

INTRODUCTION TO THE WORLD OF IPV6 NETWORKS

Communication Networks

Dr. László Bokor
Ph.D., assistant professor
Department of Networked Systems and Services, BME
bokorl@hit.bme.hu

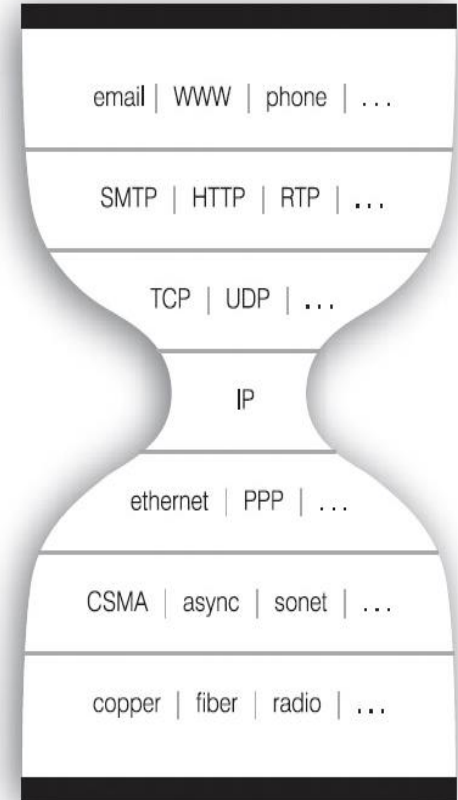


- Introduction to the world of IPv6 networks
 - Development of the IPv6-based Internet
 - Protocol architecture of IPv6, considerations of design
 - Network trends on IPv6

What is the Internet

- Ain't got a better definition than
 - „connection of networks”
 - „IP-using functioning networks”
 - let the philosophers define what it really is...
- Important: need IP for the use of the Internet
 - Internet Protocol
 - Versions: IPv4 and IPv6
 - although today the first one is dominant, we shall focus on the second one during the semester
 - today IPv4, but soon IPv6

- Imagine it as a post office
- IP is
 - packet-based protocol
 - datagram like, unreliable
 - transmission based on address
 - anyone can read it (especially those who should not)
- IP hourglass effect
 - today almost everything goes over IP
 - e.g. voice and text traffic of service providers



source:

<http://www.w3.org/DesignIssues/diagrams/layers/IP-hourglass-zittrain.png>

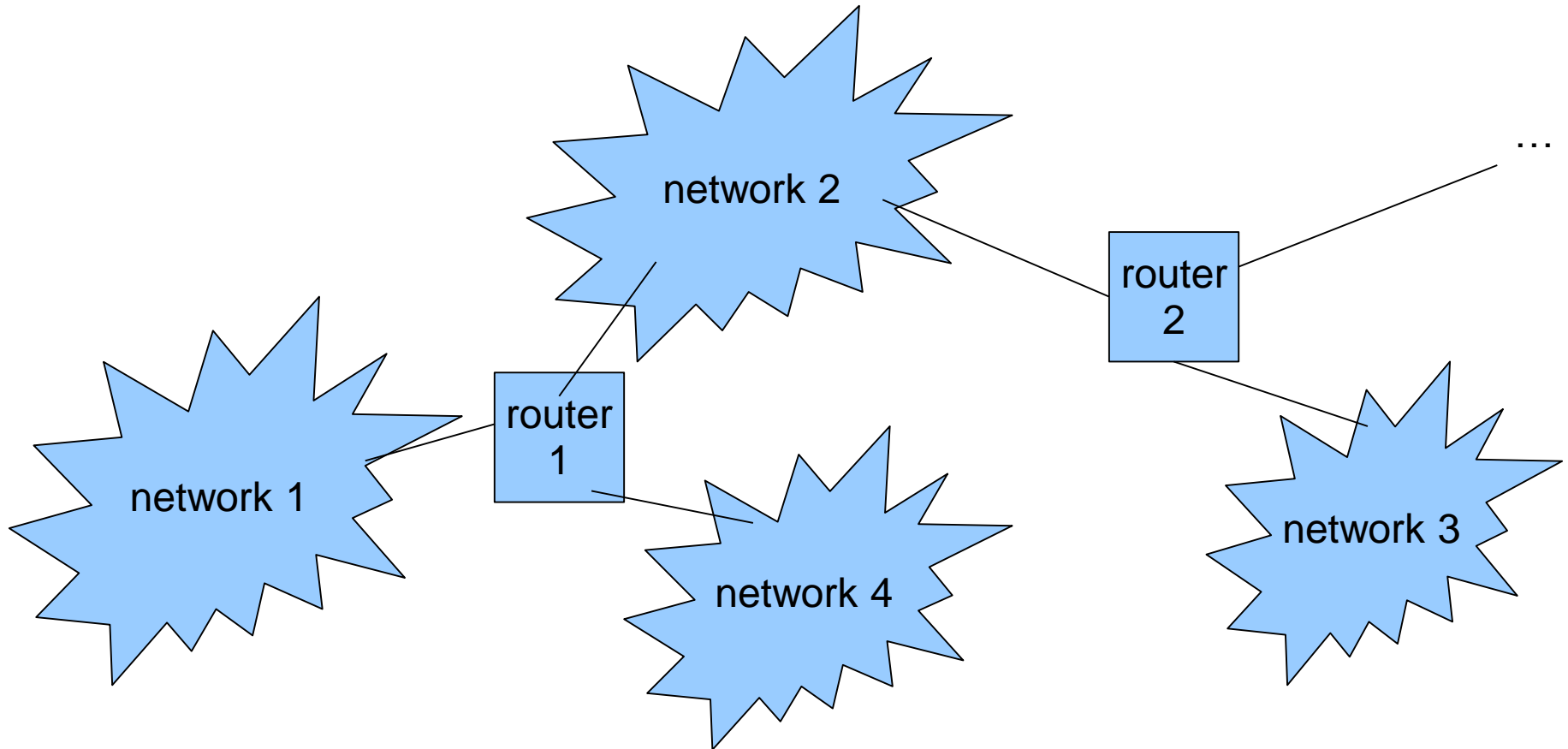
OSI-ARPA reference model

ISO/OSI layers		TCP/IP model	Sample protocols	Devices
7	Application	Application	SOAP, XML	XML Appliances
6	Presentation		HTTP, HTTPS FTP Telnet SMTP LDAP NTP	Content Service Switch Layer 4-7 Switches
5	Session			
4	Transport	Transport	TCP, UDP	Router, Layer-3 Switch
3	Network			
2	Data Link	Network	IP, ICMP, IGMP, IPX	Switches, Bridges
1	Physical	Link	Network Interface: Ethernet, Token Ring, FDDI	Hubs, Repeaters

source: <http://soamag.com/l38/0410-3.php>

- No idea of the routes of packets at the time of sending
 - Best effort
- Every packet includes sender and recipient address
- The protocol does NOT guarantee
 - successful transmission
 - correct arrival destination
 - undamaged packets(handling and correcting packets is the task of upper layers)

IPv4 packet architecture and transmission



- The task of IP is the transmission of packets
- 32 bit addresses
- decimal numbers separated by dots (e.g. 152.66.248.201)
- Every address consists of a network and a station address
 - the network number identifies the network of the station
 - the station number identifies the station
 - e.g. 152.66.248.201

- Space:
 - 32 bit addresses, $2^{32} \approx 4,3 \cdot 10^9$ different addresses
 - Originally classes of A, B, C, D and E
- Operators:
 - Internet Assigned Numbers Authority (IANA)
 - Regional Internet Registry (RIR)
 - Asia Pacific Network Information Centre (APNIC)
 - Réseaux IP Européens Network Coordination Centre (RIPE NCC)
 - American Registry for Internet Numbers (ARIN)
 - Latin American and Caribbean Internet Addresses Registry (LACNIC)
 - African Network Information Centre (AfriNIC)

Regional Internet Registries



source: http://www.apnic.net/__data/assets/image/0011/5204/RIR_map.png

IPv4 address classes, Class A

- Originally the network number was 7 bit, the station number was 24 bit long
 - few but populous network expected
 - '0' + 7 bit + 24 bit (e.g. 68.23.44.198)
- This was titled Class A
- It was later revealed that less populous but numerous networks are rather required

IPv4 address classes, Class B and C

- Two more classes introduced: B and C
- More space for the network (14 and 21 bit) and less for the station (16 and 8 bit)
 - more but smaller address space
 - problems: while Class B provides way too many stations to connect, Class C is insufficient
 - Class B and C is depleted
 - '10' + 14 bit + 16 bit, '110' + 21 bit + 8 bit
 - e.g. 131.33.59.253, 193.224.53.106

- Class D = Internet Multicast
 - '1110' + 28 bit unstructured
 - one address = one multicast group
 - IGMP (Internet Group Management Protocol) introduced (RFC 1112)
- Class E
 - for further use
 - '1111' + 28 bit
 - today distributed as Class B and C

IPv4 special addresses

- 127.0.0.1 => loopback interface (IP stack testability + misc. functions)
- Special station addresses
 - all 0 = network ID (e.g. 152.66.0.0)
 - all 1 = broadcast (e.g. 152.66.255.255)
- Reserved addresses (e.g. NAT)
 - 10.0.0.0/8, 172.224.0.0/12 and 192.168.0.0/16
- Other convections (e.g 12 is DNS)

IPv4 addresses – subnet mask

- Wouldn't be wise to use 24/16/8 bit station addresses in one physical network
 - e.g. 152.66.0.0 => the whole university as one physical network?
- IPv4 addresses reserved for companies/organizations often divided further
 - first part is subnet ID
 - second is station ID
 - part of the addressing architecture [RFC950]
 - division by subnet mask

- Two ways to select the mask
 - one number after slash: how many bits represent the network ID from the beginning (including class bits)
 - IP address form: in the 32 bit mask, the bits belonging to the network or subnet number are 1, the ones to the station are 0
- If the first 8 bits are reserved for subnet in a Class B network, the subnet mask is 255.255.255.0, or /24

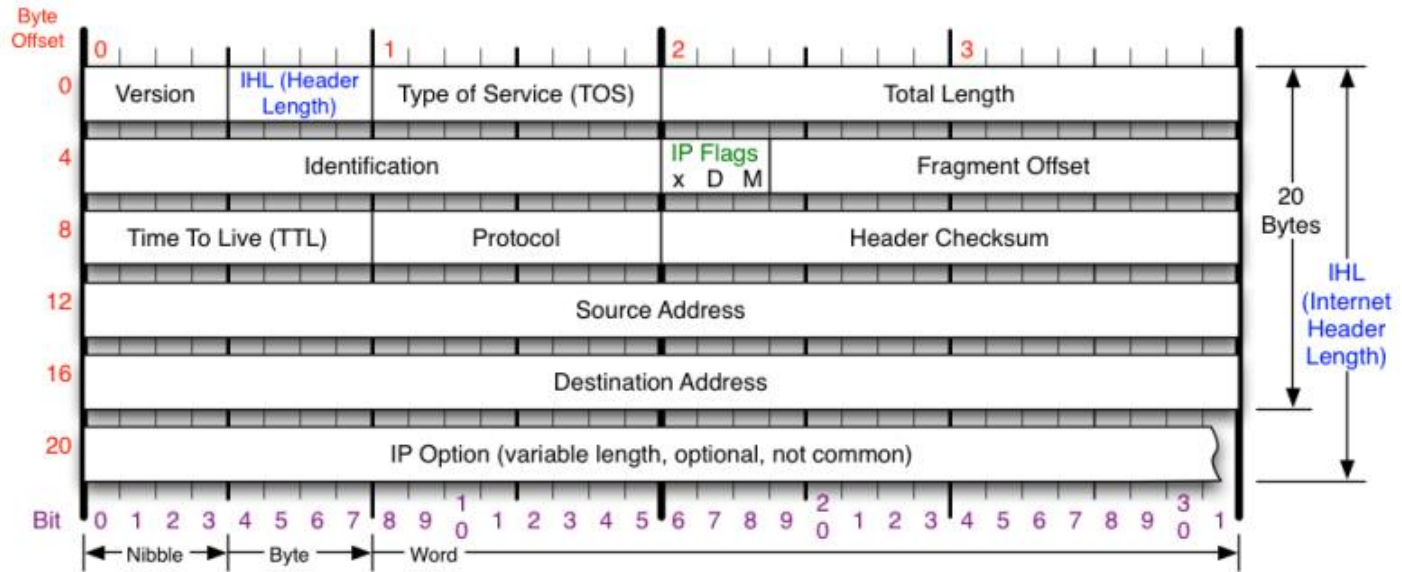
- Example: the 152.66.0.0/16 domain of the university is divided into 152.66.x.0/24 subnets
- There are also subnets of variable lengths
 - there can be a subnet with identification by the upper 8 bits of the station number, and another one with the upper 12 bits, in the same domain
- Subnets must be made prefix free
 - Subnet numbers must be distributed in a way that the first 8 bits are sufficient to decide the type (8 or 12 bit subnet number), see classes

IPv4 – variable length subnets

- Might come handy when we have several subnets with different volumes and in case of fix assignment the address space is insufficient
- This requires more complicated routing protocols and administration

- Requirements for stations on the same subnet
 - must be neighbours on the same link
 - ability to communicate without router
- Can be several subnets on 1 link
- A subnet cannot contain more than 1 link
 - transit between links only through routers
- If there are more subnets on 1 link, then the stations on different subnets shall communicate through router, although they could do it directly

IPv4 header format



Version

Version of IP Protocol. 4 and 6 are valid. This diagram represents version 4 structure only.

Header Length

Number of 32-bit words in TCP header, minimum value of 5. Multiply by 4 to get byte count.

Protocol

IP Protocol ID. Including (but not limited to):

1 ICMP	17 UDP	57 SKIP
2 IGMP	47 GRE	88 EIGRP
6 TCP	50 ESP	89 OSPF
9 IGRP	51 AH	115 L2TP

Total Length

Total length of IP datagram, or IP fragment if fragmented. Measured in Bytes.

Fragment Offset

Fragment offset from start of IP datagram. Measured in 8 byte (2 words, 64 bits) increments. If IP datagram is fragmented, fragment size (Total Length) must be a multiple of 8 bytes.

Header Checksum

Checksum of entire IP header

IP Flags

x D M

x 0x80 reserved (evil bit)
 D 0x40 Do Not Fragment
 M 0x20 More Fragments follow

RFC 791

Please refer to RFC 791 for the complete Internet Protocol (IP) Specification.

source: <http://nmap.org/book/tcpip-ref.html>

- Version (at the moment 4, IPv4)
- Header Length
- Type of Service (ToS)
 - defines the type of the service (as expected)
 - makes a limited datagram treatment selection possible
 - since it is not required by the specification, at the moment it doesn't really matter on the current Internet
- Total Length

IPv4 header format

- Identifier (important in case of fragmentation)
- Flags and Fragment Offset: defragmentation on the recipient station, with the aid of Identifier
- Header Checksum
- Source Address
- Destination Address

- Time To Live (TTL)
 - maximal on datagram send time
 - represents the lifespan of the datagram
 - decremented by each and every network device
 - if reaches zero, the packet must be thrown
 - this prevents a packet from circulating forever in a loop
- Protocol: contains upper level protocol code (e.g. TCP, UDP)

- Occurs when the MTU (Max. Transm. Unit) is smaller on the next link than the packet size
 - the router (or the sender) fragments the packet
 - every fragment contains information that identifies the byte position in the original packet
 - provides the packet with unique ID which gets inherited by the fragments so the receiver can assemble it

- The router sends the fragments one by one
- Bigger fragments can get fragmented further
- The recipient assembles the fragments
- If the sender sets the DF (Do not Fragment) bit, then it shall not be fragmented
 - in case MTU is too small for transit, the packet is thrown and the sender is informed in an ICMP message

- IP demands that
 - the network must be capable of transmitting IP packets of at least 68 bytes
 - that's the minimal MTU
 - e.g. Ethernet/WLAN: MTU = 1500 byte
 - stations must be capable of receiving at least 576 byte packets (in one piece or fragmented)
 - it is recommended that the stations use this size (576 byte) for communication
 - out of date: try to use maximal MTU

- Options implement uncommon IP functions
 - doesn't need to be a part of every packet
 - stations must interpret and process them
 - not the implementation, but their presence is optional
 - security: information needed for authentication

- Source routing: The packet travels through the path (list of stations) determined by the sender. The two types are
 - strict: the packet only travels through the given stations; if two stations, which are listed to be neighbours but are not, the packet is lost and Source routing failed ICMP packet is sent to the sender
 - loose: if two stations, which are listed to be neighbours but are not, the packet is forwarded to the next station, but on a path selected by routers

Options

- Pathlock: the IP address of stations reached by the packet is recorded in the packet
- Time stamp
- Stream ID: 16 bit identifier, especially important when cooperating with stream oriented networks

- ICMP = IP control protocol
 - seems to be an upper level protocol
 - in fact it is a part of IP
 - the ICMP protocol of IPv4 and IPv6 are totally different (ICMPv4 and ICMPv6)
 - an ICMP packet is an IP packet with protocol ID 1

- ICMP can be
 - Recipient unreachable.
 - router transmits to sender if
 - there is no such recipient
 - targeted destination is infinitely far away
 - fragmentation needed apart from set DF bit
 - recipient can also send it, if for example there is no stack supporting the correspondent protocol
 - TTL expired.
 - can be sent by router, if TTL reaches zero
 - can be sent by recipient, if the time for waiting fragments expired and not all the fragments have arrived

- Incorrect message sent.
- Packets are sent too fast. This can be sent by router or recipient before resources are depleted, thus the response to this message can still arrive.
- Redirect. The packets are redirected to a different route because it is shorter. This can be sent by routers to improve the network.

- Echo and Echo reply. Reachability can be tested by these messages. Station must respond to Echo by Echo reply. It is also used by ping.
- Timestamp request and response. Used to monitor station timers.
- Network number discovery. A station can call someone in it's network (network number can be left unfilled) to get the number. The response contains a fully completed address.

- With the introduction of RFC 1256, ICMP router discovery was added to the original ICMP functions.
 - Router Advertisement is periodically sent on the link, containing numerous parameters, so the stations get to know the routers on the link
 - Router Solicitation can be sent to acquire these information explicitly, without waiting

IPv4



source: <http://www.gomonews.com/wp-content/uploads/2010/07/ipv4-ipv6-feature.jpg>

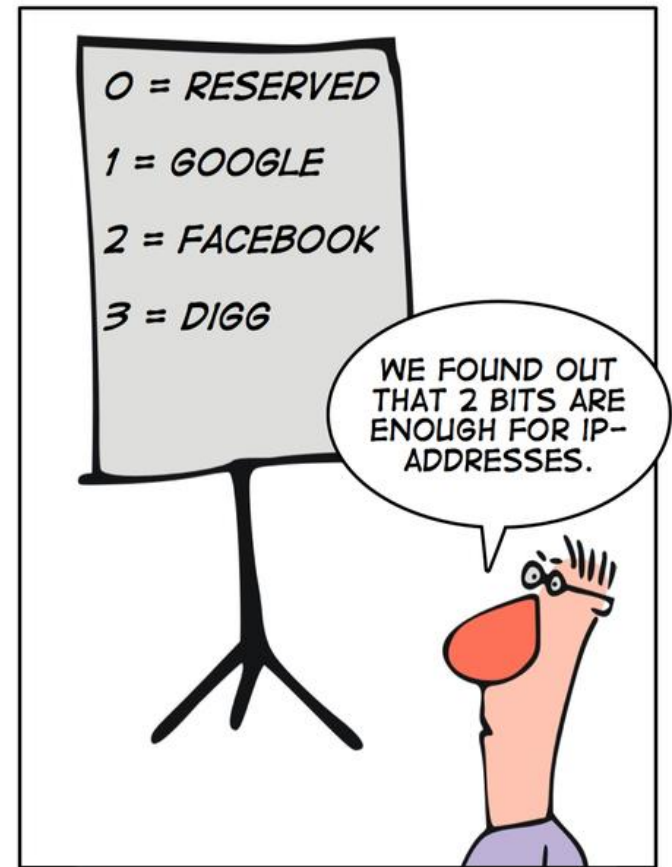
Or as heard on RIPE 55 meeting:

<http://www.youtube.com/watch?v=y36fG2Oba0>

(szöveg: <https://apps.db.ripe.net/search/query.html?searchtext=POEM-RIPE55-SONG#resultsAnchor>)

Famous statements

- **“I think there is a world market for maybe five computers.”**
 - Thomas Watson, chairman of IBM, 1943
- **“640K ought to be enough for anybody.”**
 - Bill Gates, 1981
- **“32 bits should be enough address space for Internet.”**
 - Vint Cerf, 1977 (Honorary Chairman of IPv6 Forum 2000)



IPv7

source:

<http://geekandpoke.typepad.com/.shared/image.html?/photos/uncategorized/2007/05/17/ip1.jpg>

Addressing problems

- IPv4 contains no information on geographical distances, although it would come handy in routing
- Large sites require more Class C blocks, making interdomain routing tables increase faster, than router memory
- Divided address space handling is expensive and complex (must be maintained on each and every router)
- Addresses are depleating (time's ticking)



source: <http://www.sum-it.com/?p=247>



source: <http://media.bestofmicro.com/IPv4,A-R-255699-1.jpg>

- CIDR – Classless Inter-Domain Routing
- NAT - Network Address Translator
- Obtaining unused addresses for redistribution
- Distributing unused Class A addresses
- Rearrangement of the current address usage structure

Classless Inter-Domain Routing

- CIDR differs from the conception of address classes
- Instead, it is a generalization of the subnet prefix, subnet mask conception.
 - routers use the subnet mask (which needs to be stored) to determine the border between network and station address, instead of using the first 3 bits
 - routers capable of CIDR ignore address classes, they only deal with the mask

- Experts state that if CIDR had not been introduced in 1994/1995, routing tables would have expanded to uncontrollable proportions by now, making today's Internet unable to operate
- Most routers are already equipped with this technology, even IANA address distribution is based on CIDR

- NAT is one of the most wide spread way to connect to the Internet
 - based on RFC 1918
 - a recommendation on address distribution of IP-based networks without connection to the Internet
 - instead of globally unique addresses, local is sufficient
- NAT provides a cure against address depletion

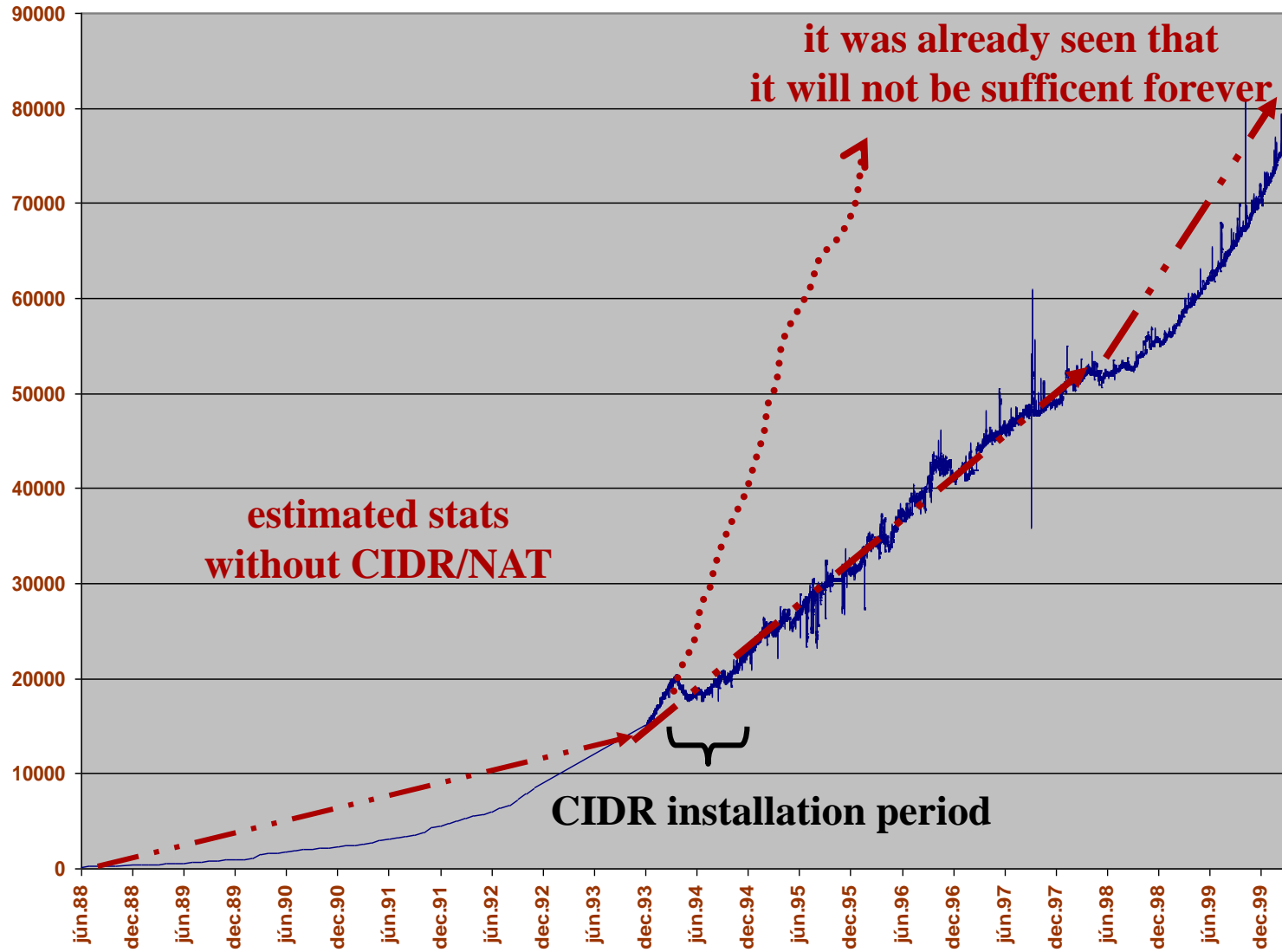
- IANA 3 separated 3 domains for this purpose:
 - 10.0.0.0/8 (10.0.0.0—10.255.255.255)
 - 172.16.0.0/12 (172.16.0.0—172.31.255.255)
 - 192.168.0.0/16 (192.168.0.0—192.168.255.255)
 - an organization which does not want to have its network connected to the Internet can choose from these addresses
 - this means there is no need to turn to IANA for IP addresses
 - IANA assures that these addresses will not be distributed

- The organization wishing to NAT will ask for an IP address from its Internet service provider (this shall be on the exterior of NAT)
- Stations of the network get labelled from one of the previously mentioned address domains (this shall be on the interior of NAT)
- A NAT-using module
 - dynamically substitutes the internal addresses with external addresses in outgoing packets
 - reverses this process in replies

- Advantages
 - reduces addresses required on the Internet
 - increases security (internal network invisible to the world)
 - the organization can keep its address network structure even in case of moving to a new Internet service provider
- Disadvantages
 - communication can only be initiated by internal stations
 - maintenance of NAT server requires some tricks
 - the union of two NAT networks can be rather complicated
 - violates the end-to-end concept

- The disadvantages of class-based address space revealed themselves with haste
 - Not enough granulatiry
- New ideas:
 - Classless Inter Domain Routing (CIDR)
 - adjustable address space
 - introduction of netmask
 - Network Address Translation (NAT)
 - private networks
 - address spaces behind routers
 - arbitrary usage of three address domains

Effects of CIDR and NAT



Reobtaining unused addresses

- IANA proposal [RFC 1917]
 - a network which shall never connect to the Internet should give the reserved IP addresses back
 - Internet Service Providers with too many unused network prefixes should give them back
- My proposal is that y'all should gimme all yo money
 - questionable success

Distributing unused A and E addresses

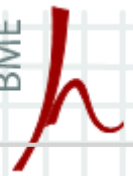
- Part of Class A is reserved for misc. purposes
 - 64.0.0.0/2 domain not distributed
 - recommendation for the distribution of this domain, since it's a fat slice of the entire IP address domain
- Class E addresses
 - distributed as Class B and C addresses

Modification of address obtaining

- Address space allocation
 - the organization asks for a domain from IANA
 - if obtained, it can be kept as long as desired (or as long as affordable)
- IETF recommendation
 - the organization only borrows the domain
 - after a specific time, it needs to let go and ask for another
 - this way allocation becomes dynamic, although comes with several great disadvantages

Address obtaining – disadvantages

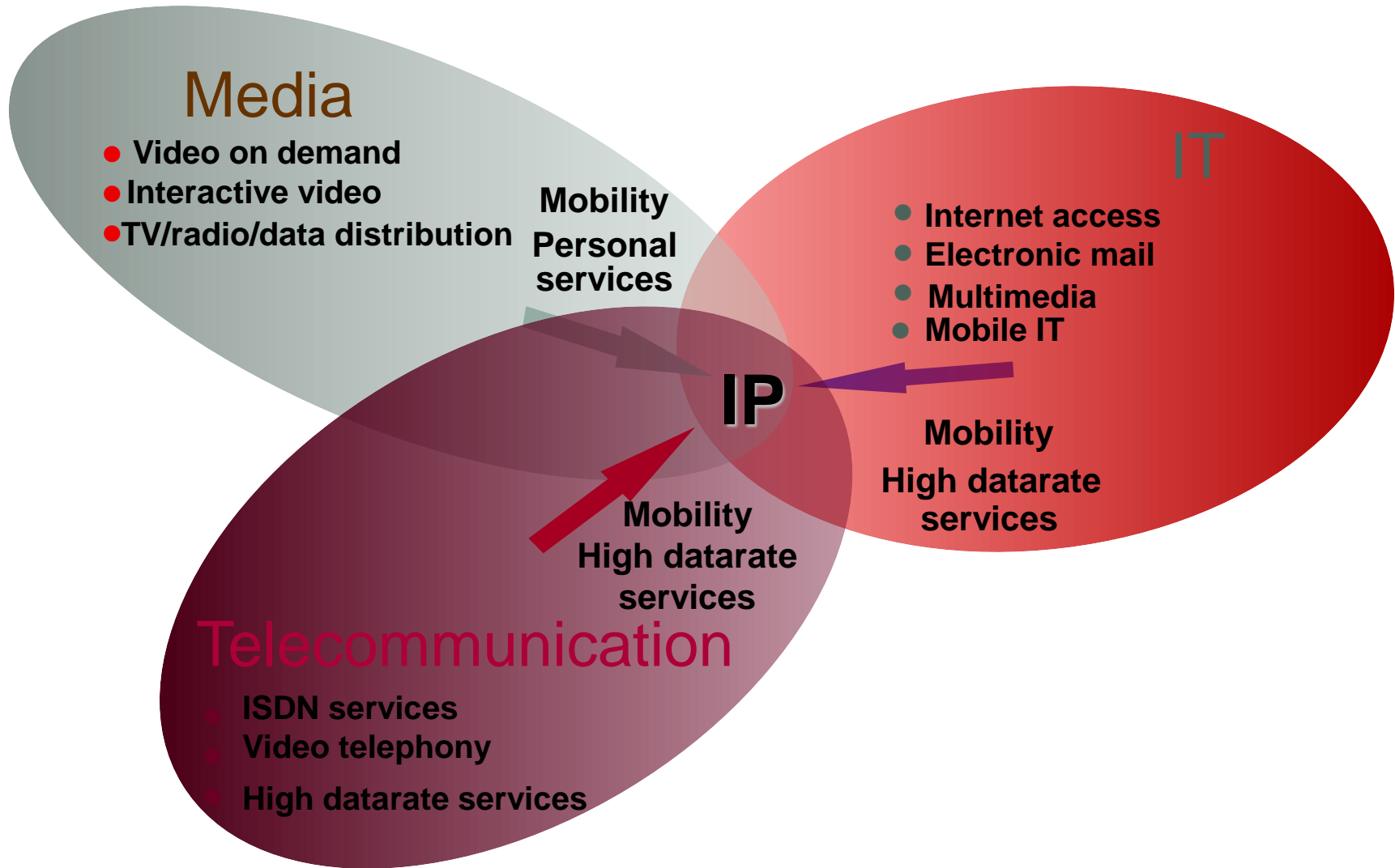
- Rule of CIDR technology: address distribution needs to reflect network topology
 - reallocation makes it chaotic
 - more and more new detours in routing tables
 - dynamism comes with the price of the reduction of efficiency of packet routing
- This method would not be very popular among the society of the Internet
 - IETF Procedures for Internet/Enterprise Renumbering (PIER)



- IPv4 address changes during movement to a new network
- Connections break up during handover
- New protocol needed in case of roaming for the association of the new IP address and the unique ID in HLR
- WAP and GPRS users dramatically increase mobile Internet usage
- IPv4 address space is about to be depleted

- Home Networking
 - tons of unique IP addresses required on equipment level
- Service support
 - new services cannot harness the fix fields of IP header
 - no build-in IPv4 security algorithm
 - QoS would be improved by IP level traffic flow attributes

Tendencies



Depletion of address domains

- Reason: Insufficient scaling decades ago
- Making things worse:
 - low efficiency of address usage
 - demographical issues
 - constant connection access
 - mobile devices
 - virtualization (multiple system on one hardware)
- Helping the situation:
 - CIDR
 - NAT
 - virtual domains distributed by name
 - strict rules of RIR distribution
 - unused large domains reobtained

Depletion of address domains

- Despite economical efforts:
 - 2011-02-01: IANA distributed 2 from 7 existing /8 networks to APNIC
 - code red: last 5 /8 domains distributed between the 5 RIR
 - 2011-02-03: glorious and majestic deliverance of the last free domains, IPv4 finally got depleted



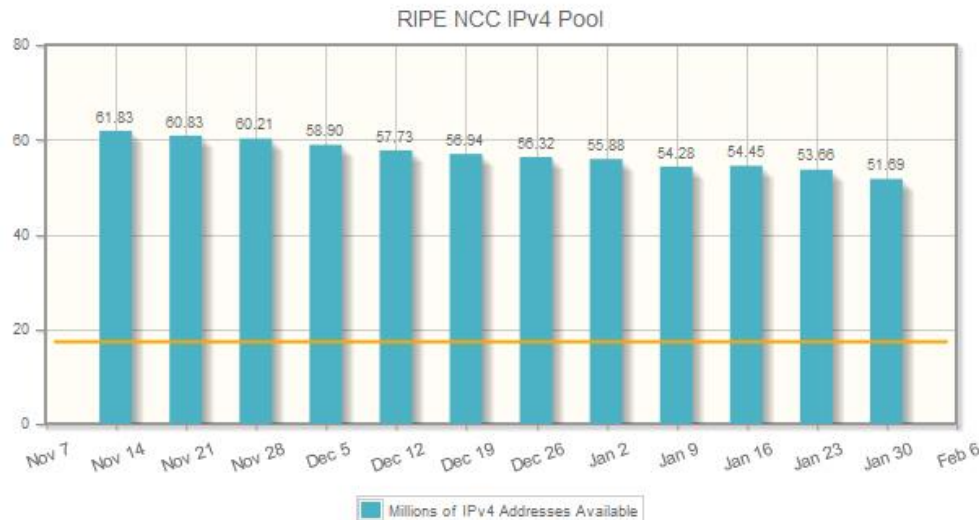
source:
<http://prensa.lacnic.net/news/en/6th-edition-february-2011>

Depletion of address domains

- Some free domains at RIR still can be found
- Although their number continuously decreases
- RIR-shopping: one RIR can purchase domains from another
- Saving for a rainy day, multinational company style: Microsoft spent 13\$/IP on domains in 2011-03

RIPE NCC IPv4 Available Pool - Graph

30 Jan 2012



Forrás:
<http://www.ripe.net/internet-coordination/ipv4-exhaustion/ipv4-available-pool-graph>

What will happen to the Internet?

- „Don't panic!” by Douglas Adams
- Searching for a solution since 1993
- In 1998, a standardized solution was born (and it's not 42):
 - Internet Protocol version 6
 - IETF RFC2460
- Apart from hopes and expectations, due to CIDR and NAT, IPv6 was forced to stay in the background
- However, the protocol kept developing, upgrading:
 - IPv6 protocol stack development, testing
 - projects worth mentioning: KAME, Nautilus6, Tipster6 (Hungarian)
 - our department had its own fair share of development and testing (e.g.: IST-PHOENIX, IST-ANEMONE, ICT-OPTIMIX, EUREKA-Celtic BOSS)

Features IPv6

- The most important: 128 bits address space, ginormous: one address for each and every m² of the planet ($6,5 \cdot 10^{23}$)
- Streamlined header
- Optional headers
- Built-in security
- Built-in mobility support
- Autoconfiguration
- Multicasting
- Anycasting
- Neighbour discovery
- Etc



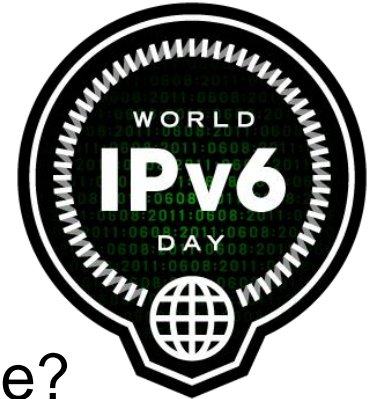
World IPv6 Day

- 2011-02-03: by the day of the depletion of IPv4, the time for IPv6 had come
- 2011-06-08: protocol testing
- By Hungarian time:
2011.06.08 2:00 – 2011.06.09 2:00
- What is this day all about?
 - a global flight with the wings of IPv6, supported by the Internet Society (isoc.org)
 - this day the major web organizations and industrial companies launched IPv6 on their services
 - by doing so, they made the ride to IPv6 less bumpy



What did World IPv6 Day achieve?

- Why do we need days like that?
 - it is the solution of the future, sooner or later everyone has to accept it
 - big companies come with big encouragement for others
- How does it motivate testing and convergence?
 - common goal for ISP, hardware manufacturers, website maintainers and for OS designers
 - we gotta survive this together
 - global scaling is one of the main ideas behind World IPv6 Day



Who participated?

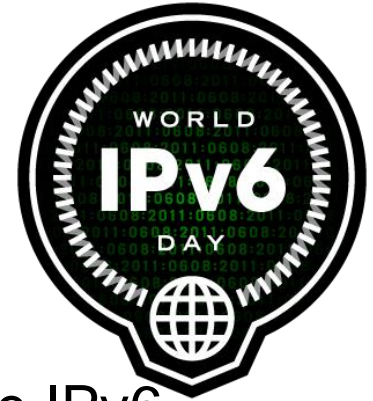
- Who participated on the very first global test of IPv6?
 - Google
 - Youtube
 - Yahoo
 - Microsoft
 - Akamai
 - Cisco
 - W3C.org
 - Facebook
 - Etc

- Full list on <http://www.worldipv6day.org/participants/index.html>



What's missing?

- What do we need for the true global usage of IPv6?
 - ISP need to provide IPv6 availability to users
 - web service providers need to provide their services on IPv6
 - OS designers need to provide service packs (only a few OS involved)
 - backbone network maintainers need to provide IPv6 access to their peers (Hungarian backbone is mostly IPv6 compatible)
 - new firmware needed from hardware, router and modem manufacturers



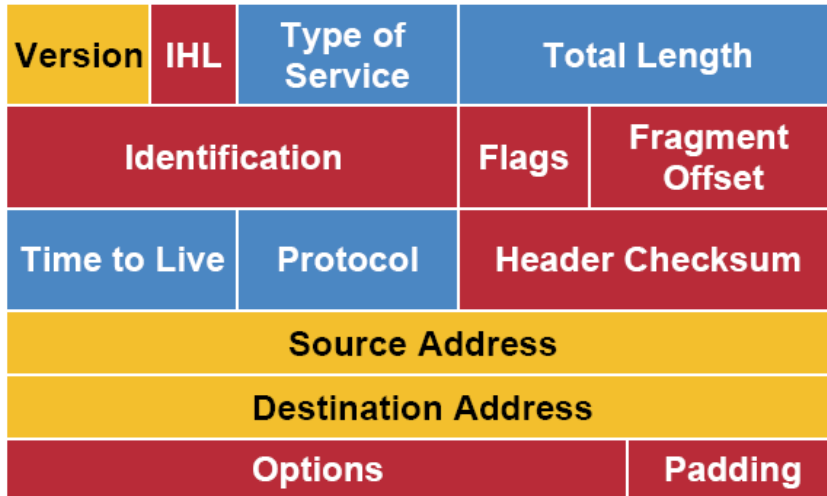
What to expect?

- Longer test periods
- Continuous convergence of users and service providers to IPv6
- 20 years of coexistence of IPv4 and IPv6 expected
- Cooperation in these years must be dealt with

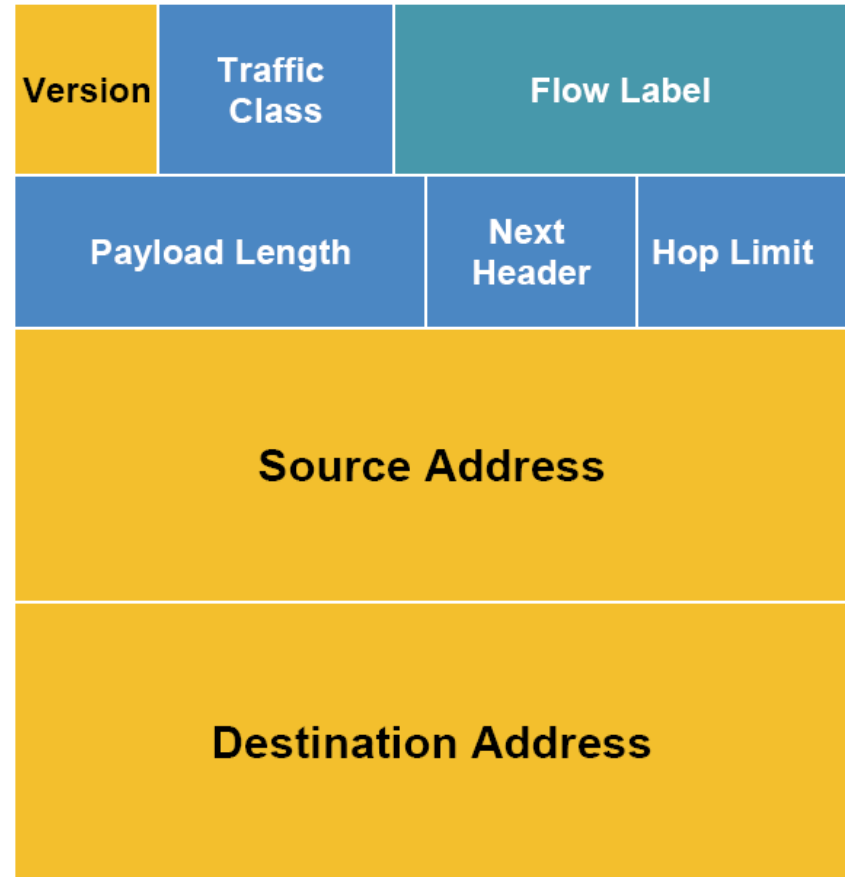
- At the moment IPv4 (from the beginning of the '70s)
 - great experience (30++ years)
 - base of the Internet since 1983
 - continuous development
- IPv6?
 - getting standardized since 1990
 - standardized in 1995 (draft)
 - several missing components
 - low level of experience
 - continuous development

IPv4 and IPv6 header comparison

IPv4 Header



IPv6 Header



- Legend**
- Field's Name Kept from IPv4 to IPv6
 - Fields Not Kept in IPv6
 - Name and Position Changed in IPv6
 - New Field in IPv6

source: <http://343networks.files.wordpress.com/2010/06/ipv4-ipv6-header.gif>

IPv6 header – missing

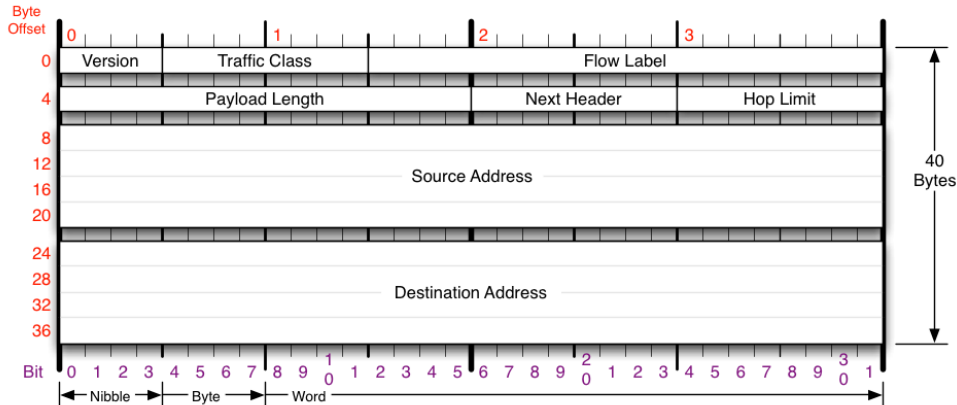
- The following IPv4 fields are gone
 - header length (fix 40 byte)
 - identifier
 - flags
 - fragment offset
 - header checksum
- The middle 3 were needed for fragmentation control, which does not exist in IPv6
- Checksum = slow

IPv6 header – changed

- Type-of-Service => Traffic class
 - priority control
- Protocol Type => Next header
 - TCP, UDP, also options header, see later
- Time To Live (TTL) => Hop Limit
- Sender and recipient address (longer)
- New field: Flow label
 - more effective packet delivery

IPv6 header

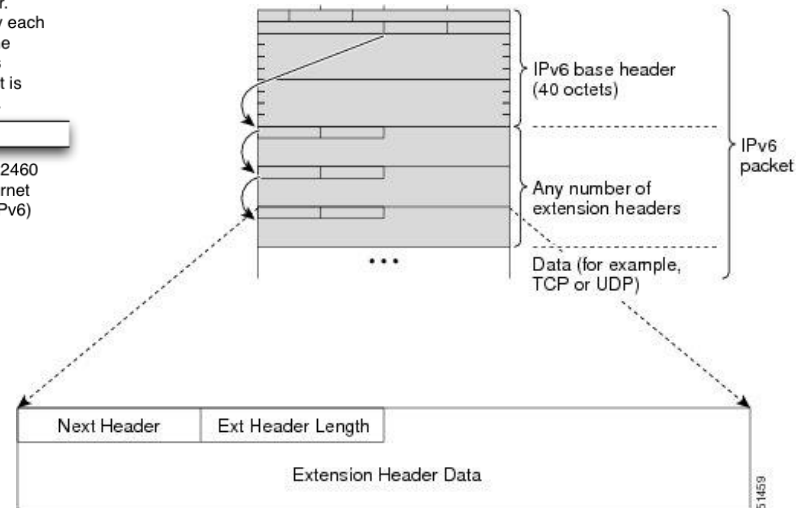
IPv6 Header



Version	Payload Length	Next Header	Hop Limit
Version of IP Protocol. 4 and 6 are valid. This diagram represents version 6 structure only.	16-bit unsigned integer. Length of the IPv6 payload, i.e., the rest of the packet following this IPv6 header, in octets. Any extension headers are considered part of the payload.	8-bit selector. Identifies the type of header immediately following the IPv6 header. Uses the same values as the IPv4 Protocol field.	8-bit unsigned integer. Decremented by 1 by each node that forwards the packet. The packet is discarded if Hop Limit is decremented to zero.
Traffic Class	Source Address	Destination Address	RFC 2460
8 bit traffic class field.	128-bit address of the originator of the packet.	128-bit address of the intended recipient of the packet (possibly not the ultimate recipient, if a Routing header is present).	Please refer to RFC 2460 for the complete Internet Protocol version 6 (IPv6) Specification.
Flow Label			
20 bit flow label.			

Copyright 2008 - Matt Baxter - mjb@fatpipe.org - www.fatpipe.org/~mjb/Drawings/

source: http://www.siongboon.com/projects/2006-03-06_serial_communication/IP-Header-v6.png



source: <http://www.cisco.com/en/US/i/000001-100000/50001-55000/51001-51500/51459.jpg>

- In this order
 - Hop-by-Hop Options header (jumbogram)
 - Destination Options header (internal destination)
 - Routing header (routing type)
 - Fragment header
 - Authentication header
 - Encrypted Security Payload header
 - Destination Options header (final destination)
 - (Upper layer header)
- Next header tells what's coming next

- In every options header:
 - next header, header length, options
- Involved nodes need to process it
 - Hop-by-hop = step by step
- Hop-by-hop options: Jumbogram and Router Alert Option
 - Jumbogram = very big packet
 - payload length = 16 bit (max. 64 kbyte)
 - 32 bit (max. 4 Gbyte)
 - Router Alert Option: this informs the routers that there is information to be processed inside the packet (e.g. resource allocation control protocol for QoS)

Destination Options header

- This is also an Options header
 - same format
- Destination needs to process it
- Can occur two times
 - if routing header is used
 - first (before routing header)
 - stations listed in routing headerben will process it
 - last (after routing header)
 - final destination processes it

- (used to be Source routing option in IPv4)
- loose/strict path selection
- In the header:
 - next header
 - node (address) number
 - routing type (loose/strict)
 - remaining segment number (next address)
 - (Reserved)
 - addresses of stations that must be reached

Fragment header

- Only the source can fragment, the router cannot
 - if jumbogram, then no
- In the header:
 - next header
 - reserved
 - fragment offset (from where to fragment)
 - e.g. IPv6 header cannot be fragmented
 - identification
 - which fragment

Authentication header

- Authentication purpose (origin)
 - was it really sent by the sender?
 - has it been modified?
- In the header:
 - next header, payload length, reserved
 - SPI (Security Parameter Index)
 - Sequence number – even in case of UDP
 - Authentication data

- Purpose: encryption (confidentiality)
 - only those can read who should
- In the header:
 - SPI
 - sequence number
 - encrypted data
(payload, padding, padding hossza, next header)
 - Authentication data
- Two types: transport and tunnel

IPv6 addressing – some numbers

- IPv4 – 32 bit
 - $2^{32} = 4,29 \cdot 10^9$ addresses (theoretically)
 - more than 6,5 billion people on Earth
 - 2 113 389 networks
- IPv6 – 128 bit
 - $2^{128} = 3,4 \cdot 10^{38}$ addresses (theoretically)
 - $6,65 \cdot 10^{23}$ addresses per m^2
 - $2^{45} / 48$ network (global unicast 001)
 - $3,5 \cdot 10^{15}$ networks
 - each comes with 65 535 /64 further networks

- By destination
 - unicast
 - multicast
 - anycast
- By routability
 - global
 - non-global
 - link-local
 - unique local IPv6 address (used to be site-local)



IPv6 address

- 128 bit = 8 x 16 bit in hexadecimal format
e.g. 2001:0db8:0000:0000:0002:b3ff:fe1e:8329
- Opportunities for simplification
 - leaving the zeros from the beginning
2001:db8:0:0:2:b3ff:fe1e:8329
 - double colon: instead of zeros
2001:db8::2:b3ff:fe1e:8329
only once
- Prefixes: IPv6 address/prefix form
 - 2001:db8:0:56::/64

- Distributable
 - 2000:: - FE80:: - FEC0:: - not used anymore
 - FC00:: - FF00::
- Special
 - :: unspecified address (like 0.0.0.0 in IPv4)
 - ::1 loopback

- IPv4 address embedded into IPv6 (out of date)
 - IPv4 compatible IPv6 address `::ipv4_address`
e.g. from 62.2.84.115 to `::3e02:5473`
- IPv4 address mapped to IPv6 `::FFFF:ipv4_address`
 - widespread and used
e.g. from 62.2.84.115 to `::ffff:3e02:5473`
 - IPv4 part can be decimal
e.g.: `::ffff:62.2.84.115`
- 6to4 address `2002:public_ipv4_address::/48`
- ISATAP addresses
 - dual stack nodes over IPv4
- Teredo addresses (behind NAT)

Global unicast addresses

- Starts with binary 001 (2000::- n bit for global route prefix (e.g. by geographical position)
- 64- n bit subnet identifier
- 64 bit interface identifier
- e.g.:
 - Hungarnet: 2001:738:: - our university: 2001:738:2001:: - our department: 2001:738:2001:4020::

- Link-local: never to be routed
 - no configuration needed
 - ideal for ad-hoc and routerless networks, or for neighbour discovery
- Form: FE80::[64_bit_Interface_ID]
 - e.g. if the Ethernet card hardware address is 00:1A:6B:3A:9F:BC, then the link-local address is FE80::**2**1A:6B**FF:FE**3A:9FBC
- Modified EUI-64 algorithm:
 - first convert the 48 bit MAC address into 64 bites EUI-64 by inserting **FF:FE** bits into the middle
 - then the 7th bit gets inverted: 00->**02** in the example
- Unique local IPv6 addresses:
 - identifying by FC00::/7 prefix
 - [7_bit_prefix][1_bit_L_flag][40_bit_global_ID]:[16_bit_subnet_ID]:[64_bit_Interface_ID]
 - L bit = 1 : local association by a pseudo-random Global ID algorithm
 - L bit = 0 : centralized, no method defined

- Invented for heavy loaded devices
 - addresses one station (typically the nearest) from a cluster of computers
- Freely from the unicast domain
- Subnet-router anycast
 - $[n_bit_subnet_prefix]:[128-n_bit_0]$
 - the first router on the link will process it
- Reserved subnet anycast address
 - on the last 7 bits, e.g. 126 (7E): mobile IPv6 Home-Agent anycast

- FF[0RPT][4_bit_scope][Group_ID]
 - 0RPT flags (bits)
 - R=0 rendezvous point not embedded
 - P=0 multicast address without prefix information
 - T=0 well known multicast address (1: temporary)
 - Scope examples
 - 1: Interface-local scope (~multicast loopback address, will not leave the node if addressed by this)
 - 2: Link-local scope (cannot be further routed)
 - 5: Site-local scope
 - E: Global scope

- Every node
 - same link as the sender FF02::1
 - same site as the sender FF05::1
- Every router
 - same link as the sender FF02::2
 - same site as the sender FF05::2
- Every DHCP client FF02::1:2
- Every DHCP server FF05::1:3
- Every NTP server
 - same site as the sender FF05::101
 - on the Internet FF0E::101

- IPv6 lets more addresses coexist on one interface, addresses can be different in their
 - scope (link-local, global)
 - state (precedence, cancelled)
- Which one to use?
 - there is a standard for this (RFC 3484)
 - e.g. source address selection:
 - advantageous if same address as the destination
 - advantageous if same or greater scope than the destination
 - native instead of 6to4 or ISATAP, if possible
 - home address preferred over temporary
 - advantageous if external interface
 - advantageous if public address
 - same longest prefix as destination

- Much better than ICMPv4
 - multicast management (instead of IGMP)
 - neighbor discovery (instead of ARP, RARP)
 - discovery of adjacent stations, routers, reachable neighbours and altering addresses
 - echo request/echo reply (ping)
 - packet too big (instead of fragment header)
- Two types of messages
 - error
 - informative

- Destination unreachable
 - IP datagram cannot be transmitted
 - no route to destination, address/port unreachable, administratively forbidden
- Packet Too Big
 - MTU on next link is smaller than packet size
- Time Exceeded
 - if hop counter reaches zero
- Parameter problem
 - if a parameter cannot be interpreted

ICMPv6 informative messages

- Echo request / echo reply
- Multicast discovery messages
 - router
 - listener
- Router discovery
- Neighbor discovery
- Router renumbering
- messages considering mobility support
 - see details later

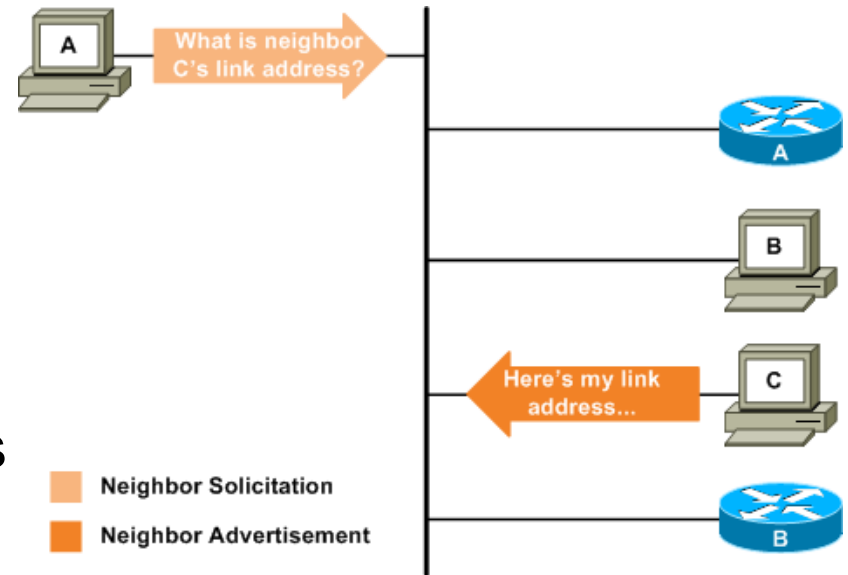
ICMPv6: echo request and reply

- Same as in ICMPv4
- The content of echo request needs to be copied into echo reply
- Used by ping6

- Tasks
 - address autoconfiguration (stateless)
 - network prefix, automatic router discovery
 - duplicated IP address query
 - MAC address discovery
 - neighbour router discovery
 - identification of unreachable neighbours (NUD)
 - detection of MAC address alteration
- Neighbor solicitation and router advertisement
 - MAC address dissolution (ARP in IPv4)
 - identification of neighbour reachability
 - identification of duplicated IP addresses
- ICMP redirect
- Multicast Listener Discovery (MLD – RFC 3810)
- Multicast Router Discovery (MRD – RFC 4286)
- Inverse Neighbor Discovery (IND)
 - RARP in IPv4
- Vulnerability, security
 - SEcure ND (SEND)

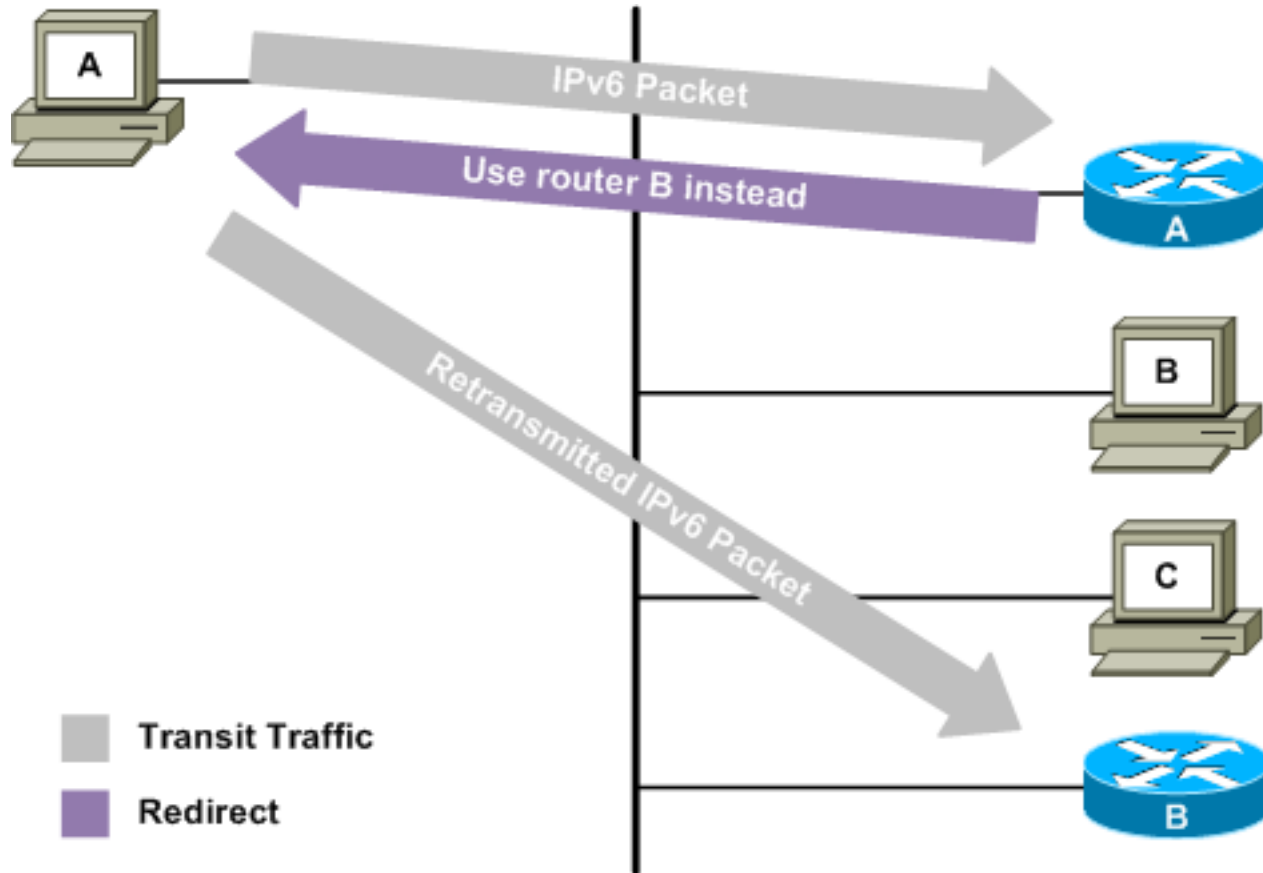
ICMPv6 example: Neighbor solicitation

- A neighbor solicitation:
 - type: 135
 - code: unused
 - checksum
 - reserved
 - target address: MAC address that needs to be resolved
 - options: e.g. source link-layer address: recipient MAC address



source:
http://media.packetlife.net/media/blog/attachments/87/neighbor_solicitation.png

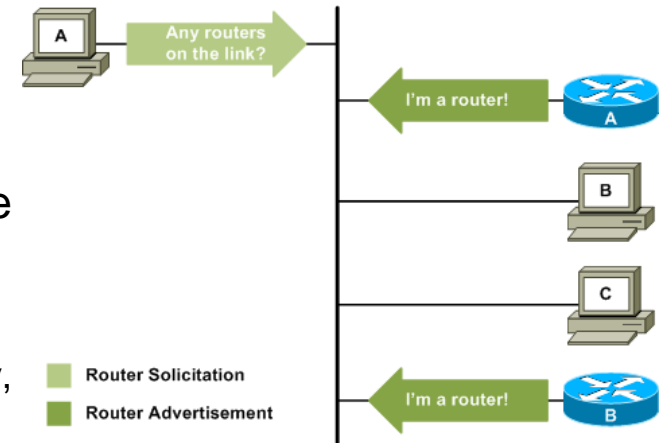
ICMPv6 example: Redirect



source:
<http://media.packetlife.net/media/blog/attachments/88/redirection.png>

ICMPv6: Router Discovery

- Two messages
 - Router advertisement
 - Router solicitation
- In router advertisement message:
 - Current Hop limit (hop limit recommendation to the nodes on the link)
 - Autoconfig flags
 - M: 0-SLAAC, 1-DHCPv6
 - O: 1-for options other than address and default gateway, use DHCPv6
 - H: 1-Home link (Home Agent flag)
 - Prf (2 bit): Preference between routereks (RFC 4191)
 - Router lifetime
 - 0- not default router
 - how many seconds is the default router available for
 - Reachable time
 - in term of neighbours: after receiving the availability information, how long should the host be considered to be available
 - used by Neighbor Unreachability Detection
 - Retransmission timer
 - time (ms) between retransit of NS messages
 - used by address resolving and Neighbor Unreachability Detection
 - Options (source MAC address, MTU, prefix)



source:

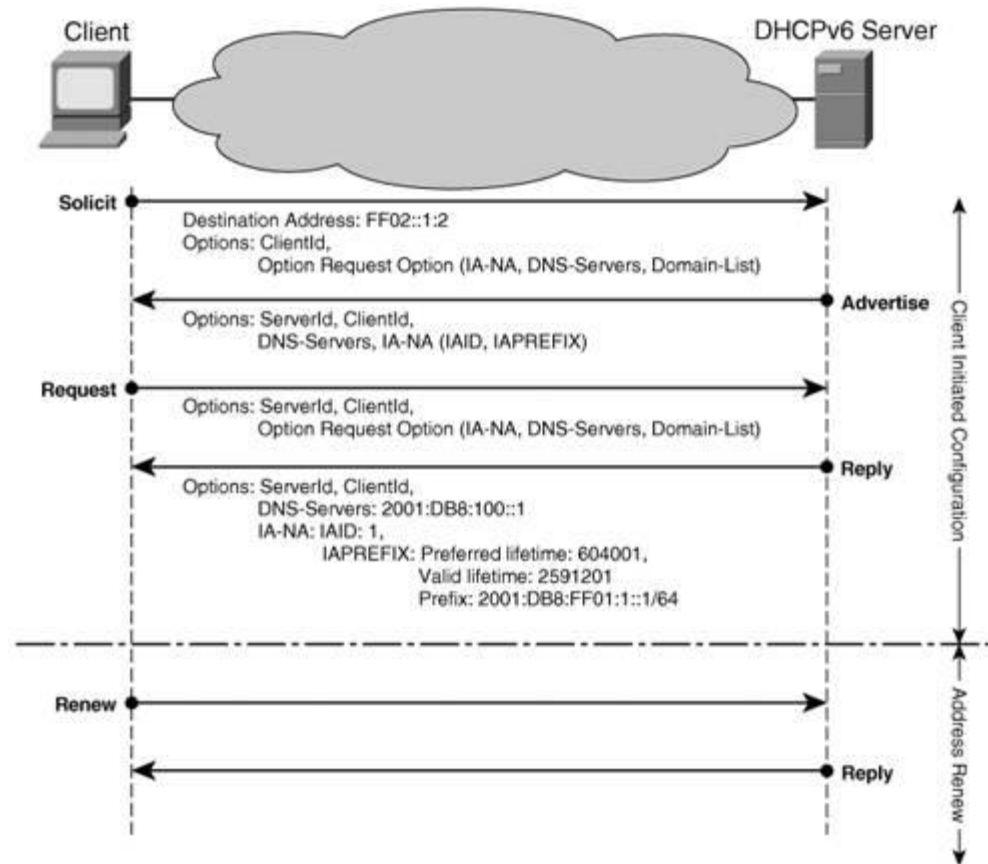
http://media.packetlife.net/media/blog/attachments/86/router_solicitation.png

- Basic address configuration types:
 - Manual
 - Stateless Address Autoconfiguration (SLAAC, RFC 4862)
 - Router Advertisement advertises the address and the default gateway
 - Stateless DHCPv6 (RFC 3736)
 - Router Advertisement advertises the address and the default gateway, but the O flag is in 1, so with DHCPv6 the node only gets DNS, NTP, etc
 - Stateful DHCPv6 (RFC 3315)
 - Router Advertisement M flag=1, A flag=0 (autoconfig off), so addressing and other information from DHCPv6, but default gateway still from Router Advertisement
 - DHCPv6-PD (Prefix delegation, RFC 3633)
 - association entire networks to the router

Stateless Address Autoconfiguration

- Link-local address generation (based on MAC address)
- Link-local address testing (if it's unique or not, with ND DAD)
 - sending NS for ones address resolution: if no response, then OK
 - if OK, then proceed, if not OK, then go back
- Link-local address setup on the interface
- Connecting the router
 - sending Router Solicitation to All Router Multicast address
 - replies with Router Advertisement, the fact of SLAAC notated by flags
- Global address configuration (with the help of prefix options)
- Run DAD on global address
- If OK, address setup

DHCPv6 operation details



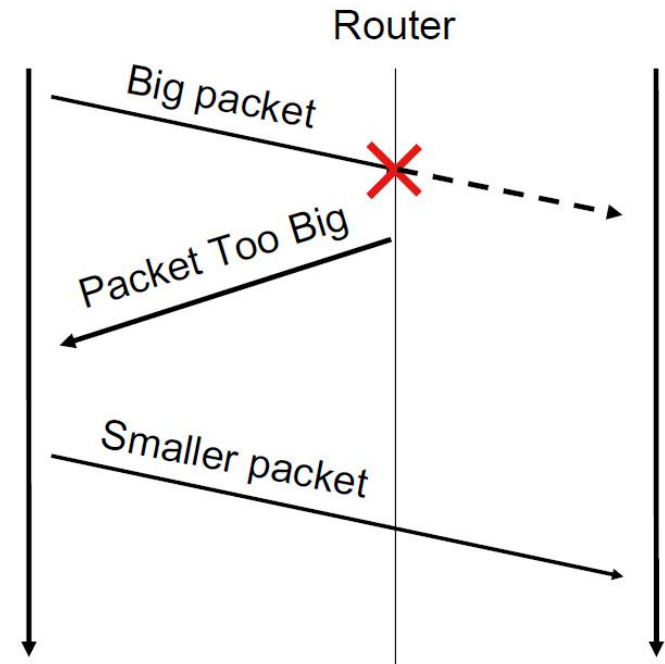
source:

[http://www.iphelp.ru/doc/3/Cisco.Press,..Deploying.IPv6.Networks.\(2006\).BBL/1587052105/images/03fig01.jpg](http://www.iphelp.ru/doc/3/Cisco.Press,..Deploying.IPv6.Networks.(2006).BBL/1587052105/images/03fig01.jpg)

- Used to overwrite subnet prefix
 - helps network administration
 - authentication needed
 - ordinal number versus circulating messages
- Two messages:
 - Router renumbering command
 - Match prefix part (what to be modified) => use prefix part (modify to what)
 - the old prefix becomes the new
 - Router renumbering result
 - feedback about the result

ICMPv6: Path MTU discovery

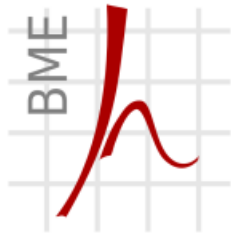
- no fragmentation in IPv6
- If the packet is too big ($>$ MTU):
 - thrown by the router
 - sends a ICMPv6 message to the sender (PTB)
 - PTB includes the MTU of the next link
- Method:
 - echo request to the address
 - start with big MTU, then decrement
 - try with the new MTU
 - never goes under 1280 byte
 - GOTO the beginning



source:
http://ripe60.ripe.net/presentations/Stasiewicz-Measurements_of_IPv6_Path_MTU_Discovery_Behaviour.pdf

Questions?

THANK YOU!



Híradástechnikai Tanszék



Dr. László Bokor
Ph.D., assistant professor
Department of Networked Systems and Services, BME
bokorl@hit.bme.hu

