Rechnernetze - Computer Networks

Lecture 9: Internetworking
Addressing

Prof. Dr.-Ing. Markus Fidler

Institute of Communications Technology
Leibniz Universität Hannover

June 14, 2024

Outline

- Internet protocol version 4
  - Fragmentation
  - Addressing
- Dynamic host configuration protocol
- Network address translation
- Internet protocol version 6

Network layer protocols and functionalities

Routing protocols

- maintaining routing tables
- algorithms for route computation
- exchange of routing information

Internet protocol (IP)

- addressing
- processing of datagrams

Internet control message protocol (ICMP)

- signalling
- error notifications

Internet protocol and Internet control message protocol

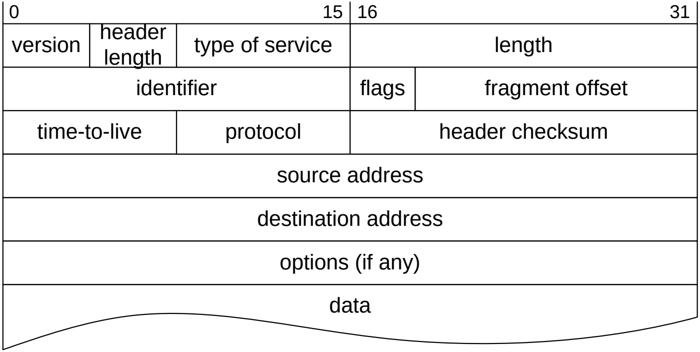
Specification

- Internet protocol (IP), RFC 791
- Internet control message protocol, RFC 792
- RFC 791 and RFC 792 together are Internet standard 5
- IP-addresses are specified in further RFCs, e.g RFC 796

Characteristics

- IP specifies basically
  - the network layer header
  - a procedure for fragmentation of datagrams
- ICMP is used for error messages and network tests
  - among routers as well as
  - between hosts and routers
- errors may be due to
  - erroneous IP datagrams
  - unavailability of networks, hosts, routers, protocols, or services

Internet protocol header



- header length in 32 bit words; length including header in byte
- identifier, flags, and fragment offset for fragmentation
- addressing: source and destination address
- (de)multiplexing: protocol specifies next higher protocol

Characteristics

- ICMP messages are transmitted as data in IP datagrams
- ICMP messages contain type and value (code) and in case of an error message the first 8 byte of the IP datagram that caused the error

type	value	code
0	0	echo reply (ping)
3	1	destination host unreachable
3	2	destination protocol unreachable
3	3	destination port unreachable
3	4	fragmentation needed but don't fragment bit set
4	0	source quench (not used)
8	0	echo request (ping)
11	0	time-to-live expired
12	0	bad IP header

Ping and traceroute

Ping

- task: verify connectivity to a destination host
- the source host sends an ICMP echo request to the destination host
- the destination host sends an ICMP echo reply to the source

Traceroute

- task: identify all hops along a network path
- the source host sends several IP datagrams with increasing time-to-live field, i.e. $TTL = 1 \ldots n$
- each router along the path decrements the time-to-live field
- if the time-to-live field reaches zero the respective router
  - discards the IP datagram
  - sends an ICMP time-to-live expired message to the source
  - the ICMP time-to-live expired message contains the IP-address of the router

Fragmentation

Maximum size of an IP datagram

- is determined by the maximum transmission unit (MTU) of the data link/physical layer of the current hop
- can differ from hop to hop

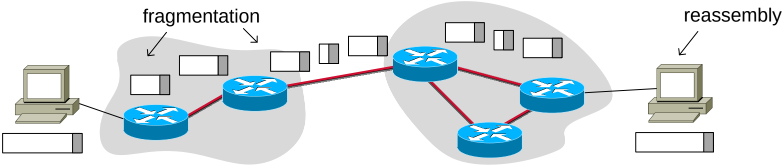
Typical MTU sizes

- Ethernet: 1500 byte
- X.25: 576 byte
- Wifi: 2312 byte

Fragmentation continued

Fragmentation

- if the IP datagram is too large it is broken down and sent as a number of fragments
- each fragment has an IP header with source and destination address for routing
- fragmentation can be used at any hop, reassembly is done only at the destination



Fragmentation continued

Fragmentation makes use of a number of fields in the IP header

- identification: unique value for each IP datagram; is copied into each fragment
- flags: more fragments bit; is turned on except for the last fragment of a segment
- fragment offset: offset of the current fragment from the beginning of the original datagram in units of eight byte
- length: is used as the total length of the IP datagram (fragment) including the header

Fragmentation: example

		length	identifier	more frag.	offset		e.g. UDP	payload
segment	...	2981	17	0	0	...	8 byte	2953 byte

←20 byte IP header→ ←IP data part→

		length	identifier	more frag.	offset		e.g. UDP	payload
1. fragment	...	1500	17	1	0	...	8 byte	1472 byte
2. fragment	...	1500	17	1	185	...		1480 byte
3. fragment	...	21	17	0	370	...		1 byte

Fragmentation: example continued

		length	identifier	more frag.	offset		data
1. fragment	...	572	17	1	0	...	552 byte
2. fragment	...	572	17	1	69	...	552 byte
3. fragment	...	396	17	1	138	...	376 byte
4. fragment	...	572	17	1	185	...	552 byte
5. fragment	...	572	17	1	254	...	552 byte
6. fragment	...	396	17	1	323	...	376 byte
7. fragment	...	21	17	0	370	...	1 byte

IPv4 addressing

IPv4 address

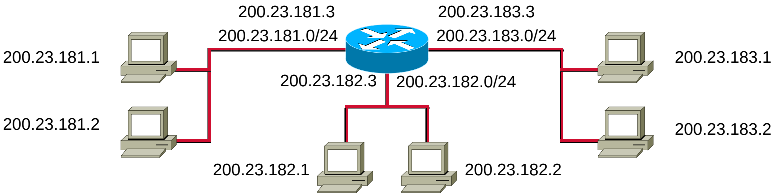
- 32 bit unsigned integer (how many addresses exist?)
- initially worldwide unique (how many people live on earth?)
- identifies a network interface (not a host or router)
- dotted decimal notation: 200.23.181.217
  - higher bits mark the network address
  - lower bits mark the host address

Network interface

- connects a node to the physical transmission medium
  - hosts usually have only one active network interface and do not perform any relaying of data
  - routers have at least two but usually more network interfaces and perform routing
- has an additional address at the data link layer
  - MAC-address, e.g., in case of Ethernet
  - address resolution protocol (ARP): maps IP to MAC address

Subnets

- the set of network interfaces with identical network address form a subnet (RFC 950)
- hosts with identical network address are directly connected (without any intermediate routers)
- network address notation: 200.23.181.0/24 denotes a 24 bit network address



Network mask and default gateway

Network mask

- set all bits of the network address to one
- set all bits of the host address to zero
  - e.g. subnet 200.23.181.0/24 has netmask 255.255.255.0
- test whether a destination is located in a foreign subnet
  - if (src. address. & netmask) $\neq$ (dest. address. & netmask)

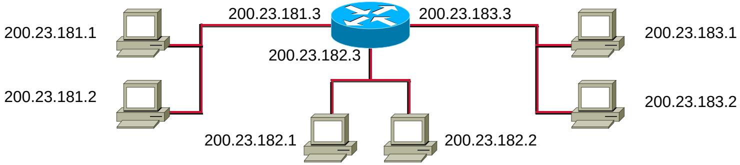
Default gateway

- network interface of a router in the host's subnet
  - e.g. host 200.23.181.1 has default gateway 200.23.181.3
- configured manually or dynamically on hosts
- hosts forward IP datagrams with foreign network address to the default gateway

How to identify subnets

To identify subnets

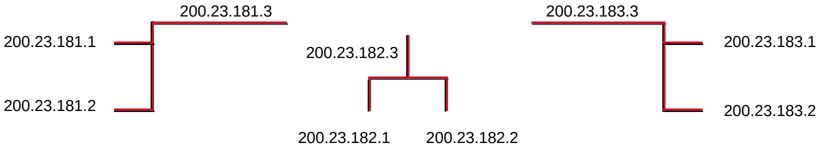
- remove all layer 3 nodes, i.e., all hosts and routers
- a number of isolated networks remain
- each isolated network is a subnet



How to identify subnets

To identify subnets

- remove all layer 3 nodes, i.e., all hosts and routers
- a number of isolated networks remain
- each isolated network is a subnet



Classful IPv4 addressing

Traditional apportionment of the IP address space

class	composition	number of nets	number of hosts	share
A	0net.host.host.host	$2^7 - 2 = 126$	$2^{24} - 2 = 16\,777\,214$	50 %
B	10net.net.host.host	$2^{14} = 16\,384$	$2^{16} - 2 = 65\,534$	25 %
C	110net.net.net.host	$2^{21} = 2\,097\,152$	$2^8 - 2 = 254$	12.5 %
D	1110 multicast	-	-	6.25 %

- two network addresses are reserved (RFC 6890)
  - address 0.0.0.0/8: this host on this network
  - address 127.0.0.0/8: localhost or loopback address
- two host addresses are reserved in each subnet
  - host address of all zeros: refers to the network
  - host address of all ones: broadcast

Classless interdomain routing (CIDR)

Problem of classful IP addressing

- unflexible and wasteful use of IP addresses
- e.g. given a company with 1000 hosts
  - class B nets are to large, e.g. 64534 addresses are wasted
  - class C nets are to small, e.g. 254 addresses do not suffice

Solution: classless interdomain routing

- no predefined border between network and host address
- address format a.b.c.d/x
- x denotes the number of bits of the network address
- e.g. for 1000 hosts a subnet where $x = 22$ is sufficient

Further remarks

- subnetting can be applied to classful IP addresses, RFC 950
- private nets 10.0.0.0/8, 172.16.0.0/12, and 192.168.0.0/16 are not routed (RFC 6890)

Allocation of IP addresses

Internet corporation for assigned names and numbers (ICANN)

- highest instance for administration of IP addresses and DNS (domain name system) entries
- regional sub-organizations
  - AfriNIC: African network information centre
  - APNIC: Asia Pacific network information centre
  - ARIN: American registry for Internet numbers
  - LACNIC: Latin American and Caribbean IP address registry
  - RIPE NCC: Réseaux IP Européens

Internet service providers (ISPs)

- obtain address spaces
- allocate blocks of their address space to customers
- take care of routing between their customers and the Internet
- maintain DNS entries and delegate DNS zones

Allocation of IP addresses continued

- list of address spaces allocated by RIPE
  `http://www.ripe.net/ripe/docs/`
- whois data base of RIPE
  `http://www.ripe.net/perl/whois`
  - information about network operators
  - information about autonomous systems
- administration of DNS root zones
  `https://www.iana.org/domains/root/db`

Allocation of IP addresses continued

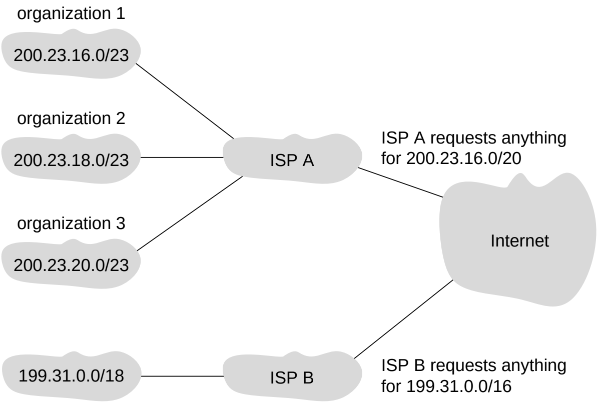
ISP owns continuous address space

- e.g. 200.23.16.0/20
  - 20 bit for addressing the network
    11001000 00010111 00010000 00000000
  - network mask is 255.255.240.0
  - 12 bit remain for addressing hosts

ISP divides the address space into eight equal-sized blocks

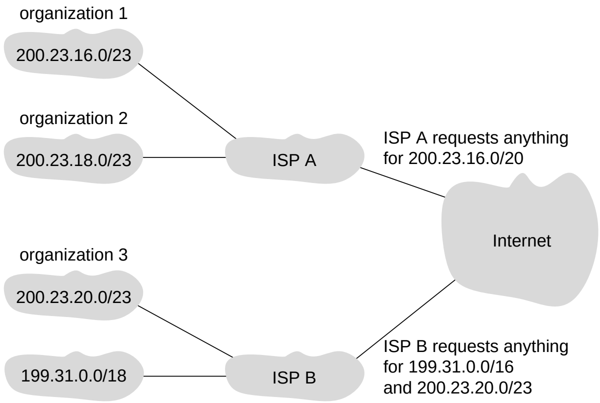
- 3 bit for addressing the networks of customers
  - customer 1's network: 200.23.16.0/23
    11001000 00010111 00010000 00000000
  - customer 2's network: 200.23.18.0/23
    11001000 00010111 00010010 00000000
  - customer 3's network: 200.23.20.0/23
    11001000 00010111 00010100 00000000
  - ...
  - network mask is 255.255.254.0
  - 9 bit remain for addressing hosts

Route aggregation



- how can e.g. organization 3 move to ISP B without having to change all IP addresses?

Longest prefix matching



- the problem is solved by longest prefix matching that is used by routers to determine the next hop

Outline

Configuring IP addresses

An organization that has a block of IP addresses assigns these to

- the router interfaces in the organization (usually manually)
- the hosts
  - manually
  - automatically: dynamic host configuration protocol (DHCP)

DHCP provides plug-and-play functionality, i.e. hosts can

- obtain an IP address on request
  - temporary address: address may change whenever the host connects to the network
  - static address: address does not change (uses MAC-address)
- obtain additional information such as
  - subnet mask
  - default gateway (first hop router)
  - DNS server (server that resolves URLs for IP addresses)

Dynamic host configuration protocol (DHCP)

Typical DHCP use cases

- nomadic hosts that connect to different networks, e.g. WLAN
- address management for residential ISPs
  - allows saving addresses if only a certain fraction of users are online at the same time

Characteristics of DHCP

- specified in RFC 2131
- uses UDP
- client server architecture
  - hosts (clients) request IP addresses from the DHCP server
  - the DHCP server manages a block of IP addresses and assigns addresses dynamically to hosts

Dynamic host configuration protocol (DHCP)

The DHCP protocol uses four steps

- **DHCP server discovery**: a newly arriving client sends a
  - DHCP discover message
  - with a random transaction ID
  - using UDP destination port 67
  - to destination IP address 255.255.255.255 (broadcast)
  - with source address 0.0.0.0 (this host)
- **DHCP server offer(s)**: each DHCP server responds a
  - DHCP offer message
  - referring to the transaction ID
  - to IP address 255.255.255.255
  - with an IP address for the client, network mask, and lease time
- **DHCP request**: the client chooses an offer and responds a DHCP request message to the respective server
- **DHCP acknowledge**: the server confirms with a DHCP acknowledgement message

Outline

Internet protocol version 4
- Fragmentation
- Addressing

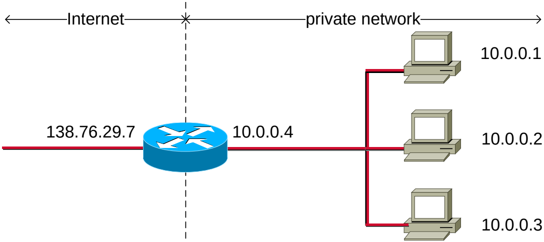
Dynamic host configuration protocol

Network address translation

Internet protocol version 6

Approach

- not all hosts in a local network require global IP addresses
- private networks, e.g. 10.0.0.0/24, are sufficient for local communication; recall private IP addresses are not routed
- mapping of private to worldwide unique IP addresses is done by network address translation (NAT) router
- NAT is specified in RFC 2663, RFC 3022



Motivation

- many local, e.g. home, networks without permanent connection to the Internet
- hosts must not be reachable from outside (exception servers)
- can save cost of allocating IP addresses to each and every host

Typical use case

- a single host from a local network dials in at an ISP
- the host is assigned a public (worldwide unique) IP address
- the host is configured as a NAT router
- the host maps all private IP addresses of the local network to its own address
- to achieve a one-to-one reversible mapping additional fields need to be used to mark the packets:
  - transport layer (L4) port fields in UDP and TCP

NAT router

A NAT-enabled router

- usually obtains its IP address from the DHCP server of an ISP
- usually runs a DHCP server to provide IP addresses to the home network

- appears to the outside world (Internet) as a single device
- hides the home network's internals from the outside world

- all outgoing datagrams have the same source IP address
- all incoming datagrams have the same destination IP address

- maintains a translation table for the home network's IP addresses
- uses port fields of the transport layer (L4) to mark packets

Mode of operation

Modification of outgoing datagrams (from the local network)

- substitute the source IP address by the global IP address of the NAT router
- substitute the source port by an arbitrary, available source port
- send modified datagram into the Internet

Entry into the NAT translation table

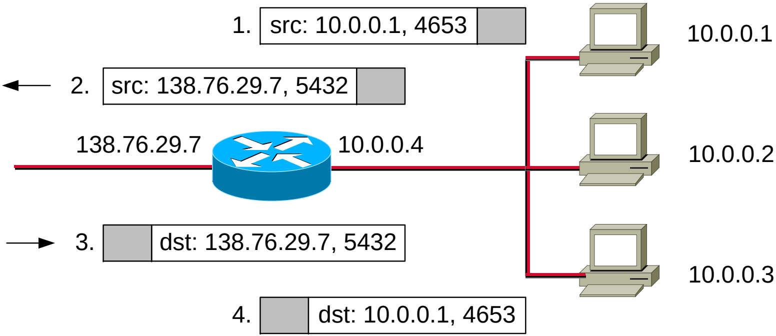
- one-to-one mapping of source IP address and source port to NAT IP address and new source port

Modification of incoming datagrams (from the Internet)

- check NAT translation table for entries that match the destination IP address and the destination port field
- exchange the destination IP address and port for the original IP address and port from the NAT translation table
- send modified datagram into the local network

Translation table

	translation table	
	global addresses	local addresses
2.	138.76.29.7, 5432	10.0.0.1, 4653
	...	...



Critique

NAT violates a number of networking principles and best-current Internet practices

- port numbers are originally meant to address application layer processes and not hosts (respectively interfaces)
- routers are supposed to work only on layers one up to three but not on the transport layer
- NAT violates layering, it mixes network and transport layer functionality
- NAT routers are intermediate systems that violate the end-to-end argument
- IP address shortage should be fixed by IPv6 and not by patching IPv4

Outline

- Internet protocol version 4
  - Fragmentation
  - Addressing
- Dynamic host configuration protocol
- Network address translation
- Internet protocol version 6

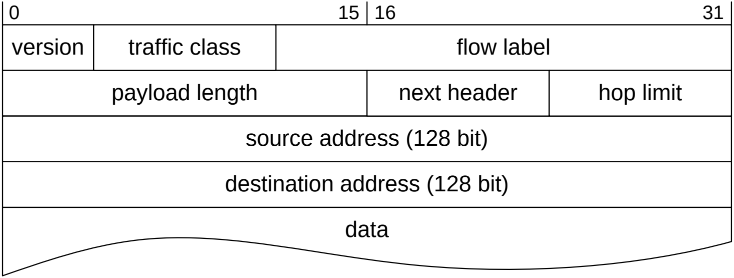
Main differences of IPv6 compared to IPv4

IP version 6

- specified in RFC 2460, Dec. 1998
- 40 byte header: a number of fields have been dropped
- larger address space: source and destination address are extended to 128 bit (before 32 bit)
- no fragmentation at intermediate routers: fragmentation at routers is considered inefficient
- no checksum: recomputation of the checksum at each router (due to change of the TTL) is considered too costly
- flow label: notion of flows, flows can be labelled

IPv6 header

Institut für Kommunikations-Technik



- version: 0x06
- traffic class: formerly type of service
- flow label: new notion of a flow
- payload length: formerly length
- next header: formerly protocol
- hop limit: formerly time-to-live

Omitted fields

- no header length, no options field
  - options can, however, be specified as a next header
- no identifier, flags, fragment offset
  - fragmentation only by hosts using a fragment extension header
  - routers do not fragment at all
  - routers instead generate ICMP packet too big messages
  - new ICMP version for IPv6, RFC 4443
- no header checksum