Multi Media Seminar

Thema

Routing im Internet

Istvan Bognar

10. Juli 1998

Betreuer: Walter Lange

Routing im Internet

Inhaltsverzeichnis

1 Einführung

Häufig wird das Internet auch als Netz der Netze bezeichnet. Dies hat seine Berechtigung darin, daß das Internet eigentlich nicht einzelne Computer, sondern Computernetze miteinander verbindet. Dabei stellt das Internet keine speziellen Anforderungen, wie diese Netze realisiert sind. Diese können z.B. als Ethernet, Token Ring usw. realisiert sein.

1.1 Wie ist das Netz aufgebaut?

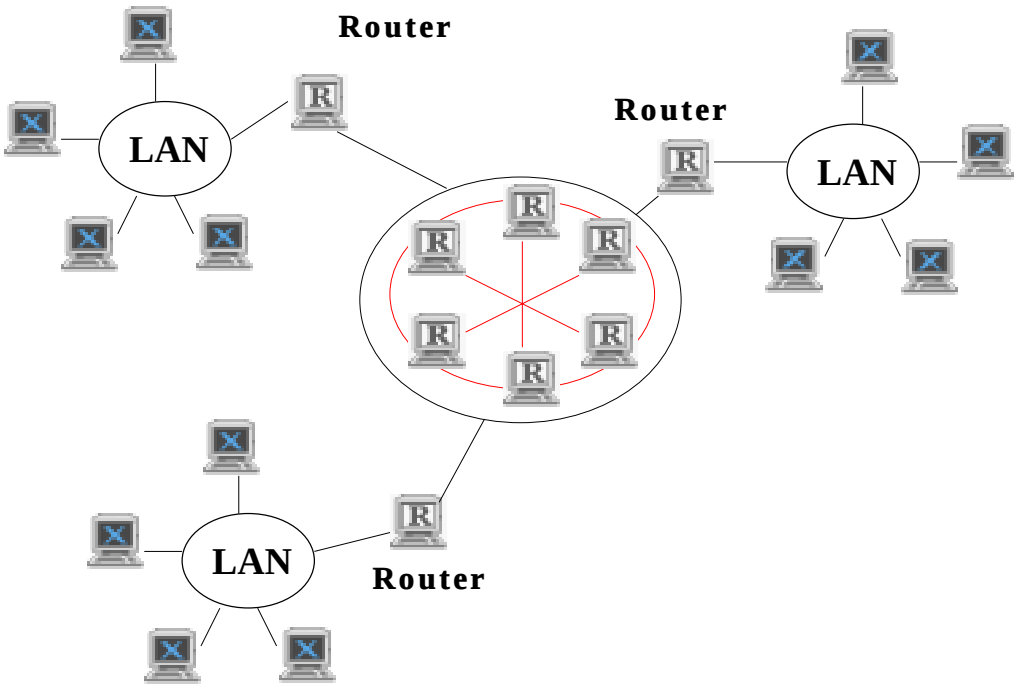

Abbildung 1: Internet - Logische Struktur

Die Verbindung vom lokalen Netz zum Internet wird durch sogenannte Gateways realisiert. Dies sind Rechner, die einerseits in das lokale Netz eingebunden sind, andererseits eine Verbindung zu einem anderen Netzwerk haben. Dieses Netzwerk ist wieder mit (mindestens) einem anderen verbunden usw. Alle auf diese Weise verbundenen Netze bilden zusammen das Internet.

Da nicht jedes Netz direkt mit jedem verbunden ist, verläuft die Kommunikation in der Regel indirekt über unbeteiligte dritte Netzwerke. Der Weg, den ein Internetpaket auf seinem Weg zum Empfänger nimmt, wird dabei von den Routern bestimmt. Dies sind spezialisierte Rechner, die aus jedem Datenpaket

die Internetadresse des Empfängers auslesen und anhand dieser Adresse bestimmen, an welchen Computer das Paket weitergereicht wird. Für diese Aufgabe unterhält jeder Router Routingtabellen, in denen für jede Adresse der nächste anzusprechende Rechner (und evtl. Ausweichrechner) eingetragen sind. Die Routingtabellen enthalten natürlich nicht die Adressen aller im Internet erreichbaren Computer. Die Namenshierarchie des DNS[1] erlaubt - ähnlich einer Postanschrift - eine einfache Aufteilung nach Land, Teilnetz usw.

1.2 Routing

Beim statischen Routing hat jeder Knoten eine Tabelle mit folgenden Einträgen: Quelle, Ziel und ausgehende Verbindungsstrecke. Ein eingehendes Paket enthält die Zieladresse (oder die Verbindungsnummer bei Virtual-Circuits) und die Routingentscheidung reduziert sich auf einen Zugriff auf die Routing-Tabelle. Wenn sich die Netzwerktopologie ändert, so wird in dem Netzwerkkontrollzentrum die globale Routingtabelle neu berechnet und an alle Knoten verteilt. Bekannte Algorithmen, wie zum Beispiel Dijkstra's „Kürzester-Pfad-Algorithmus", können benutzt werden, um das Optimum zu berechnen, falls die Netzwerktopologie als Graph mit Knoten und gewichteten Kanten betrachtet wird. Die Vorteile des statischen Routings sind die Einfachheit der Algorithmen und die Geschwindigkeit im laufenden Betrieb. Nachteile sind die fehlende Flexibilität, der Verlust der Optimalität der Routingentscheidung, wenn die Netzwerktopologie geändert wird bzw. sich die Netzlast ändert und Performance- und Zuverlässigkeitsprobleme, die typisch für alle zentralisierten Algorithmen sind. Viele X.25 Netzwerke benutzen weltweit das statische Routing, wie auch SNA (Systems Network Architecture) von IBM.

1.2.1 Shortest Path Algorithmen

Gegeben sei ein Graph G mit einer Menge 'N' von Knoten und einer Menge 'E' von Kanten:

1. Initialisiere die Menge 'P' von bekannten Wegen mit dem Startknoten N0 .
2. Finde alle unmittelbaren Nachbarknoten der Knoten, die bereits ein Element der Menge der bekannten Wege 'P' sind; jeder Weg, der in so einem Knoten endet, ist ein Kandidatenweg.
3. Sortiere die Kandidatenwege nach ihrer Länge (d.h. Summe der gewichteten Kanten).
4. Füge den kürzesten Kandidatenweg zu 'P' hinzu.

[1] Domain Name System - ordnet den logischen Namen im Internet (z.B. inca.informatik.uni-tuebingen.de $\Leftrightarrow$ 134.2.14.43) die Internetadresse als 32-bit-Zahl zu

5. Sind noch unverbundene Knoten übrig, so gehe zu 2.

Wenn der Algorithmus terminiert, ist P die Menge der kürzesten Pfade vom Startknoten N0 zu jedem anderen Knoten.

Dijkstra's Algorithmus wird oft als 'Shortest Path First' (SPF) bezeichnet. Ist die komplette Topologie des Netzwerkes gegeben und allen Kanten eine gute metrische Gewichtung zugeordnet, kann man den SPF Algorithmus verwenden, um die optimalen Wege zu ermitteln. Dies wird typischerweise beim statischen Routing, besonders in X.25 und SNA Netzwerken gemacht.

1.2.2 Distance Vector Routing

Der erste weit verbreitete Routing Algorithmus im Internet wird 'Distance Vector Routing' genannt. Jeder Knoten pflegt sein Wissen über die kürzesten Entfernungen zu anderen Knoten im Netzwerk. Eine Tabelle enthält einen Eintrag für jeden bekannten Knoten, die Entfernung zu diesem und den Verbindungsweg, um ihn zu erreichen. Die Tabelle wird deshalb auch 'Distance Vector' genannt.

Im folgenden wird vereinfachend angenommenan, daß Entfernungen durch die Anzahl der Hops auf der Strecke gemessen wird. Beispielsweise ist die Entfernung immer eins für direkte Nachbarn im Netzwerk. Wenn ein Knoten die Nachricht von einem Nachbarn erhält, daß sich eine Entfernung geändert hat, so erneuert er seine lokale Routing-Tabelle wie folgt:

- Falls ein unbekannter Knotenname erscheint, so füge einen Eintrag in die lokale Tabelle ein, mit der Entfernung von i+1 (eine Weglänge um den Nachbar zu erreichen).
- Falls aus der Nachricht hervorgeht, daß es einen kürzeren Pfad zu einem bekannten Knoten gibt, so aktualisiere den Eintrag in der Tabelle entsprechend.
- Falls ein ankommender Eintrag zu einem längeren Pfad als dem bekannten führt, ignoriere ihn.
- Falls ein ankommender Eintrag die Entfernung hat, so entferne den Zielknoten aus der lokalen Tabelle (ein Knoten oder Verbindungsweg ist auf dem Pfad über diesen Nachbar unerreichbar geworden).

Es ist einfach, den Distanz Vektor Routing Algorithmus so zu erweitern, daß gewichtete Verbindungen berücksichtigt werden. Anstatt die Kosten für jede Verbindung auf eins zu setzen, können verschiedene 'Gewichte' für die Verbindungen verwendet werden, die z.B. verschiedene Verzögerungen, Bandweiten oder einfach die aktuelle Belastung darstellen. Es ist auch möglich alternative Wege parallel

zu verwenden, die mit Hilfe von Wahrscheinlichkeiten ausgewählt werden, die indirekt proportional zu dem Gewicht der Verbindungen sind.

Distanz Vector Routing ist sehr einfach zu implementieren, sehr effizient und weitverbreitet im Internet. Jedoch hat es zwei große Probleme:

- unvollständig übertragene Nachrichten verbreiten sich nur langsam und
- in Übergangsperioden können Schleifen beim Routing auftreten.

1.3 Backbones

Bei großem Datenverkehr reicht es natürlich nicht aus, benachbarte Netzwerke direkt miteinander zu verbinden. Die Versendung von Daten müßte dann immer über viele unbeteiligte Netzwerke erfolgen; die Geschwindigkeit des Datenaustausches würde erheblich leiden. Darum wurden leistungsfähige Hauptverbindungen geschaffen, sogenannte Backbones. An diese Backbones werden lokale Netze an Übergabepunkten angeschlossen. Wichtige internationale und nationale Backbones sind:

- Der NSFNET-Backbone der National Science Foundation in den USA. Nach der Umstellung auf Übertragungsraten zwischen 155 und 622 Mbit/s wird dieser Backbone unter der Bezeichnung vBNS geführt. Er verbindet die großen Supercomputer-Zentren der USA. Eine Karte findet sich unter http://www.gov.mci.net/vBNS/network_map.html.
- Der EuropaNet-Backbone von Dante Ltd. , der die europäischen Wissenschaftsnetze miteinander verbindet. Eine Karte gibt es unter http://www.-dante.net/pics/EuropaMap.gif
- Der europäische Ebone, der 72 Internet-Provider in 31 Ländern miteinander verbindet. In Deutschland sind ECRC und NTG/XLINK angeschlossen. Nähere Informationen unter http://www.ebone.net/objectives.html.
- Das deutsche Wissenschaftsnetz (WIN), das sich mittlerweile zum Breitband-Wissenschaftsnetz mit Übertragungsraten bis zu 34 Mbit/s gemausert hat. Karten des Netzes sind zu besichtigen unter http://www.dfn.de/win/-home.html.
- Das Xlink-Netz des Internetproviders NTG/XLink, ein Tochterunternehmen der Bull AG Deutschland. Weitere Informationen und Karte unter http://www.xlink.net/xlink/topologie.
- Das ECRC-Netz, ein Joint-Venture der Firmen Bull (Frankreich), ICL (Großbritannien) und Siemens, das seit 1984 besteht. Eine Karte der Verbindungen auf http://www.ecrc.de/networking/infrastructure/map.html.

1.4 Technische Struktur

Um Nachrichten zwischen zwei Partnern zu übertragen, gibt es grundsätzlich zwei verschiedene Verfahren: verbindungsorientiert (Punkt-zu-Punkt) und paketorientierte Datenübertragungsverfahren.

- Beim verbindungsorientierten Verfahren wird - ähnlich wie beim Telefongespräch - für die Dauer der Übertragung eine (physische) Verbindung aufgebaut, die während der gesamten Datenübertragung aufrechterhalten werden muß.

- Im Gegensatz dazu besteht bei der paketorientierten Übertragung keine feste physische Verbindung zwischen den zwei Partnern. Die Daten werden in einzelne Blöcke aufgeteilt, denen jeweils ein Kopfteil (Header) vorangestellt wird. Dieser Header enthält u.a. Adressinformationen, die es ermöglichen, jedes dieser Pakete einzeln dem Empfänger über verschiedene Stationen hinweg zuzuleiten. Das läuft so ähnlich ab, wie beim Verschicken eines Briefes per Post. Die Daten kommen in den Umschlag, der Umschlag enthält die Headerinformation Empfängeradresse. Der Absender transportiert den Brief bis zum nächsten Postamt oder Briefkasten. Von dort wird er über verschiedene Verteilzentren weitergeleitet, bis er endlich dem Empfänger zugestellt wird. Die paketorientierte Datenübertragung ist daher, wie oben bereits beschrieben, effektiver und wesentlich weniger störungsanfällig.

Zur Darstellung der komplexen Aufgaben der Datenübertragung haben sich Schichtenmodelle durchgesetzt. Sie beschreiben die einzelnen Funktionen von der physischen Übertragungsschicht (elektrische und mechanische Parameter) bis hinauf zur Anwendungsschicht, die abstrahierend von allen darunter liegenden Schichten, dem Benutzer einen Dienst wie z.B. E-Mail zur Verfügung stellt.
Das bekannteste Schichtenmodell ist das 1983 von der ISO normierte 7-Schichten-OSI-Modell.

Die konkreten Regeln, nach denen die Kommunikation zwischen Rechnern abläuft, werden als Protokolle bezeichnet. Diese Protokolle sind normalerweise ebenfalls in Schichten aufgeteilt, d.h., Teilprotokolle übernehmen genau abgegrenzte Aufgaben ihrer Schicht und bieten der darüberliegenden Schicht einen entsprechenden Dienst an, der - ohne die genaue Funktionsweise zu kennen - in Anspruch genommen werden kann. Allerdings werden die Protokolle meist nicht in alle sieben ISO/OSI-Schichten zerlegt.

1.5 Protokolle

Die Internet-Protokolle können ebenfalls in ein Schichtenmodell eingeordnet werden. Es hat allerdings nur vier Ebenen. Die folgende Tabelle zeigt diese auch als

Internet Protocol Stack bezeichneten Ebenen und im Vergleich dazu die ISO/OSI-Schichten. Außerdem sind einige wichtige Anwendungs-Protokolle und die Zuordnung zu den Schichten dargestellt.

<table>
<tr><th colspan="7">Internet-Protocol-Stack und ausgewählte Internet-Dienste</th></tr>
<tr><th colspan="3">Anwendung | Darstellung | Sitzung</th><th>Transport</th><th>Netzwerk</th><th>Sicherung</th><th>Bittransport</th></tr>
<tr><td colspan="3">Prozess/Applikation</td><td>Host-to-Host</td><td>Internet</td><td colspan="2">LAN/Netzzugriff</td></tr>
<tr><td>File Transfer</td><td colspan="2">File Transfer Protocol (FTP) RFC 959</td><td rowspan="9">Transmission Control Protocol (TCP) RFC 793</td><td rowspan="2">Address Resolution Protocol (ARP) RFC 826</td><td rowspan="9">Ethernet Token Ring FDDI usw.</td><td rowspan="9">Übertragungsmedium</td></tr>
<tr><td>Electronic Mail</td><td colspan="2">Simple Mail Transfer Protocol (SMTP) RFC 821</td></tr>
<tr><td>Terminal Emulation</td><td colspan="2">Telnet Protocol (Telnet) RFC 854</td><td rowspan="7">Internet Protocol (IP) RFC 791</td></tr>
<tr><td>Usenet News</td><td colspan="2">Network News Transport Protocol (NNTP) RFC 977</td></tr>
<tr><td>Domain Name Service</td><td colspan="2">Domain Name System (DNS) RFC 1034</td></tr>
<tr><td>Gopher</td><td colspan="2">Internet Gopher Protocol RFC 1436</td></tr>
<tr><td>WAIS</td><td colspan="2">Z 39.50 RFC 1625</td></tr>
<tr><td>World Wide Web</td><td colspan="2">Hypertext Transfer Protocol (HTTP)</td><td rowspan="2">User Datagram Protocol (UDP) RFC 768</td><td rowspan="2">Internet Control Message Protocol RFC 792</td></tr>
<tr><td>Archie</td><td colspan="2">Prospero Protocol</td></tr>
</table>

1.5.1 Internet Protokoll (IP)

Das Internet-Protokoll ist das Basis-Kommunikationsprotokoll im Internet. Es überträgt die Daten

1. paketorientiert
2. verbindungslos (Jedes Paket wird für sich übertragen und für das Internet-Protokoll besteht kein Zusammenhang zwischen den Paketen.)

3. nicht garantiert (Im IP gibt es keinen Mechanismus, der für die wiederholte Übertragung verlorener Pakete sorgt.)

Die maximale Paketgröße des IP beträgt 65535 Bytes, die Mindestgröße, die jedes Gateway und jeder Router verarbeiten können muß, liegt bei 576 Byte. Im Header jedes Pakets legt das Internet-Protokoll u.a. Sende- und Empfangsadresse und eine Prüfsumme der Header-Felder ab. Interessant ist das Time-To-Live-Feld im Header, das die maximale Zeitdauer angibt, die das Paket im Netz verbringen darf. Wird die Zeit überschritten, wird das Paket verworfen.

1.5.2 Transmission Control Protocol (TCP)

Das Transmission Control Protocol steht in der Schichten-Architektur oberhalb des Internet-Protocols. Es benutzt den IP-Paketverschickungsdienst und stellt zusätzlich Mechanismen bereit, die überprüfen, ob ein Datenpaket tatsächlich beim Empfänger angekommen ist. Geht ein Datenpaket verloren, wird die Wiederholung der Übertragung angefordert. Daher benötigt TCP eine (virtuelle) Verbindung zur Gegenstation.

TCP arbeitet also im Gegensatz zu IP

1. verbindungsorientiert und

2. garantiert.

TCP segmentiert zunächst den Datenstrom (Standard-Segmentgröße 536 Byte) und numeriert diese Segmente. Die empfangende Station kann nun anhand dieser Segmentnummer den Empfang des Paketes bestätigen. Die Empfangsbestätigung wird allerdings nicht für jedes Paket einzeln abgewartet. Spätestens nach dem Senden einer in der sogenannten Fenstergröße (Sliding Window) genau festgelegten Zahl von Paketen muß aber die Bestätigung für eines der versandten Pakete eintreffen. Erst dann darf das nächste Paket versandt werden. Der Empfänger bestätigt mit einem Paket immer auch den Empfang der vorhergegangenen Pakete. Im Fehlerfall müssen alle Pakete seit der letzten Empfangsbestätigung erneut übertragen werden.

1.5.3 User Datagram Protocol (UDP)

Nicht alle Datendienste benötigen auf der Transportebene eine gesicherte Verbindung zwischen den Kommunikationspartnern. Das Transportprotokoll kann wesentlich einfacher gestaltet werden, wenn z.B. das Netzwerk selbst zuverlässig genug ist. UDP ist ein einfacher verbindungsloser Dienst, der gegenüber dem Internet-Protokoll zusätzlich Portnummern und eine Prüfsumme anbietet. Dadurch ist der Protokollaufwand gegenüber TCP wesentlich geringer und die Netzbelastung entsprechend niedriger.

1.5.4 IP Version 6

IP Version 6 (oder kurz: IPv6) ist die neue Version des Internet Protokolls IP. Weltweit wird gerade die IP Version 4 verwendet. Erweiterungen für Multimedia Daten Ströme sind die Hauptgründe für die neue Version, aber nicht die einzigsten. Andere Gründe sind ein größerer Adreßraum sowie Authentifizierungs- und Verschlüsselungsmerkmale.

Ein wichtiges Designziel von IPv6 ist die Kompatibilität zu IPv4. In einem so großen und heterogenen Netzwerk wie dem Internet würde es unmöglich sein, alle IP Knoten in einer Nacht-und-Nebel-Aktion umzuwandeln. Neue IPv6 Hosts und Routers werden auch mit dem alten IPv4 Hosts koexistieren können und ermöglichen so einen langsamen Übergang.

IPv6 beruht auf den gleichen Merkmalen wie die älteren IP Versionen: es ist verbindunglos (ein Datagram Protokoll) und es hat keine Fehler- oder Flußkontrolle in der Vermittlungsschicht. Es hat aber sehr attraktive neue Merkmale:

- Ein Adreßraum mit 128 Bits (anstatt mit 32 Bits) erlaubt es, viel mehr Rechner zu adressieren und ermöglicht mehre Hierarchieebenen zu verwalten.

- Ein verbessertes Multicast Adressierungsschema erlaubt die Einschränkung des Multicast Routings auf bestimmte Domainen. Durch das Bereichsfeld innerhalb der Multicast Adresse wird der Bereich der Gültigkeit einer Adresse eingeschränkt, zum Beispiel innerhalb eines Intranets eines Unternehmens. Ein zusätzliches Flag Feld erlaubt die Unterscheidung zwischen permanenten und temporären Gruppenadressen.

- Das neue Flow-Label Feld im Header erlaubt die Identifizierung von allen Paketen, die zu dem gleichen Datenstrom gehören (bezeichnet als Fluß (flow) in IP). Ein Fluß ist eine Reihe von Paketen, gesendet von einem Host zu einer einzelnen Adresse oder zu mehreren Adressen (Multicast). Auf diese Weise können alle Router entlang des Pfades die Pakete eines Flow identifizieren und sie auf einer speziellen Art behandeln. Sie können zum Beispiel Pakete, die zu einem Audio Strom gehören mit einer höheren Priorität behandeln, als solche, die zu einem Dateitransfer gehören. Die Flow ID ist die Schlüsseleigenschaft von IPv6 für die Mittelreservierungen und Quality of Service (QoS) auf der IP Ebene im Internet. In den älteren Versionen von IP war es nicht möglich Pakete zu identifizieren, die zu einem besonderen Multimediastrom (die Quell- und Zieladresse sind sicherlich nicht ausreichend) gehören und Mittelreservierungen und QoS waren unmöglich zu implementieren.

- Neue Verfahren zur Authentifizierung, Integrität und Datenverschlüsselung werden eingeführt.

Das neue Flow - Label Feld beinhaltet 28 Bits, vier Bits für die Verkehrsklasse, und 24 Bits für die Flow ID. Die Verkehrsklasse ist sehr ähnlich zu der Verkehrsklasse in ATM. Der Verkehrsklasse für einen ununterbrochenen Strom (z.B. Video oder Audio) wird eine höhere Priorität zugewiesen, als der Verkehrsklasse von traditionellen Flow-kontrollierten Datenströmen (z.B. ein TCP Strom). Wie bereits erwähnt, zerstört der Sliding Window Flow Control Algorithmus sowieso den ununterbrochen Datenfluß. Somit macht es keinen Sinn, den Paketen eines Sliding-Window Strom eine sehr hohe Priorität innerhalb von Routern zu geben. Ein Router wird typischerweise Pakete mit niedriger Priorität mit höherer Wahrscheinlichkeit verwerfen.

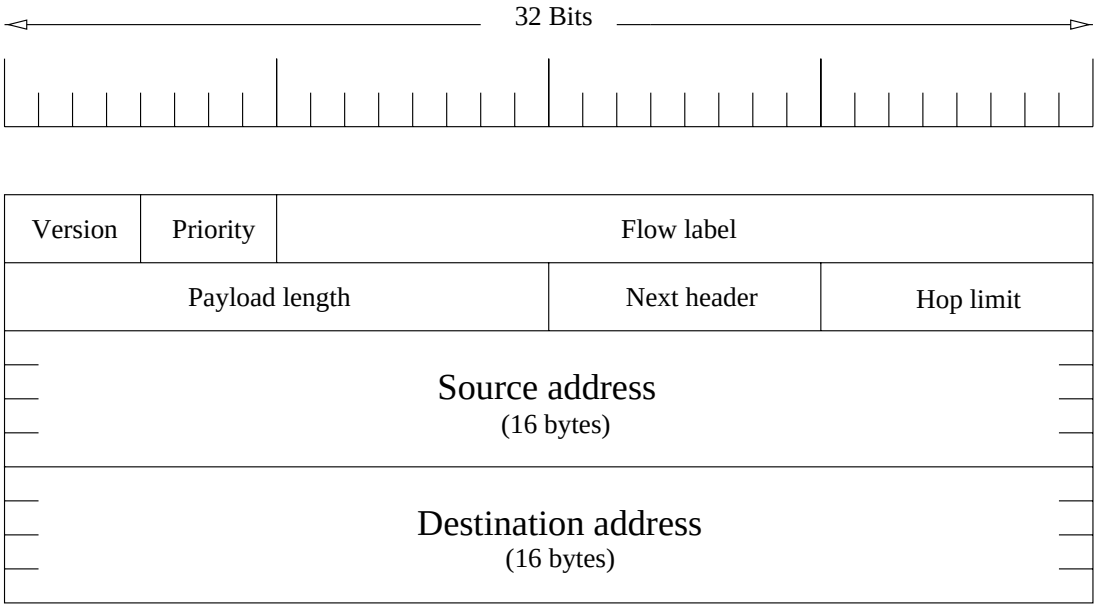

Abbildung 2: IPv6 Header

Die Flow ID ist eine von dem Quellknoten generierte pseudo Zufallsnummer. Zusammen mit der Quelladresse (welche sich in dem Header befindet) bildet es einen globalen eindeutigen Identifizierer für den Fluß. Wie lange pflegt man Informationen über den Fluß (z.B. Prioritätszuweisungen und Mittelreservierungen) in einem IPv6 Knoten? Das Protokoll ist verbindungslos und so kann man diese Informationen nicht bei Verbindungsende löschen. Der Vorschlag für IPv6 ist, den Zustand zu halten bis ein Timeout erreicht wird (d.h., es kamen innerhalb einer bestimmten Zeitspanne keine Pakete mehr von diesen Fluß an) oder bis ein gesondertes Kontrollpaket uns explizit auffordert, die Zustandsinformationen

wegzuwerfen. Dieses Konzept wird als *Soft State* bezeichnet. Die Internet Architekten behaupten, daß im Fall von sehr großen, heterogenen Netzwerken der Soft state robuster gegen Fehler und leichter zu verwalten ist, als der *Hard State* einer traditionellen virtuellen Verbindung.

1.5.5 RTP, ein Echtzeit Transfer Protokoll

RTP (Real Time Protocol) ist ein Transport Protokoll für Multimedia Daten Ströme im Internet. Es wurde entwickelt, um auf Multicast IP aufzusetzen und um Zeitinformationen und Stromsynchronisation bereitzustellen (tatsächlich werden RTP Pakete typischer weise über UDP, welches eine Anwendungs-Programmierschnittstelle zu IP bietet, versendet). Es ist ein einfaches Protokoll, ohne Fehlerkorrektur oder Flußkontrolle. Im Prinzip könnte RTP auch mit anderen Protokollen zusammenarbeiten. Die Anforderungen für die darunterliegenden Protokolle sind minimal.

Das Format des RTP Headers wird in Abb. 3 gezeigt. Das Zeitstempel Feld wird sowohl für die Intra Strom- als auch für die Inter Strom-Synchronisation verwendet. Es zeichnet die Entstehungszeit der ersten Bytes des Pakets auf. Die Zeitstempel- auflösung hängt vom Typ des Datenstroms ab. Zum Beispiel wird oft eine Zeitstempelfrequenz von 65,536 Hz für digitales Video verwendet, während Audio Ströme mit einem Zeitstempel, der der Sampling Rate entspricht, versehen werden. Mehrere nachfolgende RTP Pakete können den selben Zeitstempel tragen, wenn sie zur der selben Anwendungsdateneinheit des Stromes gehören, z.B. zu ein und demselben Videobild. Der Zeitstempel wird von dem Empfänger verwendet, um die Synchronisation mit Echtzeit und/oder anderen Multimedia Strömen zu gewährleisten. Der Synchronisation Quellenbezeichner (SSRC) ist eine Zufallsnummer, die vom Sender generiert wird und für die Zeitdauer einer RTP Sitzung eindeutig ist. Kollidierende SSRC Bezeichner werden entdeckt und aufgelöst. Der SSRC erlaubt dem Empfänger die eindeutige Identifizierung eines Datenstroms. RTP führt zwei Arten von Zwischenknoten zwischen Sender und Empfänger ein: Mixer und Übersetzer. Ein Mixer empfängt RTP Pakete von einem oder mehreren Sendern, fügt sie zu einem neuen RTP Paket zusammen und leitet sie weiter. Der Strom von kombinierten Paketen bekommt eine neue SSRC Nummer. Die SSRC Nummern von den beteiligten Sendern wird dem Paket als Quellenbezeichnung (CSRCs) beigefügt. Weil die Pakete von verschiedenen beteiligten Sendern ohne passende Synchronisierung ankommen können (sie können auf verschiedene Pfade durch das Netz gereist sein), ändert der Mixer die interne Struktur des Stromes. Ein Beispiel eines Mixers ist die Kombination von mehreren Audio Quellen einer Konferenz in einen einzigen Audio Strom, der zu allen Empfängern gesendet wird. Im Gegensatz zum Mixer ändert ein Übersetzer nur den Inhalt der Pakete, ohne die Ströme miteinander zu verbinden. Ein Video Verschlüsselungskonverter oder ein Firewall Filter sind Beispiele dafür.

RTP besitzt auch ein Kontrollprotokoll, RTCP. Es wird verwendet um Übertra-

gungsberichte zu allen Mitgliedern zu senden und die Performance und Qualität des Stromes zu überwachen. Es wird außerdem vom Sender dazu verwendet, die Eindeutigkeit der SSRC Nummern sicherzustellen.

Trotz der Tatsache, daß RTP ein sehr junges Protokoll ist, haben es einige Internet Anwendungen bereits implementiert, besonders die MBONE Werkzeuge *vic* und *vat*. Anders als TCP werden normalerweise RTP Algorithmen nicht als getrennte Schicht implementiert, sondern sind ein Teil des Anwendungskodes. Es wird erwartet, daß die nächste Generation von WWW Browsern das RTP für live Video und Audio Ströme verwenden.

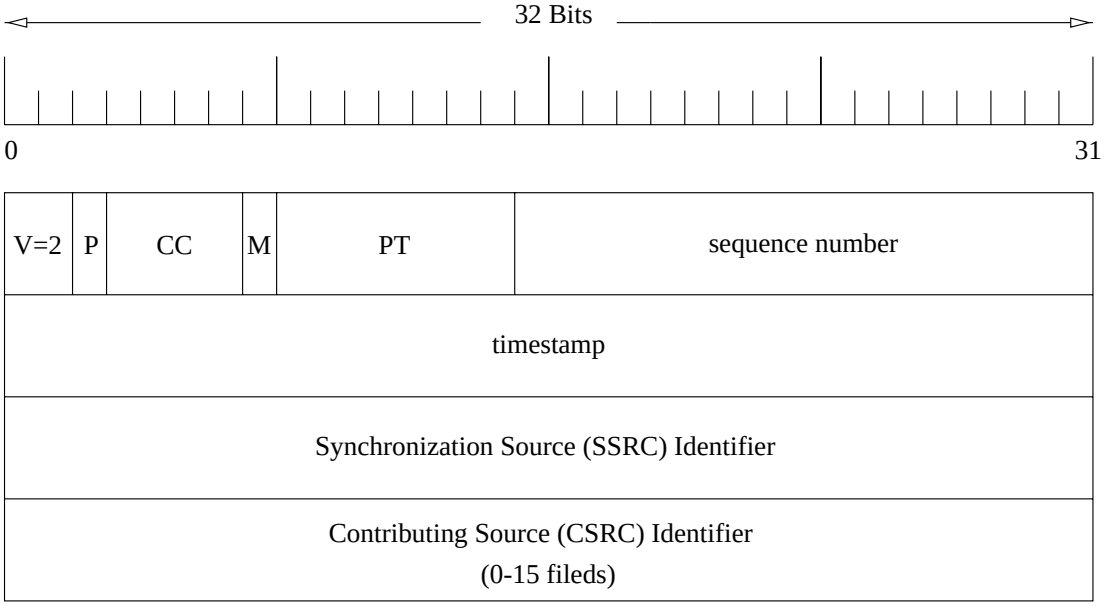

Abbildung 3: RTP Header

V = Version des Protokolls

P = Füllbytes an/aus (Wenn auf eins, dann sind Füllbytes am Ende des Pakets)

CC = CSRC Zähler (Anzahl von CSRC Feldern im Paket)

M = Markierung (zeigt wichtige Ereignisse an, wie das Ende des Datenstroms)

PT = Paket Typ (Typ des Payloads; definiert im einem RFC)

1.6 Der Weg eines Internetpakets

Wenn man von Tübingen aus eine Verbindung zum Rechner ftp.uni-stuttgart.de aufnimmt, könnten die Datenpakete etwa folgenden Weg nehmen:

```
prompt> traceroute ftp.uni-stuttgart.de
Tracing route to info2.rus.uni-stuttgart.de [129.69.18.15]:

  1. router11.zdv.uni-tuebingen.de (134.2.2.254) 2.527 ms 2.569 ms 2.483 ms
  2. router10.zdv.uni-tuebingen.de (134.2.250.254) 14.136 ms 22.945 ms
  3. Tuebingen1.BelWue.DE (129.143.62.1) 58.924 ms 45.926 ms 22.928 ms
  4. Tuebingen11.BelWue.DE (129.143.62.3) 24.288 ms 3.856 ms 6.153 ms
  5. Hohenheim1.BelWue.DE (129.143.1.197) 12.918 ms 7.598 ms 9.844 ms
  6. Stuttgart1.BelWue.DE (129.143.1.193) 116.171 ms 62.102 ms 62.867 ms
  7. BelWue-GW.Uni-Stuttgart.DE (129.143.70.9) 106.832 ms 14.443 ms 14.892
     ms
  8. info2.rus.uni-stuttgart.de (129.69.18.15) 12.785 ms 23.914 ms 16.696 ms

Trace complete.
```

Wie die Ausgabe des Programmes traceroute zeigt, geht die Verbindung zum ftp-server über acht Stationen. (Ein Test mit einem Server in Australien ergab insgesamt 20 Stationen.) Ohne auf die Einzelheiten des TCP/IP-Protokolls einzugehen, passiert in etwa folgendes bei der Übertragung:

- Der Rechner bzw. das TCP-Modul zerlegt die Nachricht in kleine Teile und verschnürt diese zu einzelnen Datenpaketen, die neben der eigentlichen Nachricht u.a. Absende- und Empfängeradresse enthalten.
- Das IP-Modul erhält das Paket als nächstes und ermittelt die Internetadresse des Empfängers (hier 129.69.18.15). Dazu werden ein oder mehrere Name Server befragt, die die Auflösung der Namen bewerkstelligen (DNS). Es stellt sich auch heraus, daß der Rechner eigentlich gar nicht ftp sondern info2 heißt. Bei ftp handelt es sich nur um einen Alias-Namen.
- Danach wird überprüft, wohin ein Paket mit einer solchen Adresse als nächstes geschickt werden soll (z.B. in das lokale Netz oder nach draußen). Dazu existieren auf den Gateways und Routern sogenannte Routing-Tabellen.
- router11 überprüft seine Routing-Tabelle und stellt fest, daß das Paket nicht zum lokalen Netz gehört. Das Paket wird daher über weitere Router an das Gateway nach draußen geschickt: BelWue-GW.Uni-Stuttgart.DE.
- Im achten Schritt schließlich im lokalen Netz der Uni-Stuttgart und wird das Paket dort lokal an den Zielrechner ausgeliefert.

Die Nachricht wird also zwischen den verschiedenen Rechnernetzen weitergeleitet, bis sie das Netz des Empfängers erreicht. Der gerade erreichte Knoten kennt immer nur den nächsten Schritt des weiteren Weges.

1.7 Aufbau eines Routers

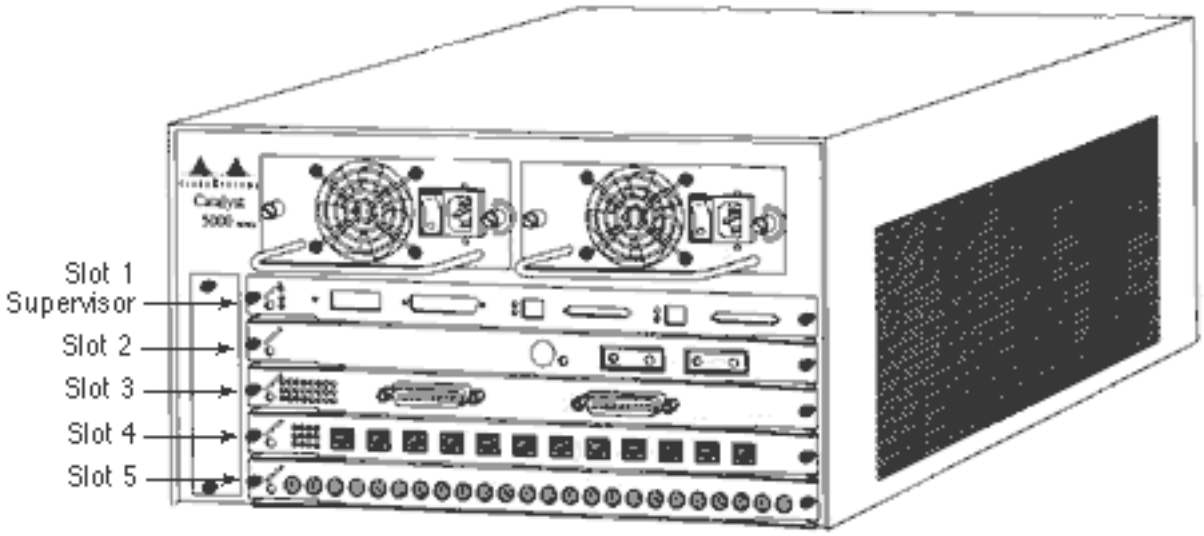

Abbildung 4: Hardware

Von einem Router wird erwartet, daß er extrem lange störungsfrei funktioniert. Daher werden oft zwei redundante Netzteile verwendet, damit trotz Ausfall der Router weiterhin seine Aufgaben ausführen kann. Die Netzteile wiederum sind als Einschübe konzipiert, damit sie schnell und einfach ausgetauscht werden können. Darüber hinaus besitzen die meisten Router weitere Einschübe (Slots) in die verschiedene Netzwerkkarten installiert werden koennen. Dies ist wichtig, damit der Router z.B. die Möglichkeit hat die ankommenden Daten vom LAN (z.B. Ethernet) auf ein anderes Interface (z.B. ISDN-Leitung) weiterzuleiten. Die Gehäuse haben üblicherweise eine Einbaugröße von 19 Zoll. Dies hat sich als Standardgröße etabliert und viele Router von verschiedenen Subnetzen werden in sog. „Racks“ zusammengefaßt.

1.8 Adressierung im Internet

Um einen bestimmten Dienst auf einem bestimmten Rechner in Anspruch nehmen zu können, muß jeder Computer im weltweiten Netz eindeutig identifizierbar sein. Darüberhinaus muß auch jedes im Netz bereitgestellte Dokument eindeutig gekennzeichnet sein, um es sicher auffinden zu können.

1.8.1 Adressierung von Rechnern

Um die eindeutige Identifizierbarkeit zu gewährleisten, wird jedem Computer im Internet eine weltweit eindeutige 32-Bit-Zahl als Adresse zugeteilt. Jede Adresse

besteht dabei zunächst aus zwei Teilen: der Netzidentifikation (net-id) und der Rechneridentifikation (host-id). Es gibt fünf verschiedene Klassen von Adressen, von denen derzeit jedoch nur drei eine Rolle spielen:

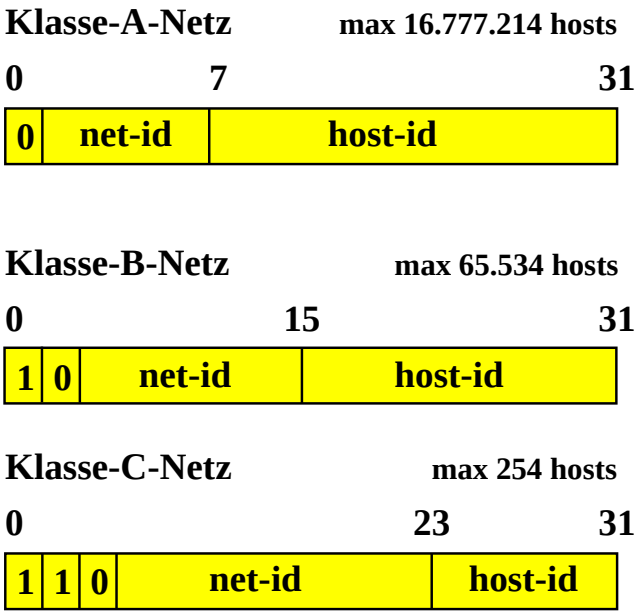

Abbildung 5: Netzklassen

Je mehr Bit für die host-id zur Verfügung stehen, umso mehr Rechner können angeschlossen werden. Andererseits stehen offensichtlich nur begrenzt Netzwerkadressen bzw. net-id's zur Verfügung. Bei dem rasanten Wachstum des Internet wundert es nicht, daß die bisherige Adressierungsart an ihre Grenzen stößt. Seit einigen Jahren wird bereits ein neuer Adressierungsstandard diskutiert. Favorisiert wird die Ausweitung des Schemas auf 48 Bit.

Die Internetadressen werden normalerweise so geschrieben, daß jedes Byte (8 Bit) als Dezimalzahl (zwischen 0 und 255) dargestellt wird und von dem Nachbarbyte durch einen Punkt getrennt wird. So wird aus der unlesbaren binären Zahl 01111111 00000000 00000000 00000000 die schon besser verständliche Adresse: 127.0.0.0, übrigens eine besondere Adresse. Die Klasse A-Adresse 127 ist für Loopback reserviert, d.h., alle Pakete, die an diese Adresse gehen, kommen unmittelbar zurück. Bestimmte Netzwerkfunktionen können damit getestet werden.

Auch die Dezimalschreibweise der Internet-Adressen ist nicht gerade benutzerfreundlich. Darum wurde das Domain Name System (DNS) geschaffen, das es erlaubt, den IP-Adressen logische Namen zuzuordnen. In den Anfängen des Internet erfolgte dies noch über eine zentral gehaltene Datei (/etc/hosts unter UN-

IX), die vom Network Information Center (NIC) regelmäßig an alle Rechner jeder Domain verschickt wurde. Das Aktualisieren und Verschicken einer solchen Datei war aber schon bald zu aufwendig.

Außerdem mußte auch eine stärker gegliederte Namenshierarchie eingeführt werden. So setzt sich heute der Name eines Rechners allgemein so zusammen:

host.subdomain.domain.topleveldomain

wobei die Subdomain je nach Größe und Organisation des Netzwerks auch entfallen kann. Z.B. bei inca.informatik.uni-tuebingen.de . Die Bezeichnungen bedeuten:

- host: der Name des Rechners im lokalen Netzwerk, hier: inca
- subdomain: Subnetzwerk-Name, hier: informatik
- domain: Name des Netzwerks, hier: uni-tuebingen
- top-level-domain: Übergeordnete Netzwerkhierarchie, entweder Länderkennung hier: de für Deutschland oder Organisationszuordnung z.B. edu für Bildungseinrichtung, com für Firma.

Eine Liste der Länderkennungen kann unter http://www.nw.com/zone/iso-country-codes eingesehen werden. Die wichtigsten Organisationsbezeichnungen als top-level-domains sind:

- com: Firma (commercial organization)
- edu: Bildungseinrichtung (educational institution)
- gov: Regierungsstelle (government)
- int: Internationale Organisation
- mil: Militärische Organisation
- net: Netzwerk-Organisation
- org: Nicht-profitorientierte Organisation

Das Domain Name System setzt sich aus drei Hauptkomponenten zusammen:

1. Der Domain Name Space: Das ist der baumartig strukturierte Namensraum mit den top-level-domains als Wurzel und den Hosts als Blättern. Der Domain Name Space ist in Zonen aufgeteilt, die aus einem Knoten im Baum und allen darunterliegenden Knoten bestehen.

2. Name Server: Das sind Rechner bzw. Programme, die die Informationen über einen Teil (bzw. eine Zone) des Domain Name Space verwalten. Da Name Server auf verschiedenen Ebenen des Domain Name Space existieren, überlappen sich die von den Name Servern verwalteten Zonen. Jeder Name Server kennt seinen nächsthöheren und nächsttieferen Nachbarn. Aus Sicherheitsgründen sind in jeder Zone mindestens zwei Name Server aktiv, die dieselbe Information liefern.

3. Resolver: Das sind Programme, die durch Anfrage an den Name Server den logischen Rechnernamen (z.B. inca.informatik.uni-tuebingen.de) in die Internetadresse (32-bit-Zahl) umwandeln oder umgekehrt. Kann der direkt zugeordnete Name Server die Anfrage nicht beantworten, können die Nachbarn des Name Servers angesprochen werden. Durch die Baum-Struktur ist der sukzessive Zugriff auf alle Name Server gewährleistet.

Wichtig ist natürlich, daß auch diese Namen weltweit eindeutig sind. Daher werden alle Domain-Namen von Network Information Centers registriert. Innerhalb der Domain können Rechner- und Subdomainnamen frei gewählt werden, solange sie im Netzwerk eindeutig sind.

1.9 Wer verwaltet und organisiert das Internet

Angesichts der Komplexität des Internet und seiner internationalen Verflechtungen ist es notwendig, Weiterentwicklung und Betrieb des Netzes national und international zu koordinieren. Im Laufe der Jahre sind zu diesem Zweck zahlreiche Organisationen entstanden, haben sich wieder aufgelöst oder mit anderen Organisationen verschmolzen. Eine vollständige Darstellung dieser Organisationen und ihrer Strukturen ist daher fast unmöglich. Ich beschränke mich auf einige wichtige internationale und deutsche Organisationen, die hauptsächlich den (nicht profitorientierten) Forschungs- und Bildungsbereich des Internet umfassen. Viele kommerzielle Anbieter haben sich daneben z.B. im Rahmen des Commercial Internet Exchange (CIX) organisiert.

1.9.1 Internationale Organisationen

- ISOC (Internet Society), gegründet 1991, ist eine internationale Nicht-Regierungs-Organisation für die Kooperation und Koordination der Internet-Aktivitäten. Mitglieder sind Einzelpersonen, Unternehmen, Non-Profit-Organisationen und Regierungsstellen. ISOC ist mit ihrenUnterorganisationen die zentrale weltweite Koordinierungsinstanz des Internet. Informationen unter info.isoc.org.

- IAB (Internet Architecture Board) ist technisches Beratungsgremium für die ISOC. Es ist zuständig für die Weiterentwicklung und Festlegung von Internet-Standards. Es hat dazu im wesentlichen zwei Unterabteilungen:

  - IETF (Internet Engineering Taskforce), die für die kurzfristigen technischen Entwicklungen zuständig ist. Eine Übersicht über Aufgaben und Geschichte der IETF gibt das Dokument http://www.ietf.org/-tao.html.
  - IRTF (Internet Research Task Force), die für die langfristigen technischen Weiterentwicklungen zuständig ist.

- TERENA (Trans European Research and Education Networking Association) entstand 1994 aus der Fusion von RARE und EARN (European Academic and Research Network). TERENA organisiert die nationalen und europäischen Forschungs- und Bildungseinrichtungen im Internet. Deutsches Mitglied ist der DFN. TERENA bietet Dienste für Provider über das RIPE NCC an. TERENA hat sich zum Ziel gesetzt, die Entwicklung der internationalen Informations- und Telekommunikationsinfrastruktur zum Nutzen von Forschung und Bildung voranzutreiben. Eine Kurzbeschreibung der Aufgaben und der Politik findet sich unter www.rare.nl/info/mission.html.

- RIPE (Réseaux IP Européens), gegründet 1989, betreibt seit 1992 das RIPE NCC (Network Coordination Centre) und bietet darin verschiedene Dienste für die angeschlossenen Internet-Provider an, so z.B. die regionale Internet-Adressen-Registratur (IP-registry) für Europa.

1.10 Internet-Standards

Obwohl das Internet keine zentrale Autorität als Lenkungsbehörde kennt, müssen Strukturen und Dienste vereinheitlicht werden, um die Funktionsfähigkeit des Netzes sicherzustellen. Spezifikationen, Vorschläge, Ideen, Richtlinien und Standards werden als sogenannte Request for Comments (RFC) veröffentlicht. Ein RFC-Editor koordiniert als Mitglied des IAB die Veröffentlichung und Verteilung im Netz.
Ein Dokument, das sich zum Internet-Standard entwickeln soll, muß drei Stadien erfolgreich absolvieren:

- Der RFC-Editor prüft den Inhalt darufhin, ob die beschriebenen Spezifikationen ausreichen, ob technische Probleme zu erwarten sind usw., und erklärt das Dokument gegebenenfalls zum proposed standard.

- Wenn mindestens zwei unabhängige Implementierungen des Vorschlages existieren, die erfolgreich zusammenarbeiten können, und wenigstens ein halbes Jahr vergangen ist, kann das Dokument zum draft standard hochgestuft werden.

- Nach mindestens weiteren vier Monaten kann der draft standard zu einem Internet-Standard werden und wird damit fester Bestandteil der Internetspezifikation.

Nicht alle RFC-Dokumente haben das Ziel der Standardisierung. In vielen Dokumenten werden auch Informationen für den Endbenutzer herausgegeben. Es gibt eine eigene Gruppe von RFC, die den Zusatz FYI (For Your Information) haben. Sie enthalten Informationen, die besonders für Anfänger von Interesse sind.

Alle RFC sind fortlaufend numeriert. Die RFC können an verschiedenen Stellen im Netz bezogen werden. Ein Server in Deutschland ist z.B. ftp://ftp.nic.-de/pub/doc/rfc. Eine Suchfunktion nach RFC-Nummern und Stichworten bietet http://www.nexor.co.uk/public/rfc/index/rfc.html.

Literatur

[1] A. Badach, E. Knauer
High Speed Internetworking
Addison-Wesley

[2] M. Zitterbart, C. Schmidt
Internet-working
Thomson's Aktuelle Tutorien (TAT), 1995

[3] A. S. Tanenbaum
Modern Operating Systems
Prentice Hall 1992

[4] P. Gortmaker
Linux Ethernet-HOWTO v2.65
ftp://sunsite.unc.edu/pub/Linux/docs/HOWTO/