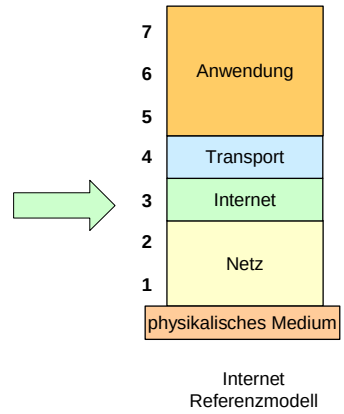
3.2a Internet-Referenzmodell: Internetschicht - Protokolle

Version: Jan. 2003

- Netzstruktur, Routeraufbau
- IP (Internet Protocol)
- IPv4/v6 Formate, Eigenschaften der Protokolle
- Fragmentierung
- Adressierung, Adressumwandlung (email, IEEE, ATM)
- ICMP (Internet Control Message Protocol)
- Routing (autonome Systeme, Distanzvektor-Routing, Linkzustands-Routing, Dijkstra Algorithmus)
- MPLS (Multi Protocol Label Switching): Netzstruktur, Vermittlungsformat, Vermittlungspfade
- QoS (Quality-of-Service, Servicequalität): Integrated Services und Differentiated Services

3.2a Internet-Referenzmodell: Internetschicht

- Netzstruktur, Routeraufbau
- IP (Internet Protocol)
- IPv4/v6 Formate
- Eigenschaften der Protokolle
- Fragmentierung
- IMCP (Internet Control Message Protocol)



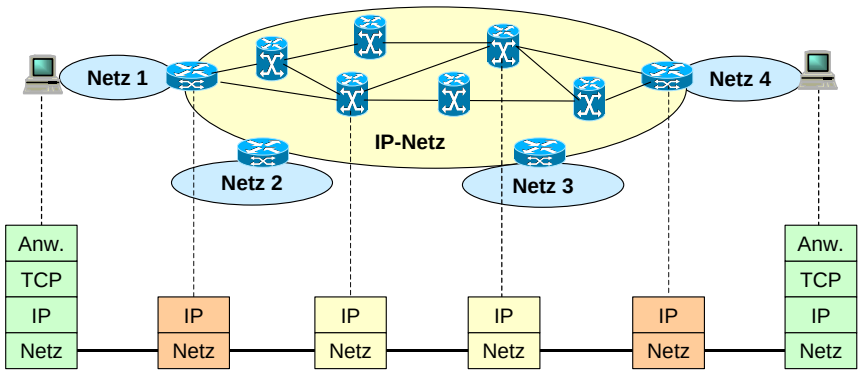
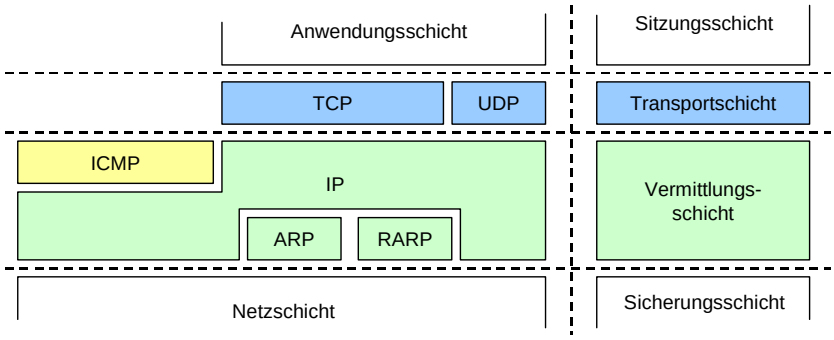


Bild: IP-Vernetzung

Aufgaben und Eigenschaften von IP

IP (Internet Protocol) ist ein Protokoll der Vermittlungsschicht, das Datagramme (datagrams) vom absendenden Endsystem (Quelle, source) zum empfangenden Endsystem (Ziel, destination) überträgt. IP leistet eine verbindungslose und unzuverlässige Übertragung. Verbindungslos bedeutet, dass die einzelnen Datagramme unabhängig von anderen Datagrammen behandelt werden, unzuverlässig drückt aus, dass Datagramme verloren gehen können, sowie in falscher Reihenfolge oder mehrfach beim Empfänger eintreffen können. IP leistet einen Best-Effort Dienst, d. h., Datagramme werden so gut und rasch wie möglich übertragen, es gibt dafür aber keinerlei Garantien. IP verwendet IP-Adressen, die ein Endsystem global eindeutig kennzeichnen. IP-Adressen besitzen eine innere Struktur, die Netz-ID und Host-ID unterscheidet. Für das Routing ist jedoch nur die Netz-ID von Bedeutung.

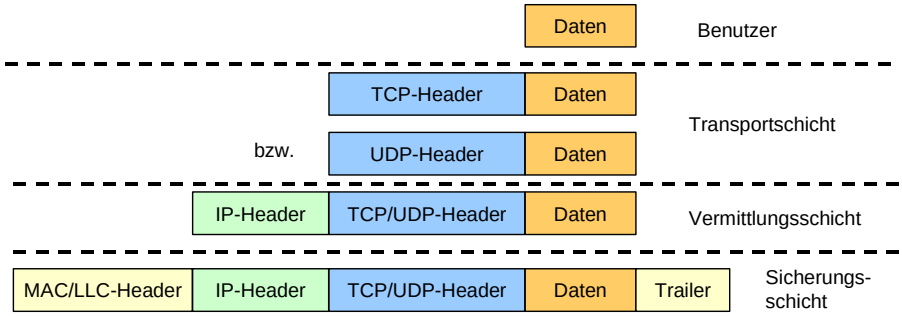


TCP: Transmission Control Protocol
UDP: User Data Protocol
IP: Internet Protocol
ICMP: Internet Control Message Protocol
ARP: Address Resolution Protocol
RARP: Reverse ARP

Die Bezeichnung TCP/IP wird häufig als Synonym für die gesamte Protokollfamilie verwendet.

Obwohl ICMP den IP-Dienst nutzt, wird es dennoch der Vermittlungsschicht zugeordnet.

Bild: Die TCP/IP-Protokollfamilie



Die Verschachtelung von PDUs (Protocol Data Units) geschieht auf der gleichen Weise wie bei OSI-Protokollen.

Bild: PDU Verschachtelung

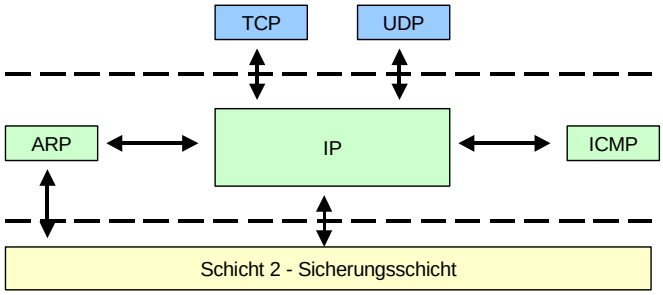


Bild: Zusammenspiel der Protokollinstanzen

Senden:
- Die TCP- bzw. UDP-Instanz übergibt Daten mit der IP-Adresse des Empfängers zur Übertragung an die IP-Instanz.
- IP-Instanz beauftragt ARP-Instanz mit Ermittlung der entsprechenden Schicht-2-Adresse.
- IP-Instanz übergibt Pakete (PDUs) mit der ermittelten Schicht-2 Adresse an die Instanz der Sicherungsschicht.

Empfangen: IP-Instanz reicht empfangene Daten an die TCP- bzw. UDP-Instanzen weiter.

Kontrolle: Im Falle von Problemen während der Übermittlung können diese den Partnerinstanzen über ICMP mitgeteilt werden (wobei ICMP zur Übertragung der Meldungen IP benutzt).

0		Reserved
1	ICMP	Internet Control Message Protocol
2	IGMP	Internet Group Management Protocol
3	GGP	Gateway-to-Gateway Protocol
4	IP	IP encapsulation
6	TCP	Transmission Control Protocol
8	EGP	Exterior Gateway Protocol
9	IGP	Interior Gateway Protocol
17	UDP	User Datagram Protocol
41	SIP	Session Initiation Protocol
46	RSVP	Resource Reservation Protocol
50	ESP	Encapsulation Security Payload for IPv6
51	AH	Authentication Header for IPv6
89	OSPF	Open Shortest Path First Routing Protocol

Bild: Protocols Numbers in IP-Paketen

Das **Internet Protocol (IP)** ist verbindungslos, d.h. jedes Datenpaket ist selbständig und enthält alle notwendigen Informationen, es von einem Endsystem zu einem anderen Endsystem zuzustellen. Die aktuelle Version ist Version 4, Version 6 ist ebenfalls spezifiziert, hat aber noch keine weite Verbreitung gefunden.

Die beiden Hauptfunktionen des Internet Protokolls sind
Adressierung und Routing,
Fragmentierung.

Adressierung höherer Protokoll-Schichten
Die IP-Schicht erhält ein Datenpaket von der unterliegenden Schicht und leitet es dann nach oben an die entsprechende höhere Protokoll-Schicht weiter. Sind nun mehrere Abnehmer, d.h. höhere Protokoll-Schichten vorhanden, dann muss eine Art Adressierung vorhanden sein. Die Information dazu steckt im Protokollnummer-Feld des IP-Protokollkopfes. Mit seiner Hilfe wird das entsprechende Transport-Protokoll adressiert.

IP Routing

Wichtig sind die beiden Adressen für Quelle und Senke-, es sind bei IPv4 32-Bit-Adressen. In dynamischen Routing Protokollen kann es - trotz aller Vorkehrungen - zu Schleifen kommen. Das kann dann dazu führen, dass ein Datenpaket ständig zwischen einigen Routern herumgereicht wird. Um zu verhindern, dass dieser Vorgang unendlich lange anhält, wird mit jedem Durchlauf durch einen Router der Inhalt eines speziellen Feldes im Paketkopf (Time-to-Live), das einen Zähler darstellt, vermindert. Eigentlich sollte eine reale Zeit eingetragen werden. In der Realität wird das Feld aber als hop count verwendet, also bei jedem Router-Durchlauf um eins vermindert. Wenn der Zählerstand null erreicht ist, wird das Paket verworfen.

Quality-of-Service (Qos)

Im klassischen Internet ist jeder Verkehr gleich viel Wert, es gibt keine Bevorzugung. Allerdings hatten schon die Erfinder des Internet Protokolls ein Feld vorgesehen, mit dem Qualitätsstufen unterschieden werden können. Dieses Type of Service genannte Feld hat eine Struktur, die es erlaubt, Prioritätsstufen zu unterscheiden und gewisse Kriterien an den Transport des Paketes zu legen.

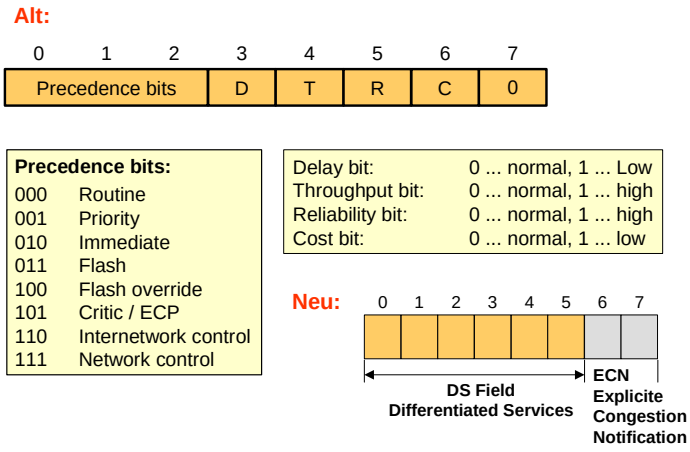


Bild: Type of Service (TOS)

Die Bezeichnungen der acht Prioritätsstufen, mit Precedence bezeichnet, wurden direkt der Originaldokumentation entnommen. Allgemein gilt, je höher der Wert, desto wichtiger ist das Paket. Es wurden Regeln entwickelt, welche Werte bei den Standarddiensten wie telnet und ftp einzusetzen sind. Die Kriterien für den Transport tragen die Bezeichnungen Zuverlässigkeit (Reliabilty), Durchsatz (Throughput), Verzögerung (Delay) und Kosten. Sie sind abstrakt und relativ zu sehen. Es wurden nie Regeln aufgestellt, wie diese Werte in einem Netz zu behandeln sind.

Im Zusammenhang mit neuen Richtungen für QoS wurde das Feld neu eingeteilt, um Mechanismen wie IntServ (Integrated Services), DiffServ (Differential Services) und MPLS (Multiprotocol Label Switching zu unterstützen. Siehe Vorlesungsteil 3.2c.

Fragmentierung

IP-Pakete dürfen eine maximale Größe von 65.535 Bytes haben. Die darunterliegenden Schicht-2-Protokolle haben aber in der Regel kleinere Paketgrößen, die von 576 Bytes bei X.25, über 1500 Bytes bei Ethernet bis zu 4500 Bytes bei FDDI reichen. Dieser Wert wird Maximum Transmission Unit (MTU) genannt. Die minimale MTU eines Paketes beträgt bei IPv4 576 Bytes, dabei bedeutet dies nicht, dass es keine kleineren Pakete geben dürfte, sondern dass jede Implementierung Pakete dieser Größe ohne Fragmentierung verarbeiten können muss.

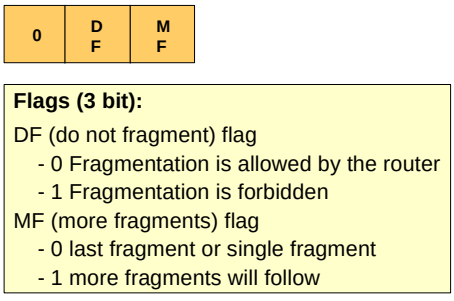


Bild: Flags im IP Datagram Header

Ist jetzt ein IP-Paket zu übertragen, das größer ist als die Schicht 2 transportieren kann, dann fragmentiert IP das Paket, teilt es also in eine Reihe kleinerer Pakete auf. Dieser Vorgang kann im Ursprung oder bei IPv4 auch in einem beliebigen Zwischenknoten (Router) stattfinden. Ein einmal fragmentiertes Paket wird erst am Ziel wieder zusammengesetzt. Wichtig ist, dass das IP-Fragment wieder wie ein normales IP-Paket aussieht. Daher ist auch das Längenfeld entsprechend anzupassen und die Prüfsumme neu zu berechnen. Wenn ein Fragment verloren geht, dann ist das ganze IP-Paket unbrauchbar, es findet keine Sicherung statt.

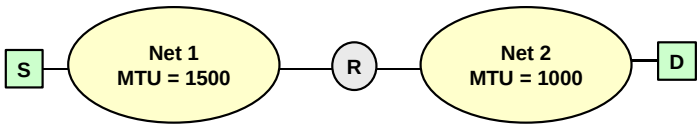


Bild: Maximum Transmission Unit (MTU)

- Jedes Subnetz hat eine maximale Rahmenlänge
  - Ethernet: 1518 Byte,
  - FDDI: 4500 Byte,
  - Token Ring: 2 to 4 Kbyte.
- Übertragungseinheit = IP datagram (data + header)
- Jedes Subnetz hat eine maximale IP Datagrammlänge: MTU (Maximum Transmission Unit).

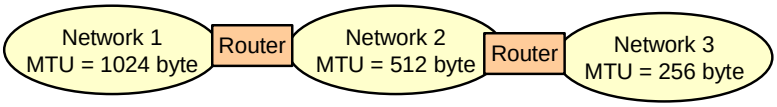


Bild: Fragmentation

- Datagramme länger als MTU (Maximum Transmission Unit) werden fragmentiert.
- Der Originalheader wird als Kopie in jedem Fragment mitgeführt.
- Bei jedem Fragment wird der Fragmentteil modifiziert (fragment flag, fragment offset, length).
- Auch einige Optionsfelder werden kopiert.

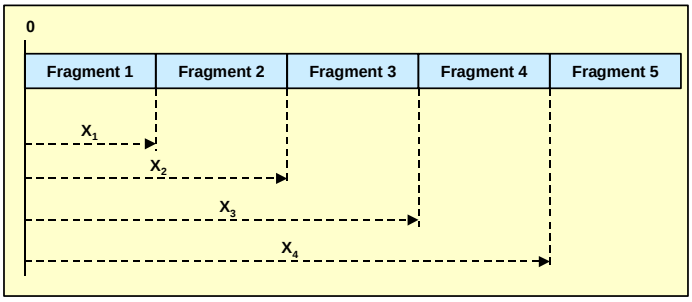


Bild: Payload-Fragmentierung und Fragment Offset

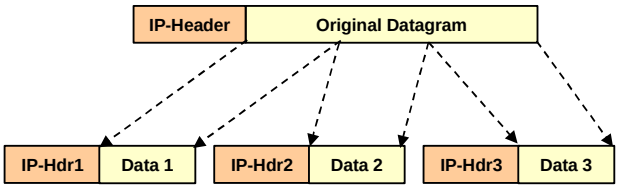


Bild: IP Fragment Offset

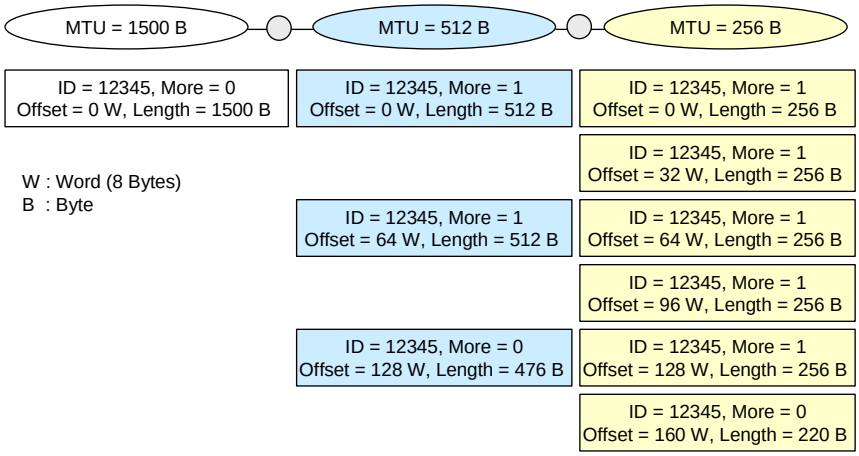


Bild: Beispiel einer Fragmentierung

- Die Reassemblierung der Fragmente findet nur am Ziel statt.
- Teilweise reassemblierte Datagramme werden nach einem Timeout verworfen.
- Fragmente können auf dem Weg zum Ziel weiter fragmentiert werden.
- Die Subfragmente haben den gleichen Format wie die Fragmente.
- Es ist nicht möglich festzustellen, wie oft fragmentiert wurde.
- Die minimale MTU auf dem Weg zum Ziel ist der Pfad MTU.

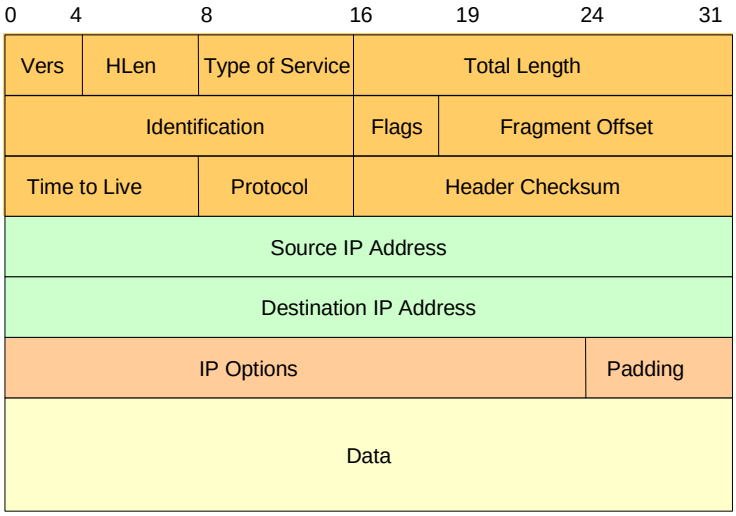


Bild: IPv4 Datagram Format

Aufbau des IP-Headers (lPv4

IP-Version 4 (kurz lPv4) ist die Version von IP mit der größten Verbreitung. Bedeutung der Felder des IP-Headers:

- **Version:** Wert 4.
- **IHL (Internet Header Length):** Länge des Headers in 32-Bit-Einheiten. Die Mindestlänge beträgt 20 Byte, sie wird bei der Verwendung von Optionen um jeweils 4 Byte vergrößert.
- **TOS (Tvpe of Service):** 3 Bit-langes Precedence Feld; einzelne Bits D (low delay), T (high throughput), R (high reliability); zwei nicht benutzte Bits. Wegen der Best-Effort-Eigenschaft von IP wurde dieses Feld praktisch nicht benutzt. Es ist jedoch im Zusammenhang mit Quality-of -Service neu definiert worden und damit extrem wichtig geworden.
- **Gesamtlänge:** Datagrammlänge (Header und Daten) in Byte.
- **Identification:** alle Fragmente eines Datagramms enthalten hier denselben Wert, der ihre Zusammengehörigkeit dokumentiert. Die **Fragmentierung** (fragmentation) eines Datagramms kann notwendig werden, wenn ein zwischen Quelle

und Ziel liegendes Netz dieses nicht übertragen kann, weil seine Länge größer ist als die für das Netz gültige MTU (Maximum Transmission Unit). Die MTU wird von der Hardware des jeweiligen Netzes bestimmt. Bei der Fragmentierung wird ein Datagramm in Fragmente zerlegt, jedes Fragment erhält einen IP-Header. Die Header zusammengehöriger Fragmente unterscheiden sich nur in den Feldern, die mit der Fragmentierung im Zusammenhang stehen. Der Empfänger macht die Fragmentierung im Prozess der **Reassemblierung** (reassembly) wieder rückgängig.

- **F (Fragmentierung)** besteht aus drei Bits: M (More fragments) kündigt weitere Fragmente an, D (Do not fragment) weist eine Zwischenstation an, nicht zu fragmentieren, und es gibt noch ein nicht definiertes Bit.
- **FO (Fragment Offset):** gibt die laufende Nummer des ersten Byte eines Fragments relativ zum ersten Byte des gesamten Datagramms an. Enthält den Wert null, wenn keine Fragmentierung verwendet wird.
- **TTL (Time-To-Live):** Zähler, der beim Senden des Datagramms auf einen Anfangswert gesetzt und von jedem Zwischensystem (bei jedem hop) dekrementiert wird. Beim Erreichen des Wertes null wird das Datagramm vernichtet. Dadurch wird eine Überlastung des Netzes durch nicht zustellbare Datagramme vermieden.
- **PR (Protocol):** Nummer des Protokolls, das oberhalb der Vermittlungsschicht verwendet wird (ICMP = 1, IGMP = 2, TCP = 6, UDP =17, SIP = 41, RSVP = 46, OSPF = 89).
- **Header-Prüfsumme:** schützt verfälschte Datagramme gegen Zustellung an die falsche Adresse. Zur Berechnung werden 16-Bit-Werte in Einerkomplement-Arithmetik addiert. Die Prüfsumme ergibt sich als Einerkomplement der berechneten Summe.
- **Quellen- und Zieladresse:** IP-Adressen der Länge 4 Byte.
- **Optionen:** Die Optionen liefern zusätzliche Möglichkeiten zur Steuerung und Überwachung der IP-Übermittlung. Die maximale Länge der Optionen beträgt 44 Byte.
- **Padding:** Nicht benutzte Bits werden mit **Füllbits** (padding bits) aufgefüllt.
- **Data:** Variables Daten Feld. Daten + header < 64 Kbyte (= 65 535 Byte).

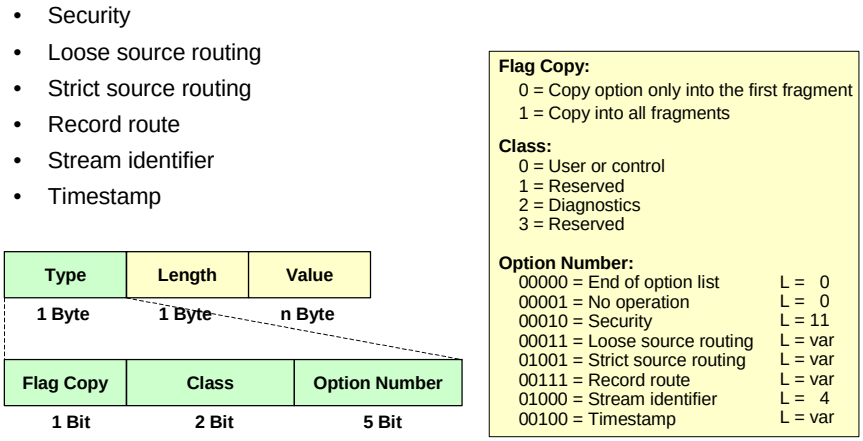


Bild: IP Options

Identifikation eines Paketes
Falls mehrere Versionen eines Protokolls im Einsatz sind, ist eine eindeutige Kennzeichnung nötig. Beim Internet Protokoll ist diese in den ersten 4 Bit des Paketes untergebracht und erlaubt es so, unterschiedliche im Einsatz befindliche Protokoll-Versionen zu unterscheiden. Bei Paketen variabler Länge muss eine Information über die Größe des Paketes vorhanden sein. Naheliegend ist die Angabe einer Gesamtlänge. Ist zusätzlich der Paketkopf variabel, muss auch dessen Länge angegeben werden, damit die Nutzinformation korrekt an die höhere Protokollschicht übergeben werden kann.

Schutz des Paketes
Prüfsummen lassen einen Rückschluss über die Qualität der Übertragung zu. Je nach Aufwand kann es sich um reines Erkennen von Bitfehlern bis zur Korrektur mehrfacher Fehler reichen. Hier ist eine Abwägung zwischen Aufwand und Nutzen notwendig. Im vorliegenden Fall wird eine einfache Prüfsumme nur für den Protokollkopf vorgesehen. Ist auch eine Schutz des Informationsinhaltes des Paketes notwendig, so ist dies auf einer höheren Protokoll-Schicht anzusiedeln.

Zusatzprotokolle im Zusammenhang mit IP
Direkt im IP residieren drei Zusatzprotokolle, die dem Betrieb des Internet Protokolls selbst dienen. Sie werden in normalen IP-Paketen transportiert, dann aber nicht an eine höhere Protokollschicht weitergeleitet. Es sind dies:

Internet Control Message Protocol (ICMP),
Address Resolution Protocol (ARP) und
Reverse Address Resolution Protocol (RARP).

Internet Control Message Protocol (ICMP)
Dieses Protokoll residiert innerhalb IP und wird in IP-Paketen transportiert, die den Protokoll-Wert 1 haben. In 8 Bytes werden die ICMP-spezifischen Informationen gesendet. Mit den Feldern Type und Code werden die Nachrichten unterschieden. CRC ist die Prüfsumme über die ICMP-Nachricht. Im Datenfeld werden die für die spezifische Nachricht notwendigen Daten ausgetauscht. Wenn eine Fehlermeldung gesendet wird, dann wird immer der IP-Protokollkopf und die ersten 8 Byte des Datenfeldes von demjenigen Paket zurückgesendet, das den Fehler verursacht hat. Damit kann der Empfänger eine genaue Analyse des Fehlerfalls durchführen. Für IPv6 wurde ein eigenes ICMP spezifiziert, das sehr viel weniger Nachrichten beinhaltet.

Address Resolution Protocol (ARP)
Die IP-Datenpakete beinhalten eine IP-Adresse-, die MAC-Schicht (z. B. Ethernet) arbeitet aber mit einer eigenen Adresse, der MAC-Adresse. Für die Abbildung aufeinander wurde eine eigene Prozedur geschaffen, das Address Resolution Protocol (ARP).

Reverse Address Resolution Protocol (RARP)
Es gibt auch den umgekehrten Vorgang: Die Suche nach einer IP-Adresse bei gegebener MAC-Adresse. Das Protokoll dazu ist das Reverse Address Resolution Protocol (RARP). Da die MAC-Adresse fest auf der Netzkarte einprogrammiert ist, die IP-Adresse normalerweise konfigunert wird, bietet RARP einem Host, der seine IP-Adresse nicht speichern kann (z. B. Diskless-Workstation), die Möglichkeit, diese Adresse aus einem Server abzufragen. Allerdings wurde für diesen Anwendungsfall ein mächtigeres Protokoll entwickelt: das Dynamic Host Configuration Protocol (DHCP).

Das Protokoll ICMP
In jedem Netz treten von Zeit zu Zeit Fehler auf, die an Verursacher oder davon Betroffene gemeldet werden müssen. Diese Aufgabe wird in Netzen mit der Protokollfamilie TCP/IP vom Protokoll ICMP (Internet Control Message Protocol) übernommen. Hierfür stellt das ICMP eine Vielzahl von sog. ICMP-Nachrichten zur Verfügung. Das ICMP wurde bereits im Jahr 1981 im RFC 792 spezifiziert. Der Funktionsumfang von ICMP wurde später im RFC 1256 erweitert. An dieser Stelle ist hervorzuheben, dass es sich hier um das ICMP für das Protokoll IP der Version 4 gehandelt hat. Das hier kurz dargestellte Protokoll ist daher als ICMP für IP4 zu bezeichnen. Das Protokoll ICMP für IPv6 existiert ebenfalls.

Zu den wichtigsten Aufgaben des Protokolls ICMP gehören:

- Unterstützung der Diagnose.
- Hilfsprogramm ping: Üblicherweise wird in Netzen zum Feststellen der Erreichbarkeit des Kommunikationspartners das Programm ping verwendet. Dieses Hilfsprogramm sendet ICMP-Echo-Anforderung an eine IP-Adresse und wartet auf ICMP-Echo-Antworten. Das Programm ping meldet die Anzahl der empfangenen Antworten und die Zeitspanne zwischen Senden der Anfrage und Eingang der Antwort.
- Hilfsprogramm traced (bzw. traceroute) als weiteres Analysewerkzeug wird zum Verfolgen von Routen eingesetzt. Es sendet Echo-Anforderung an eine IP-Adresse und analysiert die eingehenden ICMP-Fehlermeldungen.
- Unterstützung der Aufzeichnung von Zeitmarken (Timestamps) sowie Ausgabe von Fehlermeldungen bei abgelaufenen Timestamps von IP-Paketen.
- Verwaltung von Routing-Tabellen.
- Berichtigung der Flußkontrolle, um eine Überlastung des Routers bzw. des Zielrechners zu vermeiden (ICMP Source Quench).
- Mitwirken bei der Auffindung der zulässigen Größe von IP-Paketen, d.h. von MTU (Maximum Transfer Unit).

- RFC 792
- Used by IP to send error and control messages
- ICMP uses IP to send its messages
- ICMP does not report errors on ICMP messages
- ICMP message are not required on datagram checksum errors
- ICMP reports error only on the first fragment

ICMP (Internet Control Message Protocol, RFC 792) wird immer zusätzlich zu IP benötigt. Es tauscht Nachrichten für die Steuerung der Datenübertragung sowie Fehlermeldungen zwischen Routem und Hosts aus. ICMP-Nachrichten werden zur Übertragung in einem IP-Datagramm gekapselt. Das ICMP wird normalerweise als Teil der Schicht 3 betrachtet, aber ausnahmsweise werden die Daten dieses Protokolls in IP-Paketen transportiert. Somit werden die ICMP-Nachrichten wie Daten eines Transportprotokolls in IP-Paketen transportiert, obwohl ICMP kein Transportprotokoll, sondern ein Hilfsprotokoll auf Schicht 3 ist. Dem ICMP wurde die Protokollnummer 1 im IP-Header zugeordnet.

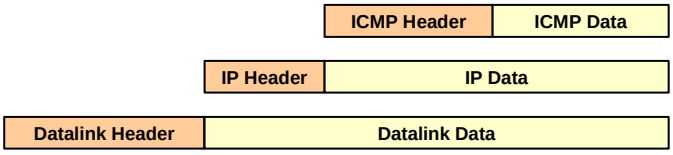


Bild: Internet Control Message Protocol (ICMP)

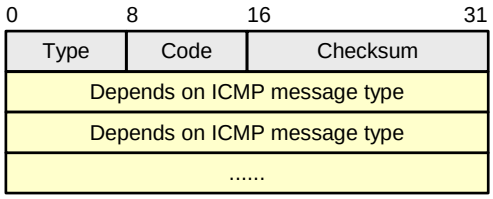
ICMP-Nachrichten
Da ICMP unterschiedliche Informationen zu transportieren hat, enthalten die ICMPNachrichten einen Header, der in allen Nachrichten immer gleich ist. Die Bedeutung von ICMP-Nachrichten (-Angaben), die direkt nach dem Header folgen, ist von einzelnen Fehlern bzw. Diagnosesituationen abhängig. Für die Struktur des Teils ICMP Data bei den einzelnen ICMP-Nachrichtentypen ist auf die Dokumente RFC 792 und RFC 1256 zu verweisen.

Die einzelnen Angaben im ICMP-Header lauten:

- Type: Unterscheidung von einzelnen ICMP-Nachrichten.
- Code: Eine weitere Unterteilung der Nachricht innerhalb eines Typs. Beispielsweise in der Nachricht "Destination unreachable" wird dem Absender eines IP-Pakets mitgeteilt, warum es nicht übermittelt werden konnte; z.B.
  0 = Netz nicht erreichbar,
  1 = Rechner nicht erreichbar,
  2 = Protokoll nicht erreichbar,
  3 = Port nicht erreichbar,
  4 = Fragmentierung erforderlich und DF-Bit gesetzt
- Checksum (Prüfsumme): eine Prüfsumme, die nur die ICMP-Daten auf Fehler überprüft.

Falls eine Fehlermeldung zu einem Rechner in einer ICMP-Nachricht ankommt, so stellt sich die Frage, auf welches IP-Paket und weiches Protokoll sich die Fehlermeldung bezieht. Abhängig vom Typ (und manchmal auch Code) werden in den ICMP-Nachrichten noch weitere Informationen als ICMP-Daten (Fehler-, Diagnose-Angaben etc.) direkt nach dem Header übermittelt. Die Bedeutung von ICMP-Daten ist von einzelnen Fehler- bzw. Diagnose-Situationen abhängig. Die ICMP-Fehlermeldungen beinhalten neben der Fehlermeldung auch immer den IP-Header und die ersten 64 Bits des diese fehlerhafte Situation verursachenden IP-Pakets.

Empfängt ein Rechner beispielsweise eine ICMP-Nachricht mit Typ = 3 und Code = 1 (d.h. Destination Unreachable Message), so kann er nach der Type- und Code-Angabe genau bestimmen, was die Ursache des Fehlers ist. In diesem Fall wird dem Absender eines IP-Pakets mitgeteilt, dass der Zielrechner nicht erreichbar ist. Der Header dieses IP-Pakets und dessen weitere 64 Bits sind in der Destination-Unreachable-Nachricht als ICMP-Daten enthalten.



Type field: type of the ICMP message
Code field: corresponding error specification

Bild: ICMP Message Format

Der ICMP-Header ist je nach Nachrichtentyp unterschiedlich aufgebaut. Er enthält unteren andern die folgenden Felder:

- **Type:** Gibt den Typ einer ICMP-Meldung an.
- **Code:** Kann den Typ näher beschreiben.
- **Prüfsumme:** Prüfsumme über die gesamte ICMP-Nachricht. Verwendet denselben Algorithmus wie IP.
- **Identifier und Sequenznummer:** kennzeichnet zusammengehörige Anfragen und Antworten (Typ 0 oder 8).
- **Optionale Daten:** Hier stehen Daten, die in einem echo request vom Sender zum Empfänger übertragen wurden. Im echo reply werden die Daten unverändert zurückgeschickt.

Nachrichten des Typs destination unreachable senden den IP-Header und die ersten 64 Bits des nicht zustellbaren Datagramms an dessen Absender zurück.

0	Echo reply	15	Information request
1	Unassigned	16	Information reply
2	Unassigned	17	Address mask request
3	Destination unreachable	18	Address mask reply
4	Source quench	19	Reserved for security
5	Redirect	20 - 29	Reserved for robustness experiment
6	Alternate host address	30	Traceroute
7	Unassigned	31	Datagram conversion error
8	Echo request	32	Mobile host redirect
9	Router advertisement	33	IPv6 (Where are you)
10	Router selection	34	IPv6 (I am here)
11	Time exceeded	35	Mobile registration request
12	Parameter problem	36	Mobile registration reply
13	Time stamp request	37 - 255	Reserved
14	Time stamp reply		

Bild: Type Field Numbers

- Echo reply (0) and echo request (8):
  test (ping) the accessibility of an IP node (host, router)

- Destination unreachable (3), time exceeded (11) and parameter problem (12):
  notification of errors related to
  - accessibility
  - time-to-live
  - reassembly of fragments
  - errors in the IP header

- Source quench (4): **flow control**

- Redirect (5): **path optimization**

- Time stamp request (13), time stamp reply (14), information request (15), information reply (16), address mask request (17) and address mask reply (18):
  diagnostics and maintenance

Bild: Verwendung von ICMP Messages

- Echo request and reply messages (used for ping)
- Time stamp request and reply messages
- Time exceeded message
- Address mask request and reply messages to discover the subnet mask currently used in a network
- Information request and reply message (e.g. used for discovery of the network ID of an IP address)
- Redirect message
- Router advertisement message
- Source Quench: Slow down
- Time Exceeded: Time to live field in one of your packets became 0 or reassembly timer expired at the destination
- Fragmentation Required: Datagram longer than MTU and “No Fragment bit” was set
- Address Mask Request/Reply: What is the subnet mask on this net?

Bild: ICMP Messages

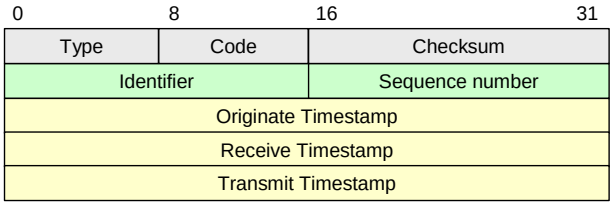
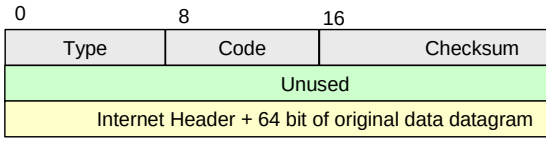
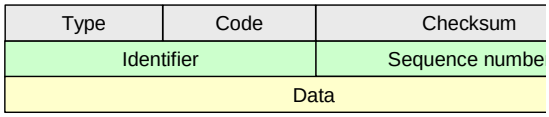


Bild: ICMP Messages

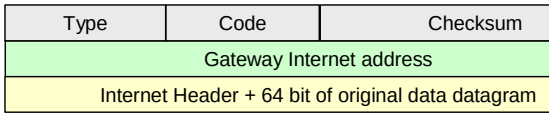
Destination Unreachable Message



Echo or Echo Reply Message (“Ping”)



Redirect Message



Code	Meaning
0	Network unreachable
1	Host unreachable
2	Protocol unreachable
3	Port unreachable
4	Fragmentation need and do not fragment bit set
5	Source route failed
6	Destination network unknown
7	Destination host unknown
8	Source host isolated
9	Communication with destination network administratively prohibited
10	Communication with destination host administratively prohibited
11	Network unreachable for type of service
12	Host unreachable for type of service

Bild: Destination Unreachable

ICMP-Nachrichten

Da ICMP unterschiedliche Informationen zu transportieren hat, enthalten die ICMPNachrichten einen Header, der in allen Nachrichten immer gleich ist. Die Bedeutung von ICMP-Nachrichten (-Angaben), die direkt nach dem Header folgen, ist von einzelnen Fehlern bzw. Diagnosesituationen abhängig. Für die Struktur des Teils ICMP Daten bei den einzelnen ICMP-Nachrichtentypen ist auf die Dokumente RFC 792 und RFC 1256 zu verweisen.

Die einzelnen Angaben im ICMP-Header lauten:

- **Type:** Unterscheidung von einzelnen ICMP-Nachrichten.
- **Code:** Eine weitere Unterteilung der Nachricht innerhalb eines Typs. Beispielsweise in der Nachricht "Destination unreachable" wird dem Absender eines IP-Pakets mitgeteilt, warum es nicht übermittelt werden konnte; z.B.
  0 = Netz nicht erreichbar,
  1 = Rechner nicht erreichbar,
  2 = Protokoll nicht erreichbar,
  3 = Port nicht erreichbar,
  4 = Fragmentierung erforderlich und DF-Bit gesetzt
- **Checksum (Prüfsumme):** eine Prüfsumme, die nur die ICMP-Daten auf Fehler überprüft.

Falls eine Fehlermeldung zu einem Rechner in einer ICMP-Nachricht ankommt, so stellt sich die Frage, auf welches IP-Paket und weiches Protokoll sich die Fehlermeldung bezieht. Abhängig vom Typ (und manchmal auch Code) werden in den ICMP-Nachrichten noch weitere Informationen als ICMP-Daten (Fehler-, Diagnose-Angaben etc.) direkt nach dem Header übermittelt. Die Bedeutung von ICMP-Daten ist von einzelnen Fehler- bzw. Diagnose-Situationen abhängig. Die ICMP-Fehlermeldungen beinhalten neben der Fehlermeldung auch immer den IP-Header und die ersten 64 Bits des diese fehlerhafte Situation verursachenden IP-Pakets.

Empfängt ein Rechner beispielsweise eine ICMP-Nachricht mit Typ = 3 und Code = 1 (d.h. Destination Unreachable Message), so kann er nach der Type- und Code-Angabe genau bestimmen, was die Ursache des Fehlers ist. In diesem Fall wird dem Absender eines IP-Pakets mitgeteilt, dass der Zielrechner nicht erreichbar ist. Der Header dieses IP-Pakets und dessen weitere 64 Bits sind in der Destination-Unreachable-Nachricht als ICMP-Daten enthalten.

Die Struktur des Teils ICMP Data bei den einzelnen ICMP-Nachrichtentypen ist in den Dokumente RFC 792 und RFC 1256 dokumentiert..

Type	Function	Type	Function
0	Echo Reply	15	Information Request
1	-	16	Information Reply
2	-	17	Adress Mask Request
3	Destination Unreachable	18	Adress Mask Reply
4	Source Quench	19	Reserviert (Security)
5	Redirect	20-29	Reserviert (for Robustness Experiments)
6	Alternate Host Address	30	Trace Route
7	-	31	Datagram Conversion Error
8	Echo	32	Mobile Host Redirect

9	Router Advertisement	33	IPv6 - Where are You
10	Router Selection	34	IPv6 - Am Here
11	Time Exceeded	35	Mobile Registration Request
12	Parameter Problem	36	Mobile Registration Reply
13	Timestamp	37-255	-
14	Timestamp Reply		

ICMP-Fehlermeldungen

Der häufigste Einsatz von ICMP liegt in der Meldung verschiedener Arten von fehlerhaften Situationen. Ein Rechner oder ein Router gibt eine ICMP-Fehlermeldung zurück, wenn er feststellt, dass ein Fehler oder eine außergewöhnliche Situation während der Weiterleitung bzw. der Übergabe an ein Transportprotokoll JCP oder UDP) eines IP-Pakets aufgetreten ist. Diese außergewöhnlichen Situationen, die eine ICMP-Fehlermeldung verursachen, sind:

- Destination Unreachable Message (Ziel nicht erreichbar) Ein IP-Paket kann nicht an den Zielrechner übergeben werden. In diesem Fall wird die Nachricht Destination Unreachable an den Quellrechner gesendet, um darauf hinzuweisen, dass der Empfänger nicht erreichbar ist. Die Ursachen hierfür sind unterschiedlich. Eventuell existiert der Zielrechner nicht mehr, oder es ist kein passendes Protokoll im Zielrechner geladen.
- Time Exceeded Message (Zeit überschritten) Befindet sich ein IP-Paket so lange im Netz, dass die "Time To Live" im IP-Header abgelaufen ist, so wird die Nachricht Time Exceeded vom Router, in dem das betreffende IP-Paket "vernichtet' wurde, an den Quellrechner zurückgeschickt.
- Parameter Problem Message (Ungültige Parameter) Ein oder mehrere Parameter im Header des IP-Pakets enthalten ungültige Angaben bzw. unbekannte Parameter. In diesem Fall wird die Nachricht Parameter Problem verschickt.
- Source Quench Message (Übertragungsrate reduzieren) Ist ein Rechner nicht in der Lage, die zu schnell ankommenden IP-Pakete rechtzeitig zu verarbeiten, wird die Nachricht Source Quench an die Quelle gesendet, damit diese die Sendung von IP-Paketen für einen gewissen Zeitraum unterbricht.
- Redirect Message (Umleitung im Netz) Bemerkt ein Router, dass es für ein IP-Paket eine bessere Route gibt als über diesen Router, so kann er dem Quellrechner eine Empfehlung mit der Nachricht Redirect geben, weitere IP-Pakete zum gleichen Zielrechner über einen anderen Router zu verschicken. Die IP-Adresse dieses Routers wird im Feld ICMP-Data übermittelt.

ICMP-Anfragen

Zusätzlich zu den ICMP-Meldungen, die in den fehlerhaften Situationen generiert werden, gibt es eine Reihe weiterer ICMP-Nachrichten, die für die Anfrage von Informationen und zur Antwort auf eine ICMP-Anfrage verwendet werden können. Hierzu gehören:

- Echo Request / Reply Message (Echo-Funktion): Die häufigsten Anfragemeldungen sind die ICMP-Nachrichten für die Implementierung des Programms ping zum Versenden von Diagnose-Nachrichten. Die Nachrichten Echo Request/Reply werden für die Implementierung einer sog. "Bist Du noch da"-Funktion verwendet. Hierbei wird von dem ping-Programm ein Echo-Request zu einem bestimmten Ziel (Rechner bzw. Router) gesendet. Das Ziel muss auf den Echo Request mit einem Echo Reply antworten. Die Nachricht Echo Request ist die einzige ICMP-Nachricht, auf die jeder IP-fähige Rechner antworten muss.
- Timestamp RequestIReply Message (Zeitmarkenanfrage) Ein Rechner oder ein Router gibt eine Zeitmarkenanfrage ab, um von einem anderen Rechner oder Router eine Zeitmarke zu erhalten, die das aktuelle Datum und die Uhrzeit angibt. Ein Rechner oder Router, der eine Zeitmarkenanfrage in der Nachricht Timestamp Request empfängt, antwortet mit der Nachricht Timestamp Reply. Die Nachrichten Timestamp Request und Reply verwendet werden, um die Laufzeit eines IP-Pakets über das Netz zu messen.
- Information Request/Reply Message (Informationsanfrage): Diese Nachrichtentypen sollen es einem Rechner ermöglichen, seine IP-Adresse (z.B. von einem Adress-Server) abzufragen. Da die dynamische Vergabe von IP-Adressen heutzutage mit dem Protokoll DHCP (Dynamic Host Configuration Protocol) gemacht wird, hat diese ICMP-Funktion an Bedeutung verloren.
- Address Mask Request/Response (Abfrage der Subnetz-Maske): Diese Nachrichtentypen ermöglichen es einem Rechner, die zu verwendende Subnetz-Maske abzufragen. In einem Subnetz, in dem diese Funktion unterstützt wird, sind ein oder mehrere Rechner als Subnetz-Masken-Server gekennzeichnet. Ein Rechner, der seine Subnetz-Maske zu ermitteln versucht, sendet eine Abfrage in der Nachricht Address Mask Request, auf die ein Subnetz-Masken-Server mit einer Nachricht Address Mask Response antwortet, in der die zu verwendende Subnetz-Maske enthalten ist.

Jedem Rechner in einem Subnetz muss die IP-Adresse eines Routers als Grenzübergang zu anderen Subnetzen bekannt sein. Diese Adresse wird üblicherweise bei der IP-Konfiguration eines Rechners als Default Gateway angegeben. Das ICMP stellt zwei Nachrichten zur Verfügung, die es ermöglichen, einen Router zu entdecken.

Diese Nachrichten, die die Entdeckung von Routern in einem Subnetz dokumentieren, sind:
• Router Solicitation (Suche nach einem Router)
• Router Advertisement (Router-Bekanntmachung)

Ein Rechner kann während seiner Konfigurationsphase eine Nachricht Router Solicitation an alle Systeme (Rechner, Router) in demselben Subnetz verschicken. Diese Nachricht bedeutet "Ich suche einen Router" und enthält im IP-Header eine IP-Multicast-Adresse 244.0.0.1 bzw. eine Limited Broadcast-Adresse 255.255.255.255. Der Router antwortet mit der Nachricht Router Advertisement, in der er seine IP-Adresse von diesem physikalischen Port bekannt macht, auf dem die Nachricht Router Solicitation empfangen wurde.
Einem physikalischen Port im Router können mehrere IP-Adressen zugeordnet werden, so dass in der Nachricht Router Advertisement alle IP-Adressen des entsprechenden Router-Ports enthalten sein können.

Pfad-MTU Ermittlung

Eine wichtige Funktion von ICMP besteht in der Unterstützung der Feststellung der Maximum Transfer Unit (MTU) für ein entferntes, über Router zu erreichendes IP-Netz. Dieses Verfahren wird als Path MTU (PMTU) Discovery bezeichnet, und ist laut RFC 1191 in Routern zu unterstützen.

Die PMTU wird im Zusammenspiel zwischen IP-Sender und den in der PMTU Übertragungsstrecke liegenden Routern entsprechend folgendem Ablauf festgestellt:

1. Die IP-Instanz des Senders generiert zunächst IP-Pakete mit gesetztem Don't Fragment-Bit (DF = 1) und der maximalen MTU des lokalen Netzes. Diese MTU entspricht in der Regel auch der des in diesem Netz liegenden IP-Interfaces des Default Gateway und somit des ersten Hops.
2. Überschreitet ein erzeugtes IP-Paket die MTU eines Transfernetzes, so dass der zugehörige Router es eigentlich fragmentieren müsste, wird es von diesem verworfen und der Sender erhält die ICMP-Nachricht Destination Unreachable mit dem Statuscode "fragmentation needed and DF set". Ferner fügt der Router die maximal mögliche IP-Paketgröße (in Bytes) in die ICMP-Nachricht ein.
3. Der Sender ist somit aufgefordert, seine ursprüngliche MTU auf die nun bekannte Obergrenze zu reduzieren und die Datagramme erneut zu übertragen.
4. Dieses Verfahren kann periodisch wiederholt werden, um z. B. wechselnden Routen zu entsprechen.

Protokolle ARP (Address Resolution Protocoll) und RARP (Reverse ARP)
Diese Protokolle sind Hilfsprotokolle bei der Adressierung von IP-Paketen. Das Protokoll ARP hat die Aufgabe, für eine Ziel-Adresse die korrespondierende MAC-Adresse zu ermitteln. Das Protokoll RARP ermöglicht, für eine MAC-Adresse die entsprechende IP-Adresse zu bestimmen. RARP wird vorwiegend von Rechnern ohne Festplatte (z. B. Netz-Computer NCs) genutzt, die als die Stationen am LAN dienen und ihre IP-Adresse nicht selbst speichern können. In Routern wird oft eine zusätzliche Lösung für das Protokoll ARP eingesetzt, die als Proxy ARP bezeichnet wird.

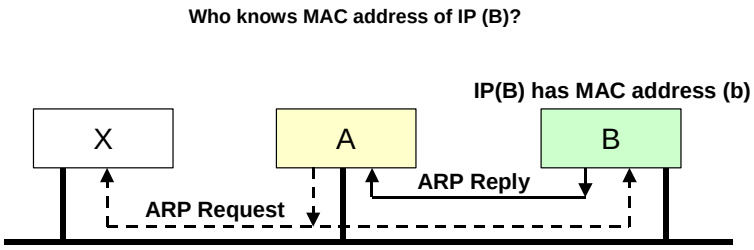


Bild: ARP: Address Resolution Protocol

ARP (Address Resolution Protocol, RFC 826): übersetzt IP-Adressen in Hardwareadressen. ARP beschränkt sich dabei auf ein physisches Teilnetz. Die Übersetzung wird dynamisch vorgenommen, indem ein Host A mittels Broadcast eine Anfrage-PDU mit der zu übersetzenden IP-Adresse aussendet. Wenn Host B darin seine IP-Adresse erkennt, antwortet er mit seiner Hardwareadresse.

Die bei ARP-Anfragen erhaltenen Adressübersetzungen werden in einem Cache gespeichert. Beim Senden von Paketen wird die Adressübersetzung zuerst im Cache gesucht. Nur falls sie dort nicht gespeichert ist, wird eine ARP-Anfrage gesendet. Einträge im Cache bleiben nur für eine bestimmte Zeit gültig.

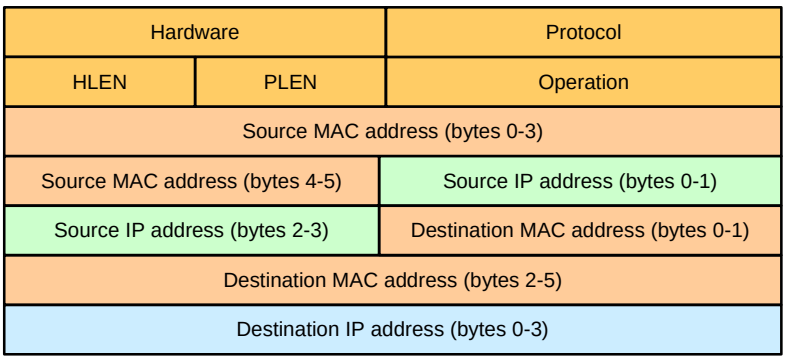


Bild: ARP Request/Reply

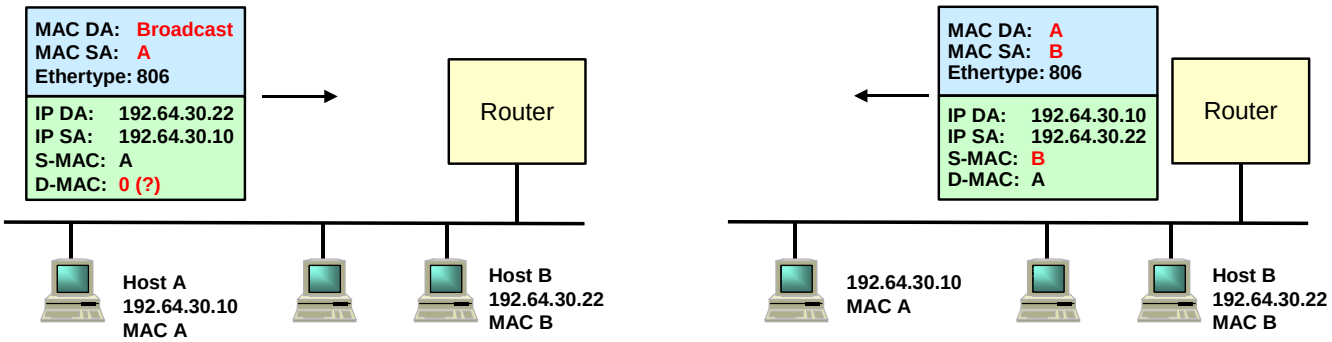
ARP-Pakete haben folgende Felder:

Hardware: LAN-Typ (z.B. Ethernet, IEEE 802.x LANs) in welchem das Paket generiert wurde.
Protokoll: Netzprotokoll von welchem die Operation angefordert wurde. Das Protokoll IP hat den Wert x'0800'. Das Protokoll ARP unterstützt auch andere Netzprotokolle.
Hardware Address Length: Länge der Hardwareadresse in Bytes (normalerweise MAC-Adresse mit 6 Bytes).
Protocol Address Length: Länge der Protokolladresse in Bytes. Bei IP-Adressen 4 Bytes).

Operation: 1- ARP Request, 2- ARP Reply (RFC 826), 3- RARP Request, 4- RARP Reply (RFC 903).
Sender MAC Address: Hier ist die MAC-Adresse des Absenders enthalten.
Sender Protocol Address: Dieses Feld enthält die IP-Adresse des Absenders.
Target MAC Address: Hier wird die gesuchte MAC-Adresse (in ARP-Reply) angegeben.
Target Protocol Address: Dieses Feld enthält die IP-Adresse, für die die MAC-Adresse ermittelt wird.

Protokoll ARP

Es sind zwei Adressierungsstufen zu unterscheiden. Einerseits müssen die Hardwarekomponenten (Endsysteme, Router) in jedem Netz eindeutig identifiziert werden. Hierfür verwendet man physikalische Netzadressen. In LANs mit einem gemeinsamen Medium werden die Netzadressen als MAC-Adressen bezeichnet. Da diese Adressen unstrukturiert sind und somit keine Lokationshinweise enthalten, werden sie auch als Nummern von LAN-Adapterkarten gesehen. Andererseits müssen die Daten in Form von IP-Paketen zwischen zwei Kommunikationspuffern in Endsystemen ausgetauscht werden. Diese Kommunikationspuffer sind im logischen LAN-Modell an der Grenze zwischen den Schichten 3 und 4 zuzuordnen. Liegt ein IP-Paket in einem Endsystern am LAN zum Senden vor, so wird dieses Paket in einen MAC-Frame eingebettet. Im Header des MAC-Frames ist eine entsprechende MAC-Adresse des Zielsystems enthalten. Somit muss eine Tabelle mit den Zuordnungen IP-Adresse zu MAC-Adresse in LAN-Endsystemen vorhanden sein.
Früher wurde dieses Problem in jedem Rechner durch statische Tabellen gelöst, in die man manuell alle Zuordnungen zwischen MAC- und IP-Adressen eintragen musste. In dieser Form war der Verwaltungsaufwand sehr hoch und das ganze System unflexibel. In der heutigen TCP/IP-Welt werden diese Zuordnungen mit dem Protokoll ARP realisiert. Das Protokoll ARP ist ein Hilfsprotokoll zur Ermittlung einer physikalischen Interface-Adresse (MAC-Adresse) für ein höheres Protokoll (z.B. IP), d.h., es ist für die Zuordnung von MAC-Adressen zu Protokoll-Adressen verantwortlich.
Das Protokoll ARP legt eine dynamisch organisierte Adressermittlungs- Tabelle mit IP-Adressen und den zugehörigen MAC-Adressen an. Oft wird diese Tabelle auch ARP-Cache genannt.
Wenn das Protokoll IP die Anforderung erhält, ein Paket an eine IP-Adresse im gleichen Subnetz zu senden, sucht es zuerst im ARP-Cache nach der korrespondierenden MAC-Adresse. Falls kein Eintrag vorhanden ist, wird versucht, mit Hilfe von ARP die gesuchte MAC-Adresse zu ermitteln.



Hierfür wird ein ARP-Request als ein MAC-Broadeast verschickt. In dieser Nachricht werden die restlichen Endsysteme in demselben Subnetz gebeten, die gesuchte Adresszuordnung IP-Adresse zu MAC-Adresse zukommen zu lassen. Ein Endsystem schickt immer eine Antwort als ein ARP-Reply (MAC-Unicast) mit der gesuchten Zuordnung zurück. Anschließend wird dieses Paar von ARP in seinem Cache abgelegt.

Die Ermittlung einer MAC-Adresse im Endsystem A nach dem Protokoll ARP erfolgt wie folgt. Die Broadcast-Nachricht ARP-Request enthält die IP-Adresse der angeforderten MAC-Adresse und wird in allen Endsystemen im LAN gelesen. Sobald ein Endsystem die eigene IP-Adresse im ARP-Request erkennt (hier Endsystem B), antwortet es mit einem ARP-Reply. Die beim Endsystem A eingehende Antwort wird im ARP-Cache vermerkt und steht damit für spätere Übertragungen zu Verfügung. Falls innerhalb einiger Sekunden keine Antwort eingeht, wird die Anforderung wiederholt.

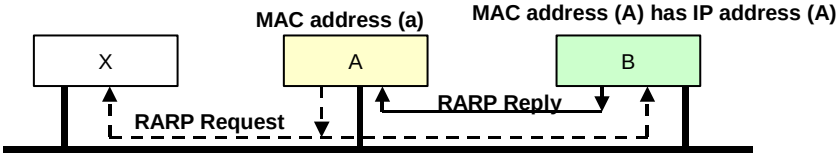
Damit nicht bei jeder Übertragung erneut Anforderungen ARP-Request gesendet werden müssen, kopiert auch das Endsystem B, d.h. das auf ARP-Request antwortet, die Zuordnung von IP-Adresse und MAC-Adresse des ARP-Request-Absenders (Endsystems A) in seinen eigenen ARP-Cache. Bei einer eventuellen Übertragung in Gegenrichtung (von A zu B) ist es daher nicht mehr nötig, eine ARP-Anforderung in umgekehrter Richtung zu senden, da die MAC-Adresse der IP-Adresse, der gerade geantwortet wurde, bereits bekannt ist.

Den Aufbau von Nachrichten ARP-Request und -Reply (ARP-PDU) zeigt das Bild. Es ist hier hervorzuheben, dass diese Nachrichten direkt in MAC-Frames transportiert werden. Sie werden somit auf dem MAC-Level übermittelt. Die Folge dessen ist, dass der ARP-Request von Routern nicht weitergeleitet werden kann, da Router auf dem IP-Level operieren und somit auf MAC-Broadcast-Nachrichten nicht reagieren. Diese Tatsache hat in der Praxis einen Nachteil. Als Folge dessen ist eine Proxy ARP Lösung notwendig.

In manchen TCP/IP-Implementierungen werden für Einträge im ARP-Cache ein Zeitlimit (time-out) gesetzt. Falls der Eintrag innerhalb dieses Zeitraums, oft 15 Minuten, nicht verwendet wird, wird er gelöscht. Einige Systeme arbeiten wiederum mit einem zeitgesteuerten Aktualisierungsprinzip. Alle 15 Minuten wird dann eine Anforderung ARP-Request gesendet, um sicherzustellen, dass die Cache-Einträge dem aktuellen Systemzustand entsprechen. Da MAC-Adressen normalerweise nur verändert werden, wenn eine Adapterkarte bzw. der ganze Rechner ausgetauscht wird, scheint dieses Prinzip von keiner großen Bedeutung zu sein.

In den Token-Ring-LANs, falls mehrere LANs miteinander vernetzt werden, muss das sogenannte Source Routing in Endsystemen unterstützt werden. Um das Source Routing unterstützen zu können, enthält der ARP-Cache in Endsystemen am Token-Ring eine zusätzliche Spalte mit der Angabe des nächsten Router-Abschnittes Next-RD (RD: Route Designator).

Probleme kann es mit ARP geben, wenn in einem Netz zwei Stationen die gleiche IP-Adresse besitzen. In einem solchen Fall kann keine exakte Zuordnung zwischen IP-Adresse und MAC-Adresse getroffen werden, d.h. die Daten werden nicht korrekt weitergeleitet, oder es wird aufgrund einer nicht identifizierten Verbindung eine Fehlermeldung produziert. In einem gut organisierten Netz ist mit diesem Problem nur ganz selten zu rechnen.



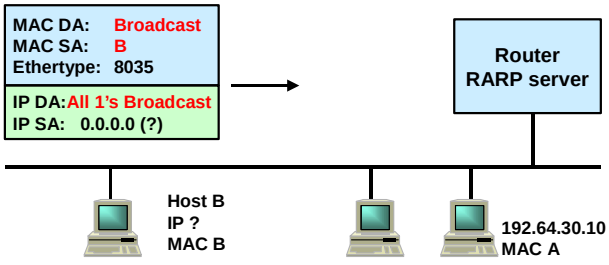
Problem: How can the IP address be resolved? (Initialization)

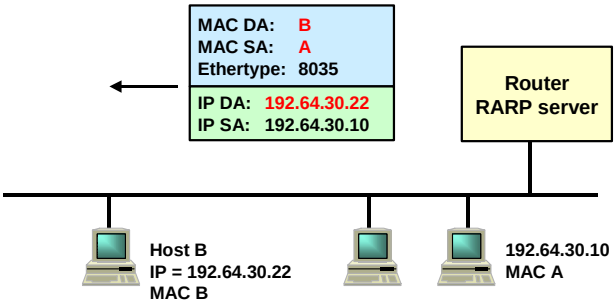
Solution: manual entry of IP address
IP from configuration file (not possible at diskless station)
IP address global fixed at interface card

Protocol: A sends broadcast information with physical address (a)
RARP server B takes the physical address, maps it to
the IP address of A and sends the IP address (A) back to A

Bild: Reverse Address Resolution Protocol (RARP)

RARP (Reverse Address Resolution Protocol)
Die RARP-Anfrage eines Hosts wird mittels Broadcast an alle anderen Teilnehmer des Teilnetzes gesendet. Diejenigen Stationen, die als RARP-Server aufgesetzt sind (z. B. Host Y), kennen die IPAdresse von Host A und teilen diese in einer RARP-Antwort mit. Falls mehrere RARP-Server aktiv sind, empfängt A mehrere Antworten, von denen jedoch nur die erste ausgewertet wird.



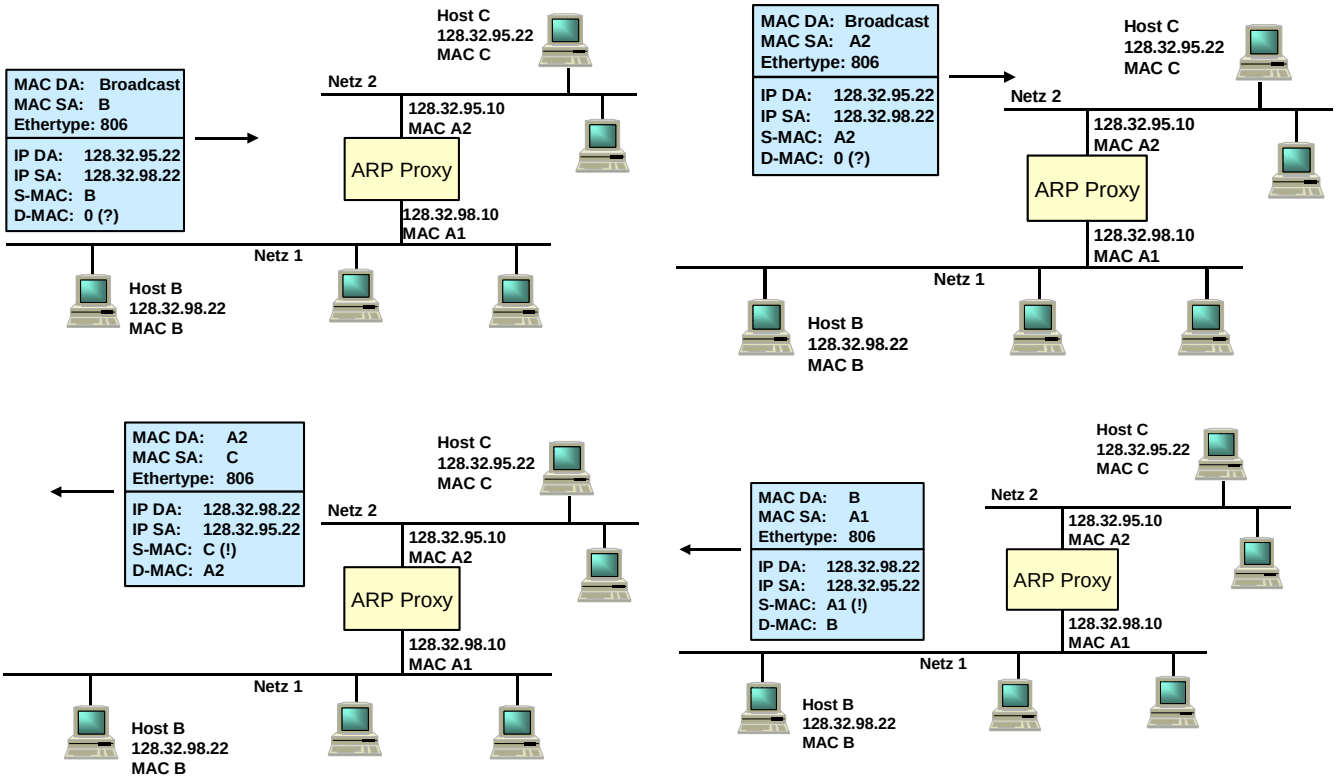


Protokoll RARP

Das Protokoll RARP (Reverse Address Resolution Protocol) ist für Stationen gedacht, die ihre IP-Adresse nicht selbst speichern können (z. B. Rernote-Boot-Stationen ohne Festplatte). RARP ist das Gegenstück zu ARP, d.h. RARP bietet Funktionen, die es ermöglichen, aus einer bekannten MAC-Adresse die zugehörige IP-Adresse zu finden. Bei RARP ist es notwendig, einen speziellen Server festzulegen, in dem eine RARP-Tabelle enthalten ist. Der Server sucht in dieser Tabelle nach der IP-Adresse, die mit der angeforderten MAC-Adresse übereinstimmt und gibt die gesuchte IP-Adresse als RARP-Antwort (Reply) bekannt.

Das RARP-Prinzip setzt voraus, dass mindestens ein Rechner als RARP-Server fungiert und dass dieser Server über eine Tabelle verfügt, in der allen MAC-Adressen eine eindeutige IP-Adresse zugeordnet ist.

Der Aufbau von RARP-Nachrichten ist wie bei ARP. Beim Protokoll RARP werden im Feld Operation die Werte 3 für RARP-Request und 4 für RARP-Reply verwendet. Wenn ein RARP-Request gesendet wird, kennt das aussendende Endsystem nur die eigene MAC-Adresse und kann daher auch nur diese Adresse im MAC-Frame angeben. In der Antwort RARP-Reply vom Server wird die gesuchte IP-Adresse eingetragen. In dieser Antwort kann auch die IP- und MAC-Adresse des RARP-Servers angegeben werden. Dies ist allerdings nicht erforderlich.



Proxy-ARP

Unter Proxy-ARP ist eine Lösung zu verstehen, die u.a. ermöglicht,

- einer Netz-ID bzw. einer Subnetz-ID mehrere physikalische Netze zuzuordnen;
- beim Subnetting auch diese Endsysteme (Hosts) weiter einzusetzen, die das Subnetting nicht unterstützen. Das Konzept von Proxy-ARP findet sich in RFC 1027.

Zunächst wird der Einsatz von Proxy-ARP an zwei Beispielen illustriert, in denen sich eine (Sub)Netz-ID auf mehrere physikalische Netze bezieht.

Betrachtet wird das Prinzip, nach dem ein Shared Medium LAN (hier beispielsweise Ethernet LAN) mit dem ISDN integriert werden kann, dass diese beiden physikalisch unterschiedlichen Netze logisch als ein Subnetz gesehen werden können. Hier sind externe Rechner über das leitungsver- Subnet2 mittelnde ISDN an ein Ethernet LAN angebunden. Aus organisatorischen Gründen müssen diese externen Rechner transparent, also mit IP-Adressen des lokalen Subnetzes (d.h. Ethernet LANs) eingebunden werden. Für die Übermittlung der IP-Pakete zwischen den externen Rechnern und dem Router wird das Protokoll PPP (Point-to-Point Protocol) verwendet.

Im LAN werden die Endsysteme mit den MAC-Adressen als LAN-Hardware-Adressen identifiziert. Außerdem ist das LAN ein roadcast-orientiertes Netz, während das ISDN ein leitungsvermittelndes Netz darstellt, in dem das Broadcast nicht unterstützt werden kann. Das Protokoll ARP setzt ein broadcastorientiertes Netz voraus. Somit lässt sich dieses Protokoll im ISDN nicht realisieren. Um die beiden Netze LAN und ISDN so zu integrieren, dass sie ein Subnetz bilden, ist die Proxy ARP-Funktion im Router nötig. Diese Funktion soll es ermöglichen, für die LAN-Endsysteme die ISDN-Endsysteme unter einer MAC-Adresse x, d.h. des Router-Ports seitens des LANs, zu verbergen. Die Proxy-ARP-Funktion besteht in diesem Fall darin, dass eine besondere ARP-Tabelle im Router an dessen LAN-Port mit der MAC-Adresse x enthalten ist. In dieser Tabelle werden die IPAdressen von ISDN-Endsystemen eingetragen, und deren IP-Adressen wird die MAC-Adresse x des Routers von der LAN-Seite zugeordnet. Mit einer solchen ARP-Tabelle wird den LAN-Endsystemen "mitgeteilt", dass die ISDN-Endsysteme unter der MAC-Adresse x des Routers zu erreichen sind.

Liegt bei einem LAN-Endsystem ein IP-Paket, das an ein ISDN-Endsystem z. B. mit der IP-Adresse y gesendet werden soll, so prüft dieses LAN-Endsystem zunächst, ob das Ziel sich im gleichen Subnetz befindet. Da dies gerade der Fall ist, wird das IP-Paket in einem MAC-Frame direkt an das Ziel gesendet. Ist die Ziel-MACAdresse dem Quell-Endsystem unbekannt, so sendet es nach dem Protokoll ARP eine Broadcast-Nachricht ARP-Request an alle Systeme in dessen Subnetz. Diese Broadcast-Nachricht wird auch vom Router empfangen, der mit einem ARP-Reply antwortet, indem der IP-Adresse y des ISDN-Endsystems die MAC-Adresse x zugeordnet wird. Nach dem Empfang von ARP-Reply vermerkt das Quell-Endsystem in seinem ARP-Cache, dass der IP-Adresse y die MAC-Adresse x entspricht. Somit wird der MAC-Frame im nächsten Schritt direkt an den Router abgeschickt. Der Router leitet gemäß der Routing-Tabelle das empfangene IP-Paket an den ISDN-Port weiter.

Wie man an diesem Beispiel sieht, ist es mit Hilfe der Proxy-ARP-Funktion möglich, mehrere physikalische Netze mit Hilfe eines Router so zu koppeln, dass sie ein heterogenes Netz bzw. Subnetz bilden und damit nur eine (Sub)Netz-ID besitzen. Es ist hierbei darauf hinzuweisen, dass eine Proxy-ARP-Lösung eine Not-Lösung ist, wenn man kein Subnetting realisieren kann. Wäre Subnetting möglich, so sollte man dem Ethemet LAN eine Subnetz-ID und dem ISDN eine weitere Subnetz-ID zuweisen. Bei einer derartigen Lösung ist auch die Proxy-ARP-Funktion im Router nicht nötig.

Unterschiedliche LANs können mit der Proxy-ARP-Funktion ein Subnetz bilden. In diesem Fall stellen zum Beispiel Ethernet und Token-Ring zwei getrennte Broadcast-Netze dar. An dieser Stelle ist hervorzuheben, dass ARP-Nachrichten die Nachrichten der MAC-Schicht sind, so dass sie über den Router nicht weitergeleitet werden können. Dies bedeutet, dass eine Broadcast-Nachricht aus dem Ethernet-Teil das Token-Ring nicht erreichen kann. Umgekehrt können die Broadcast-Nachrichten aus dem Token-Ring die Ethernet-Seite nicht erreichen.

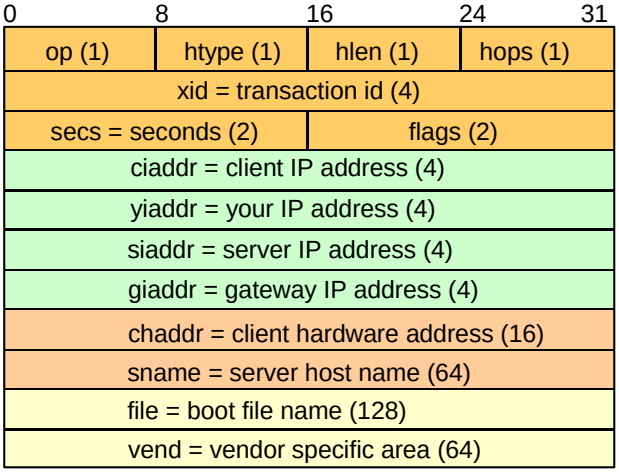
Mit der Proxy-ARP-Funktion im Router kann ein solcher Effekt erreicht werden, dass die Endsysteme am Ethernet den Eindruck gewinnen, die Token-Ring-Endsysteme wären am Ethernet angeschlossen. Umgekehrt wird den Token-Ring-Endsystemen vorgemacht, dass sich ihre Kommunikationspartner am Token-Ring statt am Ethernet befänden. Eine solche Täuschung ist mit Hilfe entsprechender ARP-Tabellen möglich. Eine Tabelle seitens des Ethernet, signalisiert den Ethernet-Endsystemen, dass die Token-Ring-Endsysteme unter der MAC-Adresse g zu erreichen sind. Dabei handelt es sich um die MAC-Adresse des Ethernet-Ports im Router. Die zweite ARP-Tabelle seitens des Token-Ring-LANs signalisiert den Endsystemen am Token-Ring, dass die Ethemet-Endsysteme unter der MAC-Adresse h erreichbar sind.

In einigen Fällen kann es sinnvoll sein, Endsystemen auf unterschiedlichen Medien (z. B. Ethernet und Token-Ring, bzw. Ethernet und FDDI) das gleiche IP-(Sub-)Netz zu definieren. In diesem Fall stellt die Proxy-ARP-Funktionalität ein geeignetes Instrument zur Kopplung dieser Frontend Netze an das Hochgeschwindigkeitsnetz Backend-Netz zur Verfügung, wo z. B. ein Host an seinem FDDI-Interface zwei getrennte IP-Adressen in den IPNetzen A und B zugewiesen bekommt. Über diese IP-Adressen ist er dann sowohl für Stationen am Ethernet über das IP-Netz A wie auch für Rechner am Token-Ring am IP-Netz B transparent erreichbar. Die Router mit Proxy-ARP-Funktionaliät gewährleisten hierbei nicht nur die Umsetzung der MAC-Adressen, sondern auch die notwendige Fragmentierung der IP-Pakete entsprechend der maximalen MTU für das jeweilige LAN.

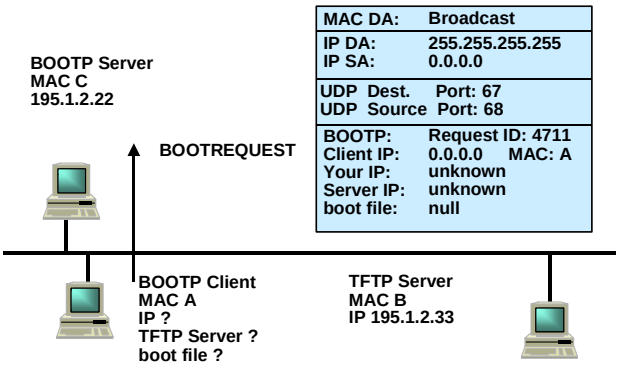
Proxy-ARP ist insbesondere dann hilfreich, wenn Endsysteme (Hosts) den Internet-Standard für die Subnetz-Adressierung nicht unterstützen, d.h. sie unterstützen kein benutzerspezifisches Subnetting. In diesem Fall handelt es sich um die alte Generation von Endsystemen (Hosts).

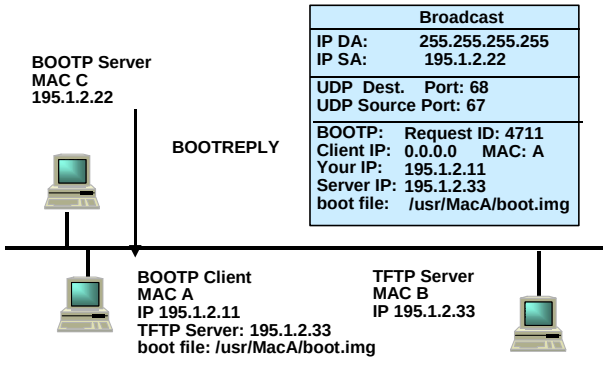
BOOTP

BOOTP (Bootstrap Protocol, RFC 951, RFC 1542, RFC 1532) ist eine Alternative zu RARP, die es Endsystemen ohne Festplatte erlaubt, ihre eigene IP-Adresse (und weitere Startinformationen, z. B. die Adresse eines Routers und eines Name Server, sowie eine Subnetz-Maske) durch Anfrage bei einem Server herauszufinden, eine Datei in den Arbeitsspeicher zu laden und auszuführen. BOOTP benutzt TFTP für die Dateiübertragung und UDP. BOOTP wird von Anwendungsprogrammen genutzt und ist damit ein Protokoll der Anwendungsschicht.

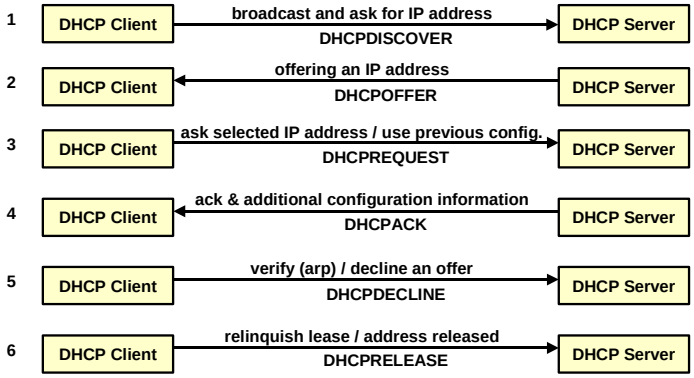


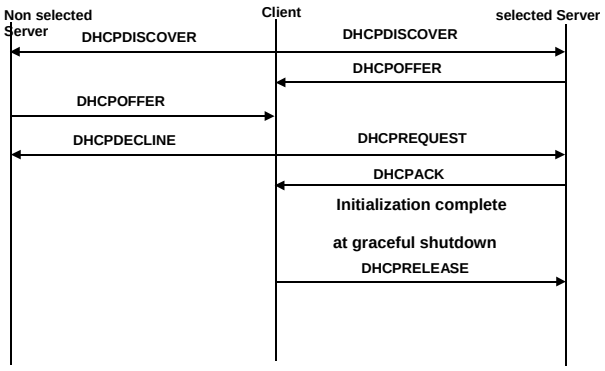
Operation code: 1: BOOTREQUEST, 2: BOOTREPLY
hardware address type, same assigned numbers as ARP,,1‘ = 10Mbit/s Ethernet
hardware address length
number of hops; client sets to zero, incremented by gateways in case of cross-gateway booting
transaction id; used to match this boot request with the responses it generates
seconds; filled in by client, seconds elapsed since client started trying to boot
–•flags: MSB is broadcast bit; shall be set if client is unable to receive unicast messages until he knows its IP address
client IP address; filled in by client in bootrequest if known
your (client) IP address; filled by server if client doesn't know its own address (ciaddr was 0)
server ip address; IP address of server holding the boot file (TFTP server), returned in bootreply by BOOTP server
gateway IP address; used in optional cross-gateway booting
client hardware address; filled-in by client
optional server host name, null terminated string
boot file name; null terminated string; 'generic' name or null in bootrequest, fully qualified directory-path name in bootreply
optional vendor-specific area, see RFC 2132 for more details, e.g. information about subnet mask, list of routers in preference order, time server, name server, DNS-servers, host name, boot file size ...





DHCP (Dynamic Host Configuration Protocol, RFC 2131, RFC 1531, RFC 1541) ist eine erweiterte Version des BOOTP-Protokolls, die zusätzlich die Fähigkeit einer automatischen und dynamischen Belegung mit wiederverwendbaren (also nicht fest einer Station zugeordneten) IP-Adressen und von Konfigurationsoptionen bietet. Eine dynamische Belegung mit IP-Adressen ist beispielsweise für drahtlose LAN erforderlich. Die IP-Adresse und die zugehörige Subnet-Maske werden für eine bestimmte Zeit (Lease-Dauer) einem Knoten zur Verfügung gestellt. DHCP ist interoperabel mit Endsystemen, die BOOTP benutzen, und soll BOOTP langfristig ablösen. DHCP ist - wie BOOTP - ein Protokoll der Anwendungsschicht





Dynamische Vergabe und Ermittlung von IP-Adressen

Durch die Vergabe von IP-Adressen können Rechner in IP-Netzen und speziell im Internet angesprochen werden. Im intuitiven Umgang sind IP-Adressen jedoch nicht "sprechend" genug. Es ist sinnvoll, statt einer IP-Adresse einen Rechner über seinen Namen zu adressieren. Dies kann im Prinzip durch eine statische Tabelle - die Host-Datei - erfolgen; sobald aber eine Vielzahl Rechner in entfernten IP-Netzen, d.h. speziell im Internet, erreicht werden sollen, wird die Pflege der Host-Dateien schnell unhandlich.
Um das Problem der dynamischen Namensauflösung im Internet zu lösen, wurde das Domain Name System (DNS) geschaffen. Das Domain Name System stellt eine verteilte Datenbank dar, die im Grunde genommen mit ihrem Informationsgehalt das Internet abbildet.
Entsprechend der Bedeutung des DNS für das Internet hat sich die Vergabe dynamischer IP-Adressen im Intranet entwickelt. Über das Dynamic Host Configuration Protocol (DHCP) kann eine dynamische und konsistente Vergabe von IP-Adressen und anderen wichtigen IP-Informationen für Rechner im Intranet erreicht werden. Dies wird im folgenden Abschnitt erläutert.

Protokoll DHCP

Um die Rechner (wie z.B. PCs) ohne Festplatte als Endsysteme in TCP/IP-Netzen zu starten und automatisch zu konfigurieren, wurde das Protokoll BOOTP (BOOT Protocol) entwickelt (RFC 1532). Ein Rechner ohne Festplatte ist normalerweise nicht in der Lage, seine IP-Adresse, die benötigten Programme seines Betriebssystems oder den TCP/IP-Programmcode in ausgeschaltetem Zustand zu speichern. Das Protokoll BOOTP soll solche Rechner in die Lage versetzen, alle für den Betrieb am TCP/IP-Netz benötigten Informationen von einem BOOTP-Server abzurufen. Dabei handelt es sich um einen einen Rechner im Netz, der auf eingehende BOOTPAnforderungen ständig wartet und die Antworten auf die Anforderungen erzeugt. Da heutzutage Rechner ohne Festplatte am Netz nur selten sind, hat das Protokoll BOOTP an Bedeutung verloren.

Das Protokoll DHCP (Dynamic Host Configuration Protocol) kann als eine BOOTP neue und erweiterte Generation des Protokolls BOOTP gesehen werden. Mit Hilfe des Protokolls DHCP ist es möglich, die IP-Adressen und anderen zusätzliche Konfigurationsparameter jenen Rechnern automatisch zuzuweisen, die für die Nutzung von DHCP konfiguriert sind. In diesen Rechnern muss das Protokoll DHCP implementiert werden, weswegen sie zwansgsläufig über eine Festplatte verfügen. Mit Hilfe des Protokolls DHCP ist es somit möglich, sämtliche TCP/IP-Konfigurationsparameter zentral zu verwalten und zu warten. Insofern besteht auch die Möglichkeit die Endsysteme in TCP/IP-Netzen nach dem Prinzip Plug and Play zu installieren.
Jeder einzelne Rechner in einem Netz muss sowohl über einen eindeutigen Namen als auch eine eindeutige IP-Adresse verfügen, um mit anderen Rechnern kommunizieren zu können. Die IP-Adressen können dem Rechner entweder manuell oder automatisch zugewiesen werden. Bei der manuellen Zuweisung handelt es sich um sogenannte statische IP-Adressen, die ein Administrator manuell konfigurieren und bei Bedarf neu zuordnen muss. Bei der dynamischen Zuweisung wird einem Rechner automatisch eine IP-Adresse zugewiesen, wenn er eingeschaltet wird. In diesem Fall spricht man von dynamischen IP-Adressen.
Durch den Einsatz des Protokolls DHCP lassen sich vor allem jene Probleme beseitigen, die mit dem manuellen Konfigurieren von IP-Adressen verbunden sind. Die erste Version des Protokolls DHCP wurde Ende 1993 im RFC 1541 veröffentlicht und als Standard im März 1997 durch den RFC 2131 mit einer neuen DHCP-Version abgelöst. Die neue DHCP-Version erweitert die alte um einige Besonderheiten.

Das Protokoll DHCP funktioniert nach dem Client/Server-Prinzip. Ein DHCP-Server ist ein Rechner, in dem sämtliche Konfigurationsparameter für die Rechner (oft nur innerhalb eines Subnetzes) abgespeichert worden sind. Die Rechner, die auf den DHCP-Server zugreifen, um bestimmte Konfigurationsangaben abzufragen, werden als DHCP-Clients bezeichnet. Wenn ein DHCP-Client gestartet wird, fordert er von einem DHCP-Server die Information über dessen Konfigurationsparameter (wie IP-Adresse und Subnet Mask) an. Optional kann der Client auch zusätzliche Angaben wie z.B. die Adressen von Routern (Default-Gateway), Domain Name Server (DNS) beim Server abrufen. Diese zusätzlichen Konfigurationsparameter

werden als Optionen beim Protokoll DHCP definiert. Die Beschreibung von allen derartigen Optionen enthält das Dokument RFC 1533.

Es ist hervorzuheben, dass einige Rechner weiterhin manuell konfiguriert werden können. Oft handelt es sich hierbei um die Rechner mit der Kommunikationssoftware der "alten Generation", so dass das Protokoll DHCP nicht unterstützt werden kann (Nicht-DHCP-Client).

Beim DHCP-Protokoll können sogenannte DHCP-Relay-Agenten implementiert werden. Ein solcher Agent hat die Aufgabe, DHCP-Nachrichten in andere Subnetze weiterzuleiten, die nicht über einen eigenen DHCP-Server verfügen. Ein Relay-Agent wird entweder in einen IP-Router oder in einen für diesen Zweck konfigurierten Rechner impleinentiert. Der Einsatz von Relay-Agenten hat den Vorteil, dass nicht für jedes Subnetz ein eigener DHCP-Server zur Verfügung gestellt werden muss. Andererseits besteht die Gefahr, dass beim Ausfall eines DHCP-Servers einige Clients nicht in der Lage sind, am Netzbetrieb teilzunehmen. Es ist deswegen erforderlich, sowohl redundante DHCP-Server als auch redundante DHCP-RelayAgenten immer einzuplanen. Aus diesen Gründen läßt das Protokoll DHCP mehrere DHCP-Server sowie mehrere DHCP-Relay-Agenten zu

Wir betrachten den Einsatz des Protokolls DHCP beim Remote-Access auf ein LAN über das ISDN. Nehmen wir an, dass es sich um ein Access großes und bundesweit agierendes Versicherungsunternehmen handelt. Die einzelnen Filialen dieses Unternehmens werden über das ISDN an ein zentrales LAN angebunden. Eine Besonderheit dieser Vernetzung besteht einerseits darin, dass die Anzahl der PCs am ISDN sehr groß ist (z.B. über 500) und dass diese PCs nur sporadisch auf den zentralen Server im LAN des Unternehmens zugreifen. Andererseits ist es nicht sinnvoll, einen DHCPServer am ISDN zu installieren, so dass die Funktion eines DHCP-RelayAgenten im Router notwendig ist. Dieses Unternehmen verfügt über einen Pool von IP-Adressen der Klasse C. Falls nur zwei Subnetze organisiert werden, d.h. das LAN an der Zentrale als ein Subnetz und die Remote-PCs am ISDN als weiteres Subnetz, so kann die maximale Anzahl von Hosts in jedem Subnetz nur 62 betragen. In diesem Fall könnte folgende Lösung in Frage kommen: Die 62 IP-Adressen werden als ein Pool von Adressen im DHCP-Server im LAN für alle Remote-PCs zur Verfügung gestellt und den einzelnen PCs am ISDN nach Bedarf dynamisch zugewiesen. Da die Remote-PCs nur sporadisch auf das LAN zugreifen und die "Belegung" der IP-Adresse nicht lange dauert, kann man davon ausgehen, dass alle PCs (ca. 500) mit den 62 Adressen zufriedenstellend bedient werden.

Bei der dynamischen Zuweisung wird einem Rechner eine IP-Adresse für einen bestimmten Zeitraum zugeteilt. Dieser Zeitraum wird mit dem englischen Wort Lease bezeichnet. Der Rechner kann aber auch von sich aus die Adresse vorher wieder freigeben, wenn er sie selbst nicht mehr benötigt. Der Vorteil besteht darin, dass eine von einem DHCP-Client nicht mehr benötigte IP-Adresse an einen beliebigen anderen DHCP-Client vergeben werden kann.

Wenn beim DHCP-Server eine Anforderung eintrifft, wählt er die IP-Adresse aus einem Pool von EP-Adressen aus und bietet sie dem DHCPClient an. Falls der Client die angebotene IP-Adresse akzeptiert, wird sie ihm für einen festgelegten Zeitraum (Lease) zur Verfügung gestellt. Wenn keine IP-Adressen mehr im Pool beim DHCP-Server vorhanden sind, kann einem Client auch keine Adresse zur Verfügung gestellt werden, so dass er nicht initialisiert werden kann.

Aufbau von DHCP-Nachrichten

Zwischen einem DHCP-Client und einem DHCP-Server werden festgelegte DHCP-Nachrichten übermittelt, für die das verbindungslose Protokoll UDP eingesetzt wird. Der DHCP-Client stellt einen Anwendungsprozess im einem Rechner dar und ist über den Well Known Port 68 zu erreichen. Der DHCP-Server ist ein Anwendungsprozess in einem dedizierten Rechner und erreichbar über den Well Known Port 67. Dese Port-Nummern werden im UDP-Header angegeben.

Die folgenden Felder werden in DHCP-Nachrichten verwendet:

- op (1 Byte), Operation: Angabe, ob es sich um eine Anforderung (Request) oder eine Antwort handelt.
- htype (1 Byte): Hier wird der Netztyp g'emäß RFC 1340 (Assigned Number) mitgeteilt (z.B. 6 = IEEE 802 x-LANs).
- hlen (1 Byte): Länge der Hardware-Adresse, d.h. physikalischen Netzadresse (6 für eine MAC- Adresse).
- hops (1 Byte, optional). Hier wird die Anzahl von Routern mit der DHCP-Relay-Funktion auf dem Datenpfad angegeben.
- xid (4 Bytes), Transaktions-ID: Dies ist die Identifikation für die Transaktion zwischen dem Client und Server, um den DHCP-Clients im Server die Antworten zu den richtigen Anforderungen (Requests) zuordnen zu können.
- secS (2 Bytes), Sekunden: Wird vom Client ausgefüllt und bedeutet die Zeit in Sekunden, die seit Beginn des Vorgangs abgelaufen ist.
- flags (2 Bytes): Das höchstwertige Bit dieses Feldes zeigt an, ob ein Client in der Lage ist, die IP-Pakete zu empfangen. Ist dies der Fall, verfügt der Client noch über eine gültige IP-Adresse. Die restlichen Bits dieses Feldes werden zur Zeit nur auf 0 gesetzt und sind für zukünftige Zwecke reserviert.
- ciaddr (4 Bytes), Client-IP-Adresse: Wird vom Client ausgefüllt, falls er eine IP-Adresse besitzt.
- yiaddr (4 Bytes), Your-IP-Adresse: Hier wird die IP-Adresse eingetragen, die der Server dem Client zugewiesen hat.
- siaddr (4 Bytes), Server-IP-Adresse: Hier wird die IP-Adresse des Servers angegeben (z.B. in der Nachricht DHCP-OFFER), der bei der nächsten Anforderung benutzt werden soll.

- giaddr (4 Bytes, optional), IP-Adresse des Gateways bzw. Routers mit der DHCP-Relay-Funktion.
- chaddr (16 Bytes), Client-NIAC-Adresse.
- sname (64 Bytes, optional), Server-Name: Ein Client, der den Namen eines Servers kennt, von dem er Konfigurationsparameter haben will, trägt hier diesen Namen ein und stellt somit sicher, dass nur der angegebene Server auf dessen Anforderung antwortet. Enthält dieses Feld "Alle Bits 0" 1, so kann jeder DHCP-Server im Netz antworten.
- file (128 Bytes, optional), File-Naine: Der File-Name ist ein alphanumerischer String (Zeichenfolge). Diese Angabe ermöglicht einem 1)HCP-Client, eine bestimmte Datei zu bestimmen, die er vom Server abrufen will. Der Server ist somit in der Lage, die richtige Datei auszuwählen und sie z.B. mittels des Protokolls FTP dem Client zukommen zu lassen.
- options (312 Bytes, optional): Zusätzliche bzw. herstellerspezifische Konfigurationsparameter. Dieses Feld enthält sogenannte DHCP-Optionen, die im Dokument RFC 1533 festgelegt werden.

DHCP im Einsatz

Der Einsatz des Protokolls DHCP zur automatischen Konfiguration der IP-Adressen bedeutet, dass der Benutzer eines Rechners keine IP-Adressierungsinformationen mehr von einem Administrator benötigt, um TCP/IP-Parameter zu konfigurieren. Der DHCP-Server stellt allen DHCP-Clients die erforderlichen Konfigurationsinformationen zur Verfügung.

Das Protokoll DHCP lässt mehrere DHCP-Server zu. Ein wichtiger Grund dafür ist die Server-Verfügbarkeit. Fällt ein Server aus, werden seine Funktionen automatisch durch andere Server übernommen.

Es sind vier Phasen nötig, um einem Rechner eine IP-Adresse zuweisen zu können:

- Anforderungsphase: Der Client sendet die Nachricht DHCP-DISCOVER in einem IP-Broadcast-Paket (Ziel-IP-Adresse = 255.255.255.255) als eine Anforderung, um von einem Server die benötigten IP-Adressierungsinformationen (IP-Adresse, Subnet Mask etc.) zu bekommen. Ein Wert, der unbedingt in dieser Nachricht angegeben werden muss, ist die MACAdresse des Clients (im Feld: chaddr).
  Die Client-Anforderung DHCP-DISCOVER als Broadcast wird normalerweise auf das eigene Subnetz eingeschränkt. Diese Client-Anforderung kann aber über eventuell vorhandene DHCP-Relay-Agenten in die weiteren Subnetze weitergeleitet werden. Der Einsatz von DHCPRelay-Agenten hat dann eine große Bedeutung, wenn nicht alle Subnetze über ihre eigenen DHCP-Server verfügen.

- Angebot-Phase: Jeder DHCP-Server kann mit einer Nachricht DHCP-OFFER dem Client sein Angebot von IP-Adressierungsinförmationen zukommen lassen. Der Server versucht zuerst, dem Client direkt das Angebot zu senden. Aber dies ist nicht immer möglich. Hierbei sind zwei Fälle zu unterscheiden:

  1. Der Client wird gerade initialisiert, so dass er noch über keine eigene IP-Adressse verfügt. In diesem Fall sendet der Server sein Angebot als Broadcast-Nachricht (IP-Adresse 255.255.255.255). Diese DHCPNachricht enthält bereits die MAC-Adresse des betreffenden Clients, so dass nur der "richtige" Client diese Nachricht lesen darf.

  2. Der Client verfügt bereits über eine IP-Adresse, doch die Lease-Dauer geht zu Ende, so dass er diese Adresse auf die nächste Lease-Periode "verlängern" möchte. In diesem Fall wird das Angebot vom Server direkt an den Client gesendet.

- Auswahlphase: In dieser Phase wählt der Client die IP-Adressierungsinformationen des ersten von ihm empfangenen Angebots aus und sendet eine BroadcastNachricht DHCP-REQUEST, um das ausgewählte Angebot anzufordern. In der Nachricht DHCP-REQUEST ist der Name des ausgewählten DHCP-Servers enthalten (=> Feld sname). Es kann hier auch die angebotene IP-Adresse mit Hilfe der Option Requested IP Address (Angeforderte IP-Adresse) bestätigt werden.

  Die Nachricht DHCP-REQUEST wird als Broadcast verschickt, um allen übrigen DHCP-Servem, die möglicherweise seine Angebote für den Client reserviert hatten, mitteilen zu können, dass sich der Client für einen anderen Server entschieden hat. Diese übrigen Server können die reservierten Parameter wieder freigeben, um sie anschließend anderen Clients anzubieten.

- Bestätigungsphase: Dieser DHCP-Server, der vom Client ausgewählt wurde, antwortet mit der Nachricht DHCP-ACK, die alle Konfigurationsparameter für den Client enthält. Nach dem Empfang von DHCP-ACK und nach dem Eintragen von Parametern wird beim Client der Konfigurationsvorgang beendet. In dieser Phase können eventuell noch die weiteren "verspäteten" Angebote eintreffen. Sie werden nun vom Client einfach ignoriert.

Das Protokoll DHCP stellt die weiteren drei Nachrichten zur Verfügung:

- DHCP-NAK: Diese Nachricht wird in der Bestätigungsphase verwendet und dann von einem ausgewählten DHCP-Server an einen Client gesendet, um darauf zu verweisen, dass die in der Nachricht DCHP-REQUEST geforderten Konfigurationsparameter abgelehnt wurden. Dies kann dann erfolgen, wenn:
  - ein Client versucht, die Lease für seine bisherige IP-Adresse zu verlängem und diese IP-Adresse nicht mehr verfügbar ist.

- die IP-Adresse ungültig ist, weil der Client in ein anderes Subnetz "umgezogen" ist.

- DHCP-RELEASE: Mit dieser Nachricht teilt ein DHCP-Client einem Server mit, dass einige Parameter (z.B. IP-Adresse) nicht mehr benötigt werden. Damit werden diese Parameter freigegeben und stehen anderen Clients zur Verfügung; dies müssen die Remote-PCs am ISDN tun.
- DHCP-DECLINE: Mit dieser Nachricht teilt ein DHCP-Client dem Server mit, dass einige "alte" Parameter (wie z.B. dessen MAC-Adresse) ungültig sind.
- DHCP-INFORM: Diese Nachricht ist nur in neuen Protokoll DHCP enthalten (RFC 2131). Diese Nachricht kann ein Client nutzen, dem eine statische IP-Adresse manuell zugeteilt wurde, doch er möchte dynamisch zusätzliche Konfigurationsparameter vom DHCP-Server zugeteilt bekommen.

Alle DHCP-Clients versuchen, ihre Lease zu erneuern, sobald die Leasedauer zu 50 Prozent abgelaufen ist. Um seine Lease zu erneuern, sendet der Client eine Nachricht DHCP-REQUEST direkt an den DHCP-Server, von dem er zuvor die Konfigurationsparameter erhalten hat. Der DHCP-Server bestätigt dies dem Client mit einer Nachricht DHCP-ACK, in der eine neue Lease-Dauer und alle aktualisierten Konfigurationsparameter enthalten sind. Wenn der Client diese Bestätigung erhält, aktualisiert er entsprechend seine Konfigurationsparameter.

Versucht ein Client, seine Lease zu erneuern, ist jedoch der gewünschte DHCP-Server nicht erreichbar, kann der Client die Parameter (IP-Adresse) dennoch weiter verwenden, weil noch 50% der Lease-Dauer verfügbar ist. Wenn die Lease nach dem Ablauf von 50% der Dauer nicht vom ursprünglichen DHCP-Server erneuert werden konnte, versucht der Client nach Ablauf von 87,5% der Lease-Dauer, einen anderen DHCP-Server in Anspruch zu nehmen. Hierfür sendet der Client eine Broadcast-Nachricht DHCPREQUEST. Jeder beliebige DHCP-Server kann darauf antworten:

- mit einer Nachricht DHCP-ACK, wenn er diese Lease erneuert hat, oder
- mit einer Nachricht DHCP-NAK, wenn er den DHCP-Client zur Neuinitialisierung und Übernahme einer neuen Lease für eine andere IP-Adresse zwingen will.

Wenn ein DHCP-Client neu gestartet wird, versucht er zuerst vom ursprünglichen DHCP-Server dieselbe IP-Adresse zu erhalten. Er erreicht dies, indem er einen DHCP-REQUEST als Broadcast verschickt und die zuletzt erhaltende IP-Adresse angibt. Wenn dies keinen Erfolg hat und die Lease-Dauer noch nicht zu Ende ist, kann der DHCP-Client dieselbe IP-Adresse über die verbleibende Lease-Dauer noch verwenden.

Wenn die Lease-Dauer abläuft oder eine Nachricht DHCP-NAK empfangen wird, muss der DHCP-Client unmittelbar die Verwendung der IP-Adresse einstellen und einen neuen Prozess der Vergabe von neuen IP-Adressen starten. Ist die Lease bei einem Client abgelaufen, der keine neue Lease erhalten hat, wird die TCP/IP-Kommunikation so lange eingestellt, bis eine neue IP-Adresse zugewiesen werden kann.

Implementierung von mehreren DHCP-Servern

Wenn in einem Netz mehrere DHCP-Server benötigt werden, muss ein eindeutiger Bereich von IP-Adressen für jedes Subnetz eingeplant werden. Ein Pool von IP-Adressen ist eine Folge von IP-Adressen, die für die Vergabe an Clients zur Verfügung stehen. Um sicherzustellen, dass Clients möglichst immer eine IP-Adresse erhalten, ist es wichtig, für jedes Subnetz mehrere Bereiche auf den verschiedenen DHCP-Servem zu reservieren.

Im allgemeinen sollte man die verfügbaren IP-Adressen folgendermaßen auf die DHCP-Server verteilen:

- Jeder DHCP-Server sollte über einen Bereich mit ca. 75% der für das eigene Subnetz bestimmten IP-Adressen verfügen.
- Jeder DHCP-Server sollte für jedes Remote-Subnetz über einen Bereich mit ca. 25% der für dieses Remote-Subnetz bestimmten IP-Adressen verfügen.

Wenn der DHCP-Server eines Clients nicht verfügbar ist, kann dieser Client immer noch eine IP-Adresse von einem anderen DHCP-Server zugeteilt bekommen, der sich in einem anderen Subnetz befindet, unter der Voraussetzung, dass ein DHCP-Relay-Agent im Router implementiert ist.

Bei der Installation eines DHCP-Servers sind folgende Punkte zu beachten:

- Bevor ein DHCP-Server die IP-Adressen an DHCP-Clients vergeben kann, muss er über einen Bereich von gültigen IP-Adressen verfügen.
- Deshalb ist es notwendig, jedem DHCP-Server eine eindeutige statische IP-Adresse (manuell) zuzuweisen. Der- DHCP-Server selbst kann kein DHCP-Client sein.
- Nicht-DHCP-Clients besitzen die statischen IP-Adressen, die manuell angegeben werden müssen.

- Die statischen IP-Adressen dürfen im Pool von für DHCP-Clients verfügbaren IP-Adressen nicht enthalten sein. Anderenfalls könnte der DHCP-Server dieselbe Adresse einem anderen Client zuweisen, was zu Problemen aufgrund doppelt vorhandener Adressen führen würde.
- Falls die IP-Adressen mit einen DHCP-Server den Clients in mehreren Subnetzen vergeben werden, müssen alle Router, die die einzelnen Subnetze verbinden, auch als DHCP-Relay-Agenten dienen.
- Werden mehrere Subnetze mit Routern vernetzt, aber diese Router unterstützen nicht die DHCP-Relay-Funktion, ist in jedem Subnetz, das DHCP-Clients enthält, zumindest ein DHCP-Server erforderlich.

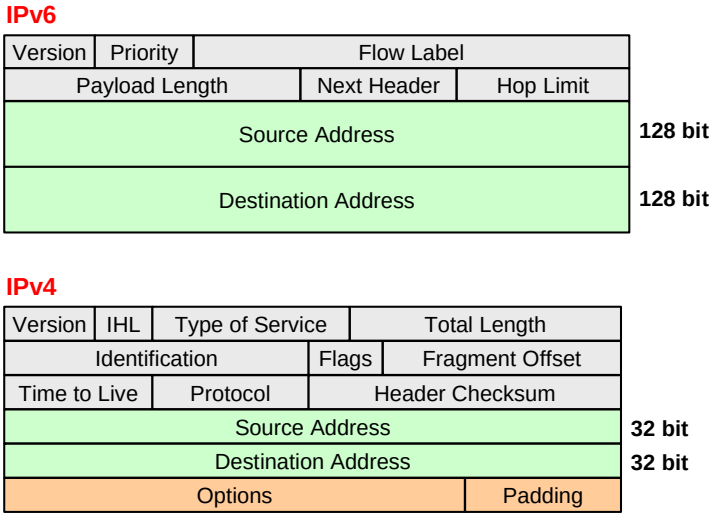
Das Protokoll DHCP ist Gegenstand mehrerer Internet-Dokumente RFCs 1533, 1534, 1541, 1542 und 2131).

Adressierung höherer Protokoll-Schichten
Im Jahre 1992 war abzusehen, dass das bestehende Internet Protokoll (Version 4) an seine Grenzen stoßen wird, vor allem bezüglich des Adressraumes. Deshalb wurde eine Arbeitsgruppe der IETF unter dem Begriff IP next generation (IPng) gegründet, die Vorschläge für ein neues Protokoll erarbeitet hat - heute unter dem Namen IP Version 6 (IPv6) bekannt. Ziele für diese neue Version waren:

Unglücklicherweise wurde in den frühen Internet-Tagen eine Nummerierung für das Versions-Feld angegeben, bei der die Nummer 4 für IP, die Nummer 5 für das sogenannte Stream Protocol (ST) und 6 für ein Simple Internet Protocol (SIP) vergeben wurde. Anscheinend wurde letzteres nie realisiert, so dass die neue Version die Nummer 6 erhalten hat.

- IP - Internet Protokoll
- größerer Adressraum,
- automatische Konfiguration (z. B. der Adressen),
- leichteres Routing,
- bessere Netzstrukturierung,
- verbesserte Sicherheitsfunktionen,
- Unterstützung für Echtzeit- und Multimedia-Dienste.

Gelöst wurde dieses durch einen vereinfachten Protokollkopf. Gegenüber den 13 Feldern bei IP4 (ohne Optionen) sind es nur noch 8 Felder bei IPv6 (ohne Erweiterungs-Köpfe). Für zusätzliche Funktionen können bei Bedarf Erweiterungs-Köpfe angefügt werden. Also wenn z. B. eine Sicherheitsfunktion nicht benötigt wird, dann verbraucht sie auch keinen Platz.



Aufbau des IP-Headers (lPv6)
Das rasche Wachstum des Internet führt zusammen mit den Beschränkungen von IPv4 (begrenzter Adressraum, starre Klasseneinteilung) zu Engpässen, die durch die Version 6 von IP beseitigt werden sollen.

Bild: IPv6 Header

Die Felder im 40 Byte langen lPv6-Header haben die folgende Bedeutung:

- **V (Version):** enthält den Wert 6.
- **TClass** (Traffic Class): Dringlichkeit des Pakets.
- **Flow Labet:** Zur Angabe des Typs der enthaltenen Daten.
- **Payload Length:** Länge der Nutzdaten im Feld, das auf den Header (bzw. die Extension Headers) folgt. Die Länge wird in Byte (Byte) angegeben, der maximale Wert beträgt 64 KByte. Durch ein Erweiterungsfeld (mit Extension Header des Typs Fragmentation) sind größere Werte möglich.
- **NH (Next Header):** Zahl, definiert den Typ eines nächsten Headers, der unmittelbar nach dem Feld Zieladresse folgen

kann.

- **HL (Hop Length):** Zahl (anfänglicher Maximalwert ist 254), die in jedem Zwischensystem dekrementiert (um 1 verringert) wird. Das Datagramm wird vernichtet, falls der Wert 0 wird, bevor das Ziel erreicht ist.
- **Quellen- und Zieladresse:** Die Länge der Adressen beträgt 16 Byte (128 bit), damit beträgt die Größe des Adressraums 3,4 - 10-38. Die Adressen sind hierarchisch aufgebaut um den Routing-Aufwand zu verringern. Eine sog. Cluster-Adresse bezeichnet eine geografische Region des Netzes.

Das Prioritätsfeld unterscheidet zwei Verkehrsarten:

Lastgesteuerter Datenverkehr, der einer Flusskontrolle unterworfen werden darf, wie sie z. B. TCP bereitstellt, und Echtzeit-Verkehr, der eine konstante Datenrate und eine konstante Verzögerungszeit erfordert.

Innerhalb jeder Verkehrsart gibt es 8 Prioritätsklassen, die jeweils nur für die Verkehrsart gelten, also keine übergreifende Bedeutung haben. (So stellt 6 eine höherer Priorität als 2 dar, aber über die Priorität von 13 gegenüber 2 ist keine Aussage zu machen.)

Mit dem Flow Label wird ein Ansatz In Richtung einer Verbindung innerhalb des verbindungslosen IP gemacht: Angenommen, mehrere Pakete gehören irgendwie zusammen, z. B. sie transportieren alle Anteile einer Datei, dann können diese Pakete mit einem einheitlichen Wert für das Flow Label gekennzeichnet werden; dieser Wert wird zufällig bestimmt. Router merken sich Flows und behandeln alle Pakete, die zu einem Flow gehören, in gleicher Weise - wie, das muss vorher über andere Mechanismen (spezielles Protokoll oder manuell) ausgehandelt werden. Da es kein definiertes Ende eines Flows gibt, können die Router nur über die Alterung den Zustand eines Flows wieder löschen.

Das Feld Next Header zeigt auf den ersten Erweiterungs-Kopf, soweit ein solcher vorhanden ist, oder auf das Transportprotokoll, vergleichbar dem Protokoll-Feld in IP4. Einige Werte zeigt Tabelle 5.4, Verweise auf Erweiterungs-Köpfe sind markiert. (Diese Werte entsprechen - bis auf die Erweiterungs-Köpfe - den Werten des Protokoll-Feldes in IP4.)

Das Feld Hop Limit entspricht grob der Funktionalität des Time-to-Live Feldes in IPv4. Es hatte sich herausgestellt, dass die Angabe einer wirklichen Zeit keinen Vorteil bringt, mehr noch, die gängigen Implementierungen sowieso schon das Hop Limit als reinen Hop Count", also die Anzahl Router auf dem Weg, verwendet haben.

Wenn Erweiterungs-Köpfe notwendig sind, dann werden diese durch das Feld Next Header angekündigt. Dabei ist eine Reihenfolge vorgegeben.

Hop-by-Hop Options Header. Dieser Kopf ist der einzige, der von jedem Router auf dem Weg zum Ziel ausgewertet wird, alle anderen Erweiterungs-Köpfe haben nur Ende-zu-Ende-Relevanz.

Routing Header. Damit wird der Weg durch das Netz explizit angegeben. Dazu enthält der Routing Header eine Tabelle von Router-Adressen. (Man nennt dies Source Routing", da die Quelle den kompletten Weg vorgibt.)

Fragment Header. Im Gegensatz zu IP4 wird in IPv6 keine Fragmentierung auf dem Weg zum Ziel durchgeführt, sondern diese Aufgabe wird an die sendende Station delegiert. Sollte auf dem Weg ein Router ein zu großes Paket erhalten, so sendet er eine ICMP-Nachricht an den Sender mit der Meldung, dass zu fragmentieren ist.

Der Sender wird dann sein Datenpaket in Fragmente zerlegen, jedes in ein eigenes lPv6-Paket verpacken und mit einem Fragment-Header versehen. Dieser enthält die notwendigen Elemente, wie sie in IP4 schon im Protokollkopf vorgesehen sind: Identification, Offset und More-Fragments-Flag.

Die minimale Paketgröße wurde von 576 Byte bei lPv4 auf 1280 Byte bei lPv6 angehoben.

Authentication Header bei IPv6 und Encapsulation Security Header dienen der Authentisierung und Verschlüsselung der Daten.

Destination Option Header transportiert zusätzliche Empfänger-Informationen. Das Format des Inhalts muss vorher vereinbart werden.

Base Header

Base Header Next = TCP	TCP Segment

Base Header and One Extension Header

Base Header Next = Route	Route Header Next = TCP	TCP Segment

Base Header and Two Extension Headers

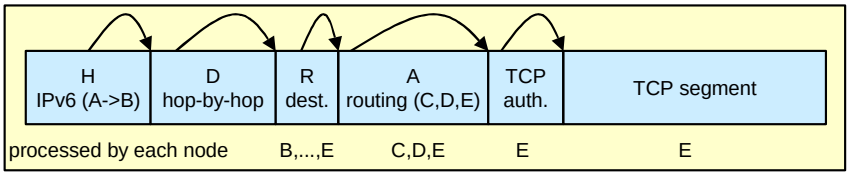
Base Header Next = Route	Route Header Next = Auth	Auth Header Next = TCP	TCP Segment

Bild: Extension Header

Next Header	Routing Type	Num. Address	Next Address
Reserved	Strict/Loose Bit Mask		
Address 1			
Address 2			
Address n			

- Strict ⇒ Discard if Address [Next-Address] <> neighbor
- Type = 0 ⇒ Current source routing
- Type > 0 ⇒ Policy based routing (later)
- New Functionality: Provider selection, Host mobility, Auto-readdressing (route to new address)

Bild: Routing Header



- next header can be IPv4 ”tunneling”
- header extension: hop-by-hop options (HH0), routing, fragment, destination options (DO)
- DO, HHO: type-length-value TLV options
- HHO:
  - looked at by each node, immediately after header
  - jumbo payload option (32 bit)
- routing header:
  - fixed header may not contain final address if routing header!
  - mixed loose/strict source route (bitmask)
  - swap destination address and next address from routing header
- fragment header: like IPv4 (32 bit identification, offset, more fragments flag)
- explicit MTU message rather than try-until-fit

Bild: IPv6 Packet Structure

Zusätzliche Information kann in bis zu sechs weiteren Headern *(extension header)* angegeben werden. Dafür stehen sechs mögliche Typen zur Verfügung:

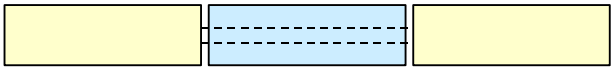
- **Hop-by-Hop Options:** Diese müssen von jedem Router ausgewertet werden.
- **Routing:** Für Source Routing (analog zu lPv4), bis zu 24 lPv6-Adressen können angegeben werden.
- **Fragmentation.** identifiziert ein Fragment aus einer Folge von Fragmenten (analog zu lPv4).
- **Authentirication:** enthält eine Prüfsumme, welche die Authentifikation des Senders erlaubt.
- **(ESP) Encapsulating Secure Payload:** enthält die Schlüsselnummer und die verschlüsselten Nutzdaten
- **Destination Options:** interessiert nur den Empfänger.

Die Header müssen, falls vorhanden, in einer bestimmten Reihenfolge angegeben werden. Jeder Header verweist in seinem Feld NH auf den Typ des nächsten Headers, der sich unmittelbar anschließt. Im letzten Extension Header steht im Feld NH die Nummer des auf der Transportschicht verwendeten Protokolls (analog zum PR-Feld in IPv4).

Die Herausforderung liegt damit im Zusammenwirken von Netzen auf Basis lPv4 mit Netzen auf Basis lPv6: Regeln für den Transport des Protokolls durch Domänen des anderen Protokolls sowie Übergänge sind notwendig und wurden inzwischen

auch beschrieben.

- Tunneling is used to cross islands with different protocols = Encapsulation



- IPv6 routers can encapsulate the original datagram in another IPv6, fragment it, and send it to the final destination.

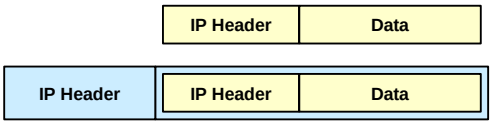


Bild: Tunneling

Koexistenz von IP4 und lPv6 bzw. Migration

Ein Wechsel von IP4 zu lPv6 an einem Stichtag ist wegen der Größe des Internet nicht möglich und auch nicht wünschenswert. Deshalb müssen die beiden Versionen über längere Zeit koexistieren. Dafür gibt es zwei Möglichkeiten:

- **Tunneln:** lPv6-Pakete werden durch IP4 unverändert weitergegeben. Die Pakete werden also nur in Systemen (Routern) ausgewertet, die auf lPv6 ausgelegt sind.
- **Parallele Implementierung** von lPv4 und lPv6 *(dual stack operation):* Systeme, auf denen beide Protokollstapel zur Verfügung stehen, können mit beiden Versionen kommunizieren.

Eine **Migration** von lPv4 zu IPv6 ist damit für einzelne Systeme, unabhängig von ihrer Umgebung, möglich. Bezüglich Sicherheit ist **IPSec** fester Bestandteil von IPv6.

Multicast-Verkehr

Multicast-Pakete sind immer UDP-Pakete und werden nach dem Best-Effort Prinzip verschickt. Es gibt damit - wie bei Datagrams üblich - keine Garantie für das Ankommen, die Fehlerfreiheit beim Empfänger und die Paketreihenfolge beim Empfänger der Pakete. TCP verbietet sich hier, denn TCP tauscht Status-Informationen zwischen einem Sender und einem Empfänger aus. Im Fehlerfall wird ein Multicast-Paket verworfen ohne eine Fehlermeldung (ICMP-Nachricht) zu generieren. Als Fehler können auftreten: der Host erhält ein Paket für eine Gruppe, der er nicht beigetreten ist; der Host erhält ein Paket für eine Gruppe, der er beigetreten ist, allerdings nicht über die Schnittstelle, über die er das Paket erwartet; das Paket enthält als Sendeadresse eine Gruppenadresse.

Die Kommunikationsprozeduren können in Senden, Empfangen und Routen unterschieden werden.

Senden: Die normale Sende-Prozedur kann (nahezu unverändert) übernommen werden. Die Adresse ist eine Multicast-Adresse. Der Sender kennt die einzelnen Mitglieder der Multicast-Gruppe, an die er sendet, nicht. Er muss auch nicht zwangsläufig Mitglied dieser Gruppe sein.

Empfangen: Wenn nicht eine feste Konfiguration vorgegeben wird (z. B. per Management), dann muss eine Prozedur vorhanden sein, mit der IP Hosts einer Gruppe beitreten bzw. die Gruppe wieder verlassen. Die IP-Schicht muss eine Tabelle von Multicast-Adressen der Gruppen beinhalten, der der Host beigetreten ist.

Routen: Im Netz müssen multicast-fähige Router vorhanden sein. Das können die normalen Router mit einer entsprechenden Erweiterung sein, oder den normalen Routern werden Multicast-Router beigestellt.

Gruppenmitgliedschaft

Die Mitgliedschaft eines Hosts in einer Gruppe kann dynamisch sein, der Host kann sich zu jeder Zeit der Gruppe anschließen (join) oder die Gruppe wieder verlassen (leave). Ein Sender muss nicht Mitglied einer Gruppe sein, wenn er an diese Gruppe Daten senden will.

Eine Gruppe kann permanent eingerichtet sein (per Management), wobei eine solche Gruppe auch zeitweise keine Mitglieder haben kann. Die Gruppe kann aber auch dynamisch eingerichtet werden; sie lebt nur solange, wie sie Mitglieder hat. Als Adresse kann sie eine der noch nicht verwendeten Multicast-Adressen benutzen.

Mit dem Internet Group Management Protocol (IIGMP) teilen IP Hosts ihre Multicast Mitgliedschaften an benachbarte Multicast-Router mit. Das Protokoll kann auch zwischen Routern eingesetzt werden. Es wird, vergleichbar zum ICMP, als Bestandteil von IP angesehen.

Multicast-Routing

Für das Weiterleiten von IP Multicast Paketen werden Multicast Router benötigt. Dabei findet keine Umrechnung von Multicast-Adressen in normale Adressen statt! Der Multicast-Router unterhält eine Routing-Tabelle aufgrund, derer er weiß, auf welchen Schnittstellen er Pakete mit diesen Adressen aussenden muss.

Man ging ursprünglich davon aus, dass es viele Multicast-Teilnehmer geben wird. Deshalb werden in diesem Vorschlag zuerst Broadcasts ausgesendet, um dann bei Nichtbenutzung Wege wieder aus dem Baum zu nehmen. Dieser klassische Ablauf des Multicast-Routing wird mit broadcast-and-prune bezeichnet, allgemein nennt man diese Art Protokolle dense-mode-Protokolle.

Der zweite, mehr realistischere Ansatz, geht davon aus, dass wenige Multicast-Mitglieder im Netz weit verstreut liegen und selbst aktiv werden, wenn sie sich einer Gruppe anschließen, also eine Join-Nachricht senden. Dieser Ansatz wird sparse mode genannt und zwei Protokolle sind heute im Einsatz: MOSPF und PIM-SM.

Beim MOSPF handelt es sich um eine Multicast-Erweiterung des OSPF. Bei OSPF verteilt ein Knoten Topologie-Information an seine Nachbarknoten (flooding). MOSPF verteilt zusätzlich Informationen über Empfänger von Gruppen. Damit kann dann ein Routing-Baum aufgebaut werden.

Protocol Independant Multicast (PIM) kennt zwei Varianten: Dense Mode (DM) und Sparse Mode (SM), letzteres wird hier betrachtet. Bei PIM-SM werden sogenannte Rendezvous Points (RP) im Netz eingerichtet, ein Vorgang, der per Management stattfindet und Ergebnis einer Netzplanung sein muss. Die Information, welche Router im Netz als RPs dienen, muss über ein geeignetes Protokoll ausgetauscht werden.

Empfänger senden Join"-Nachrichten an die RPs, die darauf einen Baum einrichten. Ein Sender sendet dann sogenannte Register-Pakete gezielt zu einem RP (es sind Multicast-Pakete verpackt in Unicast-Paketen.) Wenn der RP Gruppenmitglieder in seinem Baum hat, dann sendet er als Antwort eine Join-Nachricht in Richtung zur Quelle. Dieser Prozess wird solange fortgesetzt, bis der erste Router nach der Quelle eine join Nachricht erhält.

Der Vorteil der Sparse-Mode Protokolle ist, dass nur solche Router in den Prozess eingebunden sind, die zwischen der Quelle und einem Gruppenempfänger liegen. Und natürlich tragen auch nur solche Links Multicast-Verkehr, die zwischen Quelle und Gruppenempfänger liegen.

Das Multicast-Routing hat eine Reihe neuer Probleme gebracht.

- **Skalierbarkeit:** Mit der Zunahme der Anzahl Gruppen-Empfänger wurden die Routing-Tabellen immer größer und unübersichtlicher. Lösungen sind die Einführung einer Routing-Hierarchie und die Zusammenfassung von Routen (route aggregation).
- **Inter-Domain-Routing:** Die ersten Überlegungen gingen immer von einer flachen Hierarchie aus. Das Internet ist aber inzwischen stark strukturiert. Es wird durch eine Vielzahl von miteinander verbundenen Teilnetzen, sogenannten Autonomous Systems (AS), gebildet. Routing zwischen AS unterliegt strengeren Kriterien als innerhalb eines AS, da dort üblicherweise unterschiedliche Betreiber aneinander stoßen. Daher wurde die Notwendigkeit einer neuen Klasse von Routing-Protokollen zwischen Teilnetzen, den Inter-Domain Protokollen, erkannt. Für Multicast-Verkehr gab es kein solches Inter-Domain Routing-Protokoll. Deshalb wurde eine Multicast-Erweiterung des Border Gateway Protocols (BGP) vorgenommen (als BGMP bezeichnet). Das reicht allerdings noch nicht aus, denn nur innerhalb einer Domain kennt ein Empfänger den Sender von Multicast-Nachrlchten. Diese Information muss nun irgendwie über Domain-Grenzen transportiert werden. Dazu wird gerade ein neues Protokoll erarbeitet, das Multicast Source Discovery Protocol (MSDP).
- **Managebarkeit:** im Internet gibt es keine Instanz, die Gruppen verwaltet. Bisher basierte jegliche Verwaltung nur auf E-Mail-Listen, über die Interessierte miteinander kommunizierten.

IP-Multicasting

Die Adreß-Klasse D beinhaltet IP-Multicast-Adressen, die den Adressbereich 224.0.0.0 bis 239.255.255.255 überstreichen dürfen, was einem Umfang von ca. 250 Millionen potentiellen Adressen entspricht.
Aufgabe des IP-Multicasting ist es, eine Information, z.B. einen Video- oder Audio-Datenstrom - z.B. über das RTP-Protokoll -nicht einfach an n Teilnehmer von der Quelle aus zu replizieren, sondern über eine bestehende Multicast-Infrastruktur über geeignete Router gezielt an m lokale Multicast-Router zu verteilen, die diese dann an die k Zielempfänger der entsprechenden Multicast-Gruppen weiterleiten. Diese Faktorierung des Datenverkehrs entlastet das Transportnetz (z. B. Internet), erfordert aber zusätzliche Intelligenz sowohl in den Routersystemen als auch in den Endgeräten. Der Absender selbst muss kein Mitglied einer Multicast-Gruppe sein.

Im Zusammenhang mit dem IP-Multicasting sind daher vier Problemfelder zu lösen:

- Adressen-Mapping: Abbildung von IP-Multicast auf MAC-Multicast-Adressen.

- Endsystem-Registrierung: Der Empfänger eines Multicast-IP-Pakets muss einer ensprechenden IP-Multicast-Gruppe angehören und kann dieser beitreten bzw. sie verlassen. Dies ist Aufgabe des Internet Group Message Protocols (IGMP).
- Multicast-Routing: Informationen über die Multicast-Gruppen müssen verteilt und gepflegt werden. Hierzu stehen die Protokolle Protocol Independent Mulicast PIM, das Multicast Open Shortest Path First MOSPF sowie das Distance Vector **Routing Multicast Protocol DVMRP** zur Verfügung. Diese Protokolle wollen wir in Kapitel 9 vorstellen.

Eine reale Implementierung der Multicast-Infrastraktur stellt das Multicast Backbone Mbone dar, das 1992 ins Leben gerufen wurde Das MBone kann als logisches Netz innerhalb des Internet verstanden werden, dessen "M-Router" Multicast-Routing unterstützen. Sind die IP-Pakete zusätzlich über Netze mit Standardrouter zu übertragen, werden die Multicast-Pakete in Unicast-Pakete eingebettet (IP-/IP-Encapsulation) und somit über diese Netze getunnelt.

Ein simples Routingschema wird mit Hilfe des Time To Live (TTL)-Bits im IP-Header realisiert. Befindet sich die Multicast-Gruppe in einem lokalen IP-Netz, wird der TTL-Wert auf eins gesetzt, bei Empfängern in entfernten Netzen auf einen höheren Wert (Defaultwerte: 15, 63 oder 127). Die M-Router entscheiden aufgrund dieses TTL-Wertes, ob der Multicast weiterzuleiten ist oder nicht. Wie üblich, wird bei jedem Hop die TTL um 1 herabgesetzt. Im Gegensatz zum Standard-IP erhält der Multicast-Absender jedoch beim Erreichen des Wertes 0 für TTL keine ICMP-Benachrichtigung (Destination Unreachable), sondern das Multicast-Paket wird einfach verworfen.

Internet Group Management Protocol (IGMP)
Das IGMP ebenso wie das ICMP auf der Schicht 3 des OSI/ISO-Referenzmodells angesiedelt und unterstützt IP4 und die Multicast-Routing-Protokolle. IGMP liegt mittlerweile in der Version 2 vor (IGNIPv2, RFC 2236), eine Version 3 mit umfangreicheren Eigenschaften wird alsbald erwartet. IGMP ist auf allen Endsystemen und Routern zu implementieren, die Multi- IGMP-casting unterstützen (RFC 1054).

Die IGMP-Nachricht wird in ein IP-Paket mit der Protokollkennung 2 eingefügt und weist folgende Kenngrößen auf:
- **Type** (Nachrichtentyp für die Endsystem/M-Router-Kommunikation)
  -**Membership Query** (x'11')'): mit den Möglichkeiten einer allgemeinen (Welche Gruppen gibt es?) und einer Gruppenspezifischen Anfrage (Welche Mitglieder gibt es in der gewünschten Gruppe?) je nach Gruppenadresse.
  -Version 1 Membership Report (x'12') aus Kompatibilitätsgründen.
  -Distance Vector Multicast Routing Protocol (DVMRP-V3-)Nachricht (0x'13')
  -Version 2 Membership Report (x'16') zur Aufnahme von Endsystemen in Multicast-Gruppen (Membership Request).
  -Leave Group (x'17')
  -Version 3 Membership Report (x'22')
- **Maximum Response Time (MRT)** unterstützt die Membership Query in bezug auf die M-Router, die erst nach Ablauf dieser Zeit Updates der Routing-Information vornehmen müssen.
- **Checksum**: 16-Bit-Checksumme der gesamten IGMP-Nachricht.
- **Gruppenadresse**: Für die Membership Query wird diese Adresse zunächst auf 0 und anschließend in der Antwort auf den Wert der Multicast-Gruppe gesetzt. In einem Membership Report oder einer Leave Group-Nachricht entspricht diese natürlich der entsprechenden Multicast-Gruppe.
- **Sonstige Felder:** sind möglich und im Hinblick auf neuere Protokoll-Versionen vorgesehen.

Ein M-Router sendet in festen Zeitintervallen (100 Sekunden) IGMP-Nachrichten vom Typ Membership Query an alle Endsysteme im LAN über die Multicast-Adresse 224.0.01 (All-Host Group) mit einem TTL von 1. Die Multicast-Endsysteme antworten hierauf mit einem Membership Report und teilen dem M-Router mit, zu welcher Gruppe sie gehören.
Soll ein Endsystem in eine Multicast-Gruppe eingefügt werden, so sendet dieses einen MAC-Multicast mit einem IGMP Host Membership Report und der entsprechenden Multicast-Adresse aus, wobei die Zuordnung zwischen IP- und MAC-Multicast-Adresse erfolgt.
Ein Endsystem kann auf die gleiche Weise über eine IGMP Leave Group-Nachricht die Gruppe verlassen. Als Konsequenz daraus antwortet das Endsystem anschließend nicht mehr auf die zugehörige Membership Query des M-Routers.
Es obliegt dem M-Router, die Buchführung für die aktiven Multicast-Gruppen in seinem LAN und den zugehörigen Endsystemen vorzunehmen. Für den M-Router spielt es keine Rolle, ob in einer bestehenden Multicast-Gruppe zwanzig oder lediglich ein Mitglied vorhanden sind. Änderungen von Multicast-Gruppen teilt er jedoch über die ihm zur Verfügung stehenden Multicast-Routing-Protokolle mit.