

## Internetworking

*Andreas Polze*

Humboldt-Universität zu Berlin  
Fachbereich 16  
PSF 1297  
1086 Berlin

email : apolze@informatik.hu-berlin.de

Bei der bisherigen Betrachtung lokaler Netze wurde angenommen, daß alle Stationen — *hosts, end systems* — an einen *Netzwerktyp* angeschlossen sind. Hier sollen Probleme angesprochen werden, die bei Kommunikation über Netzwerke verschiedenen Typs entstehen.

Werden mehrere Netzwerke von einer Anwendung benutzt, so spricht man von *Internetworking*. Der Terminus *Internet* bezeichnet gleichzeitig ein weltweites Computernetzwerk, das Anfang der 70er Jahre aus dem ARPANET des DoD hervorgegangen ist und seitdem exponentiell wächst.

## Terminologie

Weltweite Netzwerke sind nur durch Kombination von LANs und WANs möglich geworden. Spezielle Geräte stellen die physische Verbindung und eine Protokollumwandlung her.

In OSI-Terminologie:

- *intermediate system(IS), internetworking unit(IWU)*

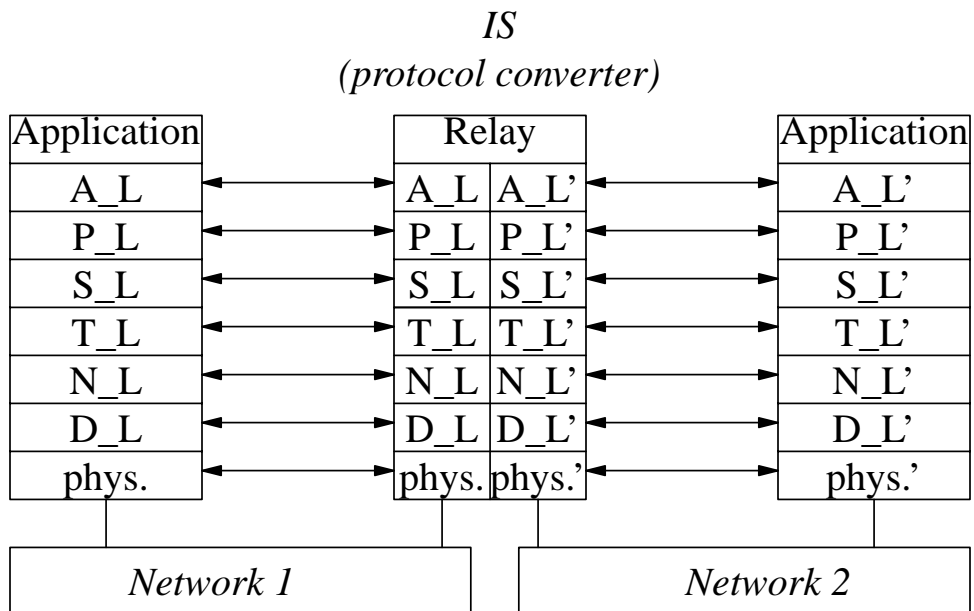
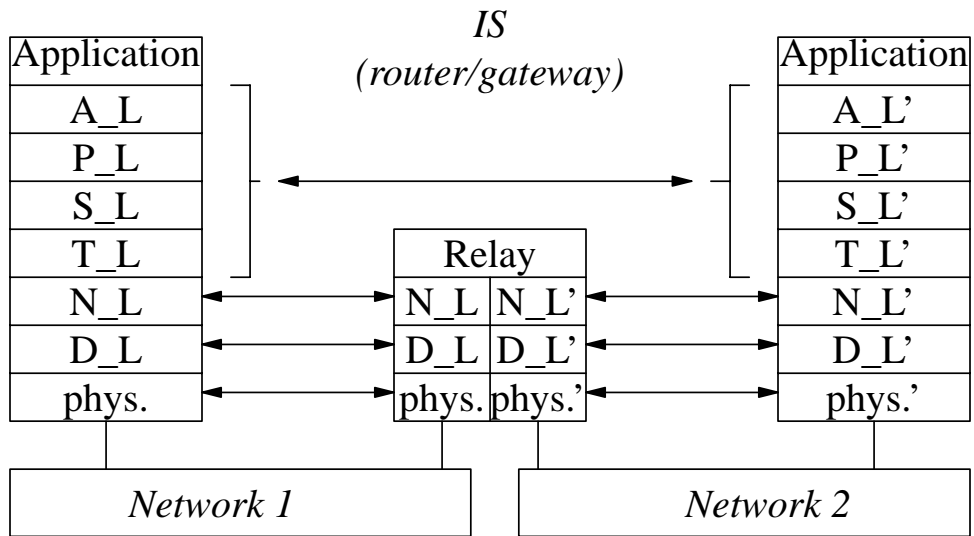
oder alternativ:

- **router**
- **gateway**
- **protocol converter**

Ein **protocol converter** ist ein IS, das zwei Netzwerke mit vollständig verschiedenen Protokollen verbindet.

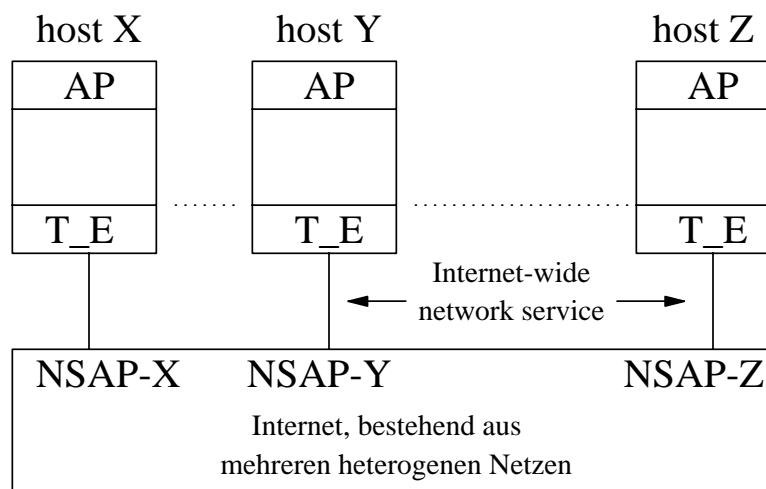
Ein **router** vollführt seine Aktivitäten auf der Ebene des *network layers*.

In einem **offenen System** sind die oberen Protokollebenen für alle Endsysteme identisch.



Aus Sicht des Internet Benutzers soll das Internet (ein Transport Protokoll) einen wohldefinierten Netzwerkdienst am *network service access point* (NSAP) anbieten.

- Verschiedene Netzwerke und -typen sollen transparent sein.
- Kein erkennbarer Unterschied zwischen Internet, einzelmem LAN oder WAN.



NSAP-X: Network service access point at system-X

AP: Application process

T\_E: Transport protocol entity

## **Anforderungen und Merkmale eines Internets**

Folgende Leistungen beschreiben ein Internet:

- Netzwerkdienste
- Adressierung
- Routing
- Qualität des Dienstes
- Maximale Paketgrößen
- Flußsteuerung, Überflutungskontrolle
- Fehlermeldungen

Zunächst sollen Probleme skizziert werden, die aus den Leistungsanforderungen an ein Internet erwachsen. Später sollen die Protokolle aufgeführt werden, mit denen diese Probleme für *das Internet* gelöst werden.

## Netzwerkdienste

LANs und WANs unterscheiden sich durch ihre Adressierung, Fehlerraten und Verzögerungszeiten.

Für LANs gilt:

- MAC sublayer Adressen identifizieren Stationen.
  - MAC Adressen und transparente **bridges** werden benutzt, um frames zwischen Systemen zu *routen*.
  - Normalerweise wird ein verbindungsloses Protokoll benutzt.
- Viele LANs bieten **connectionless network service** (CL-NS).

Anders dagegen gilt für WANs:

- Adressen der Link-Ebene haben nur lokale Signifikanz.
  - Adressen der Netzwerk-Ebene werden zur Identifikation von Systemen und für Routing benutzt.
  - Aufgrund der Verzögerungen und der Fehlerraten wird häufig ein verbindungsorientiertes Protokoll verwendet.
- Die meisten WANs bieten einen **connection-oriented network service** (CO-NS).

Da Internet-Nutzer an verschiedene Netze angeschlossen sein können, muß bei der Nutzung eines Dienstes zunächst der Typ angegeben werden.

In einem Internet, das aus verschiedenen Subnetztypen besteht, ist es nötig, den lokal verfügbaren Typ eines Dienstes in andere Typen umzuwandeln.

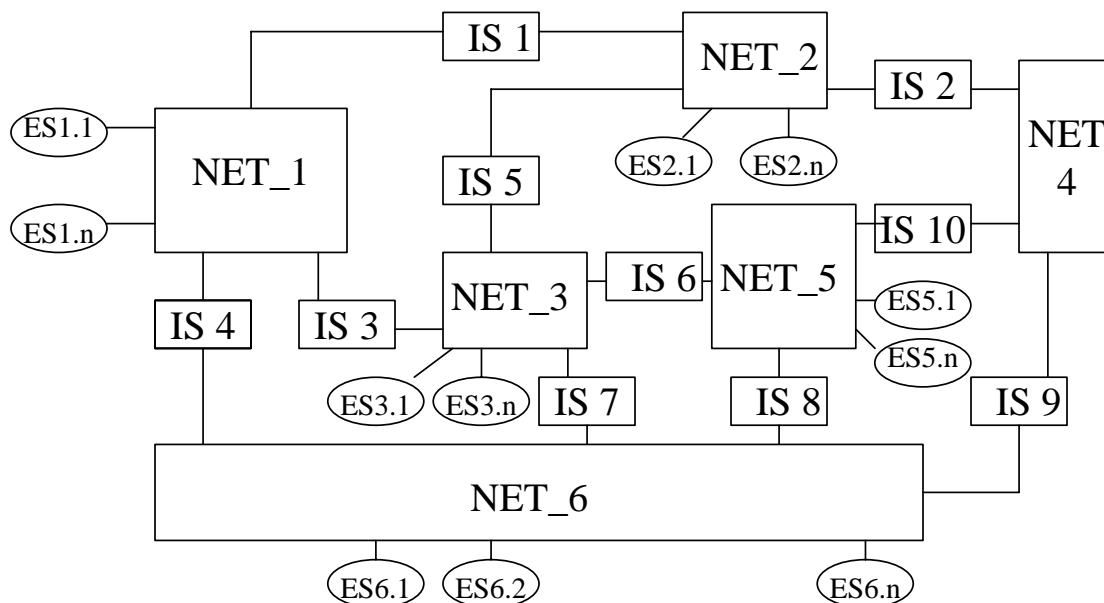
## Adressierung

Die NSAP (network service access point) Adresse wird zur Identifizierung eines Nutzers in einem *end system* benutzt.

- Diese Adressen sind nur in einem LAN/WAN eindeutig.
  - Es ist idR. nicht möglich, den *network point of attachment* als Basisadresse zu verwenden (zu viele Netze existieren bereits).
- In einem Internet ist ein vollständig neues Konzept von NSAP-Adressen nötig.



## Routing



Wie kann eine NPDU von Endsystem 1.1 zum Endsystem 5.1 gesendet werden ?

- In einem Internet können Endsysteme (ES) verschiedenen Netzwerken angehören.
- Die Adresse des *network point of attachment* (NPA) eines Zielsystems läßt sich aus der Adresse des *network service access point* nicht ablesen.
- Ein ES kann Pakete an ein *intermediate system* (IS) direkt senden, wenn es dessen NPA kennt.
- Wenn ein IS alle ES in allen angeschlossenen Netzen kennt, so können Pakete direkt zugestellt werden.

- Zusätzlich kann ein IS Pakete an andere IS im selben Netz senden.

Diese Eigenarten gestatten die Auswahl eines Weges durch mehrere verbundene Netze (Routing). Ein paar Fragen müssen zuvor beantwortet werden:

- Wie kann ein ES die NPA Adressen aller IS in seinem Netz erfahren?
- Wie kann ein IS die NPA Adressen aller angeschlossenen ES erfahren?
- Nach welchen Kriterien wählt ein ES ein bestimmtes IS als Empfänger eines Paketes?
- Wie erfährt ein IS die Adressen anderer IS im selben Netz.
- Wie wählt ein IS ein anderes IS, um ein Paket an ein bestimmtes Ziel-ES zu senden?

## Qualität des Dienstes

Ein *quality of service* (QOS) Parametersatz ist mit jeder Dienstprimitive assoziiert, die am NSAP empfangen wird. Dieser Parametersatz beschreibt die Leistungen, die der Netzwerkbenutzer vom Dienstanbieter erwartet.

Die QOS Parameter beschreiben:

- Die erwartete Verzögerung.
- Die Sicherheit gegenüber Modifikationen.
- Die Kosten eines Dienstes.
- Die Priorität des Dienstes.

In einem verbindungsorientierten Dienst findet eine Verständigung über die QOS Parameter beim Verbindungsaufbau statt.

Im verbindungslosen Fall muß der Dienstbenutzer die vom Dienstanbieter erbrachten QOS Parameter kennen.

## Maximale Paketgröße

Die maximale Paketgröße variiert in verschiedenen Netzwerktypen. Sie wird beeinflusst durch Faktoren wie:

- Bitfehlerrate.
- Übertragungsverzögerung: Je größer ein Paket ist, umso länger blockiert es einen Link (IS).
- Puffergrößen.
- Zusätzliche Bearbeitungskosten (overhead).

Für jedes einzelne Netzwerk ist die maximale Paketgröße bekannt.

In einem Internet muß jedes IS in der Lage sein, Pakete zu “zerstückeln” (*segmentation, fragmentation*), und wieder “zusammenzusetzen” (*reassembly*). Andernfalls darf die “kleinste maximale Paketgröße” nicht überschritten werden.

## Flußsteuerung und Überflutungskontrolle

Verschiedene Netze können Pakete in unterschiedlichen Übertragungsraten versenden. Ein “langsamer Empfänger” muß einen “schnellen Sender” bremsen können.

→ Eine Flußsteuerung ist nötig.

Schlimmer ist eine Überflutung eines “mittleren” Knoten. Diese Überflutung kann auch Pakete zur Flußsteuerung verzögern, so daß der “schnelle Sender” das Problem nicht erkennt.

verbindungsorientiertes Netz (X.25):

Flußsteuerung auf Basis eines “virtual circuits”. Ein *send-Window* erlaubt das Verschicken von n Paketen, dann muß auf eine Bestätigung gewartet werden.

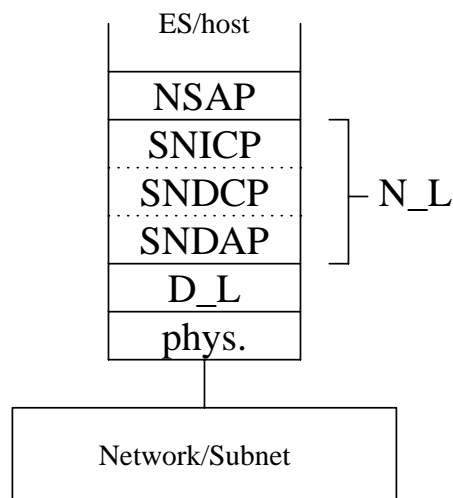
verbindungsloses Netz:

Flußsteuerung wird nicht vom Netz direkt unterstützt, sie geschieht auf *end-to-end* Basis.

## Substruktur der Netzwerkschicht

In jedem ES und IS müssen *routing*- und *relaying*-Funktionen transparent auf der Ebene der Netzwerkschicht ausgeführt werden. Im ISO Referenzmodell wird die Netzwerkschicht daher in 3 Subschichten mit eigenen Protokollen unterteilt.

- subnetwork independent convergence protocol (SNICP).
- subnetwork dependent convergence protocol (SNDACP).
- subnetwork dependent access protocol (SNDAP).



## Das Internet — etwas Geschichte

- Mitte der 70er Jahre begann die Entwicklung von TCP/IP durch DARPA (DoD's Advanced Research Project Agency).
- ARPANET - *packet-switched network*, über Punkt-zu-Punkt Verbindungen, Funk- und Satellitenkanäle, 1978/79.
- Erste Stücke des **Internets** entstanden 1980, als die ersten Rechner am ARPANET auf die TCP/IP Protokolle umgestellt wurden.
- Im Januar 1983 forderte DARPA, daß alle Rechner am ARPANET TCP/IP nutzen.
- Bolt Beranek & Newman, Inc.(BBN) wurde gegründet, um die TCP/IP Protokolle in das UNIX der *Berkeley Software Distribution* der *University of California, Berkeley* einzubauen (BSD4.2).
- DARPA konnte so mehr als 90% der wissenschaftlichen Rechner erreichen.

- BSD4.2 bot erstmals *socket*-Programmierschnittstelle. Netzwerkverbindungen wurden ähnlich einfach wie Dateien im UNIX handhabbar.
- Seit 1985/86 verbindet NSFnet alle sechs Supercomputer-Zentren in den USA und schließt sie an das ARPANET an.
- 1987 wuchs das Internet monatlich um 15%.
- März 1992, San Diego, 23. Internet Engineering Taskforce: Class-B Netze drohen für die nächsten Jahre auszugehen, das Ende der verfügbaren Class-C Netze ist für die Jahrtausendwende abzusehen, Routing-Tabellen explodieren.
- In Zukunft sollen 10 Mrd. Rechner vernetzt werden können.
- Internet Wachstum:

Zeitpunkt	Hosts(Knoten)	Domains(Zonen)
10/89	160 000	4 800
10/90	313 000	9 300
7/91	535 000	16 000

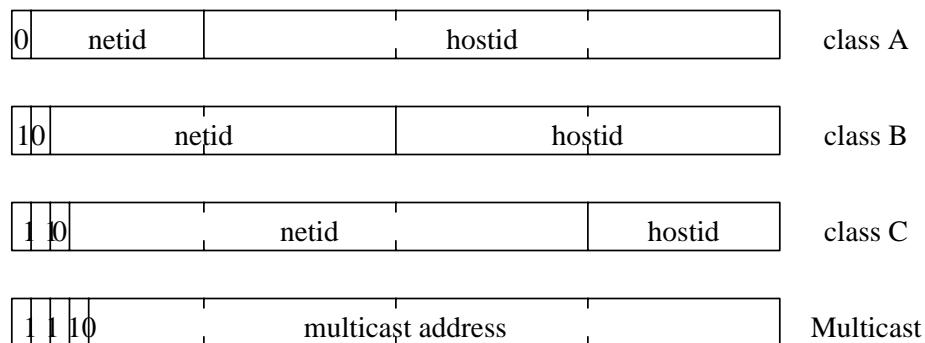


## Internet Adressen

Zwei Adressen beschreiben einen host/ES am Internet. In der ISO-Terminologie sind das die *network service access point* (NSAP) Adresse und die *subnet point of attachment* (SNPA) Adresse.

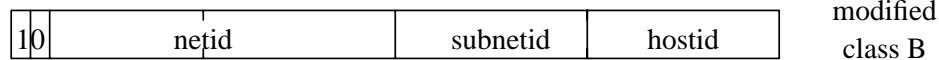
Innerhalb TCP/IP sind dies die IP Adresse und die NPA Adresse. IP Adressen sind Internet-weit eindeutig. NPA Adressen unterscheiden sich für verschiedene Netz- und Subnetztypen.

Es existieren vier Klassen von Adressen:



Für Institutionen mit einer großen Anzahl verschiedener LANs, die nicht mehr mit *bridges* verbunden werden können (Segmentation, Reassembly) wurde das Konzept der **Subnets** eingeführt.

*Routing* im lokalen Fall kann so von Internet-weitem *routing* getrennt werden:



Adreß-Masken definieren die Subadreß-Grenzen für ein spezielles Netzwerk.

Internet Adressen werden ebenso wie Adreß-Masken als “gepunktete Dezimalzahlen” geschrieben.

Beispiele sind:

255.255.255.0 - Adreß-Maske  
16.2.0.1 - gatekeeper.dec.com

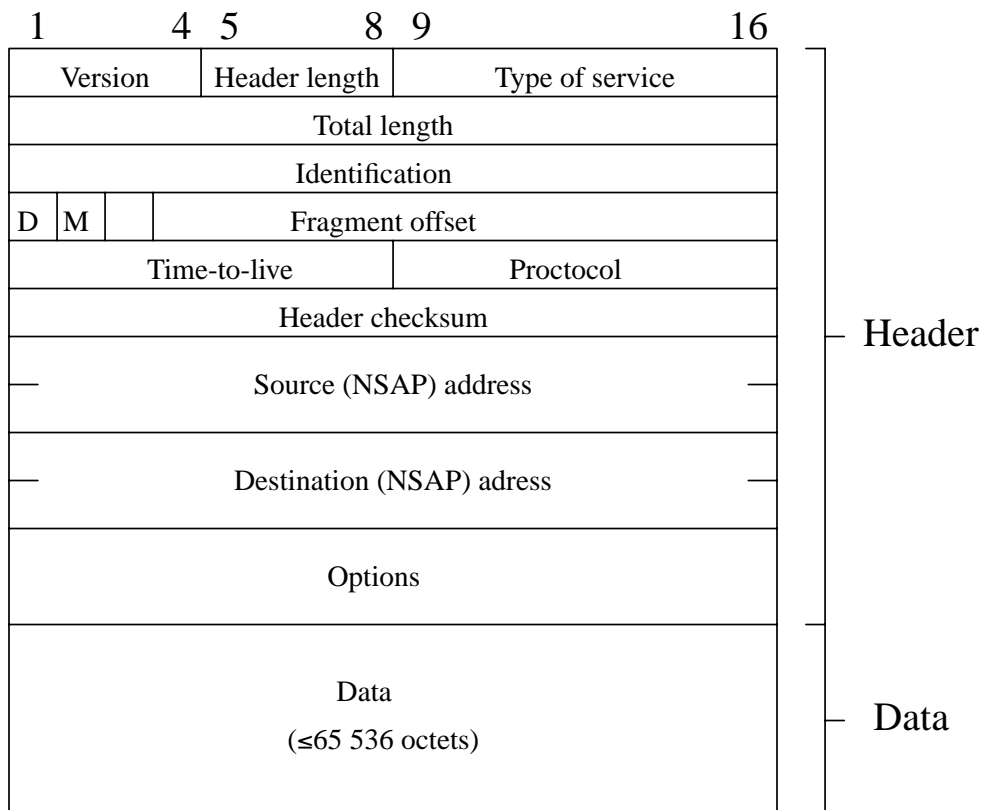
Ein 1-bit in einer Adreß-Maske kennzeichnet die betreffende position als Netzwerkadresse (netid + subnetid).

Um Eindeutigkeit für Netz- und Subnetzadressen zu gewährleisten, werden diese Adressen von einer zentralen Stelle, dem **Network Information Centre (NIC)** vergeben.

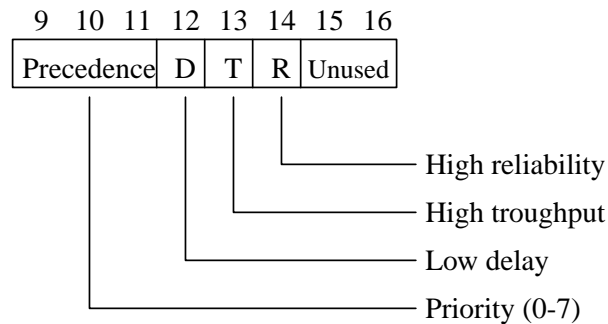
## Datagramme

Die *IP protocol data unit* wird als **Datagramm** bezeichnet. Ein Datagramm ist maximal 64 Kbytes groß (octets - so ist die Ordnung der Bits eindeutig festgelegt).

Ein Datagramm hat folgende Struktur:



Der *Type-of-Service* spielt dieselbe Rolle, wie der QOS-Parameter in der ISO Terminologie.



Einige Felder mit spezieller Bedeutung:

- *don't fragment (D-) bit*: Ist dieses Bit gesetzt, so kommt das Datagram entweder in einem Stück oder gar nicht beim Empfänger an.
- *more fragments (M-) bit*: Diese Bit wird beim Zusammensetzen von Fragmenten (Reassembly) ausgewertet.
- *Time-to-live*: Hier wird eine Zeit angegeben (in sec.). Jedes Gateway verringert den Wert, erreicht er Null, so wird das betreffende Datagram weggeworfen.
- *options*: Diese Feld wird mitunter für Debugging, Fehlermeldungen und *route redirection* benutzt.

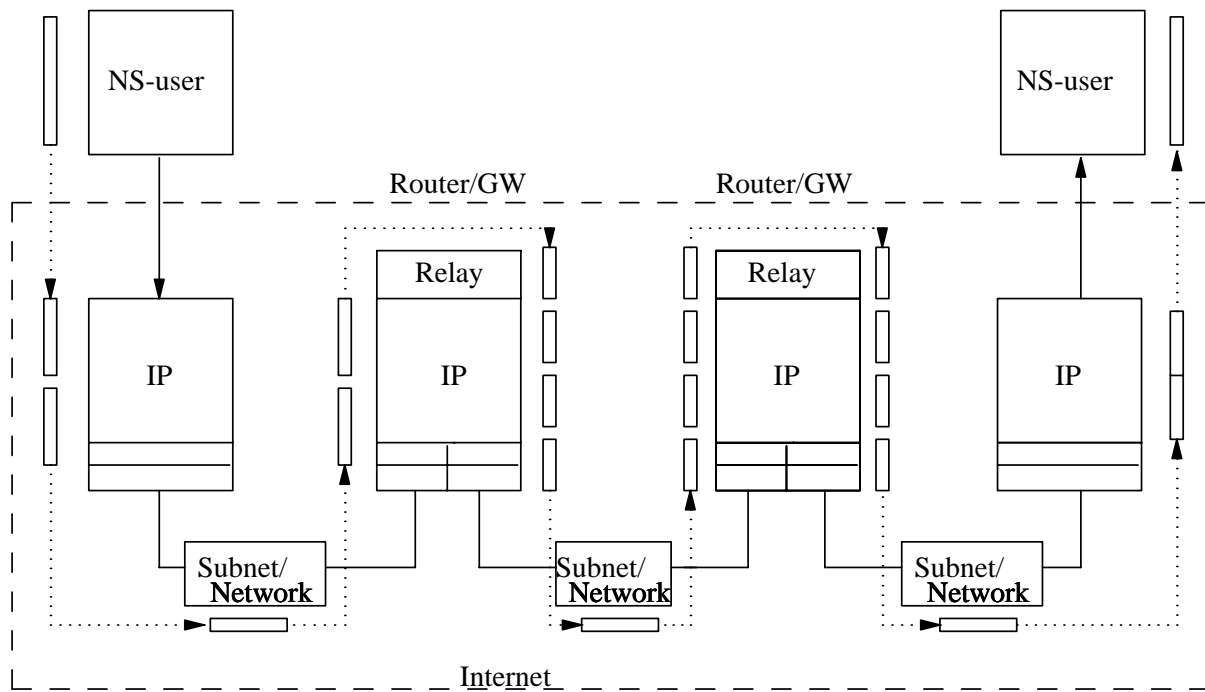
## Protokollfunktionen

### 1. Fragmentation and reassembly

Abhängig von der maximalen Paketgröße eines Netzwerks muß ein Datagramm u.U. “zerstückelt” werden.

Zwei Ansätze sind denkbar:

- *intranet fragmentation*: Alle Fragmente eines Datagramms werden in jedem Router zusammengesetzt. Abhängig von der maximale Paketgröße im anderen Subnetz werden neue Fragmente erzeugt.
- *internet fragmentation*: Jeder Router zerlegt Pakete, sofern sie größer als die maximale Paketgröße sind. Erst beim Empfänger werden Fragmente eines Datagramms zusammengesetzt.



*Internet fragmentation* hat Vorteile, wenn einzelne Fragmente eines Datagramms verlorengehen. Nur diese Fragmente müssen nochmals übertragen werden.

## 2. Routing

Um Datagramme von einem Ziel zu einem Empfänger zu senden, muß die *network point of attachment* (NPA) Adresse des Empfängers bekannt sein. Liegt diese NPA Adresse in einem entfernten Subnetz, so muß die Adresse des nächsten Gateways auf dem Weg dorthin bekannt sein.

Zwei Ansätze gibt es für *routing* in einem Internet:

- **Zentrales Routingschema.**
- **Verteiltes Routingschema** mit **Routingtabellen** in jedem System.

Die IP-Routing Prozedur geht folgendermaßen vor:

- 1) IP-Zieladresse (NSAP) eines Datagramms lesen.
- 2) NPA Adresse des host/ES oder des nächsten Gateways anhand der Routingtabellen ermitteln.

Spezielle Protokolle beschreiben den Aufbau und die Wartung von Routingtabellen.

Für Routingzwecke wird das Internet als eine Menge von *autonomen Systemen* angesehen, die an das *core backbone network* angeschlossen sind.

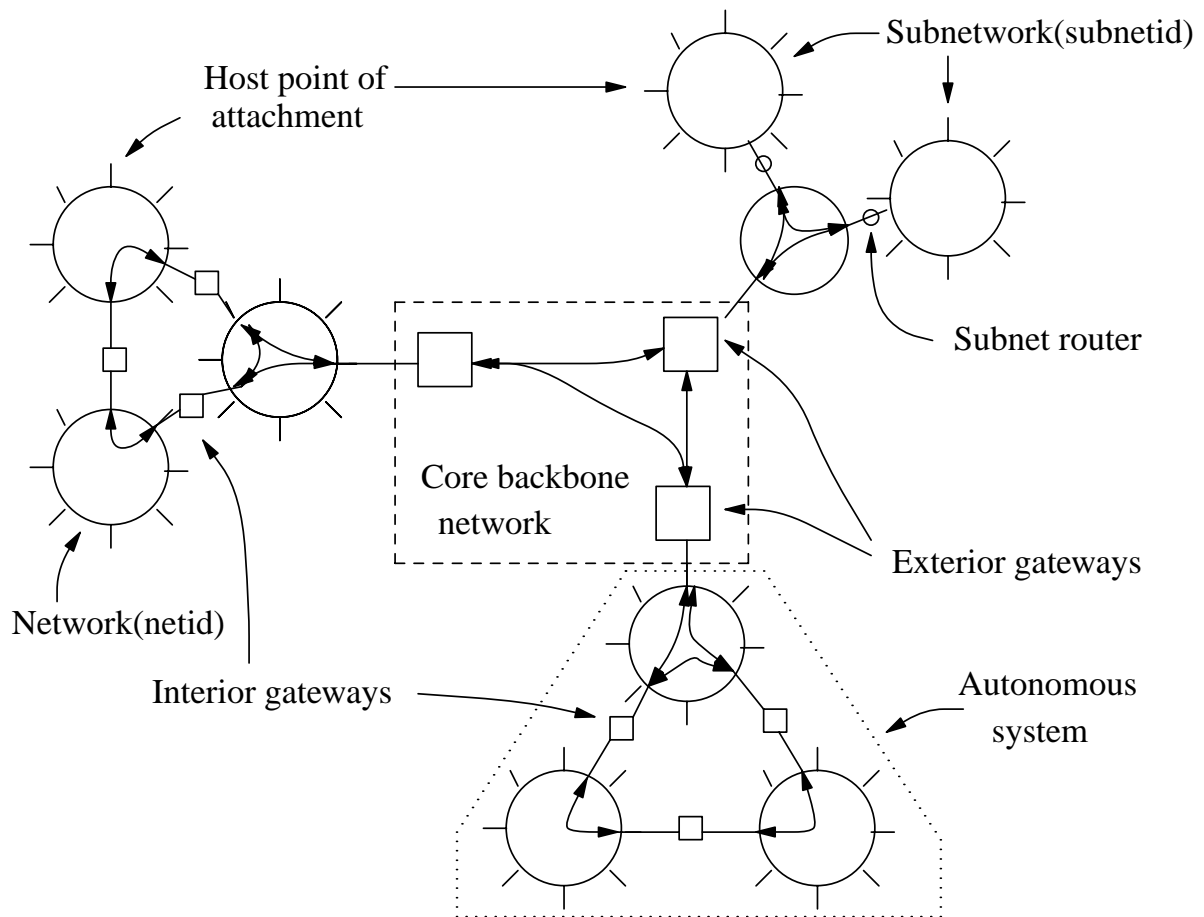
- Das **interior gateway protocol (IGP)** wird innerhalb eines autonomen Systems benutzt.
- Das **exterior gateway protocol (EGP)** beschreibt den Anschluß eines autonomen Systems an das *core backbone network*.

Das Internet aus einer großen Anzahl verschiedener autonomer Systeme besteht, existieren im Internet mehrere IGP's.

Das Internet EGP ist dagegen ein Internet-weiter Standard.



## Die allgemeine Internet Struktur und Terminologie:

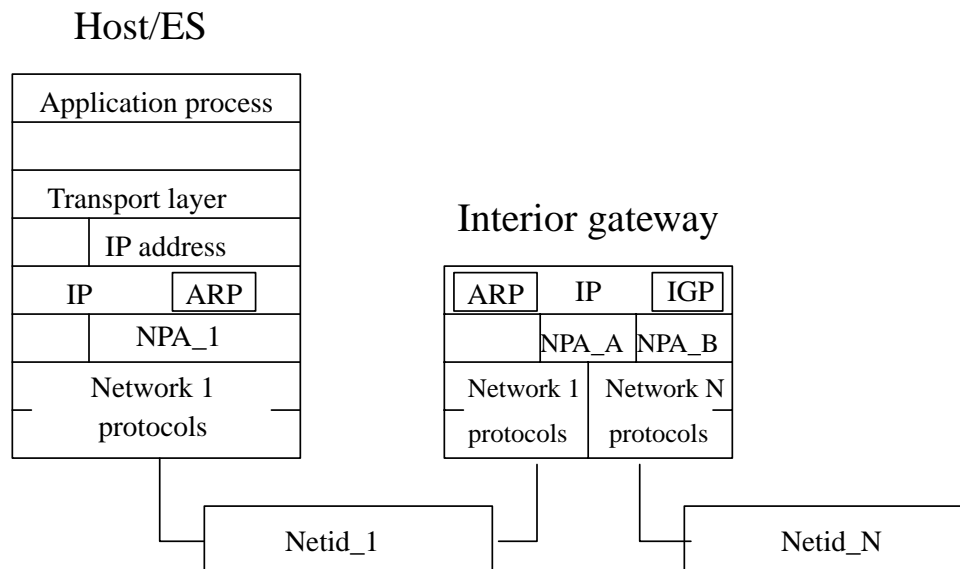


Um ein Überfließen von Routingtabellen zu verhindern und die Menge der zu übertragenden Routinginformationen einzuschränken, ist im Internet die Routinginformation hierarchisch organisiert:

- Hosts/ES besitzen genügend Information, um Datagramme an andere hosts/ES oder *interior gateways* zu senden.
- *Interior gateways* können Datagramme im selben *autonomous system* weiterleiten.
- *Exterior gateways* können Datagramme an andere *exterior gateways* oder an *interior gateways* im selben *autonomous system* weiterleiten.

## Adress resolution protocol

- *Interior gateway* muß Paar aus *NPA/IP*-Adressen für jeden host/ES im selben autonomen System kennen.
- Diese Information kennt jeder host/ES statisch, sie wird beim boot-Vorgang *broadcasted*.
- *Interior gateways* haben *local routing tables*.



- Das IP in jedem host/ES versucht ebenfalls Tabellen aller lokalen *hostid/NPA*-Paare aufzubauen.
- Bevor ein Datagramm versendet wird, wird versucht, die IP Adresse des Ziels durch die NPA Adresse zu ersetzen.
- Das SNDAP erhält die NPA Adresse.
- Eine **ARP request message** wird ausgesendet, wenn Paar IP/NPA Adresse lokal unbekannt ist.
- *Interior gateway* antwortet mit **ARP reply message**.

Besonderes Problem bei *diskless*-Nodes.

- **RARP** — Zuordnung von Hardware Adresse zu IP Adresse.

## Interior gateway protocols

Das weitverbreitetste Protokoll heißt **IP routing information protocol (RIP)**.

Einige Eigenarten von **RIP**:

- Verteiltes Routing Protokoll.
- Basiert auf dem Abstand zweier Gateways als *routing metric*.
- Durchmesser eines Netzes ist als Maximale Zahl von *hops* im Netz definiert.
- Unglücklicherweise ist  $16 = \infty$ .
- **RIP** ist in vielen Routern (Proteon, cisco, UB..) und in BSD Unix Systemen (routed) implementiert.

Problematisch an **RIP** ist, daß unterschiedliche Geschwindigkeiten auf einzelnen Links nicht berücksichtigt werden.

**RIP** ist vollständig dokumentiert in RFC-1058.

**Hello** ist ein anderes Routing Protokoll.

- Ebenso wie bei RIP werden Listen von Netzen und deren Metriken betrachtet.
- Metrik ist durch Verzögerung in Millisekunden gegeben.
- **Hello** ist günstig für Links mit verschiedenen Geschwindigkeiten.
- **Hello** kommt besser als RIP mit Überflutungssituationen klar.
- Diese Protokoll wurde früher auf dem NSFnet backbone für Routing verwendet.

Ein interessanter Effekt von **Hello** ist die Möglichkeit, zeitliche Bedingungen für IP Verbindungen einzuhalten.

**Hello** ist vollständig beschrieben in RFC-891.

## Exterior gateway protocol

Innerhalb seines autonomen Systems kommuniziert ein *exterior gateway* über das IGP.

→ Jedes *exterior gateway* kennt alle *netids* in “seinem” autonomen System und ihre Entfernung zum Gateway.

Gegenüber dem *core backbone network* wird das **exterior gateway protocol (EGP)** benutzt.

- Ein *exterior gateway* trägt die eindeutige Identität “seines” autonomen Systems.
- Eine **reachability table** enthält jene *exterior gateways*, mit denen über das *core backbone network* kommuniziert werden kann.
- Im BSD Unix wird das **EGP** durch das Programm *gated* implementiert.
- *gated* “spricht” neben EGP auch RIP und Hello, problematisch ist allerdings eine Umsetzung von Metriken.

Die Funktionalität des EGP umfaßt:

- Ermittlung eines Nachbarn.
- Test der Erreichbarkeit eines Nachbarn.
- Aktualisierung von Routinginformation.

EGP Nachrichten und ihre Bedeutung:

Function	EGP message	Bedeutung
Neighbour acquisition	Acquisition request	Aufforderung an ein Gateway, es möge Nachbar werden.
	Acquisition confirm	Gateway stimmt zu.
	Acquisition refuse	Gateway verweigert Nachbar zu werden.
	Cease request	Aufforderung an ein Gateway, Nachbarschaft zu beenden.
	Cease confirm	Abbruch der Nachbarschaft zugestimmt.
Neighbour reachability	Hello	Aufforderung an einen Nachbarn, eine etablierte Verbindung zu bestätigen.
	I-heard-you	Verbindung bestätigt.
Routing update	Poll request	Aufforderung, Erreichbarkeitsinformation für ein Netz zu aktualisieren.
	Routing update	Erreichbarkeitsinformation für ein Netzwerk.
Error response	Error	Antwort auf jede inkorrekte Nachricht.



### 3. Fehlermeldungen

Das **internet control message protocol (ICMP)** wird von hosts/ES und Gateways benutzt.

Das **ICMP** bietet folgende Funktionalität:

- Fehlermeldungen.
- Erreichbarkeitstest.
- Überflutungskontrolle.
- Erkennen von Routenänderungen.
- Durchsatzmessungen.
- Subnetzadressierung.

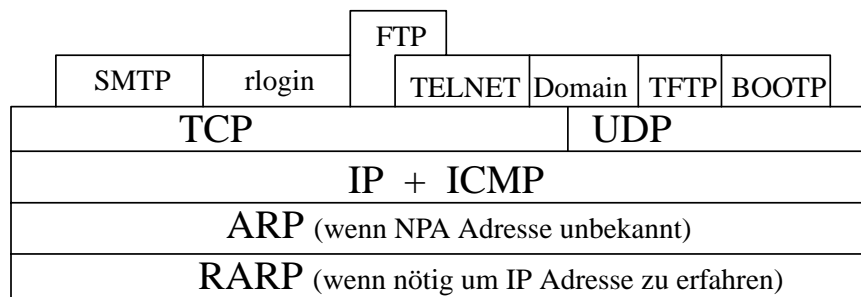
## ICMP Nachrichten und ihre Benutzung:

Function	ICMP message	Benutzung
Error reporting	Destination unreachable	Ein Datagramm wurde aus dem angegebenen Grund weggeworfen.
	Time exceed	<i>Time-to-live</i> Parameter hat Null erreicht.
	Parameter error	Ein Parameter im Kopf eines Datagramms ist falsch.
Reachability testing	Echo request/reply	Wird benutzt, um Erreichbarkeit eines Hosts oder Gateways zu testen.
Congestion control	Source quench	Benutzt, um einen Sender zu geringerer Senderate zu bewegen.
Route exchange	Redirect	Benutzt von Gateways, um einen Host aufzufordern, eine Route über ein anderes Gateway für sein Ziel zu verwenden.
Performance measuring	Timestamp request/reply	Benutzt, um Übertragungsverzögerung zwischen zwei Hosts zu ermitteln.
Submit addressing	Adress mask request/reply	Benutzt von Hosts, um die mit einem Subnetz assoziierte Adreßmaske zu erfahren.

## Protokolle oberhalb von IP

Mit dem Internet Protokoll sind auf Ebene 4 des OSI Modells zwei Transportprotokolle verbunden: **TCP** und **UDP**.

Diese Protokolle benutzen *IP* und *ICMP* folgendermaßen:



## Literatur

Douglas Comer

*Internetworking with TCP/IP, principles, protocols and architecture*

Prentice-Hall International Editions, 1988.

Fred Halsall

*Data Communications, Computer Networks and Open Systems*

Addison Wesley, 3rd. Edition, 1992.

E. Krol, University of Illinois, Urbana

*The Hitchhikers Guide to the Internet*

RFC-1118.

D.Clark et al., MIT

*Towards the Future Internet Architecture*

RFC-1287.