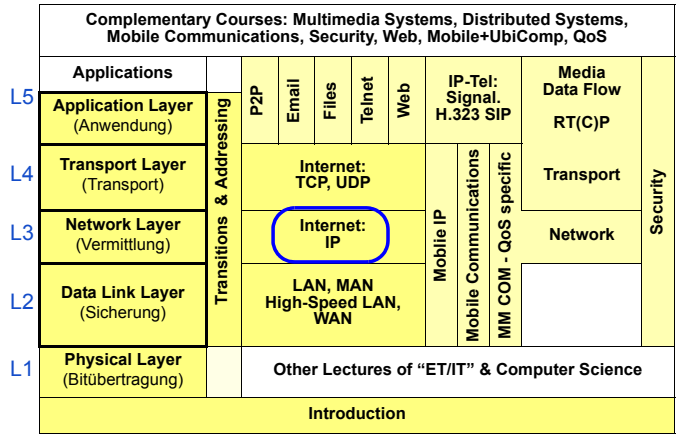
Communication Systems
Network Layer:
Internet Protocols

Prof. Dr.-Ing. **Lars Wolf**

TU Braunschweig
Institut für Betriebssysteme und Rechnerverbund

Mühlenpfordtstraße 23, 38106 Braunschweig, Germany
Email: wolf@ibr.cs.tu-bs.de

Scope



Overview

1. History and Architecture

ARPANET

- **initiated and financed by ARPA (Advanced Research Projects Agency of the U.S. Department of Defense (DoD))**
- **objective:**
  - originally: network to survive nuclear war
  - later: connecting scientific and military institutions
- **1969: experimental network with 4 nodes, followed by rapid growth, BBN first contractor**
- **development of the INTERNET**
  - standardized protocols for communication between networks: TCP/IP (1983)
  - linking military subnetworks (MILNET, MINET)
  - linking satellite networks (SATNET, WIDEBAND)
  - linking the LANs of the universities
- **fast spreading of TCP/IP technology as a part of UNIX → ARPANET growing rapidly (1987 : 15% per month)**
- **1987: 20.000 computers, more than 100.000 users**
- **1990: ARPANET replaced, MILNET still exists**
- **services: E-mail, file transfer, remote login, WWW. . .**

The Internet and its Tasks

Internet (Internet Society)

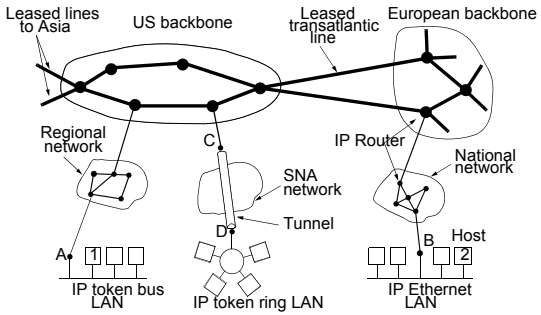
- **mid-80s** **a multiple of networks was designated as the "Internet"**
- **Jan. 1992:** **founding of the (actual) Internet Society**
  objective: **to spread the use of the Internet (protocols and services)**
- **IAB: Internet Architecture Board**
  - founded in 1983 to involve researchers in the ARPANET
  - today it is the supreme Internet board
- **IAB oversees/nominates**
  - IETF (Internet Engineering Taskforce)
    - divided into approx. 70 working groups
    - actual governing board
  - IRTF (Internet Research Taskforce)
- **RFCs (Requests for Comments)**
  - recommendations, today numbering approx. 3300

Tasks in the INTERNET

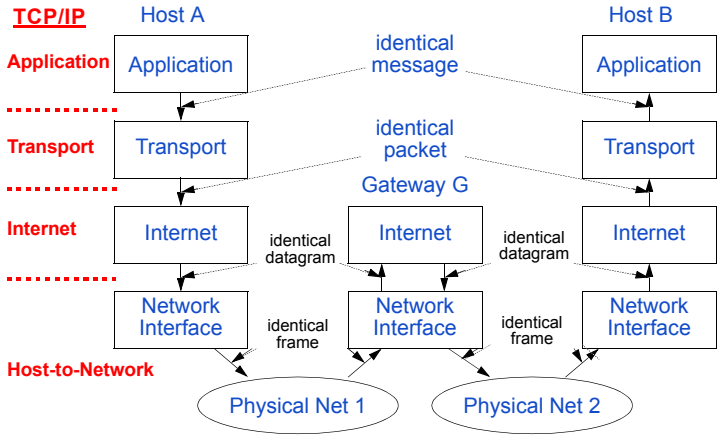
- **connect different networks over gateways**
- **definition of**
  - protocols that work on all subnetworks
  - standardized addressing pattern for a very large network
  - global routing architecture

Subnets in the INTERNET

- **Ethernet LANs**
  - mainly large Campus networks
- **LAN Rings**
  - mainly smaller/ experimental networks
- **Arpanet**
  - network with specific protocols, partially connected over leased lines
- **NSFNet (National Science Foundation Network)**
  - backbone consisting of leased high-speed lines
  - connecting the NSF supercomputers with each other and to regional networks and campus networks
  - since 1995 AOL, now a multitude of backbones in USA
- **CSNET (X.25 NET)**
  - public packet relay network by X.25

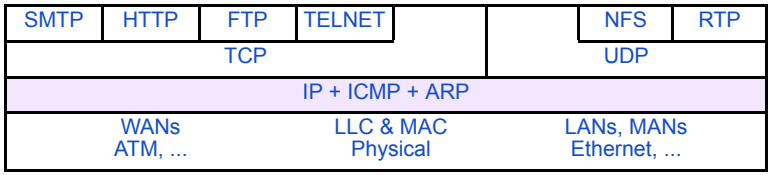


Internet Model



- **ISO-OSI presentation and session do not exist**
- **data link and physical layer combined**

Well-Known Internet Protocols



ARP	=	**Address Resolution Protocol**
FTP	=	**File Transfer Protocol**
HTTP	=	**Hypertext Transfer Protocol**
IP	=	**Internet Protocol**
ICMP	=	**Internet Control Message Protocol**
LLC	=	**Logical Link Control**
MAC	=	**Media Access Control**
NFS	=	**Network File System**
SMTP	=	**Simple Mail Transfer Protocol**
TELNET	=	**Remote Login Protocol**
TCP	=	**Transmission Control Protocol**
UDP	=	**User Datagram Protocol**
RTP	=	**Real-Time Transport Protocol**

Internet Architecture

No formal architecture

No unchangeable principles:

> ***The principle of constant change is perhaps the only principle of the Internet that should survive indefinitely. [RFC 1958, Architectural Principles of the Internet, June 1996]***

The Internet approach in very general terms (from RFC 1958):

- **the goal is connectivity**
- **the tool is the Internet Protocol**
- **the intelligence is end-to-end rather than hidden in the network**

Some Internet Guidelines

- **make sure it works**
  - do not finalize design / standard before multiple prototypes are interoperable
- **keep it simple (stupid ... KISS)**
  - when in doubt, use simplest solution: leave unnecessary features out
- **make clear choices**
  - if there are alternatives for the same thing, choose one (avoid too many options)
  - re-use good solutions if applicable but avoid duplication of functionality
- **exploit modularity**
  - like layers in protocol stacks
- **expect heterogeneity**
  - hardware, applications, etc.
- **avoid static options and parameters**
  - negotiations among sender and receiver
- **look for good design, but not necessarily perfect:**
  - adopt almost complete solution now, don't wait until perfect solution be found
- **be strict when sending and tolerant when receiving**
  - follow specifications precisely when sending, but tolerate faulty input from net
- **consider scalability**
  - many nodes per site and many millions of sites
- **consider cost and performance**

2. Internet Protocol (IP)

IP basics

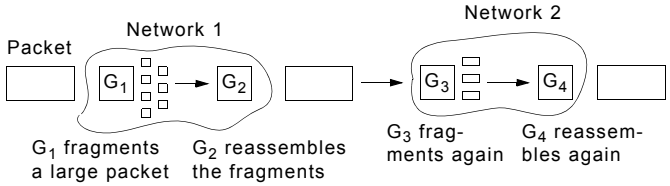
- **defined for the first time in 1981**
  - J. Postel
  - RFC 791, September 1981
- **packet length**
  - in theory: up to 64 kBytes
  - in real life: approx. 1500 Bytes

connectionless service (datagram)

- **provide best-efforts (not guaranteed) way to transport datagrams**
- **from source to destination**
- **without regard whether these machines are on the same network or whether there are other networks in between**

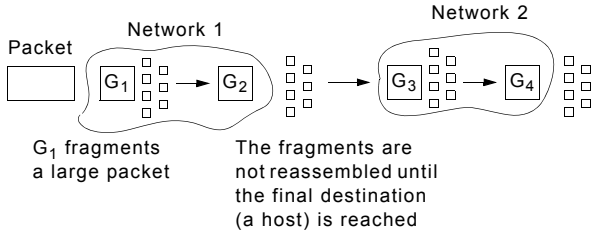
2.1 IP: Segmentation/Reassembling

1. Transparent segmentation

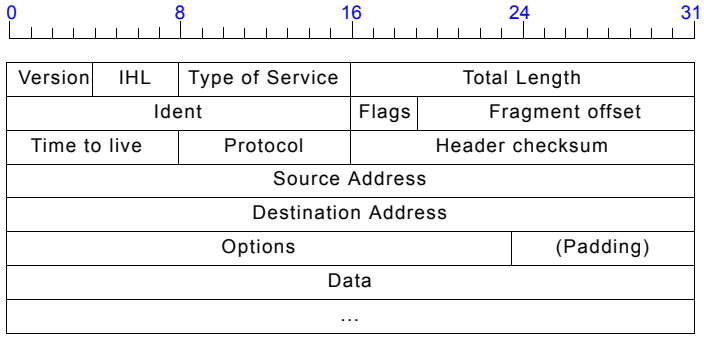


2. Non-transparent segmentation

USED IN **IP**:



2.2 IPv4 Datagram Format



Comment

- **transmission in "Big Endian" order (from left to right, highest version bit first)**
  - Big Endian
    - e.g. IBM PowerPC and SUN SPARC computers
  - Little Endian
    - e.g. Intel Pentium computer
    - conversion during receiving and sending

IPv4 Datagram Format (2)

Version

- **protocol version: presently IP v.4**
- **IP v.5 (realtime data transfer): ST-2**
- **IP v.6: successor to IPv.4**

Header Length (IHL) (in 32 bit words)

- at least 5 words with 32 bit each = 20 bytes
- at most 15 words with 32 bit each = 60 bytes

Type of Service

- **simple QoS: a combination of reliability and delay**
  - precedence (3 bit):
    - priority 0 (normal) ...7 (network control)
    - influences the queuing scheme (and not routing)
  - D (1 bit): Delay, e.g. no satellite transmission
  - T (1 bit): Throughput, e.g. no telephone line
  - R (1 bit): Reliability, e.g. no radio channels
  - C (1 bit): low Cost, defined later on
  - 1 bit unused
    comment: C & D activated: e.g. invalid
- **in practical use: ignored by routers**
- **redefined for Differentiated Services (DiffServ)**

IPv4 Datagram Format (3)

Total length

- **full length including the data**
- **stated in bytes**
- **all hosts must be prepared to accept datagrams of up to 576 bytes**
- **recommendation:**
  - send larger datagrams only if assured that destination can handle these
- **max. 65.535 byte, often approximately 1500 byte**

Identification

- **necessary for destination to determine datagram a fragment belongs to**
- **all fragments of a datagram contain same identification value**

Flags

- **1 bit unused**
- **DF (1 bit): don't fragment**
  - packets may have a length of up to 576 byte
- **MF (1 bit): more fragments**
  - last fragment marked 0

IPv4 Datagram Format (4)

Fragment offset

- **offset of this fragment, i.e. the position within a datagram**
- **stated in multiples of 8 bytes (elementary frament unit)**
- **13 bits ⇒ max. 8192 fragments / datagram ⇒ max. datagram len. 65536 bytes**

Time To Live (TTL)

- **life cycle in seconds, max. 255 sec**
- **when 0: drop packet, feedback to sender**
- **must be decremented per hop, in practical use: counts hops (not seconds)**

Protocol : type of the higher level protocol for transmission

No.	Abbreviation	Protocol
0		reserved
1	ICMP	Internet Control Message
2	IGMP	Internet Group Management
3	GGP	Gateway to Gateway
4	IP	IP in IP
5	ST	Stream
6	TCP	Transmission Control
...	...	...

IPv4 Datagram Format (5)

Header Checksum
- **to detect errors generated by bad memory words inside an IS**
- **observed each time when datagram is received (both in IS and ES) if necessary datagram is dropped**
- **certain summation of the header words**
  - addition of all 16-bit halfwords in one's complement arithmetic and use one's complement of result (assume this field as zero upon arrival)
- **must be recomputed at each hop (due to change in Time-to-Live field)**

Source Address
- **sender's IP address**

Destination Address
- **receiver's IP address**

IPv4 Datagram Format (6)

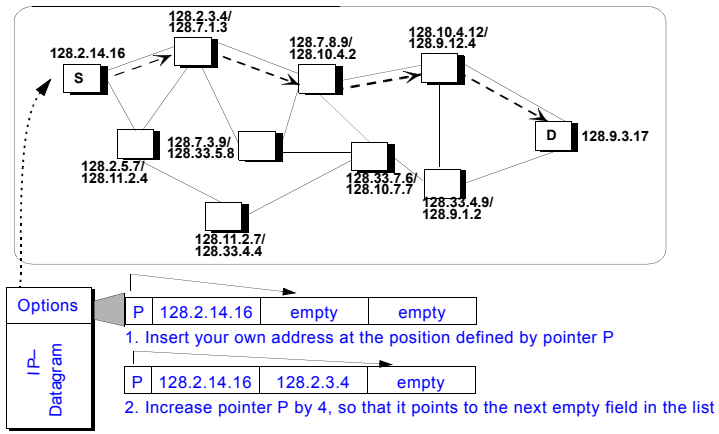
Options
- **options for routing, testing and debugging**
- **conceptual design: as an enhancement for future versions**
- **variable length: each begins with 1-byte identification code**
- **included (e.g. in 1996):**
  - security: security degree exclusion of routes, but ignored in practice
  - strict source routing: the exact route is specified
  - loose source routing: part of the route is given, i.e., list of routers to visit
  - record route: store IP addresses of routers, but nowadays headers are too small for this purpose
  - timestamp: like record route, but also timestamp added at router

Padding
- **fill up to the word limit**

Data
- **field for user data**

IPv4 Datagram Format (7)

Example
- **option: Record Route**

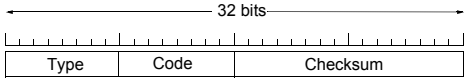


3. Internet Control Message Protocol (ICMP)

History
- **J. Postel**
- **RFC 792, Sept. 1981**

are sent as IP packets
- **i.e., the first 32 bits of the IP data field are ICMP headers**

Header structure

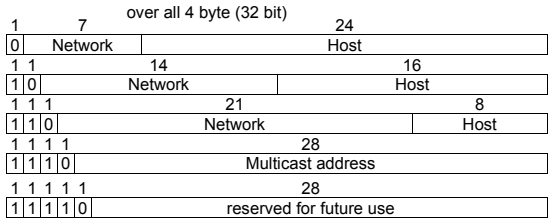


Type
- **16 types, a.o.**
  - destination or port or protocol unreachable
  - time exceeded
  - fragmentation necessary but DF (don't fragment) DF is set
  - redirect, packet seems to be routed wrong
  - echo request and echo reply (for Ping program)
  - source quench (previously used for congestion control: Choke packet)

Code
- **states cause if type is "destination unreachable"**
  - e. g. net, host, protocol, port unreachable or
  - fragmentation needed, source route failed

4. Internet Addresses and Internet Subnetworks

Global addressing concept for ES (and IS) in the Internet
- **unique 32 bit address with net-ID (subnetwork-Id), ES-Id**
- **i.e., each network interface (not ES) has its own unique address**
- **5 classes**

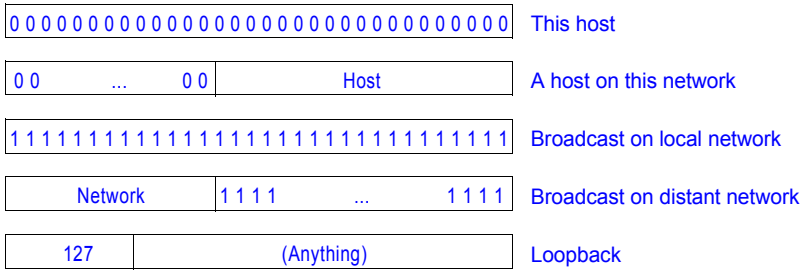


ICANN (Internet Corporation for Assigned Numbers and Names)
- **manages network numbers**
- **delegates parts of the address space to regional authorities**

Network addresses typically written in dotted decimal notation
- **e.g., 134.169.34.18**
- **lowest 0.0.0.0 (0 means this host or network)**
- **highest 255.255.255.255 (broadcast on local network)**

Internet Addresses and Internet Subnetworks (2)

Special IP addresses



Internet Addresses and Internet Subnetworks (3)

Networks grow and should be somehow structured
- **several networks instead of one preferable**
- **but getting several address areas is hard**
  - since address space is limited
  - e.g., university may have started with class B address, doesn't get second one

Problem:
- **class A, B, C refer to one network, not collection of LANs**

⇒ **allow a network to be split into several parts**
- for internal use
- still look like single network to outside world

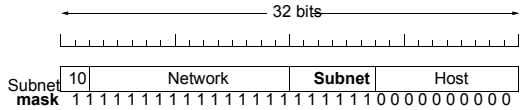
⇒ **Provide for subnetworks**

Internet Addresses and Internet Subnetworks (4)

Subnets:
- **e.g., Ethernet-based LAN**

Idea:
- **local decision for subdividing host share into subnetwork portion and end system portion**

- **example: class B address: max. 63 subnetworks**



Use subnet mask to indicate split between network + subnet and host part

routing with 3 levels of hierarchy
- **algorithm in router (by masking bits: AND between address and subnet mask):**
  - packet to another network (yes, then to this router)
  - packet to local ES (yes, then deliver packet)
  - packet to other subnetwork (yes, then reroute to appropriate router)

CIDR: Classless InterDomain Routing

IPs growth leads to lack of addresses

- **in principle many addresses due to 32-bit address space**
- **but inefficient allocation due to class-based organization**
  - class A network with 16 million addresses too big for most cases
  - class C network with 256 addresses is too small
  - most organizations are interested in class B network, but there are only 16384
    - (in reality, class B too large for many organizations)

large number of networks leads to large routing tables

⇒ **Introduction of CIDR (Classless InterDomain Routing) (RFC1519)**

- **allocate IP addresses in variable-sized blocks (without regard to classes)**
- **e.g., request for 2000 addresses leads to assignment of 2048 address block starting on 2048 byte boundary**

dropping classes makes forwarding more complicated

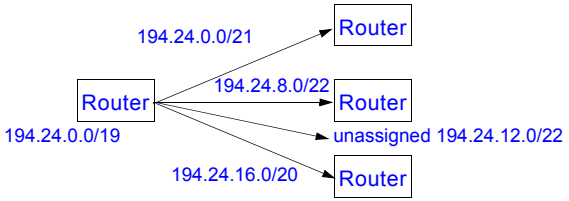
CIDR: Classless InterDomain Routing (2)

Search for longest matching prefix

- **if several entries with different subnet mask length may match then use the one with the longest mask**
- **i.e., AND operation for address and mask must be done for each table entry**

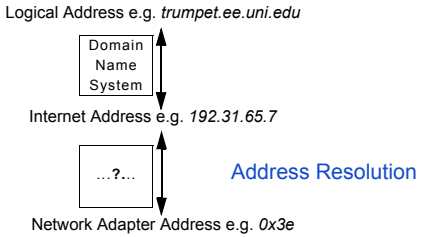
E.g., potentially several 'class C' networks can be characterized by one prefix

Entries may be aggregated to reduce routing tables



5. Address Resolution

Addressing levels



Host identification and routing specification within a subnetwork

- **are based on the (local) physical network addresses of the end systems**
- **(e.g. station address of the adapter card)**

Problem:

- **INTERNET address (32 bit) must be mapped onto the physical network address, ususally 48 bit (ADDRESS RESOLUTION)**

Address Resolution: Methods

Address resolution in

- **source ES, if destination ES is local (direct routing)**
- **Gateway, if destination ES is not local**

Solutions:

1. **Direct HOMOGENEOUS ADDRESSING**
   - if the physical address can be dialed by the user, then the dial-up is:
     - physical address = Hostid of the INTERNET address
2. **If the physical address is pre-defined or if it has to have a different format, one of the following has to be used**
   - a mapping table from the configuration data base (IPaddr → HWaddr),
     - e.g. in the Gateway,
     - may become maintenance nightmare
   - the Address Resolution Protocol (ARP)
     - mainly applied in LANs with broadcasting facility

Address Resolution Protocol (ARP)

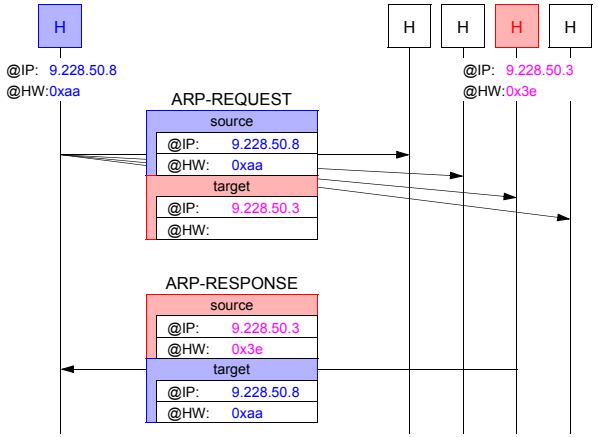
Process:

1. **broadcast ARP request datagram on LAN**
   - including receiver's INTERNET address (desired value)
   - sender's physical (HW) and INTERNET address (IP)
2. **every machine on LAN receives this request and checks address**
3. **reply by sending ARP response datagram**
   - machine which has requested address responses
   - including the phyiscal address
4. **enter the pair (I,P) into buffer for future requests**

Refinement:

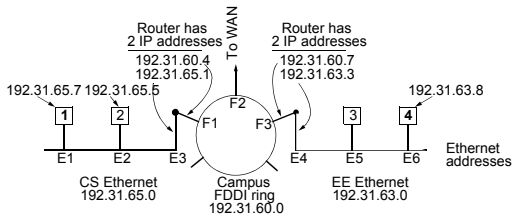
- **the receiver of the ARP request stores the sender's (I,P) pair in its cache**
- **send own table during the boot process (but may be too old)**
- **entries in ARP cache should time out after some time (few minutes)**

Address Resolution Protocol (ARP) (2)



Address Resolution Protocol (ARP) (2)

End system not directly available by broadcast



Example: ES 1 to ES 4

- **ARP would not receive a response**
  - Ethernet Broadcast is not rerouted over a router

Solution: proxy ARP

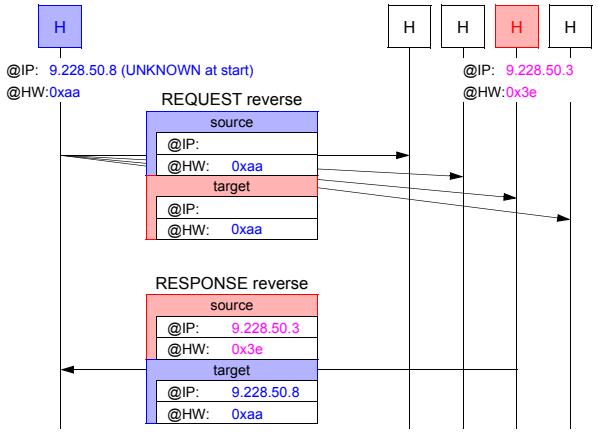
- **the local router knows all remote networks with their respective routers**
  - responds to local ARP
- **local ES 1 sends data for ES 4 always to the local router, this router forwards the data (by interpreting the IP address contained in the data)**

Solution: remote network address is known

- **local ES 1 sends data to the appropriate remote router**
- **local router forwards packets**

Reverse Address Resolution Protocol (RARP)

RFC 903: Retrieve Internet address from knowledge of hardware address



- **RARP server responds**
- **RARP server has to be available on the LAN**

Application: diskless workstation boots over the network

DHCP: Dynamic Host Configuration Protocol

DHCP has largely replaced RARP (and BOOTP)
- **extends functionality**

DHCP
- **simplifies installation and configuration of end systems**
- **allows for manual and automatic IP address assignment**
- **may provide additional configuration information (DNS server, netmask, default router, etc.)**

DHCP server is used for assignment
- **request can be relayed by DHCP relay agent, if server on other LAN**

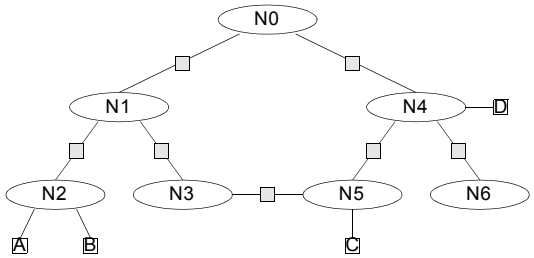
Client broadcasts DHCP DISCOVER packet
- **server answers**

Address is assigned for limited time only
- **before the 'lease' expires, client must renew it**
- **allows to reclaim addresses of disappearing hosts**

6. IP Route Determination: Internal and External

1. **Direct Routing/Interior Protocols:**

   Both source and destination ES are located in the same subnetwork
   - source ES sends datagram to the destination ES
   - identification done by the local address → mapping
   - routing is completely defined by the subnetwork routing algorithm

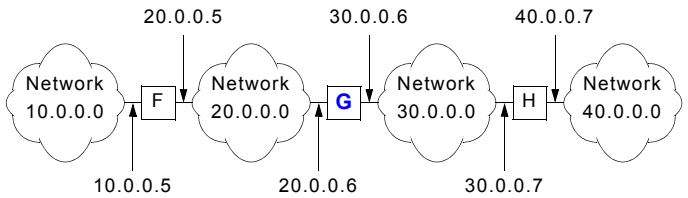


2. **Indirect Routing/Exterior Protocols:**

   Source and destination ES are located on different networks
   - source ES sends datagram to the next router
   - each router determines the next router on the path to the destination ES
   - routing decision is based only on the netid part of the Internet address, i.e. hostid is not queried

IP Routing

Routing tables of the Gateways



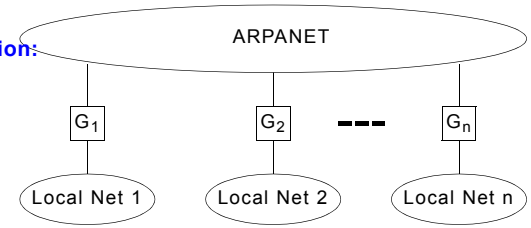
	TO REACH HOSTS ON NETWORK	ROUTE TO THIS ADDRESS
G:	20.0.0.0	DELIVER DIRECT
	30.0.0.0	DELIVER DIRECT
	10.0.0.0	20.0.0.5
	40.0.0.0	30.0.0.7

Gateways may have incomplete information

⇒ **default paths**

6.1 IP Routing: Gateway-to-Gateway Protocol (GGP)

Original Implementation:

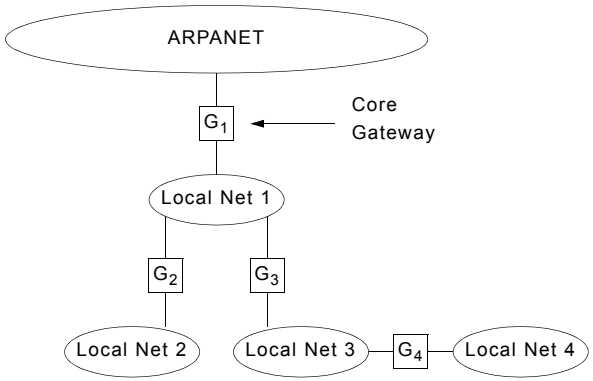


Core Gateways
- **connect LANs to the backbone, know the routes to all networks**
- **exchange routing information with each other**
- **Gateway-to-Gateway Protocol (GGP):**
  - distributed routing definition (group **"DISTANCE-VECTOR-PROCEDURE"**)
  - metrics: simply by distance

Problems: particularly poor scalability
- **several backbones**
- **not all networks are connected directly to the backbone**
- **all Gateways communicate with each other**

IP Routing: Internet Enhancements

Hidden networks

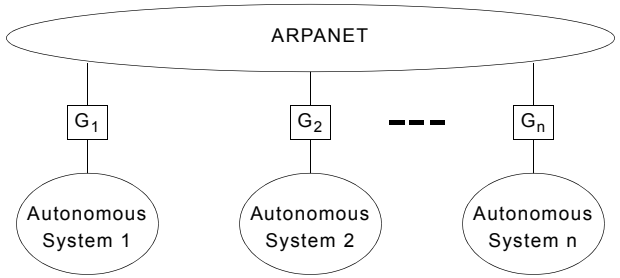


Core Gateways have to be informed about hidden networks

⇒ **Autonomous systems**

IP Routing Definition: Autonomous Systems

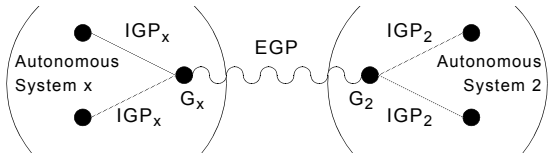
Autonomous systems



Autonomous system (AS) = administrative entity
- **collects routing information on networks in the AS**
- **defines gateways,**
  - that transmit routing information to other AS (Exterior Gateways)

Exterior Gateway (actually router)
- **transmits only information about network accessibility to its AS**
- **reason: each AS can monitor exactly, to whom the information about accessibility is given to**

6.2 Interior Gateway Protocol



In general
- **individual solutions possible**

Presently preferred procedures
- (old) Routing Information Protocol (RIP)
- Open Shortest Path First (OSPF)
- Interior Gateway Routing Protocol (IGRP) and Enhanced IGRP (EIGRP)
  - by Cisco
  - based on ISO connectionless L3 protocol

other variant
- **HELLO by Dave Mills**
  - distributed routing algorithm
  - distance: Delay
  - requires synchronized clocks

Routing Information Protocol (RIP)

Background (regarding the originally used protocol)
- **developed as a part of Berkeley UNIX**
- **since 1988, RIP Version 1, RFC 1058**

Principle
- **distributed routing algorithm: Distance-Vector-Procedure**

i.e.
- **IS periodically sends a list containing ESTIMATED DISTANCES to each destination to its neighbors**
  - distance:
    - number of hops: 0 .. 15 (15 corresponds to infinite)
  - periodical:
    - every 30 sec; after 180 sek without packet → distance infinite

RIP Version 2
- **G. Malkin, RFC 1387, 1388 and 1389 (RIP-MIB)**
  - uses multicast if necessary to distribute data
  - not broadcast (so that all ES also receive this)
- **networks without broadcast or multicast (ISDN, ATM)**
  - "triggered" updates
  - to be sent only if the routing table changes

Open Shortest Path First (OSPF)

Background: since 1990 Internet Standard, RFC 1247

- **transition from distance-vector to link-state-procedures**

Principle (link-state-procedure)

- **IS measures "distance" to the immediately adjacent IS, distributes the information, calculates the optimal route**
  1. **determine the address of adjacent IS**
  2. **measure the "distance" (delay, ..) to adjacent IS**
     - OSPF permits different metrics
     - selection per packet possible (RFC 1349)

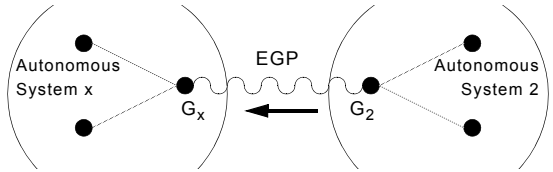
OSPF No.	Meaning
0	normal service
2	minimize financial costs
4	maximize reliability
8	maximize throughput
16	minimize delay

  3. **prepare local link-state information as a packet**
  4. **distribute information to all IS**
  5. **compute route from the information of all IS, e.g., with Dijkstra's "shortest path first" algorithm → name "Open Shortest Path First"**

6.3 Exterior Gateway Protocol

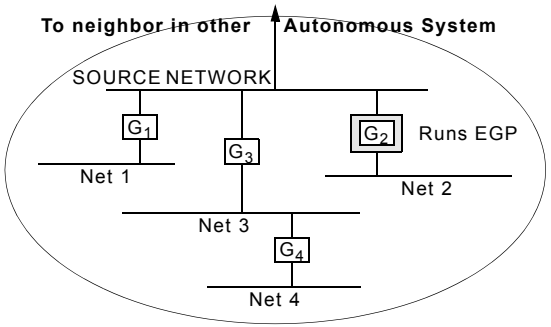
Example:

- **G2 sends to EGP neighbors, e.g. G_x routing update message**
- **(G1 : Net1; G3 : Net3, Net4; G2 : Net2)**



Example:

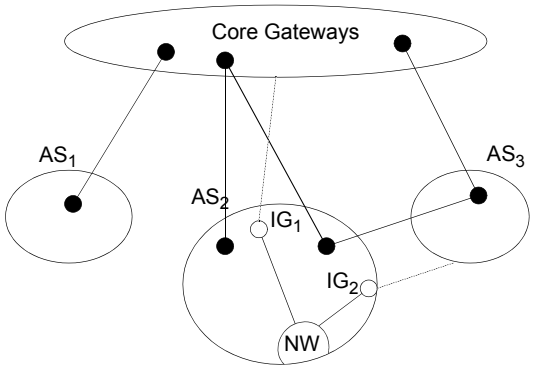
- **of autonomous system 2**
- **Exterior Gateways**



Exterior Gateway Protocol: Circumstances

Requirements, basic conditions

- **political**
- **economical**
- **security-related**



- **examples**
  - avoid certain autonomous systems
  - avoid certain countries
  - stay within one country (before going via foreign country)
  - data of company A should not pass through company B

- **exchange information on accessibility**
  - including at least one Core Gateway
  - possibly with other AS

Border Gateway Protocol (BGP)

Internet Exterior Gateway Protocol (RFC 1654)

Configurations

- **possibly several Exterior Gateways per AS, variations:**
  - branch (topology):
    - all of the external traffic is routed over this single, external router
  - networks with various connections
    - linked to many end systems
    - can pass on traffic if necessary
  - transit networks
    - networks with increased capacity and
    - often linked to many AS

Algorithm

- **Distance-Vector-Procedure**
  - IS periodically sends a list containing the estimated distances to each destination to its neighbours
- **BGP specifically**
  - IS periodically sends a list containing the estimated distances to each destination AND THE EXACT PATH to each destination to its neighbours
  - with this, the procedure can be optimized
  - but it is based on a limited number of routers

6.4 Example: IP Router

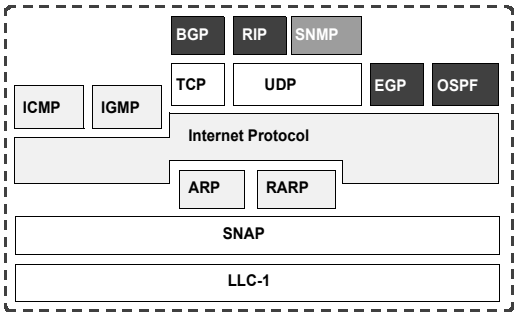
Network layer protocols
- IP (Internet Protocol)
- ARP (Address Resolution Protocol),
- RARP (Reverse ARP)
- ICMP (Internet Control Message Protocol)
- IGMP (Internet Group Management Protocol)

Routing protocols
- RIP (Routing Information Protocol)
- BGP (Border Gateway Protocol)
- EGP (Exterior Gateway Protocol)
- OSPF (Open Shortest Path First)

Network management protocols
- SNMP (Simple Network Management Protocol)

Transport protocols
- UDP (User Datagram Protocol)
- TCP (Transmission Control Protocol)

and
- SNAP (Subnet Access Point)
- LLC (Logical Link Control)



7. Internet Multicast

Multicast
- means to create trees, to address them, to modifiy them ...

IP Multicast Model
- Shared Tree
  - tree may be used by several senders
- Source Tree
  - tree is used by exactly one sender

Properties / Fields of Activity /Topics
- group addressing
- routing

and (more transport than network layer issues)
- reliable multicast
  - temporally limited and error free
- and in doing so
  - the return of ACK/NACK messages
    - but should not cause congestion of the system

Internet Multicast: Concepts

Virtual Overlay Network
- isolated solutions capable of multicasting
- connected worldwide through several tunnels
- logical tree structure

Data transport by UDP

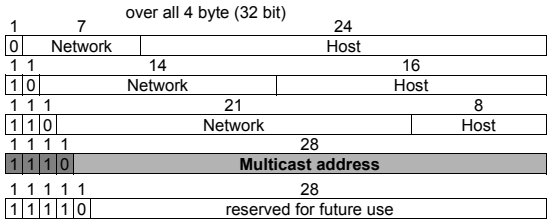
Dynamic, anonymous group model:
- no restrictions regarding the participants (location/number)
- dynamic group membership
- one host may be a member of several groups at the same time
- sender does not have to be a member of the group
- no restrictions regarding the groups's duration

Further information

`http://www.mbone.de/`

Internet Multicast: Addressing

Address class D



- 28 bit, i.e. > 250 Mio. groups
- data transmission to all group members (but unreliable)

Types of group addresses
- permanent
  - e.g. all ES and IS on one LAN,
  - all IS (router) on one LAN, ...
- temporary

Internet Group Management Protocol (IGMP)
- RFC 1112
- dynamic definition of group memberships

8. IP Version 6 (IPv6)

Motivation / main problem:

- addressing (presently 32 bit) and
- many other shortcomings in IP (QoS, mobility, ..)

Background & Status

- 1990: Call for Proposals
- 1992: 21 variants, with 7 possible candidates
- 1993: combination of 2 candidates: S. Deering and Francis (Xerox, Palo Alto)
- result: RFC 1883-87 protocol, addressing, ICMP, RFC 1825-29, newer ones appeared later (RFC2460-2466)
- since 2000: possibility to expand but still debate about its future, when it really will appear, ...

IPv6 Objectives

support billions of ES
- longer addresses

reduce routing tables

simplify protocol processing
- simplified header

increase security
- security means integrated

support realtime data traffic (quality of service)
- flow label, traffic class

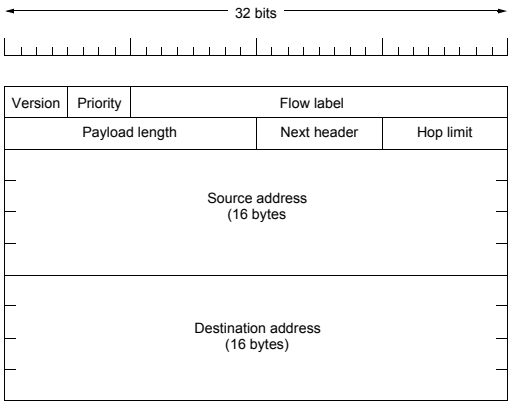
multicasting

support mobility (roaming)

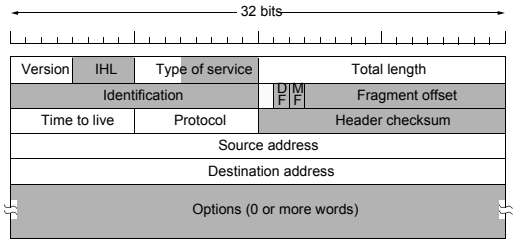
open for change (future)
- extension headers

coexistence with existing protocols

IPv6 Header

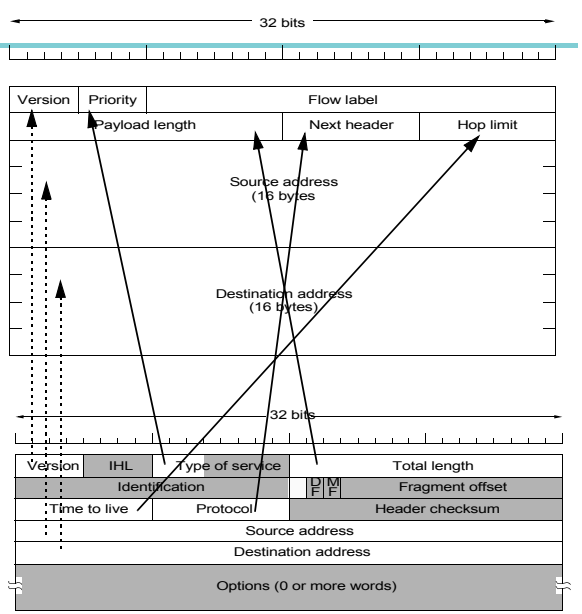


IPv4 Header



Shadowed: differences or not existing in IPv6

IPv6 vs. IPv4



IPv6 Header Fields

General comments:

- **header with fixed length**
- **64 bit alignment (IPv4: 32 bit)**

Version

Version	Description
0 -3	not in use
4	Internet Protocol, IP (presently used)
5	Stream Protocol, ST
6	**IPv6**
7	IPv77, TP/IX CATNIP
8	Pip
9	TUBA
10-15	not in use

Header length: (eliminated from v.4)

- **efficiency during processing**

IPv6 Header Fields (3)

TYPE OF Service:

- **Precedence replaced by traffic class**
- **D T R C-Bits (QoS) eliminated/replaced by "Flow Label"**

Flow Label

- **Definition may still change (experimental)**
- **Flow = Tupel (source ID, dest ID, No.)**
- **pre-defined**
- **handling defined by external auxiliary protocol**

~~Total Length~~ → Payload Length

- **length including the data (but without the 40 byte header)**
  - actually a maximum of 65.535 byte (plus 40 byte header)
- **possibly extension via "Jumbogram" (but then no fragmentation)**
  - a maximum of 65.535 byte may not be enough for a major data transmission

IPv6 Header Fields (4)

~~IDENT ification~~, ~~FLAGS~~, ~~FRAGMENT OFFSET~~

- **minimum packet size of IPv6 increased (from 576 to 1280)**
- **if still too large packet, then error message instead of fragmentation**
  - L4 should then take over this task and
  - transfer the PDU with the appropriate size to L3

~~Protocol~~ → Next Header

- **contains T4 protocol identification**
- **options (presently):**

Extension Header	Description
Hop-by-hop options	Miscellaneous information for routers
Routing	Full or partial route to follow
Fragmentation	Management of datagram fragments
Authentication	Verification of the sender's identity
Encrypted security payload	Information about the encrypted contents
Destination options	Additional information for the destination

~~TIME to live~~ = Hop limit

- **life cycle in number of hops, max. 255**
  this may not be sufficient, presently usually approx. 32 hops

IPv6 Header Fields (5)

~~**HEADER CHECKSUM**~~
- **L2 and L4 have sufficient mechanisms**
- **communication channels better nowadays, at the expense of the performance**

Source and Destination ADDRESS
- **32 bit → 128 bit**
- **128 bit addresses: 2^{128} different addresses, approx. $3*10^{38}$, approx. $7*10^{23}$ IP addresses per m^2 on earth (land and water)**

~~**OPTIONS**~~

Addressing: Anycast

Definition
- **previously**
  - unicast, broadcast and multicast
- **now (new)**
  - anycast
- **a system within a pre-defined group is to be accessed**

Application
- **search for the nearest web-server**
- **locate the nearest router of a multicast group in order to participate in group communication**

IP Addressing: The Future

Prefix (binary)	Usage	Fraction
0000 0000	Reserved (including IPv4)	1/256
0000 0001	Unassigned	1/256
0000 001	OSI NSAP addresses	1/128
0000 010	Novell Netware IPX addresses	1/128
0000 011	Unassigned	1/128
0000 1	Unassigned	1/32
0001	Unassigned	1/16
001	Unassigned	1/8
010	Provider-based addresses	1/8
011	Unassigned	1/8
100	Geographic-based addresses	1/8
101	Unassigned	1/8
110	Unassigned	1/8
1110	Unassigned	1/16
1111 0	Unassigned	1/32
1111 10	Unassigned	1/64
1111 110	Unassigned	1/128
1111 11100	Unassigned	1/512
1111 111010	Link local use addresses	1/1024
1111 111011	Site local use addresses	1/1024
1111 1111	Multicast	1/256

IP Addressing: The Future (2)

i.e.,
- **provider based: approx. 16 mio. companies allocate addresses**
- **geographically based: allocation as it is today**
- **link, site-used: address has only local importance (security, Firewall concept)**

Anycast
- **send data to one member of a group**
- **for example to the member which is the nearest one geographically**