Communication Networks
Chapter 11: The Internet

- Internet Protocol (version 4)
- Auxiliary Protocols in the Internet Layer
- Internet: The New Generation
- Transport Layer
- Applications

Repetition: The Internet Protocol Reference Model

<table>
<thead>
<tr>
<th>Layer</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application Layer</td>
<td>Application protocols implement application-specific functionalities</td>
</tr>
<tr>
<td>Transport Layer</td>
<td>Data transmission end-to-end</td>
</tr>
<tr>
<td>Internet Layer</td>
<td>Routing and addressing</td>
</tr>
<tr>
<td>Data Link Layer</td>
<td>Interface to the physical medium Network-to-Host (N2H)</td>
</tr>
<tr>
<td>Physical Layer</td>
<td></td>
</tr>
</tbody>
</table>

The OSI layers 5 to 7 are integrated into one application layer.
The Internet Layer

- **Transport Layer (TCP, UDP, ...)**
- **Internet Layer**
  - **Routing Protocols**
    - Routing tables
    - RIP, OSPF, BGP, ...
  - **IP**
    - Addressing
    - Packet format
    - Packet processing
  - **Routing Table**
  - **ICMP**
    - Error/state messages
    - Control
  - **ARP/RARP**
    - Address translation
    - IP <-> MAC
  - **IGMP / DHCP / NAT / CIDR / RSVP / ...**

Network to Host Layer

The Internet Protocol

- **History:**
  - Initiated by the American Department of Defense (DoD)
  - Deployed in the so called ARPANET in 1969 (originally 4 hosts!).
- **Current state and further developments:**
  - IP is the most commonly used layer-3-protocol nowadays.
  - 1998, the project IPng (IP next generation) of the IETF (Internet Engineering Task Force) led to the new version IPv6.
Characteristics of IP

- Packet-switched
- Connectionless (Datagrams)
- Unsecured data transmission:
  - Datagram may get lost
  - Datagram may get duplicated
  - Datagrams may overtake each other during transmission
  - Datagrams might circulate endlessly (theoretically…)
  - If layer-2-entity experiences some error which it cannot correct, then the IP entity will not be able to correct it either
  - With the protocol ICMP (Internet Control Message Protocol) a means to report fatal errors is available
- No flow control
- No explicit congestion control
- IP is used in both private and public networks worldwide
- Global unique (hierarchical) IP addresses are necessary

Interworking with IP
IP Addresses

- Originally 5 address classes (32 bit length):
  1. Class A for networks with up to 16 Mio. nodes
  2. Class B for networks with up to 65.536 nodes
  3. Class C for networks with up to 256 nodes
  4. Class D for multicast
  5. Class E, reserved for future applications

<table>
<thead>
<tr>
<th>Network ID</th>
<th>Node ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

IP Subnetwork Addresses

- IP-Address (e.g. class B):
  Network Part | Local Part
  | | |

  Network Part | Subnetwork Part | End System
  | | |

- Subnetwork masks identify the part of the IP address, that belongs to the network path and the subnetwork part (binary „1”s).
- Example:
  IP Address: 129.13.3.64
  Subnetwork Mask: 255.255.255.0
  Network ID: 129.13.3.0
  Subnetwork ID: 255.255.255.0
  End System: 64

- The network ID can be derived from the address class
- If the subnetwork mask only covers the network part, there will be no subnetwork part (e.g. 255.255.0.0)
IP Subnetworking: Example

IP Datagram

Communication Networks - 11. The Internet
**IP Datagram: Fields**

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version</td>
<td>Version number of IP (4)</td>
</tr>
<tr>
<td>Header Length</td>
<td>Length of IP header (in words of 32 bits)</td>
</tr>
<tr>
<td>TOS / DiffServ</td>
<td>Quality of service information</td>
</tr>
<tr>
<td>Total Length</td>
<td>Length of total datagram</td>
</tr>
<tr>
<td>Identifier</td>
<td>Datagram identifier (no sequence number!)</td>
</tr>
<tr>
<td>Flags</td>
<td>Needed for segmenting</td>
</tr>
<tr>
<td>Fragmentation Offset</td>
<td>Needed for Reassembly</td>
</tr>
<tr>
<td>Time to Live</td>
<td>Maximum lifetime for the datagram</td>
</tr>
<tr>
<td>Protocol</td>
<td>Protocol of the higher layer (e.g. 6=TCP, 17=UDP)</td>
</tr>
<tr>
<td>Header Checksum</td>
<td>Error detection for header</td>
</tr>
<tr>
<td>Source/Destination Address</td>
<td>Source and destination system</td>
</tr>
<tr>
<td>Options</td>
<td>Additional information for special functions</td>
</tr>
<tr>
<td>Padding</td>
<td>for 32-Bit-Alignment</td>
</tr>
<tr>
<td>Data</td>
<td>User information</td>
</tr>
</tbody>
</table>

**IP Routing**

- Each IP entity has its routing table
- According to the destination address, the table entry is chosen that defines the next hop:
  - Check the host addresses
  - Check the network addresses
  - Look for default entry

<table>
<thead>
<tr>
<th>Destination is...</th>
<th>Route</th>
<th>MAC frame is addressed to...</th>
</tr>
</thead>
<tbody>
<tr>
<td>... directly connected</td>
<td>Direct Route</td>
<td>Destination System</td>
</tr>
<tr>
<td>... not directly connected</td>
<td>Indirect Route</td>
<td>Next Router</td>
</tr>
</tbody>
</table>
Routing Example

- The following destination nodes shall be addressed
  1. 129.13.35.73 (sioux.telematik.informatik.uni-karlsruhe.de)
  2. 132.151.1.19 (www.ietf.org)

- The system's routing table looks like this:

<table>
<thead>
<tr>
<th>Destination</th>
<th>Gateway</th>
<th>Flags</th>
<th>Refs</th>
<th>Use</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>i70lr0</td>
<td>UGS</td>
<td>1</td>
<td>13320</td>
<td>tu0</td>
</tr>
<tr>
<td>127.0.0.1</td>
<td>localhost</td>
<td>UH</td>
<td>7</td>
<td>242774</td>
<td>lo0</td>
</tr>
<tr>
<td>129.13.3</td>
<td>i70r35</td>
<td>UGS</td>
<td>0</td>
<td>6</td>
<td>tu0</td>
</tr>
<tr>
<td>129.13.35</td>
<td>mohave</td>
<td>U</td>
<td>11</td>
<td>3065084</td>
<td>tu0</td>
</tr>
<tr>
<td>129.13.41</td>
<td>i70r35</td>
<td>UGS</td>
<td>2</td>
<td>4433</td>
<td>tu0</td>
</tr>
<tr>
<td>129.13.42</td>
<td>i70r35</td>
<td>UGS</td>
<td>0</td>
<td>4</td>
<td>tu0</td>
</tr>
</tbody>
</table>

IP-Multicasting

- An IP datagram shall be sent to more than one end system (optional).
- The group of the receiving nodes is managed using IGMP (Internet Group Management Protocol).
- The address is a class D address
  - Starting with 1110.
  - Followed by a 28 bit long group ID.
- Example

![Diagram of IP-Multicasting](image)
IP Functions: Header Check

- Checking the IP datagram header:
  - Checking the correct header length
  - Checking the IP version number
  - Checking the correct length of the datagram
  - Checking the header checksum
  - Checking the time to live
  - Checking the protocol ID
  - Checking the address classes of both source and destination address

- If one check fails, the IP datagram will be discarded and an ICMP error message will be sent to the sender of the IP datagram.

IP Function: Source Routing

- Usually, the IP entity is responsible for forwarding the datagram.
- The sending protocol entity on top of IP may also define the path (a list of intermediate nodes) the packet must take to reach its destination:
  - Use pointer P to address the next router
  - This pointer then replaces its address with the outgoing address and increments pointer P by four [bytes]

  - **Strict Source Routing**
    - The sender has to define the complete path for the packet to the destination.

  - **Loose Source Routing**
    - The sender only specifies a subset of all the nodes on the path to the destination
    - Additional intermediate nodes will be chosen by the IP entities on the routers accordingly
    - Using the “Route Recording” option, the path of the packet is logged in the options field of the header
IP Function: Source Routing – Example

In the datagram’s option field, the path is logged:

Options

IP Data-gram

1. Add your address in the field given by pointer P...

2. ... and increment pointer P by 4 [byte] so that it points to the next empty entry in the list.
IP Function: Time Stamp

- Each router adds a time stamp to the options field, which defines the point in time, at which the router processed the datagram.
  - Statements on the network load are possible
  - Efficiency of the routing algorithms can be deduced
- A 4 bits long field defines the possibilities:
  - Flag = 0: Just include time stamps, but no addresses.
  - Flag = 1: Both time stamps and addresses are logged (Route Recording)
  - Flag = 3: Addresses are given by the sender (Source Routing), routers add their time stamps.

IP Function: Segmentation and Reassembly

- To support different maximum transmission units (MTUs), IP provides segmentation and reassembly:
  - Example Ethernet: 1.500 byte user information
- Information for segmentation and reassembly are stored in the header of the datagram:
  - Flags
    - Bit 0: Reserved
    - Bit 1: 0 = datagram may be segmented, 1 =datagram may no be segmented
    - Bit 2: 0 = last segment, 1 = more segments to come
  - Fragment Offset
    - Gives the position, where the segment has to be included into the original message (unit 8 bytes)
IP Function: Segmentation and Reassembly – Example

- Example for computing the fragment offsets

| Fragment 1 Header | Data 1 600 bytes | Fragment Offset: 0 |
| Fragment 2 Header | Data 2 600 bytes | Fragment Offset: 600/8 = 75 |
| Fragment 3 Header | Data 3 200 bytes | Fragment Offset: 1200/8 = 150 |

IP: Resume

- Although IP is the most important protocol in the Internet, it does not provide all the required functionalities. Thus, there must be set of other protocols extending the functionality of the Internet layer.

- The number of IP addresses has currently reached its limits (though there are theoretically $2^{32} = 4,294,967,296$ addresses)
  - New scheme for addressing is required
  - The IP protocol has to be changed radically
  - Currently, there are some „patches“ to still work with the limited address space, but soon we run out of IP addresses!
Requests for Comments RFCs


The Internet Layer

<table>
<thead>
<tr>
<th>Transport Layer (TCP, UDP, ...)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet Layer</td>
</tr>
<tr>
<td>Routing Protocols</td>
</tr>
<tr>
<td>- Routing tables</td>
</tr>
<tr>
<td>- RIP, OSPF, BGP, ...</td>
</tr>
<tr>
<td>IP</td>
</tr>
<tr>
<td>- Addressing</td>
</tr>
<tr>
<td>- Packet format</td>
</tr>
<tr>
<td>- Packet processing</td>
</tr>
<tr>
<td>Routing Table</td>
</tr>
<tr>
<td>ICMP</td>
</tr>
<tr>
<td>- Error/state messages</td>
</tr>
<tr>
<td>- Control</td>
</tr>
<tr>
<td>ARP/RARP</td>
</tr>
<tr>
<td>- Address translation</td>
</tr>
<tr>
<td>- IP &lt;-&gt; MAC</td>
</tr>
<tr>
<td>IGMP / DHCP / NAT / CIDR / RSVP / ...</td>
</tr>
</tbody>
</table>

Network to Host Layer
**ICMP – Internet Control Message Protocol**

- If packets are discarded due to overload, the router will not notify the sender about it.
- Fatal problems (e.g. link failure) will be reported via ICMP in order to prevent subsequent faults.

**ICMP: Error Messages**

- **Destination unreachable:** A packet could not be delivered, because the destination node is not available e.g. due to link or router failure.
- **Time exceeded:** A packet has been discarded because its time to live counter has been decremented to zero.
- **Parameter problem:** A packet has been discarded because the node could not deal with one of the parameters in the packet header (e.g. version number).
- **Source quench:** An overloaded node requests the sender to decrease the transmission rate. In IP Version 4 routers are not allowed to originate a source quench and are not obligated to act on a received source quench.
- **Redirect:** A packet should better be sent over another path.

**Error Messages contain a field that explains the reason for the error in detail, e.g. network unreachable error or host unreachable error in the message destination unreachable.**
ICMP: Status Requests

- **Echo and Echo Reply:**
  Serves to monitor the activity of an Internet node. The receiver of an echo message sends an echo reply message with the received information to the sender.

- **Timestamp and Timestamp Reply:**
  Serves to estimate packet round trip times. There are several fields in the packet containing time stamps which characterize the packet processing at the receiver and the delays in the network.

![Diagram of ICMP message exchange](image)

ICMP: Packet Format

- **Transmission of ICMP messages**
  - ICMP messages are transmitted in the data portion of an IP datagram. The value of the "protocol" field in the IP header is set to "1".

- **Format of ICMP messages**
  - **Type:** message type (e.g., 3 means "Destination Unreachable")
  - **Code:** detailed explanation of the message (e.g., 1 means "host unreachable")
  - **Checksum:** error detection mechanism for the complete ICMP message
  - **Info:** depends on message type (e.g., fields for time stamps).
### ICMP: traceroute

Traceroute to www.ietf.org [4.17.168.6]:

<table>
<thead>
<tr>
<th>Step</th>
<th>TTL in IP datagram</th>
<th>IP Address</th>
<th>Response Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10 ms</td>
<td>141.24.95.253</td>
<td>&lt;10 ms</td>
</tr>
<tr>
<td>2</td>
<td>90 ms</td>
<td>golden-gate.rz.tu-ilmenau.de</td>
<td>&lt;10 ms</td>
</tr>
<tr>
<td>3</td>
<td>90 ms</td>
<td>ar-ilmenau1.g-winn.dfn.de</td>
<td>&lt;10 ms</td>
</tr>
<tr>
<td>4</td>
<td>90 ms</td>
<td>cr-leipzig1.g-winn.dfn.de</td>
<td>&lt;10 ms</td>
</tr>
<tr>
<td>5</td>
<td>90 ms</td>
<td>cr-frankfurt1.g-winn.dfn.de</td>
<td>&lt;10 ms</td>
</tr>
<tr>
<td>6</td>
<td>90 ms</td>
<td>so-6-0-0-ar2.FRA2.gblx.net</td>
<td>&lt;10 ms</td>
</tr>
<tr>
<td>7</td>
<td>90 ms</td>
<td>pos3-0-622M.crm1.BOS1.gblx.net</td>
<td>62.163.32.73</td>
</tr>
<tr>
<td>8</td>
<td>90 ms</td>
<td>pos5-0-2488M.crm1.BOS1.gblx.net</td>
<td>[4.4.212.165.130]</td>
</tr>
<tr>
<td>9</td>
<td>90 ms</td>
<td>so0-4-0-622m.br1.BOS1.gblx.net</td>
<td>206.132.247.78</td>
</tr>
<tr>
<td>10</td>
<td>90 ms</td>
<td>208.51.74.62</td>
<td>&lt;10 ms</td>
</tr>
<tr>
<td>11</td>
<td>90 ms</td>
<td>so-3-0-0-bstrm1-nbr1.bbnplanet.net</td>
<td>[4.24.4.218]</td>
</tr>
<tr>
<td>12</td>
<td>90 ms</td>
<td>p7-0-cambridge1-nbr1.bbnplanet.net</td>
<td>[4.24.6.30]</td>
</tr>
<tr>
<td>13</td>
<td>90 ms</td>
<td>p2-0-cambridge1-crl.bbnplanet.net</td>
<td>[4.1.80.5]</td>
</tr>
<tr>
<td>14</td>
<td>100 ms</td>
<td>s0.foretec.bbnplanet.net</td>
<td>[4.1.138.146]</td>
</tr>
<tr>
<td>15</td>
<td>110 ms</td>
<td>4.17.168.6</td>
<td>110 ms</td>
</tr>
</tbody>
</table>

End of trace.

---

### Address Resolution Protocol (ARP)

- **Task:**
  - Mapping: IP Address $\Rightarrow$ Layer-2-Address (MAC Address)
  - Example: PC „Sioux“:
    - IP-Address: 129.13.35.73 $\Rightarrow$ Ethernet Address: 08-00-2b-a2-80-dd

- **Modus operandi:**
  - ARP receives an IP address for address resolution.
  - ARP sends a broadcast packet in the LAN containing the IP address.
  - All stations in the LAN receive this packet, but only the station that recognizes its IP address sends a reply.
  - The reply is stored in the originating station to avoid continuous ARP requests.
  - However, the entry must be deleted after a certain time.
**ARP: Example**

1. Wanted: Hardware Address for 129.13.35.73
   - Station 129.13.35.71 looks for station 129.13.35.73
2. I am station 129.13.35.73 and my MAC address is 08-00-2b-a2-80-dd

**ARP: Packet Format**

- Transmission of an ARP message:
  - An ARP message is transferred in the information part of a data link frame.

<table>
<thead>
<tr>
<th>Bit 16</th>
<th>Bit 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of physical address</td>
<td>e.g. Ethernet</td>
</tr>
<tr>
<td>Protocol of network layer</td>
<td>e.g. IP</td>
</tr>
<tr>
<td>Length of physical Address</td>
<td>e.g. 6 byte or 4 byte</td>
</tr>
<tr>
<td>Length of layer 3 address</td>
<td>e.g. 1 for „request“</td>
</tr>
<tr>
<td>Type of message</td>
<td>e.g. Ethernet addr: Sender; IP addr: Sender</td>
</tr>
<tr>
<td>Address fields</td>
<td>Ethernet addr: Receiver; IP addr: Receiver</td>
</tr>
</tbody>
</table>

⇒ Length and structure of the address fields depend on the type of the addresses.
Reverse Address Resolution Protocol (RARP)

- **Task:**
  - Mapping MAC address $\Rightarrow$ IP address.
  - Used for diskless work stations that boot from a file server. Therefore, they need their IP address which they do not know when being switched on.
  - The file server only holds one operation system for all diskless work stations.

- **Modus operandi:**
  - Station sends out a broadcast with its own MAC address given by the network card.
  - RARP server receives the broadcast and looks up the according IP address in a configuration file.
  - RARP server sends this IP address back to the MAC address given by the sender of the broadcast frame.

IGMP: Internet Group Management Protocol

- **Problem:**
  - How does a router know, which multicast packets have to be forwarded over which network interfaces?

- **Example:**
  - Packets addressed to the groups 1 and 3 have to be forwarded to the LAN, packets belonging to other groups can be discarded.

- **Solution:**
  - Manual entry of all group addresses in the configuration of the router $\Rightarrow$ high effort in administration for dynamically changing groups
  - Self organized learning of the groups by exchanging information $\Rightarrow$ such a procedure is defined by the Internet Group Management Protocol (IGMP)
IGMP: Protocol Operation (I)

1. Router periodically sends out „Membership Query“ broadcasts on all interfaces. By setting the „Time To Live“ (TTL) to 1, the broadcast is limited to the LAN.

2. After having received such a query, a station starts a timer, initialized with a random number, for all groups it belongs to. If the timer expires, the station replies with a „Membership Report“ message addressed to the group ID with TTL 1.

IGMP: Protocol Operation (II)

3. All group members receive this „Membership Report“ message and stop their according timer. Thus, redundant replies are prevented.

4. The router receives all „Membership Report“ messages and accordingly updates its routing table. If a router repeatedly does not receive a „Membership Report“ message for a certain group, the entry will be deleted from the table.

⇒ If a station joins a group, it immediately sends a „Membership Report“ message to all routers connected to the LAN. For fault tolerance, this message should be repeated until the station will receive an according „Membership Query“ message.
**IGMP: Packet Format**

- Transmission of IGMP messages
  - IGMP messages are transmitted in the information field of an IP datagram. The protocol ID is 2.

- Format of IGMP messages
  - **IGMP-Version**: Version of the IGMP protocol (currently version 3).
  - **IGMP-Type**: Type of the message (e.g. 11 = query, 22 = report).
  - **Unused**: set to 0.
  - **Checksum**: error detection of the complete IGMP message.
  - **Group Address**: Set to 0 in a query message, otherwise it contains the group address the report is related to.

---

**CIDR: Classless Inter-Domain Routing**

- **Up to now**:
  - 3 address classes for IP unicast
  - Many addresses cannot be used because they have been assigned to a certain network

- **Example**:
  - A small company needs about 100 IP addresses and applies for a class C address.
  - 254 addresses are assigned → 154 addresses remain unused

- **CIDR**: Substitute the fixed classes by variable network prefixes with a length from 13 to 27 bits
  - Example: 129.24.12.0/14: The first 14 bits of the IP address identify the network, the remaining 18 bits can be used for the subnetworks / end systems.

- Hierarchical routing now possible:
  - Backbone router, e.g. part of a transatlantic link, only considers the first 13 bits:
    - Smaller routing tables
    - Small computation effort
  - Routers of connected providers then consider e.g. the first 15 bits
  - Finally, a router in a company network with 128 hosts considers the first 25 bits
NAT: Network Address Translation

- Problem:
  - Addresses have to be globally unique, even when CIDR is used.

- Idea:
  - In a company network, only those nodes require a globally unique address that want to communicate with another node somewhere outside the company’s network.
  - This globally unique address can be temporarily assigned.
  - Addresses are administered in an address pool by the gateway.
  - If there are too few globally unique addresses, a port mapping can help in distinguishing the different communication associations.

NAT: Process

- Local IP addresses that are unique in the LAN only
- Internet Traffic
- Gateway
- Address Translation
- Address Pool
- Globally unique IP addresses:
  - 128.211.114.51 — assigned to 10.0.0.5
  - 128.211.114.52
  - 128.211.114.53
DHCP: Dynamic Host Configuration Protocol

- Task
  - Simplify the installation and administration of networked stations.
  - Provide necessary information about IP address, DNS server address, domain name, subnetwork mask, gateway, etc.
  - Automatically integrate a new station into the LAN / the Internet.
  - Based on User Datagram Protocol!

- Client/Server Model
  - Client sends a request (via broadcast) to a DHCP server (possibly over a DHCP relay)
  - DHCP server provides the required information

DHCP: Process

1. Server is directly connected

2. Server is not directly connected: use of a DHCP relay
PPP (Point-to-Point Protocol)

- The biggest part of the Internet is based on point-to-point links:
  - Links between routers in the WAN
  - Subscriber line to the Internet Provider
  - Predecessor SLIP (serial line IP, RFC 1055):
    - No error detection
    - Support of IP only
    - No dynamic address assignment
    - No authentication
  - PPP (RFC 1661 ff.) offers among other:
    - Layer 2 frame with error detection and frame boundaries
    - Control protocol (LCP, Link Control Protocol) for connection setup, connection test, connection negotiation and connection teardown
    - Possibility to negotiate options of network layer independently of the network layer protocol
    - Separate NCP (Network Control Protocol) for all supported protocols on the network layer

PPP: Packet Format

- Packet format according to HDLC:
  - Character-oriented (instead of bit-oriented), therefore character stuffing
  - Typically, only unnumbered frames are used
  - At high error rates (e.g., over a radio link), a more reliable mode with sequence numbers and acknowledgments can be selected
  - PPP support the following protocols in the payload field: IP, AppleTalk, IPX (among others)
  - The maximum payload length is set to 1,500 bytes if not negotiated otherwise
  - Length of the packet header can also be negotiated
Recapitulation

- The Internet layer is not consisting of IP alone!
- Functions that have to be provided by other protocols are
  - Notification of errors
  - Translation of addresses
  - Configuration of stations
  - Management of group members
  - Exchange of routing information

RFCs (1)

- **ARKKO, J.; PIGNATARO, C.:** IANA Allocation Guidelines for the Address Resolution Protocol (ARP), April 2009 (RFC 5494).
- **CHESHIRE, S.:** IPv4 Address Conflict Detection, July 2008 (RFC 5227).
- **DROMS, R.:** Dynamic Host Configuration Protocol, März 1997 (RFC 2131).
- **EGEYANG, K.; SRIXURESH, P.:** Traditional IP Network Address Translator (Traditional NAT), Januar 2001 (RFC 3022).
- **FULLER, V.; LI, T.:** Classless Interdomain Routing (CIDR): The Internet Address Assignment and Aggregation Plan. August 2006 (RFC 4632).
RFCs (2)


Motivation for a „New” Internet Protocol

- Problems with addressing
  - IP address space is not adequate any more
  - Class-B-addresses are not available any more
  - CIDR is only an interim solution
  - There is no real hierarchical addressing
  - Routing tables grow rapidly, which slows down routing
- No real integration of security mechanisms
- QoS requirements caused by multimedia applications
History of IPv6

1993 Call for Proposals for IP next generation (IPng) in RFC 1550
1994 Proposal: SIPP (Simple Internet Protocol Plus), a combination of three submissions
   first prototypic implementations → soft migration is required
1996 First IPv6 backbone (6Bone), first products in store
1998 IPv6 becomes draft standard in RFC 2460
2006 6Bone was shut down
2016 Global Internet infrastructure is still not upgraded to IPv6…

Overview of the Main Characteristics of IPv6

- Enhanced addressing scheme
- New IP packet header format
  - Simple structure
  - Better integration of different options
- Segmentation only end-to-end
- Autoconfiguration of IP systems
- QoS support
- Integration of multicast
- Security mechanisms
IPv6 Address

- 128 bits long addresses
  - Theoretically $3.4 \times 10^{38}$ addresses
  - Optimistic estimation: $700 \times 10^{21}$ per m$^2$
  - Pessimistic estimation (RFC1715): 1,700 per m$^2$
- New notation
  - 8 hexadecimal numbers containing 4 digits separated by colons
    - 5800:0000:0000:0000:0000:0000:0056:0078
  - Sequences of „0“ can be omitted
    - 5800::56:78
- IPv6 addresses can contain topology information for localization

IPv6 Addresses: Aggregatable Unicast Address

- **Top-Level Aggregation** (TLA)
  - Huge Internet Service Providers (ISP) with transit networks, to which other ISPs are connected
- **Next-Level Aggregation** (NLA)
  - Organisations on a lower level
  - Several NLA-levels possible
- **Site-Level Aggregation** (SLA)
  - Individual address hierarchy of a single organization / company

<table>
<thead>
<tr>
<th>3</th>
<th>13</th>
<th>8</th>
<th>Public Topology 24</th>
<th>Site Topology 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>TLA ID</td>
<td>Res.</td>
<td>NLA ID</td>
<td>SLA ID</td>
</tr>
<tr>
<td>Interface ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
IPv6 Addresses: Special Unicast Addresses

- **Local Unicast Address**
  - Addresses in the link-local prefix are only valid on a single link (comparable to the auto-configuration addresses).
  - Unique local addresses (ULA's) are intended for local communication. They are routable only within a set of cooperating sites.

- **Compatible Unicast Address**
  - IPv4-mapped addresses: Prefix (96 „0“-Bits) + IPv4-Address
  - IPv4-translated addresses: Prefix (80 „0“-Bits + 16 „1“-Bits) + IPv4-Address as used by the Stateless IP/ICMP Translation (SIIT) protocol
  - IPX- or OSI-compatible

- **Unspecified Address**
  - 0::0 (or ::) e.g. for booting

- **Loopback Address**
  - 0::1 (or ::1) corresponds to v4 address 127.0.0.1

IPv6 Addresses: Anycast

- New type of address in IPv6
- Part of the unicast address space
- A group of stations is addressed, only the station with the lowest load / smallest distance / best communication path... replies
- For each anycast address, there is an entry in the routing table
- Anycast addresses are relevant for routers
- Application example: Distribution of a web server over several physical stations
IPv6 Addresses: Multicast

- All routers and end systems support multicast
- There is a predefined multicast group for control and management tasks
- IGMP was integrated into ICMPv6
- The multicast address additionally contains
  - Flags (for the distinction temporary/permanent)
  - Scope (to define the radius in which the packet is distributed)

<table>
<thead>
<tr>
<th>Flags</th>
<th>Scope</th>
<th>Group Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 bits</td>
<td>4 bits</td>
<td>112 bits</td>
</tr>
</tbody>
</table>

Comparison of Addressing in IPv4 and IPv6

<table>
<thead>
<tr>
<th>Type of Address</th>
<th>IPv4</th>
<th>IPv6</th>
<th>Outgoing Interfaces</th>
<th>Destination Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unicast</td>
<td>Obligatory</td>
<td>Obligatory</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Multicast</td>
<td>Optional</td>
<td>Obligatory</td>
<td>Group</td>
<td>All members of the group</td>
</tr>
<tr>
<td>Broadcast</td>
<td>Obligatory</td>
<td>—</td>
<td>All</td>
<td>All</td>
</tr>
<tr>
<td>Anycast</td>
<td>—</td>
<td>Obligatory</td>
<td>Group</td>
<td>1</td>
</tr>
</tbody>
</table>
### Packet Headers – Comparison of IPv4 and IPv6

<table>
<thead>
<tr>
<th>IPv4 Header</th>
<th>IPv6 Header</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>V</strong></td>
<td><strong>V</strong></td>
</tr>
<tr>
<td><strong>HL</strong></td>
<td><strong>HL</strong></td>
</tr>
<tr>
<td><strong>TOS</strong></td>
<td><strong>TOS</strong></td>
</tr>
<tr>
<td><strong>Total Length</strong></td>
<td><strong>Total Length</strong></td>
</tr>
<tr>
<td><strong>Identifier</strong></td>
<td><strong>Segmentation Info</strong></td>
</tr>
<tr>
<td><strong>TTL</strong></td>
<td><strong>Header Checksum</strong></td>
</tr>
<tr>
<td><strong>Protocol</strong></td>
<td><strong>Flow Label</strong></td>
</tr>
<tr>
<td><strong>IP Source Address</strong></td>
<td><strong>Payload Length</strong></td>
</tr>
<tr>
<td><strong>IP Destination Address</strong></td>
<td><strong>Next hdr.</strong></td>
</tr>
</tbody>
</table>

- **V:** Version
- **HL:** Header Length
- **TOS:** Type of Service
- **TTL:** Time To Live
- **C:** Class
  - ◦ deleted
  - ◦ shifted

### Extension Headers

- **Chain of extension headers**
  - ◦ Small minimal packet header
  - ◦ According to the requirements of the application and to the characteristics of the network, extension headers may be added in a given sequence
  - ◦ New options or extensions can be added easily
- **Routers need not process all the extension headers**
- **Types of extension headers:**
  - ◦ Hop-By-Hop Options
  - ◦ Routing
  - ◦ Fragment
  - ◦ Authentication Header (AH)
  - ◦ Encapsulating Security Payload (ESP)
  - ◦ Destination Options
  - ◦ No Next Header
### Examples for Extension Headers

Each IPv6 packet can have an optional extension header, which can be used for routing, fragmenting, etc.

- **IPv6-Header + TCP-Header + Data**
- **IPv6-Header + Routing Header + TCP-Header + Data**
- **IPv6-Header + Routing Header + Fragment Header + TCP-Header + Data**

**NH = Next Header**

### ICMPv6

- **ICMPv6** substitutes the protocols ICMP(v4), IGMP(v4) and ARP(v4)
- **ICMPv6** is transmitted using IPv6 → Value of „Next Header“ field: 58
- Returns as many information as possible of the concerned packet without violating the minimal MTU (576 bytes)
- Works like TCP and UDP with a pseudo header:

```
<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Checksum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

- Depending on Type
- Additional Data depending on Type
**IPv6: Segmentation**

- Only the sender may fragment the packet
- If the packet is too big, the router will send an ICMPv6 message „packet too big”
- The Maximum Transfer Unit (MTU) can be determined by information given in the ICMPv6 packet:

```
MTU=1500  MTU=576  MTU=1500
SYN; MSS=1440  SYN; ACK  MSS=1440
Data; 1440 byte
ICMPv6 error  Packet too big
MTU=576
```

**IPv6: Support of Mobile Nodes**

- Mobile nodes need a new configuration when they enter a new network
- Using the autoconfiguration mechanism, they receive a new valid IP address
- Still, their old IP address can be used to communicate with the mobile device
- Thus, packets have to be forwarded from their original network to the new network
IPv6: Automatic Address Configuration

- „Plug & Play”
  - Acquisition of an IP address
  - Discovery of duplicated IP addresses
  - Address resolution
  - Determination of location-dependent parameters (subnet-ID, MTU, DNS server, ...)
  - Router discovery
  - Support of mobile devices

- Neighbor Discovery
  - Special ICMP messages:
    - Router Solicitation / Advertisement
    - Neighbor Solicitation / Advertisement

IPv6: Security Mechanisms

- IPsec
  - Security on IP level
  - Encryption
  - Authentication

- Implemented in special extension headers
  - Authentication Header
    - Data integrity
    - Sender identity
  - Security Encapsulation Header
    - Confidentiality
    - Integrity
    - Authenticity
**IPv6 and Multimedia**

- IPv6 is ready for multimedia streams
  - Flow Label
    - Packets with the same destination get the same flow label
    - These packets can be processed equally
  - Priority
    - Packets can be processed according to their priority
    - Rough differentiation
      - Non real time
      - Real time

- Special mechanisms required in routers
  - No guarantees for special QoS requirements!

**Migration to IPv6**

- Currently, most nodes communicate over IPv4
- How can billions of nodes be migrated to IPv6?
  - Convert all nodes at one point in time – impossible!
  - Thus, slow graceful migration to IPv6: Both standards co-exist for a certain time!
- Mechanisms
  - Tunneling
  - Dual Stack
  - Protocol translation

According to the number of nodes working with IPv6
Migration Approach: Tunneling

- IPv6 router encapsulates IPv6 packets in IPv4 packets and forwards them to the next IPv6 router.
  - IPv6 communication between IPv6 routers (tunnel end points)
  - Other (IPv4) routers do not notice IPv6 communication
- Addressing
  - Automatic (IPv4-compatible addresses) or
  - Configurable (statically configured addresses for tunnel end points)

Migration Approach: Dual Stack

- Both end nodes and routers possess two protocol stacks: IPv4 and IPv6
- The type of address given in the reply of the DNS server decides which stack should be used
- DNS needs to support both protocols
- Number of IPv4 addresses can be reduced by and by

<table>
<thead>
<tr>
<th>Applications</th>
<th>UDP for IPv4</th>
<th>TCP for IPv4</th>
<th>UDP for IPv6</th>
<th>TCP for IPv6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Socket Interface</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network Interface</td>
<td></td>
<td></td>
<td>IPv4 etc.</td>
<td>IPv6 etc.</td>
</tr>
</tbody>
</table>

Prof. Jochen Seitz
**Migration Approach: Protocol Translation**

- Translation of IPv4 packets into IPv6 packets
- Application layer is independent of this translation process
- Examples
  - Stateless IP/ICMP Translator (SIIT)
  - Network Address Translation – Protocol Translation (NAT-PT)
  - Socket-based IPv4/IPv6 Gateway
  - Bump In The Stack (BIS)

**IPv6 in Praxis**

- Operation systems usually support IPv6
- Many current devices also support IPv6
- But
  - Usually, IPv4 is used (protection of investment)
  - Add-ons to IPv4 allow the application of old technology
  - Applications still not need the new features of IPv6 or are satisfied with the add-ons to IPv4
- IPv6 is still subject to special research networks
  - 6bone, the first IPv6-Backbone was shut down on June 6, 2006
  - Internet2 as platform to develop new applications based on IPv6 has not proven to be very successful
RFCs