (mit einem Sender)
- Empfänger abonnieren Kanäle entsprechend der zur Verfügung stehenden Bandbreite
- Dadurch empfangergesteuerte Staukontrolle möglich
 - Ohne Feedback, d.h. massiv skalierbar
 - Mechanismus fest eingebaut, erlaubt aber beliebige Verfahren

■ Coding

- Wenn Daten durch Kodierung auf unterschiedliche Kanäle aufgeteilt werden können
 - Erlaubt dem Empfänger, die Qualität zu steuern
 - Ursprünglich für Audio-/Video-Streams verwendet
 - Für Bulk-Daten: Unterschiedliche Redundanz mit FEC

37

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
<http://tm.kit.edu/>

- Staukontrolle ohne Feedback: empfangergesteuert durch Beitreten oder Verlassen der Kanäle
- Ebenfalls weitere Überarbeitung, um Experimental-Status in Standards Track zu überführen

7.4 Übungen



- | | |
|---|--|
| 7.1 Welche Gründe gibt es für die Einführung halb-zuverlässiger Multicastdienste? | 7.8 Wie unterscheiden sich Unicast und Multicast hinsichtlich der Sicherheitsmechanismen? |
| 7.2 Erläutern Sie die Unterschiede einer Ordnungserhaltung bei Unicast und Multicast. | 7.9 Wie unterscheiden sich die Token-Mechanismen von MTP und RMP? |
| 7.3 Weshalb wird ein einziges Multicastprotokoll nicht ausreichend sein? | 7.10 Kategorisieren Sie die vorgestellten Protokolle hinsichtlich des realisierten Multicast-Dienstes. |
| 7.4 Vergleichen Sie den Sender-gesteuerten mit dem Empfänger-gesteuerten Beitritt. | 7.11 Wie garantiert MTP globale Ordnungserhaltung? |
| 7.5 Erläutern Sie die Vor- und Nachteile von Empfängerlisten. | 7.12 Vergleichen Sie den Aufwand für Übertragungswiederholungen bei MTP und RMP. |
| 7.6 Bei welchen Typen von Gruppen bieten sich Empfängerlisten an? | 7.13 Diskutieren Sie die Vor- und Nachteile von PGM. |
| 7.7 Stellen Sie die Verfahren zur Multicast-Fehlerkorrektur gegenüber. | 7.14 Wie werden unterschiedliche Übertragungsraten bei ALC realisiert? |

38

Next Generation Internet SS2010 – 7. Multicast Transport (R0)



Institut für Telematik, Fakultät für Informatik
<http://tm.kit.edu/>

7.5 Literaturhinweise



- [Atwo04] J. Atwood: A Classification of Reliable Multicast Protocols, IEEE Network Magazine, May/June 2004
- [DiDC97] C. Diot, W. Dabbous, J. Crowcroft; Multipoint communications: A survey of protocols, functions and mechanisms; IEEE Journal on Selected Areas in Communications, Vol. 15, No. 3, April 1997
- [Paul98] S. Paul; Multicasting on the Internet and its Applications; Kluwer Academic Publishers, 1998
- [ToKP97] D. Townsley, J. Kurose, S. Pingali; A comparison of sender-initiated and receiver-initiated reliable multicast protocols; IEEE Journal on Selected Areas in Communications, Vol. 15, No. 3, April 1997
- [WiZi01] R. Wittmann, M. Zitterbart; Multicast Communication: Protocols and Applications; Morgan Kaufmann Publishers, 2001
- [RFC 2887] M. Handley, S. Floyd, B. Whetten, R. Kermode, L. Vicisano und M. Luby. The Reliable Multicast Design Space for Bulk Data Transfer. RFC 2887 (Informational), August 2000. URL: <http://www.ietf.org/rfc/rfc2887.txt>



Literaturhinweise (2)



- [RFC 3048] B. Whetten, L. Vicisano, R. Kermode, M. Handley, S. Floyd und M. Luby. Reliable Multicast Transport Building Blocks for Oneto-Many Bulk-Data Transfer. RFC 3048 (Informational), Januar 2001. URL: <http://www.ietf.org/rfc/rfc3048.txt>
- [RFC 3451] M. Luby, J. Gemmell, L. Vicisano, L. Rizzo, M. Handley und J. Crowcroft. Layered Coding Transport (LCT) Building Block. RFC 3451 (Experimental), Dezember 2002. URL: <http://www.ietf.org/rfc/rfc3451.txt>
- [RFC 3738] M. Luby und V. Goyal. Wave and Equation Based Rate Control (WEBRC) Building Block. RFC 3738 (Experimental), April 2004. <http://www.ietf.org/rfc/rfc3738.txt>
- [RFC 3940] B. Adamson, C. Bormann, M. Handley und J. Macker. Negative acknowledgment (NACK)-Oriented Reliable Multicast (NORM) Protocol. RFC 3940 (Experimental), November 2004. URL: <http://www.ietf.org/rfc/rfc3940.txt>
- [RFC 3941] B. Adamson, C. Bormann, M. Handley und J. Macker. Negative- Acknowledgment (NACK)-Oriented Reliable Multicast (NORM) Building Blocks. RFC 3941 (Experimental), November 2004. URL: <http://www.ietf.org/rfc/rfc3941.txt>



Literaturhinweise (3)



[RFC 4654] J. Widmer und M. Handley. TCP-friendly Multicast Congestion Control (TFMCC): Protocol Specification. RFC 4654 (Experimental), August 2006. URL: <http://www.ietf.org/rfc/rfc4654.txt>

[RFC 5052] M. Watson, M. Luby und L. Vicisano. Forward Error Correction (FEC) Building Block. RFC 5052 (Experimental), August 2007. URL: <http://www.ietf.org/rfc/rfc5052.txt>

